

# Quantum Mechanics over Sets: Making the case for the objective indefiniteness interpretation of QM

David Ellerman  
University of Ljubljana, Slovenia

December 11, 2021

## Abstract

Quantum mechanics over sets (QM/ $\mathbb{Z}_2$  or QM/Sets) is:

1) a pedagogical or ‘toy’ model of finite-dimensional quantum mechanics (QM/ $\mathbb{C}$ ) that reproduces in the simplified setting of vector spaces over  $\mathbb{Z}_2$  the essentials of projective measurements, the double-slit experiment, the indeterminacy principle, entanglement, Bell’s Theorem, the statistics of indistinguishable particles, and so forth,

2) a non-commutative extension of finite probability theory that includes "superposition events" in addition to the usual classical events (subsets of the outcome set) and multiple equi-cardinal bases, and

3) a new objective indefiniteness approach to the old problem of interpreting quantum mechanics by highlighting the mathematics of indefiniteness vs. definiteness (i.e., partitions) and by showing what is important and what is not important (e.g., the wave interpretation) in the usual mathematics of QM/ $\mathbb{C}$ .

## Contents

<b>1</b>	<b>Introduction: QM/Sets</b>	<b>3</b>
1.1	The vector space $\mathbb{Z}_2^n$ . . . . .	3
1.2	Kets, bra-kets, Dirac brackets, and the norm in QM/Sets . . . . .	6

1.3	Numerical attributes and random variables in QM/Sets . . . . .	10
1.4	The linearization methodology behind QM/Sets . . . . .	13
<b>2</b>	<b>Philosophical interlude: the Literal or Objective Indefiniteness Interpretation of QM</b>	<b>15</b>
<b>3</b>	<b>Measurement and dynamics in QM/Sets</b>	<b>22</b>
3.1	Measurement and partitions . . . . .	22
3.2	Non-commutativity and the Indeterminacy Principle in QM/Sets . . .	27
3.3	‘Dynamics’ in QM/Sets and the double-slit experiment . . . . .	32
<b>4</b>	<b>Entanglement in QM/Sets</b>	<b>35</b>
4.1	Direct products and correlation in QM/Sets . . . . .	35
4.2	Bell’s Theorem in QM/Sets . . . . .	39
<b>5</b>	<b>Density matrices in QM/Sets</b>	<b>44</b>
5.1	Logical entropy using density matrices . . . . .	44
5.2	Revisiting measurement with density matrices . . . . .	52
5.3	Classical events and superposition events . . . . .	59
5.4	Modeling measurement in QM/Sets . . . . .	63
5.5	Measuring measurement with logical entropy . . . . .	67
<b>6</b>	<b>Indistinguishability of particles in quantum mechanics</b>	<b>70</b>
<b>7</b>	<b>Group representation theory with sets and vector spaces</b>	<b>81</b>
7.1	Group representations define partitions . . . . .	81
7.2	Where do the fully distinct eigen-alternatives come from? . . . . .	82
7.3	Attributes and observables . . . . .	85
7.4	Irreps of vector space representations . . . . .	89
<b>8</b>	<b>Concluding remarks</b>	<b>101</b>
<b>9</b>	<b>Appendix: Transporting vector space structures</b>	<b>106</b>

# 1 Introduction: QM/Sets

## 1.1 The vector space $\mathbb{Z}_2^n$

The topic of this book is a ‘non-commutative’ extension of finite probability theory with superposition events in addition to classical events and different basis sets for  $\mathbb{Z}_2^n$  as different equi-cardinal outcome sets<sup>1</sup>. This also provides a ‘pedagogical’ (or ‘toy’) model of finite-dimensional quantum mechanics (QM). Finite-dimensional QM is expressed in the  $n$ -dimensional vector space  $\mathbb{C}^n$  (for some  $n$ ) over the field of complex numbers  $\mathbb{C}$ ; the pedagogical model is expressed in the  $n$ -dimensional vector space  $\mathbb{Z}_2^n$  over the finite field  $\mathbb{Z}_2 = \{0, 1\}$ . The two models might be referred to as QM/ $\mathbb{C}$  ("QM over  $\mathbb{C}$ ") and QM/ $\mathbb{Z}_2$  ("QM over  $\mathbb{Z}_2$ " or QM/Sets). The point of developing this model is however not just pedagogical. QM/ $\mathbb{Z}_2$  is rich enough so that many of the characteristic quantum features of QM/ $\mathbb{C}$  can be modeled within this simple framework of QM/ $\mathbb{Z}_2$ —such as projective measurements, the double-slit experiment, Bell’s Theorem, the indeterminacy principle, the statistics of indistinguishable particles, and so forth. That helps to show what is important (and what is not important, e.g., "wave" imagery) in obtaining those peculiarly quantum features in the full theory.

The mathematical operations in  $\mathbb{Z}_2$  are all modulo 2 so  $1+1 = 2 = 0$ . The addition and multiplication tables could hardly be simpler as shown in Table 1.1:

+	0	1	and	×	0	1
0	0	1		0	0	0
1	1	0		1	0	1

Table 1.1: Addition and multiplication in  $\mathbb{Z}_2$ .

Instead of the  $n$ -dimensional vector space  $\mathbb{C}^n$ , the model is developed in the  $n$ -dimensional vector space  $\mathbb{Z}_2^n$ . A vector in  $\mathbb{Z}_2^n$  (expressed in a certain basis) is just an ordered  $n$ -tuple of zeros and ones. For instance for  $n = 3$ , there is the standard or computational basis of three vectors:  $(1, 0, 0)$ ,  $(0, 1, 0)$ , and  $(0, 0, 1)$ . The connection to sets (as in "quantum mechanics over sets" or QM/Sets) comes from associating each computational basis vector with a given (‘universe) set such as  $U = \{u_1, u_2, u_3\}$  or

---

<sup>1</sup>This is a simpler approach to "non-commutativity" than in the rather advanced "non-commutative probability theory" of Dan Voiculescu [64] following Alan Connes [14]. In that approach, a given probability space has an algebra of complex-valued random variables on it, and then one can generalize to an abstract version of those algebras where the elements are still called "random variables" and which may not commute.

$U = \{a, b, c\}$ . There are always  $2^n$  vectors in  $\mathbb{Z}_2^n$  so once we associate a set of  $n$  basis vectors (standard or not) with a set  $U$  of  $n$  elements, then the  $2^n$  vectors in  $\mathbb{Z}_2^n$  can be associated with the  $2^n$  subsets in the power set  $\wp(U)$ . For instance, if we take  $U = \{a, b, c\}$  and make the associations:  $\{a\}$  with  $(1, 0, 0)$ ,  $\{b\}$  with  $(0, 1, 0)$ , and  $\{c\}$  with  $(0, 0, 1)$ , then the subsets  $\{a\}$ ,  $\{b\}$ , and  $\{c\}$  of  $U$ , can be taken as the corresponding basis vectors in  $\wp(U) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$  considered as a vector space over  $\mathbb{Z}_2$ . When considered as  $n$ -tuples in  $\mathbb{Z}_2^n$ , the vectors add together component-wise mod  $(2)$  so  $(0, 1, 1) + (1, 1, 0) = (1, 0, 1)$  (since  $1 + 1 = 0$  in the middle coordinate). When the vectors are considered as subsets in  $\wp(U)$ , then the addition of vectors is the symmetric-difference operation: for subsets  $S, T \in \wp(U)$ ,  $S + T = S \cup T - S \cap T$ , i.e., the union of the subsets with the common elements taken out. Thus if  $S = \{b, c\}$  and  $T = \{a, b\}$ , then  $S + T = \{a, b, c\} - \{b\} = \{a, c\}$ . With this correspondence between zero-one vectors in  $\mathbb{Z}_2^n$  and subsets in  $\wp(U)$ , e.g.,  $(1, 0, 1) \longleftrightarrow \{a, c\}$ , we have a vector space isomorphism  $\mathbb{Z}_2^n \cong \wp(U)$ . That is why "quantum mechanics over  $\mathbb{Z}_2$ " or QM/ $\mathbb{Z}_2$  is also "quantum mechanics over sets" or QM/Sets. Thus QM/ $\mathbb{Z}_2$  and QM/Sets will be used interchangeably.

There have been several ([52], [33], and [61]) earlier attempts to build a model of QM with  $\mathbb{C}$  replaced by  $\mathbb{Z}_2$ . Since there are no inner products on vectors spaces over finite fields, the main problem is how to define the Dirac brackets which ordinarily take values in the base field. Building the 'right' model of QM over  $\mathbb{Z}_2$  means making the 'best' decisions as to what features to retain or not retain from QM/ $\mathbb{C}$ . The previous attempts retained the feature of the Dirac brackets taking values in  $\mathbb{Z}_2$  so they could only have the values of 0 and 1—which excludes a non-trivial probability calculus. For instance, the best-known Schumacher-Westmoreland model does "not make use of the idea of probability" [52, p. 919] and have instead only a modal interpretation ( $1 =$  possibility and  $0 =$  impossibility). There is a fourth category-theoretic model where the objects are sets [1] but it also has the "brackets" taking 0, 1 values.<sup>2</sup>

With so many attempts to build a 'toy model' over the simplest field  $\mathbb{Z}_2$ , how can one judge whether the right decisions have been made about "what to leave in, what to leave out"? In the case of this model, the answer is very simple. QM/Sets is also an extension of finite probability theory to include superposition events in addition to classical events where the vector space  $\mathbb{Z}_2^n$  has multiples bases and non-commutative

---

<sup>2</sup>The Spekkens toy model [56] does not use vector spaces at all or utilize sets so it is not directly comparable.

random variables ('observables'). This relates the quantum mechanical probability machinery of QM/C back to simpler and more familiar concepts. In ordinary Laplace-Boole finite probability theory, the finite universe set  $U$  is the set of equiprobable outcomes, subsets  $S \subseteq U$  as events, and the probability of an event is  $\Pr(S) = \frac{|S|}{|U|}$ , the normalized cardinality of  $S$ . Given that an event  $S$  occurs, the probability that another event  $T \subseteq U$  occurs is the conditional probability:  $\Pr(T|S) = \frac{|T \cap S|}{|S|}$ . Given a real-valued random variable on  $U$ ,  $f : U \rightarrow \mathbb{R}$ , the probability of getting the value  $r \in f(U)$  given  $S$  is:  $\Pr(r|S) = \frac{|f^{-1}(r) \cap S|}{|S|}$ .

In the non-commutative extension of finite probability theory, each subset  $S \subseteq U = \{u_1, \dots, u_n\}$  of the outcome set (or sample space) defines two types of events, the usual classical event  $S$  and the superposition events [23] associated with  $S$ . Intuitively, the classical event  $S$  treats all the outcomes in  $S$  as being fully distinguished and each with a probability  $\frac{1}{|S|}$  in the case of equiprobable outcomes or  $\frac{p_i}{\sum_{u_j \in S} p_j} = \frac{p_i}{\Pr(S)}$  in the case of point probabilities  $p = (p_1, \dots, p_n)$ . The superposition event  $S$  is a  $\mathbb{Z}_2$ -version of the pure superposition state in QM/C that treats the elements in  $S$  as cohering together and being undifferentiated from each other. The best way to clearly differentiate the two types of events is to use density matrices as will be eventually done. But before bringing in that level of complication, it suffices to borrow the differentiation between "pure" states and "mixed" states. The *superposition event*  $S \in \wp(U)$  is a vector in the vector space  $\wp(U) \cong \mathbb{Z}_2^n$  and is called a *pure* state while the *classical event*  $S \subseteq U$  is a *mixed* state that is a lottery-like probability-weighted (convex) combination or mixture with weights  $\Pr(\{u_i\})$  of the pure outcomes  $\{u_i\} \subseteq U$ . In short, a pure state is a vector in a vector space  $\wp(U)$  over  $\mathbb{Z}_2$  and a mixed state is a probability-weighted set of pure states (which is thus not an element in the vector space).

We can foreshadow the differentiation between pure states and mixed states using density matrices by invoking incidence matrices. The *incidence matrix* of a binary relation  $R \subseteq U \times U$  is the  $n \times n$  matrix  $In(R)$  where  $In(R)_{i,j} = 1$  if  $(u_i, u_j) \in R$  and  $= 0$  otherwise. Then the superposition event  $S \in \wp(U)$  would be associated with the matrix  $In(S \times S)$  while the classical event  $S$  would be associated with the incidence matrix for the diagonal  $\delta S = \{(u_i, u_i) : u_i \in S\}$ . For instance, if  $U = \{a, b, c\}$  and  $S = \{a, b\}$ , then:

$$\text{In}(S \times S) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \text{In}(\delta S) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

With equiprobable outcomes, the corresponding density matrices would be obtained by dividing through by  $\frac{1}{|S|}$ .

One of the interesting results is that the superposition event  $S \in \wp(U)$  and the classical event  $S \subseteq U$  cannot be distinguished by any probabilities calculated using equiprobable outcomes or point-probabilities on  $U$ . At first, this may seem like a problem in the model of non-commutative finite probability theory or QM/Sets, but it is a feature rather than a bug since the corresponding fact holds in QM/C [3, p. 176]. One can only differentiate the states by measuring in a different basis in QM/C, and the same holds in QM/Sets using a different basis set for  $\mathbb{Z}_2^n$ .

Unless otherwise specified, we will take the notation  $S, T \in \wp(U)$  to refer to the pure superposition events—even though the same probabilistic formulas will hold for the corresponding mixed classical events  $S, T \subseteq U$ .

One of the main mathematical notions that will be used throughout is the notion of a partition [19]. A *partition*  $\pi = \{B_1, \dots, B_m\}$  or just  $\pi = \{B\}_{B \in \pi}$  on the universe set  $U$  is a set of disjoint nonempty subsets of  $U$  whose union is  $U$ .<sup>3</sup> Each *block*  $B \in \pi$  can be viewed as a pure state  $B \in \wp(U)$  and the whole partition can be associated with the mixed state of the blocks each with probability  $\text{Pr}(B)$ .

## 1.2 Kets, bra-kets, Dirac brackets, and the norm in QM/Sets

Given a universe set  $U = \{a, b, c\}$ , the  $U$ -basis in  $\wp(U)$  is the set of singletons:  $\{\{a\}, \{b\}, \{c\}\}$ . The  $U$ -basis will be taken as the computational basis by identifying those singletons respectively with the standard basis  $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$  in  $\mathbb{Z}_2^3$ . But  $\{a, b\} = (1, 1, 0)$ ,  $\{b, c\} = (0, 1, 1)$ , and  $\{a, b, c\} = (1, 1, 1)$  also form a basis set since:

$$\begin{aligned} \{b, c\} + \{a, b, c\} &= \{a\}; \\ \{b, c\} + \{a, b\} + \{a, b, c\} &= \{b\}; \text{ and} \\ \{a, b\} + \{a, b, c\} &= \{c\}. \end{aligned}$$

---

<sup>3</sup>For the sake of completeness, the empty partition  $\emptyset$  might be mentioned which is the inverse-image of the empty function  $\emptyset \rightarrow X$  for any set  $X$ , but for our purposes, we will ignore it.

That different basis could be considered as the singletons in an equicardinal universe set  $U' = \{a', b', c'\}$  where  $\{a'\} = \{a, b\} = (1, 1, 0)$ ,  $\{b'\} = \{b, c\} = (0, 1, 1)$ , and  $\{c'\} = \{a, b, c\} = (1, 1, 1)$ . One important thing to understand is that, say,  $\{a'\} = \{a, b\}$  represent the same abstract vector expressed in two different bases. The following table represents the  $2^3 = 8$  vectors in  $\mathbb{Z}_2^3$  expressed in three different bases. Since an abstract vector is often called a "ket" in Dirac's treatment of quantum mechanics [16], the rows in Table 1.2, a "ket table," represent three different representations of the corresponding abstract vector.

$\mathbb{Z}_2^3$	$U = \{a, b, c\}$	$U' = \{a', b', c'\}$	$U'' = \{a'', b'', c''\}$
(1, 1, 1)	$\{a, b, c\}$	$\{c'\}$	$\{a'', b'', c''\}$
(1, 1, 0)	$\{a, b\}$	$\{a'\}$	$\{b''\}$
(0, 1, 1)	$\{b, c\}$	$\{b'\}$	$\{b'', c''\}$
(1, 0, 1)	$\{a, c\}$	$\{a', b'\}$	$\{c''\}$
(1, 0, 0)	$\{a\}$	$\{b', c'\}$	$\{a''\}$
(0, 1, 0)	$\{b\}$	$\{a', b', c'\}$	$\{a'', b''\}$
(0, 0, 1)	$\{c\}$	$\{a', c'\}$	$\{a'', c''\}$
(0, 0, 0)	$\emptyset$	$\emptyset$	$\emptyset$

Table 1.2: Ket table giving a vector space isomorphism:

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

How many different bases are there in  $\mathbb{Z}_2^3$ ? Gauss's formula for the number of ordered bases in  $\mathbb{Z}_2^n$  is  $(2^n - 1)(2^n - 2)(2^n - 2^2) \dots (2^n - 2^{n-1})$  [40, p. 71] and the number of different unordered bases is obtained by dividing by  $n!$ . In the case of  $n = 3$ , there are  $(2^3 - 1) = 7$  choices for a non-zero vector to generate a one-dimensional subspace with 2 elements. Then there are another  $(2^3 - 2) = 6$  choices to generate a two-dimensional subspace with 4 elements. And finally, there are  $(2^3 - 4) = 4$  choices for the elements to generate the whole space. Hence there are  $(2^3 - 1)(2^3 - 2)(2^3 - 4) = 7 \times 6 \times 4$  ordered bases of  $\mathbb{Z}_2^3$  so dividing by  $3! = 6$  gives  $7 \times 4 = 28$  different bases for  $\mathbb{Z}_2^3$ , three of which are listed in Table 1.2.

In Dirac's ket notation, the abstract vector expressed as  $\{a, b\}$  in the  $U$ -basis would be denoted:  $|\{a, b\}\rangle$  or for  $S \in \wp(U)$ , just as  $|S\rangle$ . Since a row of the ket table gives the different representations of the *same* ket in the different bases, we have, for instance:  $|\{a, c\}\rangle = |\{a', b'\}\rangle = |\{c''\}\rangle$ . It should be noted that computations can be carried on in any basis and will get the same result. For instance, in the  $U$ -basis,

$|\{a, c\}\rangle + |\{b, c\}\rangle = |\{a, b\}\rangle$  (with the cancellation of  $\{c\}$ ) which in the  $U'$ -basis is:  $|\{a', b'\}\rangle + |\{b'\}\rangle = |\{a'\}\rangle$  (with the cancellation of  $\{b'\}$ ) since  $|\{a, b\}\rangle = |\{a'\}\rangle$ .

In QM/ $\mathbb{C}$ , the Dirac brackets are defined as the inner product on  $\mathbb{C}$  which takes values in the ordered subfield  $\mathbb{R}$ . But there is no inner product over finite fields since the positivity requirement  $\langle v, v \rangle > 0$  for  $v \neq 0$  requires an ordered subfield such that adding positive elements preserves the ordering. But in any field of characteristic  $p$  (a prime), adding 1 to each side of  $0 < 1$  will eventually lead to  $p - 2 < p - 1$  so adding another 1 yields  $p - 1 < p = 0$ , so there cannot be such an ordering. Hence the need to model QM over  $\mathbb{Z}_2$  requires another definition of the Dirac brackets than an inner product. One has to decide "what to leave in, what to leave out" from QM/ $\mathbb{C}$ . For instance, the Schumacher-Westmoreland model of QM with the base field  $\mathbb{Z}_2$  decided to retain the feature that the Dirac brackets take values in the base field so instead of probabilities, they only had the modal values of 0 as impossible and 1 as possible [52]. A different tact is taken in QM/Sets. In QM/ $\mathbb{C}$ , the Dirac bracket  $\langle \psi | \phi \rangle$  for normalized states  $\psi$  and  $\phi$  is called the "overlap" where the minimal overlap is  $\langle \psi | \phi \rangle = 0$  for states that are orthogonal (or disjoint) and maximal overlap is  $\langle \psi | \phi \rangle = 1$  when they are the same state. And in QM/Sets, there is an obvious notion of overlap, the cardinality of the intersection, that takes its values in the natural numbers  $\mathbb{N}$ . That is, for  $S, T \in \wp(U)$ :

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

The idea of a function of vectors in  $\mathbb{Z}_2^n$  taking values outside of  $\mathbb{Z}_2$  is hardly new. The natural-number-valued *Hamming distance function*  $d_H(S, T) = |S + T|$  on vector spaces over  $\mathbb{Z}_2$  is used in coding theory [46] and logical information theory [25] so that for  $S, T \in \wp(U)$ :

$$d_H(S, T) = |S + T| = |S \cup T| - |S \cap T| = |S - T| + |T - S|.$$

The ket  $|T\rangle$  denotes the ket of  $T \in \wp(U)$  and is in that sense basis-independent, but the 'bra'  $\langle S |_U$  must be taken as basis-dependent as indicated by the subscript  $U$ . And the intersection  $S \cap T$  requires that both  $S$  and  $T$  be subsets of  $U$ . Moreover the brackets or bra-kets are not necessarily linear since for  $T' \in \wp(U)$ ,  $\langle S |_U T + T' \rangle = |S \cap (T + T')| = |S \cap (T \cup T' - T \cap T')|$  which is not equal to  $\langle S |_U T \rangle + \langle S |_U T' \rangle$  unless  $T \cap T' = \emptyset$ , i.e.,  $T$  and  $T'$  are disjoint.



For a general finite universe set  $U$  with the disjoint basis  $\{\{u\}\}_{u \in U}$ , the ket-bra operator  $|\{u\}\rangle \langle \{u\}|_U () : \wp(U) \rightarrow \wp(U)$  takes  $|S\rangle$  to  $|\{u\}\rangle \langle \{u\}|_U |S\rangle = |\{u\} \cap S| |u\rangle$  which is  $\{u\}$  if  $u \in S$  and 0 otherwise. Hence  $|\{u\}\rangle \langle \{u\}|_U ()$  is the projection operator  $\{u\} \cap () : \wp(U) \rightarrow \wp(U)$ . Then the sum of these projection operators over the whole  $U$ -basis is the identity operator:

$$\sum_{u \in U} |\{u\}\rangle \langle \{u\}|_U () = I() : \wp(U) \rightarrow \wp(U).$$

In QM/ $\mathbb{C}$ , given an orthonormal (ON) basis  $\{v_i\}_{i=1}^n$  of the Hilbert space  $V$ , the sum of the ket-bra projection operators is also the identity operator:

$$\sum_{i=1}^n |v_i\rangle \langle v_i| = I : V \rightarrow V$$

Completeness of the ket-bra sum.

Since  $\langle \{u\} |_U S\rangle = \langle S |_U \{u\}\rangle = |S \cap \{u\}| = \chi_S(u)$ , any bra-ket  $\langle S |_U T\rangle$  can be resolved using the ket-bra sum:

$$\sum_{u \in U} \langle S |_U \{u\}\rangle \langle \{u\} |_U T\rangle = \sum_{u \in U} \chi_S(u) \chi_T(u) = |S \cap T| = \langle S |_U T\rangle$$

which is the QM/ $\mathbb{Z}_2$  version of the QM/ $\mathbb{C}$ :

$$\langle \psi | \varphi \rangle = \sum_i \langle \psi | v_i \rangle \langle v_i | \varphi \rangle$$

Resolution of unity by ket-bra sum.

In QM/ $\mathbb{C}$ , the magnitude or norm of a vector  $\psi$  is often denoted  $|\psi| = \sqrt{\langle \psi | \psi \rangle}$  but that conflicts with our notation  $|S|$  for cardinality, so we will use  $\|\psi\| = \sqrt{\langle \psi | \psi \rangle}$  for the norm in QM/ $\mathbb{C}$  and the corresponding norm in QM/ $\mathbb{Z}_2$  is:

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

Norm in QM/ $\mathbb{Z}_2$

which takes values in the reals  $\mathbb{R}$ . Applied to the resolution of unity:

$$\|S\|_U^2 = \langle S |_U S \rangle = \sum_{u \in U} \langle S |_U \{u\}\rangle \langle \{u\} |_U S \rangle = |S|$$

which in QM/ $\mathbb{C}$  is:

$$\|\psi\|^2 = \langle \psi | \psi \rangle = \sum_i \langle \psi | v_i \rangle \langle v_i | \psi \rangle = \sum_i \langle v_i | \psi \rangle^* \langle v_i | \psi \rangle$$

where  $\langle v_i|\psi\rangle^* = \langle \psi|v_i\rangle$  is the complex conjugate of  $\langle v_i|\psi\rangle$ .

In QM/ $\mathbb{C}$ , a vector can be normalized at any time, but in QM/ $\mathbb{Z}_2$ , normalization is only done when probabilities are computed so to better draw out the analogies, we will not necessarily assume a vector  $\psi$  is normalized. When a state  $\psi$  is measured in the (non-degenerate) measurement basis  $\{v_i\}$ , then the probability of obtaining  $v_i$  is given by the Born Rule:

$$\Pr(v_i|\psi) = \frac{\|\langle v_i|\psi\rangle\|^2}{\|\psi\|^2}$$

and the corresponding ‘Born Rule’ formula in QM/ $\mathbb{Z}_2$  is:

$$\Pr(u|_U S) = \frac{\|\langle \{u\}|_U S\rangle\|_U^2}{\|S\|_U^2} = \frac{|\{u\} \cap S|}{|S|} = \begin{cases} 1/|S| & \text{if } u \in S \\ 0 & \text{if } u \notin S \end{cases}.$$

Thus the QM/ $\mathbb{Z}_2$  probabilities just reproduce the usual finite probabilities for an outcome  $u$  to occur conditioned on the classical event  $S \subseteq U$  with an equiprobable outcome set  $U$  or the probability of getting the outcome  $u$  when ‘measuring’ the superposition state  $S \in \wp(U)$  using the  $U$ -basis.

### 1.3 Numerical attributes and random variables in QM/Sets

The role of observables, i.e., Hermitian (or self-adjoint) operators, in QM/ $\mathbb{C}$  is taken in QM/ $\mathbb{Z}_2$  by real-valued attributes  $f : U \rightarrow \mathbb{R}$  on a basis set for  $\mathbb{Z}_2^n$  which is also just a real-valued random variable (r.v.) on the Laplacian (equiprobable) outcome space  $U$ . The range for the numerical attribute or random variable could also just be the integers or natural numbers; it doesn’t matter in QM/ $\mathbb{Z}_2$  any more than it matter in finite probability theory. The point is that these r.v.’s are not restricted to having values in the base space  $\mathbb{Z}_2$ . An attribute that did take values in  $\mathbb{Z}_2$  would just be the characteristic function for some subset  $S \in \wp(U)$ ,  $\chi_S : U \rightarrow \mathbb{Z}_2$  and thus it would define a linear operator on  $\mathbb{Z}_2^n$  which is just the projection operator  $S \cap () : \wp(U) \rightarrow \wp(U)$ . Thus the analogues of the QM/ $\mathbb{C}$  observables in QM/Sets are the r.v.s  $f : U \rightarrow \mathbb{R}$  which are much more general than the linear operators on  $\mathbb{Z}_2^n$  (which are just the projection operators).

Since the observables in QM/ $\mathbb{C}$  are operators, we have to define the QM/Sets analogues of eigenvectors, eigenvalues, and eigenspaces for the general r.v.s  $f : U \rightarrow \mathbb{R}$ . Given a Hermitian operator  $F : V \rightarrow V$ , the eigenvectors and eigenvalues of  $F$  are the vectors  $v \in V$  and reals  $\lambda \in \mathbb{R}$  that satisfy the eigenvector equation:

$$Fv = \lambda v.$$

For  $r \in \mathbb{R}$  and  $S \in \wp(U)$ , let  $rS$  "formally"<sup>4</sup> stand for the result of assigning  $r$  to the elements of  $S$ . For  $f : U \rightarrow \mathbb{R}$  and  $f \upharpoonright S$  being the restriction of  $f$  to  $S$ , then the corresponding eigenvector equation in QM/Sets is:

$$f \upharpoonright S = rS$$

Eigenvector equation in QM/Sets.

Thus an *eigenvalue* of  $f$  is any  $r$  that satisfies  $f \upharpoonright S = rS$  for some  $S \in \wp(U)$ , and an *eigenvector* of  $f$  is any  $S \in \wp(U)$  that satisfies  $f \upharpoonright S = rS$  for some  $r$ . Thus the eigenvectors of  $f$  are just the subsets  $S \in \wp(U)$  where  $f$  is constant, the level-sets of  $f$ , and the eigenvalues are those constant values, i.e., the elements of the image  $f(U) \subseteq \mathbb{R}$ . The *eigenspace* of  $f$  for an *eigenvalue*  $r$  is the subspace  $\wp(f^{-1}(r)) \subseteq \wp(U)$  of all the subsets of  $U$  where  $f$  has the constant value  $r$ , just as in QM/C, the eigenspace  $V_\lambda$  is the subspace of all eigenvectors of  $F$  with the eigenvalue  $\lambda$ . In QM/C, the eigenspaces  $V_\lambda$  and  $V_{\lambda'}$  for  $\lambda \neq \lambda'$  are orthogonal or disjoint in the sense that the only common vector is the zero vector, and the same holds for eigenvalues  $r \neq r'$  in QM/Sets:  $\wp(f^{-1}(r)) \cap \wp(f^{-1}(r')) = \{\emptyset\}$ . If  $P_\lambda : V \rightarrow V$  is the projection operator to the eigenspace  $V_\lambda$ , then the completeness of those projection operators is expressed by:

$$\sum_\lambda P_\lambda = I : V \rightarrow V$$

(sum is over the eigenvalues  $\lambda$  of  $F$ ). Similarly in QM/Sets, let  $P_r = f^{-1}(r) \cap () : \wp(U) \rightarrow \wp(U)$  be the projection operator to the eigenspace  $\wp(f^{-1}(r))$  and then the sum of the projection operators to the eigenspaces is the identity as illustrated in Figure 1.1:

$$\sum_{r \in f(U)} f^{-1}(r) \cap () = I() : \wp(U) \rightarrow \wp(U).$$

---

<sup>4</sup>We say "formally" since the operation of multiplying a real number times a set is not defined so we interpret it as the result of defining a function on  $S \in \wp(U)$  as having the constant value  $r$ .

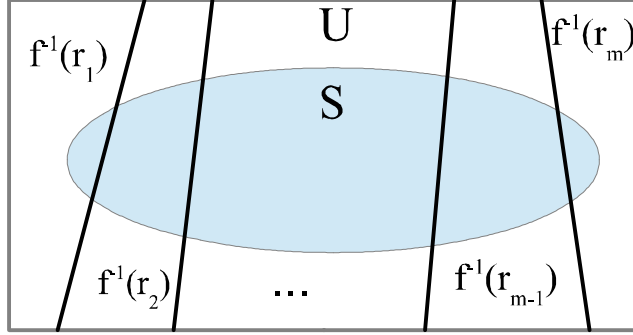


Figure 1.1: Subset  $S$  expressed as union over  $r \in f(U)$  of disjoint intersections  $f^{-1}(r) \cap S$ .

In QM/ $\mathbb{C}$ , the eigenspaces of  $F$  form a direct-sum decomposition of  $V$ , symbolized:

$$V = \sum_{\lambda} \oplus V_{\lambda}$$

which means that every vector  $v \in V$  can be uniquely expressed as a sum of vectors  $v_{\lambda} \in V_{\lambda}$  from the eigenspaces. Reasoning in terms of sets, a r.v.  $f : U \rightarrow \mathbb{R}$  induces the inverse-image partition  $\{f^{-1}(r)\}_{r \in f(U)}$  on  $U$  and any subset  $S$  can be uniquely expressed as the union of (disjoint) subsets where  $f$  is constant:  $S = \bigcup_{r \in f(U)} f^{-1}(r) \cap S$ . This fact translates directly into  $\wp(U)$  being the direct-sum decomposition of the eigenspaces of  $f$ :

$$\wp(U) = \sum_{r \in f(U)} \oplus \wp(f^{-1}(r)).$$

Thus we see how the whole machinery of eigenvectors, eigenvalues, eigenspaces, and direct-sum decompositions of eigenspaces are developed in QM/Sets for numerical attributes or r.v.s—even though the only actual operators are the projections.

To approach the probability calculus for r.v.s  $f : U \rightarrow \mathbb{R}$ , the QM/ $\mathbb{C}$  equation:  $\|\psi\|^2 = \langle \psi | \psi \rangle = \sum_{\lambda} \|P_{\lambda}(\psi)\|^2$  is expressed in QM/Sets as:  $\|S\|_U^2 = \langle S |_U S \rangle = \sum_r \|f^{-1}(r) \cap S\|_U^2 = \sum_r |f^{-1}(r) \cap S| = |S|$ . Then we normalize to have probabilities that sum to one:  $\sum_{\lambda} \frac{\|P_{\lambda}(\psi)\|^2}{\|\psi\|^2} = 1$  for  $\psi \neq 0$  and  $\sum_{r \in f(U)} \frac{\|f^{-1}(r) \cap S\|_U^2}{\|S\|_U^2} = \sum_r \frac{|f^{-1}(r) \cap S|}{|S|} = 1$  for  $S \neq \emptyset$ . Then when measuring  $\psi$  by the observable  $F$ , the Born Rule probability of getting the eigenvalue  $\lambda$  is:

$$\Pr(\lambda | \psi) = \frac{\|P_{\lambda}(\psi)\|^2}{\|\psi\|^2}$$

and the corresponding probability for getting the eigenvalue  $r$  of the r.v.  $f$  when conditioned by  $S$  is:

$$\Pr(r|S) = \frac{\|f^{-1}(r) \cap S\|_U^2}{\|S\|_U^2} = \frac{|f^{-1}(r) \cap S|}{|S|}.$$

Again we see that with a fixed basis  $U$ , the probability calculus of QM/Sets is just the ordinary finite probability calculus for values of a random variable  $f$  on the equiprobable outcome set  $U$ —except that  $S$  can be a superposition event  $S$  in the vector space  $\wp(U)$  or a classical event  $S$  which is just a subset of  $U$ .

For example, consider the roll of a pair of fair dice so the equiprobable outcome space is  $U = \{1, \dots, 6\} \times \{1, \dots, 6\}$ . The numerical attribute or r.v. is  $f : U \rightarrow \mathbb{N}$  where  $f(d_1, d_2) = d_1 + d_2$  (the sum of the two dies faces showing up). Then  $f^{-1}(7) = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$  so for any conditioning event  $S \in \wp(U)$ , the probability of getting a seven is:  $\Pr(7|_U S) = \frac{\|f^{-1}(7) \cap S\|_U^2}{\|S\|_U^2} = \frac{|f^{-1}(7) \cap S|}{|S|}$ . If  $S$  was the set of outcomes where at least one die came up with a 1, then  $|S| = 11$ , and  $\Pr(7|_U S) = \frac{|f^{-1}(7) \cap S|}{|S|} = \frac{2}{11}$ .

For the attribute  $f : U \rightarrow \mathbb{R}$ , the spectral decomposition of  $f$  is formally  $f \upharpoonright () = \sum_r r [f^{-1}(r) \cap ()]$ , so we have:

$$\langle S|_U f \upharpoonright ()|S \rangle = \langle S|_U \sum_r r f^{-1}(r) \cap ()|S \rangle = \sum_r r \langle S|_U f^{-1}(r) \cap S \rangle = \sum_r r |f^{-1}(r) \cap S|.$$

Then normalizing we have the average value of the r.v.  $f$  on the subset  $S \in \wp(U)$ :

$$\langle f \rangle_S = \frac{\langle S|_U f \upharpoonright ()|S \rangle}{\langle S|_U S \rangle} = \sum_r r \frac{|f^{-1}(r) \cap S|}{|S|} = \sum_r r \Pr(r|S) = \text{average of } f \text{ on } S.$$

The corresponding formula in QM/ $\mathbb{C}$  for the average value of the observable  $F$  when measuring the state  $\psi$  is:

$$\langle F \rangle_\psi = \frac{\langle \psi|F|\psi \rangle}{\langle \psi|\psi \rangle}.$$

## 1.4 The linearization methodology behind QM/Sets

There is a linearization methodology at work in developing QM/ $\mathbb{Z}_2$  concepts from QM/ $\mathbb{C}$  concepts. The linearization method develops the correlation between set concepts and vector-space concepts. For instance, if we apply a set concept to a basis

set of a vector space, then it will generate the corresponding vector-space concept. Table 1.3 gives a brief summary of some set concepts and the corresponding linearized vector-space concepts.

Set concept	Vector space concept
Cardinality	Dimension
Subset	Subspace
Partition	Direct-sum decomposition
Numerical attribute $f$	Linear operator $F$
Value $r$ of $f$	Eigenvalue $\lambda$ of $F$
Constant set of $f$	Eigenvector of $F$
Set of constant $r$ -sets $\wp(f^{-1}(r))$	Eigenspace of $\lambda$
Direct product of sets	Tensor product of spaces

Table 1.3: Linearization of set concepts to corresponding vector-space concepts.

The correspondence can be traversed in the other direction to arrive at the corresponding set concept starting with a vector-space concept of quantum mechanics. The methodology behind the development of QM/ $\mathbb{Z}_2$  is then to express the set concept in the vector-space setting of  $\mathbb{Z}_2^n$  where the vectors are interpreted as subsets of a universe set—but where there are different basis sets for  $\mathbb{Z}_2^n$  so many of the characteristic features of QM/ $\mathbb{C}$  can be modeled in QM/ $\mathbb{Z}_2$ . That is essentially how QM/ $\mathbb{Z}_2$  is generated. The Table 1.4 gives a brief summary of that correspondence between QM/ $\mathbb{Z}_2$  and QM/ $\mathbb{C}$ .

	QM/ $\mathbb{Z}_2$	QM/ $\mathbb{C}$
Bra-kets	$\langle S _U T \rangle =  S \cap T $	$\langle \psi \varphi \rangle = \text{overlap of } \psi, \varphi$
Resolution of unity	$\langle S _U T \rangle = \sum_{u \in U} \langle S _U \{u\} \rangle \langle \{u\} _U T \rangle$	$\langle \psi \varphi \rangle = \sum_i \langle \psi v_i \rangle \langle v_i \varphi \rangle$
Projection op.	$S \cap () : \wp(U) \rightarrow \wp(U)$	$P : V \rightarrow V$
Observable	$f : U \rightarrow \mathbb{R}$	$F : V \rightarrow V$
Eigenvalue Eqn.	$f \upharpoonright S = rS$	$Fv = \lambda v$
Eigenspaces	$\wp(f^{-1}(r))$ for eigenvalue $r$	$\{v : Fv = \lambda v\}$ for eigenvalue $\lambda$
Spectral Decomp.	$f = \sum_{r \in f(U)} r \chi_{f^{-1}(r)}$	$F = \sum_{\lambda} \lambda P_{\lambda}$
Completeness	$\sum_r f^{-1}(r) \cap () = I$	$\sum_{\lambda} P_{\lambda} = I$
Orthogonality	$r \neq r', f^{-1}(r) \cap f^{-1}(r') = \emptyset$	$\lambda \neq \lambda', P_{\lambda} P_{\lambda'} = 0$
Direct-Sum Decomp.	$\wp(U) = \sum_r \oplus \wp(f^{-1}(r))$	$V = \sum_{\lambda} \oplus V_{\lambda}$
Norm	$\ S\ _U = \sqrt{\langle S _U S \rangle} = \sqrt{ S }$	$\ \psi\  = \sqrt{\langle \psi \psi \rangle}$
Probability	$\text{Pr}(r S) = \frac{\ f^{-1}(r) \cap S\ _U^2}{\ S\ _U^2} = \frac{ f^{-1}(r) \cap S }{ S }$	$\text{Pr}(\lambda \psi) = \frac{\ P_{\lambda}(\psi)\ ^2}{\ \psi\ ^2}$
Average value	$\langle f \rangle_S = \frac{\langle S _U f \upharpoonright () S \rangle}{\langle S _U S \rangle}$	$\langle F \rangle_{\psi} = \frac{\langle \psi F \psi \rangle}{\langle \psi \psi \rangle}$

Table 1.4: Summary of analogues in QM/ $\mathbb{Z}_2$  and QM/ $\mathbb{C}$ .

## 2 Philosophical interlude: the Literal or Objective Indefiniteness Interpretation of QM

While QM/Sets can be viewed purely as a pedagogical model, it also may bear on the questions about the philosophical interpretation of QM/ $\mathbb{C}$ . Classical physics is compatible with the common-sense view of reality that might be referred to as "definite properties all the way down."

It is now rather widely accepted that this common-sense view of reality is not compatible with quantum mechanics. If we think in terms of only two positions, *here* and *there*, then in classical physics a particle is either definitely *here* or *there*, while in QM, the particle can be "neither definitely here nor there." [66, p. 144] This is usually misrepresented in the popular literature as the particle being "both *here* and *there* at the same time." This is not an epistemic or subjective indefiniteness of location; it is an ontological or objective indefiniteness. The notion of *objective indefiniteness* in QM has been most emphasized by Abner Shimony ([53], [54], [55]).

From these two basic ideas alone – indefiniteness and the superposition

principle – it should be clear already that quantum mechanics conflicts sharply with common sense. If the quantum state of a system is a complete description of the system, then a quantity that has an indefinite value in that quantum state is objectively indefinite; its value is not merely unknown by the scientist who seeks to describe the system. ...Classical physics did not conflict with common sense in these fundamental ways.[53, p. 47]

These statements ... may collectively be called "the Literal Interpretation" of quantum mechanics. This is the interpretation resulting from taking the formalism of quantum mechanics literally, as giving a representation of physical properties themselves, rather than of human knowledge of them, and by taking this representation to be complete. [55, pp. 6-7]

Peter Mittelstaedt has also emphasized blurred, unsharp, or "incompletely determined" [47, p. 171] quantum states, and Shimony's and Mittelstaedt's students have continued this line of approach [8].<sup>5</sup>

Other philosophers of physics have suggested related ideas: inherent indefiniteness [28, p. 202], value indefiniteness [57], ontic vagueness [41], indeterminate properties [36], quantum vagueness [31], and so forth. Indeed the idea that a quantum state is in some sense "blurred" or "like a cloud" is now rather commonplace even in the popular literature. It is the standard view that a description of a superposition quantum state is a *complete* description, which implies that the indefiniteness of a superposition state is in some sense objective.

In QM/Sets, the fully distinct elements  $\{u\} \in \wp(U)$  might be called "eigen-elements" and a subset  $S$  is a "superposition" of the eigen-elements  $\{u\} \in \wp(S)$  –thinking of the collecting together  $\{u, u', \dots\} = S$  of the elements of  $S$  as their "superposition"–which is their vector addition when we interpret  $\wp(U)$  as a vector space over  $\mathbb{Z}_2$ . With distinctions, the indistinct element  $S$  might be refined into one of the eigen-elements  $\{u\} \in \wp(S)$ .

Abner Shimony as well as Shimon Malin [44] in the description of a superposition state as being objectively indefinite, sometimes adopted Heisenberg's [34] language of "potentiality" and "actuality" to describe the relationship of the eigenstates that are superposed to give an objectively indefinite superposition. This terminology is,

---

<sup>5</sup>See also the discussion by Falkenburg [26].



however, somewhat misleading since the indefinite superposition state is perfectly actual; it is only the multiple eigenvectors in the superposition that are "potential" until "actualized" by some further distinctions. In a measurement, an actual indefinite superposition state becomes an actual more-definite state. Since the measurement goes from actual indefinite state to actual more-definite state, the potential-to-actual language of Heisenberg should be interpreted as indefinite-to-more-definite. Consider a three-element universe  $U = \{a, b, c\}$  in QM/Sets and a partition  $\pi = \{\{a\}, \{b, c\}\}$ . The block  $S = \{b, c\}$  is objectively indefinite between  $\{b\}$  and  $\{c\}$  so those singletons are its potential more-definite states in the sense that a distinction could sharpen  $\{b, c\}$  to either  $\{b\}$  or  $\{c\}$ .

Note that this objective indefiniteness is not well-described as saying that indefinite pre-distinction superposition  $\{b, c\}$  is "simultaneously both a definite  $\{b\}$  and a definite  $\{c\}$ "; instead it is indefinite between  $\{b\}$  and  $\{c\}$ . That is, a superposition should *not* be thought of like a double exposure photograph which has two fully definite images (e.g., simultaneously a picture of say  $\{b\}$  and  $\{c\}$ ). That imagery is a holdover from classical wave imagery, e.g., in Fourier analysis, where definite eigen-waveforms are superposed to give a definite superposition waveform. That is the wrong imagery to understand QM. The definite superposition waveform should be thought of as an indefinite waveform that is indefinite between the eigen-waveforms that are superposed. The superposition of eigen-waveforms should be thought as the mathematical way to define indefiniteness between them rather than the way to define a new definite waveform. There is an alternative and better way to mathematically depict indefiniteness, namely density matrices (with their non-zero off-diagonal elements), which will be considered later. They can be defined over other fields, e.g., the rationals  $\mathbb{Q}$  in QM/Sets and thus avoid the misleading wave-imagery of QM/ $\mathbb{C}$ .

In terms of the photo imagery, instead of a double-exposure photograph, a superposition representation might be thought of as "a photograph of clouds or patches of fog." (Schrödinger quoted in: [30, p. 66]). Schrödinger distinguishes a "photograph of clouds" from a blurry photograph presumably because the latter might imply that it was only the photograph that was blurry while the underlying objective reality was sharp. The "photograph of clouds" imagery for a superposition connotes a clear and complete picture of an objectively "cloudy" or indefinite reality.

What about "wave-particle duality"? States that are indistinct for an observable are represented as weighted vector sums or superpositions of the eigenstates

that might be rendered more definite by further distinctions. This indistinctness-represented-as-superpositions is usually interpreted as "wave-like aspects" of the particles in the indefinite state. Hence the distinction-making measurements take away the indistinctness—which is usually interpreted as taking away the "wave-like aspects," i.e., "collapse of the wave packet." But there are no actual physical waves in quantum mechanics (e.g., the "wave amplitudes" are complex numbers); only particles with indefinite states for certain observables. Thus the "collapse of the wave packet" is better described as the "collapse of indefiniteness" to achieve more definiteness. And the "wave-particle duality" is actually the contrast between *particles* in indefinite or definite states.

Consider the standard double-slit experiment. When there is no distinction between the two slits, then the position attribute of the traversing particle is indefinite between top slit and bottom slit (*not* "going through both slits"), which is usually interpreted as the "wave-like aspects" that show interference. But when a distinction is made between the slits, e.g., inserting a detector in one slit or closing one slit, then the distinction reduces the indefiniteness to definiteness so the indefiniteness disappears, i.e., the "wave-like aspects" such as interference disappear.

What about quantum dynamics? What about the (time-dependent) Schrödinger wave equation? Since measurements, or, more generally, interactions between quantum systems may make distinctions, we might ask the following question. What is the evolution of a quantum system that is *isolated* so that no distinctions are made in the sense that the degree of indistinctness between state vectors is not changed? Two states  $\psi$  and  $\varphi$  in a Hilbert space are *fully distinct* if they are orthogonal, i.e.,  $\langle \psi | \varphi \rangle = 0$ . Two states are *fully indistinct* if  $\langle \psi | \varphi \rangle = 1$ . In between, the *degree of indistinctness* can be measured by the *overlap*  $\langle \psi | \varphi \rangle$ , the inner product of the state vectors. Hence the evolution of an isolated quantum system where the degree of indistinctness does not change is described by a linear transformation that preserves inner products, i.e., by a unitary transformation. Hence the partition-based distinctness-indistinctness analysis accounts for the unitary behavior of an isolated quantum system not undergoing any interaction that "could, *in principle*, distinguish the alternative *final* states" [29, p. 3-9] in a superposition.

The connection between unitary transformations and the solutions to the Schrödinger "wave" equation is given by Stone's Theorem [60]: there is a one-to-one correspondence between strongly continuous 1-parameter unitary groups  $\{U_t\}_{t \in \mathbb{R}}$  and Hermitian

operators  $H$  on the Hilbert space so that  $U_t = e^{iHt}$ .

In simplest terms, a unitary transformation describes a rotation such as the rotation of the unit vector in the complex plane as shown in Figure 2.1.

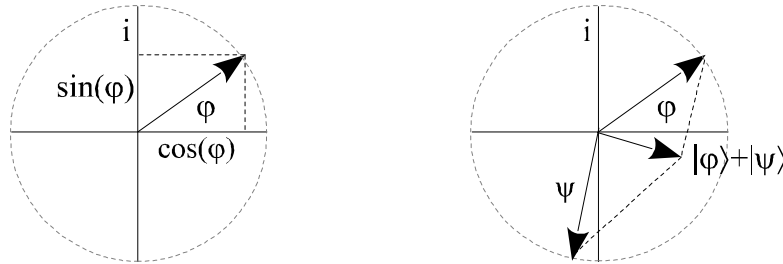


Figure 2.1: Rotating vector and addition of vectors.

The rotating unit vector traces out the cosine and sine functions on the two axes, and the position of the arrow can be compactly described as a function of  $\varphi$  using Euler's formula:

$$e^{i\varphi} = \cos(\varphi) + i \sin(\varphi).$$

Such complex exponentials and their superpositions are the "wave functions" of QM/ $\mathbb{C}$ . The "wave functions" describe the evolution of particles in indefinite states in isolated systems where there are no interactions to change the degree of indistinctness between states, i.e., the context where Schrödinger's equation holds. Classically it has been assumed that the mathematics of waves must describe physical waves of some sort, and thus the puzzlement about the "wave functions" of QM having complex amplitudes (in  $3N$ -dimensional space for systems of  $N$  particles) and no corresponding physical waves. But there is *another* interpretation; "wave" mathematics is the mathematics of indefiniteness-preserving evolution, e.g., superposition represents indefiniteness and unitary evolution represents the indistinctness-preserving evolution of an isolated system.

Thus the objective indefiniteness approach to interpreting QM/ $\mathbb{C}$  provides an explanation for the appearance of the wave mathematics (which implies interference as well as the quantized solutions to the "wave" equation that gave QM/ $\mathbb{C}$  its misleading name) when, in fact, there are no actual physical waves involved.

The role of the "waves" in ordinary quantum mechanics can be further clarified by viewing quantum dynamics in QM/Sets. The set version of a linear transformation

that takes an orthonormal basis to an orthonormal basis (i.e., a unitary transformation) is a linear transformation that takes a basis to a basis (i.e., a nonsingular transformation). Thus we can take a nonsingular transformation  $A : \wp(U) \rightarrow \wp(U)$  as a one period dynamic law in QM/Sets on the subsets of a set  $U$ .

In QM/ $\mathbb{C}$ , suppose the Hamiltonian  $H$  has an orthonormal basis of energy eigenstates  $\{|E_j\rangle\}$  so  $H|E_j\rangle = E_j|E_j\rangle$ . Then the application of the 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:

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

Thus  $U(t)$  transforms the orthonormal basis  $\{|E_j\rangle\}$  into the orthonormal basis  $\{|E'_j\rangle\} = \{e^{iE_j t} |E_j\rangle\}$ . 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  $\|c_j\|^2 = \|c_j e^{iE_j t}\|^2$ . 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  $|a_k\rangle$  (the summing of the rotating unit vectors in Figure 2.1). 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. Suppose we start with a state  $S \in \wp(U)$  for  $U = \{u_1, \dots, u_n\}$  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  $\langle u_j | U S \rangle = b_j \in \mathbb{Z}_2$ . Then the "dynamics" of a nonsingular transformation  $A : \mathbb{Z}_2^n \rightarrow \mathbb{Z}_2^n$  takes the basis  $\{|u_j\rangle\}$  to another basis  $\{|u'_j\rangle\}$  (where  $A|u_j\rangle = |u'_j\rangle$ ) which is the set version of  $U(t)$  taking the orthonormal basis  $\{|E_j\rangle\}$  to the orthonormal basis  $\{|E'_j\rangle\}$  where  $|E'_j\rangle = e^{iE_j t} |E_j\rangle$ . Thus  $|S\rangle$  is transformed into  $|S'\rangle = \sum_j b_j |u'_j\rangle$  with the same  $b_j$ 's so that  $\Pr(u_j|S) = \frac{b_j^2}{|S|} = \frac{b_j^2}{|S'|} = \Pr(u'_j|S')$  and  $\langle S | U T \rangle = \langle S' | U' T' \rangle$  (where for  $T \in \wp(U)$ ,  $A|T\rangle = |T'\rangle$  for some  $T' \in \wp(U')$ ). But the state  $|S'\rangle = \sum_j b_j |u'_j\rangle$  could be measured in another  $U''$ -basis  $\{|u''_j\rangle\}$  where  $|u'_j\rangle = \sum_k \alpha_k^j |u''_k\rangle$  so that:

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

Then under time evolution, there is interference in the coefficient  $\sum_j b_j \alpha_k^j$  of each eigenstate  $|u''_k\rangle$ . This suffices to give the interference phenomena that are ordinarily seen as characteristic of wave-like behavior but there is no mathematics of waves in QM/Sets.

The Table 2.1 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}  u'_j\rangle$	$ E_j\rangle \xrightarrow{U}  E'_j\rangle = e^{iE_j t}  E_j\rangle$
$ S\rangle = b_1  u_1\rangle + b_2  u_2\rangle \rightarrow b_1  u'_1\rangle + b_2  u'_2\rangle$	$ \psi_0\rangle = c_1  E_1\rangle + c_2  E_2\rangle \rightarrow c_1  E'_1\rangle + c_2  E'_2\rangle$
$ u'_j\rangle = \sum_k \langle u''_k   U^j  u'_j\rangle  u''_k\rangle = \sum_k \alpha_k^j  u''_k\rangle$	$ E_j\rangle = \sum_k \alpha_k^j  a_k\rangle;  E'_j\rangle = e^{iE_j t} \sum_k \alpha_k^j  a_k\rangle$
$b_1  u_1\rangle + b_2  u_2\rangle \rightarrow \sum_k (b_1 \alpha_k^1 + b_2 \alpha_k^2)  u''_k\rangle$	$c_1  E_1\rangle + c_2  E_2\rangle \rightarrow \sum_k (c_1 e^{iE_1 t} \alpha_k^1 + c_2 e^{iE_2 t} \alpha_k^2)  a_k\rangle$

Table 2.1: Showing the role in interference in QM/Sets and in QM/ $\mathbb{C}$

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 complex numbers  $\mathbb{C}$  are the algebraically complete extension of the reals where Hermitian operators will always find a complete set of eigenvectors, and the complex numbers also provide the natural mathematics to express the action of waves. Another context where the algebraic completeness of  $\mathbb{C}$  is important (and not particularly related to waves) is the role of group representations in vector spaces over  $\mathbb{C}$  for the irreducible representations that model elementary particles [59].

Thus QM/Sets allows us to separate the QM behavior due to superposition and interference apart from the specifically *wave-version* of that interference in QM/ $\mathbb{C}$ . 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/ $\mathbb{C}$  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.

### 3 Measurement and dynamics in QM/Sets

#### 3.1 Measurement and partitions

A partition (and their vector-space counterpart of direct-sum decompositions) is the mathematical concept to model distinctions and indistinctions, definiteness and indefiniteness, and distinguishability and indistinguishability. A *partition*  $\pi$  on a set  $U$  is a set of nonempty subsets  $\pi = \{B, B', \dots\}$  called "blocks" such that the blocks are disjoint and the union of the blocks is  $U$ . Each numerical attribute  $f : U \rightarrow \mathbb{R}$  induces the inverse-image partition  $f^{-1} = \{f^{-1}(r)\}_{r \in f(U)}$ . Thus a partition on  $U$  can be considered as an abstraction from a numerical attribute on  $U$  where we abstract away from the different values assigned to the blocks of the partition. Given another partition  $\sigma = \{C, C', \dots\}$  on  $U$ , the partition  $\sigma$  is *refined* by  $\pi$ , written  $\sigma \preceq \pi$ , if for every block  $B \in \pi$ , there is a block (necessarily unique)  $C \in \sigma$  such that  $B \subseteq C$ . Intuitively,  $\pi$  can be obtained from  $\sigma$  by chopping up some blocks of  $\sigma$  to obtain the blocks of  $\pi$ . Refinement is a partial order for the partitions on  $U$  meaning that it is reflexive, i.e.,  $\pi \preceq \pi$ , anti-symmetric, i.e.,  $\sigma \preceq \pi$  and  $\pi \preceq \sigma$  implies  $\sigma = \pi$ , and transitive, i.e., if  $\sigma \preceq \pi$  and  $\pi \preceq \tau$ , then  $\sigma \preceq \tau$ . The least refined partition is the *indiscrete partition*  $\mathbf{0}_U = \{\{U\}\}$  whose only block is  $U$  itself and the most refined partition is the *discrete partition*  $\mathbf{1}_U = \{\{u\}\}_{u \in U}$  whose blocks are the singletons of the elements of  $U$ . The partial order of refinement is picture in Figure 3.1 for  $U = \{a, b, c\}$ .

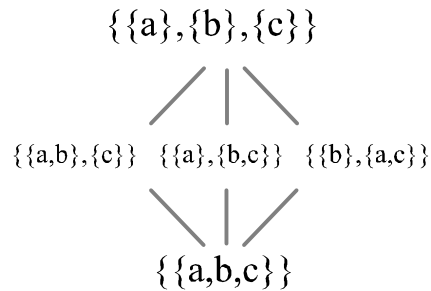


Figure 3.1: Partial order of partitions on  $U = \{a, b, c\}$ .

All the logical operations in the Boolean logic of subsets (usually presented as "propositional logic") can also be defined for partitions [19] but the only one we will use is the *join*  $\pi \vee \sigma$  which is the set of nonempty intersections  $B \cap C$  for  $B \in \pi$

and  $C \in \sigma$ . For instance, the join of  $\{\{a, b\}, \{c\}\}$  and  $\{\{a\}, \{b, c\}\}$  is the discrete partition  $\{\{a\}, \{b\}, \{c\}\} = \mathbf{1}_U$ .<sup>6</sup>

An ordered pair  $(u, u') \in U \times U$  is called a *distinction* or *dit* of a partition  $\pi$  if  $u$  and  $u'$  are in different blocks of  $\pi$ . The set of all distinctions of  $\pi$  is its *ditset*  $\text{dit}(\pi) \subseteq U \times U$ . An ordered pair  $(u, u')$  is called an *indistinction* or *indit* of  $\pi$  if  $u$  and  $u'$  are in the same block of  $\pi$ . The set of all indistinctions of  $\pi$  is its *inditset*  $\text{indit}(\pi) \subseteq U \times U$ . Each partition  $\pi$  defines an equivalence relation on  $U \times U$  where  $u \sim u'$  if they are in the same block of  $\pi$  so that equivalence relation is just the inditset  $\text{indit}(\pi)$ . Since each pair  $(u, u')$  are either in the same block or in different blocks of  $\pi$ , the ditset and inditset of  $\pi$  are complements in  $U \times U$ :  $\text{dit}(\pi) = U \times U - \text{indit}(\pi)$ . The ditsets of partitions on  $U$  might be called the *partition relations* on  $U \times U$ , but are also called *apartness relations* in computer science or constructive mathematics.

With the equiprobable assumption on the elements of  $U$ , each block in a partition  $\pi$  has a probability assigned to it,  $\Pr(B) = \frac{|B|}{|U|}$ . Since each block  $B \in \pi$  is a "state"  $B \in \wp(U)$ , the partition  $\pi$  can be seen as a probability weighed mixture of states in QM/Sets, and it corresponds to a *mixed state* in QM/C. The vectors in  $\wp(U)$ , e.g., the blocks  $B \in \pi$ , correspond to *pure states* in QM/C.

In QM/C, a measurement of a pure state superposition creates distinctions between eigenvectors that were previously superposed. In QM/Sets, let  $f : U \rightarrow \mathbb{R}$  be a random variable on  $U$  and consider a (pure) state  $S \in \wp(U)$ . The *measurement of  $S$  by the "observable"  $f$*  creates the partition on  $S$  consisting of the nonempty blocks  $f^{-1}(r) \cap S$  for  $r \in f(U)$  each with probability  $\Pr(r|S) = \frac{|f^{-1}(r) \cap S|}{|S|}$ . In this manner, a measurement of a pure state creates a mixed state. Prior to the measurement, the nonempty subsets  $f^{-1}(r) \cap S$  of  $S$  are the eigenvectors of  $f$  that are superposed in  $S$ , and the  $f$ -measurement distinguishes them. In terms of the partition join, the  $f$ -measurement on  $S$  is just the result on  $S$  of joining the partition  $\{S, S^c\}$  (where  $S^c$  is the complement of  $S$  in  $U$ ) with  $\{f^{-1}(r)\}_{r \in f(U)}$ .

One of the core issues in the philosophy of quantum mechanics is the nature of measurement. Thus it should be noted that in QM/Sets, the measurement of a pure state  $S \in \wp(U)$  by an observable  $f : U \rightarrow \mathbb{R}$  is 'formally' the same as the process in ordinary finite probability theory of sampling the r.v.  $f$  on the conditioning event  $S$  which returns a value  $r$  of the r.v. with probability  $\Pr(r|S)$ . When an  $f$ -

---

<sup>6</sup>In some of the older literature ([7]; [32]), the partial order is written the other way around so the join then becomes the meet.

value  $r$  is returned, then it means that the event  $f^{-1}(r) \cap S$  occurred so we could say that the event  $S$  projected to the event  $f^{-1}(r) \cap S$  by the projection operator  $f^{-1}(r) \cap () : \wp(U) \rightarrow \wp(U)$ . Thus these operations in finite probability theory just mimic (projective) measurement in QM/ $\mathbb{C}$ . [12, p. 221]

These analogies between ordinary finite probability theory and QM/ $\mathbb{C}$  are not entirely new. Hermann Weyl hinted at them in his informal description of quantum mechanics. Weyl referred to a partition as a "grating" or "sieve." Arthur Eddington made a very early use of the sieve idea:

In Einstein's theory of relativity the observer is a man who sets out in quest of truth armed with a measuring-rod. In quantum theory he sets out armed with a sieve.[17, p. 267]

This passage was quoted by Weyl [67, p. 255] in his treatment of gratings. Weyl then considered *both* set partitions and vector space partitions (direct-sum decompositions) as the respective types of gratings.[67, pp. 255-257] He started with a numerical attribute on a set, e.g.,  $f : U \rightarrow \mathbb{R}$ , which defined the set partition or "grating" [67, p. 255] with blocks having the same attribute-value, e.g.,  $\{f^{-1}(r)\}_{r \in f(U)}$ . Then he moved to the QM/ $\mathbb{C}$  case where the universe set, e.g.,  $U = \{u_1, \dots, u_n\}$ , or "aggregate of  $n$  states has to be replaced by an  $n$ -dimensional Euclidean vector space" [67, p. 256]. The appropriate notion of a vector space partition or "grating" is a "splitting of the total vector space into mutually orthogonal subspaces" so that "each vector  $\vec{x}$  splits into  $r$  component vectors lying in the several subspaces" [67, p. 256], i.e., a direct-sum decomposition of the space. After referring to a partition as a "grating" or "sieve," Weyl notes that "Measurement means application of a sieve or grating" [67, p. 259], e.g., in QM/Sets, the application (i.e., join) of the set-grating or partition  $\{f^{-1}(r)\}_{r \in f(U)}$  to the pure state  $\{S\}$  to give the mixed state  $\{f^{-1}(r) \cap S\}_{r \in f(U)}$ . In the pure state  $S$ , the  $f$ -eigenvectors  $f^{-1}(r) \cap S$  are superposed in an indefinite superposition but are distinguished (or "decohered") by the  $f$ -measurement to become more definite. This process might be illustrated as the result of an indefinite blob of dough falling through one of the differently shaped holes in a grating, and then taking on that shape as in Figure 3.2.



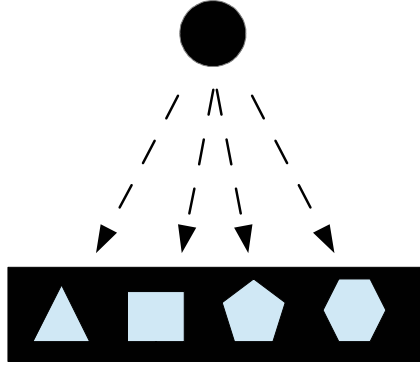


Figure 3.2: Measurement as passing an indefinite shape through a grating to acquire a definite shape.

A measurement occurs when a superposition state  $S$  interacts with something that makes distinctions—a partition  $\{f^{-1}(r)\}_{r \in f(U)}$ —that does not allow the superposed states to continue in that indefinite form so that only one of the superposed states  $f^{-1}(r) \cap S$  emerges from the measurement interaction.

**Example:** A measurement by a numerical attribute  $f$  is said to be *nondegenerate* if all the eigenspaces are one-dimensional. For  $U = \{a, b, c\}$ , let  $f : U \rightarrow \mathbb{R}$  be  $f(a) = 1$ ,  $f(b) = 2$ , and  $f(c) = 3$  so the one-dimensional eigenspaces are:  $\wp(f^{-1}(1)) = \{\emptyset, \{a\}\}$ ,  $\wp(f^{-1}(2)) = \{\emptyset, \{b\}\}$ , and  $\wp(f^{-1}(3)) = \{\emptyset, \{c\}\}$ . Then the measurement of  $S = U$  will give the mixed state of  $\{a\}$  with probability  $\Pr(\{a\} | U) = \frac{|\{a\} \cap U|}{|U|} = \frac{1}{3}$  and similarly for  $\{b\}$  and  $\{c\}$ . If the measurement or trial returns the eigenvalue of 3, then the resultant state is the projection  $f^{-1}(3) \cap U = \{c\}$  as illustrated in Figure 3.3.

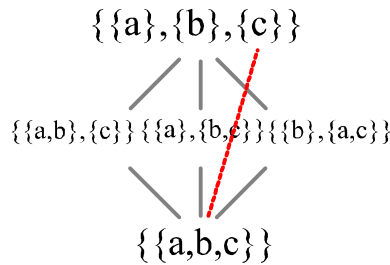


Figure 3.3: Nondegenerate measurement.

**Example:** Consider a simple example where we start with  $S = U = \{a, b, c\} \in \wp(U)$ . Regardless of the values of the r.v.s on  $U$ , there are essentially only four inverse-image partitions that refine the indiscrete partition as is seen in Figure 3.3. Consider

the observable, numerical attribute, or random variable  $\chi_{\{b,c\}} : U \rightarrow \mathbb{Z}_2$  whose inverse-image partition is  $\chi_{\{b,c\}}^{-1} = \{\{a\}, \{b, c\}\}$ . The two eigenspaces are  $\wp(\{a\}) = \{\emptyset, \{a\}\}$  and  $\wp(\{b, c\}) = \{\emptyset, \{b\}, \{c\}, \{b, c\}\}$  which form a direct-sum decomposition of  $\wp(U)$ , i.e., each subset of  $U$  is a sum of a unique pair of subsets, one from each eigenspace. The "measuring apparatus" makes distinctions by joining the attribute's inverse-image partition

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

with the pure state representing the indefinite entity  $U = \{a, b, c\}$  which yields the mixed state of  $\{a\}$  with probability  $\Pr(\{a\} | U) = \Pr(0, U) = \frac{|\{a\} \cap U|}{|U|} = \frac{1}{3}$  and  $\{b, c\}$  with probability  $\Pr(\{b, c\} | U) = \Pr(1 | U) = \frac{|\{b, c\} \cap U|}{|U|} = \frac{2}{3}$ . Suppose that the measurement or the trial in ordinary probability theory returns a 1 so the indefinite state  $\{a, b, c\}$  projects to the more definite state  $\chi_{\{b,c\}}^{-1}(1) \cap U = \{b, c\}$ . This is called a *degenerate* result since the eigenspace  $\wp(\chi_{\{b,c\}}^{-1}(1)) = \wp(\{b, c\}) = \{\emptyset, \{b\}, \{c\}, \{b, c\}\}$  doesn't have dimension one. Another measurement is needed to make more distinctions. Either  $\chi_{\{a,b\}}$  or  $\chi_{\{a,c\}}$  would suffice to decohere  $\{b, c\}$  into singletons with one-dimensional eigenspaces. Taking, for example, the other attribute as  $\chi_{\{a,b\}}$ , the join of the two attributes' partitions is discrete:

$$\chi_{\{b,c\}}^{-1} \vee \chi_{\{a,b\}}^{-1} = \{\{a\}, \{b, c\}\} \vee \{\{a, b\}, \{c\}\} = \{\{a\}, \{b\}, \{c\}\} = \mathbf{1}_U.$$

The second measurement turns the pure state  $\{b, c\}$  into the mixed state of  $\{b\}$  with probability  $\Pr(1 | \{b, c\}) = \frac{|\{a,b\} \cap \{b,c\}|}{|\{b,c\}|} = \frac{1}{2}$  and of  $\{c\}$  with the probability  $\Pr(0 | \{b, c\}) = \frac{|\{c\} \cap \{b,c\}|}{|\{b,c\}|} = \frac{1}{2}$ . Let us suppose that the measurement or trial returns a 0 so  $\{b, c\}$  projects to  $\chi_{\{a,b\}}^{-1}(0) \cap \{b, c\} = \{c\} \cap \{b, c\} = \{c\}$ . The results of these two measurements and outcomes is pictured in Figure 3.4.

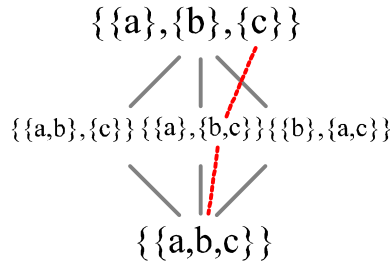


Figure 3.4: Results of the two measurements or trials.

Since the join  $\chi_{\{b,c\}}^{-1} \vee \chi_{\{a,b\}}^{-1} = \mathbf{1}_U$  was the discrete partition on  $U$ , those two attributes are said to be a "complete set", the QM/ $\mathbb{Z}_2$  version of Dirac's Complete Set of Commuting Operators or CSCOs [16], since they suffice to reduce  $U$  to singletons each with probability  $\frac{1}{3}$  as we found for the singleton  $\{c\}$  with probability  $\frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$ . Then each singleton can be characterized by the ordered pair of eigenvalues along the path, i.e.,  $|\{a\}\rangle = |0, 1\rangle$ ,  $|\{b\}\rangle = |1, 1\rangle$ , and  $|\{c\}\rangle = |1, 0\rangle$ .

### 3.2 Non-commutativity and the Indeterminacy Principle in QM/Sets

In finite probability theory with a fixed outcome space  $U$ , all random variables on  $U$  would have a definite value on any trial and in that sense they "commute." But we have reproduced a model of finite probability theory using a basis  $U$  of the vector space  $\mathbb{Z}_2^n$  which, however, has many more bases. By considering numerical attributes defined on different basis sets for  $\mathbb{Z}_2^n$ , we can reproduce the non-commutativity of operators in QM/ $\mathbb{C}$ .

Since we don't have linear operators on  $\mathbb{Z}_2^n$  aside from projection operators, the non-commutativity has to be defined using numerical attributes on different basis sets. Consider  $f : U \rightarrow \mathbb{R}$  with eigenvalues  $\{r\}_{r \in f(U)}$  and  $g : U' \rightarrow \mathbb{R}$  with eigenvalues  $\{s\}_{s \in g(U')}$  for two different basis sets  $U$  and  $U'$  for  $\mathbb{Z}_2^n$ . Each numerical attribute defines a set of eigenspaces  $\{f^{-1}(r)\}_{r \in f(U)}$  and  $\{g^{-1}(s)\}_{s \in g(U')}$  on the *same* space  $\mathbb{Z}_2^n$ . Given two partitions  $\pi$  and  $\sigma$  on the same set  $U$ , their *join*  $\pi \vee \sigma$  is the partition on  $U$  whose blocks are all the non-empty intersections  $B \cap C$  for blocks  $B \in \pi$  and  $C \in \sigma$  [19]. Taking the non-zero intersections gives a set of subspaces of the form  $f^{-1}(r) \cap g^{-1}(s)$  for some of the  $r \in f(U)$  and  $s \in g(U')$ . The sets of eigenspaces of  $f$  and of  $g$  each span the whole space  $\mathbb{Z}_2^n$ . The non-zero subspaces of the form  $f^{-1}(r) \cap g^{-1}(s)$  are the subspaces of simultaneous eigenvectors of  $f$  and  $g$  so let  $\mathcal{SE}$  be the subspace of  $\mathbb{Z}_2^n$  that they span.

In QM/ $\mathbb{C}$ , given two Hermitian operators  $F, G : V \rightarrow V$ , the space  $\mathcal{SE}$  of simultaneous operators can be constructed in the same way.

**Proposition 1**  $\mathcal{SE}$  is the kernel of the commutator:  $\mathcal{SE} = \ker([F, G]) = \ker(FG - GF)$ .

Proof: Let  $F, G : V \rightarrow V$  be two Hermitian operators on a finite dimensional vector space  $V$  and let  $v$  be a simultaneous eigenvector of the operators, i.e.,  $Fv = \lambda v$

and  $Gv = \mu v$ . Then  $[F, G](v) = (FG - GF)(v) = (\lambda\mu - \mu\lambda)v = 0$  so the space  $\mathcal{SE}$  spanned by the simultaneous eigenvectors is contained in the kernel  $\ker([F, G])$ , i.e.,  $\mathcal{SE} \subseteq \ker([F, G])$ . Conversely, if we restrict the two operators to the subspace  $\ker([F, G])$ , then the restricted operators commute on that subspace. Then it is a standard theorem of linear algebra [37, p. 177] that the subspace  $\ker([F, G])$  is spanned by simultaneous eigenvectors of the two restricted operators. But if a vector is a simultaneous eigenvector for the two operators restricted to a subspace, they are the same for the operators on the whole space  $V$ , since the two conditions  $Fv = \lambda v$  and  $Gv = \mu v$  only involves a vector in the subspace. Hence  $\ker([F, G]) \subseteq \mathcal{SE}$ .  $\square$

The operators commute if  $[F, G] = 0$  so the kernel of the commutator is the whole space  $V$ . This shows what the definitions should be in QM/Sets.

$$\begin{aligned} f, g \text{ commute or are compatible if } \mathcal{SE} &= \mathbb{Z}_2^n; \\ f, g \text{ are incompatible if } \mathcal{SE} &\neq \mathbb{Z}_2^n; \text{ and} \\ f, g \text{ are conjugate if } \mathcal{SE} &= \{\emptyset\}, \end{aligned}$$

*Example of incompatibility:* Let  $U = \{a, b, c\}$  and  $U' = \{a', b', c'\}$  where  $\{a'\} = \{a, b\}$ ,  $\{b'\} = \{b, c\}$ , and  $\{c'\} = \{a, b, c\}$  all as in Table 1.1 with the  $U$ -basis being the computational basis. For  $f : U \rightarrow \mathbb{R}$ , let  $f(a) = 1$ ;  $f(b) = f(c) = 2$ , so  $\wp(f^{-1}(1)) = \{\emptyset, \{a\}\}$  and  $\wp(f^{-1}(2)) = \{\emptyset, \{b\}, \{c\}, \{b, c\}\}$ . For  $g : U' \rightarrow \mathbb{R}$ , let  $g(a') = g(b') = 3$ ;  $g(c') = 5$ ; so  $\wp(g^{-1}(3)) = \{\emptyset, \{a'\}, \{b'\}, \{a', b'\}\} = \{\emptyset, \{a, b\}, \{b, c\}, \{a, c\}\}$  and  $\wp(g^{-1}(5)) = \{\emptyset, \{c'\}\} = \{\emptyset, \{a, b, c\}\}$ . Hence the two direct-sum decompositions of eigenspaces are:

$$\begin{aligned} &\wp(f^{-1}(1)) \oplus \wp(f^{-1}(2)) \\ &= \{\emptyset, \{a\}\} \oplus \{\emptyset, \{b\}, \{c\}, \{b, c\}\} \\ &\quad \text{and} \\ &\wp(g^{-1}(3)) \oplus \wp(g^{-1}(5)) \\ &= \{\emptyset, \{a, b\}, \{b, c\}, \{a, c\}\} \oplus \{\emptyset, \{a, b, c\}\}. \end{aligned}$$

Each r.v. has two eigenspaces so there are four intersections and the only non-zero intersection is:

$$\begin{aligned} \wp(f^{-1}(2)) \cap \wp(g^{-1}(3)) &= \{\emptyset, \{b\}, \{c\}, \{b, c\}\} \cap \{\emptyset, \{a, b\}, \{b, c\}, \{a, c\}\} = \\ &\{\emptyset, \{b, c\}\} = \mathcal{SE} \end{aligned}$$

so  $f$  and  $g$  are incompatible but not conjugate.

*Example of compatibility:* Let  $U$  and  $f$  be the same but  $U^* = \{a^*, b^*, c^*\}$  where  $\{a^*\} = \{a\}$ ,  $\{b^*\} = \{b\}$ , and  $\{c^*\} = \{b, c\}$  with  $g : U^* \rightarrow \mathbb{R}$  defined as  $g(a^*) = 3$ ,  $g(b^*) = 4$ , and  $g(c^*) = 5$ . Then the three ‘eigenspaces’ of  $g$  (i.e., subspaces determined by  $g$ ) are  $\wp(g^{-1}(3)) = \{\emptyset, a^*\} = \{\emptyset, a\}$ ,  $\wp(g^{-1}(4)) = \{\emptyset, \{b\}\}$ ;  $\wp(g^{-1}(5)) = \{\emptyset, \{b, c\}\}$ . The direct-sum decompositions are:

$$\begin{aligned} & \wp(f^{-1}(1)) \oplus \wp(f^{-1}(2)) \\ &= \{\emptyset, \{a\}\} \oplus \{\emptyset, \{b\}, \{c\}, \{b, c\}\} \\ & \quad \text{and} \\ & \wp(g^{-1}(3)) \oplus \wp(g^{-1}(4)) \oplus \wp(g^{-1}(5)) \\ &= \{\emptyset, a\} \oplus \{\emptyset, \{b\}\} \oplus \{\emptyset, \{b, c\}\}. \end{aligned}$$

Then the intersections with the eigenspaces of  $f$  contain  $\{a\}$ ,  $\{b\}$ , and  $\{b, c\}$  and those three vectors form a basis for  $\wp(U) \cong \mathbb{Z}_2^3$  so  $\mathcal{SE} = \mathbb{Z}_2^3$  and thus  $f$  and  $g$  commute.

*Example of conjugacy:* Consider the  $U$ -basis  $= \{\{a\}, \{b\}, \{c\}, \{d\}\}$  and  $\hat{U}$ -basis  $= \{\{\hat{a}\}, \{\hat{b}\}, \{\hat{c}\}, \{\hat{d}\}\} = \{\{b, c, d\}, \{a, c, d\}, \{a, b, d\}, \{a, b, c\}\}$  of  $\mathbb{Z}_2^4$  where  $\{\hat{a}\} = \{b, c, d\}, \dots, \{\hat{d}\} = \{a, b, c\}$ . Let  $f = \chi_{\{a, b\}} : U \rightarrow \mathbb{Z}_2$  so  $f(a) = f(b) = 1$  and  $f(c) = f(d) = 0$  with  $\wp(f^{-1}(1)) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$  and  $\wp(f^{-1}(0)) = \{\emptyset, \{c\}, \{d\}, \{c, d\}\}$ . Let  $g = \chi_{\{\hat{b}, \hat{c}\}} : \hat{U} \rightarrow \mathbb{Z}_2$  so  $g(\hat{b}) = g(\hat{c}) = 1$  and  $g(\hat{a}) = g(\hat{d}) = 0$  with  $\wp(g^{-1}(1)) = \{\emptyset, \{\hat{b}\}, \{\hat{c}\}, \{\hat{b}, \hat{c}\}\} = \{\emptyset, \{a, c, d\}, \{a, b, d\}, \{b, c\}\}$  and  $\wp(g^{-1}(0)) = \{\emptyset, \{\hat{a}\}, \{\hat{d}\}, \{\hat{a}, \hat{d}\}\} = \{\emptyset, \{b, c, d\}, \{a, b, c\}, \{a, d\}\}$ . The two direct-sum decompositions are:

$$\begin{aligned} & \wp(f^{-1}(1)) \oplus \wp(f^{-1}(0)) \\ &= \{\emptyset, \{a\}, \{b\}, \{a, b\}\} \oplus \{\emptyset, \{c\}, \{d\}, \{c, d\}\} \\ & \quad \text{and} \\ & \wp(g^{-1}(1)) \oplus \wp(g^{-1}(0)) \\ &= \{\emptyset, \{a, c, d\}, \{a, b, d\}, \{b, c\}\} \oplus \{\emptyset, \{b, c, d\}, \{a, b, c\}, \{a, d\}\}. \end{aligned}$$

Then it is easily checked that all the four intersections of eigenspaces have only  $\emptyset$  in common so  $\mathcal{SE} = \{\emptyset\}$  and  $f$  and  $g$  are conjugate.

Since the two numerical attributes in the conjugacy example are characteristic functions, they define linear operators on  $\mathbb{Z}_2^4$  and thus we can compute their commutator as usual once restated in the computational basis. The matrix to convert a 0, 1-vector written in the  $\hat{U}$ -basis to the same 0, 1-vector written in the  $U$ -basis is:

$$\begin{aligned}
C_{U \leftarrow \hat{U}} &= \begin{bmatrix} \langle \{a\} |_U \{\hat{a}\} \rangle & \langle \{a\} |_U \{\hat{b}\} \rangle & \langle \{a\} |_U \{\hat{c}\} \rangle & \langle \{a\} |_U \{\hat{d}\} \rangle \\ \langle \{b\} |_U \{\hat{a}\} \rangle & \langle \{b\} |_U \{\hat{b}\} \rangle & \langle \{b\} |_U \{\hat{c}\} \rangle & \langle \{b\} |_U \{\hat{d}\} \rangle \\ \langle \{c\} |_U \{\hat{a}\} \rangle & \langle \{c\} |_U \{\hat{b}\} \rangle & \langle \{c\} |_U \{\hat{c}\} \rangle & \langle \{c\} |_U \{\hat{d}\} \rangle \\ \langle \{d\} |_U \{\hat{a}\} \rangle & \langle \{d\} |_U \{\hat{b}\} \rangle & \langle \{d\} |_U \{\hat{c}\} \rangle & \langle \{d\} |_U \{\hat{d}\} \rangle \end{bmatrix} \\
&= \begin{bmatrix} \langle \{a\} |_U \{b, c, d\} \rangle & \langle \{a\} |_U \{a, c, d\} \rangle & \langle \{a\} |_U \{a, b, d\} \rangle & \langle \{a\} |_U \{a, b, c\} \rangle \\ \langle \{b\} |_U \{b, c, d\} \rangle & \langle \{b\} |_U \{a, c, d\} \rangle & \langle \{b\} |_U \{a, b, d\} \rangle & \langle \{b\} |_U \{a, b, c\} \rangle \\ \langle \{c\} |_U \{b, c, d\} \rangle & \langle \{c\} |_U \{a, c, d\} \rangle & \langle \{c\} |_U \{a, b, d\} \rangle & \langle \{c\} |_U \{a, b, c\} \rangle \\ \langle \{d\} |_U \{b, c, d\} \rangle & \langle \{d\} |_U \{a, c, d\} \rangle & \langle \{d\} |_U \{a, b, d\} \rangle & \langle \{d\} |_U \{a, b, c\} \rangle \end{bmatrix} = \\
&\quad \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}.
\end{aligned}$$

For instance,  $\{a\} = \{\hat{b}, \hat{c}, \hat{d}\}$  so that vector in  $\hat{U}$ -basis is the column vector  $[0, 1, 1, 1]^t$  ( $t$  represents the transpose) and

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \stackrel{\text{mod}(2)}{=} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \{a\} \text{ in } U\text{-basis.}$$

Using some mathematical software, the inverse to the conversion matrix  $C_{U \leftarrow \hat{U}}$  computed in the rationals is:

$$C_{\hat{U} \leftarrow U} = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{bmatrix} \stackrel{\text{mod}(2)}{=} \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} = C_{\hat{U} \leftarrow U}$$

and then converting everything mod (2) yields back the same matrix so it is its own inverse. Then the projection matrix  $P_s$  for  $s = 1$  (or  $P_{s=1}$ ) to the eigenspace  $\wp(g^{-1}(1))$  for the  $g$ -operator  $\hat{b}, \hat{c} \rightarrow 1; \hat{a}, \hat{d} \rightarrow 0$  can be converted to the  $U$ -basis as:

$$C_{U \leftarrow \hat{U}} P_s C_{\hat{U} \leftarrow U}$$

$$\begin{aligned}
&= \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 2 & 1 & 1 & 2 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 2 \end{bmatrix} \stackrel{\text{mod}(2)}{=} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.
\end{aligned}$$

Then we can compute the commutator of  $P_r$  for  $r = 1$  (or  $P_{r=1}$ ) which projects to  $\wp(f^{-1}(1)) = \{\emptyset, \{a\}, \{b\}, (a, b)\}$  and  $P_s$  for  $s = 1$  (or  $P_{s=1}$ ) which projects to  $\wp(g^{-1}(1)) = \{\emptyset, \{a, c, d\}, \{a, b, d\}, \{b, c\}\}$  in the computational  $U$ -basis:

$$\begin{aligned}
&[P_{r=1}, P_{s=1}] = P_{r=1}P_{s=1} - P_{s=1}P_{r=1} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \stackrel{\text{mod}(2)}{=} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}
\end{aligned}$$

which has a determinant of 1 so the commutator is a non-singular transformation with a zero space kernel  $\mathcal{SE}$ .

Heisenberg's indeterminacy (or uncertainty) principle in QM/Sets brings out what might be considered the essence of the principle. If  $f$  and  $g$  are conjugate attributes, then any vector (always nonzero) in  $\mathbb{Z}_2^n$  that is a (sharp) eigenvector for one attribute must be an indefinite superposition of eigenvectors for the other attribute. A vector cannot be an eigenvector for both attributes since, by assumption, the space of simultaneous eigenvectors for conjugate attributes is the zero space. For instance in the example, an eigenvector like  $\{a\}$  in the  $U$ -basis is a superposition of eigenvectors  $\{\hat{b}, \hat{c}, \hat{d}\}$  in the  $\hat{U}$ -basis and an eigenvector like  $\{\hat{b}\}$  in the  $\hat{U}$ -basis is a superposition  $\{a, c, d\}$  in the  $U$ -basis.

At the other extreme of commuting attributes, they act very much like ordinary r.v.s on a fixed basis (of simultaneous eigenvectors). A set of ordinary set partitions on

a universe  $U$  is said to be *complete* if their join is the discrete partition  $\mathbf{1}_U = \{\{u\}\}_{u \in U}$  whose blocks have cardinality one. Similarly a set of commuting attributes is said to be *complete* (a CSCO) [16] if all the non-zero intersections of all their eigenspaces are of dimension one. Then each of the simultaneous eigenvectors is uniquely characterized by the set of eigenvalues of those intersection eigenspaces. For instance in the example of compatibility above,  $f : U \rightarrow \mathbb{R}$  and  $g : U^* \rightarrow \mathbb{R}$  are complete so each of simultaneous eigenvectors  $\{a\}$ ,  $\{b\}$ , and  $\{b, c\}$  can be characterized by their pair  $(r, s)$  eigenvalues:  $|\{a\}\rangle = |1, 3\rangle$ ,  $|\{b\}\rangle = |2, 4\rangle$ , and  $|\{b, c\}\rangle = |2, 5\rangle$ . The same holds for a complete set of commuting operators or CSCO in QM/ $\mathbb{C}$ .

### 3.3 ‘Dynamics’ in QM/Sets and the double-slit experiment

Since this treatment of non-commutativity in QM/Sets is the first place that different bases play a key role, it is a good place to point out that the probability calculations in a certain basis always treat each basis vector as a singleton even though it may not be a singleton when represented in another basis. For instance, in the  $U'$ -basis of Table 1.1,  $\{a'\} = \{a, b\}$ ,  $\{b'\} = \{b, c\}$ , and  $\{c'\} = \{a, b, c\}$  are all disjoint for the  $U'$  calculations, e.g.,  $\langle \{a'\} |_{U'} \{b', c'\} \rangle = |\{a'\} \cap \{b', c'\}| = 0$  whereas in the computational basis,  $\{b', c'\} = \{a\}$  and  $\{a'\} = \{a, b\}$  so  $\langle \{a, b\} |_{U'} \{a\} \rangle = |\{a, b\} \cap \{a\}| = 1$ . This is important in motivating the treatment of ‘dynamics’ in QM/Sets.

In QM/ $\mathbb{C}$ , the unitary dynamics can be characterized as taking a ON basis to an ON basis or as simply preserving the values of the brackets (or inner products)  $\langle \psi | \phi \rangle$ . The analogue of taking an ON basis to an ON basis in QM/ $\mathbb{Z}_2$  is just taking a basis set to a basis set, i.e., a non-singular linear transformation  $\mathbb{Z}_2^n \rightarrow \mathbb{Z}_2^n$ . Taking all basis subsets always as singletons like the  $U$ -basis  $\{\{a\}, \{b\}, \{c\}\}$  and the  $U'$ -basis  $\{\{a'\}, \{b'\}, \{c'\}\}$ , then the non-singular linear transformation  $\{a\} \mapsto \{a'\}$ ,  $\{b\} \mapsto \{b'\}$ ,  $\{c\} \mapsto \{c'\}$  would preserve the size of intersections each expressed in its own basis. In that manner, non-singular transformation also preserve the brackets, e.g.,  $\langle \{a\} |_{U'} \{a, b\} \rangle = \langle \{a'\} |_{U'} \{a', b'\} \rangle$ , just as unitary transformations preserve the brackets in QM/ $\mathbb{C}$ . This is how QM/Sets mimics the probability calculations of QM/ $\mathbb{C}$ . Hence non-singular transformations will be taken as the ‘dynamics’ in QM/ $\mathbb{Z}_2$  (which was also done in modal QM [52]).

For the double-slit experiment, we take a dynamics that most closely mimics the spreading of waves. For  $U = \{a, b, c\}$ , consider the dynamics:  $\{a\} \rightarrow \{a, b\}$ ;  $\{b\} \rightarrow$



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

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

To model the two slit experiment, we take  $a$ ,  $b$ , and  $c$  as three vertical positions and we have a vertical diaphragm with slits at  $a$  and  $c$  as in Figure 3.5. 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.

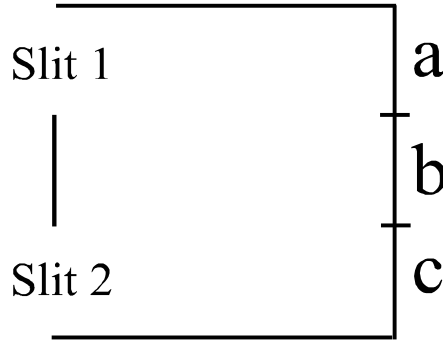


Figure 3.5: Setup for two-slit experiment.

**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 vertical position is measured again. Prior to the measurement at the slits, the particle is in a superposition state between slit 1 at  $a$  and slit 2 at  $b$ . The probability that the particle is detected at slit 1 or at slit 2 is:

$$\begin{aligned} \Pr(\{a\} \text{ at slits} \mid \{a, c\} \text{ at slits}) &= \frac{\|\langle \{a\} \mid_U \{a, c\} \rangle\|_U^2}{\|\{a, c\}\|_U^2} = \frac{|\{a\} \cap \{a, c\}|}{|\{a, c\}|} = \frac{1}{2}, \\ \Pr(\{c\} \text{ at slits} \mid \{a, c\} \text{ at slits}) &= \frac{\|\langle \{c\} \mid_U \{a, c\} \rangle\|_U^2}{\|\{a, c\}\|_U^2} = \frac{|\{c\} \cap \{a, c\}|}{|\{a, c\}|} = \frac{1}{2}. \end{aligned}$$

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

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

If 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:

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

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 (see Figure 3.6) if it starts at the slits in the superposition state  $\{a, c\}$  and the measurements were made at the slits:

$$\begin{aligned}\Pr(\{a\} \text{ at wall} \mid \{a, c\} \text{ at slits}) &= \frac{1}{2} \frac{1}{2} = \frac{1}{4}; \\ \Pr(\{b\} \text{ at wall} \mid \{a, c\} \text{ at slits}) &= \frac{1}{2} \frac{1}{2} + \frac{1}{2} \frac{1}{2} = \frac{1}{2}; \\ \Pr(\{c\} \text{ at wall} \mid \{a, c\} \text{ at slits}) &= \frac{1}{2} \frac{1}{2} = \frac{1}{4}.\end{aligned}$$

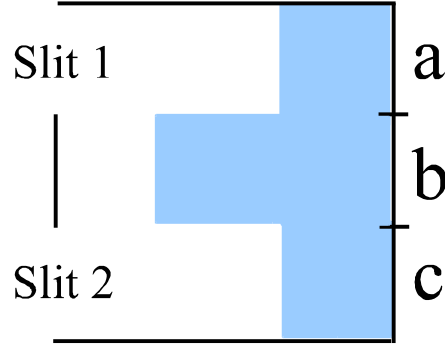


Figure 3.6: Particle distribution at the wall with detection at slits.

**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  $\{a, b\} + \{b, c\} = \{a, c\}$  where the superposition at  $\{b\}$  cancels out. Then the final probabilities will just be probabilities of finding  $\{a\}$ ,  $\{b\}$ , or  $\{c\}$  when the measurement is made only at the wall on the right are pictured in Figure 3.7:

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

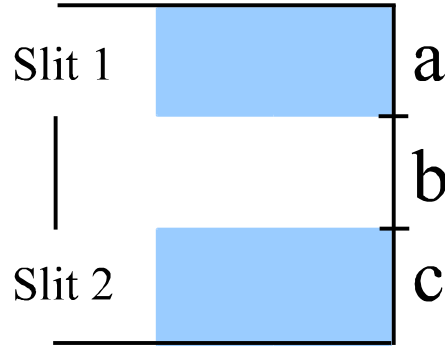


Figure 3.7: Particle distribution with no detection at slits.

The particle distribution shows the usual stripes of the double-slit experiment in QM/C. 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 amplitudes are now added together *as part of the evolution* (instead of the probabilities adding as in the first case). This allows the interference of the two results and thus the cancellation of the  $\{b\}$  term in  $\{a, b\} + \{b, c\} = \{a, c\}$ . The addition is, of course, mod 2 (where  $-1 = +1$ ) so, in "wave language," the "wave crest" and "wave trough" that add at the location  $\{b\}$  cancel out.

## 4 Entanglement in QM/Sets

### 4.1 Direct products and correlation in QM/Sets

We are now in a better position to understand what the author's mathematical mentor, Gian-Carlo Rota, would call the methodology or "yoga" of translating set concepts into vector-space concepts. Apply the set concepts to basis sets of vector spaces and then see what vector-space concept is generated. That correlation of concepts is not the only possible one. For instance, a set partition can always be represented as the inverse image of a set function  $f : U \rightarrow \pi$  where each  $u \in U$  is taken to the block containing it in a partition  $\pi$  on  $U$  so that  $f^{-1} = \pi$ . Then one might take the vector space version of a set partition as the partition on  $V$  that is the inverse image of a linear functional  $\phi : V \rightarrow \mathbb{R}$ . The two methods give quite different results. But the

development of QM/Sets is based on the first methodology. By that methodology, we take a set partition on a basis set of a vector space and see what vector-space concept it generates—and we see that it generates a set of disjoint (i.e., only intersection is the zero vector) vector spaces whose direct-sum is the whole space, i.e., a direct-sum decomposition. Moreover, we see that when the vector spaces are over  $\mathbb{Z}_2$ , then the original set concept and the vector-space concept are ‘essentially’ the same—as in a set partition  $\pi = \{B\}_{B \in \pi}$  on  $U$  and the direct-sum decomposition  $\{\wp(B)\}_{B \in \pi}$  of  $\wp(U)$ . Starting with a vector-space concept in QM/C, we consider the corresponding set concept which is then restated in  $\mathbb{Z}_2^n$  where the vectors are interpreted as subsets—and thus QM/Sets is generated.

In QM/C, the notion of entanglement arises by considering the tensor product of two quantum systems—each represented by a state in a vector space. What is the set-version of the tensor product of vector spaces? For sets, the notion of product is the Cartesian or direct product of two sets:  $X \times Y = \{(x, y) : x \in X, y \in Y\}$ . Applying the methodology, we take the direct product of a basis set  $\{v_1, \dots, v_m\}$  of a  $V$  and a basis set  $\{w_1, \dots, w_n\}$  of  $W$  (over the same base field). Instead of representing each ordered pair as  $(v_i, w_j)$ , we denote that it as  $v_i \otimes w_j$  and then we generate a vector space from those basis pairs. That vector space is the tensor product  $V \otimes W$ . Thus the tensor product of vector spaces is the vector-space version of the direct product of sets—even though there is also the notion of the direct product of vector spaces  $V \times W$ . The vector space direct product  $V \times W$  is just the set of ordered pairs  $\{(v, w) : v \in V, w \in W\}$  with component-wise addition and it has a basis set of  $\{(v_1, 0), (v_2, 0), \dots, (v_m, 0)\} \cup \{(0, w_1), (0, w_2), \dots, (0, w_n)\}$  so it has dimension  $\dim(V) + \dim(W) = m + n$ . In contrast, the tensor product  $V \otimes W$  has dimension  $|\{v_1, \dots, v_m\} \times \{w_1, \dots, w_n\}| = \dim(V) \times \dim(W) = mn$ . We have seen that the direct or Cartesian product of sets induces or lifts to the tensor product of vector spaces. Moreover, when the vector spaces are over  $\mathbb{Z}_2$ , then the set concept and the vector-space concept are ‘essentially’ the same. That is, given two finite sets  $X = \{x_1, \dots, x_m\}$  and  $Y = \{y_1, \dots, y_n\}$  the cardinality of  $X \times Y$  is  $|X| \times |Y| = mn$  and the dimension of  $\wp(X) \otimes_{\mathbb{Z}_2} \wp(Y)$  is  $\dim(\wp(X)) \times \dim(\wp(Y)) = mn$  where the basis elements  $x_i \otimes y_j$  of the tensor product are correlated with the ordered pairs  $(x_i, y_j) \in X \times Y$  in the vector space isomorphism:  $\wp(X) \otimes_{\mathbb{Z}_2} \wp(Y) \cong \wp(X \times Y)$  as vector spaces over  $\mathbb{Z}_2$ . Hence for the sake of simplicity, we will just use the direct product  $X \times Y$  as the universe set to analyze entanglement in QM/Sets.

In QM/ $\mathbb{C}$ , 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*. In QM/Sets, a subset  $S \subseteq X \times Y$  is said to be *separated* or a *product set* 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<sup>7</sup> are defined in the usual way:

$$\begin{aligned}\Pr(x) &= \sum_y \Pr(x, y) \\ \Pr(y) &= \sum_x \Pr(x, y).\end{aligned}$$

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

$$\begin{aligned}\Pr(x, y) &= \Pr(x) \Pr(y). \\ \text{Independent distribution}\end{aligned}$$

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

**Proposition 2** *A subset  $S \subseteq \wp(X \times Y)$  is entangled iff the equiprobable distribution on  $S$  is correlated (non-independent).*

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  $|S_X||S_Y| = |S| = N$ . 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  $S \subsetneq S_X \times S_Y$  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.  $\square$

---

<sup>7</sup>The marginal distributions are the set versions of the reduced density matrices of QM/ $\mathbb{C}$ .

Consider the set version of one qubit space where  $U = \{a, b\}$ . The product set  $U \times U$  has 15 nonempty subsets. Each  $\wp(U)$  has 3 nonempty subsets so  $3 \times 3 = 9$  of the 15 subsets are "separated" subsets leaving 6 "entangled" subsets listed in Table 4.1.

$S \in \wp(U \times U)$	$v \in \wp(U) \otimes \wp(U)$
$\{(a, a), (b, b)\}$	$\{a\} \otimes \{a\} + \{b\} \otimes \{b\}$
$\{(a, b), (b, a)\}$	$\{a\} \otimes \{b\} + \{b\} \otimes \{a\}$
$\{(a, a), (a, b), (b, a)\}$	$\{a\} \otimes \{a\} + \{a\} \otimes \{b\} + \{b\} \otimes \{a\}$
$\{(a, a), (a, b), (b, b)\}$	$\{a\} \otimes \{a\} + \{a\} \otimes \{b\} + \{b\} \otimes \{b\}$
$\{(a, b), (b, a), (b, b)\}$	$\{a\} \otimes \{b\} + \{b\} \otimes \{a\} + \{b\} \otimes \{b\}$
$\{(a, a), (b, a), (b, b)\}$	$\{a\} \otimes \{a\} + \{b\} \otimes \{a\} + \{b\} \otimes \{b\}$

Table 4.1: The six entangled subsets of  $\wp(U \times U)$  and corresponding vectors in  $\wp(U) \otimes_{\mathbb{Z}_2} \wp(U)$ .

The first two might be called *Bell states* which are the two graphs of bijections  $U \longleftrightarrow U$ . The degree of entanglement can be measured by the logical divergence between  $\Pr(u, u')$  and  $\Pr(u) \Pr(u')$  for  $u, u' \in U$  where for two probability distributions  $p = \{p_1, \dots, p_n\}$  and  $q = \{q_1, \dots, q_n\}$ , the *logical divergence* between them is defined as the square of the Euclidean distance between them:  $d(p||q) = \sum_i (p_i - q_i)^2$ [20].

Consider the Bell state  $S = \{(a, b), (b, a)\}$ . Then  $\Pr(u, u') = \frac{1}{|S|} = \frac{1}{2}$  if  $(u, u') \in S$  and  $\Pr(u, u') = 0$  for  $u, u' \in U$  otherwise so  $\Pr(a, b) = \frac{1}{2} = \Pr(b, a)$  while  $\Pr(u, u') = 0$  otherwise. Since  $X = Y = U$ , we need to, in general, distinguish the marginals on the left and right. In this symmetrical case, the marginals on the left and right are:  $\Pr_L(a) = \sum_{u \in U} \Pr(a, u) = \frac{1}{2} = \sum_{u \in U} \Pr(u, a) = \Pr_R(a)$  and  $\Pr(b)_L = \sum_{u \in U} \Pr(b, u) = \frac{1}{2} = \sum_{u \in U} \Pr(u, b) = \Pr_R(b)$ . Then the logical divergence of the Bell state is:

$$\begin{aligned}
d(\Pr(-, -) || \Pr_L(-) \Pr_R(-)) &= \sum_{(u, u') \in U \times U} (\Pr(u, u') - \Pr(u)_L \Pr_R(u'))^2 \\
&= (\Pr(a, a) - \Pr(a)_L \Pr(a)_R)^2 + (\Pr(a, b) - \Pr(a)_L \Pr(b)_R)^2 + \\
&\quad (\Pr(b, a) - \Pr(b)_L \Pr(a)_R)^2 + (\Pr(b, b) - \Pr(b)_L \Pr(b)_R)^2 \\
&= \left(-\frac{1}{2}\frac{1}{2}\right)^2 + \left(\frac{1}{2} - \frac{1}{2}\frac{1}{2}\right)^2 + \left(\frac{1}{2} - \frac{1}{2}\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\frac{1}{2}\right)^2 \\
&= \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} = \frac{1}{4}.
\end{aligned}$$

Consider the entangled state  $S = \{(a, a), (a, b), (b, a)\}$  which is not a Bell state. Then  $\Pr(u, u') = \frac{1}{|S|} = \frac{1}{3}$  if  $(u, u') \in S$  and 0 otherwise. The marginals on the left are

$\Pr_L(a) = \sum_{u \in U} \Pr(a, u) = \frac{2}{3}$  and  $\Pr_L(b) = \sum_{u \in U} \Pr(b, u) = \frac{1}{3}$  and the marginals on the right are:  $\Pr_R(a) = \sum_{u \in U} \Pr(u, a) = \frac{2}{3}$  and  $\Pr_R(b) = \sum_{u \in U} \Pr(u, b) = \frac{1}{3}$ . Then the logical divergence is:

$$\begin{aligned} d(\Pr(-, -) \parallel \Pr_L(-) \Pr_R(-)) &= \sum_{(u, u') \in U \times U} (\Pr(u, u') - \Pr_L(u) \Pr_R(u'))^2 \\ &= \left(\frac{1}{3} - \frac{2}{3} \frac{2}{3}\right)^2 + \left(\frac{1}{3} - \frac{2}{3} \frac{1}{3}\right)^2 + \left(\frac{1}{3} - \frac{1}{3} \frac{2}{3}\right)^2 + \left(-\frac{1}{3} \frac{1}{3}\right)^2 \\ &= \frac{1}{81} + \frac{1}{81} + \frac{1}{81} + \frac{1}{81} = \frac{4}{81}. \end{aligned}$$

And for one of the separated states, say,  $S = \{a\} \times \{a, b\} = \{(a, a), (a, b)\}$ , we have  $\Pr(a, a) = \Pr(a, b) = \frac{1}{2}$  and 0 otherwise.  $\Pr_L(a) = \sum_{u \in U} \Pr(a, u) = 1$  and  $\Pr_R(a) = \frac{1}{2}$  while  $\Pr_L(b) = 0$  and  $\Pr_R(b) = \frac{1}{2}$ . Hence the logical divergence is:

$$\begin{aligned} d(\Pr(-, -) \parallel \Pr_L(-) \Pr_R(-)) &= \sum_{(u, u') \in U \times U} (\Pr(u, u') - \Pr_L(u) \Pr_R(u'))^2 \\ &= \left(\frac{1}{2} - \frac{1}{1} \frac{1}{2}\right)^2 + \left(\frac{1}{2} - \frac{1}{1} \frac{1}{2}\right)^2 + \left(0 - 0 \frac{1}{2}\right)^2 + \left(0 - 0 \frac{1}{2}\right)^2 = 0. \end{aligned}$$

This logical divergence  $d(\Pr(-, -) \parallel \Pr_L(-) \Pr_R(-))$ , as a measure of entanglement, gives the maximal entanglement  $\frac{1}{4}$  to the Bell states, an intermediate entanglement  $\frac{4}{81}$  to the entangled states that are not Bell states, and zero entanglement to the separated states.

For an entangled subset  $S$ , a sampling  $u$  of left-hand system will change the probability distribution for a sampling of the right-hand system  $u'$ ,  $\Pr(u'|u) \neq \Pr(u')$ . In the case of maximal entanglement (e.g., the Bell states), when  $S$  is the graph of a bijection between  $U$  and  $U$ , e.g.,  $\{(a, b), (b, a)\}$  is the graph of the bijection  $U \cong U$  where  $a \longleftrightarrow b$ , the value of the right-hand  $u'$  is determined by the value of left-hand  $u$  (and vice-versa). The Bell state  $\{(a, b), (b, a)\}$  in  $\wp(U \times U)$  or  $\{a\} \otimes \{b\} + \{b\} \otimes \{a\}$  in  $\wp(U) \otimes_{\mathbb{Z}_2} \wp(U)$  is the QM/Sets version of the entangled Bell state  $|L \downarrow\rangle \otimes |R \uparrow\rangle + |L \uparrow\rangle \otimes |R \downarrow\rangle$  in the standard Bell-type experiment in QM/ $\mathbb{C}$  where two particles are separated, one going left and the other going right. When one measures the spin of the left-hand particle, then it determines the spin of the right-hand particle.

## 4.2 Bell's Theorem in QM/Sets

This treatment of Bell's Theorem in QM/Sets is based on the simple expository example developed by Bernard D 'Espagnat [15]. A simple version of a Bell inequality can be derived in the case of  $\mathbb{Z}_2^2$  with three bases  $U = \{a, b\}$ ,  $U' = \{a', b'\}$ , and  $U'' = \{a'', b''\}$ , and where the kets are listed in Table 4.2.

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$	$\emptyset$	$\emptyset$	$\emptyset$

Table 4.2: Ket table for  $\wp(U) \cong \wp(U') \cong \wp(U'') \cong \mathbb{Z}_2^2$ .

The different basis vectors can be thought of as spin-up and spin-down along three different  $A$ ,  $B$ , and  $C$  axes, e.g.  $\{a\} = \{A^+\}$  and  $\{b\} = \{A^-\}$ ,  $\{a'\} = \{B^+\}$  and  $\{b'\} = \{B^-\}$ , and  $\{a''\} = \{C^+\}$  and  $\{b''\} = \{C^-\}$ , but we will stick to our usual notation.

Attributes or observables 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  $U''$ .

A more technical way to show incompatibility is to exploit the vector space structure of  $\mathbb{Z}_2^2$  and to see if the projection matrices for  $\{a\} \cap ()$  and  $\{b'\} \cap ()$  commute. The basis conversion matrices between the  $U$ -basis and  $U'$ -basis are:

$$\mathcal{C}_{U \leftarrow U'} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \text{ and } \mathcal{C}_{U' \leftarrow U} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

The projection matrix for  $\{a\} \cap ()$  in the  $U$ -basis is, of course,  $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$  and the projection matrix for  $\{b'\} \cap ()$  in the  $U'$ -basis is  $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ . Converting the latter to the  $U$ -basis to check commutativity gives:

$$\begin{aligned} & \mathcal{C}_{U \leftarrow U'} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \mathcal{C}_{U' \leftarrow U} \\ &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}. \end{aligned}$$

Hence the commutativity check is:



$$\begin{aligned} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \neq \\ \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \end{aligned}$$

so the two operators for the observables  $\chi_{\{a\}}$  and  $\chi_{\{b\}}$  do not commute. In a similar manner, it is seen that the three observables are mutually incompatible.

Given a ket in  $\mathbb{Z}_2^2 \cong \wp(U) \cong \wp(U') \cong \wp(U'')$ , and using the equiprobability assumption on different basis sets, the probabilities of getting the different outcomes for the various observables in the different given states are given in the Table 4.3.

Given state \ Outcome of test	$\chi_{\{a\}} = 1$	$= 0$	$\chi_{\{b\}} = 0$	$= 1$	$\chi_{\{a''\}} = 1$	$= 0$
$\{a, b\} = \{a'\} = \{a''\}$	$\frac{1}{2}$	$\frac{1}{2}$	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

Table 4.3: State-outcome table.

The tensor product of vector spaces over  $\mathbb{Z}_2$  is the Cartesian or direct product of sets, e.g.,  $\wp(U) \otimes \wp(U) \cong \wp(U \times U)$ , and 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  $\{a\} \times \{b\} = \{(a, b)\} \in \wp(U \times U)$  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 4.4 that suffices for our calculations.

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

Table 4.4: Partial ket table for  $\wp(U \times U) \cong \wp(U' \times U') \cong \wp(U'' \times U'')$ .

As before, we can classify each vector or subset as separated or entangled and we can furthermore see how that is independent of the basis. For instance  $\{(a, a), (a, b)\}$  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, there is a probability distribution on  $U \times U' \times U''$  where  $\Pr(a, a', a'')$  (for instance) is defined as the probability of getting the result  $\{a\}$  if a  $U$ -measurement is performed on the left-hand system, and if instead a  $U'$ -measurement is performed on the left-hand system then  $\{a'\}$  is obtained, and if instead a  $U''$ -measurement is performed on the left-hand system then  $\{a''\}$  is obtained. Thus we would have  $\Pr(a, a', a'') = \frac{1}{2} \frac{2}{3} \frac{2}{3} = \frac{2}{9}$ . In this way the 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 [15]. Consider the following marginals:

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

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:

$$\begin{aligned}
\Pr(a, a') + \Pr(b', b'') &\geq \Pr(a, b'') \\
&\text{Bell inequality.}
\end{aligned}$$

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

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

the probabilities would be the same.<sup>8</sup> That is, under that assumption, the probabilities,  $\Pr(a) = \frac{1}{2} = \Pr(b)$ ,  $\Pr(a') = \frac{2}{3} = \Pr(a'')$ , and  $\Pr(b') = \frac{1}{3} = \Pr(b'')$  are the same regardless of whether we are measuring the left-hand or right-hand system of that composite state. Thus with those left-right independent measurements,  $\Pr(a, a') = \frac{1}{2} \frac{2}{3} = \frac{1}{3}$  which is now interpreted as the probability of getting  $\{a\}$  in a left-hand  $U$ -measurement and then getting  $\{a'\}$  in a right-hand  $U'$ -measurement, and similarly  $\Pr(b', b'') = \frac{1}{3} \frac{1}{3} = \frac{1}{9}$ , and  $\Pr(a, b'') = \frac{1}{2} \frac{1}{3} = \frac{1}{6}$ , so the above Bell inequality would still hold. But we can use QM/ $\mathbb{Z}_2$  to compute the probabilities for those different measurements on the two systems to see if the independence assumption is compatible with QM/ $\mathbb{Z}_2$ .

---

<sup>8</sup>The same holds for the other "Bell state":  $\{(a, b), (b, a)\}$ .

To compute  $\Pr(a, a')$ , 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  $\Pr(a, a') = \frac{1}{2} \frac{1}{2} = \frac{1}{4}$ .

To compute  $\Pr(b', b'')$ , 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 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  $\Pr(b', b'') = 0$ .

Finally we compute  $\Pr(a, b'')$  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  $\Pr(a, b'') = \frac{1}{2}$ .

Then we plug the probabilities into the Bell inequality:

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

$$\frac{1}{4} + 0 \not\geq \frac{1}{2}$$

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/ $\mathbb{Z}_2$  so the effects of the QM/ $\mathbb{Z}_2$  measurements are said to be "nonlocal"—or, more simply put, the outcome of the left-measurement affects the outcomes of the right-measurement of the entangled state.

## 5 Density matrices in QM/Sets

### 5.1 Logical entropy using density matrices

Our informal comments about "making distinctions" and distinguishability can be formalized using the logical theory of information [20] that is the quantitative version of the logic of partitions [19]. In category theory, the notion of a partition on a set

is dual to the notion of a subset of a set, so the logic of partitions is in that sense dual to the usual Boolean logic of subsets—that is usually presented in the special case of "propositional logic." As George Boole emphasized, the quantitative version of Boolean logic is finite probability theory so that would suggest a certain duality between the two quantitative notions of probability and information. Gian-Carlo Rota suggested precisely this connection in his Fubini Lectures [51], in his lectures at MIT, and in the exposition of *Combinatorics: The Rota Way* [39]. The “lattice of partitions plays for information the role that the Boolean algebra of subsets plays for size or probability” [39, p. 30] or symbolically:

$$\frac{\text{information}}{\text{partitions}} \approx \frac{\text{probability}}{\text{subsets}}.$$

Since “Probability is a measure on the Boolean algebra of events” that gives quantitatively the “intuitive idea of the size of a set”, we may ask by “analogy” for some measure to capture a property for a partition like “what size is to a set.” Rota goes on to ask:

How shall we be led to such a property? We have already an inkling of what it should be: it should be a measure of information provided by a random variable. Is there a candidate for the measure of the amount of information? [51, p. 67]

Logical information theory answers that question by taking a distinction of a partition as the analogue of an element of a subset. Recall that a *distinction* or *dit* of a partition  $\pi = \{B_1, \dots, B_m\}$  on  $U$  is an ordered pair  $(u, u') \in U \times U$  where  $u$  and  $u'$  are in different blocks  $B, B' \in \pi$ . An *indistinction* or *indit* of  $\pi$  is an ordered pair  $(u, u')$  of elements of  $U$  which are in the same block. The *indit set*  $\text{indit}(\pi) \subseteq U \times U$  is the equivalence relation determined by  $\pi$ , i.e.,  $\text{indit}(\pi) = \bigcup_{i=1}^m B_i \times B_i$ . The *ditset*  $\text{dit}(\pi)$  of  $\pi$ , the set of  $\pi$ -distinctions, is  $\text{dit}(\pi) = U \times U - \text{indit}(\pi)$ . Hence analogy of the number of elements in a subset, i.e., the size of the subset, is the number of distinctions in a partition, i.e., the size of the ditset  $\text{dit}(\pi)$ . For equiprobable points in a universe  $U$ , the finite probability of an event  $S \subseteq U$  is the normalized size of the subset:  $\Pr(S) = \frac{|S|}{|U|}$ . In like manner, the *logical entropy*  $h(\pi) = \frac{|\text{dit}(\pi)|}{|U \times U|}$  of a partition  $\pi$  is the normalized number of distinctions in the partition. Each block  $B \in \pi$  has the probability  $\Pr(B) = \frac{|B|}{|U|}$ , so the logical entropy can be expressed as:

$$h(\pi) = \frac{|\text{dit}(\pi)|}{|U \times U|} = \frac{|U \times U| - \sum_{i=1}^m |B_i \times B_i|}{|U \times U|} = 1 - \sum_{i=1}^m \Pr(B_i)^2$$

which is the probability that two independent (i.e., "with replacement") draws from  $U$  will get a dit of  $\pi$ . Thus logical entropy is a (probability) measure on  $U \times U$ , the two-draw probability of a distinction.

In contrast to logical entropy, the better-known Shannon entropy is not a (non-negative) measure in the sense of measure theory and all the Shannon concepts of simple, joint, conditional, and mutual entropy result from a non-linear but monotonic dit-bit transform of the corresponding concepts in the logical information theory ([20]; [25]). Both logical entropy and Shannon entropy extend to quantum information theory where the quantum version of Shannon entropy is called "von Neumann entropy." The contrast between logical entropy and Shannon/von-Neumann entropy is important because logical entropy connects directly with the process of measurement in QM/ $\mathbb{C}$  and QM/ $\mathbb{Z}_2$ ; logical entropy measures measurement. Moreover, logical entropy is the quantitative version of the logic of partitions, and partitions turn out to be a key concept to make more sense out of quantum mechanics. This is because partitions are the natural mathematical tool to model indefiniteness and definiteness. A partition distinguishes the elements in different blocks of the partition and does not distinguish the elements within each block. As partitions become more refined through the join operation, they make more distinctions. A complete set of partitions on a universe set is a set of partition whose join is the discrete partition, and such set notions are just the QM/Sets version of the "complete set of commuting operators" (CSCO) whose combined measurements (essentially the same as the join operation) reduce a pure superposition state into a completely decomposed mixed state—the quantum version of the discrete partition. It is not an over-statement to say that these analogies between the mathematics of QM/ $\mathbb{C}$  and the mathematical machinery of partitions provided the genesis for the development of QM/Sets.

The mathematics of QM/ $\mathbb{C}$  can be reformulated using density matrices in place of vectors as representing states [49, p. 102], and logical information theory is also easily reformulated using density matrices so we turn to that development.

A binary relation  $R \subseteq U \times U$  on  $U = \{u_1, \dots, u_n\}$  can be represented by an  $n \times n$  *incidence matrix*  $In(R)$  where

$$In(R)_{ij} = \begin{cases} 1 & \text{if } (u_i, u_j) \in R \\ 0 & \text{if } (u_i, u_j) \notin R. \end{cases}$$

Taking  $R$  as the equivalence relation  $\text{indit}(\pi)$  associated with a partition  $\pi = \{B_1, \dots, B_m\}$ , the non-zero off-diagonal elements indicate that the corresponding elements along the diagonal are in the same block of the partition and thus they "cohere" together in that sense. For instance, for  $\pi = \{\{a, b\}, \{c\}\}$ , then the incidence matrix for the equivalence relation  $\text{indit}(\pi)$  defined by the partition is:

$$In(\text{indit}(\pi)) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where the non-zero off-diagonal elements indicate that  $\{a\}$  and  $\{b\}$  cohere together in  $\pi = \{\{a, b\}, \{c\}\}$  while  $\{c\}$  is a fully distinguished singleton. When  $\pi$  is joined with another partition, say,  $\sigma = \{\{a\}, \{b, c\}\}$ , then the join  $\pi \vee \sigma$  is in general more refined than either  $\pi$  or  $\sigma$  since the join is the least upper bound on  $\pi$  and  $\sigma$  in the refinement partial order on the partitions of  $U = \{a, b, c\}$ . That refinement is indicated by some non-zero off-diagonal elements in the incidence matrices for  $\pi$  and  $\sigma$  being zeroed in the join. If two elements of  $U$  are distinguished by either partition, then they are distinguished in the join since  $\text{dit}(\pi) \cup \text{dit}(\sigma) = \text{dit}(\pi \vee \sigma)$ . In terms of inditsets, that is  $\text{indit}(\pi) \cap \text{indit}(\sigma) = \text{indit}(\pi \vee \sigma)$  which means that the incidence matrix for the join can be computed as the entry-wise conjunction, i.e.,  $1 \wedge 1 = 1$  and  $1 \wedge 0 = 0 \wedge 1 = 0 \wedge 0 = 0$ .

$$In(\text{indit}(\pi)) \wedge In(\text{indit}(\sigma)) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \wedge \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = In(\text{indit}(\mathbf{1}_U))$$

which in this case is just the incidence matrix of the discrete partition since  $\pi \vee \sigma = \mathbf{1}_U$  since the only equivalence classes are the singletons. In the example, the partitions were chosen so that the blocks in the partition were nicely represented by blocks along the diagonal in the matrix, but that is not the case in general. For instance the partition  $\tau = \{\{a, c\}, \{b\}\}$  has the incidence matrix:

$$In(\text{indit}(\tau)) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

But the important point still holds; as more distinctions are made, that is indicated by non-zero off-diagonal elements being zeroed in the post-distinction incidence matrix. That property carries over to density matrices which in equiprobable QM/Sets are just incidence matrices of partitions normalized to be of trace one.

Given a partition  $\pi = \{B\}_{B \in \pi}$  on a finite set  $U = \{u_1, \dots, u_n\}$ , the ket  $|B\rangle$  for the subset  $B \in \wp(U)$  is represented (in the  $U$ -basis) as the column vector

$$|B\rangle = \begin{bmatrix} \chi_B(u_1) \\ \chi_B(u_2) \\ \vdots \\ \chi_B(u_n) \end{bmatrix}.$$

If we multiply the column vector  $|B\rangle$  times its row vector transpose  $(|B\rangle)^t$ , we get an  $n \times n$  matrix that is the incidence matrix for the binary relation  $B \times B \subseteq U \times U$ :

$$\begin{aligned} |B\rangle(|B\rangle)^t &= \begin{bmatrix} \chi_B(u_1) \\ \chi_B(u_2) \\ \vdots \\ \chi_B(u_n) \end{bmatrix} \begin{bmatrix} \chi_B(u_1) & \chi_B(u_2) & \cdots & \chi_B(u_n) \end{bmatrix} \\ &= \begin{bmatrix} \chi_B(u_1) & \chi_B(u_1)\chi_B(u_2) & \cdots & \chi_B(u_1)\chi_B(u_n) \\ \chi_B(u_2)\chi_B(u_1) & \chi_B(u_2) & \cdots & \chi_B(u_2)\chi_B(u_n) \\ \vdots & \cdots & \ddots & \vdots \\ \chi_B(u_n)\chi_B(u_1) & \chi_B(u_n)\chi_B(u_2) & \cdots & \chi_B(u_n) \end{bmatrix} = In(B \times B). \end{aligned}$$

The trace (sum of diagonal elements) of  $In(B \times B)$  is  $|B|$  so dividing through by  $|B|$  will normalize the matrix to trace one and that is a density matrix:

$$\rho(B) = \frac{1}{|B|} |B\rangle(|B\rangle)^t = \frac{1}{|B|} In(B \times B).$$

The probability-weighted sum of the block density matrices  $\rho(B)$  for  $B \in \pi$  with  $\Pr(B) = \frac{|B|}{|U|}$  gives the density matrix for the partition  $\pi$ :

$$\rho(\pi) = \sum_{B \in \pi} \Pr(B) \rho(B) = \sum_{B \in \pi} \frac{|B|}{|U|} \frac{1}{|B|} In(B \times B) = \frac{1}{|U|} \sum_{B \in \pi} In(B \times B).$$

The probability weights preserve the property of having trace one. For  $\pi = \{\{a, b\}, \{c\}\}$ , the density matrix is:



$$\rho(\pi) = \frac{1}{|U|} \text{In}(\text{indit}(\pi)) = \frac{1}{|U|} \text{In}(\cup_{B \in \pi} B \times B) = \frac{1}{3} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}.$$

Since QM/ $\mathbb{C}$  can be reformulated using density matrices, so can QM/Sets. In QM/ $\mathbb{C}$ , the density matrix of a normalized state  $|\psi\rangle$  expressed in the an orthonormal basis  $\{v_i\}$  of eigenvectors for an observable  $F : V \rightarrow V$  is (where  $()^\dagger$  indicates the conjugate transpose and  $()^*$  indicates the complex conjugate):

$$\begin{aligned} \rho(\psi) &= |\psi\rangle \langle\psi| = \begin{bmatrix} \langle v_1|\psi\rangle \\ \langle v_2|\psi\rangle \\ \vdots \\ \langle v_n|\psi\rangle \end{bmatrix} \begin{bmatrix} \langle v_1|\psi\rangle^* & \langle v_2|\psi\rangle^* & \cdots & \langle v_n|\psi\rangle^* \end{bmatrix} \\ &= \begin{bmatrix} \|\langle v_1|\psi\rangle\|^2 & \langle v_1|\psi\rangle \langle v_2|\psi\rangle^* & \cdots & \langle v_1|\psi\rangle \langle v_n|\psi\rangle^* \\ \langle v_2|\psi\rangle \langle v_1|\psi\rangle^* & \|\langle v_2|\psi\rangle\|^2 & \cdots & \langle v_2|\psi\rangle \langle v_n|\psi\rangle^* \\ \vdots & \cdots & \ddots & \vdots \\ \langle v_n|\psi\rangle \langle v_1|\psi\rangle^* & \langle v_n|\psi\rangle \langle v_2|\psi\rangle^* & \cdots & \|\langle v_n|\psi\rangle\|^2 \end{bmatrix} \end{aligned}$$

where the diagonal elements add to one since  $|\psi\rangle$  was assumed to be normalized. Moreover,  $\rho(\psi)$  equals its conjugate transpose, i.e.,  $\rho(\psi) = \rho(\psi)^\dagger$  and it is idempotent (i.e., a projection matrix) since  $\rho(\psi) \rho(\psi) = |\psi\rangle \langle\psi| \psi\rangle \langle\psi| = |\psi\rangle \langle\psi| = \rho(\psi)$ .

Given a probability distribution  $(p_1, \dots, p_m)$  over a set of density matrices  $\rho(\psi_i) = |\psi_i\rangle \langle\psi_i|$  for  $i = 1, \dots, m$ , the general definition of a density matrix is the probability-weighted sum  $\rho = \sum_{i=1}^m p_i \rho(\psi_i)$  which retains the property of having trace one. Abstractly, density matrices  $\rho$  can be characterized as having trace one and being positive in the sense that for any other state  $|\varphi\rangle$ ,  $\langle\varphi|\rho|\varphi\rangle = \sum_i p_i \langle\varphi|\psi_i\rangle \langle\psi_i|\varphi\rangle = \sum_i p_i \|\langle\varphi|\psi_i\rangle\|^2 \geq 0$  [49, p. 101]. The density matrices  $\rho(\psi)$  constructed from one normalized state are called *pure* and can be characterized by the property of being idempotent (i.e., being projection matrices)  $\rho(\psi) \rho(\psi) = \rho(\psi)$  so that  $\text{tr}[\rho(\psi)^2] = 1$ ; otherwise a density matrix is *mixed* if  $\rho^2 \neq \rho$  and then  $\text{tr}(\rho^2) < 1$ .

Let  $\hat{O} : V \rightarrow V$  be any linear operator on  $V$  and let  $\{v_i\}$  be an orthonormal basis for  $V$ . Then what is the trace of the matrix representation of  $\hat{O}$  in the basis. The effect of applying the operator to any vector  $v$  is (using the resolution of unity  $\sum_i |v_i\rangle \langle v_i| = I$ ):

$$\hat{O}|v\rangle = \sum_{i,j} |v_i\rangle \langle v_i|\hat{O}|v_j\rangle \langle v_j|v\rangle = \sum_i \left( \sum_j \langle v_i|\hat{O}|v_j\rangle \langle v_j|v\rangle \right) |v_i\rangle$$

so the  $i, j$  entry in the matrix representation of  $\hat{O}$  is  $\langle v_i | \hat{O} | v_j \rangle$  and thus the trace is  $\text{tr} [\hat{O}] = \sum_i \langle v_i | \hat{O} | v_i \rangle$ . When a pure state density matrix  $\rho(\psi)$  expressed in the O.N. eigen-basis  $\{v_i\}$  of an observable  $F$ , then the probability of the  $F$ -measurement of  $\psi$  returning an eigenstate  $v_i$  is:

$$\begin{aligned} \text{Pr}(v_i|\psi) &= \frac{\|\langle v_i|\psi\rangle\|^2}{\|\psi\|^2} = \|\langle v_i|\psi\rangle\|^2 \\ &= \langle v_i|\psi\rangle \langle \psi|v_i\rangle = \sum_j \langle v_j|\rho(\psi) P_{v_i}|v_j\rangle = \text{tr}[\rho(\psi) P_{v_i}] \end{aligned}$$

where  $P_{v_i}$  is the projection to  $v_i$  and  $P_{v_i}|v_j\rangle = \delta_{ij}|v_i\rangle$ . Then for  $P_\lambda$  as the projection to the eigenspace of the eigenvalue  $\lambda$ , we have:

$$\begin{aligned} \text{Pr}(\lambda|\psi) &= \|P_\lambda(\psi)\|^2 = \sum_{Fv_i=\lambda v_i} \|\langle v_i|\psi\rangle\|^2 \\ &= \sum_{Fv_i=\lambda v_i} \langle v_i|\psi\rangle \langle \psi|v_i\rangle = \sum_j \langle v_j|\rho(\psi) P_\lambda|v_j\rangle = \text{tr}[\rho(\psi) P_\lambda] \end{aligned}$$

since  $P_\lambda|v_j\rangle = |v_j\rangle$  if  $Fv_j = \lambda v_j$  and 0 otherwise.

For the observable  $F$  and state  $\psi$ , we have:

$$\begin{aligned} \text{tr}[\rho(\psi) F] &= \sum_i \langle v_i|\rho(\psi) F|v_i\rangle = \sum_i \langle v_i|\rho(\psi) \lambda|v_i\rangle = \\ &= \sum_\lambda \lambda \sum_{Fv_i=\lambda v_i} \langle v_i|\psi\rangle \langle \psi|v_i\rangle = \sum_\lambda \lambda \text{Pr}(\lambda|\psi) = \langle F \rangle_\psi \end{aligned}$$

so the average value of the observable  $F$  when measuring the state  $\psi$  is:

$$\langle F \rangle_\psi = \text{tr}[\rho(\psi) F].$$

All these results in QM/ $\mathbb{C}$  carry over, *mutatis mutandis*, to QM/Sets. Let  $f : U \rightarrow \mathbb{R}$  be a random variable on the equiprobable outcome set  $U = \{u_1, \dots, u_n\}$ . A superposition event  $S \in \wp(U)$  has a density matrix:

$$\rho(S) = \frac{1}{|S|} \begin{bmatrix} \chi_S(u_1) & \chi_S(u_1)\chi_S(u_2) & \cdots & \chi_S(u_1)\chi_S(u_n) \\ \chi_S(u_2)\chi_S(u_1) & \chi_S(u_2) & \cdots & \chi_S(u_2)\chi_S(u_n) \\ \vdots & \cdots & \ddots & \vdots \\ \chi_S(u_n)\chi_S(u_1) & \chi_S(u_n)\chi_S(u_2) & \cdots & \chi_S(u_n) \end{bmatrix} = \frac{1}{|S|} \text{In}(S \times S).$$

The projection matrix  $P_T$  for  $T \subseteq U$  corresponding to the projection operator  $T \cap () : \wp(U) \rightarrow \wp(U)$  is the diagonal matrix with the characteristic function  $\chi_T$  along the diagonal:

$$P_T = \begin{bmatrix} \chi_T(u_1) & 0 & \cdots & 0 \\ 0 & \chi_T(u_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \chi_T(u_n) \end{bmatrix}$$

Then the matrix  $P_T \rho(S)$  has the  $i^{\text{th}}$  diagonal entry:  $(P_T \rho(S))_{ii} = \frac{1}{|S|} \chi_T(u_i) \chi_S(u_i) = \frac{1}{|S|} \chi_{T \cap S}(u_i)$  so the trace is:

$$\text{tr}[P_T \rho(S)] = \text{tr}[\rho(S) P_T] = \frac{1}{|S|} \sum_{i=1}^n \chi_{T \cap S}(u_i) = \frac{|T \cap S|}{|S|} = \Pr(T|S).$$

In particular for  $T = f^{-1}(r)$ , we have:

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

Previously in QM/Sets we defined an eigenvector  $S$  of  $f : U \rightarrow \mathbb{R}$  with the eigenvalue  $r$  is they satisfied the formal equation:  $f \upharpoonright S = rS$  (where  $rS$  stands for the result of assigning  $r$  to each element in  $S$ ). Another version of the same definitions can be restated using matrices that can combine with density matrices. Let  $f$  as a matrix be the diagonal matrix  $[f]$  with the eigenvalues  $f(u_i)$  along the diagonal:

$$[f] = \begin{bmatrix} f(u_1) & 0 & \cdots & 0 \\ 0 & f(u_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & f(u_n) \end{bmatrix}.$$

Each subset  $S \in \wp(U)$  is represented by the column vector  $|S\rangle = [\chi_S(u_1), \dots, \chi_S(u_n)]^t$  so the eigenvectors  $|S\rangle$  for  $[f]$  are the ones that, for some  $r \in \mathbb{R}$ , satisfy the equation:

$$[f] |S\rangle = r |S\rangle.$$

Then the average value of the random variable  $f$  conditioned on the event  $S \in \wp(U)$  is:

$$\begin{aligned} \langle f \rangle_S &= \text{tr}[[f] \rho(S)] = \sum_{i=1}^n ([f] \rho(S))_{ii} = \frac{1}{|S|} \sum_i f(u_i) \chi_S(u_i) \\ &= \frac{1}{|S|} \sum_{r \in f(U)} \sum_{u_i \in f^{-1}(r)} r \chi_S(u_i) = \frac{1}{|S|} \sum_{r \in f(U)} r |f^{-1}(r) \cap S| = \sum_{r \in f(U)} r \Pr(r|S). \end{aligned}$$

The previous development of QM/Sets has assumed for the sake of simplicity that the outcomes in  $U$  are equiprobable, but density matrices make it easy to incorporate different point probabilities such as  $p = (p_1, \dots, p_n)$  assigned to the points of  $U = \{u_1, \dots, u_n\}$  so that  $\Pr(S) = \sum_{u_i \in S} p_i$ . Since a point  $u_i$  with  $p_i = 0$  does not figure into the calculations, we will assume that all probabilities  $p_i > 0$ . Then the normalized column vector for  $\emptyset \neq S \in \wp(U)$  is:  $|S\rangle = \frac{1}{\sqrt{\Pr(S)}} [\chi_S(u_1) \sqrt{p_1}, \dots, \chi_S(u_n) \sqrt{p_n}]^t$  so that its density matrix is:

$$\frac{1}{\Pr(S)} \begin{bmatrix} \rho(S) = |S\rangle \langle S| = \\ \chi_S(u_1) p_1 & \chi_S(u_1) \chi_S(u_2) \sqrt{p_1 p_2} & \cdots & \chi_S(u_1) \chi_S(u_n) \sqrt{p_1 p_n} \\ \chi_S(u_2) \chi_S(u_1) \sqrt{p_2 p_1} & \chi_S(u_2) p_2 & \cdots & \chi_S(u_2) \chi_S(u_n) \sqrt{p_2 p_n} \\ \vdots & \cdots & \ddots & \vdots \\ \chi_S(u_n) \chi_S(u_1) \sqrt{p_n p_1} & \chi_S(u_n) \chi_S(u_2) \sqrt{p_n p_2} & \cdots & \chi_S(u_n) p_n \end{bmatrix}.$$

Then for any  $T \in \wp(U)$ ,

$$\text{tr}[P_T \rho(S)] = \text{tr}[\rho(S) P_T] = \frac{1}{\Pr(S)} \sum_{i=1}^n \chi_{T \cap S}(u_i) p_i = \frac{\Pr(T \cap S)}{\Pr(S)} = \Pr(T|S)$$

so that for  $T = f^{-1}(r)$ :

$$\Pr(r|S) = \Pr(f^{-1}(r)|S) = \text{tr}[P_{f^{-1}(r)} \rho(S)] = \frac{1}{\Pr(S)} \sum_{i=1}^n \chi_{f^{-1}(r) \cap S}(u_i) p_i = \frac{\Pr(f^{-1}(r) \cap S)}{\Pr(S)}.$$

And, similarly:

$$\begin{aligned} \langle f \rangle_S &= \text{tr}[[f] \rho(S)] = \sum_{i=1}^n ([f] \rho(S))_{ii} = \frac{1}{\Pr(S)} \sum_i f(u_i) \chi_S(u_i) p_i \\ &= \frac{1}{\Pr(S)} \sum_{r \in f(U)} \sum_{u_i \in f^{-1}(r)} r \chi_S(u_i) p_i = \frac{1}{\Pr(S)} \sum_{r \in f(U)} r \Pr(f^{-1}(r) \cap S) = \\ &= \sum_{r \in f(U)} r \Pr(r|S). \end{aligned}$$

Thus for a fixed outcome space  $U$  with point probabilities, QM/Sets reproduces finite probability theory in a QM setting using superposition events.

## 5.2 Revisiting measurement with density matrices

Given a partition  $\pi = \{B\}_{B \in \pi}$  on  $U = \{u_1, \dots, u_n\}$  with point probabilities  $p = (p_1, \dots, p_n)$ , the pure density matrix for a block  $B \in \wp(U)$  is given by:

$$\rho(B)_{ij} = \begin{cases} \frac{1}{\Pr(B)} \sqrt{p_j p_k} & \text{if } u_i, u_j \in B \\ 0 & \text{otherwise.} \end{cases} \quad \text{so that } \text{tr}[\rho(B)] = \frac{1}{\Pr(B)} \sum_{u_i \in B} p_i = 1.$$

Then the probability-weighted sum gives the mixed density matrix for partition:

$$\rho(\pi) = \sum_{B \in \pi} \Pr(B) \rho(B) \quad \text{where } \rho(\pi)_{ij} = \begin{cases} \sqrt{p_i p_j} & \text{if } (u_i, u_j) \in \text{indit}(\pi) \\ 0 & \text{if } (u_i, u_j) \in \text{dit}(\pi). \end{cases}$$

so that  $\text{tr}[\rho(\pi)] = \sum_{i=1}^n \rho(\pi)_{ii} = \sum_{i=1}^n \sqrt{p_i p_i} = 1$ , where  $\text{indit}(\pi)$  is the equivalence relation  $\cup_{B \in \pi} B \times B$  on  $U \times U$  defined by  $\pi$ .

One of the most important facts about density matrices is the characterization of the trace of the density matrix squared. In the case at hand of  $\rho(\pi)$ , we have:

$$(\rho(\pi)^2)_{ii} = \sum_{u_k \text{ s.t. } (u_i, u_k) \in \text{indit}(\pi)} \sqrt{p_i p_k} \sqrt{p_k p_i} = p_i \sum_{u_k \in B \text{ where } u_i \in B} p_k = p_i \Pr(B)$$

so that:  $\sum_{u_i \in B} p_i \Pr(B) = \Pr(B) \sum_{u_i \in B} p_i = \Pr(B)^2$ , and hence:

$$\text{tr}(\rho(\pi)^2) = \sum_{i=1}^n (\rho(\pi)^2)_{ii} = \sum_{B \in \pi} \sum_{u_i \in B} p_i \Pr(B) = \sum_{B \in \pi} \Pr(B)^2.$$

Then we have the density-matrix version of the logical entropy of a partition:

$$h(\pi) = 1 - \sum_{B \in \pi} \Pr(B)^2 = 1 - \text{tr}(\rho(\pi)^2).$$

With point probabilities  $\{p_i\}$ ,  $\Pr(B)^2 = (\sum_{u_i \in B} p_i)^2 = \sum_{u_i \in B} p_i^2 + \sum_{u_i, u_j \in B, u_i \neq u_j} p_i p_j$  and  $1 = (\sum_{i=1}^n p_i)^2 = \sum_{i=1}^n p_i^2 + \sum_{i \neq j} p_i p_j$  so that:

$$\begin{aligned} h(\pi) &= 1 - \sum_{B \in \pi} \Pr(B)^2 \\ &= \sum_{i=1}^n p_i^2 + \sum_{i \neq j} p_i p_j - \sum_{B \in \pi} \sum_{u_i \in B} p_i^2 - \sum_{B \in \pi} \sum_{u_i, u_j \in B, u_i \neq u_j} p_i p_j \\ &= \sum_{i \neq j} p_i p_j - \sum_{B \in \pi} \sum_{u_i, u_j \in B, u_i \neq u_j} p_i p_j = \sum_{i \neq j} p_i p_j - \sum_{(u_i, u_j) \in \text{indit}(\pi), u_i \neq u_j} p_i p_j = \\ &\quad \sum_{(u_i, u_j) \in \text{dit}(\pi)} p_i p_j. \end{aligned}$$

Thus with point probabilities, the logical entropy is still the probability that in two independent draws from  $U$ , a distinction of  $\pi$  is obtained.

In QM/ $\mathbb{C}$ , a normalized vector  $\psi \in V$  expressed in the O.N. basis  $\{v_i\}_{i=1}^n$  of eigenvectors for an observable  $F$  is  $|\psi\rangle = \sum_{i=1}^n \langle v_i | \psi \rangle |v_i\rangle$  and its density matrix is the projection matrix:  $\rho(\psi) = |\psi\rangle \langle \psi|$  where  $\rho(\psi)_{ij} = \langle v_i | \psi \rangle \langle \psi | v_j \rangle$  so  $\left\| \rho(\psi)_{ij} \right\|^2 = \|\langle v_i | \psi \rangle\|^2 \|\langle v_j | \psi \rangle\|^2$  which is the probability of an  $F$ -measurement of  $\psi$  returning  $v_i$  times the probability of an  $F$ -measurement of  $\psi$  returning  $v_j$ .

The *quantum logical entropy* of a density matrix  $\rho$  in QM/ $\mathbb{C}$  has the same definition:  $h(\rho) = 1 - \text{tr}[\rho^2]$ . This formula has been also called the “mixedness” [38, p. 5] of the state  $\rho$  since  $\rho$  is a mixed state if  $\text{tr}[\rho^2] < 1$ . The important paper of Manfredi and Feix [45] also advocates the same formula  $1 - \text{tr}[\rho^2]$  for the quantum notion of entropy which they denote as  $S_2$ .

In QM/ $\mathbb{C}$ , the operation of a projective  $F$ -measurement on a density matrix  $\rho$  is described by the:

$$\rho \rightsquigarrow \hat{\rho} = \sum_{\lambda} P_{\lambda} \rho P_{\lambda}$$

Lüders mixture operation in QM/ $\mathbb{C}$  ([3, p. 279]; [42])

which transforms the pre-measurement density matrix  $\rho$  into the post-measurement density matrix  $\hat{\rho}$ . The measurement of a normalized pure state  $|\psi\rangle$  results in the state  $P_{\lambda}|\psi\rangle$  with the probability  $p_{\lambda} = \text{tr}[P_{\lambda}\rho(\psi)] = \text{Pr}(\lambda|\psi)$  where  $\rho(\psi) = |\psi\rangle\langle\psi|$ . The projected resultant state  $P_{\lambda}|\psi\rangle$  has the density matrix  $\frac{P_{\lambda}|\psi\rangle\langle\psi|P_{\lambda}}{\text{tr}[P_{\lambda}\rho(\psi)]} = \frac{P_{\lambda}\rho(\psi)P_{\lambda}}{\text{tr}[P_{\lambda}\rho(\psi)]}$  so the mixed state describing the probabilistic results of the measurement is [?, p. 101 or p. 515]:

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

Since we have density matrices and projection matrices to the eigenspaces  $\wp(f^{-1}(r))$  of a random variable  $f : U \rightarrow \mathbb{R}$  in QM/Sets, the corresponding Lüders mixture operation is:

$$\rho \rightsquigarrow \hat{\rho} = \sum_{r \in f(i)} P_{f^{-1}(r)} \rho P_{f^{-1}(r)}$$

Lüders mixture operation in QM/Sets.

Consider first the operation on a pure state density matrix  $\rho(S)$  for  $\emptyset \neq S \in \wp(U)$ :

$$\rho(S)_{ij} = \begin{cases} \frac{1}{\text{Pr}(S)} \sqrt{p_i p_j} & \text{if } u_i, u_j \in S \\ 0 & \text{otherwise.} \end{cases}.$$

To check on purity, we can use the criterion of purity if  $\text{tr}[\rho(S)^2] = 1$  (and a mixture otherwise when  $\text{tr}[\rho(S)^2] < 1$ ), where:

$$\begin{aligned} (\rho(S)^2)_{ii} &= \frac{1}{\text{Pr}(S)^2} \sum_{u_k \text{ s.t. } u_i, u_k \in S} \sqrt{p_i p_k} \sqrt{p_k p_i} \\ &= \frac{p_i}{\text{Pr}(S)^2} \sum_{u_k \in S \text{ where } u_i \in S} p_k = \frac{p_i \text{Pr}(S)}{\text{Pr}(S)^2} = \frac{p_i}{\text{Pr}(S)} \text{ if } u_i \in S \text{ and } 0 \text{ otherwise,} \end{aligned}$$

so the only non-zero elements along the diagonal are for  $u_i \in S$ , and thus

$$\text{tr} [\rho(S)^2] = \sum_{u_i \in S} (\rho(S)^2)_{ii} = \sum_{u_i \in S} \frac{p_i}{\text{Pr}(S)} = \frac{\text{Pr}(S)}{\text{Pr}(S)} = 1.$$

The projection matrix  $P_{f^{-1}(r)}$  is given by:

$$P_{f^{-1}(r)} = \begin{bmatrix} \chi_{f^{-1}(r)}(u_1) & 0 & \cdots & 0 \\ 0 & \chi_{f^{-1}(r)}(u_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \chi_{f^{-1}(r)}(u_n) \end{bmatrix}.$$

On each block of the partition  $\{f^{-1}(r)\}_{r \in f(U)}$ , the operation of the projection matrices is:

$$(P_{f^{-1}(r)} \rho(S) P_{f^{-1}(r)})_{ij} = \begin{cases} \frac{1}{\text{Pr}(S)} \sqrt{p_j p_k} & \text{if } u_i, u_j \in f^{-1}(r) \cap S \\ 0 & \text{otherwise.} \end{cases}$$

so the sum over the eigenvalues  $r \in f(U)$  gives

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

where

$$(\hat{\rho}(S))_{ij} = \begin{cases} \frac{1}{\text{Pr}(S)} \sqrt{p_i p_j} & \text{if } u_i, u_j \in f^{-1}(r) \cap S \text{ for some } r \in f(U) \\ 0 & \text{otherwise.} \end{cases}.$$

Checking the trace:

$$\text{tr} [\hat{\rho}(S)] = \sum_{i=1}^n (\hat{\rho}(S))_{ii} = \frac{1}{\text{Pr}(S)} \sum_{r \in f(U), u_i \in f^{-1}(r) \cap S} p_i = \frac{1}{\text{Pr}(S)} \sum_{u_i \in S} p_i = 1.$$

The non-zero off-diagonal elements in  $\hat{\rho}(S)$  correspond to the pairs  $\{u_i, u_j\}$  where  $u_i, u_j \in f^{-1}(r) \cap S$  for some  $r$  so they ‘cohere’ together or are rendered indistinct in the superposition state  $f^{-1}(r) \cap S$ .

[T]he off-diagonal terms of a density matrix ... are often called *quantum coherences* because they are responsible for the interference effects typical of quantum mechanics that are absent in classical dynamics. [3, p. 177]

If we distinguish the elements of  $S$  by their  $f$ -eigenvalues, then we have a partition of nonempty blocks  $\{f^{-1}(r) \cap S\}_{r \in f(U)}$ , which is the partition  $f^{-1} = \{f^{-1}(r)\}_{r \in f(U)}$  restricted to  $S$ . All and only the non-zero off-diagonal elements of  $\rho(S)$  are zeroed which correspond to some  $u_i$  and  $u_j$  with different  $f$ -eigenvalues. Another way to obtain  $\hat{\rho}(S)$  is to start with the pure state density matrices  $\rho(f^{-1}(r) \cap S)$  when that intersection is non-empty and then construct the mixed state density matrix which weighs the pure state density matrices with the probabilities  $\Pr(r|S) = \frac{\Pr(f^{-1}(r) \cap S)}{\Pr(S)}$  so that we have the mixed state density matrix:

$$\rho = \sum_{r \in f(U)} \frac{\Pr(f^{-1}(r) \cap S)}{\Pr(S)} \rho(f^{-1}(r) \cap S).$$

For the pure state density matrices, the  $i, j$ -entries are:

$$\rho(f^{-1}(r) \cap S)_{ij} = \begin{cases} \frac{1}{\Pr(f^{-1}(r) \cap S)} \sqrt{p_i p_j} & \text{if } u_i, u_j \in f^{-1}(r) \cap S \\ 0 & \text{otherwise.} \end{cases}$$

and thus the  $i, j$ -entries in  $\rho$  are:

$$\rho_{ij} = \frac{\Pr(f^{-1}(r) \cap S)}{\Pr(S)} \frac{1}{\Pr(f^{-1}(r) \cap S)} \sqrt{p_i p_j} = \frac{1}{\Pr(S)} \sqrt{p_i p_j}$$

if  $u_i, u_j \in f^{-1}(r) \cap S$  for some  $r \in f(U)$ , and 0 otherwise, so  $\rho = \hat{\rho}(S)$ .

This means that when the projective measurement of a quantum state  $\psi$  by an observable  $F$  in QM/ $\mathbb{C}$  is restated in QM/Sets, then it is just the stochastic sampling of the random variable  $f$  conditioned on the superposition event  $S$ —with the same probabilities for the classical event  $S \subseteq U$ .

If we started with  $S = U$ , then  $\rho(U) = \rho(\mathbf{0}_U)$  for the indiscrete partition  $\mathbf{0}_U$  and the measurement operation produces  $\hat{\rho}(\mathbf{0}_U) = \rho(f^{-1})$ , the density matrix for the partition  $f^{-1} = \{f^{-1}(r)\}_{r \in f(U)}$  where:

$$\rho(f^{-1}) = \begin{cases} \sqrt{p_i p_j} & \text{if } u_i, u_j \in f^{-1}(r) \\ 0 & \text{otherwise} \end{cases}$$

which is the special case of  $\rho = \sum_{r \in f(U)} \frac{\Pr(f^{-1}(r) \cap S)}{\Pr(S)} \rho(f^{-1}(r) \cap S)$  for  $S = U$ .

The key thing to note is that the measurement transition  $\rho(\mathbf{0}_U) \rightsquigarrow \hat{\rho}(\mathbf{0}_U) = \rho(f^{-1})$  is just the transition of the join operation on partitions since  $\mathbf{0}_U \vee f^{-1} = f^{-1}$ . This can be checked by considering a measurement of the mixed state  $\rho(f^{-1})$  by the compatible observable  $g : U \rightarrow \mathbb{R}$ , which is analogous to performing two measurements on the original state  $\psi$  by two commuting observables  $F$  and  $G$  in QM/ $\mathbb{C}$ . The Lüders mixture operation gives:



$$\hat{\rho}(f^{-1}) = \sum_{s \in g(U)} P_{g^{-1}(s)} \rho(f^{-1}) P_{g^{-1}(s)}$$

which gives:

$$\left\{ \begin{array}{l} (\hat{\rho}(f^{-1}))_{ij} = \\ \sqrt{p_i p_j} \text{ if } u_i, u_j \in g^{-1}(s) \cap f^{-1}(r) \text{ for some } r \in f(U) \text{ and } s \in g(U) \\ 0 \text{ otherwise.} \end{array} \right.$$

The claim is that this is just the density matrix  $\rho(f^{-1} \vee g^{-1})$  for the join of the partitions  $\{f^{-1}(r)\}_{r \in f(U)}$  and  $\{g^{-1}(s)\}_{s \in g(U)}$  whose blocks are the nonempty intersections  $f^{-1}(r) \cap g^{-1}(s)$ , block which can be characterized by the ordered pair  $(r, s)$  of eigenvalues. The density matrix  $\rho(f^{-1} \vee g^{-1})$  is formed as the probability-weighted sum of the pure state density matrices for the nonempty blocks  $\rho(f^{-1}(r) \cap g^{-1}(s))$  where:

$$\rho(f^{-1}(r) \cap g^{-1}(s))_{ij} = \left\{ \begin{array}{l} \frac{1}{\Pr(f^{-1}(r) \cap g^{-1}(s))} \sqrt{p_i p_j} \text{ if } u_i, u_j \in f^{-1}(r) \cap g^{-1}(s) \\ 0 \text{ otherwise.} \end{array} \right.$$

so in the probability-weighted sum:

$$\rho(f^{-1} \vee g^{-1}) = \sum_{r,s \text{ s.t. } f^{-1}(r) \cap g^{-1}(s) \neq \emptyset} \Pr(f^{-1}(r) \cap g^{-1}(s)) \rho(f^{-1}(r) \cap g^{-1}(s))$$

and thus the  $i, j$ -entry is:

$$\rho(f^{-1} \vee g^{-1})_{ij} = \left\{ \begin{array}{l} \sqrt{p_i p_j} \text{ if } u_i, u_j \in f^{-1}(r) \cap g^{-1}(s) \\ 0 \text{ otherwise.} \end{array} \right. = (\hat{\rho}(f^{-1}))_{ij}.$$

Hence the combined effect of projective measurements of  $U$  by two commuting r.v.s  $f$  and  $g$  is just the single measurement by random variable  $(f, g) : U \rightarrow \mathbb{R} \times \mathbb{R}$  where  $(f, g)(u) = (f(u), g(u))$  whose inverse-image partition is the join  $f^{-1} \vee g^{-1} = \{f^{-1}(r) \cap g^{-1}(s) \neq \emptyset\}_{(r,s) \in (f,g)(U)}$ . Hence a complete set of commuting r.v.s is a set such that the join of their inverse-image partitions is the discrete partition  $\mathbf{1}_U = \{\{u\}\}_{u \in U}$ —all of which is the QM/Sets version of Dirac's Complete Set of Commuting Operators (CSCO) where the simultaneous eigenvectors of all the operators will distinguish or decompose all the simultaneous eigenvectors which than only generate one-dimensional eigenspaces. In each case, those simultaneous eigenvectors can be characterized by the ordered  $m$ -tuple of eigenvalues if there are  $m$  r.v.s or operators in the complete set.

The main results can be illustrated with some simple examples. Let  $U = \{a, b, c\}$  with the respective point probabilities  $p = (p_1, p_2, p_3)$  and let  $f : U \rightarrow \mathbb{R}$  be the r.v.  $f(a) = f(b) = 1$  and  $f(c) = 0$  so that  $f = \chi_{\{a,b\}}$  and  $f^{-1} = \{\{a, b\}, \{c\}\}$ . Then we have the density matrices:

$$\begin{aligned} \rho(U) = \rho(\mathbf{0}_U) &= \begin{bmatrix} p_1 & \sqrt{p_1 p_2} & \sqrt{p_1 p_3} \\ \sqrt{p_2 p_1} & p_2 & \sqrt{p_2 p_3} \\ \sqrt{p_3 p_1} & \sqrt{p_3 p_2} & p_3 \end{bmatrix}, \\ \rho(\{a, b\}) &= \begin{bmatrix} \frac{1}{p_1+p_2} p_1 & \frac{1}{p_1+p_2} \sqrt{p_1 p_2} & 0 \\ \frac{1}{p_1+p_2} \sqrt{p_2 p_1} & \frac{1}{p_1+p_2} p_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ \text{and } \rho(\{c\}) &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{p_3} p_3 = 1 \end{bmatrix} \text{ so} \\ \rho(f^{-1}) &= \frac{1}{p_1+p_2} \rho(\{a, b\}) + \frac{1}{p_3} \rho(\{c\}) = \begin{bmatrix} p_1 & \sqrt{p_1 p_2} & 0 \\ \sqrt{p_2 p_1} & p_2 & 0 \\ 0 & 0 & p_3 \end{bmatrix}. \end{aligned}$$

Then the Lüders mixture operation of  $f$ -measurement on  $\rho(\mathbf{0}_U)$  should yield  $\hat{\rho}(\mathbf{0}_U) = \rho(\mathbf{0}_U \vee f^{-1}) = \rho(f^{-1})$ .

$$\begin{aligned} \hat{\rho}(\mathbf{0}_U) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_1 & \sqrt{p_1 p_2} & \sqrt{p_1 p_3} \\ \sqrt{p_2 p_1} & p_2 & \sqrt{p_2 p_3} \\ \sqrt{p_3 p_1} & \sqrt{p_3 p_2} & p_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &+ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_1 & \sqrt{p_1 p_2} & \sqrt{p_1 p_3} \\ \sqrt{p_2 p_1} & p_2 & \sqrt{p_2 p_3} \\ \sqrt{p_3 p_1} & \sqrt{p_3 p_2} & p_3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} p_1 & \sqrt{p_1 p_2} & 0 \\ \sqrt{p_1 p_2} & p_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & p_3 \end{bmatrix} = \begin{bmatrix} p_1 & \sqrt{p_1 p_2} & 0 \\ \sqrt{p_2 p_1} & p_2 & 0 \\ 0 & 0 & p_3 \end{bmatrix} = \rho(f^{-1}). \checkmark \end{aligned}$$

A further measurement of  $\rho(f^{-1})$  by  $g : U \rightarrow \mathbb{R}$  where  $g(a) = g(c) = 1$  and  $g(b) = 0$  so  $g = \chi_{\{a,c\}}$  should yield  $f^{-1} \vee g^{-1} = \mathbf{1}_U$ . The Lüders mixture operation yields:

$$\begin{aligned}
\hat{\rho}(f^{-1}) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_1 & \sqrt{p_1 p_2} & 0 \\ \sqrt{p_2 p_1} & p_2 & 0 \\ 0 & 0 & p_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&+ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_1 & \sqrt{p_1 p_2} & 0 \\ \sqrt{p_2 p_1} & p_2 & 0 \\ 0 & 0 & p_3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} p_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & p_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & p_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} p_1 & 0 & 0 \\ 0 & p_2 & 0 \\ 0 & 0 & p_3 \end{bmatrix} = \rho(\mathbf{1}_U). \checkmark
\end{aligned}$$

### 5.3 Classical events and superposition events

The density matrix  $\rho(\mathbf{1}_U)$  is the mixed state density matrix for the classical event  $U \subseteq U$  where all the eigenvectors  $\{u_i\} \in \wp(U)$  are decomposed, decohered, and distinguished to create the classical event  $U$  [3, p. 176]. We previously defined the pure state density matrix  $\rho(S)$  for a subset  $S \in \wp(U)$  which construed  $S$  as a superposition event, not the classical event associated with  $S \subseteq U$ . The density matrix machinery allows us to differentiate between the superposition event  $S \in \wp(U) \cong \mathbb{Z}_2^{|U|}$  and the classical event  $S \subseteq U$  associated with a subset  $S$ . These are the two ways, the classical way and the quantum way, to interpret a subset  $S$ . The density matrix associated with the classical event  $S \subseteq U$  is the diagonal matrix with the diagonal entries:

$$\rho(\delta S)_{ii} = \begin{cases} \frac{p_i}{\Pr(S)} & \text{if } u_i \in S \\ 0 & \text{otherwise} \end{cases}.$$

The distinction between the superposition event represented by pure state  $\rho(S)$  and the classical event represented by the completely decomposed mixed state  $\rho(\delta S)$  does not show up in the probability formulas for measuring in the  $U$ -basis since the two states cannot be distinguished by such measurements. At first, this might seem like a bug in QM/Sets but it is a feature since the same lack of distinguishability occurs in QM/C. Substituting our notation:

[A]s long as one measures the polarization in the basis  $\{|h\rangle, |v\rangle\}$ , it is not possible to detect a difference between  $\rho(S)$  and  $\rho(\delta S)$ : The probabilities  $\Pr(h)$  and  $\Pr(v)$  of detecting horizontal and vertical polarization, respectively, are in both cases equal to  $\|c_h\|^2$  and  $\|c_v\|^2$ . [3, p. 176]

To distinguish the two states, one needs to measure them in a different basis, and that can also be done in QM/Sets.

In a fair-coin-tossing example where  $U = \{H, T\}$  is taken as the outcome set, there is another basis  $U' = \{H', T'\}$  where  $\{H'\} = \{H, T\}$  and  $\{T'\} = \{T\}$  which is a basis since  $\{H'\} + \{T'\} = \{H, T\} + \{T\} = \{H\}$  (mod 2 addition) and  $\{T'\} = \{T\}$ . The third basis is for  $U'' = \{H'', T''\}$  where  $\{H''\} = \{H\}$  and  $\{T''\} = \{H, T\}$ . Those are the only basis sets since by Gauss's formula for  $n = 2$ , there are  $(2^2 - 1)(2^2 - 2^1) \frac{1}{2!} = 3$  (unordered) bases of  $\mathbb{Z}_2^2$ . Since we have different bases for  $\mathbb{Z}_2^2$ , we can consider a ket as an abstract vector that can be represented in different bases, e.g.,  $\{H\}$ ,  $\{H', T'\}$ , and  $\{H''\}$  all represent the same abstract vector in different bases. Then we can form a ket-table where each row represents a ket. In  $\mathbb{Z}_2^2$ , there are  $2^2 - 1 = 3$  non-zero abstract vectors, each corresponding to a row in the ket-table Table 5.1.

$U$ -basis	$U'$ -basis	$U''$ -basis
$\{H, T\}$	$\{H'\}$	$\{T''\}$
$\{H\}$	$\{H', T'\}$	$\{H''\}$
$\{T\}$	$\{T'\}$	$\{H'', T''\}$

Table 5.1: Ket-table for  $\mathbb{Z}_2^2$ .

Each ket or abstract vector is a superposition in one basis and a singleton event in the other two bases.

We noted previously that we could not distinguish the classical mixture event  $U$  associated with  $\rho(\delta U)$  from the superposition event associated with  $\rho(U)$  when measured in the  $U$ -basis. For instance, the probability of getting heads in the two cases is:

$$\begin{aligned} \Pr(H|\rho(\delta U)) &= \text{tr} [P_{\{H\}}\rho(\delta U)] = \text{tr} \left[ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \right] = \text{tr} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix} = \frac{1}{2} \\ \Pr(H|\rho(U)) &= \text{tr} [P_{\{H\}}\rho(U)] = \text{tr} \left[ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \right] = \text{tr} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix} = \frac{1}{2}. \end{aligned}$$

But the two events can be distinguished when measured in a different basis such as the  $U'$ -basis.

The vector  $\{H\}$  is expressed in the  $U$ -basis by the column vector  $\begin{bmatrix} 1 \\ 0 \end{bmatrix}_U$  (the subscript indicating the basis) and in the  $U'$ -basis by the column vector  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}_{U'}$ . The basis conversion matrix is

$$C_{U \rightarrow U'} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \text{ so } \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}_U = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{U'}.$$

Hence converting the superposition  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}_U$  or  $\{H, T\}$  to the  $U'$ -basis gives:

$$C_{U \rightarrow U'} \begin{bmatrix} 1 \\ 1 \end{bmatrix}_U = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}_U = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{U'} \text{ or } \{H'\} \text{ so its density matrix (computing in the reals) is } \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{U'} \begin{bmatrix} 1 & 0 \end{bmatrix}_{U'} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}_{U'}.$$

The classical mixed event  $U$  is the half-half mixture of  $\{H\}$  and  $\{T\}$ . The basis conversion for  $\{H\}$  gives  $C_{U \rightarrow U'} \begin{bmatrix} 1 \\ 0 \end{bmatrix}_U =$

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}_U = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{U'} \text{ so the associated real density matrix is:}$$

$$\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}_{U'} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}_{U'} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}_{U'}$$

and for  $\{T\}$ ,  $C_{U \rightarrow U'} \begin{bmatrix} 0 \\ 1 \end{bmatrix}_U = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}_U = \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{U'}$  so its real density matrix is:

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}_{U'} \begin{bmatrix} 0 & 1 \end{bmatrix}_{U'} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}_{U'}.$$

Their half-half mixture has the density matrix in the  $U'$ -basis:

$$\frac{1}{2} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}_{U'} + \frac{1}{2} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}_{U'} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}_{U'}.$$

We then measure by the partition  $\sigma = \{\{H'\}, \{T'\}\}$  with half-half probabilities so the probability of  $H'$  for the superposition event  $\{H, T\}$  or  $\{H'\}$  in the  $U'$ -basis is:

$$\text{tr} \left[ P_{\{H'\}} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}_{U'} \right] = \text{tr} \left[ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}_{U'} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}_{U'} \right] = \text{tr} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}_{U'} = 1$$

and for the classical mixture of half  $\{H\}$  and half  $\{T\}$  which in the  $U'$ -basis is the mixture of half  $\{H', T'\}$  and half  $\{T'\}$ , is:

$$\text{tr} \left[ P_{\{H'\}} \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}_{U'} \right] = \text{tr} \left[ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}_{U'} \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}_{U'} \right] = \text{tr} \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ 0 & 0 \end{bmatrix}_{U'} = \frac{1}{4}.$$

The first calculation makes intuitive sense since the superposition  $\{H, T\}$  in the  $U$ -basis is the singleton event  $\{H'\}$  in the  $U'$ -basis, so measuring in the  $U'$ -basis for the event  $\{H'\}$  will give  $\{H'\}$  with probability 1. The second calculation makes intuitive sense since it is half-half in the classical mixture whether we get the  $\{T\} = \{T'\}$  event or the  $\{H\} = \{H', T'\}$  event and then the probability of getting  $H'$  is zero for the  $\{T'\}$  event and  $\frac{1}{2}$  for the  $\{H', T'\}$  event so the overall probability of  $\{H'\}$  is  $\frac{1}{4}$ . Thus the two events, the classical mixture of half  $\{H\}$  and half  $\{T\}$ , and the superposition  $\{H, T\}$ , which cannot be distinguished by measurements in the  $U$ -basis, can be distinguished by measurement in the  $U'$ -basis.

The key to understanding the objective indefiniteness interpretation of QM is understanding the indefiniteness interpretation of the superposition states such as  $\{H, T\}$ . It is not a coin being simultaneously heads and tails (or a particle going through both slits). It represents the indefinite state prior to the distinctions being made that differentiate heads from tails.<sup>9</sup> The virtue of the density matrix formalism is that the non-zero off-diagonal elements make the distinction between the classical event of getting heads or tails with equal probability, i.e.,  $\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}_U$ , and the superpo-

sition event  $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}_U$  that in a distinguishing measurement yields head or tails with equal probability. The likelihood of confusion is particularly strong when considering the vector space  $\wp(\{H, T\}) \cong \mathbb{Z}_2^2$ . The vectors in that space are routinely thought of as subsets like the universal event  $\{H, T\} \subseteq U = \{H, T\}$ , but that classical mixture is not a (pure state) vector in  $\wp(\{H, T\}) \cong \mathbb{Z}_2^2$ ; it is the equal mixture of the pure state vectors  $\{H\}$  and  $\{T\}$ . The superposition event  $\{H, T\}$  is a vector in  $\wp(\{H, T\}) \cong \mathbb{Z}_2^2$  and the density matrices give the differences. The density matrix for the superposition event  $\{H, T\}$  has the non-zero off-diagonal elements which represent that the basis vectors  $\{H\}$  and  $\{T\}$  in the  $U$ -basis are cohered together so as to be indistinct between those two outcomes (like two points being in the same block of a set partition). A measurement could distinguish between or decohere those two outcomes to yield the outcomes  $\{H\}$  or  $\{T\}$  with equal probability, i.e., to yield the classical mixture

---

<sup>9</sup>In terms of our intuition-aid (developed more fully in the next chapter) of the two distinguished isosceles triangles that result from reflection about the vertical axis descending from vertex  $a$ ,  ${}_b\overset{a}{\Delta}_c$  and  ${}_c\overset{a}{\Delta}_b$ , the classical mixture is represented by those two definite outcomes, while the superposition event of those two outcomes would be the indefinite state  $\overset{a}{\Delta}$ , prior to the distinction between the triangles with labeled vertices.

of  $\{H\}$  and  $\{T\}$ .

## 5.4 Modeling measurement in QM/Sets

It might be useful to mention some of the literature from the perspective of this conceptual treatment of the measurement problem. The basic problem has been expressed with great clarity and vigor in John S. Bell's famous essay *Against "Measurement"* [5]. One aspect of the problem is the unclear criterion for separating the processes that constitute measurements from the processes of unitary evolution. In much of the literature, the separation hides behind the distinction between "macroscopic" or "classical" on the one hand, and "microscopic" or "quantum" on the other hand—as if there were precise theoretical definitions of those concepts.

The intuitive idea is that "macroscopic" or "classical" does not countenance any superpositions, so that in the interaction between a macroscopic measurement apparatus and the quantum-level superposition state being "measured," the superposition state will be reduced (or "wave-packet" collapsed) in the measurement process. But considering the measurement apparatus as a large quantum system, the interaction can be modeled by a unitary evolution that would seem to leave the indicator states of the measurement apparatus in a superposition state (like Schrödinger's cat).

Quite aside from the ill-defined nature of the macroscopic-microscopic distinction, a "measurement" can take place by distinguishing events solely at the quantum level. Feynman [29, §3.3] considered a case of a quantum-level measurement apparatus. A neutron is scattering off the nuclei of atoms in a crystal. If the nuclei have no spin, then the amplitude for the neutron to be scattered to some given point would be the superposition of the scattering amplitudes off the various nuclei since there is no distinguishing physical event to distinguish between scattering off one nucleus or another. But if all the nuclei had spin in, say, the down direction while the neutron had spin up, then in the scattering interaction, one of the nuclei might flip its spin which would be the microscopic physical event to distinguish that trajectory. Then the probability of the neutron arriving at the given point with its spin reversed (indicating that a spin flip had occurred) would be the sum of the probabilities (not the amplitudes) for those distinguished trajectories over all the nuclei. In that case, the superposition was reduced (the indefinite became definite) and the nucleus with its spin flipped plays the role of a detector registering a hit. The spin-state of the

nuclei served as a quantum-level measuring apparatus to measure which scattering trajectory was taken by the neutron to reach the detector. By illustrating a measurement decoherence at the quantum level, Feynman's treatment is independent of the Zurek-treatment of "decoherence" [70] that involves interaction with a macroscopic system.

The key concept in this treatment of the measurement problem is the notion of distinguishability. Of all the quantum theorists, Richard Feynman is perhaps the clearest in singling out the notion of distinguishability in determining whether or not a superposition of alternatives is reduced (so probabilities add) or not (so amplitudes add). The point is that distinguishing physical events do not allow a superposition to continue evolving—which was the real work done by the notions of "macroscopic" or "classical" in the conventional textbook treatment of measurement.

In his crystal-scattering example, the distinguishing physical events are the flips in the spin of nuclei in the crystal. But Feynman's warhorse example is always the two-slit experiment. An electron traverses a screen with two slits and eventually registers a hit on a far wall. The electron traverses the two-slit screen in the superposition state of being indefinite between "going through slit 1" and "going through slit 2." If there is no physical event to distinguish between the two alternative trajectories, then an electron will evolve in its indefinite state and show interference effects.

If, however, there is some way to, in principle, physically distinguish between the two alternative trajectories, then the superposition is reduced and repeated trials will show no interference effects. The key to what von Neumann [65] called a "type 1" (measurement-type) process is the physical realization of a distinction. In the two-slit experiment, one way to physically realize a distinction is to cover up one of the slits. A more subtle way to make a physical distinction is to put a detector at one or both slits. Then the registration of a passage at a detector  $D1$  or  $D2$  serve as the "non-superposable" physically distinguishing events.

The physics will differ from case to case as to why, say, "going through slit 1" and "going through slit 2" are superposable but registers at detectors  $D1$  or  $D2$  are not. The important thing here is not the specific way to physically make the distinction, but the concept of *making physical distinctions* that prevents the further evolution of a superposition.

The analysis of the measurement problem can be modeled in QM/Sets. Let  $Q = \{a, b\}$  be the quantum system and let  $M = \{0, 1, 2\}$  be the measuring apparatus so



that  $\wp(Q \times M)$  is the space for the composite system (recall that the tensor product  $\wp(Q) \otimes \wp(M) \cong \wp(Q \times M)$ ). The initial state of the quantum system is the superposition  $\{a, b\}$ . The state  $\{0\}$  is the neutral state of the measuring apparatus, and  $\{1\}$  and  $\{2\}$  are the pointer states to be correlated respectively with  $\{a\}$  and  $\{b\}$ . The key assumption is that the pointer states are distinguishing states that will not allow  $\{a, b\}$  to evolve as a superposition. Intuitively we might think of  $\{a\}$  and  $\{b\}$  as the "going through slit 1" and "going through slit 2" states. The neutral state  $\{0\}$  corresponds to the detectors being off where no measurement is being taken, while the indicator states of  $\{1\}$  and  $\{2\}$  correspond to the detectors being on and the detectors  $D1$  or  $D2$  respectively registering a passage of the particle.

In the initial interaction, the composite system is in the superposition state  $\{(a, 0), (b, 0)\}$ . Then what von Neumann would call a "type 2" (i.e., non-singular) transformation is applied in the composite system  $\mathbb{Z}_2^6$  with the action on the basis as listed in Table 5.2:

$$\begin{aligned} (a, 0) &\rightarrow (a, 1) \\ (b, 0) &\rightarrow (b, 2) \\ (a, 1) &\rightarrow (a, 0) \\ (b, 1) &\rightarrow (b, 1) \\ (a, 2) &\rightarrow (a, 2) \\ (b, 2) &\rightarrow (b, 0) \end{aligned}$$

Table 5.2: Action of non-singular transformation.

This gives a "type 2" transformation of the initial superposition state  $\{(a, 0), (b, 0)\}$  to the superposition state  $\{(a, 1), (b, 2)\}$  which correlates the quantum states with the pointer states. But it is assumed that the pointer states  $\{1\}$  and  $\{2\}$  cannot be superposed. The discrete partition  $\mathbf{1}_M = \{\{0\}, \{1\}, \{2\}\}$  on  $M = \{0, 1, 2\}$  times the indiscrete partition on  $Q = \{a, b\}$  gives the product partition

$$\mathbf{0}_Q \times \mathbf{1}_M = \{\{(a, 0), (b, 0)\}, \{(a, 1), (b, 1)\}, \{(a, 2), (b, 2)\}\}$$

of the composite system which is the least refined partition that still mathematically expresses the assumed *distinguishing nature* of the  $M$ -states in the composite system. That distinguishing nature of the  $M$ -states is then applied by the join-action of  $\mathbf{0}_Q \times \mathbf{1}_M$  on the pure state  $\{(a, 1), (b, 2)\}$  which results in the mixed state  $\{\{(a, 1)\}, \{(b, 2)\}\}$ . That join-action or Lüders mixture operation, which turns the

pure superposition state  $\{(a, 1), (b, 2)\}$  into the mixed state  $\{\{(a, 1)\}, \{(b, 2)\}\}$ , is not a type 2 process (i.e., non-singular in QM/Sets or unitary in QM/C)—since those processes turn pure states into pure states.

In terms of density matrices, the pure state  $S = \{(a, 1), (b, 2)\}$  is represented by the following density matrix (where the row labels are on the right and the columns are labeled likewise):

$$\begin{aligned} \rho(\{(a, 1), (b, 2)\}) &= \begin{bmatrix} 0 \\ 0 \\ \frac{1}{\|S\|_{Q+M}} \\ 0 \\ 0 \\ \frac{1}{\|S\|_{Q+M}} \end{bmatrix} \begin{bmatrix} 0 & 0 & \frac{1}{\|S\|_{Q+M}} & 0 & 0 & \frac{1}{\|S\|_{Q+M}} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{matrix} (a, 0) \\ (b, 0) \\ (a, 1) \\ (b, 1) \\ (a, 2) \\ (b, 2) \end{matrix}. \end{aligned}$$

The partition  $\mathbf{0}_Q \times \mathbf{1}_M$  plays the role of the inverse-image partition of the observable that assigns each pair  $(x, i) \in Q \times M$  its  $M$ -value  $i$  so the blocks define the projection matrices:

$$P_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, P_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \text{ and } P_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then the density matrix for the result of the "join-action" of the measurement is obtained by the Lüders mixture operation:

$$\hat{\rho}(\{(a, 1), (b, 2)\}) = \sum_{i=0}^2 P_i \rho(\{(a, 1), (b, 2)\}) P_i = \rho(\mathbf{0}_Q \times \mathbf{1}_M \vee (\{(a, 1), (b, 2)\}))$$

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

Join-action of Measurement:

$$\{(a, 1), (b, 2)\} \rightsquigarrow \mathbf{0}_Q \times \mathbf{1}_M \vee (\{(a, 1), (b, 2)\}) = \{\{(a, 1)\}, \{(b, 2)\}\}.$$

Thus with half-half probability, the composite system is in the state  $\{(a, 1)\}$ , i.e.,  $Q$  is in state  $\{a\}$  and  $M$  is in state  $\{1\}$ , or is in the state  $\{(b, 2)\}$ , i.e.,  $Q$  is in state  $\{b\}$  and  $M$  is in state  $\{2\}$ .

Insofar as the quantum system  $Q$  is concerned, this is, of course, the same outcomes as described by the "measurement" by an observable  $f : Q \rightarrow \mathbb{R}$  with  $f(a) = 1$  and  $f(b) = 2$  so that  $f^{-1} = \{\{a\}, \{b\}\} = \mathbf{1}_Q$ :

$$\rho(\{a, b\}) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \rightsquigarrow \hat{\rho} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

Measurement: join action of  $\{a, b\} \implies f^{-1} \vee \{\{a, b\}\}.$

The controversial part in the conventional treatment of the measurement problem is that last step of the join-action that expresses the effect of the distinguishing physical events (e.g., the spin-flip of a nucleus in the crystal-scattering example or the hit at a detector in the two-slit experiment) that make distinctions between the superposed alternatives.

## 5.5 Measuring measurement with logical entropy

One of the main results about density matrices in QM/ $\mathbb{C}$  or QM/Sets is:

**Proposition 3**  $\text{tr}[\rho^2] = \sum_{i,j} \|\rho_{ij}\|^2$ . [27, p. 77]

**Proof:** A diagonal entry in  $\rho^2$  is  $(\rho^2)_{ii} = \sum_{j=1}^n \rho_{ij} \rho_{ij}^* = \sum_{j=1}^n \|\rho_{ij}\|^2$  so  $\text{tr}[\rho^2] = \sum_{i=1}^n (\rho^2)_{ii} = \sum_{i,j} \|\rho_{ij}\|^2$ .  $\square$

The logical entropy of a density matrix is thus:  $h(\rho) = 1 - \text{tr}[\rho^2] = 1 - \sum_{i,j} \|\rho_{ij}\|^2$ . For instance, the logical entropy of  $\{\{a, b\}, \{c\}\}$  by our previous definitions is:

$$h(\{\{a, b\}, \{c\}\}) = 1 - \Pr(\{a, b\})^2 - \Pr(\{c\})^2 = 1 - (p_1 + p_2)^2 - p_3^2 = 1 - \sum_{i=1}^3 p_i^2 - 2p_1p_2.$$

By the density matrix definition, it is:

$$h(\rho(\{\{a, b\}, \{c\}\})) = 1 - \text{tr}[\rho(\{\{a, b\}, \{c\}\})^2] = 1 - \sum_{i,j=1}^3 \left\| \rho(\{\{a, b\}, \{c\}\})_{ij} \right\|^2$$

where

$$\rho(\{\{a, b\}, \{c\}\}) = \begin{bmatrix} p_1 & \sqrt{p_1p_2} & 0 \\ \sqrt{p_2p_1} & p_2 & 0 \\ 0 & 0 & p_3 \end{bmatrix}$$

so

$$\sum_{i,j} \left\| \rho(\{\{a, b\}, \{c\}\})_{ij} \right\|^2 = \sum_{i=1}^3 p_i^2 + 2p_1p_2$$

and thus the two definitions agree.

The main theorem relating logical entropy to projective management follows immediately.

**Theorem 1** *For any density matrix  $\rho$  (in  $QM/\mathbb{C}$  or  $QM/\mathbb{Z}_2$ ) and the projective measurement of  $\rho$  by an observable ( $F$  or  $f$ ), the sum of the (absolute) squares of the non-zero entries  $\rho_{ij}$  in  $\rho$  that are zeroed in transition  $\rho \rightsquigarrow \hat{\rho}$  is the difference in their logical entropies  $h(\hat{\rho}) - h(\rho)$ .*

**Proof:**  $h(\hat{\rho}) - h(\rho) = 1 - \text{tr}[\hat{\rho}^2] - (1 - \text{tr}[\rho^2]) = \sum_{i,j} \|\rho_{ij}\|^2 - \sum_{i,j} \|\hat{\rho}_{ij}\|^2. \quad \square$

Measurement makes distinctions and the measure of distinctions is logical entropy. One of the founders of quantum information theory, Charles Bennett, makes the point about distinctions, differentiation, and distinguishability.

So information really is a very useful abstraction. It is the notion of distinguishability abstracted away from what we are distinguishing, or from the carrier of information... [6, p. 155]

The theorem's direct quantitative connection between state discrimination, measurement, and quantum logical entropy reinforces the judgment of Boaz Tamir and Eliahu Cohen ([62] or [63]) that quantum logical entropy is a natural and informative entropy concept for quantum mechanics.

We find this framework of partitions and distinction most suitable (at least conceptually) for describing the problems of quantum state discrimination, quantum cryptography and in general, for discussing quantum channel capacity. In these problems, we are basically interested in a distance measure between such sets of states, and this is exactly the kind of knowledge provided by logical entropy [Reference to [18]]. [62, p. 1]

In the  $F$ -measurement of a pure state  $\psi$ , the non-zero  $i, j$ -entries in  $\hat{\rho}(\psi)$  that were not zeroed are the ones in  $\rho(\psi)$  that survived the pre- and post-projections in the Lüders mixture operation  $\hat{\rho}(\psi) = \sum_{\lambda} P_{\lambda} \rho(\psi) P_{\lambda}$  for some eigenvalue  $\lambda$ . Those are the  $i, j$ -entries where the eigenvectors  $v_i$  and  $v_j$  have the same eigenvalue  $\lambda$ . This is easily seen by looking at the pre- and post-multiplication of:

$$\rho(\psi) = \begin{bmatrix} \|\langle v_1|\psi\rangle\|^2 & \langle v_1|\psi\rangle \langle v_2|\psi\rangle^* & \cdots & \langle v_1|\psi\rangle \langle v_n|\psi\rangle^* \\ \langle v_2|\psi\rangle \langle v_1|\psi\rangle^* & \|\langle v_2|\psi\rangle\|^2 & \cdots & \langle v_2|\psi\rangle \langle v_n|\psi\rangle^* \\ \vdots & \cdots & \ddots & \vdots \\ \langle v_n|\psi\rangle \langle v_1|\psi\rangle^* & \langle v_n|\psi\rangle \langle v_2|\psi\rangle^* & \cdots & \|\langle v_n|\psi\rangle\|^2 \end{bmatrix}$$

by the projection matrices  $P_{\lambda} = \sum_{v_i: Fv_i = \lambda v_i} |v_i\rangle \langle v_i|$  which are diagonal matrices with  $(P_{\lambda})_{ii} = 1$  if  $Fv_i = \lambda v_i$  and 0 otherwise. In the pre-multiplication:

$$\begin{aligned} (\rho(\psi) P_{\lambda})_{ij} &= \sum_{k=1}^n \langle v_i|\psi\rangle \langle v_k|\psi\rangle^* (P_{\lambda})_{kj} \\ &= \langle v_i|\psi\rangle \sum_{k=1}^n \langle v_k|\psi\rangle^* (P_{\lambda})_{kj} = \begin{cases} \langle v_i|\psi\rangle \langle v_j|\psi\rangle^* = \rho(\psi)_{ij} & \text{if } v_j \text{ has eigenvalue } \lambda \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

and for the post-multiplication:

$$\begin{aligned} (P_{\lambda} \rho(\psi))_{ij} &= \sum_{k=1}^n (P_{\lambda})_{ik} \langle v_k|\psi\rangle \langle v_j|\psi\rangle^* \\ &= \langle v_j|\psi\rangle^* \sum_{k=1}^n (P_{\lambda})_{ik} \langle v_k|\psi\rangle = \begin{cases} \langle v_i|\psi\rangle \langle v_j|\psi\rangle^* = \rho(\psi)_{ij} & \text{if } v_i \text{ has eigenvalue } \lambda \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Thus all the zero  $\hat{\rho}(\psi)_{ij}$  are for eigenvectors  $v_i$  and  $v_j$  with different eigenvalues. For a given eigenvalue  $\lambda$ , the probability of the  $F$ -measurement returning an eigenvalue  $\lambda$  is  $p_{\lambda} = \Pr(\lambda|\psi) = \text{tr}[P_{\lambda} \rho(\psi)] = \sum_{Fv_i = \lambda v_i} \|\langle v_i|\psi\rangle\|^2 = \sum_{Fv_i = \lambda v_i} p_i$  where  $\|\langle v_i|\psi\rangle\|^2$  is abbreviated as  $p_i$ . Hence

$$p_{\lambda}^2 = \left( \sum_{Fv_i = \lambda v_i} \|\langle v_i|\psi\rangle\|^2 \right)^2 = \sum_{Fv_i = \lambda v_i} p_i^2 + 2 \sum_{Fv_i = \lambda v_i, Fv_j = \lambda v_j} p_i p_j$$

which is the sum of the absolute squares of the non-zero entries  $\hat{\rho}(\psi)_{ij}$  where  $v_i$  and  $v_j$  have the same eigenvalue  $\lambda$ . Indeed, the squares of the diagonal entries  $\hat{\rho}(\psi)_{ii}$  where  $Fv_i = \lambda v_i$  is the term  $\sum_{Fv_i=\lambda v_i} p_i^2$  and the sum of the absolute squares of the non-zero off-diagonal terms is  $2 \sum_{Fv_i=\lambda v_i, Fv_j=\lambda v_j} p_i p_j$ . Thus the sum  $\sum_{\lambda} p_{\lambda}^2 = \sum_{i,j} \left\| \hat{\rho}(\psi)_{ij} \right\|^2 = \text{tr} [\hat{\rho}(\psi)^2]$  recalling that the  $\hat{\rho}(\psi)_{ij}$  terms where  $v_i$  and  $v_j$  have different eigenvalues are zero. Hence we have a simple interpretation for the quantum logical entropy of the result  $\hat{\rho}(\psi)$  of measuring  $\rho(\psi)$  by an observable  $F$ :

$$h(\hat{\rho}(\psi)) = 1 - \text{tr} [\hat{\rho}(\psi)^2] = \text{the probability of getting different eigenvalues in two independent } F\text{-measurements of } \psi.$$

This is, of course, the QM/ $\mathbb{C}$  version of the previous QM/Sets result that the logical entropy  $h(f^{-1})$  of inverse-image partition  $f^{-1} = \{f^{-1}(r)\}_{r \in f(U)}$  of a random variable  $f : U \rightarrow \mathbb{R}$  is the probability that two independent trials for the r.v. will yield distinct values. The von Neumann (vN) entropy, in contrast, has no such simple interpretation, and there seems to be no such intuitive connection between pre- and post-measurement density matrices and vN entropy, although vN entropy also increases in a projective measurement [49, Theorem 11.9, p. 515].

## 6 Indistinguishability of particles in quantum mechanics

We may begin by considering the combinatorics of the classical Maxwell-Boltzmann (MB) distribution of classical particles in a number of states.

Let  $|U| = n$  be the number of states a particle can be in and let  $k$  be the number of particles. In the balls-and-boxes language of combinatorial theory,  $|U| = n$  is the number of boxes and  $k$  the number of balls. How many different ways can the balls be put into the boxes. The state space is  $\wp(U^k) = \wp\left(U \times \dots \times U\right)$   $k$  times. A distribution of the balls to the boxes means that each of the  $k$  particle-balls is assigned to one of the  $n$  state-boxes  $\{u\}$  for  $u \in U$  so there are  $n^k$  ways to distribute the balls (particles) to the boxes (states). Each distribution gives a basis element in the space  $\wp\left(U \times \dots \times U\right)$   $k$  times. Classically, the balls (and boxes) are distinguishable but the order of the balls within each box does not matter. Each of the  $n$  boxes (states) has an occupation number  $\theta_i$  where  $0 \leq \theta_i \leq k$  and  $\sum_{i=1}^n \theta_i = k$ .

The multinomial coefficient gives the number of different ways to distribute  $k$  distinguishable particles among the  $n$  boxes that have the same occupation numbers  $\theta_1, \dots, \theta_n$ . One way to arrive at the formula is to list all the  $k!$  different ways to linearly order (i.e., permute)  $k$  distinct elements, one ordering per row. Then draw vertical lines representing the boxes with the given occupation number  $\theta_i$  in the  $i^{\text{th}}$  box as illustrated in Figure 6.1.

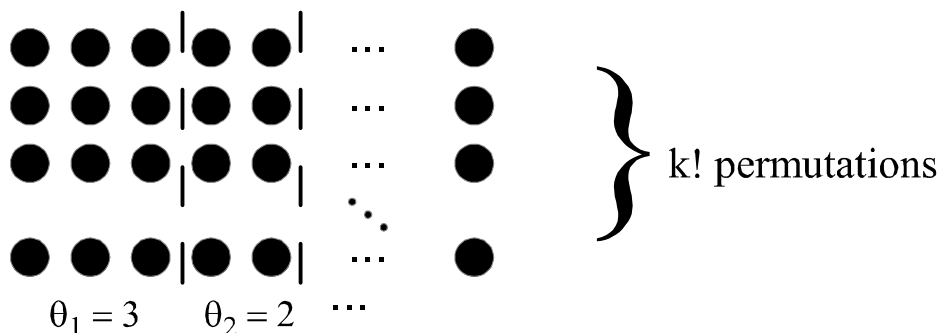


Figure 6.1: Arbitrary placement of vertical lines representing occupation numbers.

Thus there are  $k!$  ways to distribute the balls between the boxes with the given occupation numbers  $\theta_1, \dots, \theta_n$  paying attention to the ordering within each box. This seems counterintuitive at first since  $k!$  does not take into account the occupation numbers  $\theta_1, \dots, \theta_n$  but looking at the figure makes it clear that it doesn't matter what the specific occupation numbers are as long as  $\sum_{i=1}^n \theta_i = k$ . But we do not distinguish between different orderings within each box so we need to divide by  $\theta_i!$  for  $i = 1, \dots, n$  to obtain the number of different ways to distribute  $k$  balls between  $n$  boxes with the given occupation numbers  $\theta_1, \dots, \theta_n$  disregarding the ordering within each box. That yields the:

$$\binom{k}{\theta_1, \dots, \theta_n} = \frac{k!}{\theta_1! \dots \theta_n!}$$

Multinomial coefficient.

If we assume that all the  $n^k$  ways of distributing the balls (particles) between the boxes (states) are equiprobable, then the probability of having a distribution with the given occupation numbers is:

$$\text{Pr}_{MB}(\theta_1, \dots, \theta_n) = \frac{\binom{k}{\theta_1, \dots, \theta_n}}{n^k} = \frac{k!}{\theta_1! \dots \theta_n! n^k}$$

Maxwell-Boltzmann distribution.

In terms of the vector space  $\wp(U^k) = \wp\left(U \times \dots \times U\right)$ , the dimension is  $\left|U \times \dots \times U\right| = n^k$  and each basis vector  $\{(u_1, \dots, u_k)\} \in U^k$  is equiprobable. The set of basis vectors giving the occupation numbers  $\theta_1, \dots, \theta_n$  has the cardinality  $\binom{k}{\theta_1, \dots, \theta_n} = \frac{k!}{\theta_1! \dots \theta_n!}$  so the calculation of the probability in the MB distribution is a classical  $\frac{\# \text{ favorable cases}}{\# \text{ total cases}}$ .

Classically, there is no formation of new states by superposition so the subsets in  $\wp(U^k)$  are just classical collections of cases, not new states. In terms of density matrices, the classical collection of decohered or decomposed states is represented by a diagonal  $n^k \times n^k$  density matrix. In the quantum case of QM/sets, the same mathematical subsets in  $\wp(U^k)$  may be interpreted not just as a collections of basis (eigen-)states but as a new type of state, the superposition of the states in the collection. They are represented by density matrices with non-zero off-diagonal elements representing the cohering together of the corresponding diagonal elements in a superposition. The consideration of those new quantum superposition states brings us to the Bose-Einstein (BE) distribution.

In the quantum context, the balls or particles are indistinguishable or "identical." Mathematically, this means that the physical state must be identical under any permutation of the balls. One can still consider occupation numbers for the boxes, but, unlike the MB case, "different" sets of  $\theta_i$  balls in the  $i^{\text{th}}$  box are not considered different. All that counts is the number of balls in each box so we need to compute the number of sets of occupation numbers:

$$\theta_1, \theta_2, \dots, \theta_n \text{ where } \sum_{i=1}^n \theta_i = k.$$

To do that, we first compute the number of ways to distribute  $k$  balls among  $n$  boxes paying attention to the order of the balls in each box.

- First ball:  $n$  choices =  $n$  boxes;
- Second ball:  $n - 1$  empty boxes plus before or after the first ball in whatever box it was in =  $n + 1$  choices;
- Each placed ball adds another choice: before or after it as illustrated in Figure 6.2.

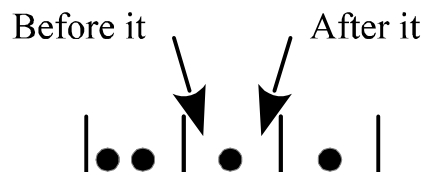




Figure 6.2: Each placed ball creates one additional place for the next ball.

The number of ways to distribute the balls paying attention to the order within each box is thus the:

$$\langle n \rangle_k = n(n+1) \dots (n+(k-1)) \text{ (} k \text{ times)}$$

Rising factorial.

The linear ordering of the balls in each box treats them as distinguishable, but indistinguishability requires that all the permutations keeping the number of balls in each box the same give the same set of occupation numbers so we divide by  $k!$ , the number of ways to permute the balls to get:

$$\langle n \rangle_k = \frac{\langle n \rangle_k}{k!} = \frac{n(n+1) \dots (n+k-1)}{k!} \text{ [alt. notation: } \binom{n}{k}]$$

Number of sets of occupation numbers  $\theta_1, \theta_2, \dots, \theta_n$  where  $\sum_{i=1}^n \theta_i = k$ .

The next step is to associate an invariant vector in  $\wp(U^k)$  with each set of occupation numbers so those vectors will span "the bosonic subspace" of  $\wp(U^k)$ .

These balls-in-boxes *distribution* problems can also be seen as *occupancy* problems. A distribution of balls to boxes can be defined as a function  $f : \{\text{balls}\} \rightarrow \{\text{boxes}\}$  so  $f(i) = c$  means the  $i^{\text{th}}$  ball is distributed to the box (or state)  $c$ . For the occupancy interpretation, we describe the same information by considering a "word" of  $k$  letters, an ordered  $k$ -tuple, with  $f(i)$  in the  $i^{\text{th}}$  position:

$$(f(1), f(2), \dots, f(k)) \text{ (} k\text{-tuple)}$$

Occupancy form of the distribution  $f : \{\text{balls}\} \rightarrow \{\text{boxes}\}$ .

For instance, in the case of  $U = \{a, b, c\}$  and  $k = 2$ , the set of occupation numbers  $(1, 1, 0)$  describing a distribution between the  $a$ ,  $b$ , and  $c$  boxes respectively is restated in occupancy form as:  $(a, b) \in U \times U$  or, in terms of subset-vectors,  $\{(a, b)\} \in \wp(U \times U)$ .

There are two ways to construct an invariant-under-permutations vector from the *occupancy*  $k$ -tuple  $(f(1), f(2), \dots, f(k))$ : allowing or not allowing repetitions.

- Without repetitions: All that is retained in the subset/vectors in  $\wp(U^k)$  from the permutations of the singleton  $\{(f(1), f(2), \dots, f(k))\}$  are the distinct ones ignoring repetitions, and the balls/particles with those basis-vectors/eigenstates are called *bosons*;

- With repetitions: All permutations of  $(f(1), f(2), \dots, f(k))$  are put into a subset/vector in  $\wp(U^k)$  where repetitions will give cancellations (since set addition is mod(2)), and balls/particles with those (non-zero) basis-vectors/eigenstates are called *fermions*.

The bosonic basis-vectors are the *minimal* (i.e., ignoring repetitions) invariant-under-permutations version of the occupancy  $k$ -tuples, e.g.,  $\{(a, b)\}$  becomes  $\{(a, b), (b, a)\}$ . For occupation numbers with repetitions such as  $(0, 2, 0)$ , the occupancy form  $\{(b, b)\}$  is already invariant under any permutation, i.e., under any transposition (since all permutations are the product of transpositions). The minimal invariant vectors constructed from the occupancy forms are in one-to-one correspondence with the sets of occupation numbers, and those vectors are the basis for the *bosonic subspace*  $V_{BE} \subseteq \wp(U^k)$  with dimension  $\dim(V_{BE}) = \langle n \rangle_k = \frac{n(n+1)\dots(n+k-1)}{k!}$ .

The transition from occupation numbers to bosonic basis vectors is illustrated in Table 6.1 with representative cases for 3 ball-bosons and 4 boxes or states with  $U = \{a, b, c, d\}$  where  $(a, b, c)$  is shortened to  $(abc)$ .

Occup. #	$\sum_{\sigma \in \mathcal{S}(k)} \{\sigma(f(1), f(2), \dots, f(k))\}$ no repetitions
(1, 1, 1, 0)	$\{(abc), (acb), (bac), (bca), (cab), (cba)\}$
(1, 1, 0, 1)	$\{(abd), (adb), (bad), (bda), (dab), (dba)\}$
(2, 1, 0, 0)	$\{(aab), (aba), (baa)\}$
(0, 2, 1, 0)	$\{(bbc), (bcb), (cbb)\}$
(3, 0, 0, 0)	$\{(aaa)\}$
(0, 0, 3, 0)	$\{(ccc)\}$

Table 6.1: Generation of permutations of occupancy  $k$ -tuple without repetitions (not all possibilities listed).

Taking each basis vector for  $V_{BE}$  as being equiprobable,

$$\Pr_{BE}(\theta_1, \dots, \theta_n) = \frac{1}{\langle n \rangle_k} = \frac{k!}{n(n+1)\dots(n+k-1)}$$

Probability  $k$  bosons in state  $(\theta_1, \dots, \theta_n)$ .

Number of  $n^k$  basis vectors of  $\wp(U^k)$  in each BE basis vector is given by the multinomial coefficient. For instance, the number of triples included in the basis vector corresponding to  $(1, 1, 1, 0)$ , i.e.,  $\{(abc), (acb), (bac), (bca), (cab), (cba)\}$ , is:

$$\binom{k}{\theta_1, \dots, \theta_n} = \frac{k!}{\theta_1! \dots \theta_n!} = \frac{3!}{1!1!1!0!} = 6.$$

In MB statistics, the probability of getting that set  $(1, 1, 1, 0)$  of occupation numbers is  $\binom{k}{\theta_1, \dots, \theta_n} / n^k = 6/4^3 = 6/64 = 3/32$  which is the same as the probability of the event  $\{(abc), (acb), (bac), (bca), (cab), (cba)\}$  occurring.

But in BE statistics, that "event" is one basis vector among  $\binom{n}{k} = \frac{n(n+1)\dots(n+k-1)}{k!} = \frac{4(5)(6)}{3!} = 20$  basis vectors so its probability is  $\frac{1}{20} = \frac{8}{160} < \frac{15}{160} = \frac{3}{32}$ , which shows that bosons "don't like to be separated into different states" compared to classical particles.

In Hermann Weyl's treatment [68], he assumed two identical (indistinguishable) particles which he artificially label "Mike" and "Ike." Each particle can be in one of three eigenstates  $A, B,$  and  $C$  so the single particle state space  $\wp(U)$  is generated by  $U = \{A, B, C\}$  and thus the space of two-particle states is generated by the nine basis elements in  $\wp(U \times U) = \wp(\{A, B, C\} \times \{A, B, C\})$ .

We define an equivalence relation on the nine basis elements where each ordered pair is equivalent to the one where Mike and Ike are permuted. There are the following six equivalence classes on the basis elements. If we take the basis set  $U \times U$  as the sample space with each pair as being equiprobable, then the probability of the equivalence classes is the *Maxwell-Boltzmann distribution* illustrated in Table 6.2.

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

Table 6.2: Maxwell-Boltzmann distribution.

In quantum mechanics, it is often said that the result obtained after permuting the particles is "indistinguishable" from the prior state, but that is also true for classical particles. It would be more accurate to say that the result in the quantum case is *identical* when the artificially labeled particles are permuted. In other words, what is classically seen as an equivalence class of numerically-distinct states obtained by permutations should really be seen in the quantum case as one "indefinite" superposition state that is identical under permutation. In that case, the sample space has

six indefinite entities which we may take as equiprobable to obtain the *Bose-Einstein distribution* as illustrated in Table 6.3 and they span a six-dimensional subspace  $V_{BE}$  of the nine-dimensional  $\varphi(U \times U)$ .

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

Table 6.3: Bose-Einstein distribution.

It may be useful to develop an imagery to distinguish between an equivalence class that is the same, i.e., invariant, under a certain group operations like permutations and an indefinite element that is identical under such permutations. Consider an isosceles triangle with labeled angles and sides, and consider the reflection that interchanges the two equal sides and angles as shown in Figure 6.3.

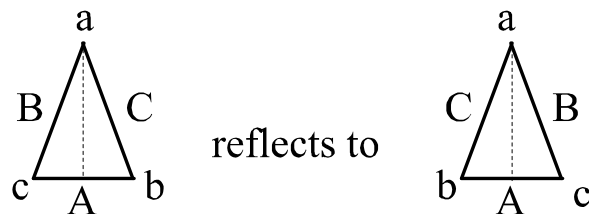


Figure 6.3: Reflection on  $A - a$  axis of isosceles triangle.

Then under that reflection operation, there is the equivalence class as pictured in Figure 6.4 that is closed or invariant under that reflection on  $a - A$ -axis operation.

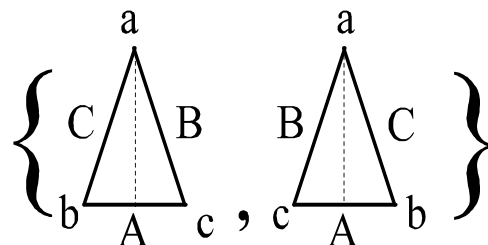


Figure 6.4: Equivalence class of definite triangles symmetric under reflection.

That equivalence class should be thought of as classical imagery—like what is represented by the diagonal density matrices. The quantum version is where the differences or distinctions between the two triangles are ‘cohered’ or rendered indefinite as indicated by non-zero off-diagonal elements in a density matrix—which might then be rendered definite in one way or the other by an appropriate measurement. With the differences still indefinite, then the reflection operation is the identity operation which might be imaged in Figure 6.5.

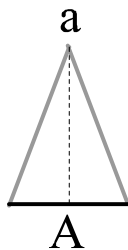


Figure 6.5: Indefinite triangle identical under reflection.

In the classical MB case, the pair of permutation-symmetric states  $\{(A, B), (B, A)\}$  should be thought of as like the equivalence class of the reflection-symmetric triangles. In the quantum BE case, the state  $\{(A, B), (B, A)\}$  should be thought of as the single indefinite superposition state where permutation is identity. Instead of being the set of two states where one state is with Mike in state  $A$  with Ike in state  $B$ , or the reverse state, it is the one superposition state where one particle is in state  $A$  and the other particle is in state  $B$ .

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

A set of occupation numbers  $(\theta_1, \dots, \theta_n)$  where  $\sum_{i=1}^n \theta_i = k$ , is given by any distribution  $f : \{\text{balls}\} \rightarrow \{\text{boxes}\}$  such that  $|f^{-1}(i^{\text{th}} \text{ box})| = \theta_i$  for  $i = 1, \dots, n$ . The occupancy version of the distribution is  $(f(1), f(2), \dots, f(k))$  and the minimal (ignoring repetitions) invariant-under-transpositions (of balls/particles) gives the corresponding BE basis vector. The *Fermi-Dirac way* to construct a vector in  $\wp(U^k)$

from the  $k$ -tuple  $(f(1), f(2), \dots, f(k)) \in U^k$  is to take the set of *all*  $k!$  permutations allowing repetitions: (where  $S(k)$  is the symmetric group of all permutations of a set of  $k$  elements)

$$\sum_{\sigma \in S(k)} \{\sigma(f(1), f(2), \dots, f(k))\}.$$

Recall that set-addition is addition of subsets mod(2) where  $-1 = +1$ . For any set of occupation numbers where some  $1 < \theta_i$ , then there is a repeating box/state, say  $f(j) = f(j')$ , so the occupancy  $k$ -tuple has the form:  $(f(1), \dots, f(j), \dots, f(j'), \dots, f(k))$ . Hence in the sum over all permutations, we have for the transposition of the  $j$  and  $j'$  term:

$$\{(f(1), \dots, f(j), \dots, f(j'), \dots, f(k))\} + \{(f(1), \dots, f(j'), \dots, f(j), \dots, f(k))\} = 0.$$

For any set of occupation numbers with some  $\theta_i \geq 2$ , the corresponding Fermi-Dirac vector contains  $\theta_i!$  copies of the same  $k$ -tuple which will cancel since  $\theta_i!$  is an even number. In contrast, the Bose-Einstein way to construct the basis vectors was to keep one representative of each  $k$ -tuple that arises (no repetition) instead of letting pairs cancel out.

The cases where each  $\theta_i$  is 0 or 1 give nonzero vectors  $\sum_{\sigma \in S(k)} \{\sigma(f(1), f(2), \dots, f(k))\} \in \wp(U^k)$  which are the *FD basis vectors* for the *fermionic subspace*  $V_{FD} \subseteq \wp(U^k)$ . Thus each FD basis vector corresponds to a  $k$ -element subset of the  $n$  boxes/states since the sum of the  $\theta_i = 1$  has to be  $k$ :

$$\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k!}$$

Number of FD basis vectors = dimension of  $V_{FD}$ .

For the previous example where  $k = 3$  and  $n = 4$ , the Fermi-Dirac construction of the corresponding vector  $\sum_{\sigma \in S(k)} \{\sigma(f(1), f(2), \dots, f(k))\}$  is as follows in Table 6.4 for some representative cases:

Occup. #	$\sum_{\sigma \in S(k)} \{\sigma(f(1), f(2), \dots, f(k))\}$ with repetitions
(1, 1, 1, 0)	$\{(abc), (acb), (bac), (bca), (cab), (cba)\}$
(1, 1, 0, 1)	$\{(abd), (adb), (bad), (bda), (dab), (dba)\}$
(2, 1, 0, 0)	$\{(aab), (aba), (aab), (aba), (baa), (baa)\} = 0$
(0, 2, 1, 0)	$\{(bbc), (bcb), (bbc), (bcb), (cbb), (cbb)\} = 0$
(3, 0, 0, 0)	$\{(aaa), (aaa), (aaa), (aaa), (aaa), (aaa)\} = 0$
(0, 0, 3, 0)	$\{(ccc), (ccc), (ccc), (ccc), (ccc), (ccc)\} = 0$

Table 6.4: Generation of permutations of occupancy  $k$ -tuple with repetitions.

Taking each basis vector for  $V_{FD}$  as being equiprobable,

$$\Pr_{FD}(\theta_1, \dots, \theta_n) = \frac{1}{\binom{n}{k}} = \frac{k!}{n(n-1)\dots(n-k+1)}$$

Probability  $k$  fermions in state  $(\theta_1, \dots, \theta_n)$  where each  $\theta_i = 1$  or  $0$ .

Note that while we have developed the Maxwell-Boltzmann, Bose-Einstein, and Fermi-Dirac statistics in the context of the QM/sets model and the different basis sets for the respective subspaces of  $\wp(U^k)$ , the results are the correct statistics (not some toy version)—just as the probability calculus of the QM/sets model is the classical finite probability calculus with the "non-commutative" difference that each basis set  $\{\{u\}\}_{u \in U}$  for  $\mathbb{Z}_2^n$  gives a calculus that computes probabilities in  $\wp(U) \cong \mathbb{Z}_2^n$ .

In the simple model of two particle-balls and three boxes or states, by allowing repetitions, we get another  $(A, A)$  in the superposition state  $\{(A, A), (A, A)\} = 0$ . Thus allowing repetitions gives only  $\binom{3}{2} = 3$  non-zero indefinite states which form the sample space of equiprobable points for the Fermi-Dirac distribution and which span a three-dimensional subspace  $V_{FD}$  of the nine-dimensional  $\wp(U \times U)$  as illustrated in Table 6.5.

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

Table 6.5: Fermi-Dirac distribution.

Since these computations in QM/Sets are essentially combinatorial, the boson-fermion distinction illustrates on "the combinatorial level, the duality between... balls into boxes (subject to certain conditions) not allowing repetitions or allowing repetitions." [58, p. 295] When repetitions are allowed (so they cancel out), then the Pauli exclusion principle is a consequence of addition mod 2 whereas in QM/C, it is a consequence of treating transpositions as being antisymmetric, i.e., transpositions that change the sign.

All the (non-zero) FD basis vectors are also BE basis vectors so  $V_{FD} \subseteq V_{BE}$  and the two corresponding dimensions are (for  $1 < k$ ):

$$\dim(V_{FD}) = \binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k!} < \frac{n(n+1)\dots(n+k-1)}{k!} = \langle n \rangle_k = \dim(V_{BE}).$$

Putting each fermion-ball into a box/state takes away that box as a choice for the next fermion to be distributed, while putting a boson-ball into a box/state adds a choice (before or after that boson) for the next boson to be distributed, and then, in both cases, the division by  $k!$  (to cancel out the ordering of choices) computes the respective number of basis vectors. The difference between the two dimensions is given by the difference between the *falling factorial*  $n(n-1)\dots(n-k+1)$  ( $k$  terms) and the *rising factorial*  $n(n+1)\dots(n+k-1)$  ( $k$  terms). This analogy between the two enumerations is obscured in the standard combinatorics texts that represent the "number of  $k$ -combinations with repetition of  $n$  distinct things" as the binomial coefficient:

$$\binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!} = \frac{(n-1)!n(n+1)\dots(n+k-1)}{k!(n-1)!} = \frac{n(n+1)\dots(n+k-1)}{k!} = \langle n \rangle_k = \left(\binom{n}{k}\right).$$

From the QM/sets or combinatorial viewpoint, the difference between bosons and fermions arises from the difference between counting with repetition and without repetition (closely related to sampling with or without replacement). Moreover, this shows up in "the reciprocity between  $\binom{n}{k}$  and  $\left(\binom{n}{k}\right)$ " [58, p. 295] which are the respective dimensions of the Fermi-Dirac subspace  $V_{FD}$  and Bose-Einstein subspace  $V_{BE}$  of  $\wp(U^k)$ . Recall  $\left(\binom{n}{k}\right)$  is alternative notation for  $\frac{n(n+1)\dots(n+k-1)}{k!} = \langle n \rangle_k = \dim(V_{BE})$ .

**Example:** The simplest intuitive example is that  $k = 2$  particles and  $n = 2$  states  $U = \{h, t\}$  like two coins with heads and tails as the states. What is the probability that one "coin" will be "heads" and the other "tails"?

- Classical coins:  $\Pr_{MB}(\{(h, t), (t, h)\}) = \frac{|\{(h, t), (t, h)\}|}{\dim(U \times U)} = \frac{\binom{k}{\theta_1, \dots, \theta_n}}{n^k} = \frac{\theta_1! \dots \theta_n!}{n^k} = \frac{2}{4} = \frac{1}{2}$ .
- Boson coins:  $\Pr_{BE}(\{(h, t), (t, h)\}) = \frac{|\{(h, t), (t, h)\}|}{\dim V_{BE}} = \frac{k!}{n(n+1)\dots(n+k-1)} = \frac{2}{2(3)} = \frac{1}{3}$ .
- Fermion coins:  $\Pr_{FD}(\{(h, t), (t, h)\}) = \frac{|\{(h, t), (t, h)\}|}{\dim V_{FD}} = \frac{k!}{n(n-1)\dots(n-k+1)} = \frac{2}{2(1)} = 1$ .

The fermion coins *have* to be in different states (i.e., with probability 1) which illustrates the Pauli exclusion principle in QM/sets. The three bosonic basis vectors are  $\{(h, t), (t, h)\}, \{(h, h)\}$ , and  $\{(t, t)\} \in \wp(U \times U)$  so the probability that two bosonic coins are in the same state is  $\frac{2}{3}$  in comparison with the classical  $\frac{1}{2}$  which illustrates in QM/sets the "social" tendency of bosons to "want" to be in the same state.



What is the key difference between the classical and quantum cases? Many texts loosely say we cannot distinguish between, say, two electrons in an atom but the same could be said of two electrons in a classical model of an atom. That is, one can easily consider two classical objects so that we could not distinguish the two cases where they were permuted. Yet the original and permuted case still count separately in the Maxwell-Boltzmann distribution. The real difference is in the quantum notion of *superposition* so that  $\{(h, t)\} + \{(t, h)\} = \{(h, t), (t, h)\}$  is one case (recall the density matrix with off-diagonal elements or the isosceles triangle with the equal sides rendered indefinite), not just two classical cases  $\{(h, t)\}$  and  $\{(t, h)\}$  (recall the diagonal density matrix or the equivalence class of two isosceles triangles symmetric under reflection) which are indistinguishable (when we can't classically distinguish between the first and second coin). The creation of one case by superposition is the point, not indistinguishability *per se*.

## 7 Group representation theory with sets and vector spaces

### 7.1 Group representations define partitions

We have seen how the non-zero off-diagonal terms in a density matrix are ‘coherences’ indicating indistinctness. Group representations provide another mathematical way to indicate indistinctness as ‘symmetry.’ Given a *set*  $G$  indexing mappings  $\{R_g : U \rightarrow U\}_{g \in G}$  on a set  $U$ , what are the conditions on the set of mappings so that it is a set representation of a group? Define the binary relation on  $U \times U$ :

$$u \sim u' \text{ if } \exists g \in G \text{ such that } R_g(u) = u'.$$

Then the conditions that make  $R$  into a group representation are the conditions that imply  $u \sim u'$  is an equivalence relation:

1. existence of the identity  $1_U \in U$  implies reflexivity of  $\sim$ ;
2. existence of inverses implies symmetry of  $\sim$ ; and
3. closure under products, i.e., for  $g, g' \in G$ ,  $\exists g'' \in G$  such that  $R_{g''} = R_{g'}R_g$ , implies transitivity of  $\sim$ .

Hence a set representation of a group might be seen as a "dynamic" way to define an equivalence relation and thus a partition on the set.[9] A symmetry group defines indistinctions. For instance, if linear translations form a symmetry group for a quantum system, then the system behavior before a linear translation is indistinct from the behavior of the translated system. Given this intimate connection between groups and partitions, it is then no surprise that group representation theory has a basic role to play in quantum mechanics and in the partition-based objective indefiniteness or literal interpretation of QM.

## 7.2 Where do the fully distinct eigen-alternatives come from?

In classical mechanics, the role of symmetry groups is to establish invariances, e.g., Noether's Theorem. But in quantum mechanics, the spaces satisfy the superposition principle (i.e., objectively indefinite superpositions) and that allows symmetry groups and group representation theory to play a much more fundamental role than simply the role of accounting for invariants in classical mechanics. What is that more fundamental role that goes beyond symmetry-induced invariance? The more fundamental role is to determine—within the constraints of the symmetries—what are all the maximally distinct eigen-alternatives.

In a quantum state space, we are *given* the observable with its distinct eigenstates so the indefinite states are linear combinations of those eigenstates. But how is the observable with the range of distinct eigen-alternatives determined?

In the set case, we are *given* the universe  $U$  of distinct eigen-alternatives  $u \in U$ , and then the indistinct entities are the subsets such as the blocks  $B \in \pi$  in a partition of  $U$ . A "measurement" is some distinction-making operation that reduces an indistinct state  $B$  down to a more distinct state  $B' \subseteq B$  or, in the nondegenerate case, to a fully distinct singleton  $\{u\}$  for some  $u \in B$ . But where do the fully distinct elements come from?

The *basic idea* is that a symmetry group defines indistinctions, so what are all the ways that there can be distinct eigen-elements that are consistent with those indistinctions? In a representation of a group by permutations on a set  $U$ , the answer is:

distinct eigen-elements consistent with symmetry group  $\approx$  orbits of group  
representation.

Two elements of  $U$  inside the same orbit cannot be considered distinct in a way consistent with the indistinction-making action of the group since they are, by definition, mapped from one to the other by an "indistinction-making" symmetry group operation. Hence the maximally distinct subsets—consistent with the indistinction-making symmetries—are the *minimal invariant subsets*, the orbits in the set case. That is how the partition ideas mesh with group representation theory. First we consider the set version, and then we lift to the vector space version of group representation theory.

Let  $U$  be a set and  $S(U)$  the group of all permutations of  $U$ . Then a *set representation* of a group  $G$  is an assignment  $R : G \rightarrow S(U)$  where for  $g \in G$ ,  $g \mapsto R_g \in S(U)$  such that  $R_1$  is the identity on  $U$  and for any  $g, g' \in G$ ,  $R_{g'R_g} = R_{g'g}$ . Equivalently, a *group action* is a binary operation  $G \times U \rightarrow U$  such that  $1u = u$  and  $g'(gu) = (g'g)u$  for all  $u \in U$ .

Defining  $u \sim u'$  if  $\exists g \in G$  such that  $R_g(u) = u'$  [or  $gu = u'$  using the group action notation], then, as previously noted, the properties of a group imply that  $\sim$  is an equivalence relation on  $U$  where the blocks are called the *orbits* and the orbit-partition of  $U$  is *Orbits*.

How are the ultimate distinct eigen-alternatives, the distinct "eigen-forms" of "substance," defined in the set case? Instead of just assuming  $U$  as the set of eigen-alternatives, we start with  $U$  as the carrier for a set representation of the group  $G$  as a group of symmetries. What are the smallest subsets (forming the blocks  $B$  in a set partition) that respect the symmetries, i.e., that are *invariant* in the sense that  $R_g(B) \subseteq B$  for all  $g \in G$ ? Those minimal invariant subsets are the orbits, and all invariant subsets are unions of orbits. Thus the orbits, thought of as points in the quotient set  $U/G \cong \text{Orbits}$  (set of orbits), are the eigen-alternatives, the "eigen-forms" of "substance," defined by the symmetry group  $G$  in the set case.

**Example 1:** Let  $U = \{0, 1, 2, 3, 4, 5\}$  and let  $G = S_2 = \{1, \sigma\}$  (symmetric group on two elements) where  $R_1 = 1_U$  and  $R_\sigma(u) = u + 3 \pmod 6$  as shown in Figure 7.1.

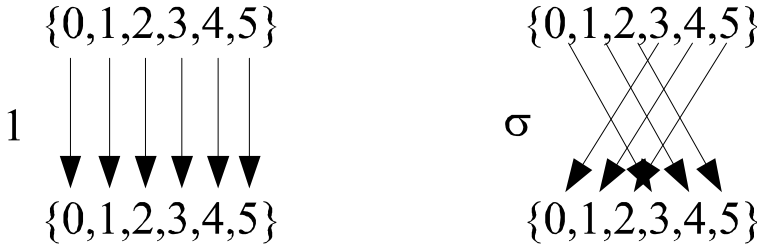


Figure 7.1: Action of  $S_2$  on  $U = \{0, 1, 2, 3, 4, 5\}$ .

There are 3 orbits:  $\{0, 3\}$ ,  $\{1, 4\}$ , and  $\{2, 5\}$ , and they partition  $U$ . Those three orbits are the points in the quotient set  $U/G$ , i.e., they are the distinct eigen-alternatives defined by the symmetry group's  $S_2$  action on  $U$ .

The set concepts "lift" to vector space concepts. A *vector space representation* of a group  $G$  on a vector space  $V$  is a mapping  $g \mapsto R_g : V \rightarrow V$  from  $G$  to invertible linear transformations on  $V$  such that  $R_{g'}R_g = R_{g'g}$ .

The lifts to the vector space representations of groups are;

- minimal invariant subsets = orbits  $\xrightarrow{\text{Lifts}}$  minimal invariant subspaces = *irreducible subspaces*,
- representation restricted to orbits  $\xrightarrow{\text{Lifts}}$  representation restricted to irreducible subspaces which gives the *irreducible representations* (the eigen-forms of "substance" in the vector space case), and
- set partition of orbits  $\xrightarrow{\text{Lifts}}$  vector space partition, i.e., direct-sum decomposition, of irreducible subspaces.

In Heisenberg's philosophical terms, the irreducible representations of certain symmetry groups of particle physics determine the fundamental eigen-forms that the substance (energy) can take.

The elementary particles are therefore the fundamental forms that the substance energy must take in order to become matter, and these basic forms must in some way be determined by a fundamental law expressible in mathematical terms. ... The real conceptual core of the fundamental law must, however, be formed by the mathematical properties of the symmetry it represents.[35, pp. 16-7]

The "irreducible representations" in the set case are just the restrictions of the representation to the orbits, e.g.,  $R \upharpoonright \{0, 3\} : S_2 \rightarrow S(\{0, 3\})$ , as their carriers. A set representation is said to be *transitive*, if for any  $u, u' \in U$ ,  $\exists g \in G$  such that  $R_g(u) = u'$ . A transitive set representation has only one orbit, all of  $U$ . Any set "irreducible representation" is transitive.

We are accustomed to thinking of some distinction-making operation as reducing a whole partition to a more refined partition, and thus breaking up a block  $B$  into

distinguishable non-overlapping subsets  $B', B'', \dots \subseteq B$ . Now we are working at the more basic level of determining the distinct eigen-alternatives, i.e., the orbits of a set representation of a symmetry group. Here we might also consider how distinctions are made to move to a more refined partition of orbits. Since the group operations identify elements,  $u \sim u'$  if  $\exists g \in G$  such that  $R_g(u) = u'$ , we would *further distinguish* elements by moving to a subgroup, i.e., fewer group elements making indistinctions so more distinctions in maximally distinct eigen-alternatives. The symmetry operations in the larger group are "broken," so the remaining group of symmetries is a subgroup. That is how *symmetry-breaking* ([13]; [48]) is accounted for in this interpretation.

**Example 1 revisited:** the group  $S_2$  has only one subgroup, the trivial subgroup of the identity operation, and its orbits are clearly the singletons  $\{u\}$  for  $u \in U$ . That is the simplest example of symmetry-breaking that gives a more distinct set of eigen-alternatives.

In any set representation, the maximum distinctions are made by the smallest symmetry subgroup which is always the identity subgroup, so that is always the waste case that takes us back to the singleton orbits in  $U$  which form the discrete partition on  $U$ .

Thus we see that symmetry-breaking is analogous to distinction-making measurement but at this more fundamental level where the distinct eigen-forms (orbits in the orbit-partition) are determined in the first place by symmetry operations.

### 7.3 Attributes and observables

An *attribute* on a set  $U$  is a function  $f : U \rightarrow \mathbb{R}$ . An attribute induces a set partition  $\{f^{-1}(r)\}_{r \in f(U)}$  on  $U$ . An attribute  $f : U \rightarrow \mathbb{R}$  *commutes* with a set representation  $R : G \rightarrow S(U)$  if for any  $R_g$ , the following diagram commutes in the sense that  $fR_g = f$ :

$$\begin{array}{ccc} U & \xrightarrow{R_g} & U \\ & \searrow f & \downarrow f \\ & & \mathbb{R} \end{array}$$

Commuting attribute.

The lifts to vector space representations are immediate:

- a real-valued attribute on a set  $\xrightarrow{\text{Lifts}}$  an observable represented by a Hermitian operator on a complex vector space; and

- the commutativity condition on a set-attribute  $\xrightarrow{\text{Lifts}}$  an observable operator  $H$  (like the Hamiltonian) commuting with a symmetry group in the sense that  $HR_g = R_gH$  for all  $g \in G$ .

The commutativity-condition in the set case means that whenever  $R_g(u) = u'$  then  $f(u) = f(u')$ , i.e., that  $f$  is an *invariant* of the group. Recall that each orbit of a set representation is transitive so for any  $u, u'$  in the same orbit,  $\exists R_g$  such that  $R_g(u) = u'$  so  $f(u) = f(u')$  for any two  $u, u'$  in the same orbit. In other words:

**Schur's Lemma** (set version): a commuting attribute restricted to an orbit is constant.

The lift to vector space representations is one version of the usual

**Schur's Lemma** (vector space version): An operator  $H$  commuting with  $G$  restricted to an irreducible subspace is a constant operator.

This also means that the inverse-image partition  $\{f^{-1}(r)\}$  of a commuting attribute is refined by the orbit partition. If an orbit  $B \subseteq f^{-1}(r)$ , then the eigenvalue  $r$  of the attribute  $f$  is associated with that orbit. Every commuting attribute  $f : U \rightarrow \mathbb{R}$  can be uniquely expressed as a spectral decomposition:

$$f = \sum_{B \in \text{Orbits}} r_B \chi_B,$$

where  $r_B$  is the constant value on the orbit  $B \subseteq U$  and  $\chi_B : U \rightarrow \mathbb{R}$  is the characteristic function of the orbit  $B$ .

There may be other orbits with the same  $f$ -value. Then we would need another commuting attribute  $g : U \rightarrow \mathbb{R}$  so that for each orbit  $B$ , there is an eigenvalue  $s$  of the attribute  $g$  such that  $B \subseteq g^{-1}(s)$ . Then the eigen-alternative  $B$  may be characterized by the ordered pair  $|r, s\rangle$  if  $B = f^{-1}(r) \cap g^{-1}(s)$ . If not, we continue until we have a *Complete Set of Commuting Attributes* (CSCA) whose ordered  $n$ -tuples of eigenvalues would characterize the eigen-alternatives, the orbits of the set representation  $R : G \rightarrow S(U)$ .

Obviously, we are just spelling out the set version whose lift is the use of a Complete Set of Commuting Operators (CSCO) to characterize the eigenstates by kets of ordered  $n$ -tuples  $|\lambda, \mu, \dots\rangle$  of eigenvalues of the commuting operators. *But* these eigenstates are not the singletons  $\{u\}$  but are the maximally distinct invariant subsets or orbits of the set representation of the symmetry group  $G$ . The basic theme is that the indefinite elements are distinguished to form more definite entities by the distinctions made by the joins of the inverse-image partitions of compatible attributes.

**Example 1 again:** Consider the attribute  $f : U = \{0, 1, 2, 3, 4, 5\} \rightarrow \mathbb{R}$  where  $f(n) = n \bmod 3$ . This attribute commutes with the previous set representation of  $S_2$ , namely  $R_1 = 1_U$  and  $R_\sigma(u) = u + 3 \bmod 6$ , and accordingly by Schur's Lemma (set version), the attribute is constant on each orbit  $\{0, 3\}$ ,  $\{1, 4\}$ , and  $\{2, 5\}$ . In this case, the blocks of the inverse-image partition  $\{f^{-1}(0), f^{-1}(1), f^{-1}(2)\}$  equal the blocks of the orbit partition, so this attribute is the set version of a nondegenerate measurement in that its eigenvalues suffice to characterize the eigen-alternatives, i.e., the orbits. By itself, it forms a complete set of attributes.

**Example 2:** Let  $U = \{0, 1, \dots, 11\}$  where  $S_2 = \{1, \sigma\}$  is represented by the operations  $R_1 = 1_U$  and  $R_\sigma(n) = n + 6 \bmod (12)$ . Then the orbits are  $\{0, 6\}$ ,  $\{1, 7\}$ ,  $\{2, 8\}$ ,  $\{3, 9\}$ ,  $\{4, 10\}$ , and  $\{5, 11\}$ . Consider the attribute  $f : U \rightarrow \mathbb{R}$  where  $f(n) = n \bmod (2)$ . This attribute commutes with the symmetry group and is thus constant on the orbits. But the blocks in the inverse-image partition are now larger than the orbits, i.e.,  $f^{-1}(0) = \{0, 2, 4, 6, 8, 10\}$  and  $f^{-1}(1) = \{1, 3, 5, 7, 9, 11\}$  so the orbit partition strictly refines  $\{f^{-1}(r)\}$ . Thus this attribute corresponds to a degenerate measurement in that the two eigenvalues do not suffice to characterize the orbits.

Consider the attribute  $g : U \rightarrow \mathbb{R}$  where  $g(n) = n \bmod (3)$ . This attribute commutes with the symmetry group and is thus constant on the orbits. The blocks in the inverse-image partition are:  $g^{-1}(0) = \{0, 3, 6, 9\}$ ,  $g^{-1}(1) = \{1, 4, 7, 10\}$ , and  $g^{-1}(2) = \{2, 5, 8, 11\}$ . The blocks in the join of the two partitions  $\{f^{-1}(r)\}$  and  $\{g^{-1}(s)\}$  are the non-empty intersections of the blocks as shown in Table 7.1.

$f^{-1}(r)$	$g^{-1}(s)$	$f^{-1}(r) \cap g^{-1}(s)$	$ r, s\rangle$
$\{0, 2, 4, 6, 8, 10\}$	$\{0, 3, 6, 9\}$	$\{0, 6\}$	$ 0, 0\rangle$
$\{0, 2, 4, 6, 8, 10\}$	$\{1, 4, 7, 10\}$	$\{4, 10\}$	$ 0, 1\rangle$
$\{0, 2, 4, 6, 8, 10\}$	$\{2, 5, 8, 11\}$	$\{2, 8\}$	$ 0, 2\rangle$
$\{1, 3, 5, 7, 9, 11\}$	$\{0, 3, 6, 9\}$	$\{3, 9\}$	$ 1, 0\rangle$
$\{1, 3, 5, 7, 9, 11\}$	$\{1, 4, 7, 10\}$	$\{1, 7\}$	$ 1, 1\rangle$
$\{1, 3, 5, 7, 9, 11\}$	$\{2, 5, 8, 11\}$	$\{5, 11\}$	$ 1, 2\rangle$

Table 7.1:  $f$  and  $g$  as a complete set of commuting attributes.

Thus  $f$  and  $g$  form a Complete Set of Commuting Attributes to characterize the eigen-alternatives, the orbits, by the kets of ordered pairs of their eigenvalues.

**Example 3:** Let  $U = \mathbb{R}^2$  as a set and let  $G$  be the special orthogonal matrix group  $SO(2, \mathbb{R})$  of matrices of the form;

$$\begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \text{ for } 0 \leq \varphi < 2\pi.$$

This group is trivially represented by the rotations in  $U = \mathbb{R}^2$ :

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

The orbits are the circular orbits around the origin. The attribute "radius"  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  where  $f(x, y) = \sqrt{x^2 + y^2}$  commutes with the representation since:

$$\begin{aligned} f(x', y') &= \sqrt{(x')^2 + (y')^2} \\ &= \sqrt{(x \cos \varphi - y \sin \varphi)^2 + (x \sin \varphi + y \cos \varphi)^2} \\ &= \sqrt{x^2 (\cos^2 \varphi + \sin^2 \varphi) + y^2 (\cos^2 \varphi + \sin^2 \varphi)} \\ &= f(x, y). \end{aligned}$$

That means that "radius" is an invariant of the rotation symmetry group. The blocks in the set partition  $\{f^{-1}(r) : 0 \leq r\}$  of  $\mathbb{R}^2$  coincide with the orbits so the eigenvalues of the radius attribute suffice to characterize the orbits.

**Example 4:** The Cayley set representation of any group  $G$  is given by permutations on  $U = G$  itself defined by  $R_g(g') = gg'$ , which is also called the *left regular representation*. Given any  $g, g' \in G$ ,  $R_{g'g^{-1}}(g) = g'$  so the Cayley representation is always transitive, i.e., has only one orbit consisting of all of  $U = G$ . Since any commuting attribute  $f : U = G \rightarrow \mathbb{R}$  is constant on each orbit, it can only be a constant function such as  $\chi_G$ .

Thus the Cayley set representation is rather simple, but we could break some symmetry by considering a proper subgroup  $H \subseteq G$ . Then using only the  $R_h$  for  $h \in H$ , we have a representation  $H \rightarrow S(G)$ . The orbit-defining equivalence relation is  $g \sim g'$  if  $\exists h \in H$  such that  $hg = g'$ , i.e., the orbits are the *right cosets*  $Hg$ .

The linearization program from set representations to vector space representations is summarized in Table 7.2 and will be illustrated in the next section.



Sets to Vector Spaces	Set representations of groups	Vector space ( $\mathbb{C}$ ) representations of groups
Representation	Group $G$ represented by permutations $R_g : U \rightarrow U$	Group $G$ represented by invertible linear ops: $R_g : V \rightarrow V$
Min. Invariants	Orbits	Irreducible subspaces
Partitions	Set partitions of orbits	Direct-sum decomposition of irreducible subspaces
Irreducible reps	Reps restricted to orbits	Reps restricted to irreducible subspaces
Commuting with rep.	Attribute $f : U \rightarrow \mathbb{R}$ commuting with rep., i.e. $fR_g = f$ for all $R_g$	Operator $H$ commuting with all $R_g$ , i.e., $HR_g = R_gH$
Invariants	Inverse-images $f^{-1}(r)$ for commuting $f$ are invariant	Eigenspaces of commuting $H$ are invariant
Schur's Lemma	Commuting $f$ restricted to orbit is constant	Commuting $H$ restricted to irreducible subspace is constant

Table 7.2: Linearizing set concepts to vector space concepts.

## 7.4 Irreps of vector space representations

Our conceptual purpose here is to describe how group representation theory, the set version or the lift to vector spaces, answers the question of determining the form of the distinct eigen-alternatives. Given a group of symmetries acting on a set or on a vector space, what are the *most distinct* subsets or subspaces that satisfy the symmetries? Those minimal invariant subobjects are the orbits in the set case and the irreducible subspaces in the vector space case. In each case, those subobjects give the appropriate type of partition (i.e., a set partition or a vector space partition, i.e., a direct-sum decomposition). And from partition logic, we know that the way to carve out more refined alternatives is the join of partitions. We have already illustrated this in the set case where the inverse-image partitions of a complete set of compatible attributes were joined so that they characterized the most distinct eigen-forms, the orbits of the representation. We will illustrate how the same partition methods of CSCOs apply in the lifted case of vector space representations.

The distinct eigen-alternatives in the set case (the orbits) cannot show much variation since there is only the question of an element of the universe set  $U$  being in or out of a subset. That is, viewing a subset  $S \subseteq U$  as a vector in  $\mathbb{Z}_2^{|U|}$ , the coefficients of each basis vector are only 0 or 1. But when we lift to group representations over a vector space with  $\mathbb{C}$  as the base field, then there is much more variation in distinct eigen-alternatives. And it is that much wider range of maximally-distinct eigen-forms, the irreducible representations that are the representations restricted to the irreducible subspaces, that are of such importance in quantum mechanics.

The partition-join method of determining the eigen-forms for vector space representations will be illustrated with several simple examples with finite dimensional vector spaces  $V$ . Recall that a *vector space representation* of a group  $G$  is given by an assignment of an invertible linear operator  $R_g : V \rightarrow V$  to each  $g \in G$  such that  $R_1 = 1_V$  and  $R_{g'}R_g = R_{g'g}$ .

**Example 5:** The multiplicative group  $S_2 \times S_2$  written additively is the Klein four-group  $G = \mathbb{Z}_2 \times \mathbb{Z}_2 = \{(0, 0), (1, 0), (0, 1), (1, 1)\}$ . The complex vector space  $\{\mathbb{Z}_2 \times \mathbb{Z}_2 \rightarrow \mathbb{C}\}$  of all complex-valued functions on the four-element set  $\mathbb{Z}_2 \times \mathbb{Z}_2$  is the *Cayley group space* of that group. A basis for the four-dimensional space  $\mathbb{C}^4$  is the set of functions  $|g'\rangle$  which take value 1 on  $g'$  and 0 on the other  $g \in G$ . Then the action of the group on this space is defined by  $R_g(|g'\rangle) = |g + g'\rangle$  (or  $|gg'\rangle$  if the group operation was written multiplicatively). Thus the group action just permutes the basis vectors in the Cayley group space and would be represented by permutation matrices. The non-identity operators have the matrices;

$$R_{(1,0)} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{matrix} (0,0) \\ (1,0) \\ (0,1) \\ (1,1) \end{matrix}; R_{(0,1)} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}; R_{(1,1)} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Since the group is Abelian, each of these operators can be viewed as an observable  $H$  that commutes with the  $R_g$  for  $g \in G$ , so its eigenspaces will be invariant under the group operations.

For  $R_{(1,0)}$ , the invariant eigenspaces with their eigenvalues and generating eigenvectors are:

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \right\} \leftrightarrow \lambda = -1, \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \right\} \leftrightarrow \lambda = 1.$$

For  $R_{(0,1)}$ , we have:

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \right\} \leftrightarrow \lambda = -1, \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \right\} \leftrightarrow \lambda = 1.$$

Since these two operators commute, their eigenspace partitions are compatible so we can take their join. The blocks of the join are a vector space partition and are automatically invariant. Since the blocks of the join are one-dimensional, those four subspaces are also irreducible and thus those two operators form a complete set of commuting operators (CSCO). The commuting operators always have a set of simultaneous eigenvectors, and we have arranged the generating eigenvectors of the eigenspaces so that they are all simultaneous eigenvectors which can, as usual, be characterized by kets using the respective eigenvalues;

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = |1, 1\rangle; \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} = |-1, 1\rangle; \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} = |1, -1\rangle; \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = |-1, -1\rangle.$$

The group representation restricted to these four irreducible subspaces give the four irreducible representations or irreps of the group. Since any vector can be uniquely decomposed into the sum of vectors in the irreducible subspaces, the representation on the whole space can be expressed, in the obvious sense, as the direct sum of the irreps.

In the set case, moving to smaller invariant subsets by making distinctions gives more distinct elements, so in the vector space case, moving to smaller invariant subspaces give more distinct alternatives. The minimal invariant subspaces, i.e., the irreducible subspaces, thus give the maximally-distinct invariant subspaces, and the representation restricted to those subspaces gives the maximally-distinct symmetry-respecting alternatives, i.e., the irreps.

This might be illustrated by giving a geometric version of the representation. Consider a rectangle  $\begin{bmatrix} a & b \\ d & c \end{bmatrix}$  under the operations of flipping on the horizontal axis  $\sigma_h : \begin{bmatrix} a & b \\ d & c \end{bmatrix} \mapsto \begin{bmatrix} d & c \\ a & b \end{bmatrix}$  and flipping on the vertical axis:  $\sigma_v : \begin{bmatrix} a & b \\ d & c \end{bmatrix} \mapsto \begin{bmatrix} b & a \\ c & d \end{bmatrix}$  as well as their composition  $\sigma_{hv} : \begin{bmatrix} a & b \\ d & c \end{bmatrix} \mapsto \begin{bmatrix} c & d \\ b & a \end{bmatrix}$  and the identity. This gives the same  $S_2 \times S_2$  group with the multiplication Table 7.3.

2nd \ 1st	1	$\sigma_h$	$\sigma_v$	$\sigma_{hv}$
1	$\begin{bmatrix} a & b \\ d & c \end{bmatrix}$	$\begin{bmatrix} d & c \\ a & b \end{bmatrix}$	$\begin{bmatrix} b & a \\ c & d \end{bmatrix}$	$\begin{bmatrix} c & d \\ b & a \end{bmatrix}$
$\sigma_h$	$\begin{bmatrix} d & c \\ a & b \end{bmatrix}$	$\begin{bmatrix} a & b \\ d & c \end{bmatrix}$	$\begin{bmatrix} c & d \\ b & a \end{bmatrix}$	$\begin{bmatrix} b & a \\ c & d \end{bmatrix}$
$\sigma_v$	$\begin{bmatrix} b & a \\ c & d \end{bmatrix}$	$\begin{bmatrix} c & d \\ b & a \end{bmatrix}$	$\begin{bmatrix} a & b \\ d & c \end{bmatrix}$	$\begin{bmatrix} d & c \\ a & b \end{bmatrix}$
$\sigma_{hv}$	$\begin{bmatrix} c & d \\ b & a \end{bmatrix}$	$\begin{bmatrix} b & a \\ c & d \end{bmatrix}$	$\begin{bmatrix} d & c \\ a & b \end{bmatrix}$	$\begin{bmatrix} a & b \\ d & c \end{bmatrix}$

Table 7.3: Multiplication table for  $S_2 \times S_2$  as symmetry group resulting from flipping on the horizontal and vertical axes.

Then the basis vectors for the four irreducible subspaces of the Cayley space (sometimes called the *irreducible basis vectors*) are given in Table 7.4:

$$\begin{aligned} \chi_1 &= \begin{bmatrix} a & b \\ d & c \end{bmatrix} + \begin{bmatrix} d & c \\ a & b \end{bmatrix} + \begin{bmatrix} b & a \\ c & d \end{bmatrix} + \begin{bmatrix} c & d \\ b & a \end{bmatrix} \\ \chi_2 &= \begin{bmatrix} a & b \\ d & c \end{bmatrix} - \begin{bmatrix} d & c \\ a & b \end{bmatrix} + \begin{bmatrix} b & a \\ c & d \end{bmatrix} - \begin{bmatrix} c & d \\ b & a \end{bmatrix} \\ \chi_3 &= \begin{bmatrix} a & b \\ d & c \end{bmatrix} + \begin{bmatrix} d & c \\ a & b \end{bmatrix} - \begin{bmatrix} b & a \\ c & d \end{bmatrix} - \begin{bmatrix} c & d \\ b & a \end{bmatrix} \\ \chi_4 &= \begin{bmatrix} a & b \\ d & c \end{bmatrix} - \begin{bmatrix} d & c \\ a & b \end{bmatrix} - \begin{bmatrix} b & a \\ c & d \end{bmatrix} + \begin{bmatrix} c & d \\ b & a \end{bmatrix}. \end{aligned}$$

Table 7.4: Four irrep basis vectors.

The same information can also be expressed in the character Table 7.5.

	$\begin{bmatrix} a & b \\ d & c \end{bmatrix}$	$\begin{bmatrix} d & c \\ a & b \end{bmatrix}$	$\begin{bmatrix} b & a \\ c & d \end{bmatrix}$	$\begin{bmatrix} c & d \\ b & a \end{bmatrix}$
$\chi_1$	1	1	1	1
$\chi_2$	1	-1	1	-1
$\chi_3$	1	1	-1	-1
$\chi_4$	1	-1	-1	1

Table 7.5: Character table for  $S_2 \times S_2$ .

The one-dimensional subspace generated by any one of the vectors is invariant under the group operations. For instance, if we apply  $\sigma_v$  to  $\chi_3$ , we get:

$$R_{\sigma_v}(\chi_3) = \begin{bmatrix} b & a \\ c & d \end{bmatrix} + \begin{bmatrix} c & d \\ b & a \end{bmatrix} - \begin{bmatrix} a & b \\ d & c \end{bmatrix} - \begin{bmatrix} d & c \\ a & b \end{bmatrix} = -\chi_3.$$

These four irrep basis vectors represent the maximally distinct (e.g., mutually orthogonal) eigen-forms that respect the symmetry operations. Each one is a superposition of the original four basis vectors which were not "symmetry-adapted." By combining the original basis vectors in this way, we get the maximally distinct eigen-forms obeying the symmetry group.

We can also use this example to illustrate how the vector space representations, as opposed to the set representations, generate more variety due to the richer base field. Instead of the Cayley space, we consider the *Cayley set* which is just the set  $G$  of group operations which we might represent by the set of four configurations obtained from the initial configuration  $\begin{bmatrix} a & b \\ d & c \end{bmatrix}$ , namely  $\left\{ \begin{bmatrix} a & b \\ d & c \end{bmatrix}, \begin{bmatrix} d & c \\ a & b \end{bmatrix}, \begin{bmatrix} b & a \\ c & d \end{bmatrix}, \begin{bmatrix} c & d \\ b & a \end{bmatrix} \right\}$ . Each group operation such as  $R_{\sigma_v}$  acts on these four elements as indicated by the row in the multiplication table. Thus we have a set representation of the group whose orbits will be the minimal invariant subsets. But in every Cayley set representation, there is only one orbit since any element can be mapped to any other element by the action of one of the four operators, i.e., the Cayley group action is transitive. Thus the only eigen-form we get from the Cayley set representation of  $S_2 \times S_2$  is the minimal invariant subset or orbit  $\left\{ \begin{bmatrix} a & b \\ d & c \end{bmatrix}, \begin{bmatrix} d & c \\ a & b \end{bmatrix}, \begin{bmatrix} b & a \\ c & d \end{bmatrix}, \begin{bmatrix} c & d \\ b & a \end{bmatrix} \right\}$  which corresponds to  $\chi_1$  in the Cayley vector space representation. Thus by working in any base field where  $+1 \neq -1$ , we immediately get a richer set of symmetry-adapted eigen-forms in addition to the always-present  $\chi_1$ .

A key part in the above derivation was the fact each  $R_g$  could be treated as an operator commuting with all the  $R_{g'}$  for  $g' \in G$  since the group was Abelian. A few extra wrinkles must be added when the group is not Abelian.

**Example 6:** The smallest non-Abelian group is the symmetric group  $S_3$  on three elements which is isomorphic to  $D_3$ , the full symmetry group for the equilateral triangle. Starting with the initial configuration  ${}_b\overset{a}{\Delta}_c$ , the non-identity symmetries are (where the "bottom side" of the triangle is "painted dark" so the triangle turns dark when flipped over and where the vertexes are labeled as in the original configuration  ${}_b\overset{a}{\Delta}_c$  when defining the flips):

$$\text{rotation counterclockwise by } 120^\circ, C_3 : {}_b\overset{a}{\Delta}_c \mapsto {}_a\overset{c}{\Delta}_b;$$

rotation counterclockwise by  $240^\circ$ ,  $C_3^2 : b \triangle_c^a \mapsto c \triangle_a^b$ ;  
 flip on bisector through vertex  $a$ ,  $C_2^{(a)} : b \triangle_c^a \mapsto c \blacktriangle_b^a$ ;  
 flip on bisector through vertex  $b$ ,  $C_2^{(b)} : b \triangle_c^a \mapsto b \blacktriangle_a^c$ ;  
 flip on bisector through vertex  $c$ ,  $C_2^{(c)} : b \triangle_c^a \mapsto a \blacktriangle_c^b$ .

In the following multiplication table, the column on the left has the inverses of the elements in the top row so that the identity will always be on the diagonal. The operations are represented in Table 7.6 by the configuration resulting from applying the operation to the initial configuration.

$2nd \setminus 1st$	$I$	$C_3$	$C_3^2$	$C_2^{(a)}$	$C_2^{(b)}$	$C_2^{(c)}$
$I$	$b \triangle_c^a$	$a \triangle_b^c$	$c \triangle_a^b$	$c \blacktriangle_b^a$	$b \blacktriangle_a^c$	$a \blacktriangle_c^b$
$C_3^2$	$c \triangle_a^b$	$b \triangle_c^a$	$a \triangle_b^c$	$b \blacktriangle_a^c$	$a \blacktriangle_c^b$	$c \blacktriangle_b^a$
$C_3$	$a \triangle_b^c$	$c \triangle_a^b$	$b \triangle_c^a$	$a \blacktriangle_c^b$	$c \blacktriangle_b^a$	$b \blacktriangle_a^c$
$C_2^{(a)}$	$c \blacktriangle_b^a$	$b \blacktriangle_a^c$	$a \blacktriangle_c^b$	$b \triangle_c^a$	$a \triangle_b^c$	$c \triangle_a^b$
$C_2^{(b)}$	$b \blacktriangle_a^c$	$a \blacktriangle_c^b$	$c \blacktriangle_b^a$	$c \triangle_a^b$	$b \triangle_c^a$	$a \triangle_b^c$
$C_2^{(c)}$	$a \blacktriangle_c^b$	$c \blacktriangle_b^a$	$b \blacktriangle_a^c$	$a \triangle_b^c$	$c \triangle_a^b$	$b \triangle_c^a$

Table 7.6: Multiplication table for  $D_3$ .

Table 7.6 indicates that the group is not Abelian. We form the *Cayley space*  $V = \{D_3 \rightarrow \mathbb{C}\} \cong \mathbb{C}^6$  with is spanned by the *standard basis* vectors  $|g\rangle$  which take the value 1 on  $g \in D_3$  and 0 elsewhere. The operators on the Cayley space  $R_{g'} : \mathbb{C}^6 \rightarrow \mathbb{C}^6$  are defined by  $R_{g'} : |g\rangle \mapsto |g'g\rangle$  which permutes the basis vectors.

In the previous example of the Abelian group  $S_2 \times S_2$ , the next step was to consider each  $R_g$  like an observable that commutes with all the  $R_{g'}$  of the representation. That is not possible for a non-Abelian group so we need to construct operators that do commute with all the group operators and with each other. If we took a subset  $C \subseteq \{R_g : g \in D_3\}$ , then the requirement that the subset "commute" with all the  $R_g$  would be  $R_g C = C R_g$  or  $R_g C R_{g^{-1}} = C$ . That means for any  $R_h \in C$  that  $R_g R_h R_{g^{-1}} \in C$ . If we define a binary relation  $R_h \sim R_{h'}$  if  $R_g R_h R_{g^{-1}} = R_{h'}$  then that *conjugacy* relation is an equivalence relation, and the equivalence classes are the *conjugacy classes*. A conjugacy class can be turned into a single operator by summing the operators in the class. For a conjugacy class  $C \subseteq \{R_g : g \in D_3\}$ , let  $\mathcal{C} = \sum_{R_g \in C} R_g$  be the *class sum operator* of the class.

A little computation shows that  $D_3$  divides up into three conjugacy classes  $\{I\}$  (always a class by itself),  $\{C_3, C_3^2\}$ , and  $\{C_2^{(a)}, C_2^{(b)}, C_2^{(c)}\}$ . If we represent the  $R_g$ 's by their configurations, then the three class sum operators are:

$$\begin{aligned}\mathcal{C}_1 &= I = {}_b\Delta_c^a \\ \mathcal{C}_2 &= C_3 + C_3^2 = {}_a\Delta_b^c + {}_c\Delta_a^b \\ \mathcal{C}_3 &= C_2^{(a)} + C_2^{(b)} + C_2^{(c)} = {}_c\blacktriangle_b^a + {}_b\blacktriangle_a^c + {}_a\blacktriangle_c^b.\end{aligned}$$

### Class sum operators

Since the effect of conjugation is only to permute the elements in a conjugacy classes and thus commute the terms in the class sum, the class sum operators commute with all the group operations and thus with each other:

$$R_g\mathcal{C}_i = \mathcal{C}_iR_g \text{ and } \mathcal{C}_i\mathcal{C}_j = \mathcal{C}_j\mathcal{C}_i \text{ for all } g \in D_3 \text{ and } i, j = 1, 2, 3.$$

The class sums are also vectors in the Cayley space  $V$  (like  ${}_c\blacktriangle_b^a + {}_b\blacktriangle_a^c + {}_a\blacktriangle_c^b$ ) and they are linearly independent since they are sums of disjoint sets of linearly independent basis vectors. The three-dimensional subspace  $V_c \subseteq V$  generated by the class sums is the *class space*. In the class space, we are back in the situation of a space generated by a set of vectors which can also be seen as operators acting on the space, and those operators commute with each other. And since they commute with each other, they will determine a set of simultaneous eigenvectors in the usual manner, and those will be the basis vectors for the irreducible subspaces (of the class space  $V_c$ ) that are the carriers for the irreps.

But the action of the class sum operators on the class sum vectors in  $V_c$  is not just a permutation so we need to construct the multiplication Table 7.7 for the class sum operators. For instance, consider the multiplication:

$$\mathcal{C}_2\mathcal{C}_2 = (C_3 + C_3^2)^2 = C_3^2 + 2C_3C_3^2 + (C_3^2)^2 = C_3^2 + 2I + C_3 = 2\mathcal{C}_1 + \mathcal{C}_2.$$

	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$
$\mathcal{C}_1$	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$
$\mathcal{C}_2$	$\mathcal{C}_2$	$2\mathcal{C}_1 + \mathcal{C}_2$	$2\mathcal{C}_3$
$\mathcal{C}_3$	$\mathcal{C}_3$	$2\mathcal{C}_3$	$3\mathcal{C}_1 + 3\mathcal{C}_2$

Table 7.7: Multiplication table for class sum operators  
(commutativity indicated by symmetry of table).

Then the operation of the class sum operators on  $V_c$  can be described by the following matrices:

$$\mathcal{C}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{matrix} \mathcal{C}_1 \\ \mathcal{C}_2 \\ \mathcal{C}_3 \end{matrix}, \mathcal{C}_2 = \begin{bmatrix} 0 & 2 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \mathcal{C}_3 = \begin{bmatrix} 0 & 0 & 3 \\ 0 & 0 & 3 \\ 1 & 2 & 0 \end{bmatrix}.$$

All the vectors of the class space  $V_c$  are eigenvectors of  $\mathcal{C}_1$  with  $\lambda = 1$ . The eigenspaces, eigenvalues, and generating eigenvectors of  $\mathcal{C}_2$  are:

$$\left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \right\} \leftrightarrow \lambda = -1, \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\} \leftrightarrow \lambda = 2.$$

For  $\mathcal{C}_3$ , we have:

$$\left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \right\} \leftrightarrow \lambda = 0, \left\{ \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\} \leftrightarrow \lambda = -3, \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \leftrightarrow \lambda = 3.$$

The generating eigenvectors have been expressed so that they are also the simultaneous eigenvectors. But there is no need in this case to take the join of the eigenspace partitions since the eigenspaces of the one operator  $\mathcal{C}_3$  are one-dimensional so it forms a CSCO by itself for the class space. The three simultaneous eigenvectors are the basis vectors for the three irreducible subspaces of  $V_c$  and they can be arranged as the character Table 7.8.

	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$
$\chi_1$	1	1	1
$\chi_2$	1	1	-1
$\chi_3$	2	-1	0

Table 7.8: Character table for  $D_3 \cong S_3$ .

For an Abelian group, the conjugacy classes are all singletons, so the procedure followed in that case is a special instance of the general procedure for non-Abelian groups using class sum operators.



Given two representations  $R = \{R_g : V \rightarrow V\}_{g \in G}$  and  $R' = \{R'_g : V' \rightarrow V'\}_{g \in G}$  of the same group  $G$ , the representations are said to be *equivalent* if there is a non-singular linear transformation  $S : V \rightarrow V'$  such that for any  $g \in G$ ,  $R_g = S^{-1}R'_gS$ . The remarkable fact is the three eigen-forms determined by the simultaneous eigenvectors given in the character table are the only inequivalent irreducible representations used in any representation of  $D_3$ , not just the Cayley representation. Any representation of  $D_3$  can be expressed as a direct sum (with repetitions) of those three inequivalent irreps.

The class space is a subspace of the whole Cayley space, but the operator  $\mathcal{C}_3$ , which was a CSCO in the class space, is not a CSCO in the whole space. As an operator on the whole space,

$$\mathcal{C}_3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} I \\ C_3^2 \\ C_3 \\ C_2^{(a)} \\ C_2^{(b)} \\ C_2^{(c)} \end{matrix}$$

it has the same eigenvalues, and the eigenspaces  $V_3$  and  $V_{-3}$  for  $\lambda = 3, -3$  are still one-dimensional, but the eigenspace  $V_0$  for  $\lambda = 0$  is four-dimensional. That four-dimensional eigenspace is invariant but it is not minimal. We need to find a new "symmetry-adapted" or "irreducible" basis so that when all the group operations  $R_g$  are expressed in that basis, then their matrices will be block-diagonal in the same pattern so that the multiplication of those block-diagonal matrices will only multiply within the blocks. Since the multiplications stay within the blocks, the columns for each block will generate an invariant subspace. When the blocks are of minimum size, then the columns will give irreducible subspaces (some of which may be equivalent). The whole space is a direct sum of those irreducible subspaces, and the original representation is said to be a direct sum of the irreps formed by restricting the representation to the irreducible subspaces.

In the case of the Cayley representation of  $D_3$ , it is the direct sum of the first two irreps and two copies of the third one. But the expression of the Cayley representation in these terms is far from unique. In many texts, the required symmetry-adapted basis is generated in a rather *ad hoc* manner. But Jin-Quan Chen and his colleagues in the

Nanjing School have developed a CSCO method (i.e., partition joins) to systematically find the irreducible basis vectors for the irreducible spaces that works not only for all representations of finite groups but for all compact Lie groups as needed in QM ([10], [11]). "[T]he foundation of the new approach is precisely the theory of the complete set of commuting operators (CSCO) initiated by Dirac..." [11, p. 2] Thus the linearized partition math of the CSCO method extends also to all compact group representations to characterize the maximally definite eigen-alternatives.

In general, to distinguish indefinite elements into more definite elements, we need more compatible partitions so the joins will have more definite blocks. In the lifted version, more compatible partitions mean more commuting operators whose eigenspace partitions can thus be joined with the given partition  $\{V_3, V_{-3}, V_0\}$  to find smaller invariant subspaces until we arrive at minimal ones. Applying the CSCO method of the Chen School, there are two sources of new commuting operators. The CSCOs of a subgroup chain  $G \supset G(1) \supset \dots \supset G(m)$  is one source.

But each group  $G$  has an *opposite group*  $\bar{G}$  (Chen calls it the *intrinsic group*) and all the Cayley space constructions can be carried out for the opposite group which is anti-isomorphic to  $G$  by the mapping  $g^{-1} \longleftrightarrow \bar{g}$ . The CSCO of the whole opposite group  $\bar{G}$  is of no help in cutting down the invariant subspaces, but the CSCOs of the subgroup chain  $\bar{G} \supset \bar{G}(1) \supset \dots \supset \bar{G}(m)$  form a second source of commuting operators to reduce the invariant subspaces.

In the case at hand,  $D_3 \cong S_3$  is the symmetric group on three elements, and there are three copies of the subgroup  $S_2$ , i.e., those generated by  $C_2^{(a)}$ ,  $C_2^{(b)}$ , or  $C_2^{(c)}$ . Arbitrarily picking  $C_2^{(c)}$ , its matrix and the matrix for  $\bar{C}_2^{(c)}$  as operators on  $V$  are their respective CSCOs:

$$C_2^{(c)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \bar{C}_2^{(c)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The three matrices  $\mathcal{C}_3$ ,  $C_2^{(c)}$ , and  $\bar{C}_2^{(c)}$  commute, and the join of the eigenspace partitions is nondegenerate, i.e., all blocks are one-dimensional. The six simultaneous eigenvectors can be characterized by the kets of eigenvalues (as usual). Grouping

the two columns together for each  $\bar{C}_2^{(c)}$  eigenvalue gives the matrix whose columns are symmetry-adapted basis vectors that will block-diagonalize the matrices of the original group representation into irreducible blocks:

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & -1 \\ 1 & 1 & -\frac{1}{2} & -1 & 1 & \frac{1}{2} \\ 1 & 1 & -\frac{1}{2} & 1 & -1 & \frac{1}{2} \\ 1 & -1 & -\frac{1}{2} & -1 & -1 & -\frac{1}{2} \\ 1 & -1 & -\frac{1}{2} & 1 & 1 & -\frac{1}{2} \\ 1 & -1 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Transposed, the new basis vectors as rows have the indicated kets of eigenvalues.

	$I$	$C_3^2$	$C_3$	$C_2^{(a)}$	$C_2^{(b)}$	$C_2^{(c)}$
$ 3, 1, 1\rangle$	1	1	1	1	1	1
$ 3, -1, -1\rangle$	1	1	1	-1	-1	-1
$ 0, 1, 1\rangle$	1	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	1
$ 0, -1, 1\rangle$	0	-1	1	-1	1	0
$ 0, 1, -1\rangle$	0	1	-1	-1	1	0
$ 0, -1, -1\rangle$	-1	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	1

Table 7.9: Symmetry-adapted basis in rows with eigenvalue labels.

The matrix  $A$  represents the symmetry-adapted or irreducible basis vectors in the standard basis, and it can be considered as the change-of-basis matrix  $A = C_{St-SA}$  to convert the symmetry-adapted basis  $SA$  to the standard basis  $St$ . Then we can use  $A$  and its inverse to convert the five non-identity group representation matrices from the standard to the symmetry-adapted basis.

For instance, the rotate-by-120° matrix in the standard basis and in the symmetry-adapted basis are:

$$C_3 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \text{ and}$$

$$\Gamma(C_3) = C_{SA \leftarrow St} C_3 C_{St \leftarrow SA} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & -1 & 0 & 0 \\ 0 & 0 & \frac{3}{4} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & \frac{3}{4} \\ 0 & 0 & 0 & 0 & -1 & -\frac{1}{2} \end{bmatrix}$$

which has the appropriate block-diagonal structure. Similarly the matrices for  $C_2^{(b)}$  are:

$$C_2^{(b)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \text{ and } \Gamma(C_2^{(b)}) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & \frac{3}{4} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & -\frac{3}{4} \\ 0 & 0 & 0 & 0 & -1 & \frac{1}{2} \end{bmatrix}$$

and similarly for the other operations.

All the six operations are turned into block-diagonal matrices of the same structure so that when the matrices are multiplied, only the corresponding blocks are multiplied. In this manner the Cayley space is partitioned into the direct sum of the irreducible subspaces  $V_1 = \{a_1\}$ ,  $V_2 = \{a_2\}$ ,  $V_3 = \{a_3, a_4\}$ , and  $V_4 = \{a_5, a_6\}$ . The irrep obtained by restricting  $\Gamma$  to  $V_1$ , i.e.,  $\Gamma^{(1)}$ , is given by the first character  $\chi_1$  and by restricting  $\Gamma$  to  $V_2$ , i.e.,  $\Gamma^{(2)}$ , is given by  $\chi_2$ . Restricting  $\Gamma$  to  $V_3$  and  $V_4$  give two equivalent irreps  $\Gamma_1^{(3)}$  and  $\Gamma_2^{(3)}$  given by  $\chi_3$ . Then  $V = V_1 \oplus V_2 \oplus V_3 \oplus V_4$  and similarly for the representations:  $\Gamma = \Gamma^{(1)} \oplus \Gamma^{(2)} \oplus \Gamma_1^{(3)} \oplus \Gamma_2^{(3)}$ .

It is easy to get lost in the detailed mathematics of group representation theory and to overlook the basic theme. The theme is that given the symmetry group, the irreps give all the different ways that maximally definite alternatives can be developed consistent with the symmetries. The irreps fill out the symmetry-adapted possibilities. As always, the indefinite is rendered definite by the distinctions made by the joins of partitions collected together in a complete set. In quantum mechanics, the definite eigen-alternatives of the observables are carved out by the joins of vector space partitions of Dirac's CSCOs. In group representation theory, the properties of the eigen-alternatives (e.g., as determined by the irreps of the symmetry group of the Hamiltonian, or as the elementary particles themselves are determined by the irreps

of groups in particle physics. For a certain symmetry group of particle physics, "an elementary particle 'is' an irreducible unitary representation of the group." [59, p. 149] Thus our partitional approach comports with "the soundness of programs that ground particle properties in the irreducible representations of symmetry transformations..." [30, p. 171] These alternatives are carved out by the joins of the vector space partitions of Chen's CSCOs—which constitute a "systematic theory ... established for the rep group based on Dirac's CSCO (complete set of commuting operators) approach in quantum mechanics." [10, p. 211]

All the development of the *mathematics* of partitions from sets to vector spaces and now to group representations over vector spaces would be there even if the physical world was perfectly classical. In that case, the mathematics of partitions would just describe some alternative reality, some hypothetical world where there was objective indefiniteness rather than the classical world of definite properties "all the way down." But that view of the physical world was overthrown by the quantum revolution, and we have found throughout that the mathematics used by quantum mechanics to describe the physical world is precisely this linearized mathematics of partitions describing a world of objective indefiniteness.

## 8 Concluding remarks

QM/Sets addresses at least three different questions.

1. QM/Sets is a pedagogical or toy model of QM/C. Some of the paradoxical aspects of QM/C can be represented in a very simple model that eliminates some of the more difficult aspects of the full mathematical model in Hilbert space. For instance, much effort is expended in the texts to present the Heisenberg indeterminacy principle as an inequality, and then it is "explained" by saying that what it really means is that a maximum precision measurement in one observable means a minimally precise measurement in a conjugate observable. QM/Sets makes the point at the outset that an eigenvector of one observable is a maximal superposition in a conjugate observable. In a similar manner, the QM/Sets treatment of the double-slit experiment makes the relevant points rather directly without any recourse to the dynamics of wave propagation in QM/C.

2. QM/Sets is a simple non-commutative extension of ordinary finite probability theory (without recourse to the sophisticated mathematics of Connes [14] and Voiculescu [64]). The key is to substitute a basis set in  $\mathbb{Z}_2^n$  for the usual given outcome space. Since there are many basis sets, many of the peculiarities of quantum probability can be modeled in this simple setting. But fixing a basis set still gives two types of events associated with a subset  $S$  of  $U$ : the classical event and the new superposition event. Mathematically, the difference is best expressed using density matrices:

- the superposition event has the pure state density matrix  $\rho(S)$  where the non-zero off-diagonal elements represent the coherence between the elements of  $S$  (since the "off-diagonal terms of a density matrix ... are often called quantum coherences because they are responsible for the interference effects typical of quantum mechanics that are absent in classical dynamics" [3, p. 177], and
- the classical event has the mixed state density matrix  $\rho(\delta S)$  which is a diagonal matrix with non-zero diagonal elements representing the classical probabilities of the elements of  $S$  conditional on  $S$ , e.g.,  $\frac{1}{|S|}$  or  $\frac{p_i}{\Pr(S)}$ . QM/Sets also reproduces the fact in QM/C that the two analogous density matrices cannot be distinguished by measurements in the basis (for the matrix representation) but can be distinguished by measurement in another basis.

3. The fact that so many distinguishing features of QM/C can be reproduced in QM/Sets shows which features of QM/C were important or essential and which were not. There is superposition but no waves in QM/Sets. The wave aspects of what used to be called "wave mechanics" could be seen as an 'artifact' of the use of the complex numbers (the natural math for waves) as the algebraically closed extension of the reals (to get a full set of eigenvectors for the observables). That view comports with the fact that the "wave equation" does not describe any actual waves—to the disappointment of Schrödinger himself and many others. The so-called "wave-particle duality" is really the duality between *particles* with indefinite or definite attributes.

From the viewpoint of interpreting quantum mechanics, QM/Sets brings to the forefront the role of partitions and thus distinctions and indistinctions. John von Neu-

mann [65] made the basic division between "type 1" processes (e.g., measurements) and "type 2" processes (unitary evolution). This categorization is easily characterized as the type 1 processes that make distinctions and the type 2 processes that preserve the degree of distinctness unchanged. The evolution of an isolated quantum system is the one that preserves the degree of indistinction between states as measured by the inner product, i.e., unitary transformations, and then the connection to the Schrödinger wave equation comes via Stone's Theorem (connecting one-parameter families of unitary transformations with self-adjoint operators on Hilbert space). And the whole 'story' about particle-wave complementarity should be recast as the contrast between definite and indefinite particle states.

There has often been hints of links between quantum mechanics and information [69]. These hints are realized with the development of the 'classical' and quantum notions of logical entropy ([20]; [25]) which define and measure information in terms of distinctions, differences, distinguishability, and symmetry-breaking. Unlike the notion of von Neumann entropy, logical entropy measures measurement in the sense that the logical information created in a projective measurement is the sum of the absolute squares of the non-zero off-diagonal terms in the density matrix that are zeroed (i.e., the coherences that are decohered or the indistinctions that are distinguished) in the measurement. Moreover, the logical entropy of the post-measurement density matrix has the simple interpretation as the probability that in two independent measurements of the same prepared state, different eigenvalues will be obtained.

The notion of measurement rests precisely on interactions that involve distinctions being made. Feynman makes one of the clearest statements of how measurement hinges on distinguishability and also how measurement has nothing to do with a human or macroscopic apparatus (and thus has nothing to do with decoherence in the sense of Zurek [70]) but only with a distinction-making interaction involving a macroscopic apparatus or not.

If you could, *in principle*, distinguish the alternative *final* states (even though you do not bother to do so), the total, final probability is obtained by calculating the *probability* for each state (not the amplitude) and then adding them together. If you *cannot* distinguish the final states *even in principle*, then the probability amplitudes must be summed before taking the absolute square to find the actual probability.[29, p. 3-9]

For instance, when a particle scatters off the atoms in a crystal, the question of whether or not it should be treated as a superposition of scattering off the different atoms or as scattering off of particular atoms with certain probabilities—hinges on distinguishability. If there was no distinction between scattering off different atoms, then no ‘measurement’ took place in the interaction and the superposition pure state evolves to a pure state. But if there was some distinction caused by scattering off an atom, then the result is the mixed state of scattering off the different atoms with different probabilities. For instance, if all the atoms had spin down and scattering off an atom flipped the spin, then a distinction was made so that constituted a measurement. Thus the difference between an interaction that constitutes a measurement or not is whether or not any distinction exists between the different alternatives. Feynman’s implicit rule might be paraphrased:

"If it makes a difference, then a difference is made."

One way to contrast the definiteness of the classical view and the indefiniteness that characterizes the quantum view is to contrast the two category-theoretically dual notions of subsets and partitions. That duality can be further traced back to the two fundamental notions of *elements* (of a subset) and *distinctions* (of a partition), or "its" and "dits" [24]. Philosophically, these two notions might be respectively the mathematical building blocks of the old metaphysical concepts of matter (or substance) and form (as in in-form-ation), e.g., in the speculations of Heisenberg [34]. The matter versus form idea [2] can be illustrated by comparing the two lattices of subsets and partitions on a set.

For  $U = \{a, b, c\}$ , start at the bottom and move towards the top of each lattice in Figure 8.1.

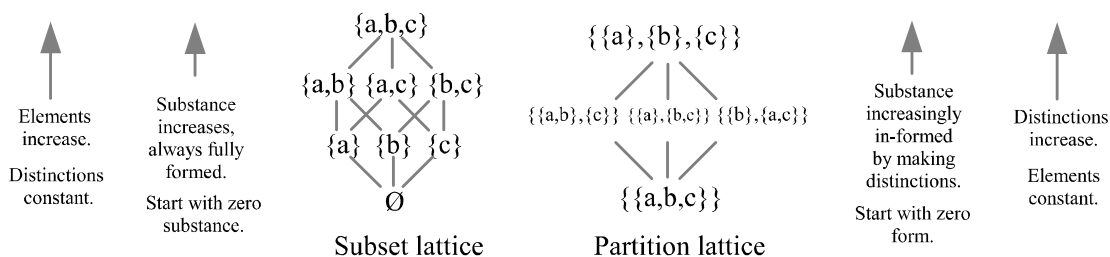


Figure 8.1: Two lattices for two creation stories.



At the bottom of the Boolean subset lattice is the empty set  $\emptyset$  which represents no substance (no elements or ‘its’). As one moves up the lattice, new elements of substance, new elements, are created that are always fully distinguished or formed until finally one reaches the top, the universe  $U$ . Thus new substance is created in moving up the lattice but each element is fully formed and thus always distinguished from the other elements.

At the bottom of the partition lattice is the indiscrete partition or "blob"  $\mathbf{0}_U = \{U\}$  (where the universe set  $U$  makes one block) which represents all the substance or matter but with no distinctions to in-form the substance (no distinctions or ‘dits’). As one moves up the lattice, no new substance is created but distinctions are created that in-form the indistinct elements as they become more and more distinct. Finally one reaches the top, the discrete partition  $\mathbf{1}_U = \{\{u\}\}_{u \in U}$  where all the elements of  $U$  have been fully in-formed or distinguished. A partition combines indefiniteness (within blocks) and definiteness (between blocks). At the top of the partition lattice, the discrete partition  $\mathbf{1}_U = \{\{u\} : \{u\} \subseteq U\}$  is the result making all the distinctions to eliminate any indefiniteness.

Thus one ends up at essentially the same place—a universe of fully formed entities—either way, but by two totally different but dual ‘creation stories’:

- Subset or "Definite all the way down" Creation Story 1: creating elements as in creating fully-formed and distinguished matter out of nothing, versus
- Partition or Objective indefiniteness Creation Story 2: creating distinctions by starting with a totally undifferentiated matter having "perfect symmetry" [50] and then, in a ‘big bang,’ start making distinctions, e.g., breaking symmetries, to give form to the matter.

Finally, the literal or objective indefiniteness interpretation of quantum mechanics interprets the mathematics as saying that reality at the quantum level is not the usual notion of "definiteness all the way down." Somehow our mental preconceptions seem to presuppose definiteness all the way down, so we have no natural intuitive ("*anschaulich*") picture of a reality that is objectively indefinite and that can ‘jump’ into a more definite state when a difference is made. Since the mathematics of quantum mechanics speaks rather clearly about the reality of objective indefiniteness, the "*problem* of interpreting QM" seems to be largely *our* problem in developing an intuitive picture of reality at the quantum level.

## 9 Appendix: Transporting vector space structures

It may be useful [21] to rigorously understand the mathematics connecting finite-dimensional QM over  $\mathbb{C}^n$  to QM/Sets over  $\mathbb{Z}_2^n$ . There is a general method to transport some structures from a vector space  $V$  over a field  $\mathbb{K}$  to a vector space  $V'$  over a *different* field  $\mathbb{K}'$ . Select a basis set  $U$  for the source space  $V$  and then consider a structure on  $V$  that can be characterized in terms of the basis set  $U$ . Then apply the free vector space over the field  $\mathbb{K}'$  construction to  $U$  to generate the target vector space  $V'$ . Since the source structure was defined in terms of the basis set  $U$ , it can be carried over or "transported" to  $V'$  via *its* basis set  $U$ .

This method can be stated in rigorous terms using category theory ([43]; [4]). The construction of the free vector space over a field  $\mathbb{K}$  is a functor from the category *Sets* of sets and functions to the category  $Vect_{\mathbb{K}}$  of vector spaces over  $\mathbb{K}$  and linear transformations. The functor will only be used here on finite sets where it takes a finite set  $U$  to the vector space  $\mathbb{K}^U$ . The primary structures being transported are direct-sum decompositions (DSD) of a finite-dimensional vector space  $V$ . A DSD a set  $\{V_i\}$  of disjoint subspaces (i.e., only overlap is zero space) so that the whole space  $V$  is their direct sum, or, in terms of category theory,  $V$  is the coproduct  $V = \oplus V_i$  of the subspaces  $\{V_i\}$ . In the category *Sets*, a set  $\{B_i\}$  of disjoint subsets of a set  $U$  is a set partition of  $U$  if  $\cup B_i = U$ , or, in terms of category theory,  $U$  is the coproduct of the disjoint subsets  $\{B_i\}$ . The free vector space over  $\mathbb{K}$  functor is a left adjoint, "left adjoints preserve colimits" [4, p. 197], and coproducts are a special type of colimit. Hence the free vector space functor carries a set partition  $\pi = \{B_i\}_{i=1,\dots,m}$  to the DSD  $\{V_i = \mathbb{K}^{B_i}\}$  of  $V = \mathbb{K}^U = \oplus \mathbb{K}^{B_i}$ .

Now start with the structure of a DSD  $\{V_i\}$  on  $V \in Vect_{\mathbb{K}}$ . What we previously called "characterizing the structure in terms of a basis set  $U$ " is rigorously interpreted to mean, in this case, finding a basis  $U$  and a partition  $\{B_i\}$  on  $U$  so that the given DSD  $\{V_i\}$  is the image of the free vector space functor, i.e.,  $V = \mathbb{K}^U = \oplus \mathbb{K}^{B_i} = \oplus V_i$ . But then the free vector space functor over a different field  $\mathbb{K}'$  can be applied to the same set partition  $\{B_i\}$  of the set  $U$  to generate a DSD  $\{V'_i = \mathbb{K}'^{B_i}\}$  of  $V' = \mathbb{K}'^U$ . That is how to rigorously describe "transporting" a set-based structure on a vector  $V$  over  $\mathbb{K}$  to a vector space  $V'$  over a different field  $\mathbb{K}'$ .

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

DSD, the union  $U = \cup B_i$  is a basis for  $V$  so  $V_i = \mathbb{K}^{B_i}$  and  $V = \mathbb{K}^U = \oplus \mathbb{K}^{B_i}$ .

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

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

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

Hence outside of characteristic functions, the real-valued functions  $f : U \rightarrow \mathbb{R}$  cannot be internalized as operators on  $\mathbb{Z}_2^U$ . But that is fine since we have the machinery in QM/Sets (including DSDs) that is sufficient to define the probabilities for classical events and superposition events with the outcome set or sample space  $U$  where  $f : U \rightarrow \mathbb{R}$  is a real-valued random variable. We have illustrated the transporting of set-based structures on  $\mathbb{C}^n$  to  $\mathbb{Z}_2^n$  using a basis set  $U$ , but in the stand-alone model QM/Sets, we cut the umbilical cord to  $\mathbb{C}^n$  and work with any other basis  $U'$  of  $\mathbb{Z}_2^n$

and real-valued random variables  $g : U' \rightarrow \mathbb{R}$  on *that* sample space.

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

## References

- [1] Abramsky, Samson and Coecke, Bob 2004. A categorical semantics of quantum protocols. In: *Proceedings of the 19th IEEE Symposium on Logic in Computer Science (LiCS'04)*, IEEE Computer Science Press, 415–425.
- [2] Ainsworth, Thomas. 2020. “Form vs. Matter.” In *The Stanford Encyclopedia of Philosophy (Summer 2020 Edition)*, edited by Edward N. Zalta. <https://plato.stanford.edu/archives/sum2020/entries/form-matter/>.
- [3] Auletta, Gennaro, Mauro Fortunato, and Giorgio Parisi. 2009. *Quantum Mechanics*. Cambridge UK: Cambridge University Press.
- [4] Awodey, Steve. 2006. *Category Theory*. Oxford: Clarendon Press.
- [5] Bell, John S. 1990. Against "Measurement". In *Sixty-Two Years of Uncertainty*. Arthur I. Miller ed., New York: Plenum Press: 17-31.
- [6] Bennett, Charles H. 2003. Quantum Information: Qubits and Quantum Error Correction. *International Journal of Theoretical Physics* 42 (2 February): 153–76.
- [7] Birkhoff, Garrett 1948. *Lattice Theory*. New York: American Mathematical Society.
- [8] Busch, Paul and Greg Jaeger 2010. Unsharp Quantum Reality. *Foundations of Physics*. 40: 1341-1367.

- [9] Castellani, Elena 2003. Symmetry and equivalence. In *Symmetries in Physics: Philosophical Reflections*. Katherine Brading and Elena Castellani eds., Cambridge: Cambridge University Press: 425-436.
- [10] Chen, Jin-Quan, Mei-Juan Gao and Guang-Qun Ma 1985. The representation group and its application to space groups. *Reviews of Modern Physics*. 57 (1): 211-278.
- [11] Chen, Jin-Quan, Jialun Ping and Fan Wang 2002. *Group Representation Theory for Physicists (2nd ed.)*. Singapore: World Scientific.
- [12] Cohen-Tannoudji, Claude, Bernard Diu and Franck Laloë 2005. *Quantum Mechanics Vol. 1*. New York: John Wiley & Sons.
- [13] Collier, John. 1996. Information Originates in Symmetry Breaking. *Symmetry: Science and Culture* 7: 247–56.
- [14] Connes, Alan. 1994. *Non-Commutative Geometry*. New York: Academic Press.
- [15] D’Espagnat, Bernard. 1979. The Quantum Theory and Reality. *Scientific American* 241 (5): 158–81.
- [16] Dirac, Paul A. M. 1958. *The Principles of Quantum Mechanics (4th ed.)*. Oxford: Clarendon.
- [17] Eddington, Arthur S. 1947. *New Pathways in Science (Messenger Lectures 1934)*. Cambridge UK: Cambridge University Press.
- [18] Ellerman, David. 2009. Counting Distinctions: On the Conceptual Foundations of Shannon’s Information Theory. *Synthese* 168 (1 May): 119–49.
- [19] 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.
- [20] Ellerman, David. 2017. Logical Information Theory: New Foundations for Information Theory. *Logic Journal of the IGPL* 25 (5 Oct.): 806–35.
- [21] Ellerman, David. 2017. Quantum Mechanics over Sets: A Pedagogical Model with Non-Commutative finite Probability Theory as Its Quantum Probability Calculus. *Synthese* 194 (12): 4863–96.

- [22] Ellerman, David. 2018. Logical Entropy: Introduction to Classical and Quantum Logical Information Theory. *Entropy* 20 (9): Article ID 679. <https://doi.org/10.3390/e20090679>.
- [23] Ellerman, David. 2021. On Abstraction in Mathematics and Indefiniteness in Quantum Mechanics. *Journal of Philosophical Logic*. <https://doi.org/10.1007/s10992-020-09586-1>.
- [24] Ellerman, David. 2021. The Logical Theory of Canonical Maps: The Elements & Distinctions Analysis of the Morphisms, Duality, Canonicity, and Universal Constructions in Set. *ArXiv.Org*, April. <https://arxiv.org/abs/2104.08583>.
- [25] Ellerman, David. 2021. *New Foundations for Information Theory: Logical Entropy and Shannon Entropy*. Cham, Switzerland: SpringerNature.
- [26] Falkenburg, Brigitte 2010. Language and Reality: Peter Mittelstaedt's Contributions to the Philosophy of Physics. *Foundations of Physics*. 40: 1171-1188.
- [27] Fano, U. 1957. Description of States in Quantum Mechanics by Density Matrix and Operator Techniques. *Reviews of Modern Physics* 29 (1): 74–93.
- [28] Feyerabend, Paul 1983 (orig. 1962). Problems of Microphysics. In *Frontiers of Science and Philosophy*. Robert G. Colodny ed., Lanham MD: University Press of America: 189-283.
- [29] Feynman, Richard P., Robert B. Leighton and Matthew Sands 1965. *The Feynman Lectures on Physics: Quantum Mechanics (Vol. III)*. Reading MA: Addison-Wesley.
- [30] Fine, Arthur 1986. *The Shaky Game: Einstein, Realism, and the Quantum Theory*. Chicago: University of Chicago Press.
- [31] French, Steven and Decio Krause 2003. Quantum Vagueness. *Erkenntnis*. 59: 97-124.
- [32] Grätzer, George 2003. *General Lattice Theory* (2nd ed.). Boston: Birkhäuser Verlag.
- [33] Hanson, Andrew J., Gerardo Ortiz, Amr Sabry, and Yu-Tsung Tai 2013. *Discrete Quantum Theories*. arXiv:1305.3292v1.

- [34] Heisenberg, Werner 1958. *Physics & Philosophy: The Revolution in Modern Science*. New York: Harper Torchbooks.
- [35] Heisenberg, Werner 1961. Planck's discovery and the philosophical problems of atomic physics. In *On Modern Physics*. New York: Clarkson N. Potter Inc.: 3-19.
- [36] Herbut, Fedor 2013. Wavefunction reality, indeterminate properties and degrees of presence. *Studies in History and Philosophy of Modern Physics*. 44: 182-190.
- [37] Hoffman, K., and R. Kunze. 1961. *Linear Algebra*. Englewood Cliffs NJ: Prentice-Hall.
- [38] Jaeger, Gregg. 2007. *Quantum Information: An Overview*. New York: Springer Science+Business Media.
- [39] Kung, Joseph P. S., Gian-Carlo Rota and Catherine H. Yan 2009. *Combinatorics: The Rota Way*. New York: Cambridge University Press.
- [40] Lidl, Rudolf, and Harald Niederreiter. 1986. *Introduction to Finite Fields and Their Applications*. Cambridge UK: Cambridge University Press.
- [41] Lowe, E. J. 1994. Vague Identity and Quantum Indeterminacy. *Analysis*. 54 (2): 110-114.
- [42] Lüders, Gerhart. 1951. Über Die Zustandsänderung Durch Meßprozeß. *Annalen Der Physik* 8 (6): 322–28.
- [43] Mac Lane, Saunders. 1998. *Categories for the Working Mathematician. 2nd Ed.* New York: Springer Science+Business Media.
- [44] Malin, Shimon. 2012. *Nature Loves to Hide: Quantum Physics and the Nature of Reality, a Western Perspective (Revised Ed.)*. Singapore: World Scientific Publishing.
- [45] Manfredi, Giovanni, and M. R. Feix. 2000. Entropy and Wigner Functions. *Physical Review E* 62: 4665–74.
- [46] 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.

- [47] Mittelstaedt, Peter 1998. The Constitution of Objects in Kant's Philosophy and in Modern Physics. In *Interpreting Bodies: Classical and Quantum Objects in Modern Physics*. Elena Castellani ed., Princeton: Princeton University Press: 168-180.
- [48] Muller, Scott J. 2007. *Asymmetry: The Foundation of Information*. Berlin: Springer Verlag.
- [49] Nielsen, Michael and Isaac Chuang 2000. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.
- [50] Pagels, Heinz. 1985. *Perfect Symmetry: The Search for the Beginning of Time*. New York: Simon and Schuster.
- [51] Rota, Gian-Carlo. 2001. Twelve Problems in Probability No One Likes to Bring up. In *Algebraic Combinatorics and Computer Science*, edited by Henry Crapo and Domenico Senato, 57–93. Milano: Springer.
- [52] Schumacher, Benjamin and Michael Westmoreland 2012. Modal Quantum Theory. *Foundations of Physics*, 42, 918-925.
- [53] Shimony, Abner 1988. The reality of the quantum world. *Scientific American*. 258 (1): 46-53.
- [54] Shimony, Abner 1989. Conceptual foundations of quantum mechanics. In *The New Physics*. Paul Davies ed., Cambridge: Cambridge University Press: 373-395.
- [55] Shimony, Abner. 1999. Philosophical and Experimental Perspectives on Quantum Physics. In *Philosophical and Experimental Perspectives on Quantum Physics: Vienna Circle Institute Yearbook 7*, 1–18. Dordrecht: Springer Science+Business Media.
- [56] Spekkens, Robert W. 2007. Evidence for the Epistemic View of Quantum States: A Toy Theory. *Physical Review A* 75 (3): 032110.
- [57] Stairs, Allen 1983. Quantum Logic, Realism, and Value Definiteness. *Philosophy of Science*. 50 (4): 578-602.
- [58] Stanley, Richard P. 1999. *Enumerative Combinatorics Vol. 2*. New York: Cambridge University Press.



- [59] Sternberg, Shlomo 1994. *Group Theory and Physics*. Cambridge: Cambridge University Press.
- [60] Stone, M. H. 1932. On one-parameter unitary groups in Hilbert Space. *Annals of Mathematics*. 33 (3): 643–648.
- [61] Takeuchi, T., L. N. Chang, Z. Lewis, and D. Minic. 2012. *Some Mutant Forms of Quantum Mechanics*. [quant-Ph] arXiv:1208.5544v1.
- [62] Tamir, Boaz, and Eliahu Cohen. 2014. Logical Entropy for Quantum States. *ArXiv.Org*, December. <http://de.arxiv.org/abs/1412.0616v2>.
- [63] Tamir, Boaz, and Eliahu Cohen. 2015. A Holevo-Type Bound for a Hilbert Schmidt Distance Measure. *Journal of Quantum Information Science* 5: 127–33. doi:10.4236/jqis.2015.54015.
- [64] Voiculescu, Dan, Kenneth Dykema, and Alexandru Nica. 1992. *Free Random Variables. A Noncommutative Probability Approach to Free Products with Applications to Random Matrices, Operator Algebras and Harmonic Analysis on Free Groups*, Providence, RI: American Mathematical Society.
- [65] von Neumann, John 1955. *Mathematical Foundations of Quantum Mechanics*. Robert T. Beyer trans., Princeton: Princeton University Press.
- [66] Weinberg, Steven 1994. *Dreams of a Final Theory*. New York: Vintage Books.
- [67] Weyl, Hermann 1949. *Philosophy of Mathematics and Natural Science*. Princeton NJ: Princeton University Press.
- [68] Weyl, Hermann. 1950. *The Theory of Groups and Quantum Mechanics*. Translated by H. P. Robertson. New York: Dover Publications.
- [69] Wheeler, John A. 1999. Information, Physics, Quantum: The Search for Links. In *Feynman on Computation*, edited by Anthony J. G. Hey, 309–36. Reading MA: Perseus Books.
- [70] Zurek, Wojciech Hubert. 2003. Decoherence, Einselection, and the Quantum Origins of the Classical. *Review of Modern Physics* 75 (3 July): 715–75.