

Follow the Math!
The mathematics of quantum mechanics as the mathematics
of set partitions linearized to (Hilbert) vector spaces.

David Ellerman
University of Ljubljana, Slovenia
orcid.org/0000-0002-5718-618X
david@ellerman.org

September 6, 2022

Abstract

The purpose of this paper is to show that the mathematics of quantum mechanics (QM) is the mathematics of set partitions (which specify indefiniteness and definiteness) linearized to vector spaces, particularly in Hilbert spaces. The key analytical concepts are definiteness versus indefiniteness, distinctions versus indistinctions, and distinguishability versus indistinguishability. The key machinery to go from indefinite to more definite states is the partition join operation at the set level that prefigures at the quantum level projective measurement as well as the formation of maximally-definite state descriptions by Dirac’s Complete Sets of Commuting Operators (CSCOs). The mathematics of partitions is first developed in the context of sets and then linearized to vector spaces where it is shown to provide the mathematical framework for quantum mechanics. This development is measured quantitatively by logical entropy at the set level and by quantum logical entropy at the quantum level. This follow-the-math approach supports the Literal Interpretation of QM—as advocated by Abner Shimony among others which sees a reality of objective indefiniteness that is quite different from the common sense and classical view of reality as being “definite all the way down.”

Keywords: partitions, direct-sum-decompositions, partition join, objective indefiniteness, definite-all-the-way-down..

Contents

1	Introduction	2
2	Partitions: The logical concept to describe indefiniteness and definiteness	3
3	Intuitive imagery for superposition	5
4	Logical entropy: the measure of distinctions	7
5	Formulation using density matrices	8
6	The Yoga of Linearization: From Sets to Vector Spaces	11
7	Measurement with Quantum Logical Entropy	14

8 Other quantum mathematics	18
8.1 Commuting and non-commuting observables	18
8.2 The two types of quantum processes	20
8.3 Measurement and the collapse postulate	20
8.4 Indistinguishability of particles	23
8.5 Group representation theory	25
9 Concluding remarks	29

1 Introduction

After a century of proliferating interpretations of quantum mechanics (QM), it is time to ask the simple question: “Where does the math of quantum mechanics come from?”¹ The purpose of this paper is to show that the mathematics of quantum mechanics comes from the mathematics of set partitions (the mathematical tools to describe indefiniteness and definiteness) linearized to vector spaces, particularly Hilbert spaces.

Classical physics exemplified the common-sense idea that reality had definite properties “all the way down.” At the logical level, i.e., Boolean subset logic, each element in the Boolean universe set is either definitely in or not in a subset, i.e., each element either definitely has or does not have a property. Each element is characterized by a full set of properties, a view that might be referred to as “definite all the way down.” This view of reality was expressed by Leibniz’s principle of the identity of indiscernible. If one could always dig deeper into a definite reality to find attributes to distinguish entities, then entities that were completely indistinguishable would logically have to be identical. However, if there is no digging deeper to find distinctions, then any remaining indefiniteness is objective.

It is now rather widely accepted that this common-sense always-definite view of reality is not compatible with quantum mechanics (QM). 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 objectively “neither definitely here nor there.” [48, p. 144] Paul Feyerabend asserted that “inherent indefiniteness is a universal and objective property of matter.” [24, p. 202] This is not an epistemic or subjective indefiniteness of location; it is an ontological or objective indefiniteness. The indefiniteness or indistinguishability cannot be resolved by digging deeper with more precision. The notion of *objective indefiniteness* in QM has been most emphasized by Abner Shimony.

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.[40, p. 47]

Since our natural common sense view of the world is “definite all the way down” (as in classical physics), how can we describe an indefinite actuality? The basic mathematical, indeed logical, concept that describes indefiniteness and definiteness is the notion of a *partition on a set*. It is the purpose of this paper to show that the mathematics of quantum mechanics is essentially the mathematics of partitions linearized to (Hilbert) vector spaces. This substantiates that key analytical concepts in QM are indefiniteness and definiteness, indistinction and distinction, and indistinguishability and distinguishability. Key machinery of QM such as projective measurement and the specification of

¹The slogan “Follow the money” means that finding the source of an organization’s or person’s money may reveal their true nature. In a similar sense, we use the slogan “Follow the math!” to mean that finding “where the mathematics of QM comes from” reveals a good deal about the key concepts and machinery of the theory.

maximally definite states by Dirac’s Complete Sets of Commuting Operators (CSCOs) will also be seen as the linearization of the corresponding machinery in the logic of partitions.

2 Partitions: The logical concept to describe indefiniteness and definiteness

Given a universe set $U = \{u_1, \dots, u_n\}$, a *partition* $\pi = \{B_1, \dots, B_m\}$ is a set of subsets $B_j \subseteq U$ (for $j = 1, \dots, m$) called *blocks* that are disjoint and whose union is U .² A *distinction* (or *dit*) of a partition π is an ordered pair of elements $(u_i, u_k) \in U \times U$ in different blocks of π , and $\text{dit}(\pi) \subseteq U \times U$ is the set of all distinctions, called the *ditset* of π . An *indistinction* (or *indit*) of π is an ordered pair of elements in the same block of U , and the set of all indistinctions $\text{indit}(\pi) \subseteq U \times U$, called an *indit set*, is the equivalence relation associated with π where the blocks are the equivalence classes. The ditset and indit set of a partition are complements, i.e., they are disjoint and their union is $U \times U$.

Each block $B_j \in \pi$ of a partition should be thought of as being indefinite or indistinct between its elements $u_i, u_k \in B_j$. Partitions naturally arise as the inverse-images $f^{-1} = \{f^{-1}(y)\}_{y \in f(U)}$ of functions $f : U \rightarrow Y$. In particular, a *numerical attribute* is a function $f : U \rightarrow \mathbb{R}$ into some set of values which we can take as the real numbers \mathbb{R} . Each block $f^{-1}(r)$ in the partition f^{-1} then represents the constant set of all elements $u_i \in U$ taking the value $f(u_i) = r \in \mathbb{R}$. When the set U is taken as the outcome set or sample space of a finite probability distribution [with equiprobable points or point probabilities $p_i = \Pr(u_i)$], then the numerical attribute is a random variable. As an aid to intuition, these simple concepts at the logical level might be seen as the elementary forms of the more developed mathematical concepts of quantum mechanics as illustrated in Table 1. These connections will be further developed in the later section on the Yoga of Linearization.

Logical concept	QM concept
Block of a partition	Eigenspace
Elements in a block	Eigenvectors with same eigenvalue
Numerical attribute	Observable (self-adjoint operator)
Elements of U	Basis eigenvectors of an observable
Values of attribute	Eigenvalues of an observable
Partition on a set	Direct-sum decomposition of eigenspaces

Table 1: Partition logical precursors for QM math concepts

Subsets of a set and partitions on a set are mathematically dual concepts [31]. The Boolean logic of subsets (usually presented in the special case of “propositional logic”) thus has a dual mathematical logic, the logic of partitions [15]. The “logical” concepts that prefigure the mathematics of QM are those of the logic of partitions.

In the Boolean logic of subsets, the powerset $\wp(U)$ (all subsets of U) forms a lattice where the partial order is set inclusion, the join (least upper bound) and meet (greatest lower bound) are union and intersection respectively, and the top and bottom of the lattice are the universe set U and the empty set \emptyset respectively.

In the dual logic of partitions, the set $\Pi(U)$ of partitions on U also forms a lattice where the partial order is refinement. Given another partition $\sigma = \{C_1, \dots, C_{m'}\}$ on U , the partition σ is *refined by* π , written, $\sigma \lesssim \pi$, if for every block of π , there is a block of σ containing it. Intuitively, the blocks of π can be obtained by chopping up the blocks of σ . If π and σ are the inverse images of random variables $f : U \rightarrow \mathbb{R}$ and $g : U \rightarrow \mathbb{R}$ respectively, the $\sigma \lesssim \pi$ means that the random variable f is *sufficient* for g , i.e., the value of f determines the value of g .

²Since our purpose is conceptual clarity, not mathematical generality, we will stick to the finite sets and dimensions throughout.

The *join* $\pi \vee \sigma$ (least upper bound in the refinement ordering) is the partition of U whose blocks are all the non-empty intersections $B_j \cap C_{j'}$. To form the *meet* $\pi \wedge \sigma$ (greatest lower bound in the refinement ordering), think of two intersecting blocks B_j and $C_{j'}$ as two overlapping blobs of mercury that unify to make a larger blob. Doing this for all overlapping blocks, the blocks of the meet are the subsets of U that are a union of certain blocks of π and simultaneously a union of certain blocks of σ (and are minimal in that respect). The top of the lattice of partitions $\Pi(U)$ is the maximally distinguished *discrete partition* $\mathbf{1}_U = \{\{u_1\}, \dots, \{u_n\}\}$ whose blocks are all the singletons of the elements of U and the bottom is the minimally distinguished *indiscrete partition* $\mathbf{0}_U = \{U\}$ (nicknamed the “Blob”) which blobs all the elements together into one indefinite “superposition.”³ The join operation is the only one we will need as it prefigures a projective measurement in quantum mechanics.

Underlying the duality between subsets (e.g., images of functions) and partitions (inverse-images of functions) is the duality between elements of subsets and distinctions of a partitions, the ‘its and dits’ duality. In the Boolean lattice $\wp(U)$ of subsets, the partial order $S \subseteq T$ for $S, T \in \wp(U)$ is the inclusion of elements. In the lattice of partitions $\Pi(U)$, the refinement partial order $\sigma \lesssim \pi$ is just the inclusion of distinctions, i.e., $\sigma \lesssim \pi$ if and only if (iff) $\text{dit}(\sigma) \subseteq \text{dit}(\pi)$. Moreover, wherever the logical partial order holds, there is an induced logical map. If $S \subseteq T$, then there is the canonical injection $S \hookrightarrow T$, and if $\sigma \lesssim \pi$, then there is the canonical surjection $\pi \twoheadrightarrow \sigma$ that carries each block $B_j \in \pi$ to the unique block $C_k \in \sigma$ such that $B_j \subseteq C_k$.⁴ The top $\mathbf{1}_U$ of the partition lattice includes all possible distinctions, i.e., $\text{dit}(\mathbf{1}_U) = U \times U - \Delta$ [where Δ is the diagonal of self-pairs (u_i, u_i)], just as the top U of the the subset lattice thus includes all possible elements. The bottom $\mathbf{0}_U$ of the partition lattice has no distinctions, i.e., $\text{dit}(\mathbf{0}_U) = \emptyset$, just as the bottom \emptyset of the subset lattice has no elements. The ditset of the join in the partition lattice is the union of the distinctions just as the join in the lattice of subsets is union of the elements of the subsets. This duality between elements and distinctions is illustrated in Table 2.

Dualities	Subset logic	Partition logic
Its or dits	Elements u of S	Distinctions (u, u') of π
Partial order	Inclusion $S \subseteq T$	$\sigma \lesssim \pi$, $\text{dit}(\sigma) \subseteq \text{dit}(\pi)$
Logical maps	$S \hookrightarrow T$	$\pi \twoheadrightarrow \sigma$
All	All elements U	All distinctions $\mathbf{1}_U$
None	No elements \emptyset	No distinctions $\mathbf{0}_U$
Join	$S \cup T$	$\text{dit}(\pi \vee \sigma) = \text{dit}(\pi) \cup \text{dit}(\sigma)$

Table 2: Its & Dits duality

Following Heisenberg,⁵ we might express this duality by going back to the ancient Greek metaphysical notions of substance (or matter) and form [1]. At this simple level, one can still discern two ‘creation stories’ corresponding to the classical (definite all the way down) and the quantum (objective indefiniteness) versions. These two stories can be represented by moving from the bottom up the two logical lattices illustrated in Figure 1 where the universe now consists three states $U = \{a, b, c\}$.

Classical creation story: In the Beginning was the Void (no substance) and then fully definite elements (“Its”) were created until the universe U was created.

Quantum creation story: In the Beginning was the Blob—all substance (energy) with no form—and then, in a Big Bang, distinctions (“Dits”) were created (e.g., symmetries were broken) as the

³Many of the older texts [5] presented the “lattice of partitions” upside down, i.e., with the opposite partial order, so the join and meet as well as the top and bottom were interchanged.

⁴In [18], it is shown that the maps in the category of Sets that are considered “canonical” are the maps induced from the given data by these logical maps in the two logical lattices that play symmetrical mathematical roles.

⁵In his sympathetic interpretation of Aristotle’s treatment of substance and form, Heisenberg refers to the substance as: “a kind of indefinite corporeal substratum, embodying the possibility of passing over into actuality by means of the form.” [27, p. 148] Heisenberg’s “potentiality” “passing over into actuality by means of the form” should be seen as the actual indefinite “passing over into” the actual definite by being objectively in-formed through the making of distinctions.

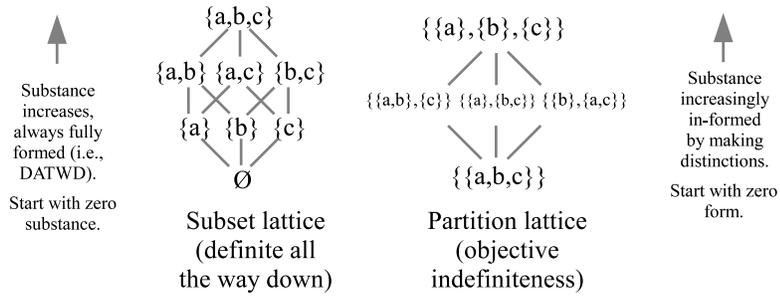


Figure 1: The two creation stories illustrated by the dual lattices

substance was increasingly in-formed to reach the universe U .

3 Intuitive imagery for superposition

How should one imagine a quantum superposition? The most misleading imagery in QM is the *classical* interpretation of superposition (Figure 2) as the addition of two definite waves to get another *definite* wave.

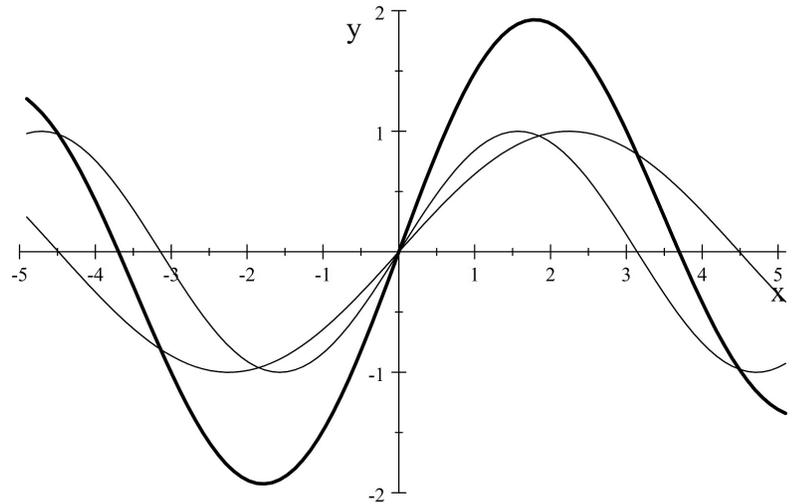


Figure 2: Imagery for *classical* superposition

But there seem to be no actual physical waves in QM, much to the disappointment of Schrödinger and others. The complex numbers are the natural mathematics to describe waves so the misleading wave formalism is always there.

Such analogies have led to the name ‘Wave Mechanics’ being sometimes given to quantum mechanics. It is important to remember, however, that the superposition that occurs in quantum mechanics is of an essentially different nature from any occurring in the classical theory, as is shown by the fact that the quantum superposition principle demands indeterminacy in the results of observations in order to be capable of a sensible physical interpretation. The analogies are thus liable to be misleading. [11, p. 14]

The complex numbers are needed in the mathematics of QM (among other reasons) since they are the algebraically-complete extension of the reals so the real-valued observables will have a full set of eigenvectors, *not* because there are any physical waves.

The quantum (as opposed to classical) interpretation of superposition is the addition of two definite states to get a new state *indefinite between the definite states*—not the ‘double-exposure’ image (e.g., not being simultaneously *here* and *there* as in so much of the popular science literature) suggested by the wave interpretation.⁶ In Figure 3, the superposition of the two definite isosceles triangles is the indefinite triangle which is indefinite on where the two definite triangles are distinct (the labeling of the equal sides) and is definite on where the two triangles do not differ (the aA -axis).⁷

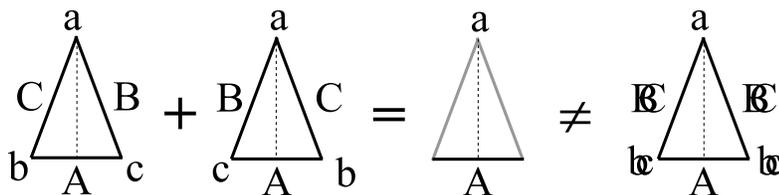


Figure 3: Imagery for *Quantum* Superposition

The intuitions in the theater of our minds have evolved to think in terms of a macroscopic spatial world, so one should not expect to have fully definite classical imagery for indefinite quantum states. The best to expect is probably a set of image ‘crutches’ to illustrate one aspect or another as in Figure 3.

At the simplest logical level, the pure and mixed quantum states for a single particle can be illustrated by partitions as in Figure 4. There are four possible eigenstates for a particle represented by a, b, c , and d . The pure state of all those eigenstates superposed is represented by the indiscrete partition $\{\{a, b, c, d\}\}$ (written in shorthand as $\{abcd\}$) and then distinctions are made (i.e., ‘measurements’ are made) to get the other partition representations of the mixed states of ‘orthogonal’ (disjoint) superpositions such as $\{\{a, c\}, \{b, d\}\}$ (or in shorthand $\{ac, bd\}$).

In terms of sets (as in Figure 4), the superposition of the eigenstates $\{a\}$ and $\{c\}$ is the state $\{a, c\}$, which is the state indefinite between $\{a\}$ and $\{c\}$; not definitely $\{a\}$ and not definitely $\{c\}$, but definitely not $\{b\}$ or $\{d\}$. A distinction between $\{a\}$ and $\{c\}$, e.g., in the join $\{ac, bd\} \vee \{a.bcd\} =$

⁶The indefiniteness interpretation of qubit $a|0\rangle + b|1\rangle$ is more routine in quantum information and computation theory [34] as opposed to the being simultaneously $|0\rangle$ and $|1\rangle$ version of superposition.

⁷There is at least an analogy between superposition in QM and abstraction in mathematics. In Frege’s example of a set of parallel directed line segments oriented in the same way, the abstraction “direction” is definite on what is common between the lines and indefinite on how they differ [19]. In QM, the emphasis is on the indefiniteness between the definite eigenstates (glass half-empty) while in abstraction, the emphasis is on the common definiteness between the instances (glass half-full).

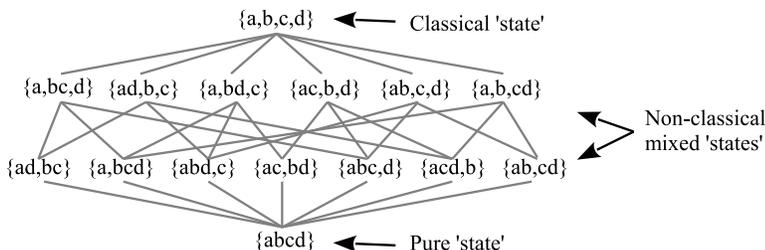


Figure 4: Lattice of partitions on 4 elements with labelled states

$\{a, bd, c\}$, would reduce the superposition $\{a, c\}$ to a mixture of the eigenstates $\{a\}$ and $\{c\}$, e.g., in the mixed state $\{a, bd, c\}$. In contrast to the misleading superposition of waves imagery, these simple examples illustrate that the superposition $\{a, c\}$ is not a definite state but is an indefinite state that could reduce (e.g., with distinctions supplied by $\{a.bcd\}$) to a mixture of the definite eigenstates $\{a\}$ or $\{c\}$.

4 Logical entropy: the measure of distinctions

Since partitions are the mathematical concept to represent distinctions and indistinctions (or definiteness and indefiniteness), there should be a quantitative measure (in the sense of measure theory) to quantify distinctions. Since U is a finite set and the set of distinctions of a partition π is finite, the obvious notion to measure distinctions is simply the cardinality of the set of distinctions $\text{dit}(\pi) \subseteq U \times U$ normalized by the size of $U \times U$. Hence the *logical entropy of a partition* $\pi = \{B_1, \dots, B_m\}$ for equiprobable points in U is:

$$h(\pi) = \frac{|\text{dit}(\pi)|}{|U \times U|} = \frac{|U \times U| - |\cup_j B_j \times B_j|}{|U \times U|} = 1 - \sum_j \frac{|B_j|^2}{|U|^2} = 1 - \sum_j \text{Pr}(B_j)^2$$

where $\text{Pr}(B_j) = |B_j|/|U|$ is the probability that a random draw from U gives an element of B_j ([14], [20], [21]). When the points of U have the probabilities p_i for $i = 1, \dots, n$, then:

$$h(\pi) = 1 - \sum_j \text{Pr}(B_j)^2$$

where $\text{Pr}(B_j) = \sum_{u_i \in B_j} p_i$. The logical entropy of π has an immediate interpretation; it is the probability that in two independent random draws from U , one will obtain a distinction of π —just as the probability of a subset $S \subseteq U$, $\text{Pr}(S)$ is the probability that in one random draw from U , one will obtain an element of S . Since the indiscrete partition makes no distinctions, its logical entropy is zero, $h(\mathbf{0}_U) = 0$. The discrete partition makes all possible distinctions so its logical entropy is $h(\mathbf{1}_U) = 1 - \sum_{i=1}^n p_i^2$, which, in the equiprobable case, is $1 - \frac{1}{n}$, the probability that the second draw is not the same as the first draw.

This definition of logical entropy fulfills a program of Gian-Carlo Rota that begins with the idea: “The lattice of partitions plays for information the role that the Boolean algebra of subsets plays for size or probability.” [30, p. 30] In Rota’s Fubini Lectures (and in his lectures as MIT), he argued that since partitions are dual to subsets, then quantitatively, information is to partitions as probability is to subsets:

$$\frac{\text{Information}}{\text{Partitions}} \approx \frac{\text{Probability}}{\text{Subsets}}.$$

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 “which will capture some property that will turn out to be for [partitions] what size is to a set.” He then asks:

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? [38, p. 67]

The underlying duality of elements and distinctions answers that question. The lattice of partitions is isomorphic to the lattice of ditsets partially ordered by inclusion (since refinement is just inclusion of ditsets), and the normalized size of subsets and ditsets (equiprobable case) gives the notions of probability $\text{Pr}(S) = \frac{|S|}{|U|}$ and logical entropy $h(\pi) = \frac{|\text{dit}(\pi)|}{|U \times U|}$ —as summarized in Table 3.

	Logical Probability Theory	Logical Information Theory
‘Outcomes’	Elements $u \in U$ finite	Dits $(u, u') \in U \times U$ finite
‘Events’	Subsets $S \subseteq U$	Ditsets $\text{dit}(\pi) \subseteq U \times U$
Equiprobable points	$\Pr(S) = \frac{ S }{ U }$	$h(\pi) = \frac{ \text{dit}(\pi) }{ U \times U }$
Point probabilities	$\Pr(S) = \sum \{p_j : u_j \in S\}$	$h(\pi) = \sum \{p_j p_k : (u_j, u_k) \in \text{dit}(\pi)\}$
Interpretation	$\Pr(S)$ = one-draw prob. of S -element	$h(\pi)$ = two-draw prob. of π -distinction

Table 3: Classical logical probability theory and ‘classical’ logical information theory.

When the point probabilities on U are given by the probability distribution $p = (p_1, \dots, p_n)$, then the logical entropy $h(\pi)$ is the product probability measure $p \times p$ (defined on $U \times U$) of the ditset $\text{dit}(\pi) \subseteq U \times U$. Logical entropy is the measure of information-as-distinctions. Since the logical entropy is the value of a measure in the sense of measure theory (unlike Shannon entropy [20]), the compound notions of logical entropy are naturally defined in the usual Venn diagram manner as illustrated in Figure 5 which includes the conditional logical entropy $h(\sigma|\pi)$ (the measure of the distinctions in σ that were not in π) and the mutual logical information $m(\pi, \sigma)$ (the measure of the distinctions common to π and σ).

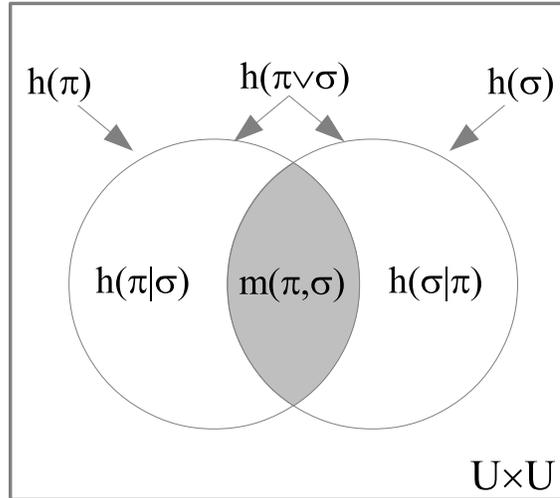


Figure 5: $h(\pi \vee \sigma) = h(\pi) + h(\sigma) - m(\pi, \sigma) = h(\pi|\sigma) + h(\sigma|\pi) + m(\pi, \sigma)$

The compound notion of logical entropy that we will make later use of in the analysis of quantum measurement is the logical entropy $h(\pi \vee \sigma)$ of the join $\pi \vee \sigma$ which is the probability measure $p \times p$ on $\text{dit}(\pi \vee \sigma) = \text{dit}(\pi) \cup \text{dit}(\sigma)$.

5 Formulation using density matrices

In quantum mechanics, the state of a system can be represented by state vectors or by density matrices [34, p. 102]. The best form for our purposes is density matrices because the relevant machinery developed above about partitions and logical entropy can be reformulated using ‘classical’ density matrices over the reals.

Given a partition $\pi = \{B_1, \dots, B_m\}$ on $U = \{u_1, \dots, u_n\}$ with point probabilities $p = (p_1, \dots, p_n)$, an $n \times n$ density matrix $\rho(B_j)$ can be defined for each block $B_j \in \pi$ as follows:

$$\rho(B_j)_{ik} = \begin{cases} \frac{\sqrt{p_i p_k}}{\Pr(B_j)} & \text{if } (u_i, u_k) \in B_j \times B_j \\ 0 & \text{otherwise} \end{cases} .$$

where $\Pr(B_j) = \sum_{u_i \in B_j} p_i$. Then these density matrices for the blocks are combined to form the density matrix $\rho(\pi)$ representing the partition π :

$$\rho(\pi) = \sum_{j=1}^m \Pr(B_j) \rho(B_j)$$

so the entries are:

$$\rho(\pi)_{ik} = \begin{cases} \sqrt{p_i p_k} & \text{if } (u_i, u_k) \in \text{indit}(\pi) \\ 0 & \text{otherwise} \end{cases} .$$

These density matrices over the reals are symmetric and have trace (sum of diagonal elements) equal to 1 since the diagonal elements are $\sqrt{p_i p_i} = p_i$ for $i = 1, \dots, n$. The probability p_i of an element u_i is recovered as $\text{tr}[P_{u_i} \rho(\pi)] = \sqrt{p_i} \sqrt{p_i} = p_i$ [where P_{u_i} is the diagonal projection matrix with entries $\chi_{\{u_i\}}()$] which is the set version of the Born Rule. Assuming only non-zero probabilities, the non-zero off-diagonal elements indicate the indistinctions of π where elements u_i and u_k ‘cohere’ together in the same block of the partition π and are called “coherences” in the case of quantum density matrices ([10, p. 303]; [2, p. 177]). Thus at the logical level, indistinctions prefigure quantum coherences.

In the formula for logical entropy $h(\pi) = 1 - \sum_j \Pr(B_j)^2$, the density matrix replaces the block probabilities and the trace replaces the summation to give the same result:

$$h(\pi) = 1 - \sum_j \Pr(B_j)^2 = 1 - \text{tr}[\rho(\pi)^2].$$

These real-valued density matrices encapsulate the ‘classical’ treatment of the mathematics of partitions that prefigures the quantum treatment. Two of these classical results translate directly into the corresponding results in quantum mechanics.

The first result is that the projective measurement in QM is classically just the join of partitions. We start with the partition π expressed by the density matrix $\rho(\pi)$ and then we think of the partition $\sigma = \{C_1, \dots, C_m\}$ on U as being the inverse-image of a numerical attribute or ‘observable’ $g : U \rightarrow \mathbb{R}$. In QM, the effect of a projective measurement on a density matrix ρ is given by the *Lüders mixture operation* ([2, p. 279]; [33]). For each block $C_j \in \sigma$, let P_{C_j} be the projection matrix that is a diagonal matrix with the diagonal elements given by the characteristic function $\chi_{C_j} : U \rightarrow 2 = \{0, 1\}$ of C_j . Then the Lüders mixture operation transforms the density matrix $\rho(\pi)$ into the density matrix $\hat{\rho}(\pi)$ according to the formula:

$$\hat{\rho}(\pi) = \sum_{C_j \in \sigma} P_{C_j} \rho(\pi) P_{C_j} .$$

Classical Lüders Mixture Operation

Theorem 1 $\hat{\rho}(\pi) = \rho(\pi \vee \sigma)$.

Proof: A nonzero entry in $\rho(\pi)$ has the form $\rho(\pi)_{ik} = \sqrt{p_i p_k}$ iff there is some block $B \in \pi$ such that $(u_i, u_k) \in B \times B$, i.e., if $u_i, u_k \in B$. The matrix operation $P_{C_j} \rho(\pi)$ will preserve the entry $\sqrt{p_i p_k}$ if $u_i \in C_j$, otherwise the entry is zeroed. And if the entry was preserved, then the further matrix operation $(P_{C_j} \rho(\pi)) P_{C_j}$ will preserve the entry $\sqrt{p_i p_k}$ if $u_k \in C_j$, otherwise it is zeroed. Hence the entries $\sqrt{p_i p_k}$ in $\rho(\pi)$ that are preserved in $P_{C_j} \rho(\pi) P_{C_j}$ are the entries where both $u_i, u_k \in B$ for some $B \in \pi$ and $u_i, u_k \in C_j$. Recall that $\text{dit}(\pi \vee \sigma) = \text{dit}(\pi) \cup \text{dit}(\sigma)$ so $\text{indit}(\pi \vee \sigma) = \text{indit}(\pi) \cap \text{indit}(\sigma)$ —since the join of partitions is just the partition corresponding to the equivalence relation resulting from intersecting two equivalence relations (indit sets). These are the entries in $\rho(\pi \vee \sigma)$ corresponding to the blocks $B \cap C_j$ for some $B \in \pi$, so summing over $C_j \in \sigma$ gives the result: $\sum_{C_j \in \sigma} P_{C_j} \rho(\pi) P_{C_j} = \hat{\rho}(\pi) = \rho(\pi \vee \sigma)$. \square

Our theme is that the vector space mathematics of QM is prefigured at the logical level by the mathematics of partitions on sets. The above Theorem shows that the standard partition operation of join is essentially the set version of the projective measurement operation in QM. Note that partitions have two separate roles in this set-based example; π represents the state being measured and σ represents the numerical attribute (or observable) being measured on that state. The join operation creates more distinctions since $\text{dit}(\pi \vee \sigma) = \text{dit}(\pi) \cup \text{dit}(\sigma)$. The off-diagonal non-zero entries in the density matrices represent indistinctions, so the distinctions that are created by joining σ with π will be indicated by those non-zero entries in $\rho(\pi)$ that are zeroed in $\hat{\rho}(\pi) = \rho(\pi \vee \sigma)$. Logical entropy measures information-as-distinctions so the non-zero off-diagonal entries that are zeroed, the indistinctions that become distinctions (i.e., the coherences that are decohered in the quantum case), will be measured by the increase in logical entropy.

Corollary 1 *The sum of all the squares $p_i p_k$ of all the entries $\sqrt{p_i p_k}$ that were zeroed in the Lüders mixture operation that transforms $\rho(\pi)$ into $\hat{\rho}(\pi) = \sum_{C_j \in \sigma} P_{C_j} \rho(\pi) P_{C_j} = \rho(\pi \vee \sigma)$ is $h(\pi \vee \sigma) - h(\pi) = h(\sigma|\pi)$.*

Proof: All the entries $\sqrt{p_i p_k}$ that got zeroed were for ordered pairs (u_i, u_k) that were indits of π but not indits of $\pi \vee \sigma$, i.e., $(u_i, u_k) \in \text{indit}(\pi) \cap \text{indit}(\pi \vee \sigma)^c = \text{dit}(\pi)^c \cap \text{dit}(\pi \vee \sigma) = \text{dit}(\pi \vee \sigma) - \text{dit}(\pi)$. The sum of products $p_i p_k$ for those pairs (u_i, u_k) is just the product probability measure on that set $\text{dit}(\pi \vee \sigma) - \text{dit}(\pi)$ which is $h(\pi \vee \sigma|\pi)$. And since $\text{dit}(\pi) \subseteq \text{dit}(\pi \vee \sigma)$, the measure on $\text{dit}(\pi \vee \sigma) - \text{dit}(\pi)$ is $h(\pi \vee \sigma|\pi) = h(\pi \vee \sigma) - h(\pi) = h(\sigma|\pi)$ (see Figure 5) which is the information-as-distinctions that σ added to the information in π . \square

Example: If the four elements of $U = \{a, b, c, d\}$ were equiprobable, the real-valued density matrix of the partition $\{abc, d\}$ is:

$$\rho(\{abc, d\}) = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{bmatrix}.$$

The main partition operation representing going from an indefinite state or partition to a more definite (i.e., more refined) one is the *join* operation, for instance: $\{ac, bd\} \vee \{abc, d\} = \{ac, b, d\}$ as in Figure 6:

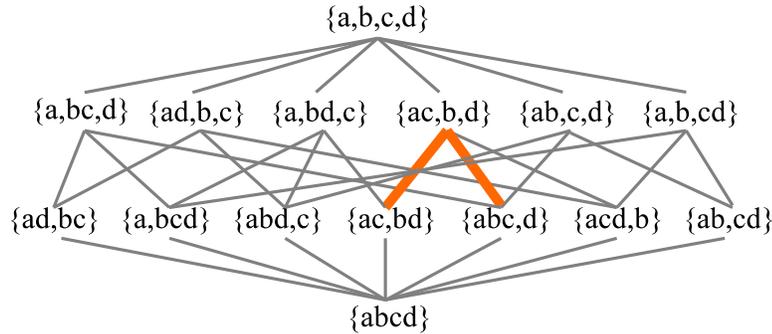


Figure 6: Partition join: $\{ac, bd\} \vee \{abc, d\} = \{ac, b, d\}$

The QM-version of the partition join is a projective measurement described by the Lüders mixture operation. Since $\{ac, bd\}$ is being joined to $\{abc, d\}$, we need the projection matrices to $\{a, c\}$ and to $\{b, d\}$ which are:

$$P_{\{a,c\}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } P_{\{b,d\}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The Lüders mixture operation pre- and post-multiplies the pre-measurement density matrix $\rho(\{abc, d\})$ by these two projection matrices and the result sums to yield the post-measurement density matrix $\hat{\rho}$:

$$\begin{aligned} \hat{\rho} &= P_{\{a,c\}}\rho(\{abc, d\})P_{\{a,c\}} + P_{\{b,d\}}\rho(\{abc, d\})P_{\{b,d\}} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{bmatrix} \\ &\hat{\rho} = \rho(\{ac, bd\} \vee \{abc, d\}) = \rho(\{ac, b, d\}). \end{aligned}$$

The logical entropy of $\rho(\{abc, d\})$ is

$$h(\rho(\{abc, d\})) = 1 - \text{tr} \left[\rho(\{abc, d\})^2 \right] = 1 - \frac{10}{16} = \frac{3}{8}$$

and the logical entropy of $\hat{\rho} = \rho(\{ac, b, d\})$ is

$$h(\rho(\{ac, b, d\})) = 1 - \text{tr} \left[\rho(\{ac, b, d\})^2 \right] = 1 - \frac{3}{8} = \frac{5}{8}.$$

In the transition from $\rho(\{abc, d\})$ to $\hat{\rho} = \rho(\{ac, b, d\})$, there were four entries of $\frac{1}{4}$ that were zeroed, so the sum of their squares is: $4 \times \frac{1}{16} = \frac{1}{4}$ which by the Corollary equals the increase in logical entropy; $\frac{5}{8} - \frac{3}{8} = \frac{1}{4}$.

Repeated joins, i.e., repeated intersections of blocks of different partitions, may eventually reach the discrete partition $\mathbf{1}_U$ whose density matrix is diagonal having no non-zero off-diagonal elements, i.e., all possible distinctions have been made (like a completely decomposed mixed state in QM). A set of partitions on U whose join is the discrete partition is said to be *complete*, and it is the partition-logical analogue of Dirac's complete set of commuting observables (CSCO) [11].

6 The Yoga of Linearization: From Sets to Vector Spaces

Our thesis is that the mathematics of QM is essentially the mathematics of partitions linearized to (Hilbert) vector spaces. There is a semi-algorithmic method—part of the folklore of mathematics (see Weyl's use of it below) to linearize concepts using sets (e.g., partitions) to the corresponding concepts over vector spaces—a method that Gian-Carlo Rota might call a “yoga” [37, p. 251]. The idea is based on taking the vector space concept corresponding to the notion of a set as a basis set of the space. Then the yoga is:

For any given set-concept, apply it to a basis set
and whatever is linearly generated is the corresponding vector space concept.

The Yoga of Linearization.

In applying the Yoga, we take U as being first a set and then a basis set of a vector space over a field \mathbb{k} .⁸ For instance, the set concept of a subset S when applied to a basis set generates a subspace $[S]$. The cardinality of the U gives the dimension of the space $[U]$ generated by the basis set U as shown in Table 4.

Set concept	Vector-space concept
Universe set U	U basis set of a space $[U] = V$ over \mathbb{k}
Cardinality of the set U	Dimension of the space V
Subset S of the set U	Subspace $[S]$ of the space V

Table 4: Some initial applications of the Yoga.

In particular, a singleton subset (representing an eigenstate as in Figure 4) generates a one-dimensional ray. The previous set example of joining the definite states $\{a\}$ and $\{c\}$ to form the indefinite superposition $\{a, c\}$ would linearize to the superpositions that we might represent as $k|a\rangle + k'|c\rangle$ for $k, k' \in \mathbb{k}$. It should also be noted that the treatment of a basis element as definite and a linear combination of basis vectors as indefinite is always a description relative to that basis. If the vector space V has inner products like a Hilbert space, then we will assume U is an orthonormal basis.

The Yoga may sometimes require a choice. Consider the set concept of a numerical attribute $f : U \rightarrow \mathbb{k}$ taking values in a field \mathbb{k} . Taking U as a basis set for a vector space V over \mathbb{k} , it defines both a linear functional $\hat{f} : V \rightarrow \mathbb{k}$ and a diagonalizable linear operator: $F : V \rightarrow V$ generated by $Fu_i = f(u_i)u_i$ on the basis vectors $u_i \in U$. The numerical attribute $f : U \rightarrow \mathbb{k}$ is recovered from the linear operator as the eigenvalue function assigning eigenvalues to the eigenvector basis set U .

For the purposes of extending the mathematics of set partitions to vector spaces to obtain the mathematical tools of QM, it is the *operator* F that is used in the Yoga, not the linear functional. The elements of the basis set U are a basis of eigenvectors and the values of the numerical attribute are the eigenvalues of the operator F . Two numerical attributes defined on the *same* set U would generate two linear operators that commute with the basis set U as a basis of simultaneous eigenvectors as illustrated in Table 5.

Set concept	Vector-space concept
$f : U \rightarrow \mathbb{k}$	$F : V \rightarrow V$ by $Fu = f(u)u$
$g : U \rightarrow \mathbb{k}$	$G : V \rightarrow V$ by $Gu = g(u)u$
f, g on same set U	F, G commuting with basis U of simult. eigenvectors

Table 5: Operators corresponding to numerical attributes.

The inverse-image of $f : U \rightarrow \mathbb{k}$ is a partition on U where each block $f^{-1}(k)$ is associated with a distinct $k \in f(U) \subseteq \mathbb{k}$. What is the vector-space version of a set partition? If we had taken the vector-space analog of $f : U \rightarrow \mathbb{k}$ as the linear functional $\hat{f} : V \rightarrow \mathbb{k}$, then the corresponding vector space concept would be the inverse-image of the functional which is a special type of *set* partition on V called a *commuting* or *permutable* partition (much studied by Gian-Carlo Rota and earlier by Dubreil and Dubreil-Jacotin [12]).

Following our choice, the vector-space notion of a partition is a *direct-sum decomposition*.⁹ Each block in the set partition $\{f^{-1}(k)\}_{k \in f(U)}$ generates a subspace which is just the eigenspace V_k of

⁸There is the linearization map (functor) which takes a set U to the vector space \mathbb{C}^U where u lifts to the basis vector $\delta_u = \chi_{\{u\}} : U \rightarrow \mathbb{C}$, but we apply the Yoga to many different concepts.

⁹In view of the duality between subsets and set partitions, there is the induced duality between subspaces and direct-sum decompositions. Hence the Birkhoff-von-Neumann quantum logic of subspaces [6] has a dual quantum logic of DSDs [17].

the linear operator $F : V \rightarrow V$ determined by the eigenvalue $k \in f(U) \subseteq \mathbb{k}$, and the vector space V is the direct sum of the eigenspaces $V = \bigoplus_{k \in f(U)} V_k$.

A direct-sum decomposition (DSD) $\{V_k\}_{k \in f(U)}$ is characterized by the fact that every nonzero vector $v \in V$ can be represented in exactly one way as the sum $v = \sum_{k \in f(U)} v_k$ of nonzero vectors $v_k \in V_k$. The vectors v_k are the projections $P_k(v)$ of v to the subspaces V_k . What happens if we run the Yoga backwards? Does that property of DSDs also characterize set partitions? A set partition is usually defined as a set of nonempty subsets $\{B_1, B_2, \dots, B_m\}$ that are disjoint and jointly exhaustive. Each block B_j defines a projection operator $B_j \cap () : \wp(U) \rightarrow \wp(U)$ on the powerset $\wp(U)$ as a vector space \mathbb{Z}_2^n over \mathbb{Z}_2 (with set addition as the symmetric difference [16]) down to the subspace $\wp(B_j)$. Then we have the result:

$\{B_1, B_2, \dots, B_m\}$ is a set partition of U
if and only if

every nonempty subset $S \subseteq U$, is uniquely expressed as the union of subsets of the B_j , $j = 1, \dots, m$
if and only if

$\{\wp(B_j)\}_{j=1}^m$ is a DSD of $\wp(U) \cong \mathbb{Z}_2^n$.

This running of the Yoga backwards raises the question of what is the set-analogue of the notion of an eigenvector and eigenvalue. For $k \in \mathbb{k}$ and $S \subseteq U$, let “ kS ” stand for “the value k assigned to the elements of S ”. Then we have:

The eigenvector/eigenvalue equation for $f : U \rightarrow \mathbb{k}$:
 $f \upharpoonright S = kS$ in analogy with $Fu = ku$.

Thus the set-notion of an eigenvector is just a constant set and the set notion of an eigenvalue is that constant value on a constant set as illustrated in Table 6.

A characteristic function $\chi_S : U \rightarrow \{0, 1\} \subseteq \mathbb{k}$ applied to a basis set for V linearizes to a projection operator $P_{[S]} : V \rightarrow V$ defined by $P_{[S]}u = \chi_S(u)u$. The constant sets of χ_S are S and $S^c = U - S$ and the eigenspaces of $P_{[S]}$ are $[S]$ and $[S^c]$ with the respective eigenvalues of 1 and 0. In general, $f : U \rightarrow \mathbb{k}$ linearized to F defined by $Fu = f(u)u$ with the eigenspaces of F being $V_k = [f^{-1}(k)]$ for $k \in f(U)$. If $P_k : V \rightarrow V$ is the projection operator to V_k , then the spectral decomposition of F is $\sum_k kP_k$ and the corresponding ‘spectral decomposition’ for f is: $f = \sum_{k \in f(U)} k\chi_{f^{-1}(k)} : U \rightarrow \mathbb{k}$.

Set concept	Vector-space concept
Partition $\{f^{-1}(k)\}_{k \in f(U)}$	DSD $\{V_k\}_{k \in f(U)} : V = \bigoplus_{k \in f(U)} V_k$
$f \upharpoonright S = kS$	$Fu_i = ku_i$
Constant set S of f	Eigenvector u_i of F
Value k on constant set S	Eigenvalue k of eigenvector u_i
Characteristic fcn. $\chi_S : U \rightarrow \mathbb{k}$	Projection op. $P_{[S]} : V \rightarrow V$
Spectral decomp. $f = \sum_{k \in f(U)} k\chi_{f^{-1}(k)}$	Spectral decomp. $F = \sum_k kP_k$
Set of k -constant sets $\wp(f^{-1}(k))$	Eigenspace of k -eigenvectors V_k
Partition join $\pi \vee \sigma$	Lüders op. $\sum_{k \in f(U)} P_k \rho P_k$

Table 6: Corresponding concepts based on partition of a set and direct-sum decomposition of a vector space.

What is the vector space version of the Cartesian or direct product of sets $U \times U'$? One yoga might say the direct product of vector spaces $V \times W$. But our Yoga is apply the set concept to basis sets and see what it generates. Let U be a basis for V and U' be a basis for W , both over the same field, then applying the Cartesian product to the basis sets gives $U \times U'$ and it (bi)linearly generates the tensor product $V \otimes W$ with the ordered pair (u, u') elements of $U \times U'$ corresponding to the basis elements $u \otimes u'$ of $V \otimes W$ as given in Table 7.

Set concept	Vector-space concept
Cartesian product $U \times U'$	Tensor product $V \otimes W$
(u, u') element of $U \times U'$	Basis element $u \otimes u'$ of $V \otimes W$

Table 7: Cartesian product of sets corresponds to the tensor product of vector spaces.

7 Measurement with Quantum Logical Entropy

We can now use the Yoga to extend the ‘classical’ notion of logical entropy for set partitions to quantum logical entropy. The set-version of logical entropy was defined on set partitions determined by numerical attributes $f : U \rightarrow \mathbb{k}$, and the corresponding quantum notion of logical entropy is given by DSDs determined by self-adjoint operators $F : V \rightarrow V$ on a Hilbert space V which have a basis U of eigenvectors with an eigenvalue function $f : U \rightarrow \mathbb{k}$ where $Fu_i = f(u_i)u_i$. An ordered pair of set elements (u_i, u_k) in the direct product $U \times U$ is a distinction or dit if they have different numerical attributes $f(u_i) \neq f(u_k)$, and, similarly, an ordered pair of eigenvectors written $u_i \otimes u_k$ in the tensor product $V \otimes V$ is a *qudit* if they have different eigenvalues $f(u_i) \neq f(u_k)$. Both the set version and vector space versions of logical entropy satisfy Andrei Kolmogorov’s dictum:

Information theory must precede probability theory, and not be based on it. By the very essence of this discipline, the foundations of information theory have a finite combinatorial character. [29, p. 39]

In the set version, logical entropy is represented by the finite combinatorial ditset $\text{dit}(\pi) \subseteq U \times U$ and in the vector space version, the quantum logical entropy is represented by the finite dimensional subspace $[\text{qudit}(F)] \subseteq V \otimes V$ (the subspace generated by the set $\text{qudit}(F)$ of qudits $u_i \otimes u_k$ of F). The following Table 8 gives the set to vector space correspondence before probabilities are introduced.

Classical Logical Information	Quantum Logical Information
$f, g : U \rightarrow \mathbb{R}$	Commuting self-adjoint ops. F, G
$U = \{u_1, \dots, u_n\}$	ON basis simultaneous eigenvectors F, G
Values $\{\phi_i\}_{i=1}^m$ of f	Eigenvalues $\{\phi_i\}_{i=1}^m$ of F
Values $\{\gamma_j\}_{j=1}^m$ of g	Eigenvalues $\{\gamma_j\}_{j=1}^m$ of G
Partition $\{f^{-1}(\phi_i)\}_{i=1}^m$	Eigenspace DSD of F
Partition $\{g^{-1}(\gamma_j)\}_{j=1}^m$	Eigenspace DSD of G
dits of $\pi : (u_k, u_{k'}) \in U^2, f(u_k) \neq f(u_{k'})$	Qudits of $F: u_k \otimes u_{k'} \in V \otimes V, f(u_k) \neq f(u_{k'})$
dits of $\sigma : (u_k, u_{k'}) \in U^2, g(u_k) \neq g(u_{k'})$	Qudits of $G: u_k \otimes u_{k'} \in V \otimes V, g(u_k) \neq g(u_{k'})$
$\text{dit}(\pi) \subseteq U \times U$	$[\text{qudit}(F)] =$ subspace gen. by qudits of F
$\text{dit}(\sigma) \subseteq U \times U$	$[\text{qudit}(G)] =$ subspace gen. by qudits of G
$\text{dit}(\pi) \cup \text{dit}(\sigma) \subseteq U \times U$	$[\text{qudit}(F) \cup \text{qudit}(G)] \subseteq V \otimes V$
$\text{dit}(\pi) - \text{dit}(\sigma) \subseteq U \times U$	$[\text{qudit}(F) - \text{qudit}(G)] \subseteq V \otimes V$
$\text{dit}(\pi) \cap \text{dit}(\sigma) \subseteq U \times U$	$[\text{qudit}(F) \cap \text{qudit}(G)] \subseteq V \otimes V$

Table 8: Ditsets and qudit subspaces without probabilities.

In the set case of logical entropy, we assumed a point probability distribution p on U and then applied the product distribution $p \times p$ to the ditset $\text{dit}(f^{-1}) \subseteq U \times U$ of the inverse-image partition $\{f^{-1}(\phi_i)\}_{i=1}^m$ to get the logical entropy $h(f^{-1}) = p \times p(\text{dit}(f^{-1}))$ which was interpreted as the probability of getting different f -values in two independent samples of the random variable f .

In the quantum case, the probabilities only enter by considering a certain state ψ with a density matrix $\rho(\psi)$ (represented in the basis of F eigenvectors) and then the density matrix $\rho(\psi) \otimes \rho(\psi)$ on $V \otimes V$. The probability of getting distinct eigenvalues in two independent measurements of the state ψ by the observable F is the trace of the projection $P_{[\text{qudit}(F)]}$ to the qudit space times the density

matrix $\rho(\psi) \otimes \rho(\psi)$, which is an $n^2 \times n^2$ matrix with the diagonal entries $(\rho(\pi) \otimes \rho(\pi))_{(j,k),(j,k)} = \rho(\pi)_{jj} \rho(\pi)_{kk} = p_j p_k$. The quantum logical entropy is then:

$$h(F : \psi) = \text{tr} [P_{[\text{qudit}(F)]} \rho(\psi) \otimes \rho(\psi)].$$

At the set level, we saw there were one-draw-probability versus two-draw-probability interpretations of logical probability and logical entropy respectively. Thus just as the quantum probability $\text{Pr}(\phi) = \text{tr} [P_\phi \rho(\psi)]$ (where $P_\phi =$ projection op. to the eigenspace V_ϕ of the eigenvalue ϕ and $\rho(\psi) =$ density matrix of state ψ represented in measurement basis) is the one-measurement probability of getting the eigenvalue ϕ , so the quantum logical entropy $\text{tr} [P_{[\text{qudit}(F)]} \rho(\psi) \otimes \rho(\psi)]$ is the two-measurement probability of getting different eigenvalues in two independent measurements of the same state.

Thus the quantum logical entropy $h(F : \psi)$ is defined in terms of a state ψ and measurements by the observable F . But in our previous treatment of density matrices, we saw that the previous logical entropy of set partitions could also be defined as $h(\rho(\pi)) = 1 - \text{tr} [\rho(\pi)^2] = h(\pi)$. Hence there is also a quantum notion defined for any density matrix as: $h(\rho) = 1 - \text{tr} [\rho^2]$ [21]. We saw that the indiscrete partition $\mathbf{0}_U$ made no distinctions so $h(\mathbf{0}_U) = 0$ and similarly for a pure state $\rho^2 = \rho$ so $h(\rho) = 1 - \text{tr} [\rho^2] = 1 - \text{tr} [\rho] = 0$ since all density matrices have trace 1.

This correspondence is illustrated in Table 9 where F and G are commuting observables so they have a basis of simultaneous eigenvectors (which is the analogue of two numerical attributes $f, g : U \rightarrow \mathbb{k}$ defined on the same set U).

'Classical' Logical Entropy	Quantum Logical Entropy
Pure state density matrix, e.g., $\rho(\mathbf{0}_U)$	Pure state density matrix $\rho(\psi)$
$U = \{u_1, \dots, u_n\}$	ON basis simultaneous eigenvectors F, G
$p \times p$ on $U \times U$	$\rho(\psi) \otimes \rho(\psi)$ on $V \otimes V$
$h(\mathbf{0}_U) = 1 - \text{tr} [\rho(\mathbf{0}_U)^2] = 0$	$h(\rho(\psi)) = 1 - \text{tr} [\rho(\psi)^2] = 0$
$h(\pi) = p \times p(\text{dit}(\pi))$	$h(F : \psi) = \text{tr} [P_{[\text{qudit}(F)]} \rho(\psi) \otimes \rho(\psi)]$
$h(\pi, \sigma) = p \times p(\text{dit}(\pi) \cup \text{dit}(\sigma))$	$h(F, G : \psi) = \text{tr} [P_{[\text{qudit}(F) \cup \text{qudit}(G)]} \rho(\psi) \otimes \rho(\psi)]$
$h(\pi \sigma) = p \times p(\text{dit}(\pi) - \text{dit}(\sigma))$	$h(F G : \psi) = \text{tr} [P_{[\text{qudit}(F) - \text{qudit}(G)]} \rho(\psi) \otimes \rho(\psi)]$
$m(\pi, \sigma) = p \times p(\text{dit}(\pi) \cap \text{dit}(\sigma))$	$m(F, G : \psi) = \text{tr} [P_{[\text{qudit}(F) \cap \text{qudit}(G)]} \rho(\psi) \otimes \rho(\psi)]$
$h(\pi) = h(\pi \sigma) + m(\pi, \sigma)$	$h(F : \psi) = h(F G : \psi) + m(F, G : \psi)$
$\text{Pr}(\phi_i) = \text{tr} [P_{f^{-1}(\phi_i)} \rho(f^{-1})]$	$\text{Pr}(\phi_i) = \text{tr} [P_{\phi_i} \rho(\psi)]$
$\text{Pr}(\phi_i) =$ one-draw prob. of ϕ_i	$\text{Pr}(\phi_i) =$ one-meas. prob. of ϕ_i
$h(f^{-1}) =$ two-draw prob. diff. f -values	$h(F : \psi) =$ two-meas. prob. diff. F -eigenvalues
$\rho(\pi) = \sum_i P_{B_i} \rho(\mathbf{0}_U) P_{B_i}$	$\hat{\rho}(\psi) = \sum_i P_{V_i} \rho(\psi) P_{V_i}$ (Lüders)
$h(\pi) = 1 - \text{tr} [\rho(\pi)^2]$	$h(F : \psi) = 1 - \text{tr} [\hat{\rho}(\psi)^2]$
$h(\pi) =$ sum sq. zeroed $\rho(\mathbf{0}_U) \rightsquigarrow \rho(\pi)$	$h(F : \psi) =$ sum ab. sq. zeroed $\rho(\psi) \rightsquigarrow \hat{\rho}(\psi)$

Table 9: Logical entropies with probabilities applied to ditsets and qudit spaces.

The last three lines of Table 9 anticipate the quantum version of the previous results about real-valued density matrices. One of the main results about density matrices (over the complex numbers where $\|\rho_{ij}\|^2$ is the absolute square of ρ_{ij}) is:

Proposition 1 $\text{tr} [\rho^2] = \sum_{i,j} \|\rho_{ij}\|^2$. [22, p. 77]

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

In general the quantum logical entropy of a density matrix ρ is: $h(\rho) = 1 - \text{tr}[\rho^2] = 1 - \sum_{ij} \|\rho_{ij}\|^2$. The terms $\|\rho_{ij}\|^2$ are the ‘indistinction’ probabilities so $h(\rho) = 1 - \sum_{ij} \|\rho_{ij}\|^2$ is, as in the classical case, the sum of the probabilities of distinctions.

The change in density matrices due to a projective measurement is given by the Lüders mixture operation. If V_i is the eigenspace for the eigenvalue ϕ_i of F and P_{V_i} is the projection matrix $P_{V_i} : V \rightarrow V$ to that subspace, then the post-measurement density matrix by the Lüders mixture operation is:

$$\hat{\rho}(\psi) = \sum_i P_{V_i} \rho(\psi) P_{V_i}.$$

The previous result about the sum of the squares of the non-zero off-diagonal elements of ρ that are zeroed in the transition $\rho \rightsquigarrow \hat{\rho}$ carries over to the quantum case of density matrices over the complex numbers.

Theorem 2 *The increase in quantum logical entropy, $h(\hat{\rho}(\psi))$ due to the F -measurement of the pure state ψ is the sum of the absolute squares of the non-zero off-diagonal terms (coherences) in $\rho(\psi)$ (represented in a basis of F -eigenvectors) that are zeroed (decohered) in the post-measurement Lüders mixture density matrix $\hat{\rho}(\psi) = \sum_i P_{V_i} \rho(\psi) P_{V_i}$.*

Proof: $h(\hat{\rho}(\psi)) - h(\rho(\psi)) = \left(1 - \text{tr}[\hat{\rho}(\psi)^2]\right) - \left(1 - \text{tr}[\rho(\psi)^2]\right) = \sum_{j,k} \left(\|\rho_{jk}(\psi)\|^2 - \|\hat{\rho}_{jk}(\psi)\|^2\right)$ since $\text{tr}[\rho^2] = \sum_{i,j} \|\rho_{ij}\|^2$ is the sum of the absolute squares of all the elements of ρ . If u_j and u_k are a qudit of F , then and only then are the corresponding off-diagonal terms in $\rho(\psi)$ zeroed by the Lüders mixture operation $\sum_{i=1}^I P_{V_i} \rho(\psi) P_{V_i}$ to obtain $\hat{\rho}(\psi)$ from $\rho(\psi)$. \square

A careful calculation shows that $h(F : \psi) = h(\hat{\rho}(\psi))$ which also equals the sum of the absolute squares of zeroed terms in the transition $\rho(\psi) \rightsquigarrow \hat{\rho}(\psi)$ (since pure states $\rho(\psi)$ have zero quantum logical entropy). These equalities can be illustrated by working through a simple example of measuring z -axis spin.

Example: Let $|\psi\rangle = \alpha_\uparrow |\uparrow\rangle + \alpha_\downarrow |\downarrow\rangle = \begin{bmatrix} \alpha_\uparrow \\ \alpha_\downarrow \end{bmatrix}$ be a pure normalized superposition state of z -spin up and z -spin down so the density matrix is $\rho(\psi) = \begin{bmatrix} p_\uparrow & \alpha_\uparrow \alpha_\downarrow^* \\ \alpha_\downarrow \alpha_\uparrow^* & p_\downarrow \end{bmatrix}$ (where α^* is the complex conjugate of α). For the observable F , let the eigenvalue function be $f : \{|\uparrow\rangle, |\downarrow\rangle\} \rightarrow \{+1, -1\}$ where $f(|\uparrow\rangle) = 1$ and $f(|\downarrow\rangle) = -1$. Then $F : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ is represented by the matrix $F = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. The tensor product $\rho(\psi) \otimes \rho(\psi)$ is the $2^2 \times 2^2$ matrix:

$$\begin{bmatrix} p_\uparrow \rho(\psi) & \alpha_\uparrow \alpha_\downarrow^* \rho(\psi) \\ \alpha_\downarrow \alpha_\uparrow^* \rho(\psi) & p_\downarrow \rho(\psi) \end{bmatrix} = \begin{bmatrix} p_\uparrow^2 & p_\uparrow \alpha_\uparrow \alpha_\downarrow^* & \alpha_\uparrow \alpha_\downarrow^* p_\uparrow & \alpha_\uparrow \alpha_\downarrow^* \alpha_\uparrow \alpha_\downarrow^* \\ p_\uparrow \alpha_\downarrow \alpha_\uparrow^* & p_\uparrow p_\downarrow & \alpha_\uparrow \alpha_\downarrow^* \alpha_\downarrow \alpha_\uparrow^* & \alpha_\uparrow \alpha_\downarrow^* p_\downarrow \\ \alpha_\downarrow \alpha_\uparrow^* p_\uparrow & \alpha_\downarrow \alpha_\uparrow^* \alpha_\uparrow \alpha_\downarrow^* & p_\downarrow p_\uparrow & p_\downarrow \alpha_\uparrow \alpha_\downarrow^* \\ \alpha_\downarrow \alpha_\uparrow^* \alpha_\downarrow \alpha_\uparrow^* & \alpha_\downarrow \alpha_\uparrow^* p_\downarrow & p_\downarrow \alpha_\downarrow \alpha_\uparrow^* & p_\downarrow^2 \end{bmatrix}.$$

The qudits of F are $|\uparrow\rangle \otimes |\downarrow\rangle$ and $|\downarrow\rangle \otimes |\uparrow\rangle$ so the projection matrix to the subspace [qudit (F)] of $\mathbb{C}^2 \otimes \mathbb{C}^2$ generated by those qudits is:

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

Hence the quantum logical entropy $h(F : \psi) = \text{tr}[P_{\text{qudit}(F)} \rho(\psi) \otimes \rho(\psi)]$ is:

$$\begin{aligned}
h(F : \psi) &= \text{tr} \left(\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_{\uparrow}^2 & p_{\uparrow}\alpha_{\uparrow}\alpha_{\downarrow}^* & \alpha_{\uparrow}\alpha_{\downarrow}^*p_{\uparrow} & \alpha_{\uparrow}\alpha_{\downarrow}^*\alpha_{\uparrow}\alpha_{\downarrow}^* \\ p_{\uparrow}\alpha_{\downarrow}\alpha_{\uparrow}^* & p_{\uparrow}p_{\downarrow} & \alpha_{\uparrow}\alpha_{\downarrow}^*\alpha_{\downarrow}\alpha_{\uparrow}^* & \alpha_{\uparrow}\alpha_{\downarrow}^*p_{\downarrow} \\ \alpha_{\downarrow}\alpha_{\uparrow}^*p_{\uparrow} & \alpha_{\downarrow}\alpha_{\uparrow}^*\alpha_{\uparrow}\alpha_{\downarrow}^* & p_{\downarrow}p_{\uparrow} & p_{\downarrow}\alpha_{\uparrow}\alpha_{\downarrow}^* \\ \alpha_{\downarrow}\alpha_{\uparrow}^*\alpha_{\downarrow}\alpha_{\uparrow}^* & \alpha_{\downarrow}\alpha_{\uparrow}^*p_{\downarrow} & p_{\downarrow}\alpha_{\downarrow}\alpha_{\uparrow}^* & p_{\downarrow}^2 \end{bmatrix} \right) \\
&= \text{tr} \left(\begin{bmatrix} 0 & 0 & 0 & 0 \\ \alpha_{\downarrow}p_{\uparrow}\alpha_{\uparrow}^* & p_{\uparrow}p_{\downarrow} & \alpha_{\uparrow}\alpha_{\downarrow}\alpha_{\uparrow}^*\alpha_{\downarrow}^* & \alpha_{\uparrow}p_{\downarrow}\alpha_{\downarrow}^* \\ \alpha_{\downarrow}p_{\uparrow}\alpha_{\uparrow}^* & \alpha_{\uparrow}\alpha_{\downarrow}\alpha_{\uparrow}^*\alpha_{\downarrow}^* & p_{\uparrow}p_{\downarrow} & \alpha_{\uparrow}p_{\downarrow}\alpha_{\downarrow}^* \\ 0 & 0 & 0 & 0 \end{bmatrix} \right) = 2p_{\uparrow}p_{\downarrow}.
\end{aligned}$$

The second way to calculate the quantum logical entropy of the post-measurement state is using the Lüders mixture operation. The measurement of that spin-observable F goes from the pure state $\rho(\psi)$ to

$$\begin{aligned}
&P_{\uparrow}\rho(\psi)P_{\uparrow} + P_{\downarrow}\rho(\psi)P_{\downarrow} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p_{\uparrow} & \alpha_{\uparrow}\alpha_{\downarrow}^* \\ \alpha_{\downarrow}\alpha_{\uparrow}^* & p_{\downarrow} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_{\uparrow} & \alpha_{\uparrow}\alpha_{\downarrow}^* \\ \alpha_{\downarrow}\alpha_{\uparrow}^* & p_{\downarrow} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} p_{\uparrow} & 0 \\ 0 & p_{\downarrow} \end{bmatrix} = \hat{\rho}(\psi).
\end{aligned}$$

The logical entropy of $\hat{\rho}(\psi)$ is:

$$h(\hat{\rho}(\psi)) = 1 - \text{tr} \left[\hat{\rho}(\psi)^2 \right] = 1 - p_{\uparrow}^2 - p_{\downarrow}^2 = 2p_{\uparrow}p_{\downarrow} = h(F : \psi).$$

The third way to calculate the quantum logical entropy of $\hat{\rho}(\psi)$ is to sum the absolute squares of the non-zero off-diagonal terms in the pure state density matrix $\rho(\psi)$ that are zeroed in the transition to the post-measurement density matrix $\hat{\rho}(\psi)$, i.e.,

$$\rho(\psi) = \begin{bmatrix} p_{\uparrow} & \alpha_{\uparrow}\alpha_{\downarrow}^* \\ \alpha_{\downarrow}\alpha_{\uparrow}^* & p_{\downarrow} \end{bmatrix} \rightsquigarrow \begin{bmatrix} p_{\uparrow} & 0 \\ 0 & p_{\downarrow} \end{bmatrix} = \hat{\rho}(\psi),$$

and that sum is:

$$2\alpha_{\uparrow}\alpha_{\downarrow}^*\alpha_{\downarrow}\alpha_{\uparrow}^* = 2p_{\uparrow}p_{\downarrow} = h(\hat{\rho}(\psi)) = h(F : \psi).$$

It might be noted that the Born rule is built into the density matrix formulation, e.g.,

$$\text{Pr}(|\uparrow\rangle) = \text{tr}[P_{\uparrow}\rho(\psi)] = \text{tr} \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p_{\uparrow} & \alpha_{\uparrow}\alpha_{\downarrow}^* \\ \alpha_{\downarrow}\alpha_{\uparrow}^* & p_{\downarrow} \end{bmatrix} \right) = \alpha_{\uparrow}\alpha_{\uparrow}^* = p_{\uparrow}.$$

Quantum measurement creates distinctions, e.g., the distinction between spin-up and spin-down in the spin example, and quantum logical entropy precisely measures those distinctions. We started ‘classically’ with a universe set of distinct elements

The set can be thought of as being originally fully distinct, while each partition collects together blocks whose distinctions are factored out. Each block represents elements that are associated with an equivalence relation on the set. Then, the elements of a block are indistinct among themselves while different blocks are distinct from each other, given an equivalence relation. [44, p. 1]

Then the Yoga of Linearization translated the concepts of information-as-distinctions to the corresponding vector space concept which would be the concepts of quantum information-as-qudits. One of the founders of quantum information theory, Charles Bennett, captured the target concept of information as distinctions, differences, 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...” [4, p. 155]

With these concepts in mind, it seems that the extension of this framework of partitions and distinctions to the study of quantum systems may bring new insights into problems of quantum state discrimination, quantum cryptography, and quantum channel capacity. In fact, in these problems, we are, in one way or another, interested in a distance measure between distinguishable states, which is exactly the kind of knowledge the logical entropy is associated with. [44, p. 1]

8 Other quantum mathematics

8.1 Commuting and non-commuting observables

One of the characteristic features of QM mathematics is the possibility that observables (expressed as self-adjoint operators or matrices) do not commute—which at first does not seem related to the partition math of indefiniteness. But the vector space version of a partition on a set is a direct-sum decomposition of a vector space. Given two self-adjoint operators $F, G : V \rightarrow V$, let $\{V_i\}_{i \in I}$ be the DSD of eigenspaces for F and $\{W_j\}_{j \in J}$ be the DSD of eigenspaces for G . As we saw for quantum measurement, the relevant partition operation is the join, so we may mimic the join operation with the two DSDs. This join-like operation yields the set of non-zero vector spaces $\{V_i \cap W_j\}$ which are the subspaces spanned by the simultaneous eigenvectors of F and G . The join of two partitions on the *same* set yields a partition of that set. Let \mathcal{SE} be the subspace of V spanned by the non-zero subspaces $\{V_i \cap W_j\}$, i.e., the subspace spanned by the simultaneous eigenvectors of F and G . The point is that \mathcal{SE} need not be the whole space. The condition specifying whether F and G commute or not is exactly the condition that $\mathcal{SE} = V$ or not. The *commutator* of F and G is: $[F, G] = FG - GF : V \rightarrow V$, and as a linear operator on V , the commutator has a kernel $\ker [F, G]$ which is the subspace of vectors v such that $[F, G]v = 0$.

Proposition 2 $\mathcal{SE} = \ker ([F, G])$.

Proof: Let $F, G : V \rightarrow V$ be two self-adjoint 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 [28, 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 vectors in the subspace. Hence $\ker ([F, G]) \subseteq \mathcal{SE}$. \square

Since the condition that the operators commute or not is $\ker ([F, G]) = V$ or not, it is equivalent to $\mathcal{SE} = V$ or not, so the commutativity condition on operators is captured by the mathematics of the vector-space version of partitions, i.e., DSDs. And the further condition of the operators being *conjugate* is when $\mathcal{SE} = \mathbf{0}$ (the subspace consisting of only the zero vector). The Heisenberg "uncertainty" principle is somewhat misnamed since "uncertainty" may imply a subjective uncertainty instead of objective indefiniteness. The "indefiniteness principle" or "indeterminacy principle" might be a better name.¹⁰ In any case, since conjugate observables have no (non-zero) simultaneous eigenvectors, $\mathcal{SE} = \mathbf{0}$, if a system is in an eigenstate of one observable, it cannot be in an eigenstate of the other observable.

Since we have shown how the mathematics of indefiniteness can be translated into vector spaces, it might be noted this conjugacy is not a peculiarly quantum concept about operators in Hilbert

¹⁰Heisenberg's German word was "Unbestimmtheit" which could well be translated as "indefiniteness" or "indeterminacy" rather than "uncertainty."

spaces but can occur in quite simple vector spaces such as $\wp(U) \cong \mathbb{Z}_2^n$ (where set addition in $\wp(U)$ is symmetric difference [16]). In particular, for n an even number as with $U = \{a, b, c, d\}$, the U -basis set $\{\{a\}, \{b\}, \{c\}, \{d\}\}$ for $\wp(U) \cong \mathbb{Z}_2^4$ has a conjugate basis of:

$$\{\hat{a}\} = \{b, c, d\}, \{\hat{b}\} = \{a, c, d\}, \{\hat{c}\} = \{a, b, d\}, \{\hat{d}\} = \{a, b, c\}$$

which constitutes the \hat{U} -basis. Two numerical attributes $f : U \rightarrow \mathbb{R}$ and $g : \hat{U} \rightarrow \mathbb{R}$ with distinct values on each basis set (e.g., 1, 2, 3, 4) would define the two DSDs on $\wp(U) \cong \mathbb{Z}_2^4$ which, expressed in the U -basis as the computational basis, are:

$$\begin{aligned} \text{DSD for } f & \text{ is } \{\{0, \{a\}\}, \{0, \{b\}\}, \{0, \{c\}\}, \{0, \{d\}\}\} \\ \text{DSD for } g & \text{ is } \{\{0, \{b, c, d\}\}, \{0, \{a, c, d\}\}, \{0, \{a, b, d\}\}, \{0, \{a, b, c\}\}\}. \end{aligned}$$

Clearly, there no non-zero eigenvectors in common between the two DSDs so the numerical attributes are conjugate. If the state was in an eigenstate of one attribute such as $\{a\}$, then it is in an indefinite superposition $\{\hat{b}\} + \{\hat{c}\} + \{\hat{d}\} = \{a\}$ according to the conjugate basis (and vice-versa). In Figure 4, we illustrated the set-level pure and mixed states using the lattice $\Pi(U)$. The conjugate \hat{U} -basis also has a similar lattice $\Pi(\hat{U})$ and moving to a more definite state $\{a\} + \{b\} + \{c\} = \{a, b, c\} \rightsquigarrow \{a\}$ in the U -basis would correspond to moving to a less definite state in the conjugate basis \hat{U} , e.g., $\{a, b, c\} = \{\hat{d}\} \rightsquigarrow \{\hat{b}, \hat{c}, \hat{d}\} = \{a\}$, as one would expect for conjugate bases.

It might be noted that these numerical attributes cannot be repackaged as linear operators with the eigenvalues in the base field since the only such linear operators on \mathbb{Z}_2^n are projection operators with eigenvalues 0 or 1. But all the concepts of compatibility, i.e., $\mathcal{SE} = V$, incompatibility, i.e., $\mathcal{SE} \neq V$, and conjugacy, i.e., $\mathcal{SE} = \mathbf{0}$, can be defined using the vector-space partitions, i.e., DSDs, in \mathbb{Z}_2^4 . Thus the mathematics behind “non-commutativity” in QM is not about operators *per se*, but about the underlying vector space partitions or DSDs—as was to be shown.

Commuting observables are like ordinary numerical attributes on the *same* set U —in that case a basis of simultaneous eigenvectors. It is only when $\mathcal{SE} = V$ that the join-like operation taking non-zero intersections of the eigenspaces can properly be called the *join* of the DSDs, otherwise it is a join-like (or proto-join) operation when $\mathcal{SE} \neq V$. As Hermann Weyl put it: “Thus combination [join] of two gratings [vector space partitions] presupposes commutability...”. [49, p. 257]

A set of ordinary set partitions on the same universe U is said to be *complete* if their join is the discrete partition $\mathbf{1}_U = \{\{u\}\}_{u \in U}$ where the subsets in the join have cardinality one. Numerical attributes defined on the same set are *compatible*, and if they defined a complete set of partitions, they would be a complete set of compatible attributes (CSCA). If the partitions arose as the inverse-images of numerical attributes (or random variables), then each element in U would be characterized by the ordered set of values of the attributes.

Similarly a set of commuting operators is said be *complete* (a CSCO) [11] if all the non-zero intersections of all their eigenspaces, i.e., the subspaces in the join, are of dimension one. Then each of the simultaneous eigenvectors is uniquely characterized by the ordered set of eigenvalues of those intersected eigenspaces. These results are summarized in Table 10.

Set concept	Vector-space concept
Set partition	Direct-sum decomposition
Partitions on $U \neq U'$, $ U = U' $	DSDs on V with $\mathcal{SE} \neq V$
Partitions on same set U	DSDs on V with $\mathcal{SE} = V$
Join of partitions on same set U	Join of DSDs is DSD of V
CSCA	CSCO

Table 10: Set and vector-space versions of commutativity.

8.2 The two types of quantum processes

Quantum concepts need to be ‘seen’ in a certain way to see the underlying mathematics of indefiniteness and definiteness as shown in the case of non-commuting operators. At first glance, the Schrödinger equation to describe the evolution of an isolated system seems to have nothing to do with distinctions. Von Neumann classified quantum processes into Type 1 (measurement) and Type 2 (evolution described by the Schrödinger equation). We have seen that a measurement or Type 1 process creates distinctions so the natural characterization of the Type 2 processes would be ones that make no distinctions.

The extent to which two quantum states are indistinct or distinct is given by their inner product, i.e., their overlap. When their inner product is zero, then there is zero indistinctness or zero overlap between the states, i.e., they are fully distinct. Hence the natural characterization of the Type 2 processes as not changing the indistinctness or distinctness between quantum states would a process that preserves inner products, i.e., a unitary transformation. Hence the division of quantum processes into Type 1 and Type 2 is just the division between the processes that makes distinctions and those that don’t.

What about the Schrödinger equation? The connection between unitary transformations and the solutions to the Schrödinger equation is given by Stone’s Theorem [43]: there is a one-to-one correspondence between strongly continuous 1-parameter *unitary* groups $\{U_t\}_{t \in \mathbb{R}}$ and self-adjoint operators H (Hamiltonian) on the Hilbert space so that $U_t = e^{iHt}$ (solutions of the Schrödinger equation).

8.3 Measurement and the collapse postulate

We have seen that quantum measurements create distinctions. Richard Feynman was perhaps the quantum theorist who most emphasized measurement as making distinctions. When a superposition of eigenstates undergoes an interaction, is there a distinction made in principle between the superposed eigenstates in the interaction? If the eigenstates are distinguished by the interaction, then a measurement takes place, the superposition is reduced (i.e., the so-called “wave function” collapses), and the probability of a later final state will add the probabilities (rather than amplitudes) of the eigenstates leading to the outcome. If there is no differences or distinctions between the superposed eigenstates undergoing the interaction, then no measurement takes place and the amplitudes are added.

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.[25, p. 3-9]

Feynman thus answers a question posed in the literature where the key concepts of distinguishability and indistinguishability are not used.

It indeed seems necessary to admit that “measurements” are ubiquitous, and occur even in places and times where there are no human experimenters. But it also seems hopeless to think that we will be able to give an appropriately sharp answer to the question of what, exactly, differentiates the ‘ordinary’ processes (where the usual dynamical rules apply) from the ‘measurement-like’ processes (where the rules momentarily change). [35, p. 64]

[I]t seems unbelievable that there is a fundamental distinction between “measurement” and “non-measurement” processes. Somehow, the true fundamental theory should treat all processes in a consistent, uniform fashion. [35, p. 245]

The “fundamental distinction” is between processes where, in the interaction, distinctions are made or are not made between the eigenstates in the superposition.

Feynman gives an example of a measurement entirely at the quantum level, and thus he undercuts the long and tortured discussion about measurement as involving a macroscopic apparatus. 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 a mixture of 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 as 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. It should be noted that this and other examples of Feynman [25] involve only quantum level interactions and thus have nothing to do with the “shifty split” [3] between microscopic and macroscopic, and thus are independent of the notion of “decoherence” based on interactions with macroscopic systems (e.g., [51]).

For instance in the double-slit experiment, if no distinction is made between the particle going through one slit or the other (i.e., no detectors at the slits), then the two parts in the superposition schematically represented by:

$$|\text{going through slit 1}\rangle + |\text{going through slit 2}\rangle$$

evolve unitarily and will show interference effects. The mathematics of the unitary (i.e., no distinctions) evolution over the complex numbers will have a complex-valued “wave interpretation” (Stone’s Theorem) without there being any physical waves; the interference results from the addition of vectors representing the parts of the evolving superposition.¹¹

There being no distinctions involved, the evolving indefinite superposition does not reach the level of definiteness of the particle going through one slit or the other—which conflicts with our ‘classical’ intuitive always-definite view of the particle’s trajectory. A better but still crude intuitive picture would be to recognize the *levels of indefiniteness* at the set level in the lattice of partitions (e.g., Figure 4)—which would prefigure levels of indefiniteness at the quantum level. There is the completely indefinite state represented by the indiscrete partition at the bottom of the lattice, and then moving up to partitions with each block representing an equally or more definite superposition, and finally reaching the discrete partition at the top where each block is a singleton ‘eigenstate’ analogous to the mixed state of fully decomposed eigenstates represented by a diagonal density matrix.

The difference between an interaction that constitutes a measurement or not is whether or not any distinction is made between the different superposed eigenstates undergoing the interaction. Hence Feynman’s implicit rule about state reduction might be paraphrased:

If the interaction would make distinctions, then distinctions are made.

In other words, if the interaction makes a difference between the superposed eigenstates, then the superposition decoheres with a (indefinite-to-definite) state reduction to one of those eigenstates.

One image for the measurement process is a ‘shapeless’ or indefinite blob of dough which then passes through a sieve or grating and acquires a definite polygonal shape as illustrated in Figure 7. This sort of grating or sieve imagery has been used before. In his popular writing, Arthur Eddington used the sieve metaphor:

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.[13, p. 267]

¹¹This is particularly clear in the pedagogical model of the double-slit experiment in QM over \mathbb{Z}_2 [16] where the interference in the evolving 0, 1-vectors has no resemblance to waves.

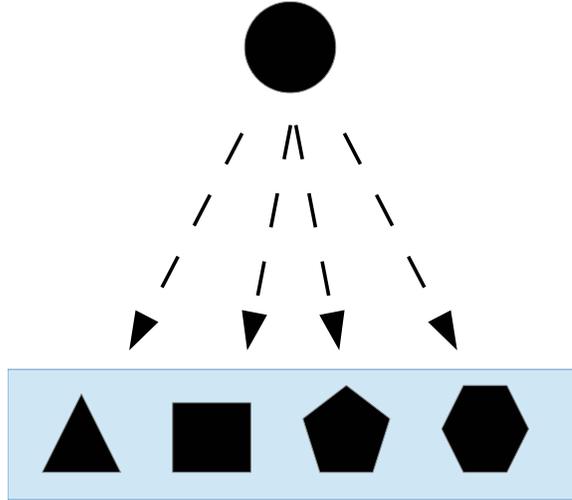


Figure 7: Measurement as an ‘indefinite’ shape passing through a sieve to get a definite polygonal shape

Hermann Weyl cited that passage [49, p. 255] in his expositional concept of gratings. Then Weyl, in effect, used the Yoga of Linearization by taking *both* set partitions and vector space partitions (direct-sum decompositions) as the respective types of gratings.[49, pp. 255-257] He started with a numerical attribute on a set (or “aggregate” in older terminology), e.g., $f : U \rightarrow \mathbb{R}$, which defined the set partition or “grating” [49, p. 255] with blocks having the same attribute-value, e.g., $\{f^{-1}(r)\}_{r \in f(U)}$. Then he implicitly applied the Yoga to reach the QM 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” [49, p. 256]. Then the notion of a vector space partition or “grating” in QM 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” [49, p. 256], i.e., a direct-sum decomposition of the space. After thus referring to a partition and a DSD as a “grating” or “sieve,” Weyl notes that “Measurement means application of a sieve or grating” [49, p. 259], i.e., an interaction that makes distinctions and thus forces an indefinite-to-definite transition, e.g., Figure 7.

Our overall goal has been to show that the mathematics of QM is the mathematics in distinctions and indistinctions or definiteness and indefiniteness expressed in terms of vector spaces. The vision of realism based on objective indefiniteness is juxtaposed to our ordinary intuitive idea that reality is definite-all-the-way-down. One aspect of “the measurement problem” has been the lack of any mathematical description of how a quantum system goes from an objectively indefinite superposition state to a more definite state during a measurement. But that question seems to arise out of imposing the fully definite framework as if it was only a transition from a perfectly definite but rather featureless state to another definite state with more discernible features—like classically going from a blank sheet of paper to a sheet with figures on it. But if quantum reality consists of objectively indefinite states, why should we expect that sort of definite-to-definite transition between states of indefiniteness—as opposed to the notion of a genuine quantum jump or leap? Leibniz’s view of reality as being definite all the way down (expressed in his identity of indiscernible) was also expressed by the slogan “*Natura non facit saltus*” (Nature does not make jumps”) [32, Bk. IV, chap. xvi]. Hence it should not be too surprising to find the opposite phenomena of jumps in a reality that is not always definite. Thus the unanswered question of the mathematical description of the ‘trajectory of a quantum jump’ may arise from an implicit assumption that reality is definite all the way down as in classical physics.

8.4 Indistinguishability of particles

The classical notion of distinguishability of particles in effect treats partitions (or DSDs) as always being refineable or definite all the way down, e.g., distinguishing particles by ‘painting different colors.’

In quantum mechanics, however, identical particles are truly indistinguishable. This is because we cannot specify more than a complete set of commuting observables for each of the particles; in particular, we cannot label the particle by coloring it blue. [39, p. 446]

Hence quantum indistinguishability *immediately* points to objective indefiniteness, as opposed to “definiteness all the way down.” If definiteness does not go “all the way down,” then the making of distinctions has to stop at some point at which the remaining indefiniteness has to be objective.

If quantum reality is not definite-all-the-way-down, then at the level where further definiteness stops (as it were), there are two possibilities. A complete state description (e.g., CSCO-defined) is sufficient to limit at most a single particle to that state or the complete description is still insufficient to limit the number of particles in that state—in neither case distinguishing between other particles of the same type.

This difference can be illustrated by the metaphor of different levels of definiteness in a mailing address. In a neighborhood of only single-family houses or vacant lots, then an address that is definite down to the street number would be sufficient to limit one or no families to each address. But in a neighborhood which had apartment houses, then addresses limited to the street number in definiteness (i.e., no apartment numbers) would allow many families at the same address.

The mathematics of quantum statistics for the two types of particles can be developed using the standard combinatorics of balls-in-boxes—which, unlike the usual treatment using symmetric and anti-symmetric wave functions, brings out the underlying role of distinctions and indistinctions. There are k balls (or particles) and they are indistinguishable. There are n boxes (or states) and they are distinguishable. If the complete state description (or address) is sufficient to limit one or no balls to that state or box, then as each ball finds an empty box, then that box is removed as a possibility for the next ball so the number of equiprobable placements of balls in boxes is given by the *falling factorial* $n(n-1)\dots(n-k+1)$ (k terms) as pictured in Figure 8.

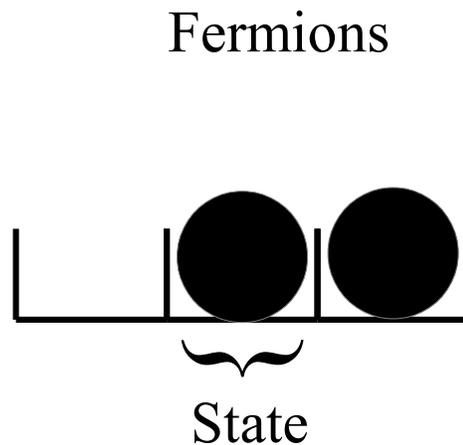


Figure 8: Fermions fill a state

If, however, there can be many balls in a box, then the placement of each ball creates two additional ‘slots,’ before and after the ball and the number of equiprobable slots increase with each

placement to give the *rising factorial* $n(n+1)\dots(n+k-1)$ (k terms) as pictured in Figure 9.

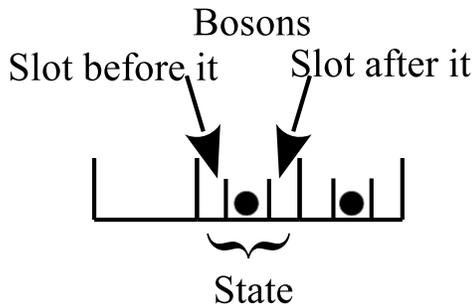


Figure 9: Each placement of a boson creates an additional possible slot

In each case, since the balls are indistinguishable, the number of equal possibilities must be divided by $k!$. Hence in the k -fold tensor product, the number of equiprobable placements is the dimension of the subspace of possible states.

- The dimension of the fermionic subspace the usual binomial coefficient:

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

The Fermi-Dirac statistics counts each state as having equal probability: $1/\binom{n}{k}$.

- The dimension of the bosonic subspace is:

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

The Bose-Einstein statistics counts each state as having equal probability $1/\langle n \rangle_k$. [23, p. 40]

An equivalent way to enumerate the number of possible states is to enumerate the number of functions of a certain type from balls to boxes.

- Fermi-Dirac statistics is based on the number of ways indistinguishable balls (particles) are allocated to distinguishable boxes (states) using *distinction-preserving* (i.e., one-to-one) functions (so two numerically distinct balls have to go to distinct boxes), while:
- Bose-Einstein statistics is based on the number of ways indistinguishable balls (particles) are allocated to distinguishable boxes (states) using *arbitrary* functions.

The classical case of Maxwell-Boltzmann (MB) statistics is where the k balls (particles) are distinguishable, the n boxes (states) are distinguishable, and the distributions of balls to boxes are by arbitrary functions. There are $k!$ different linear orders (or permutations) of the k distinguishable particles but they are grouped into n boxes with the occupation numbers of $\theta_1, \dots, \theta_n$ for the n boxes. How many distributions are there with those occupation numbers? The answer is still $k!$ independent of the occupation numbers. The proof is illustrated in Figure 10 since the $n-1$ ‘walls’ or state-dividers to make the boxes can be put in arbitrarily to get the given occupation numbers $\theta_1 + \dots + \theta_n = k$.

The ordering of the balls within each box does not matter so we need to divide through by the $\theta_i!$ for $i = 1, \dots, n$ to get the total number of possible states with those occupation numbers for the distinguishable boxes. This gives the well-known *multinomial coefficient*:

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

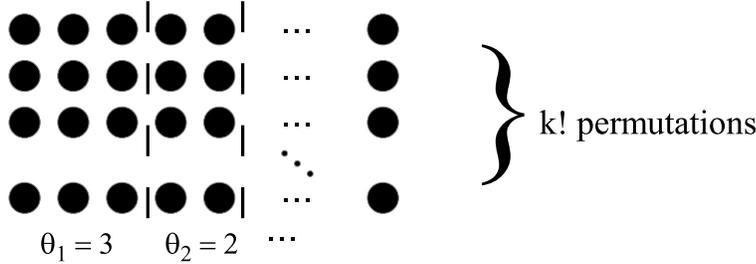


Figure 10: The number of ways k balls into n boxes with given occupation numbers is $k!$

There are n^k arbitrary functions distributing the balls in the boxes and each distribution is classically considered equiprobable so the probability of the given set of occupation numbers in the MB statistics is:

$$\binom{k}{\theta_1, \dots, \theta_n} / n^k = \frac{k!}{\theta_1! \dots \theta_n!} / n^k.$$

The difference between MB, BE, and FD statistics can be illustrated by computing the probability of flipping two coins of the same type. Hence there are $k = 2$ particles of the same type and $n = 2$ states $\{h, t\}$ 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: $\Pr_{MB}(\{(h, t), (t, h)\}) = \frac{\binom{k}{\theta_1, \dots, \theta_n}}{n^k} = \frac{\frac{k!}{\theta_1! \dots \theta_n!}}{n^k} = \frac{2}{4} = \frac{1}{2}$.
- Boson coins: $\Pr_{BE}(\{(h, t), (t, h)\}) = \frac{k!}{n(n+1)\dots(n+k-1)} = \frac{2}{2(3)} = \frac{1}{3}$.
- Fermion coins: $\Pr_{FD}(\{(h, t), (t, h)\}) = \frac{k!}{n(n-1)\dots(n-k+1)} = \frac{2}{2(1)} = 1$.

The Pauli exclusion principle is illustrated by the fact that two fermion coins *have* to be in different states (i.e., with probability 1). In the case of bosons, the two classical outcomes (h, t) and (t, h) differ only by a permutation of particles of the same type so that counts only as one state out of three equiprobable states. Since the probability of getting the same outcomes (h, h) or (t, t) is $\frac{2}{3}$ in the bosonic case in comparison with the classical MB probability of $\frac{1}{2}$, that illustrates the tendency of bosons to “be more social” (e.g., in terms of our metaphor, live in an apartment house).

8.5 Group representation theory

Group representation theory is a key part of the mathematics of QM. This immediately supports our thesis since a group representation is essentially a ‘dynamic’ or ‘active’ way to define an equivalence relation; each group operation transforms an element into an equivalent or symmetric element.

Given a set G indexing (associative) 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 R on $U \times U$ by:

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

Then the conditions that make R_g into a group representation are the conditions that imply R is an equivalence relation:

1. existence of the identity $1_U \in G$ implies reflexivity of R ;
2. existence of inverses implies symmetry of R ; 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 R .

Thus a set representation $\{R_g\}_{g \in G}$ of a group G (or group action on a set) is essentially a ‘dynamic’ way to define an equivalence relation R on the set [7]. The minimal invariant (or irreducible) subsets of the set representation are the *orbits*, and they are the equivalence classes of the equivalence relation R or blocks of the orbit partition. The restriction of a set representation of a group to an orbit is an *irreducible representation* or *irrep*.

Let $f : U \rightarrow \mathbb{R}$ be a numerical attribute on U and consider the $n \times n$ diagonal matrix $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ with the diagonal entries $(\tilde{f})_{ii} = f(u_i)$. Let $M : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be any matrix that commutes with \tilde{f} , $\tilde{f}M = M\tilde{f}$, i.e., the following diagram commutes:

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{M} & \mathbb{R}^n \\ \tilde{f} \downarrow & & \downarrow \tilde{f} \\ \mathbb{R}^n & \xrightarrow{M} & \mathbb{R}^n \end{array} .$$

Then computing the ik entries: $(\tilde{f}M)_{ik} = f(u_i)M_{ik} = M_{ik}f(u_k) = (M\tilde{f})_{ik}$. Thus $[f(u_i) - f(u_k)]M_{ik} = 0$ so if (u_i, u_k) is a dit of the partition f^{-1} , i.e., $f(u_i) \neq f(u_k)$, then $M_{ik} = 0$.¹²

A numerical attribute $f : U \rightarrow \mathbb{R}$ is said to *commute* with the set representation $R = \{R_g\}_{g \in G}$ if for any R_g , the following diagram commutes:

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

Taking $R_g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ as a permutation matrix, then f commuting with $R = \{R_g\}_{g \in G}$ means $\tilde{f}R_g = R_g\tilde{f}$ so if $(u_i, u_k) \in \text{dit}(f^{-1})$, then $(R_g)_{ik} = 0$ for all $g \in G$. A subset $S \subseteq U$ is said to be *invariant* under R if $R_g(S) \subseteq S$ for all $g \in G$. The blocks $f^{-1}(r)$ for $r \in f(U)$ for a commuting f are invariant subsets of U under R . Thus the partition f^{-1} for a commuting f is refined by the orbit partition since orbits are the *minimal* invariant subsets. A set of commuting attributes f, g, \dots, h is said to be *complete* if the join $f^{-1} \vee g^{-1} \vee \dots \vee h^{-1}$ is the orbit partition.

Given a finite group G and a finite-dimensional vector space V over \mathbb{C} , a *vector space representation* of the group G is a map $R : G \rightarrow GL(V)$ (the general linear group of invertible linear maps on V) where $g \mapsto R_g : V \rightarrow V$ from G to invertible linear maps on V such that $R_1 = I$ and $R_{g'}R_g = R_{g'g}$. Using the Yoga to lift set concepts to vector space concepts, the notion of a minimal invariant subset or orbit yields the notion of a minimal invariant subspace which is called an *irreducible subspace* of V . Just as the orbits of a set representation form a partition of U , so the irreducible subspaces of a vector space representation form a direct-sum decomposition of V . The restriction of a vector space representation of a group to an irreducible subspace is an *irreducible representation* or *irrep*.

Let $\{R_g\}_{g \in G}$ be a set representation of G on a set U that is an ON basis set for V and let $f : U \rightarrow \mathbb{R}$ be a commuting attribute. Then for any $u_i \in U$ and any $g \in G$, if $R_g(u_i) = u_k \in U$, then by commutativity, $f(u_i) = f(R_g(u_i)) = f(u_k)$. Then a unitary operator $F : V \rightarrow V$ is defined by $Fu = f(u)u$ and the group representation R_g extends to V from its definition on the basis set U . Does the operator F commute with R_g in the sense that for each R_g , the following diagram commutes?

¹²In the general vector space case, for two commuting (diagonalizable) operators, $FG = GF$, if both are represented in a basis of F -eigenvectors and (u_i, u_k) is a qudit of F , then $G_{ik} = 0$. [45, p. 4]

$$\begin{array}{ccc} V & \xrightarrow{R_g} & V \\ F \downarrow & & \downarrow F \\ V & \xrightarrow{R_g} & V \end{array} .$$

Starting with a basis element $u_i \in V$ and going around the square clockwise, $R_g(u_i) = u_k \in V$ is taken to $Fu_k = f(u_k)u_k \in V$. Going around the square counterclockwise, $Fu_i = f(u_i)u_i$ and $R_g(f(u_i)u_i) = f(u_i)u_k = f(u_k)u_k$ so the square commutes. Hence the set-concept of a commuting f extends by the Yoga to the usual concept of an observable F commuting with a vector space representation.¹³

The eigenspaces of a commuting F are invariant under R . Then the DSD of eigenspaces for a commuting F is refined (defined in the obvious way [17]) by the DSD of irreducible subspaces. A set of operators commuting with R such as F, G, \dots, H is said to be *complete* if the join-like operation on their DSDs has all its subspaces as irreducible.

Schur's Lemma (set case): A commuting f restricted to (i.e., $f \upharpoonright$) an irreducible subset (i.e., an orbit) is a constant function.

Schur's Lemma (vector space case): A commuting F restricted to (i.e., $F \upharpoonright$) an irreducible subspace is a constant operator.

The Yoga of Linearization applied to group representations is illustrated in Table 11.

Yoga	Set representations	Vector space reps.
Representation	$\{R_g : U \rightarrow U\}_{g \in G}$	$\{R_g : V \rightarrow V\}_{g \in G}$
Min. Invariants	Orbits	Irreducible subspaces
Partition	Orbit partition	DSD of irred. subspaces
Irreps	Rep. on orbits	Rep. on irred. subspaces
Commuting	$f : U \rightarrow \mathbb{R}, \forall g, fR_g = f$	$F : V \rightarrow V, \forall g, FR_g = R_gF$
Invariants	$f^{-1}(r)$ commuting f	Eigenspaces commuting F
Schur's Lemma	Comm. $f \upharpoonright$ orbit const.	Comm. $F \upharpoonright$ irred. sp. constant

Table 11: Summary of Yoga of Linearization for group representations.

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

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

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

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

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

¹³The Yoga is used to generate vector space concepts corresponding to set concepts. There is no implication that every instance of a vector space concept, e.g., a commuting F , must come from an instance of the set concept, e.g., a commuting f .

For $R_{(0,1)}$, they are:

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

Since the two operators commute, their eigenspace DSDs commute so we can take their join. The blocks of the join are a DSD and are automatically invariant. Since the blocks of the join are one-dimensional, those four subspaces are also irreducible and thus the 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 restrictions of the group representation to these four irreducible subspaces gives 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.

Moving to non-Abelian groups, not to mention Lie groups, greatly increases the mathematical complexity. But for our purposes, the point is that the key concepts of group representations for QM come out of the partitioned mathematics of definiteness and indefiniteness. The irreps give all the different ways that minimal definite alternatives are defined consistent with the indistinction-creating symmetries of the group. The properties of the eigen-alternatives are 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. The irreps fill out the symmetry-adapted possibilities.

Since the group action creates indistinctions as symmetries, moving to the action of a subgroup means less indistinctions and more distinctions, i.e., symmetry-breaking, which is the second method, in addition to the join-like operations, to move from an indefinite state to a more definite state, e.g., in the description of the Big Bang [36].

Classical physics also has symmetries so groups will play an important role, e.g., the Noether Theorems. However group representations play a more fundamental role in quantum mechanics as might be expected from their role in the mathematics of partitions. Partly, this is due to the objectively indefinite states given by the superposition principle. At a more fundamental level, it is due to the irreps (the restriction of representations to the irreducible subspaces) of certain group representations defining the elementary particles at the quantum level.

The reason is fundamentally, that the variety of states is much greater in quantum theory than in classical physics and that there is, on the other hand, the principle of superposition to provide a structure for the greatly increased manifold of quantum mechanical states. The principle of superposition renders possible the definition of the states the transformation properties of which are particularly simple. It can in fact be shown that every state of any quantum mechanical system, no matter what type of interactions are present, can be considered as a superposition of states of elementary systems. The elementary systems correspond mathematically to irreducible representations of the Lorentz group and as such can be enumerated. [50, p. 8]

Prior to the discussion of group representation theory, we started with a *given* set U of distinct elements which were then taken as a given basis set for a vector space. But those elementary given distinct elements or basis vectors are *not* given in group representation theory; they are instead determined by the group as the irreducible subspaces and irreducible representations (irreps) of the representation. The irreps are the elementary symmetry-adapted eigen-alternatives determined by the group of transformations.

For a certain symmetry group of particle physics, “an elementary particle ‘is’ an irreducible unitary representation of the group.” [42, p. 149] Thus our partitional approach comports with “the soundness of programs that ground particle properties in the irreducible representations of symmetry transformations...” [26, p. 171] These alternatives are carved out by the joins of the vector space partitions of 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” ([8, p. 211], also [9]; [47]).

This all goes back to the transformations of group representations being ‘dynamic’ or ‘active’ ways to define partitions (equivalence relations) and their vector space versions (DSDs).

9 Concluding remarks

We have taken the mathematics of QM to be sufficiently represented by:

- projective measurement and Lüders mixture operation, the von Neumann (vN) Type I process;
- commuting, non-commuting, and conjugate observables;
- evolution by the Schrödinger equation, the vN Type 2 process;
- measurement and the collapse postulate;
- quantum statistics for indistinguishable particles; and
- group representation theory applied to quantum mechanics.

And we have argued, in each case, that the mathematics of QM is the linearized to (Hilbert) vector space version of the mathematics of partitions. This is not just a coincidence or an accident. The mathematics tell us something about the unintuitive reality that QM so successfully describes.

- The key analytical concepts were the notions of definiteness versus indefiniteness, distinctions versus indistinctions, and distinguishability versus indistinguishability.
- The key machinery for moving from indefinite to more definite states was the partition-join-like operation of projective measurement—which is also quantitatively measured by quantum logical entropy (the vector space version of logical entropy at the set level).
- To arrive at a maximally definite state determination of a CSCO, the key operation was the partition-join operation on commuting DSDs (vector space partitions).
- And at the primordial level (Big Bang), the machinery to move from indefinite to more definite states was the distinction-creating symmetry-breaking change from an indistinction-creating symmetry groups to subgroups.

The ‘standard’ reality-oriented interpretations of QM (e.g., Bohmian mechanics, spontaneous collapse, or many-worlds) make little or no use of those key concepts and machinery. Our approach of showing the set-level origins of the mathematics of QM takes the formalism as being complete—without any addition of other variables, other equations, or other-worldly interpretations

of distinction-preserving (unitary) evolution and distinction-creating measurement. The reality-agnostic Copenhagen interpretation also takes the formalism of QM as being complete, so the partition mathematics approach could also be viewed as specifying the key concepts and machinery as well as supplying a dash of realism in the form of simplified images of properties and processes at the quantum level.

Objective indefiniteness at the quantum level violates our common-sense classical assumption of reality as being definite all the way down. The usual imagery of a superposition as the combination of two definite states (or waves) to yield another definite state (or wave) needs to be replaced by an indefiniteness imagery. The superposition of definite (or eigen) states (e.g., the qubit in quantum computing) should be seen as a state that is indefinite on the differences between the superposed states—which can be better understood if one takes the notion (and the underlying partitional math) of indefiniteness seriously.

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. [41, pp. 6-7]

The “Follow the Math!” approach shows that the math of QM comes from the math of partitions, and that picks out the key analytical concepts and machinery—and also allows some imagery, albeit simplified, of the nature of the physical properties and processes at the quantum level.

Statements and Declarations

There are no competing interests, no sources of funding, and no acknowledgments.

References

- [1] Ainsworth, T.: Form vs. Matter. In: Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy* (Spring 2016 Ed.) (2016) <https://plato.stanford.edu/archives/spr2016/entries/form-matter/>.
- [2] Auletta, G., Fortunato, M., Parisi, G.: *Quantum Mechanics*. Cambridge University Press, Cambridge UK (2009)
- [3] Bell, J.S.: Against “Measurement.” In: Miller, A.I. (ed.) *Sixty-Two Years of Uncertainty*. pp. 17–31. Plenum Press, New York (1990)
- [4] Bennett, C.H.: Quantum Information: Qubits and Quantum Error Correction. *International Journal of Theoretical Physics*. 42, 153–176 (2003). <https://doi.org/10.1023/A:1024439131297>
- [5] Birkhoff, G.: *Lattice Theory*. American Mathematical Society, New York (1948)
- [6] Birkhoff, G., Von Neumann, J.: The Logic of Quantum Mechanics. *Annals of Mathematics*. 37, 823–43 (1936)
- [7] Castellani, E.: Symmetry and equivalence. In: Brading, K. and Castellani, E. (eds.) *Symmetries in Physics: Philosophical Reflections*. pp. 425–436. Cambridge University Press, Cambridge (2003)
- [8] Chen, J.-Q., Gao, M.-J., Ma, G.-Q.: The representation group and its application to space groups. *Reviews of Modern Physics*. 57, 211–278 (1985)
- [9] Chen, J.-Q., Ping, J., Wang, F.: *Group Representation Theory for Physicists* (2nd ed.). World Scientific, Singapore (2002)

- [10] Cohen-Tannoudji, C., Diu, B., Laloë, F.: Quantum Mechanics: Volumes 1 and 2. John Wiley & Sons, New York (2005)
- [11] Dirac, P.A.M.: The Principles of Quantum Mechanics (4th ed.). Clarendon Press, Oxford (1958)
- [12] Dubreil, P., Dubreil-Jacotin, M.-L.: Théorie algébrique des relations d'équivalence. *J. de Mathématique*. 18, 63–95 (1939)
- [13] Eddington, A.S.: New Pathways in Science (Messenger Lectures 1934). Cambridge University Press, Cambridge UK (1947)
- [14] Ellerman, D.: Counting Distinctions: On the Conceptual Foundations of Shannon's Information Theory. *Synthese*. 168, 119–149 (2009). <https://doi.org/10.1007/s11229-008-9333-7>
- [15] Ellerman, D.: The Logic of Partitions: Introduction to the Dual of the Logic of Subsets. *Review of Symbolic Logic*. 3, 287–350 (2010). <https://doi.org/10.1017/S1755020310000018>
- [16] Ellerman, D.: Quantum mechanics over sets: a pedagogical model with non-commutative finite probability theory as its quantum probability calculus. *Synthese*. 194, 4863–4896 (2017). <https://doi.org/10.1007/s11229-016-1175-0>
- [17] Ellerman, D.: The quantum logic of direct-sum decompositions: the dual to the quantum logic of subspaces. *Logic Journal of the IGPL*. 26, 1–13 (2018). <https://doi.org/10.1093/jigpal/jzx026>
- [18] Ellerman, D.: The Logical Theory of Canonical Maps: The Elements & Distinctions Analysis of the Morphisms, Duality, Canonicity, and Universal Constructions in Set. *ArXiv.org*. (2021)
- [19] Ellerman, D.: On Abstraction in Mathematics and Indefiniteness in Quantum Mechanics. *Journal of Philosophical Logic*. 50, 813–835 (2021). <https://doi.org/10.1007/s10992-020-09586-1>
- [20] Ellerman, D.: New Foundations for Information Theory: Logical Entropy and Shannon Entropy. SpringerNature, Cham, Switzerland (2021)
- [21] Ellerman, D.: Introduction to logical entropy and its relationship to Shannon entropy. *4Open Special Issue: Logical Entropy*. 5, 1–33 (2022). <https://doi.org/10.1051/fopen/2021004>
- [22] Fano, U.: Description of States in Quantum Mechanics by Density Matrix and Operator Techniques. *Reviews of Modern Physics*. 29, 74–93 (1957)
- [23] Feller, W.: An Introduction to Probability Theory and Its Applications Vol. 1. John Wiley, New York (1968)
- [24] Feyerabend, P.: Problems of Microphysics. In: Colodny, R.G. (ed.) *Frontiers of Science and Philosophy*. pp. 189–283. University Press of America, Lanham MD (1983)
- [25] Feynman, R.P., Leighton, R.B., Sands, M.: The Feynman Lectures on Physics: Quantum Mechanics Vol. III. Addison-Wesley, Reading MA (1965)
- [26] Fine, A.: The Shaky Game: Einstein, Realism, and the Quantum Theory. University of Chicago Press, Chicago (1986)
- [27] Heisenberg, W.: *Physics & Philosophy: The Revolution in Modern Science*. Harper Torchbooks, New York (1962)
- [28] Hoffman, K., Kunze, R.: *Linear Algebra*. Prentice-Hall, Englewood Cliffs NJ (1961)
- [29] Kolmogorov, A.N.: Combinatorial Foundations of Information Theory and the Calculus of Probabilities. *Russian Math. Surveys*. 38, 29–40 (1983)

- [30] Kung, J.P.S., Rota, G.-C., Yan, C.H.: *Combinatorics: The Rota Way*. Cambridge University Press, New York (2009)
- [31] Lawvere, F.W., Rosebrugh, R.: *Sets for Mathematics*. Cambridge University Press, Cambridge MA (2003)
- [32] Leibniz, G.W.: *New Essays on Human Understanding*. Cambridge University Press, Cambridge UK (1996)
- [33] Lüders, G.: Über die Zustandsänderung durch Meßprozeß. *Annalen der Physik*. 8, 322–328 (1951)
- [34] Nielsen, M., Chuang, I.: *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge (2000)
- [35] Norsen, T.: *Foundations of Quantum Mechanics*. Springer International, Cham, Switzerland (2017)
- [36] Pagels, H.: *Perfect Symmetry: The Search for the Beginning of Time*. Simon and Schuster, New York (1985)
- [37] Rota, G.-C.: *Indiscrete Thoughts*. Birkhäuser, Boston (1997)
- [38] Rota, G.-C.: Twelve problems in probability no one likes to bring up. In: Crapo, H. and Senato, D. (eds.) *Algebraic Combinatorics and Computer Science: A Tribute to Gian-Carlo Rota*. pp. 57–93. Springer, Milano (2001)
- [39] Sakurai, J.J., Napolitano, J.: *Modern Quantum Mechanics*. Addison-Wesley, Boston (2011)
- [40] Shimony, A.: The reality of the quantum world. *Scientific American*. 258, 46–53 (1988)
- [41] Shimony, A.: Philosophical and Experimental Perspectives on Quantum Physics. In: *Philosophical and Experimental Perspectives on Quantum Physics: Vienna Circle Institute Yearbook 7*. pp. 1–18. Springer Science+Business Media, Dordrecht (1999)
- [42] Sternberg, S.: *Group Theory and Physics*. Cambridge University Press, Cambridge (1994)
- [43] Stone, M.H.: On one-parameter unitary groups in Hilbert Space. *Annals of Mathematics*. 33, 643–648 (1932)
- [44] Tamir, B., Piava, I.L., Schwartzman-Nowik, Z., Cohen, E.: Quantum logical entropy: fundamentals and general properties. *4Open Special Issue: Logical Entropy*. 5, 1–14 (2022). <https://doi.org/10.1051/fopen/2021005>
- [45] Tinkham, M.: *Group Theory and Quantum Mechanics*. Dover, Mineola NY (1964)
- [46] Von Neumann, J.: *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton (1955)
- [47] Wang, F.: A Conceptual Review of the New Approach to Group Representation Theory. In: Feng, D.H., Iachello, F., Ping, J.L., and Wang, F. (eds.) *The Beauty of Mathematics in Science: The Intellectual Path of J. Q. Chen*. pp. 1–8. World Scientific, Singapore (2004)
- [48] Weinberg, S.: *Dreams of a Final Theory*. Vintage Books, New York (1994)
- [49] Weyl, H.: *Philosophy of Mathematics and Natural Science*. Princeton University Press, Princeton (1949)

- [50] Wigner, E.P.: Invariance in Physical Theory. In: Symmetries and Reflections: Scientific Essays of Eugene P. Wigner. pp. 3–13. Ox Bow Press, Woodbridge CN (1967)
- [51] Zurek, W.H.: Decoherence, einselection, and the quantum origins of the classical. Review of Modern Physics. 75, 715–775 (2003)