

Chapter 4

The quantum logic formalism

4.1 Introduction

4.1.1 Why an algebraic structure?

The quantum logic formalism is another interesting, albeit more abstract, way to formulate quantum physics. The bonus of this approach is that one does not need to start from the assumption that the set of observables of a physical system is embodied with the algebraic structure of an associative unital algebra. As discussed in the previous section, this assumption that one can “add” and “multiply” observables is already a highly non trivial one. This algebraic structure is natural in classical physics since the observables form a commutative Poisson algebra, addition and multiplication of observables reflect the action of adding and multiplying results of different measurements (it is the Poisson bracket structure that is non trivial). In quantum physics these addition and multiplication operations on observables versus results are not equivalent anymore. Measuring the observable $C = A + B$ does not amount to measure independently the observables A and B and simply add the results, since in general the operators A and B do not commute, and cannot be measured independently. However we have seen for instance that the GNS construction relates the algebra structure of the observables to the Hilbert space structure of the pure states. In particular the superposition principle for pure states is a consequence of the existence of an addition law for the physical observables.

As will be explained in this chapter, in the quantum logic formulations the algebraic structure of the observables (the fact that they form an associative but non-commutative algebra) comes out somehow naturally from more basic assumptions one makes on the measurement operations, or tests, that one can make on physical systems. In the algebraic formulation the Hilbert space of states is reconstructed from the algebra structure of observables. Likewise in the quantum logic formulation this algebraic structure for the observables and the Hilbert space structure for the states can be derived (with some assumptions) from the possible symmetries that the measurement operations must satisfy.

4.1.2 Measurements as “logical propositions”

The idea at the root of the quantum logic approach goes back to J. von Neumann’s book [vN55, vN32] and the article by G. Birkhoff¹ and J. von Neumann (again!) [BvN36]. It starts from the remark that the observables given by projectors, i.e. operators \mathbf{P} such that $\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^\dagger$, correspond to propositions with YES or NO (i.e. TRUE or FALSE) outcome in a logical system. An orthogonal projector \mathbf{P} onto a linear subspace $P \subset \mathcal{H}$ is indeed the operator associated to an observable that can take only the values 1 (and always 1 if the state $\psi \in P$ is in the subspace P) or 0 (and always 0 if the state $\psi \in P^\perp$ belongs to the orthogonal subspace to P). Thus we can consider that measuring the observable \mathbf{P} is equivalent to perform a test on the system, or to check the validity of a logical proposition \mathbf{p} on the system.

$$\mathbf{P} = \text{orthogonal projector onto } P \leftrightarrow \text{proposition } \mathbf{p} \quad (4.1.1)$$

If the result is 1 the proposition \mathbf{p} is found to be TRUE, and if the result is 0 the proposition \mathbf{p} is found to be FALSE.

$$\langle \psi | \mathbf{P} | \psi \rangle = 1 \implies \mathbf{p} \text{ always TRUE on } |\psi\rangle \quad (4.1.2)$$

The projector $\mathbf{1} - \mathbf{P}$ onto the orthogonal subspace P^\perp is associated to the proposition **not** \mathbf{p} , meaning usually that \mathbf{p} is false (assuming the law of excluded middle)

$$\langle \psi | \mathbf{P} | \psi \rangle = 0 \implies \mathbf{p} \text{ always FALSE on } |\psi\rangle \quad (4.1.3)$$

so that

$$\mathbf{1} - \mathbf{P} = \text{orthogonal projector onto } P^\perp \leftrightarrow \text{proposition not } \mathbf{p} \quad (4.1.4)$$

In classical logic the negation **not** is denoted in various ways

$$\text{“not } \mathbf{a} \text{”} = \neg \mathbf{a}, \mathbf{a}', \bar{\mathbf{a}}, \tilde{\mathbf{a}}, \sim \mathbf{a} \quad (4.1.5)$$

I shall use the first two notations $\neg \mathbf{a}$, \mathbf{a}' .

Now if two projectors \mathbf{A} and \mathbf{B} (on two subspaces A and B) commute, they correspond to classically compatible observables A and B (which can be measured independently), and to a pair of propositions \mathbf{a} and \mathbf{b} of standard logic. The projector $\mathbf{C} = \mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A}$ on the intersection of the two subspaces $C = A \cap B$ corresponds to the proposition $\mathbf{c} = \text{“} \mathbf{a} \text{ and } \mathbf{b} \text{”} = \mathbf{a} \wedge \mathbf{b}$. Similarly the projector \mathbf{D} on the linear sum of the two subspaces $D = A + B$ corresponds to the proposition $\mathbf{d} = \text{“} \mathbf{a} \text{ or } \mathbf{b} \text{”} = \mathbf{a} \vee \mathbf{b}$.

$$A \cap B \leftrightarrow \mathbf{a} \wedge \mathbf{b} = \mathbf{a} \text{ and } \mathbf{b} \quad , \quad A + B \leftrightarrow \mathbf{a} \vee \mathbf{b} = \mathbf{a} \text{ or } \mathbf{b} \quad (4.1.6)$$

Finally the fact that for subspaces $A \subset B$, i.e. for projectors $\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A} = \mathbf{A}$, is equivalent to state that \mathbf{a} implies \mathbf{b}

$$A \subset B \leftrightarrow \mathbf{a} \implies \mathbf{b} \quad (4.1.7)$$

This is easily extended to a general (possibly infinite) set of *commuting* projectors. Such a set generates a commuting algebra of observables \mathcal{A} , which corresponds to

1. An eminent mathematician, not to be confused with his father, the famous G. D. Birkhoff of the ergodic theorem

the algebra of functions on some classical space X . The set of corresponding subspaces, with the operations of linear sum, intersection and orthocomplementation $(+, \cap, \perp)$, is isomorphic to a Boolean algebra of propositions with (\vee, \wedge, \neg) , or to the algebra of characteristic functions on subsets of X . Indeed, this is just a reformulation of “ordinary logic”² where characteristic functions of measurable sets (in a Borel σ -algebra over some set X) can be viewed as logical propositions. Classically all the observables of some classical system (measurable functions over its phase space Ω) can be constructed out of the classical propositions on the system (the characteristic functions of measurable subsets of Ω).

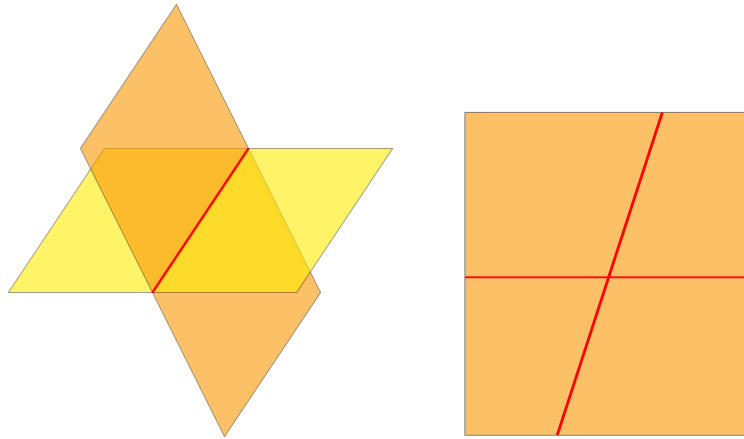


Figure 4.1: The \wedge as intersection and the \vee as linear sum of subspaces in quantum logic

In quantum mechanics all physical observables can be constructed out of projectors. For general, not necessarily commuting projectors \mathbf{A} and \mathbf{B} on subspaces A and B one still associate propositions \mathbf{a} and \mathbf{b} . The negation $\neg \mathbf{a}$, the “and” (or “meet”) $\mathbf{a} \wedge \mathbf{b}$ and the “or” (or “join”) $\mathbf{a} \vee \mathbf{b}$ are still defined by the geometrical operations \perp, \cap and $+$ on subspaces given by 4.1.6. The “implies” or \implies is also defined by the \subset as in 4.1.7

However the fact that in a Hilbert space projectors do not necessarily commute implies that the standard distributivity law of propositions

$$A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C) \quad \vee = \text{or} \quad \wedge = \text{and} \quad (4.1.8)$$

does not hold. It is replaced by the weaker condition (A, B, C are the linear subspaces associated to the projectors $\mathbf{A}, \mathbf{B}, \mathbf{C}$)

$$A \cap (B + C) \supset ((A \cap B) + (A \cap C)) \quad (4.1.9)$$

which corresponds in terms of propositions (projectors) to

$$(\mathbf{a} \wedge \mathbf{b}) \vee (\mathbf{a} \wedge \mathbf{c}) \implies \mathbf{a} \wedge (\mathbf{b} \vee \mathbf{c}) \quad (4.1.10)$$

or equivalently

$$\mathbf{a} \vee (\mathbf{b} \wedge \mathbf{c}) \implies (\mathbf{a} \vee \mathbf{b}) \wedge (\mathbf{a} \vee \mathbf{c}) \quad (4.1.11)$$

2. In a very loose sense, I am not discussing mathematical logic theory.

A simple example is depicted on fig. 4.2. The vector space V in the plane ($\dim=2$) and the subspaces A , B and C are three different coplanar lines ($\dim=1$). $B + C = V$, hence $A \cap (B + C) = A \cap V = A$, while $A \cap B = A \cap C = \{0\}$; hence $A \cap B + A \cap C = \{0\}$.

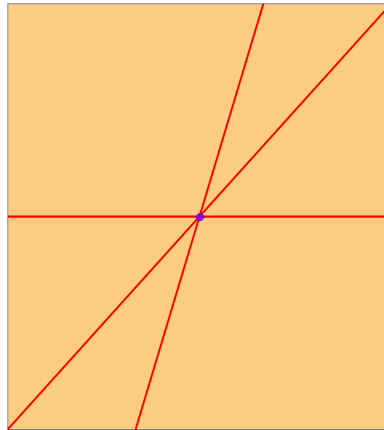


Figure 4.2: A simple example of non-distributivity

Therefore the set of projectors on a Hilbert space do not generate a Boolean algebra, but a more complicated structure, called an orthomodular lattice.

4.1.3 The quantum logic approach

The purpose of the quantum logic approach is to try to understand what are the minimal set of consistency requirements on the propositions/measurements operations, based on logical consistency (assuming that internal consistency has something to do with the physical world) and on physical requirements (in particular the assumptions of causality, reversibility and locality) and what are the consequences of these assumptions for the formulation of physical laws.

This approach was initiated by G. Birkhoff and J. von Neumann (again!) in [BvN36]. It was (slowly) developed by physicists like G. Mackey [Mac63], J. M. Jauch [Jau68] and notably by C. Piron [Pir64, Pir76], and by mathematicians like Varadarajan [Var85]. A good reference on the subject (not very recent but very valuable) is the book by E. Beltrametti and G. Cassinelli [BC81]. It is my main reference and source of understanding

The terminology “quantum logic” for this approach is historical and is perhaps not fully adequate, since it does not mean for most authors that a new kind of logic is necessary to understand quantum physics. It is in fact not a “logic” in the mathematical sense, and it relies on the standard logics used in mathematics and exact sciences. It should rather be called “quantum propositional calculus” or “quantum propositional geometry”, where the term “proposition” is to be understood as “test” or “projective measurement” on a quantum system. The mathematics of orthocomplemented and orthomodular lattices that underly the quantum logic formalism have applications in various areas of mathematics, logic and computer sciences.

At that stage I should emphasize that the quantum logic formulation is of no real use for real physical applications! Its formalism is quite heavy, and reduces for practi-

cal purposes to the standard formalism of Hilbert spaces and operator algebras, that is better suited and efficient for most if not all uses of the quantum formalism. Its interest is rather when discussing the conceptual foundations.

The quantum logic approaches do not form a unified precise and consistent framework like algebraic quantum field theory. It has several variants, most of them insist on the concept and the properties of propositions (tests), but some older one relying more on the concept of states (in particular the so called convex set approaches). Some more recent formulations of quantum physics related somehow to quantum logic and to quantum information science have some grandiose categorial formulations.

In this course I shall give a short and partial presentation of this approach, from a personal point of view³. I shall try to show where the physical concepts of causality, reversibility and locality play a role, in parallel to what I tried to do for the algebraic formalism. My presentation is a classical one based primarily on the concept of propositions. It is also conservative since I not try to use a non-classical logic system (whatever it means) but simply discuss in a classical logic framework the statements which can be made on quantum systems and the measurement operations.

4.2 A presentation of the principles

4.2.1 Projective measurements as propositions

As explained above, in the standard formulation of quantum mechanics, projectors are associated to “ideal” projective measurements (“projective measurement “of the first kind”, or “non-demolition” projective measurements). The fundamental property of such measurements is that if the system is already in an eigenstate of the projector, for instance $P|\psi\rangle = |\psi\rangle$, then after measurement the state of the system is unchanged. This means that successive measurements of P give always the same result (1 or TRUE). Without going into a discussion of measurements in quantum physics, let me stress that this is of course an idealisation of actual measurements. In general physical measurements are not ideal measurements, they may change in a non-minimal way the state of the system. While one gains some information on the system, one in general loses some other information, and measurements may and in general do destroy part or the whole of the system studied. Such general measurements and manipulation processes of a quantum system may be described by the general formalism of POVM's (Projective Operator Valued Measures).

In the following presentation, I assume that such ideal repeatable measurements are in principle for all the observable properties of a quantum system. The formalism discussed here will try to understand what is a natural and minimal set of physically reasonable and logically consistent axioms for such measurements.

4.2.2 Causality, POSET's and the lattice of propositions

Let me start from a set \mathcal{L} of propositions or tests (associated to ideal measurements) on a physical system, and from a set \mathcal{E} of states (in a similar sense as in the

3. with the usual reservation on the lecturer's qualifications

algebraic formulation, to be made more precise along the discussion). On a given state $\varphi \in \mathcal{E}$ the test (measurement) of the proposition a can give TRUE (i.e. YES or 1) or FALSE (NO or 0). It gives TRUE with some probability. In this case one has extracted information on the system, which is now (considered to be) in a state φ_a .

I note $\varphi(a)$ the probability that a is found TRUE, assuming that the system was in state φ before the test. I shall not discuss again at that stage what I mean exactly by probability, and refer the previous discussions.

4.2.2.a - Causal order relation:

The first ingredient is to assume that there an order relation $a \leq b$ between propositions. Here it will be defined by the causal relation

$$a \leq b \iff \text{for any state } \phi, \text{ if } a \text{ is found true, then } b \text{ will be found true} \quad (4.2.1)$$

Note that this definition is causal (or dynamical) from the start, as to be expected in quantum physics. It is equivalent to

$$a \leq b \iff \forall \phi, \phi_a(b) = 1 \quad (4.2.2)$$

One assumes that this causal relation has the usual properties of a partial order relation. This amounts to enforce relations between states and propositions. First one must have:

$$a \leq a \quad (4.2.3)$$

This means that if a has been found true, the system is now in a state such that a will always be found true. Second one assumes also that

$$a \leq b \text{ and } b \leq c \implies a \leq c \quad (4.2.4)$$

This is true in particular when, if the system is in a state ψ such that b is always true, then after measuring b , the system is still in the same state ψ . In other words, $\psi(b) = 1 \implies \psi_b = \psi$. This is the concept of repeatability discussed above.

These two properties makes \leq a *preorder relation* on \mathcal{L} .

One also assumes that

$$a \leq b \text{ and } b \leq a \implies a = b \quad (4.2.5)$$

This means that tests which give the same results on any states are indistinguishable. This also means that one can identify a proposition a with the set of states such that a is always found to be true (i.e. $\psi(a) = 1$). 4.2.5 makes \leq a partial order relation and \mathcal{L} a *partially ordered set* or POSET.

4.2.2.b - AND (meet \wedge):

The second ingredient is the notion of logical conjunction AND. One assumes that for any pair of test a and b , there is a *unique greater proposition*, denoted $a \wedge b$, such that

$$a \wedge b \leq a \text{ and } a \wedge b \leq b \quad (4.2.6)$$

in other word, there is a unique $a \wedge b$ such that

$$c \leq a \text{ and } c \leq b \implies c \leq a \wedge b$$

NB: This is a non trivial assumption, not a simple consequence of the previous ones. It can be justified using the notion of filters (see Jauch [Jau68]) or using the notion of questions associated to propositions (see Piron [Pir64]). Here to make the discussion simpler I just present it as an assumption. On the other hand it is very difficult to build anything without this assumption. Note that 4.2.6 implies 4.2.5.

This definition extends to any set \mathcal{A} of propositions

$$\bigwedge \mathcal{A} = \bigwedge \{a \in \mathcal{A}\} = \text{greatest } c : c \leq a, \forall a \in \mathcal{A} \quad (4.2.7)$$

I do not discuss if the set \mathcal{A} is finite or countable.

4.2.2.c - Logical OR (join \vee):

From this we can infer the existence of a logical OR (by using Birkhoff theorem)

$$a \vee b = \bigwedge \{c : a \leq c \text{ and } b \leq c\} \quad (4.2.8)$$

which extents to sets of propositions

$$\bigvee \mathcal{A} = \bigwedge \{b : a \leq b, \forall a \in \mathcal{A}\} \quad (4.2.9)$$

4.2.2.d - Trivial $\mathbf{1}$ and vacuous \emptyset propositions:

It is natural to assume that there is a trivial proposition $\mathbf{1}$ that is always true

$$\text{for any state } \phi, \mathbf{1} \text{ is always found to be true, i.e. } \phi(\mathbf{1}) = 1 \quad (4.2.10)$$

and another “vacuous” proposition \emptyset that is never true

$$\text{for any state } \phi, \emptyset \text{ is never found to be true, i.e. } \phi(\emptyset) = 0 \quad (4.2.11)$$

Naturally one has

$$\mathbf{1} = \bigvee \mathcal{L} \text{ and } \emptyset = \bigwedge \mathcal{L} \quad (4.2.12)$$

With these assumptions and definitions the set of propositions \mathcal{L} has now the structure of a complete lattice.

4.2.3 Reversibility and orthocomplementation

4.2.3.a - Negations a' and $'a$

I have not yet discussed what it is considered possible to do when a proposition is found to be false. To do so one must introduce the seemingly simple notion of negation or complement. In classical logic this is rather easy. The subtle point is that for quantum systems, where causality matters, there are two inequivalent ways to introduce the negation. These two definitions becomes equivalent only if one assumes that

propositions on quantum systems share a property that I denote “causal reversibility”, by analogy with what I did before, and that I am going to explain. If this property is satisfied, one recovers the standard negation of propositions in classical logic, and ultimately this will lead to the notion of orthogonality and of scalar product of standard quantum mechanics. Thus here again, as in the previous section, reversibility appears to be one of the essential feature of the principles of quantum physics.

Negation - définition 1: To any proposition a one can associate its negation (or complement proposition) a' defined as

$$\text{for any state } \phi, \text{ if } a \text{ is found to be true, then } a' \text{ will be found to be false} \quad (4.2.13)$$

a' can be defined equivalently as

$$a' = \bigvee \{b \text{ such that on any state } \phi, \text{ if } a \text{ is found true, then } b \text{ will be found false}\} \quad (4.2.14)$$

Negation - définition 2: It is important at that stage to realize that, because of the causal ordering in the above definition of a' , there is an alternate but symmetric definition for the negation, that I denote $'a$, which is given by

$$\text{for any state } \phi, \text{ if } 'a \text{ is found to be true, then } a \text{ will be found to be false} \quad (4.2.15)$$

or equivalently

$$'a = \bigvee \{b; \text{ such that on any state } \phi, \text{ if } b \text{ is found true, then } a \text{ will be found false}\} \quad (4.2.16)$$

These two definitions are *not equivalent*, and from the axioms that we choose up to now, each of them do not fulfill the properties of the negation in classical propositional logic⁴.

$$\neg(\neg a) = a \quad \text{and} \quad \neg(a \wedge b) = \neg a \vee \neg b$$

These problems come from the fact that the definition for the causal order $a \leq b$ does not implies that $b' \leq a'$, as in classical logic. Indeed the definition 4.2.1 for $a \leq b$ implies that for every state

$$\text{if } b \text{ is found false, then } a \text{ was found false} \quad (4.2.17)$$

while $b' \leq a'$ would mean

$$\text{if } b \text{ is found false, then } a \text{ will be found false} \quad (4.2.18)$$

or equivalently

$$\text{if } a \text{ is found true, then } b \text{ was found true} \quad (4.2.19)$$

4. The point discussed here is a priori not connected to the classical versus intuitionist logics debate. Remember that we are not discussing a logical system.

4.2.3.b - Causal reversibility and negation

In order to build a formalism consistent with what we know of quantum physics, we need to enforce the condition that the causal order structure on propositions is in fact independent of the choice of a causal arrow “if \dots , then \dots will be \dots ” versus “if \dots , then \dots was \dots ”. This is nothing but the requirement of causal reversibility that we discussed before, and it is enforced by the following simple but very important condition.

Causal reversibility: One assumes that the negation a' is such that

$$a \leq b \iff b' \leq a' \quad (4.2.20)$$

With this assumption, it is easy to show that the usual properties of negation are satisfied. The two alternate definitions of negation are now equivalent

$$a' = 'a = \neg a \quad (4.2.21)$$

and may be denoted by the standard logical symbol \neg . We then have

$$(a')' = a \quad (4.2.22)$$

and

$$(a \wedge b)' = a' \vee b' \quad (4.2.23)$$

as well as

$$\emptyset = \mathbf{1}', \quad \emptyset' = \mathbf{1} \quad (4.2.24)$$

and

$$a \wedge a' = \emptyset, \quad a \vee a' = \mathbf{1} \quad (4.2.25)$$

A lattice \mathcal{L} embodied with a complement \neg with the properties 4.2.20-4.2.25 is called an *orthocomplemented complete lattice* (in short OC lattice). For such a lattice, the couple (a, a') describes what is called a perfect measurement.

NB: Note that in Boolean logic, the implication \rightarrow can be defined from the negation \neg . Indeed $a \rightarrow b$ means $\neg a \vee b$. Here it is the negation \neg which is defined out of the implication \leq .

4.2.3.c - Orthogonality

With reversibility and complement, the set of propositions begins to have properties similar to the set of projections on linear subspaces of a Hilbert space⁵. The complement a' of a proposition a is similar to the orthogonal subspace P^\perp of a subspace P . This analogy can be extended to the general concept of orthogonality.

5. One should be careful for infinite dimensional Hilbert spaces and general operator algebras. Projectors correspond in general to orthogonal projections on closed subspaces.

Orthogonal propositions:

Two proposition a and b are orthogonal, if $b \leq a'$ (or equivalently $a \leq b'$). (4.2.26)

This is noted

$$a \perp b \quad (4.2.27)$$

Compatible propositions:

OC lattices contain also the concept of family of compatible propositions. A subset of an OC lattice \mathcal{L} is a sublattice \mathcal{L}' if it is stable under the operations \wedge , \vee and $'$ (hence it is itself an OC lattice). To any subset $\mathcal{S} \subset \mathcal{L}$ on can associate the sublattice $\mathcal{L}_{\mathcal{S}}$ generated by \mathcal{S} , defined as the smallest sublattice \mathcal{L}' of \mathcal{L} which contains \mathcal{S} .

A lattice is said to be *Boolean* if it satisfy the distributive law of classical logic

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$$

A subset \mathcal{S} of an OC lattice \mathcal{L} is said to be a subset of *compatible propositions* if the generated sublattice $\mathcal{L}_{\mathcal{S}} \subset \mathcal{L}$ is Boolean.

Compatible propositions are the analog of commuting projectors, i.e. compatible or commuting observables in standard quantum mechanics. For a set of compatible propositions, one expects that the expectations of the outcomes YES or NO will satisfy the rules of ordinary logic.

Orthogonal projection: The notion of orthogonal projection onto a subspace can be also formulated in this framework as

$$\text{projection of } a \text{ onto } b = \Phi_b(a) = b \wedge (a \vee b') \quad (4.2.28)$$

This projection operation is often called the Sasaki projection. Its dual $(\Phi_b(a'))' = b' \vee (a \wedge b)$ is called the Sasaki hook $(b \xrightarrow{S} a)$. It has the property that even if $a \not\leq b$, if for a state ψ the Sasaki hook $(a \xrightarrow{S} b)$ is always true, then for this state ψ , if a is found true then b will always be found true.

4.2.4 Subsystems of propositions and orthomodularity

4.2.4.a - What must replace distributivity?

The concept of orthocomplemented lattice of propositions is not sufficient to reconstruct a consistent quantum formalism. There are mathematical reasons and physical reasons.

One reason is that if the distributive law $A \wedge ((B \vee C) = (A \wedge B) \vee (A \wedge C)$ is known not to apply, assuming no restricted distributivity condition is not enough and leads to too many possible structures. In particular in general a lattice with an orthocomplementation \neg may be endowed with other inequivalent orthocomplementations! This is problematic for the physical interpretation of the complement as $a \rightarrow \text{TRUE} \iff a' \rightarrow \text{FALSE}$.

Another problem is that in physics one is led to consider conditional states and conditional propositions. In classical physics this would correspond to the restriction to some subset Ω' of the whole phase space Ω of a physical system, or to the projection $\Omega \rightarrow \Omega'$. Such projections or restrictions are necessary if there are some constraints on the states of the system, if one has access only to some subset of all the physical observables of the system, or if one is interested only in the study of a subsystem of a larger system. In particular such a separation of the degrees of freedom is very important when discussing locality: one is interested in the properties of the system we can associate to (the observables measured in) a given interval of space and time, as already discussed for algebraic QFT. It is also very important when discussing effective low energy theories: one wants to separate (project out) the (unobservable) high energy degrees of freedom from the (observable) low energy degrees of freedom. And of course this is crucial to discuss open quantum systems, quantum measurement processes, decoherence processes, and the emergence of classical degrees of freedom and classical behaviors in quantum systems.

4.2.4.b - Sublattices and weak-modularity

In general a subsystem is defined from the observables (propositions) on the system which satisfy some constraints. One can reduce the discussion to a single constraint a . If \mathcal{L} is an orthocomplemented lattice and a a proposition of \mathcal{L} , let me consider the subset $\mathcal{L}_{<a}$ of all propositions which imply a

$$\mathcal{L}_{<a} = \{b \in \mathcal{L} : b \leq a\} \quad (4.2.29)$$

One may also consider the subset of propositions $\mathcal{L}_{>a}$ of propositions implied by a

$$\mathcal{L}_{>a} = \{b \in \mathcal{L} : a \leq b\} = (\mathcal{L}_{<a'})' \quad (4.2.30)$$

Is this set of propositions $\mathcal{L}_{<a}$ still an orthocomplemented lattice? let me take as order relation \leq , \vee and \wedge in $\mathcal{L}_{<a}$ the same than in \mathcal{L} and as trivial and empty propositions $\mathbf{1}_{<a} = a$, $\mathbf{0}_{<a} = \emptyset$. Now, given a proposition $b \in \mathcal{L}_{<a}$, one must define what is its complement $b'_{<a}$ in $\mathcal{L}_{<a}$. A natural choice is

$$b'_{<a} = b' \wedge a \quad (4.2.31)$$

but in general with such a choice $\mathcal{L}_{<a}$ is not an orthocomplemented lattice, since it is easy to find for general AC lattices counterexamples such that one may have $b \vee b'_{<a} \neq a$.

Weak-modularity: In order for $\mathcal{L}_{<a}$ to be an orthocomplemented lattice (for any $a \in \mathcal{L}$), the orthocomplemented lattice \mathcal{L} must satisfy the weak-modularity condition

$$b \leq a \implies (a \wedge b') \vee b = a \quad (4.2.32)$$

This condition is also sufficient.

4.2.4.c - Orthomodular lattices

Orthomodularity: An OC lattice which satisfies the weak-modularity condition is said to be an *orthomodular lattice* (or OM lattice)⁶. Clearly if \mathcal{L} is OM, for any $a \in \mathcal{L}$, $\mathcal{L}_{<a}$ is

6. In French: treillis orthomodulaire, in German: Orthomodulare Verband.

also OM, as well as $\mathcal{L}_{>a}$.

Equivalent definitions: Weak-modularity has several equivalent definitions. Here are two interesting ones:

- $a \leq b \implies a$ and b are compatible.
- the orthocomplementation $a \rightarrow a'$ is unique in \mathcal{L} .

Irreducibility: For such lattices one can also define the concept of irreducibility. We have seen that two elements a and b of \mathcal{L} are compatible (or commute) if they generate a Boolean lattice. The center \mathcal{C} of a lattice \mathcal{L} is the set of $a \in \mathcal{L}$ which commute with all the elements of the lattice \mathcal{L} . It is obviously a Boolean lattice. A lattice is irreducible if its center \mathcal{C} is reduced to the trivial lattice $\mathcal{C} = \{\emptyset, \mathbf{1}\}$.

4.2.4.d - Weak-modularity versus modularity

NB: The (somewhat awkward) denomination “weak-modularity” is historical. Following Birkhoff and von Neumann the stronger “modularity” condition for lattices was first considered. Modularity is defined as

$$a \leq b \implies (a \vee c) \wedge b = a \vee (c \wedge b) \quad (4.2.33)$$

Modularity is equivalent to weak modularity for finite depth lattices (as a particular case the set of projectors on a finite dimensional Hilbert space for a modular lattice). But modularity turned out to be inadequate for infinite depth lattices (corresponding to the general theory of projectors in infinite dimensional Hilbert spaces). The theory of modular lattices has links with some W^* -algebras and the theory of “continuous geometries” (see e.g. [vN60]).

4.2.5 Pure states and AC properties

Orthomodular (OM) lattices are a good starting point to consider the constraints that we expect for the set of ideal measurements on a physical system, and therefore to study how one can represent its states. In fact one still needs two more assumptions, which seem technical, but which are also very important (and quite natural from the point of view of quantum information theory). They rely on the concept of atoms, or minimal proposition, which are the analog for propositions of the concept of minimal projectors on of pure states in the algebraic formalism.

4.2.5.a - Atoms

An element a of an OM lattice is said to be an atom if

$$b \leq a \text{ and } b \neq a \implies b = \emptyset \quad (4.2.34)$$

This means that a is a minimal non empty proposition; it is not possible to find another proposition compatible with a which allows to obtain more information on the system than the information obtained if a is found to be TRUE.

Atoms are the analog of projectors on pure states in the standard quantum formalism (pure propositions). Indeed, if the system is in some state ψ , before the measurement of a , if a is an atom and is found to be true, the system will be in a pure state ψ_a after the measure.

4.2.5.b - Atomic lattices

A lattice is said to be atomic if every non trivial proposition $b \neq \emptyset$ in \mathcal{L} is such that there is at least one atom a such that $a \leq b$ (i.e. any proposition “contains” at least one minimal non empty proposition). For an atomic OM lattice one can show that any proposition b is then the union of its atoms (atomisticity).

4.2.5.c - Covering property

Finally one needs also the covering property. The formulation useful in the quantum framework is as follows: An atomic lattice has the covering property if for any proposition a and any atom b not in the complement a' of a , then the Sasaki projection of b onto a , $\Phi_a(b) = a \wedge (b \vee a')$, is still an atom.

The original definition of the covering property for atomic lattices according to Birkhoff is: for any $b \in \mathcal{L}$ and any atom $a \in \mathcal{L}$ such that $a \wedge b = \emptyset$, $a \vee b$ covers b , i.e. there is no c between b and $a \vee b$ such that $b < c < a \vee b$.

This covering property is very important. It means that when reducing a “system” (or rather a set of states) to a “subsystem” by some constraint (projection onto a), one cannot get a non-minimal proposition on the subset out of a minimal one on the larger set. This would mean that one could get more information out of a subsystem than from the greater system. In other words, if a system is in a pure state, performing a perfect measurement can only map it onto another pure state. Perfect measurements cannot decrease the information on the system.

The covering property is in fact also related to the superposition principle. Indeed, it implies that (for irreducible lattices) for any two different atoms a and b , there must be a third atom c different from a and b such that $c < a \vee b$. Thus, in the weakest possible sense (remember we have no addition law yet) c is a superposition of a and b .

An atomistic lattice with the covering property is said to be an AC lattice. As mentioned before these properties can be formulated in terms of the properties of the set of states on the lattice rather than in terms of the propositions. This is the object of the convex set approach. I shall not discuss this here.

4.3 The geometry of orthomodular AC lattices

I have given one (possible and personal) presentation of the principles of the quantum logic formalism. It took some time since I tried to explain both the mathematical formalism and the underlying physical ideas. I now explain the main mathematical result: the definition of the set of propositions (ideal measurements) on a quantum system as an orthomodular AC lattice can be equivalently represented as the set of orthogonal projections on some “generalized Hilbert space”.

4.3.1 Prelude: the fundamental theorem of projective geometry

The idea is to extend a classical and beautiful theorem of geometry, the Veblen-Young theorem. Any abstract projective geometry can be realized as the geometry of the affine subspaces of some left-module (the analog of vector space) on a division ring K (a division ring is a non-commutative generalization of a field like \mathbb{R} or \mathbb{C}). This result is known as the “coordinatization of projective geometry”. Classical references on geometry are the books by E. Artin *Geometric Algebra* [Art57], R. Baer *Linear algebra and projective geometry* [Bae05]. More precisely, a geometry on a linear space is simply defined by a set X whose elements x are called the points of X , and a set \mathcal{L} of lines ℓ of X (at that stage a line ℓ is simply a subset of X , so \mathcal{L} is a subset of the set of subsets of X).

Theorem: If the geometry (X, \mathcal{L}) satisfies the following axioms:

1. Any line contains at least 3 points,
2. Two points lie in a unique line,
3. A line meeting two sides of triangle, not at a vertex of the triangle, meets the third side also (Veblen’s axiom),
4. There are at least 4 points non coplanar (a plane is defined in the usual way from lines),

then the corresponding geometry is the geometry of the affine subspaces of a left module M on a division ring K (a division ring is in general non-commutative field).

Discussion: The theorem here is part of the Veblen-Young theorem, that encompasses the cease when the 4th axiom is not satisfied. The first two axioms define a line geometry structure such that lines are uniquely defined by the pairs of points, but with some superposition principle. The third axiom is represented on Fig. 4.3. The fourth axiom is necessary to exclude some special case of non-Desarguesian geometries, that are of no relevance for our discussion.

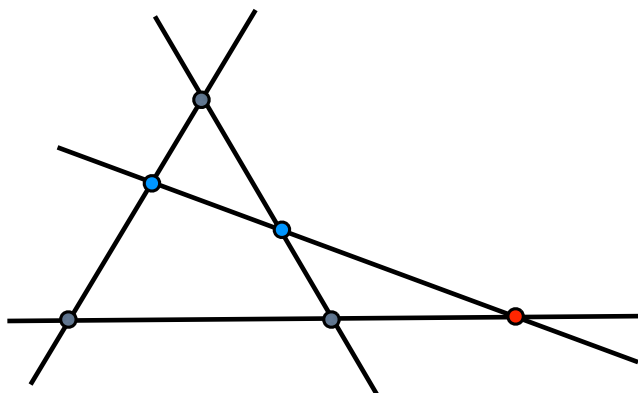


Figure 4.3: Veblen’s axiom

Let me recall that the division ring K (an associative algebra with an addition $+$, a multiplication \times and an inverse $x \rightarrow x^{-1}$) is constructed out of the symmetries of the geometry, i.e. of the automorphisms, or applications $X \rightarrow X$, $\mathcal{L} \rightarrow \mathcal{L}$, etc. which preserve the geometry. Without giving any details, let me illustrate the case of the standard real projective plane (where $K = \mathbb{R}$). The field structure on \mathbb{R} is obtained by identifying \mathbb{R} with a projective line ℓ which contains the three points 0 , 1 and ∞ . The “coordinate” $x \in \mathbb{R}$ of a point $X \in \ell$ is identified with the cross-ratio $x = (X, 1; 0, \infty)$. On Fig. 4.4 are depicted the geometrical construction of the addition $X + Y$ and of the multiplication $X \times Y$ of two points X and Y on a line ℓ .

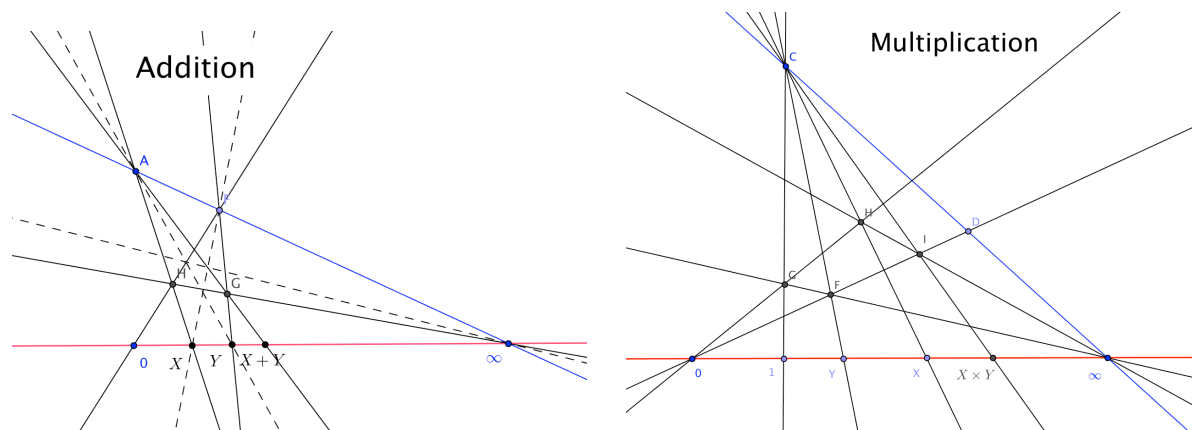


Figure 4.4: Construction of $+$ and \times in the projective real plane

4.3.2 The projective geometry of orthomodular AC lattices

4.3.2.a - The coordinatization theorem

Similar “coordinatization” theorems hold for the orthomodular AC lattices that have been introduced in the previous section. The last axioms AC (atomicity and covering) play a similar role as the axioms of abstract projective geometry, allowing to define “points” (the atoms), lines, etc., with properties similar to the first 3 axioms of linear spaces. The difference with projective geometry is the existence of the orthocomplementation (the negation \neg) which allows to define an abstract notion of orthogonality \perp , and the specific property of weak-modularity (which allows to define in a consistent way what are projections on closed subspaces). Let me first state the main theorem

Theorem (Piron): Let \mathcal{L} be a complete irreducible orthocomplemented AC lattice with length > 3 (i.e. at least three 4 different levels of proposition $\emptyset < a < b < c < d < 1$). Then the “abstract” lattice \mathcal{L} can be represented as the lattice $\mathcal{L}(V)$ of the closed subspaces of a left-module V on a division ring K with a Hermitian form f . The ring K , the module V and the form f have the following properties:

7. A module is the analog of a vector space, but on a ring instead of a (commutative) field
8. A division ring is the analog of a field, but without commutativity

- The division ring K has an involution $*$ such that $(xy)^* = y^*x^*$
- The vector space V has a non degenerate Hermitian (i.e. sesquilinear) form $f: V \times V \rightarrow K$

$$\mathbf{a}, \mathbf{b} \in V, f(\mathbf{a}, \mathbf{b}) = \langle \mathbf{a} | \mathbf{b} \rangle \in K, \quad \langle \mathbf{a} | \mathbf{b} \rangle = \langle \mathbf{b} | \mathbf{a} \rangle^* \quad (4.3.1)$$

- The Hermitian form f defines an orthogonal projection and associates to each linear subspace M of V its orthogonal M^\perp .

$$M^\perp = \{ \mathbf{b} \in V : \langle \mathbf{b} | \mathbf{a} \rangle = 0 \quad \forall \mathbf{a} \in M \} \quad (4.3.2)$$

- The closed subspaces of V are the subspaces M such that $(M^\perp)^\perp = M$.
- The Hermitian form is orthomodular, i.e. for any closed subspace, $M^\perp + M = V$.
- The OM structure $(\preceq, \wedge, ')$ on the lattice \mathcal{L} is isomorphic to the standard lattice structure (\subseteq, \cap, \perp) (subspace of, intersection of, orthogonal complement of) over the space $\mathcal{L}(V)$ of closed linear subspaces of V .
- Moreover, V and K are such that there is some element a of V with “norm” unity $f(a, a) = 1$ (where 1 is the unit element of K).

I do not give the proof. I refer to the physics literature: ([BC81] chapter 21, [Pir64, Pir76]), and to the original mathematical literature [BvN36] [MM71] [Var85].

Thus this theorem states that an OM AC lattice can be represented as the lattice of orthogonal projections over the closed linear subspaces of some “generalized Hilbert space” with a quadratic form defined over some non-commutative field K . This is very suggestive of the fact that Hilbert spaces are not abstract and complicated mathematical objects, much less natural from a physical point of view than the ordinary numbers and functions of classical mechanics (this statement is still found in some discussion and outreach presentations). On the contrary they are quite natural objects if one wants to describe ideal measurements in quantum physics. In particular the algebra (the underlying ring K and the algebraic structure of the space V) come out naturally from the symmetries of the lattice of propositions \mathcal{L} .

4.3.2.b - Discussion: which division ring K ?

The important theorem discussed before is very suggestive, but is not sufficient to “derive” standard quantum mechanics. The main question is which division algebra K and which involution $*$ and Hermitian form f are physically allowed? Can one construct physical theories based on other rings than the usual $K = \mathbb{C}$ (or \mathbb{R} or \mathbb{H})?

There is a large variety of division rings! The simplest one are the finite division rings (with a finite number of elements. The first Wedderburn theorem implies that any finite K is a direct product of Galois fields $\mathbb{F}_p = \mathbb{Z}/\mathbb{Z}_p$ (p prime). Beyond the standard fields \mathbb{C} , \mathbb{R} and \mathbb{H} , more complicated division rings are the rings of rational functions $F(X)$, and some are even larger (for instance the surreal numbers...), but still commutatives, and some are also fully non-commutatives rings.

However, the requirements that K has an involution, and that V has a non degenerate hermitian form, so that $\mathcal{L}(V)$ is a OM lattice, put already very stringent constraints on K . For instance, it is well known that finite fields like the \mathbb{F}_p (p prime) do not work.

Indeed, it is easy to see that the lattice $\mathcal{L}(V)$ of the linear subspaces of the finite dimensional subspaces of the n -dimensional vector space $V = (\mathbb{F}_p)^n$ is not orthomodular and cannot be equipped with a non-degenerate quadratic form! Check with $p = n = 3$! But still many more exotic division rings K than the standard \mathbb{R}, \mathbb{C} (and \mathbb{H}) are possible at that stage .

4.3.3 Towards Hilbert spaces

There are several arguments that point towards the standard physical solution, V must be a Hilbert space over \mathbb{R}, \mathbb{C} (or possibly \mathbb{H}). However none of these arguments is fully mathematically convincing, if most would satisfy a physicist. Remember that real numbers are expected to occur in physical theory for two reasons. Firstly the lattice structure is related to probabilities p , which are real numbers. Thus the continuous structure of the real numbers should be contained in the abstract structure (this is not really a very strong argument...) Secondly we look for a quantum theory that must be compatible with the relativistic concept of space-time, where space and time are described by continuous real variables (this is an experimental fact). Of course this is correct as long as one does not try to quantize gravity.

We have not discussed yet precisely the structure of the states ψ , and which constraints they may enforce on the algebraic structure of propositions. Remember that it is the set of states \mathcal{E} which allows to discuss the partial order relation \leq on the set of propositions \mathcal{L} . Moreover states ψ assign probabilities $\psi(a) \in [0, 1]$ to propositions a , with the constraints that if $a \perp b$, $\psi(a \vee b) = \psi(a) + \psi(b)$. Moreover the propositions $a \in \mathcal{L}$ (projective measurements) define via the Sasaki orthogonal projections π_a a set of transformations $\mathcal{L} \rightarrow \mathcal{L}$, which form a so called Baer $*$ -semi group. On the same time, propositions $a \in \mathcal{L}$ define mappings $\psi \rightarrow \psi_a$ on the states. Since, as in the algebraic formalism, convex linear combinations of states are states, \mathcal{E} generate a linear vector space E , and form a convex subset $\mathcal{E} \subset E$. Thus there is more algebraic structure to discuss than what I explained up to now. I refer to [BC81], chapters 16-19, for more details. I shall come back to states when discussing Gleason's theorem in the next section.

Assuming some "natural" continuity or completeness conditions for the states leads to theorems stating that the division ring K must contain the field of real numbers \mathbb{R} , hence is \mathbb{R}, \mathbb{C} or \mathbb{H} , and that the involution $*$ is continuous, hence corresponds to the standard involution $x^* = \bar{x}$, $x^* = \bar{x}$ or $x^* = x^*$ respectively. See [BC81], chapter 21.3.

Another argument comes from an important theorem in the theory of orthomodular lattices, which holds for lattices of projections in infinite dimensional modules.

Solèr's Theorem: (Solèr 1995) Let $\mathcal{L} = \mathcal{L}_K(V)$ be an irreducible OM AC lattice of compact linear subspaces in a left-module V over a division ring K , as discussed above. If there is an infinite family $\{v_i\}$ of orthonormal vectors in V such that $\langle v_i | v_j \rangle = \delta_{ij} f$ with some $f \in K$ then the division ring K can only be \mathbb{R}, \mathbb{C} or \mathbb{H} .

The proof of this highly non trivial theorem is given in [Sol95]. It is discussed in more details in [Hol95].

The assumptions of the theorem state that there an infinite set of mutually compatible atoms $\{a_i\}_{i \in I}$ in \mathcal{L} (commuting, or causally independent elementary propositions a_i), and in addition that there is some particular symmetry between the generators $v_i \in V$ of the linear spaces (the lines or rays) of these propositions.

The first assumption is quite natural if we take into account space-time and locality in quantum physics. Let me consider the case where the physical space in which the system is defined to be infinite (flat) space or some regular lattice, so that it can be separated into causally independent pieces \mathcal{O}_α (labelled by $\alpha \in \Lambda$ some infinite lattice). See for instant Fig. 4.5. It is sufficient to have one single proposition a_α relative to each \mathcal{O}_α only (for instance “there is one particle in \mathcal{O}_α ”) to build an infinite family of mutually orthogonal propositions $b_\alpha = a_\alpha \wedge (\bigwedge_{\beta \neq \alpha} \neg a_\beta)$ in \mathcal{L} . Out of the b_α , thanks to the atomic property (A), we can extract an infinite family of orthogonal atoms c_α .

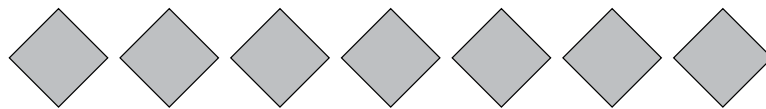


Figure 4.5: A string of causal diamonds (in space-time)

However this does not ensure the second assumption: the fact that the corresponding $v_i \in V$ are orthonormals. The group of space translations \mathfrak{T} must act as a group of automorphism on the lattice of linear subspaces $\mathcal{L} = \mathcal{L}_K(V)$ (a group of automorphisms on a OC lattice $\mathcal{L} = \mathcal{L}_K(V)$ is a group of transformations which preserves the OC lattice structure $(\leq, \wedge, ')$ or equivalently (\subseteq, \cap, \perp)). There must correspond an action (a representation) of the translation group \mathfrak{T} on the vector space V , and on the underlying field K . If the action is trivial the conditions of Soler’s theorem are fulfilled, but this is not ensured a priori. See for instance [GL12] for a recent discussion of symmetries in orthomodular geometries. However I am not aware of a counterexample where a non standard orthomodular geometry (i.e. different from that of a Hilbert space on \mathbb{C} (or \mathbb{R}) carries a representation of a “physical” symmetry group such as the Poincaré or the Galilean group of space-time transformations (representations of these groups should involve the field of real numbers \mathbb{R} in some form).

From now on we assume that a quantum system may indeed be described by projectors in a real or complex Hilbert space.

One last remark. The coordinatization theorem depends crucially on the fact that the OM lattice \mathcal{L} is atomic, hence contain minimal propositions (atoms). They are the analog of minimal projectors in the theory of operator algebras. Hence the formalism discussed here is expected to be valid mathematically to describe only type I von Neumann algebras. I shall not elaborate further.

4.4 Gleason’s theorem and the Born rule

4.4.1 States and probabilities

In the presentation of the formalism we have not put emphasis on the concept of states, although states are central in the definition of the causality order relation

\leq and of the orthocomplementation $'$. We recall that to each state ψ and to each proposition a is associated the probability $\psi(a)$ for a to be found true on the state ψ . In other words, states are probability measures on the set of propositions, compatible with the causal structure. As already mentioned, the lattice structure of propositions can be formulated from the properties of the states on \mathcal{L} .

At that stage we have almost derived the standard mathematical formulation of quantum mechanics. Propositions (yes-no observables) are represented by orthogonal projectors on a Hilbert space \mathcal{H} . Projectors on pure states correspond to projectors on one-dimensional subspaces, or rays of \mathcal{H} so the concept of pure states is associated to the vectors of \mathcal{H} .

Nevertheless it remains to understand which are the consistent physical states, and what are the rules which determine the probabilities for a proposition a to be true in a state ψ , in particular in a pure state. We remind that the states are in fact characterized by these probability distributions $a \rightarrow \psi(a)$ on \mathcal{L} . Thus states must form a convex set of functions $\mathcal{L} \rightarrow [0, 1]$ and by consistency with the OM structure of \mathcal{L} they must satisfy four conditions. These conditions define "quantum probabilities"

Quantum probabilities:

$$(1) \quad \psi(a) \in [0, 1] \quad (4.4.1)$$

$$(2) \quad \psi(\emptyset) = 0, \quad \psi(\mathbf{1}) = 1 \quad (4.4.2)$$

$$(3) \quad a \neq b \implies \exists \psi \text{ such that } \psi(a) \neq \psi(b) \quad (4.4.3)$$

$$(4) \quad a \perp b \implies \psi(a \vee b) = \psi(a) + \psi(b) \quad (4.4.4)$$

Conditions (1) and (2) are the usual normalization conditions for probabilities. Condition (3) means that observables are distinguishable by their probabilities. Condition (4) is simply the fact that if a and b are orthogonal, they generate a Boolean algebra, and the associated probabilities must satisfy the usual sum rule. These conditions imply in particular that for any state ψ , $\psi(-a) = 1 - \psi(a)$, and that if $a \leq b$, then $\psi(a) \leq \psi(b)$, as we expect.

It remains to understand if and why all states ψ can be represented by density matrices ρ_ψ , and the probabilities for propositions a given by $\psi(a) = \text{tr}(\rho_\psi P_a)$, where P_a is the projector onto the linear \mathcal{L} -subspace associated to the proposition a . This is a consequence of a very important theorem in operator algebras, Gleason's theorem [Gle57].

4.4.2 Gleason's theorem

It is easy to see that to obtain quantum probabilities that satisfy the conditions 4.4.1- 4.4.4, it is sufficient to consider atomic propositions, i.e. projections onto 1-dimensional subspaces (rays) generated by vectors $\vec{e} = |e\rangle$ (pure states) of the Hilbert space \mathcal{H} . Indeed, using 4.4.4, the probabilities for general projectors can be reconstructed (by the usual sum rule) from the probabilities for projections on rays. Denoting since there is no ambiguity for a state ψ the probability for the atomic proposition e represented by the projection $P_{\vec{e}}$ onto a vector $\vec{e} = |\vec{e}\rangle \in \mathcal{H}$ as

$$\psi(e) = \psi(P_{\vec{e}}) = \psi(\vec{e}) \quad P_{\vec{e}} = |\vec{e}\rangle\langle\vec{e}| = \text{projector onto } \vec{e} \quad (4.4.5)$$

The rules 4.4.1- 4.4.4 reduces for atomic propositions to the conditions.

Quantum probabilities for projections on pure states: For any state ψ , the function $\psi(\vec{e})$ considered as a function on the “unit sphere” of the rays over the Hilbert Space \mathcal{H} (the projective space) $\mathcal{S} = \mathcal{H}^*/K^*$ must satisfy

$$(1) \quad \psi(\vec{e}) = \psi(\lambda\vec{e}) \quad \text{for any } \lambda \in K \text{ such that } |\lambda| = 1 \quad (4.4.6)$$

$$(2) \quad 0 \leq \psi(\vec{e}) \leq 1 \quad (4.4.7)$$

$$(3) \quad \text{For any complete orthonormal basis of } \mathcal{H}, \{\vec{e}_i\}, \text{ one has } \sum_i \psi(\vec{e}_i) = 1 \quad (4.4.8)$$

Gleason’s theorem states the fundamental result that any such function is in one to one correspondence with a density matrix.

Gleason’s theorem:

If the Hilbert space \mathcal{H} over $K = \mathbb{R}$ or \mathbb{C} is such that

$$\dim(\mathcal{H}) \geq 3 \quad (4.4.9)$$

then any function ψ over the unit rays of \mathcal{H} that satisfies the three conditions 4.4.6–4.4.8 is of the form

$$\psi(\vec{e}) = (\vec{e} \cdot \rho_\psi \cdot \vec{e}) = \langle \vec{e} | \rho_\psi | \vec{e} \rangle \quad (4.4.10)$$

where ρ_ψ is a positive quadratic form (a density matrix) over \mathcal{H} with the expected properties for a density matrix

$$\rho_\psi = \rho_\psi^\dagger, \quad \rho_\psi \geq 0, \quad \text{tr}(\rho_\psi) = 1 \quad (4.4.11)$$

Reciprocally, any such quadratic form defines a function ψ with the three properties 4.4.6–4.4.8.

Gleason’s theorem is fundamental. As we shall discuss more a bit later, it implies the Born rule. It is also very important when discussing (and excluding a very general and most natural class of) hidden variables theories. So let us discuss it a bit more, without going into the details of the proof.

4.4.3 Principle of the proof

The theorem is remarkable since there are non conditions on the regularity or measurability of the function ψ . In the original derivation by Gleason [Gle57] he considers real “frame functions” f of weight W over $\mathcal{H}^* = \mathcal{H} \setminus \{0\}$ such that

$$(1) \quad f(\vec{e}) = f(\lambda\vec{e}) \quad \text{for any } \lambda \neq 0 \in K \quad (4.4.12)$$

$$(2) \quad f \text{ is bounded} \quad (4.4.13)$$

$$(3) \quad \text{For any complete orthonormal basis of } \mathcal{H}, \{\vec{e}_i\}, \sum_i f(\vec{e}_i) = W = \text{constant} \quad (4.4.14)$$

and proves that such a function must be of the form 4.4.10

$$f(\vec{e}) = (\vec{e} \cdot Q \cdot \vec{e}) \quad , \quad Q \quad \text{quadratic form such that} \quad \text{tr}(Q) = W \quad (4.4.15)$$

It is easy to see that this is equivalent to the theorem as stated above, since one can add constants and rescale the functions f to go from 4.4.12–4.4.14 to 4.4.6–4.4.8. The original proof goes into three steps

1. Real Hilbert space, $\dim(\mathcal{H}) = 3$ and f a continuous frame function \implies the theorem

This is the easiest part, involving some group theory. Any frame function f is a real function on the unit two dimensional sphere \mathcal{S}_2 and if continuous it is square summable and can be decomposed into spherical harmonics

$$f(\vec{n}) = \sum_{l,m} f_{l,m} Y_l^m(\theta, \varphi) \quad (4.4.16)$$

The theorem amounts to show that if f is a frame function of weight $W = 0$, then only the $l = 2$ components of this decomposition 4.4.16 are non zero. Some representation theory (for the $SO(3)$ rotation group) is enough. Any orthonormal (oriented) basis $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ of \mathbb{R}^3 is obtained by applying a rotation R to the basis $(\vec{e}_x, \vec{e}_y, \vec{e}_z)$. Thus one can write

$$f(\vec{n}_1) + f(\vec{n}_2) + f(\vec{n}_3) = \sum_l \sum_{m,m'} f_{l,m} D_{m,m'}^{(l)}(R) V_m^{(l)} \quad (4.4.17)$$

with the $D_{m,m'}^{(l)}(R)$ the Wigner D matrix for the rotation R , and the $V_m^{(l)}$ the components of the vectors $\vec{V}^{(l)}$ in the spin l representation of $SO(3)$, with components

$$\vec{V}^{(l)} = \{V_m^{(l)}\} \quad , \quad V_m^{(l)} = Y_{l,m}(0,0) + Y_{l,m}(\pi/2,0) + Y_{l,m}(\pi/2,\pi/2) \quad (4.4.18)$$

If f is a frame function of weight $W = 0$, the l.h.s. of 4.4.17 is zero for any $R \in SO(3)$. This implies that for a given l , the coefficients $f_{l,m}$ must vanish if the vector $\vec{V}^{(l)} \neq 0$, but are undetermined if $\vec{V}^{(l)} = 0$. An explicit calculation shows that indeed

$$\vec{V}^{(l)} \begin{cases} \neq 0 & \text{if } l \neq 2, \\ = 0 & \text{if } l = 2. \end{cases} \quad (4.4.19)$$

This establishes the theorem in case (1).

2. Real Hilbert space, $\dim(\mathcal{H}) = 3$ and f any frame function $\implies f$ continuous.

This is the most non-trivial part: assuming that the function is bounded, the constraint 4.4.14 is enough to imply that the function is continuous! It involves a clever use of spherical geometry and of the frame identity $\sum_{i=1,2,3} f(\vec{e}_i) = W$.

The basic idea is to start from the fact that since f is bounded, it has a lower bound f_{min} which can be set to 0. Then for any $\epsilon > 0$, take a vector \vec{n}_0 on the sphere such that $|f(\vec{n}_0) - f_{min}| < \epsilon$. It is possible to show that there is a neighbourhood \mathcal{O} of \vec{n}_0 such that $|f(\vec{n}_1) - f(\vec{n}_2)| < C\epsilon$ for any \vec{n}_1 and $\vec{n}_2 \in \mathcal{O}$. C is a universal constant. It follows that the function f is continuous at its minimum! Then it is possible, using rotations to show that the function f is continuous at any points on the sphere.

3. Generalize to $\dim(\mathcal{H}) > 3$ and to complex Hilbert spaces.

This last part is more standard and more algebraic. Any frame function $f(\vec{n})$ defined on unit vectors \vec{n} such that $\|\vec{n}\| = 1$ may be extended to a quadratic function over vectors $f(\vec{v}) = \|\vec{v}\|^2 f(\vec{v}/\|\vec{v}\|)$.

For a real Hilbert space with dimension $d > 3$, the points (1) and (2) implies that the restriction of a frame function $f(\vec{n})$ to any 3 dimensional subspace is a quadratic form $f(\vec{v}) = (\vec{v} \cdot Q \cdot \vec{v})$. A simple and classical theorem by Jordan and von Neumann shows that this is enough to define a global real quadratic form Q on the whole Hilbert space \mathcal{H} through the identity $2(\vec{x} \cdot Q \cdot \vec{y}) = f(\vec{x} + \vec{y}) - f(\vec{x} - \vec{y})$. For complex Hilbert spaces, the derivation is a bit more subtle. One can first apply the already obtained results to the restriction of frame functions over real submanifolds of \mathcal{H} (real submanifolds are real subspaces of \mathcal{H} such that $(\vec{x} \cdot \vec{y})$ is always real). One can then extend the obtained real quadratic form over the real submanifolds to a complex quadratic form on \mathcal{H} .

4.4.4 The Born rule

The Born rule is a simple consequence of Gleason theorem. Indeed, any state (in the general sense of statistical state) corresponds to a positive quadratic form (a density matrix) ρ . Given a minimal atomic proposition, which corresponds to a projector $P = |\vec{a}\rangle\langle\vec{a}|$ onto the ray corresponding to a single vector (pure state) $|\vec{a}\rangle$, the probability p for P of being true is

$$p = \langle P \rangle = \text{tr}(\rho P) = \langle \vec{a} | \rho | \vec{a} \rangle \quad (4.4.20)$$

The space of states \mathcal{E} is thus the space of (symmetric) positive density matrices with unit trace

$$\text{space of states} = \mathcal{E} = \{ \rho : \rho = \rho^\dagger, \rho \geq 0, \text{tr}(\rho) = 1 \} \quad (4.4.21)$$

It is a convex set. Its extremal points, which cannot be written as a linear combination of two different states, are the pure states of the system, and are the density matrices of rank one, i. e. the density matrices which are themselves projectors onto a vector $|\psi\rangle$ of the Hilbert space.

$$\rho = \text{pure state} \implies \rho = |\psi\rangle\langle\psi|, \quad \|\psi\| = 1 \quad (4.4.22)$$

One thus derives the well known fact that the pure states are in one to one correspondence with the vectors (well... the rays) of the Hilbert space \mathcal{H} that was first introduced from the basic observables of the theory, the elementary atomic propositions (the projectors P). Similarly, one recovers the simplest version of the Born rule: the probability to “observe” a pure state $|\varphi\rangle$ into the (non orthogonal) pure state $|\psi\rangle$ (to “project” $|\varphi\rangle$ onto $|\psi\rangle$) is the square of the norm of the scalar product

$$p(\varphi \rightarrow \psi) = |\langle \varphi | \psi \rangle|^2 \quad (4.4.23)$$

4.4.5 Physical observables

One can easily reconstruct the set of all physical observables, and the whole algebra of observables \mathcal{A} of the system. I present the line of the argument, without any attempt of mathematical rigor.

Any ideal physical measurement of some observable O consists in fact in taking a family of mutually orthogonal propositions a_i , i.e. of commuting symmetric projectors P_i on \mathcal{H} such that

$$P_i^2 = P_i, \quad P_i = P_i^*, \quad P_i P_j = P_j P_i = 0 \text{ if } i \neq j \quad (4.4.24)$$

performing all the tests (the order is unimportant since the projectors commute) and assigning a *real number* o_i to the result of the measurement (the value of the observable O) if a_i is found true (this occurs for at most one a_i) and zero otherwise. In fact one should take an appropriate limit when the number of a_i goes to infinity, but I shall not discuss these important points of mathematical consistency. If you think about it, this is true for any imaginable measurement (position, speed, spin, energy, etc.). The resulting physical observable O is thus associated to the symmetric operator

$$O = \sum_i o_i P_i \quad (4.4.25)$$

This amounts to the spectral decomposition of symmetric operators in the theory of algebras of operators.

Consider a system in a general state given by the density matrix ρ . From the general rules of quantum probabilities, the probability to find the value o_i for the measurement of the observable O is simply the sum of the probabilities to find the system in a eigenstate of O of eigenvalue o_i , that is

$$p(O \rightarrow o_i) = \text{tr}(P_i \rho) \quad (4.4.26)$$

and for a pure state $|\varphi\rangle$ it is simply

$$\langle \varphi | P_i | \varphi \rangle = |\langle \varphi | \varphi_i \rangle|^2, \quad |\varphi_i\rangle = \frac{1}{\|P_i |\varphi\rangle\|} P_i |\varphi\rangle \quad (4.4.27)$$

Again the Born rule! The expectation value for the result of the measurement of O in a pure state ψ_a is obviously

$$\mathbb{E}[O; \psi] = \langle O \rangle_\psi = \sum_i o_i p(O \rightarrow o_i) = \sum_i o_i \langle \psi | P_i | \psi \rangle = \langle \psi | O | \psi \rangle \quad (4.4.28)$$

This is the standard expression for expectation values of physical observables as diagonal matrix elements of the corresponding operators. Finally for general (mixed) states one has obviously

$$\mathbb{E}[O; \rho] = \langle O \rangle_\rho = \sum_i o_i p(O \rightarrow o_i) = \sum_i o_i \text{tr}(P_i \rho) = \text{tr}(O \rho) \quad (4.4.29)$$

We have seen that the pure states generate by convex combinations the convex set \mathcal{E} of all (mixed) states ψ of the system. Similarly the symmetric operators $O = O^\dagger$ generates (by operator multiplication and linear combinations) a C^* -algebra \mathcal{A} of bounded operators $\mathcal{B}(\mathcal{H})$ on the Hilbert space \mathcal{H} . States are normalized positive linear forms on \mathcal{A} and we are back to the standard algebraic formulation of quantum physics. The physical observables generates an algebra of operators, hence an abstract algebra of observables, as assumed in the algebraic formalism. We refer to the section about the algebraic formalism for the arguments for preferring complex Hilbert spaces to real or quaternionic ones.

4.5 Discussion

This chapter was a sketchy and partial introduction to the quantum logic approach for the formulation of the principle of quantum mechanics. I hope to have shown its relation with the algebraic formulation. It relies on the concepts of states and of observables as the algebraic formulation. However the observables are limited to the physical subset of yes/no proposition, corresponding to ideal projective measurements, without assuming a priori some algebraic structure between non-compatible propositions (non-commuting observables in the algebraic framework). I explained how the minimal set of axioms on these propositions and their actions on states, used in the quantum logic approach, is related to the physical concepts of causality, reversibility and separability/locality. The canonical algebraic structure of quantum mechanics comes out from the symmetries of the “logical structure” of the lattice of propositions. The propositions corresponding to ideal projective measurements are realized on orthogonal projections on a (possibly generalized) Hilbert space. Probabilities/states are given by quadratic forms, and the Born rule follows from the logical structure of quantum probabilities through Gleason’s theorem.

The quantum logic approach is of course not completely foolproof. We have seen that the issue of the possible division ring K is not completely settled. The strong assumptions of atomicity and covering are essential, but somehow restrictive compared to the algebraic approach (type II and III von Neumann algebras). It is sometime stated that it cannot treat properly the case of a system composed of two subsystems since there is no concept of “tensorial product” of two OM-AC lattices as there is for Hilbert spaces and operator algebras. Note however that one should in general always think about multipartite systems as parts of a bigger system, not the opposite! Even in the algebraic formulation it is not known in the infinite dimensional case if two commuting subalgebras \mathcal{A}_1 and \mathcal{A}_2 of a bigger C^* -algebra \mathcal{A} always correspond to the decomposition of the Hilbert space \mathcal{H} into a tensor product of two subspaces \mathcal{H}_1 and \mathcal{H}_2 (this is known as the Tsirelson conjecture).