

# On the Algebraic Structure of Quantum Mechanics

J. GUNSON

Department of Mathematical Physics  
University of Birmingham, England

Received June 1, 1967

**Abstract.** We present a reformulation of the axiomatic basis of quantum mechanics with particular reference to the manner in which the usual algebraic structures arise from certain natural physical requirements. Care is taken to distinguish between features of physical significance and those introduced for mathematical convenience. Our conclusion is that the usual algebraic structures cannot be significantly generalised without conflicting with our current experimental picture of processes occurring at the quantum level.

## 1. Introduction

The past few years have seen a widespread renewal of interest in the old problem of just why the particular algebraic structure found in quantum mechanics seems to work so well. One question we may ask is: what natural physical requirements can we find to explain the rather remarkable algebraic properties found in the usual quantum mechanical formalisms, in which self-adjoint operators, representing observable quantities, etc., act upon a Hilbert space of states? Alternatively, we may ask whether yet more general formalisms are possible, still being physically sensible? It may be that a definitive answer to these questions is not possible, in view of the difficulty in deciding just what constitutes a “natural” physical requirement.

The deepest results obtained in earlier attacks on the problem were obtained by VON NEUMANN and coworkers [1, 2, 3]. Two main approaches were used. These we may call the “Jordan algebra” and the “propositional calculus” approaches. In both of these, an attempt was made to work from a set of axioms, possessing as much direct physical significance as possible. However, it was found that the usual quantum mechanical formalism was obtained only at the expense of introducing axioms whose physical significance was far from apparent. This is clearly stated by the authors. In the Jordan algebra approach, the point in question is the *distributivity* axiom without which “an algebraic discussion is scarcely possible” [2]. In the other approach, the relevant axiom is that of the modularity of the lattice of propositions which property is “closely related to the existence of an “a priori thermodynamic weight of states” [3].

In this paper we present a reformulation of these ideas, containing some features of both the above approaches, but whose 'critical' axioms are quite different. Our main result is a set of sufficient conditions (axioms A.1 to A.12) for the usual quantum mechanical algebraic structures to result. Several of these axioms are also necessary conditions. The direct physical significance possessed by these axioms has led us to conclude that no further generalization of the usual quantum mechanical formalism is at present indicated.

At this point we would like to mention other attempts in the same direction. SEGAL [4] has constructed a system which allows non-distributive algebras of observables [5, 6], which we regard as too general. Perhaps the most important result of this investigation is the emphasis on the need for a purely algebraic (i.e. independent of any particular representation in Hilbert space) formulation of quantum mechanics. This seems particularly relevant for the quantum theory of fields, where unitarily inequivalent representations are abundant. Consequently, we have tried to make our approach as algebraic as possible. In a published set of lecture notes, MACKEY [7] has given an excellent discussion of the problem. On the basis of axioms of direct physical significance, he shows that the propositional calculus takes the form of an orthomodular partially ordered set. In lieu of a solution to the problem, he restricts the discussion to the usual quantum mechanical formalism by directly postulating the well-known structure, in his axiom VII. Most subsequent investigations, notably of ZIERLER [8, 9] and VARADARAJAN [10], have been based on the orthomodularity of the calculus of propositions. Recently, PIRON [11, 12] has made notable progress, and POOL's thesis [13] contains a detailed exposition of the basic principles.

We have attempted to make this paper as self-contained as possible, so that at each stage in the development we can distinguish between those features that are of essential physical significance and those that are introduced only for mathematical convenience. Thus our approach is as 'finitistic' as possible, though we do not restrict our development to finite-dimensional algebras and vector spaces. This freedom is essential if the usual Heisenberg commutation rules are to be permitted.

In section 2, we summarize briefly the basic concepts of state, logic, compatibility and orthomodularity. On this basis, the first seven axioms are constructed. We distinguish between pure and mixed states in section 3, and construct some very useful linear vector spaces with certain topology and order properties of direct physical significance. In section 4, we introduce the three critical axioms A.10 to A.12 and discuss their physical justification. These are then used in section 5 to derive our main result which is the equivalent, in our approach, of MACKEY's axiom VII [7]. To quote from the latter "the partially ordered

set of all questions in quantum mechanics is isomorphic to the partially ordered set of all closed subspaces of a separable infinite dimensional Hilbert space." This needs to be modified slightly if we allow super-selection rules (cf. appendix, Ref. [7]).

## 2. States and Logics

### *a) Basics*

In this section we summarize moderately well-known results in order to introduce notations and basic definitions. For a more detailed account, we recommend MACKAY's book [7].

In the basic investigations of BIRKHOFF and VON NEUMANN, the important notion of a propositional calculus associated with a given physical theory was introduced. This is an abstract formulation of the logical content of any theory in which the measuring process can be reduced to a set of experiments of the elementary true-or-false type. The elements of the calculus, taken to be in one-to-one correspondence with such equivalence classes of experiments, form an abstract set  $\mathcal{P}$  on which we impose an order relation " $\leq$ ", induced by the logical relation

$$a \leq b \Leftrightarrow a \rightarrow b \quad (2.1)$$

i.e., whenever an observation on a particular system leads to the values 'true' for  $a$ , then  $b$  gives 'true'. In the case of a classical mechanical system, whose state at a particular time may be represented by a point in phase space, the propositional calculus may be identified with the set of all Borel subsets of phase space ordered by set inclusion. The proposition associated with a given set is that which gives the value 'true' on systems whose representative point lies in the set, otherwise giving the value 'false'.

The dual natures of the concepts of state and proposition are apparent in the classical case. From an operational viewpoint, the definition of either depends in an essential way on the other one, in that a state is *defined* by the set of values taken on the set of all propositions and distinguished from other states only by this means. Conversely, propositions are defined in terms of the set of all states. This features seems unescapable in any theory directly related to experiment, leading us to formulate the first set of axioms as

**Axiom A.1.**  $f(a) = f(b)$  for all  $f \in \mathcal{S} \Leftrightarrow a = b$ ;  $a, b \in \mathcal{P}$ .

**Axiom A.2.**  $\exists \Phi \in \mathcal{P}$  such that  $f(\Phi) = 0$  for all  $f \in \mathcal{S}$ .

**Axiom A.3.**  $f(a) = g(a)$  for all  $a \in \mathcal{P} \Leftrightarrow f = g$ ;  $f, g \in \mathcal{S}$ .

In accordance with convention, we assign the numerical values 0 and 1 to the logical values 'false' and 'true'. The example of quantum mechanics shows that this set of values is too restrictive; we must allow all values between 0 and 1, i.e., a state is a probability measure on the pro-

positions. In the classical case such a state may be regarded as describing our information concerning a random member of an ensemble of systems with a given probability distribution on phase space, thus introducing an element of subjectivity into the concept of state. According to the usual interpretations, this latter feature is an essential part of quantum mechanics. In axiom A.2, the *absurd* proposition  $\Phi$  is introduced, for which every value is false.

The next pair of axioms define the partial ordering of  $\mathcal{P}$  in terms of the usual ordering of the reals.

**Axiom A.4.** *For every  $a, b \in \mathcal{P}$ , we have  $a \leq b$  if and only if  $f(a) \leq f(b)$  for all  $f \in \mathcal{S}$ .*

**Axiom A.5.** *For every  $a \in \mathcal{P}$ , there is an element  $a' \in \mathcal{P}$ , such that  $f(a) + f(a') = 1$  for all  $f \in \mathcal{S}$ .*

The axiom A.4 is an extension of ordering by implication chosen so as to accord with the classical case, when single systems are replaced by ensembles with general probability distributions. The operation of *negation* in the logic is likewise expressed in axiom A.5, again being the obvious form consistent with classical theory. The properties of the involutory order-reversing mapping  $a \rightarrow a'$  are

$$\begin{aligned} (a')' &= a & a \wedge a' &= \Phi \\ (a \vee a') &= \Phi' & (a \vee b)' &= a' \wedge b' . \end{aligned} \tag{2.2}$$

The *trivial* proposition  $\Phi'$  is usually denoted by  $I$  and called the *unit* element. The *join* ( $\vee$ ) and *meet* ( $\wedge$ ) symbols denote as usual the least upper and greatest lower bounds (when these exist) in the above order relation. In the classical case, the propositional calculus is a *Boolean lattice*, and orthocomplemented distributive lattice, in which join and meet exist for every pair of elements. There is, in general, no good reason for assuming that this lattice property always holds. However, there is at least one case in which join and meet should exist, the case of *orthogonal* elements, defined in

$$a \perp b \Leftrightarrow a \leq b' \tag{2.3}$$

where the perpendicularity symbol is used to indicate the orthogonality of  $a$  and  $b$ . This again is based in the example occurring in classical mechanics of *disjoint* subsets of phase space (and coincides in quantum mechanics with the property of orthogonality of certain pairs of closed subspaces of Hilbert space). The next axiom expresses the essential properties of this relation into our system

**Axiom A.6.** (*Orthogonality Axiom*). *If  $\{a_i; i = 1, 2, \dots, n\}$  is a set of pairwise orthogonal elements of  $\mathcal{P}$ , then there is an element  $a$  of  $\mathcal{P}$  (called the sum of the  $a_i$ ) such that*

$$f(a) = \sum_{i=1}^n f(a_i) \quad \text{all } f \in \mathcal{S}$$

defining the join  $a = a_1 \vee a_2 \vee \dots \vee a_n$ .

The restrictions imposed on  $\mathcal{P}$  by the axioms stated so far are summarized in the following theorem [7, 8, 10].

**Theorem 2.1.**  $\mathcal{P}$  is an orthomodular partially ordered set, i.e. the complementation  $a \rightarrow a'$  satisfies (2.2) and

$$a \leq b \Rightarrow (b \wedge a') \vee a \text{ exists and equals } b. \quad (2.4)$$

Equivalent useful versions of this weak modularity condition are (for proofs in the lattice case see LOOMIS [14])

$$a \leq b \Rightarrow a = b \wedge (a \vee b') \quad (2.5)$$

$$a \leq b \text{ and } b \wedge a' = 0 \Rightarrow a = b. \quad (2.6)$$

### b) Compatibility and Incompatibility

The essential distinction between the proportional calculi of the classical and quantum theories was pointed out by BIRKHOFF and VON NEUMANN [3]: the occurrence of incompatible measuring processes in quantum mechanics implies that the propositional calculus cannot be the Boolean lattice, as with classical deterministic theories. In particular, the distributive laws

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c) \quad (2.7)$$

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c) \quad (2.8)$$

cannot hold for every triplet of elements in  $\mathcal{P}$ , even when each side of the expressions exist. It is inherent in the construction of classical theories that for any state  $f$  of a single system and any pair of propositions  $a, b$ , we have  $f(a)$  and  $f(b)$  simultaneously taking one of the values true or false. In the case of probability distributions, this becomes

$$f(a) + f(b) = f(a \vee b) + f(a \wedge b). \quad (2.9)$$

In the case of a general calculus  $\mathcal{P}$ , if  $a \vee b$  and  $a \wedge b$  exist and (2.9) is satisfied, we call  $a$  and  $b$  *compatible* [11, 12], and denote the relation by  $a \leftrightarrow b$ . Various modes of expression of the compatibility relation are

$$a \leftrightarrow b \Leftrightarrow (a \wedge b) \vee b' \cong a \quad (2.10)$$

$$a \leftrightarrow b \Leftrightarrow (a \wedge b') \vee b = (b \wedge a') \vee a \quad (2.11)$$

$$a \leftrightarrow b \Leftrightarrow \exists c, d, e, \text{ with (i) } a = c \vee e, \text{ (ii) } b = d \vee e, \quad (2.12)$$

$$\text{(iii) } e \perp c, \text{ (iv) } e \perp d, \text{ (v) } c \perp d.$$

A proof of the equivalence of (2.11) and (2.12) (for lattices) has been given by FOULIS [15]. That (2.10) is equivalent to (2.12) may be seen as follows. Being given (2.10), define

$$e = a \wedge b \quad c = e' \wedge a \quad d = e' \wedge b. \quad (2.13)$$

The element  $e$  exists by assumption and the existence of  $c$  and  $d$  follows from a simple application of axiom A.6. We have then to verify the properties (i) to (v) of (2.12). Using Theorem 2.1,

- (i)  $e \leq a, a = (a \wedge e') \vee e = c \vee e,$       (ii) likewise,
- (iii)  $e \leq e \vee a' = c',$  i.e.,  $e \perp c,$       (iv) likewise,
- (v) by hypothesis,  $c = e' \wedge a \leq a \leq (a \wedge b) \vee b' = e \vee b' = d'$

giving  $c \perp d$ . Conversely,  $d \perp a \Rightarrow a \leq d' = (a \wedge b) \vee b'$ . Q.E.D. Finally (2.9) is an immediate consequence of (2.12) and from (2.9) we can obtain the relation  $a \vee (a' \wedge b') = b' \vee (a \wedge b)$  which implies (2.10).

The connection between Boolean sublattices of  $\mathcal{P}$  and mutually compatible sets of elements is expressed in the following theorem, proved by VARADARAJAN [10].

**Theorem 2.2.** *A subset  $\mathcal{B}$  of the orthomodular partially ordered set  $\mathcal{P}$  is a Boolean sublattice (with the same zero  $\Phi$  and unit  $I$  as  $\mathcal{P}$ ) if and only if (i)  $a, b \in \mathcal{B} \Rightarrow a \leftrightarrow b$  and also  $a \wedge b, a \vee b \in \mathcal{B}$ , (ii)  $a \in \mathcal{B} \Rightarrow a' \in \mathcal{B}$ .*

A simple corollary to this theorem is that  $a \leftrightarrow b$  implies  $a \leftrightarrow b'$ , etc. We are not assuming that  $\mathcal{P}$  is a lattice, and to ensure that sufficiently many joins and meets exist, we suppose that

**Axiom A.7.** *If  $a, b$  and  $c$  are three mutually compatible elements of  $\mathcal{P}$  then  $a \leftrightarrow b \vee c$  and  $a \leftrightarrow b \wedge c$ .*

The property expressed in axiom A.7 is not possessed by a general orthomodular partially ordered set, as is shown by the example of RAMSAY [16]. In accordance with the finiteness restrictions, we require:

**Axiom A.8.** *There exist at most a countable number of elements in any Boolean sublattice of  $\mathcal{P}$ , and, for every element  $a \in \mathcal{P}$ , at least one of  $(a, a')$  is finite (i.e., any maximal chain joining  $a$  to  $\Phi$  contains a finite number of elements). Except where otherwise stated, we assume that the unit  $I$  is not a finite element.*

We do not wish to imply by axiom A.8 that it is necessary always to work with such a logic, but that whichever logic is used, there is a physically equivalent one satisfying axiom A.8. A consequence of axiom A.8 is the existence of *atoms*, i.e. minimal non-zero elements in the logic  $\mathcal{P}$ .

### c) Simple Logics

Let  $\mathcal{C}$  be the *centre* of the logic  $\mathcal{P}$ , i.e. the set of all elements which are compatible with every other element. The elements of  $\mathcal{C}$  are essentially classical in nature and possess the corresponding properties. In the separable, atomistic case resulting from axiom A.8, it can be shown that  $\mathcal{C}$  is generated by the atoms in  $\mathcal{C}$ , by taking finite joins and their complements. There is a corresponding decomposition [10] of the entire logic  $\mathcal{P}$  into the *direct sum* of simple logics  $\mathcal{P}_1, \mathcal{P}_2, \dots$  in each of which

the centre consists only of the zero and unit elements. In ordinary quantum theory, the existence of superselection rules [17] implies the presence of universally compatible non-trivial observables, giving rise to the need to consider non-simple logics. However, for many purposes it is sufficient to treat only simple logics, corresponding to a single superselection class. This simplifies many of the detailed statements of theorems appearing in later sections.

#### *d) Measurability of Propositions*

In the preamble to axioms A.1 to A.8, we used an undefined notion of ‘measurability’ with reference to the elements of  $\mathcal{P}$ . This term can be justified only a posteriori, when a complete measurement theory is developed within the overall theory with logic  $\mathcal{P}$ . This is a requirement of self-consistency on the theory, which will not be considered here. In ordinary quantum mechanics, many studies of this problem have been made [18, 19], though not with wholly satisfactory results.

### 3. Linearity and Observability

#### *a) Linearity*

There are still several natural requirements of an elementary nature which one may impose on the space  $\mathcal{S}$  of states in order to limit the range of possibilities. One is to require a universal ‘mixing’ property for states. Classically, this is always permissible; if we take any two ensembles  $X$  and  $Y$ , each with given probability distributions  $x$  and  $y$ , then there is an ensemble which we may denote by  $\beta X + (1 - \beta) Y$ , with distribution  $\beta x + (1 - \beta)y$ . In other words, that states of classical systems form a convex subset of the linear vector space of all signed measures on phase space. This is a feature which we may expect to retain in more general theories as

**Axiom A.9.** *For any  $f, g \in \mathcal{S}$  and  $\beta \in [0, 1]$ , there is a state  $h \in \mathcal{S}$  such that  $h(a) = \beta f(a) + (1 - \beta) g(a)$  for all  $a \in \mathcal{P}$ .*

The uniqueness of  $h$  in A.9 is guaranteed by A.3 and it is said to be *mixture* of the states  $f$  and  $g$ . The *extreme* points (if any exist) of the convex set  $\mathcal{S}$  are the *pure states* of the system.

The convex set  $\mathcal{S}$  can be embedded in the convex cone  $\mathcal{C}_{\mathcal{S}}$  of all real multiples of the functions  $f \in \mathcal{S}$ . This cone is then simply the positive cone of an ordered linear vector space  $\mathcal{L}$  of additive functions on  $\mathcal{P}$ , where  $\mathcal{L} = \mathcal{C}_{\mathcal{S}} + (-\mathcal{C}_{\mathcal{S}})$ . The space  $\mathcal{L}$  carries the natural norm [8]

$$\|x\| = \sup\{|x(a) - x(a')|; a \in \mathcal{P}\} \quad (3.1)$$

so that the unit ball is simply the convex cover of  $\mathcal{S} \cup (-\mathcal{S})$ . The set  $\mathcal{S}$  can then be characterized in  $\mathcal{L}$  as the set of positive elements of norm one.

*b) Experimentally Distinguishable States*

There is a further topology on  $\mathcal{L}$  with direct physical relevance. As emphasized by HAAG and KASTLER [20], a weak topology is particularly appropriate for expressing the usual limitations imposed on any real experimenter, in that within a finite time only a *finite* number of propositional observations of finite accuracy can be carried out. But a subset of  $\mathcal{S}$  of the form

$$U(f; p_1, p, \dots, p_n; \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) \\ = \{g: g \in \mathcal{S}; |g(p_i) - f(p_i)| < \varepsilon_i; i = 1, 2, \dots, n\} \quad (3.2)$$

defines a weak neighbourhood of  $f \in \mathcal{S}$  in a certain weak topology which we may call  $\sigma(\mathcal{S}, \mathcal{P})$ . The positive numbers  $\{\varepsilon_i\}$  we can regard as determining the limits of accuracy of each of the experiments in the set  $\{p_i\}$  (the precise connection is not relevant). The collection of weak neighbourhoods of  $f$  when  $\{p_i\}$  runs over all finite subsets of  $\mathcal{P}$  and  $\{\varepsilon_i\}$  over all finite sets of rational positive numbers forms a base of neighbourhoods of  $f$  in  $\sigma(\mathcal{S}, \mathcal{P})$ . That this topology is Hausdorff (non-coinciding points can be placed in disjoint neighbourhoods) follows from axiom A.1, as for any  $f \in \mathcal{S}$ , there must be at least one atom  $a$  with  $f(a) \neq 0$ . The weak topology  $\sigma(\mathcal{S}, \mathcal{P})$  on  $\mathcal{L}$  is defined in a similar manner.

According to the tentative arguments of HAAG and KASTLER [20], the best any experimenter can do towards determining a state is to determine at most a weak neighbourhood in  $\mathcal{S}$ , rather than the unattainable ideal of a single point. Furthermore, we can see by means of a simple counting argument that, to cover  $\mathcal{S}$ , one needs only a finite number of weak neighbourhoods, constructed according to equation (3.2), for a given fixed set of experimental errors. In other words, there are only a finite number of experimentally distinguishable states. Mathematically, this is expressed by saying that  $\mathcal{S}$  is  $\sigma(\mathcal{S}, \mathcal{P})$ -precompact (or totally bounded). This result is more usually proved by embedding  $\mathcal{S}$  homeomorphically in the space of all functions on  $\mathcal{P}$  (taking values in  $[0, 1]$ ), this space being compact in the Tychonoff product topology.

*c) Pure States*

We cannot yet say whether  $\mathcal{S}$  possesses any extreme points at all. As any weak neighbourhood of such a pure state will inevitably contain mixed states, the distinction is never quite clear cut. For convenience, however, we can always extend  $\mathcal{S}$  by adding elements which do not affect the physical content, but which ensure a full set of pure states. The appropriate mathematical technique is that of *completion*, in the uniform structure determined by the completely regular topology  $\sigma(\mathcal{S}, \mathcal{P})$ . In the sequel, we assume that this process has been carried out and hence



that  $\mathcal{S}$  is  $\sigma(\mathcal{S}, \mathcal{P})$ -compact. The Krein-Milman theorem [21] then ensures that  $\mathcal{S}$  is the weakly closed convex hull of the set of its extreme points, thus making precise what we mean by a *full* set of pure states.

We can deduce further elementary consequences, making use of known results for cones with a compact base [21, 22]. In particular, the normed space  $\mathcal{L}$  is complete and hence is a Banach space whose closed unit ball is  $\sigma(\mathcal{L}, \mathcal{P})$ -compact. The cone  $\mathcal{C}_{\mathcal{S}}$  is  $\sigma(\mathcal{L}, \mathcal{P})$ -locally compact and  $\sigma(\mathcal{L}, \mathcal{P})$ -closed (and hence norm-closed). From these properties it follows that  $\mathcal{L}$  is canonically equivalent [23] to the Banach dual space of a normed linear subspace  $[\mathcal{P}]$  of the Banach dual  $\mathcal{L}'$  of  $\mathcal{L}$  itself. The space  $[\mathcal{P}]$  is defined as the linear span of the image of  $\mathcal{P}$  under the canonical imbedding

$$\tau: \mathcal{P} \rightarrow \mathcal{L}'; \quad \langle \tau(a), x \rangle = x(a) \quad \text{for all } x \in \mathcal{L}. \quad (3.3)$$

As no ambiguity will arise, we will omit  $\tau$  in the sequel and identify an element of  $\mathcal{P}$  with its canonical image on  $\mathcal{L}'$ .

The norm in  $[\mathcal{P}]$  is the one induced in it by the embedding in  $\mathcal{L}'$ . However, except in the finite dimensional case,  $[\mathcal{P}]$  itself is never a Banach space. For the purposes of comparison with ordinary quantum mechanics, it is convenient to introduce also the Banach space  $[\hat{\mathcal{P}}]$ , the completion of  $[\mathcal{P}]$ . We can obtain an indirect characterisation of this completion by using the method of GROTHENDIECK [21, p. 148, 24], for which the elements of  $[\hat{\mathcal{P}}]$  can be identified with those linear functionals on  $\mathcal{L}$  which are  $\sigma(\mathcal{L}, \mathcal{P})$ -continuous on bounded subsets of  $\mathcal{L}$ . The natural inclusions  $[\mathcal{P}] \subset [\hat{\mathcal{P}}] \subset \mathcal{L}'$  are all  $\sigma(\mathcal{L}', \mathcal{L})$ -dense.  $\mathcal{L}$  is the Banach space dual of both  $[\mathcal{P}]$  and  $[\hat{\mathcal{P}}]$ , and the topologies  $\sigma(\mathcal{L}, [\mathcal{P}])$  and  $\sigma(\mathcal{L}, [\hat{\mathcal{P}}])$  both induce the same topology  $\sigma(\mathcal{S}, \mathcal{P})$  on  $\mathcal{S}$ .

There is one particular pure state which needs separate treatment. Let  $[\mathcal{P}_f]$  denote the linear span of the finite elements of  $[\mathcal{P}]$  and  $[\hat{\mathcal{P}}_f]$  its completion. Then  $[\mathcal{P}_f]$  is a norm-closed subspace of  $[\mathcal{P}]$ , as the corresponding linear functional  $h$  which takes the value 0 on  $[\mathcal{P}_f]$  and 1 on  $I$  is clearly norm-continuous. The dual of  $[\mathcal{P}_f]$  is thus the factor space  $\mathcal{L}|\mathcal{R}_h$ , where  $\mathcal{R}_h$  is the one-dimensional subspace spanned by  $h$ . Clearly  $h$  is a pure state, but of no particular physical relevance, arising solely from our insistence on including non-finite elements in  $\mathcal{P}$ .

#### d) Primitive Observables

At this stage of the development, the linear subspace  $[\mathcal{P}]$  of  $\mathcal{L}'$  is not very well characterised, because we have as yet no prescription for determining how many ways there are of representing a point of  $[\mathcal{P}]$  as finite linear combinations of elements of  $\mathcal{P}$ . Thus there are several different ways that one can extend the usual definition of an observable

in terms of elementary propositions. We follow the ‘spectral resolution’ definition proposed by MACKAY [7] as being the most appropriate, though our scheme is still too general to say whether even the *sum* (in  $[\mathcal{P}]$ ) of two propositions is always an observable in this sense. In classical and ordinary quantum mechanics, sums of observables are always observables.

**Definition 3.1.** *A primitive observable is an element of  $[\mathcal{P}]$  of the form  $\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_n a_n$  where  $\{\lambda_i\}$  is any set of distinct non-zero real numbers and  $\{a_i\}$  is a set of pairwise orthogonal elements of  $\mathcal{P}$ .*

This definition of an observable, involving only a *finite* number of *compatible* propositions, is essentially the weakest one we can make. However, in the sequel, it will turn out that all the more general types of observable are limits in a certain sense of these primitive observables and so the definition is satisfactory for our purposes. We say that two primitive observables are *compatible*, when all the component propositions are pairwise compatible. The set  $\{\lambda_i\}$  is the *spectrum* of the observable. Note that the norm in  $\mathcal{L}'$  of the observable is just  $\max\{|\lambda_i|; i = 1, 2, \dots, n\}$ .

### e) Order Structures

The space  $\mathcal{L}'$  is an *ordered* linear vector space [21] if we take as positive cone the cone  $\mathcal{C}'_{\mathcal{P}}$  dual to the cone  $\mathcal{C}_{\mathcal{P}}$ . Thus  $\mathcal{C}'_{\mathcal{P}}$  consists of all continuous linear functionals on  $\mathcal{L}$  which are non-negative on  $\mathcal{C}_{\mathcal{P}}$ . The qualification “continuous” is in fact superfluous, as a theorem due to KLEE [21, p. 228; 25] shows that any positive linear functional on  $\mathcal{L}$  is automatically continuous,  $\mathcal{L}$  being a complete metric space generated by a closed positive cone.

An ordered linear space  $(E, C)$ , satisfying

- a)  $E = C - C$ , where  $C$  is the positive cone
- b)  $E$  admits a norm  $\|\cdot\|$  such that, for any  $x_1, x_2 \in C$ , we have  $\|x_1 + x_2\| = \|x_1\| + \|x_2\|$ .

c) Every positive linear functional of  $(E, C)$  is  $\|\cdot\|$ -continuous, is called a *GL-space* by MILES [26]. Thus our space  $(\mathcal{L}, \mathcal{C}_{\mathcal{P}})$  is a *GL-space*, whose dual space is thus a *GM-space* [26]. Hence the ordered linear space  $(\mathcal{L}', \mathcal{C}'_{\mathcal{P}})$  has the properties; (i) the cone  $\mathcal{C}'_{\mathcal{P}}$  contains an internal point  $e$  and (ii) the support function of the convex set  $(\mathcal{C}'_{\mathcal{P}} - e) \cap (e - \mathcal{C}'_{\mathcal{P}})$  is a norm. In our case we take  $e$  as the unit  $I$  of  $\mathcal{P}$ , in which case the unit ball in  $\mathcal{L}'$  is identical to the above convex set, i.e. the intersection of the positive and negative cones with vertices at  $-I$  and  $I$  respectively.

The positive cone  $[\hat{\mathcal{P}}_f]^+$  of  $[\hat{\mathcal{P}}_f]$  is thus  $[\hat{\mathcal{P}}_f] \cap \mathcal{C}'_{\mathcal{P}}$  and  $[\hat{\mathcal{P}}]^+$  is just the set of elements of the form  $x + \alpha I$ , where  $x \in [\hat{\mathcal{P}}_f]^+$  and  $\alpha \geq 0$ . In consequence,  $[\hat{\mathcal{P}}_f]^+$  and  $[\hat{\mathcal{P}}]^+$  possess the same set of extremal generators.

### 4. Self-adjointness and Reversibility

#### a) Introduction

We turn now to the ‘critical’ axioms which give rise to the usual algebraic structure of states and observables. An example of a propositional calculus which satisfies all the axioms stated so far (except that  $I$  is finite here) is given in Fig. 1. This is a lattice, and a simple enumeration

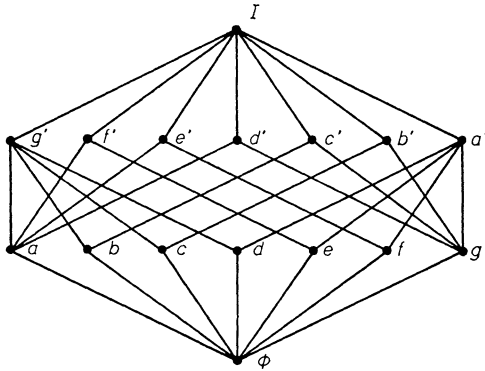


Fig. 1. An orthomodular lattice satisfying Axioms A.1 to A.9

shows there to be eight distinct pure states, spanning a four-dimensional state space. On the seven atoms  $a, b, \dots, g$ , the most general state takes the values  $1 - \gamma - \delta, \alpha, \beta, \alpha + \beta + \gamma + \delta - 1, \gamma, \delta, 1 - \alpha - \beta$  respectively. Here  $\alpha, \beta, \gamma, \delta$  are real numbers in the interval  $[0, 1]$  satisfying  $\alpha + \beta \leq 1, \gamma + \delta \leq 1, \alpha + \beta + \gamma + \delta \geq 1$ . The pure states are obtained when  $(\alpha, \beta, \gamma, \delta)$  take the following sets of values: (1000) (0100) (0010) (0001) (1010) (1001) (0110) (0101). Clearly there is no one-to-one correspondence between pure states and atoms, in contrast with the usual quantum mechanical lattices (of closed subspaces of a Hilbert space). In order to find a physical motivation for the final axioms, we have to look more closely into the interpretation of propositions as elementary measuring processes.

#### b) Elementary Observations

Much of the usual quantum theory of measurement [18] is independent of the detailed algebraic structure of the theory, in that it can be formulated in terms of the concepts of states, observables, transition probabilities, etc., already introduced. We will concentrate upon only one aspect of this theory, that of the close connection between the

observation and the preparation of a state. We always take states to refer to larger ensembles, so that the state can be determined by a complete set of measurements without effectively changing the state.

Each elementary proposition may be associated with a “filtering” process in which one imagines a portion of a given ensemble passing through the appropriate selection apparatus, the rest being rejected. The useful analogy of a semi-permeable membrane is an important feature of VON NEUMANN’S work [18]. Thereupon, a measurement of the fraction transmitted gives one physical datum concerning the state to be determined. This transmission coefficient is just what we mean by the value  $f(b)$  of the state  $f$  on the proposition  $b$ . However, not only does the selection apparatus assist in performing an observation, it can also prepare a state, the one in which the transmitted ensemble finds itself. The crucial point to be met now is: what happens when this new ensemble is subjected to a further selection process identical to the one which formed it? Certainly, the most desirable result is that of complete transmission. This accords with both the idempotent nature of propositions in the classical case and the results of experiments for many simple systems (e.g. polarization experiments). Furthermore, it seems reasonable that the most ‘selective’ processes should produce the purest states. In the optimum case, this determines a unique one-to-one correspondence between (physically distinguishable) pure states and (physically distinguishable) atomic propositions. Furthermore, we can extend this correspondence to be linear on  $\mathcal{L}_f$  into  $[\mathcal{P}_f]$ , in order that mixed states may be prepared by suitable combinations of selection processes acting ‘in parallel’. This correspondence will certainly be an algebraic isomorphism between the vector spaces  $\mathcal{L}_f$  and  $[\mathcal{P}_f]$  of finite linear combinations of pure states and atoms respectively, but need not necessarily be onto if the spaces are enlarged. The linear map

$$\sigma_f: \mathcal{L}_f \rightarrow [\mathcal{P}_f] \quad (4.1)$$

so defined is norm-reducing, as the extreme points of the unit ball in  $\mathcal{L}_f$  (the pure states and their negatives) are all mapped into elements of norm one in  $[\mathcal{P}_f]$ . Let  $\sigma$  be the extension of the map  $\sigma_f$  to the completed spaces  $\hat{\mathcal{L}}_f$  and  $[\hat{\mathcal{P}}_f]$ . The idempotence property requires that  $\langle f, \sigma(f) \rangle = 1$  for every pure state  $f$ , and the uniqueness requires that for any pure state  $g \neq f$ , we have  $\langle g, \sigma(f) \rangle < 1$ . Summarizing:

**Axiom A.10.** *There is a one-to-one continuous linear map  $\sigma: \hat{\mathcal{L}}_f \rightarrow [\hat{\mathcal{P}}_f]$  which maps the positive cone  $\mathcal{C}_{\mathcal{L}}$  onto a dense subcone of  $[\mathcal{P}_f]$ , taking pure states into atoms and such that  $\langle g, \sigma(f) \rangle \leq 1$ , where  $f, g$  are pure states, equality being attained only for  $f = g$ .*

In the sequel, the image under  $\sigma^{-1}$  of an atom  $a$  will be written  $\hat{a}$ .

*c) Reversibility and Compatibility*

The action of the filtering process associated with a proposition  $a$  may be represented by a linear operator  $E(a)$  acting on  $\mathcal{L}$ . In terms of an

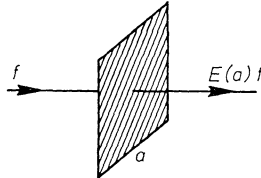


Fig. 2. Illustrating the projection operator  $E(a)$

obvious diagrammatical notation (Fig. 2), we have  $\|E(a)f\| = \langle a, f \rangle$  for any state  $f \in \mathcal{L}$ . From A.10, we find

$$E(a)\hat{a} = \hat{a}; \quad E(a)\hat{b} = \langle a, b \rangle \hat{a} \tag{4.2}$$

$$\|E(a)\hat{b}\| = 1 \Leftrightarrow a = b. \tag{4.3}$$

We can now formulate our final axioms,

**Axiom A.11.** *If  $f$  is a pure state, then, for any  $a \in \mathcal{P}$ ,  $E(a)f$  is proportional to a pure state.*

**Axiom A.12.** *If  $a \leq b$ , then  $E(a)E(b) = E(b)E(a) = E(a)$ .*

In effect, A.12 states that for the case of comparable propositions, we obtain the usual classical laws for composition of transmission probabilities (Fig. 3). Both  $\hat{g}$  and  $\hat{h}$  are pure states and  $\langle a, \hat{f} \rangle = \langle b, \hat{f} \rangle \langle a, \hat{h} \rangle$ .

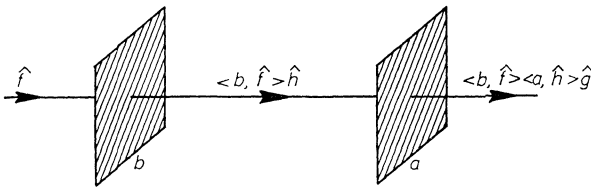


Fig. 3. Illustrating Axiom A.12 for a pure state  $\hat{f}$

This clarifies our interpretation of simultaneous observability for  $a$  and  $b$  for a given state  $\hat{f}$ , in that we can still determine  $\langle a, \hat{f} \rangle$  after  $\langle b, \hat{f} \rangle$  has been determined by processing the whole of the ensemble in the state  $\hat{f}$ .

In axiom A.11, we find a qualitative requirement which seems essential if our system is to be capable of describing a maximal set of reversible processes. Suppose all the elementary selection processes are capable (at least in principle) of being performed reversibly. Then the process splits an ensemble into two subensembles, the 'reflected' and the

‘transmitted’ ensembles, in such a manner that it is possible to reconstruct the original ensemble by reversing the entire process. A typical example is that of the splitting and recombination of polarized beams of particles under conditions such that the ‘phase relations’ (as described by ordinary quantum mechanics) are undisturbed. However, this is an experimental property which can be described and checked independently of the particular interpretation given by ordinary quantum mechanics.

One consequence of this form of reversibility is that, on starting with an ensemble already in a pure state (as determined by a complete set of measurements), then the partial ensembles created by the idempotent selection process must also be in pure states. This is clear, if we can construct a statistical theory involving the usual concepts of entropy, etc., satisfying the usual relations for reversible and irreversible processes. A lengthy discussion is given in VON NEUMANN’s book [18], based of course on the usual quantum mechanical formalism. However, the arguments used are not dependent on the particular algebraic structure associated with quantum mechanics. In order to carry through VON NEUMANN’s arguments, one needs only the existence of a ‘spectral resolution’ of a state. We show in theorems 4.12 and 4.13 that the axioms ensure this, but we have not found an argument not involving axiom A.11 itself.

*d) Self Adjointness*

We proceed to further characterise the operators  $E(a)$ .

**Lemma 4.1.** *If  $b$  is an atom and  $E(a)\hat{b} = 0$ , then  $b \perp a$ .*

*Proof.* Clearly  $\langle a, \hat{b} \rangle = 0$ , i.e.  $\langle a', \hat{b} \rangle = 1$ . From A.11,  $E(a')\hat{b} = \hat{c}$ , where  $c$  is an atom  $\leq a'$ . From A.12,  $E(c)\hat{b} = \langle c, \hat{b} \rangle \hat{c} = \hat{c}$ , hence  $b = c$ , giving the required result.

**Theorem 4.2.** (*Commutativity Theorem*). *For any  $a, b \in \mathcal{P}$ , then  $E(a)E(b) = E(b)E(a) = E(a \wedge b)$  if and only if  $a \leftrightarrow b$ .*

*Proof.* For any state  $\hat{f}$ ,

$$E(a)E(b)\hat{f} = E(a)\langle b, \hat{f} \rangle \hat{g} = \langle a, \hat{g} \rangle \langle b, \hat{f} \rangle \hat{h} \tag{4.4}$$

$$E(b)E(a)\hat{f} = E(b)\langle a, \hat{f} \rangle \hat{k} = \langle b, \hat{k} \rangle \langle a, \hat{f} \rangle \hat{l} \tag{4.5}$$

where  $g \leq b$ ,  $h \leq a$ ,  $k \leq a$ ,  $l \leq b$ . Let  $a = c + e$  and  $b = d + e$  as in (2.12). Then from (4.4) and  $c \perp b$ , we find

$$E(c)E(a)E(b)\hat{f} = E(a)E(c)E(b)\hat{f} = 0 = \langle a, \hat{g} \rangle \langle b, \hat{f} \rangle E(c)\hat{h} \tag{4.6}$$

so either (i)  $c \perp h$  or (ii)  $E(a)E(b)\hat{f} = 0$ . In case (i), the relation  $\langle a, \hat{h} \rangle = \langle c, \hat{h} \rangle + \langle e, \hat{h} \rangle = 1$  gives  $\langle e, \hat{h} \rangle = 1$ , so  $E(e)\hat{h} = \hat{m}$ , for  $m \leq e$ . But  $E(m)\hat{h} = E(m)E(e)\hat{h} = \hat{m}$ , i.e.  $\langle h, \hat{m} \rangle = 1$ , giving  $m = h$ . Thus we get

$$E(e)E(a)E(b)\hat{f} = E(e)\hat{f} = \langle e, \hat{f} \rangle \hat{n} = \langle a, \hat{g} \rangle \langle b, \hat{f} \rangle \hat{h} \tag{4.7}$$

where  $n \leq e$ . Hence

$$\langle e, \hat{f} \rangle = \langle a, \hat{g} \rangle \langle b, \hat{f} \rangle \quad \text{and} \quad h = n \tag{4.8}$$

A corresponding treatment of (4.5) gives

$$\langle e, \hat{f} \rangle = \langle b, \hat{k} \rangle \langle a, \hat{f} \rangle \quad \text{and} \quad l = n \tag{4.9}$$

and the required result follows immediately. In case (ii),

$$E(a) E(b) \hat{f} = 0 = (\langle c, \hat{g} \rangle + \langle e, \hat{g} \rangle) (\langle d, \hat{f} \rangle + \langle e, \hat{f} \rangle) \hat{h}.$$

Thus either (a)  $\langle d, \hat{f} \rangle + \langle e, \hat{f} \rangle = 0$ , whence  $f \perp e$  or  $f \leq e' = a' \vee b'$ , or (b)  $\langle c, \hat{g} \rangle + \langle e, \hat{g} \rangle = 0$ , where  $g \perp e$  or  $g \leq c'$  and also  $g \leq b$  hence  $g \leq d = b \wedge e'$ . But  $g \leq d$  implies  $(\langle d, \hat{f} \rangle + \langle e, \hat{f} \rangle) \hat{g} = E(b) \hat{f} = E(d) E(b) \hat{f} = E(d) \hat{f} = \langle d, \hat{f} \rangle \hat{g}$ , giving  $\langle e, \hat{f} \rangle = 0$ . This means that  $f \leq e' = a' \vee b'$ . In both (a) and (b), we get  $f \leq a' \vee b'$ , hence  $E(a \wedge b) \hat{f} = 0$ .

Conversely, suppose  $E(a) E(b) \hat{f} = E(b) E(a) \hat{f} = E(a \wedge b) \hat{f}$ . For  $f$  take any atom  $\leq (a \wedge b)' \wedge a$ . Then  $E(b) \hat{f} = 0$ , and, from Lemma 4.1,  $f \perp b$ . Hence  $(a \wedge b)' \wedge a \perp b$ . Similarly,  $(a \wedge b)' \wedge b \perp a$ . Thus  $a = (a \wedge b)' \wedge a + a \wedge b$ ;  $b = (a \wedge b)' \wedge b + a \wedge b$  is a decomposition satisfying 2.12, resulting in  $a \leftrightarrow b$ . This completes the proof.

**Theorem 4.3.** *If  $a \in \mathcal{P}$  and  $b$  is any atom, then  $a \vee b$  exists and  $E(a') \hat{b} = \langle a', \hat{b} \rangle \sigma^{-1}((a \vee b) \wedge a')$ .*

*Proof.* Clearly  $E(a') \hat{b} = \langle a', \hat{b} \rangle \hat{c} = \langle c, \hat{b} \rangle \hat{c}$ , where  $c \leq a'$ . Let  $d \in P$  satisfy  $d \geq a, b$ . The relation  $d \leftrightarrow a'$  and Theorem 4.2 give  $E(d) \langle c, \hat{b} \rangle \hat{c} = E(d) E(a') \hat{b} = E(a') E(d) \hat{f} = E(a) \hat{b} = \langle c, \hat{b} \rangle \hat{c}$ , i.e.  $c \leq d$  (or  $E(a') \hat{b} = 0$ , but in this case  $b \leq a$  and the theorem is trivially valid). Hence we have  $a \vee c \leq d$ , and also  $E(c') \langle c, \hat{b} \rangle \hat{c} = 0 = E(c') E(a') \hat{b} = E(a') E(c') \hat{b} = E(a' \wedge c') \hat{f}$  giving  $b \leq a \vee c$ , using theorem 4.2. Collecting these results, we get  $a, b \leq a \vee c \leq$  any upper bound of  $a, b$ . Hence  $a \vee b$  exists and equals  $a \vee c = a + c$ . From  $c \leq a'$ , orthomodularity gives  $c = (a \vee c) \wedge a' = (a \vee b) \wedge a'$ , the required result.

**Corollary 4.4.** *If  $b_1, b_2, \dots, b_n$ , are any finite set of atoms, there exists a set of mutually orthogonal atoms  $a_1, a_2, \dots, a_m$  ( $m \leq n$ ), such that  $\bigvee b_\alpha = \bigvee a_\alpha$ .*

This follows by repeated application of the preceding theorem.

**Lemma 4.5.** *If  $a$  is an atom and  $x \in \mathcal{P}$ , then  $\langle x, \hat{a} \rangle = 1 \Leftrightarrow a \leq x$ .*

For, A.11 gives  $E(x) \hat{a} = \hat{c}$ , where  $c$  is an atom majorized by  $x$ . Hence  $E(c) E(x) \hat{a} = E(c) \hat{a} = \hat{c}$ , which implies  $a = c$ . The converse is trivial.

**Definition 4.6.** *Let  $x$  be any finite proposition, then by  $[\hat{x}]$  (resp.  $[x]$ ), we denote the closed subspace of  $\mathcal{L}$  (resp.  $[\mathcal{P}]$ ) spanned by  $\{\hat{a}\}$  (resp.  $\{a\}$ ), where  $\{a\}$  is the set of atoms  $\leq x$ .*

**Lemma 4.7.** *If  $x \in [\mathcal{P}_f]$ , then  $[\hat{x}]$  is the eigenspace of  $E(x)$  with eigenvalue 1.*

*Proof.* For any atom  $a \leq x$ , we have  $E(x) \hat{a} = \hat{a}$ , so it is clear that  $[\hat{x}] \subset$  the required eigenspace. Conversely, if  $f \in \mathcal{L}$ , then  $f$  is the sum of a

suitable norm convergent series

$$f = \sum_{n=1}^{\infty} \alpha_n \hat{a}_n .$$

Suppose  $E(x)f = f$ , then, using (A.11), we see that  $f$  also possesses a norm convergent representation

$$f = \sum_{n=1}^{\infty} \alpha_n \langle x, \hat{a}_n \rangle \hat{c}_n$$

where  $c_n \leq x$  for all  $n$ . Hence,  $[\hat{x}]$  being closed,  $f \in [\hat{x}]$ .

**Lemma 4.8.** *If  $x \in [\mathcal{P}_f]$  and  $[x]^0$  is the polar of  $[x]$ , then for any state  $f$ , we have  $f - E(x)f \in [x]^0$ .*

*Proof.* For any  $a \leq x$ , axiom A.12 gives  $\langle a, E(x)f \rangle \hat{a} = E(a) E(x)f = E(a)f = \langle a, f \rangle \hat{a}$ . Thus  $\langle a, f - E(x)f \rangle = 0$ . The atoms  $a \leq x$  span  $[x]$ , and the result follows.

The results of this series of lemmas are summarized in:

**Theorem 4.9.** *If  $x \in [\mathcal{P}_f]$ , then the dual of  $[x]$  is canonically isometric to  $[\hat{x}]$ .*

*Proof.* It is a standard result [21] that the Banach dual of the closed linear subspace  $[x]$  is canonically isometric to the quotient space  $\hat{\mathcal{L}}_f/[x]^0$ . Let  $\{f\}$  be the equivalence class in this quotient containing a given pure state  $\hat{f} \in \mathcal{L}_f$ . Then  $\hat{f} = E(x)\hat{f} + (\hat{f} - E(x)\hat{f})$ , so Lemma 4.8 requires that  $E(x)\hat{f}$  lies in the same equivalence class. But  $\|E(x)\hat{f}\| \leq \|f\| = 1$ , hence  $\{f\}$  lies in the unit ball of  $\mathcal{L}_f/[x]^0$ . Taking into account Lemma 4.7 and that the unit ball in  $\mathcal{L}_f$  is  $(\mathcal{S} + (-\mathcal{S})) \cap \hat{\mathcal{L}}_f$ , we conclude that the canonical mapping  $\hat{\mathcal{L}}_f/[x]^0 \leftrightarrow E(x)\hat{\mathcal{L}}_f$  is isometric.

This theorem has important consequences, in that  $[x]$  and  $[\hat{x}]$  possess all the properties of the original spaces  $[\mathcal{P}_f]$  and  $\mathcal{L}_f$ , obtained by setting  $x = I$  (in the above we have restricted  $x$  to lie in  $\mathcal{P}_f$ , for simplicity in statement, but this is not essential). In particular, the positive cones of  $[x]$  and  $[\hat{x}]$  can be identified by using the map  $\sigma$  (4.1), as well as being dual to each other. In order to illustrate the general case, we first suppose  $x = a \vee b$ , where  $a$  and  $b$  are two different non-orthogonal atoms. Then  $0, a, b, a \vee b, (a \vee b) \wedge a' = z$  and  $(a \vee b) \wedge b' = y$  span a three-dimensional space  $\mathcal{T}$  (cf. theorem 4.3) which is a subspace of  $[a \vee b]$ . The convex hull of the set of points is a subset of the positive cone in  $\mathcal{T}$ .

**Lemma 4.10.** *The open linear segments  $a \dots b$  and  $z \dots y$  cannot both lie on the boundary of the positive cone in  $[a \vee b]$ .*

*Proof.* If  $\lambda a + (1 - \lambda)b$  ( $0 < \lambda < 1$ ) were such a point, then, by theorem 4.9, there must be an atom  $u < a \vee b$  such that  $\langle u, \lambda \hat{a} + (1 - \lambda)\hat{b} \rangle = 0$ . This gives  $\langle u, \hat{a} \rangle = \langle u, \hat{b} \rangle = 0$ . But, from Lemma 4.1, this requires that  $u \perp a$  and  $u \perp b$ , i.e.  $u \perp a \vee b$ , a contradiction.



Thus there must be a further boundary point  $c$  in the plane of  $a, b, y, z$ . Moreover,  $c$  must be an atom. For, if it were a convex combination of two or more other atoms, we can repeat the above argument to show that these atoms are orthogonal to  $c$  and also less than  $a \vee b$ , arriving at a contradiction. Thus  $a \vee b = c + w$ , where  $w = (a \vee b) \wedge c'$  (Fig. 4).

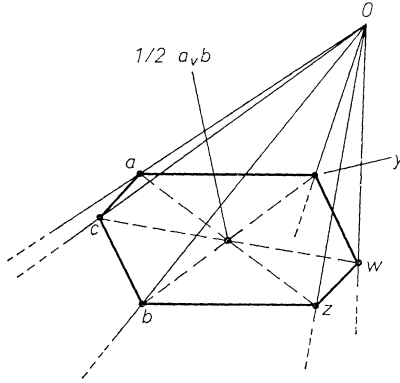


Fig. 4. Illustrating the linear relations between atoms in the proof of Lemma A.10

Repeating the above arguments for all pairs of atoms  $\langle a \vee b$ , we see that all boundary points of the positive cone of  $[a \vee b]$  are atomic. This property is essential in the proof of:

**Theorem 4.11.** (*Spectral theorem, first part*). Any element in  $[\mathcal{P}_f]$  of the form  $\lambda a + \mu b$ , where  $a$  and  $b$  are different atoms, possesses a spectral resolution of the form  $\lambda' c + \mu' d$ , where  $c \perp d$ .

*Proof.* If  $a \perp b$ , there is nothing to prove. Otherwise, following Lemma 4.10, we construct the plane containing the elements  $0, a \vee b$  and  $\lambda a + \mu b$ . This intersects the positive of  $[a \vee b]$  in a cone bounded by two atoms,  $c$  and  $d$ , say. Furthermore,  $c + d = a \vee b$  and  $c \perp d$ . Projecting onto these directions, shown dotted in Fig. 5, gives the required resolution.

**Lemma 4.12.** The bilinear form  $(\cdot, \cdot)$  on  $[a \vee b] \times [a \vee b]$  defined by  $p, q \rightarrow \langle p, \hat{q} \rangle$  is symmetric.

*Proof.* Let  $p$  and  $q$  be atoms. Then the linear functional  $z \rightarrow \langle z, \hat{q} \rangle$ ,  $z \in [a \vee b]$ , is, by definition, the element  $\hat{q}$  (cf. Theorem 4.9). Consider the linear functional  $\hat{r}$ , which, on atoms, is defined by  $\hat{r}: p \rightarrow \langle q, \hat{p} \rangle$  and elsewhere by linear extension. This is clearly a positive linear functional of norm 1 and moreover satisfies  $\langle (a \vee b) \wedge q', \hat{r} \rangle = 0$ . But the only state in  $[a \vee b]$  which is orthogonal to  $(a \vee b) \wedge q'$  is  $\hat{q}$  itself, i.e.  $\hat{q} = \hat{r}$  and hence  $\langle q, \hat{p} \rangle = \langle p, \hat{q} \rangle$ .

If  $[a \vee b]$  is a finite dimensional space, the positive cone, being its own dual under a symmetric bilinear form, is called self-adjoint [28–33]. Several studies of these cones have been made, mostly in connection with the algebraic structures that they support when the group of all automorphisms is transitive (homogeneous cones). In our case we will show that homogeneity is a consequence of the axioms.

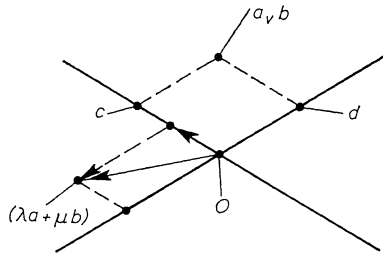


Fig. 5. Illustrating the spectral resolution of  $\lambda a + \mu b$

**Theorem 4.13.** (*Spectral theorem, second part*). Any finite element of  $[\mathcal{P}]$  possesses a spectral decomposition.

*Proof.* We need consider only a strictly positive element

$$x = \lambda_1 b_1 + \dots + \lambda_n b_n, \quad \lambda_i > 0,$$

which we take as an *internal* element of the positive cone of  $[\psi] = [b_1 \vee \dots \vee b_n]$ . Any other element can be translated into a strictly positive one by adding a suitable multiple of the appropriate order unit  $\psi$ .

The closed hyperplane  $(\psi, z) = 1$  contains all the atoms  $\leq \psi$ . The positive cone intersects this in a weakly compact convex subset  $\mathcal{S}_\psi$  of  $\mathcal{S}$ . Thus the linear function  $(x, z)$  ( $z \in \mathcal{S}_\psi$ ) attains its minimum ( $> 0$ ) on some extremal subset of  $\mathcal{S}_\psi$ . Let  $a$  be any atom on which the minimum is attained. Then the hyperplane  $H: (x - (a, x)a, z) = 0$  contains  $a$ , but if  $b (\neq a)$  is any other atom, then we have

$$(x - (a, x)a, b) \geq (a, x) - (a, x)(a, b) = (a, x)(1 - (a, b)) > 0.$$

Hence  $x - (a, x)a$  is a relatively internal element of the positive cone of  $[a' \wedge \psi]$ . Thus we have the orthogonal decomposition of  $x$  into positive elements;  $x = (a, x)a + (x - (a, x)a)$ . Repetition of this process will lead, after a finite number of steps, to a complete spectral resolution into multiples of pairwise orthogonal atoms.

### 5. Algebras of Observables

#### a) Jordan Structures

The usual Jordan product of matrices

$$\frac{1}{2}(ef + fe) = \frac{1}{2}(efe - (I - e)f(I - e) + f) \tag{5.1}$$

suggests that we introduce the following definition

$$\widehat{a \circ b} = \frac{1}{2}(E(a)\hat{b} - E(a')\hat{b} + \hat{b}) \tag{5.2}$$

of an algebraic product operation for any pair of atoms  $a, b$ .

The first main theorem of this section is:

**Theorem 5.1.** *The product  $a \circ b$  defined in (5.2) satisfies (i) commutativity,  $a \circ b = b \circ a$ , and (ii)  $a \circ (b \circ c) = b \circ (a \circ c)$  for any set of three atoms with  $a \circ b$ .*

*Proof.* (i). Using Theorem 4.3 and Lemma 4.12, we get

$$a \circ b = \frac{1}{2}((a, b)a - (a', b)(a \vee b) \wedge a' + b). \tag{5.3}$$

Using  $a \vee b = a + (a \vee b) \wedge a' = b + (a \vee b) \wedge b'$  and  $(a', b) = 1 - (a, b)$ , we can rearrange the terms into the form

$$\frac{1}{2}((b, a)(b + (b \vee a) \wedge b') - (b \vee a) \wedge b' + a)$$

which is just  $b \circ a$ . Hence the commutativity.

(ii) From the definition (5.2), the difference  $a \circ (b \circ c) - b \circ (a \circ c)$  is

$$\sigma([\frac{1}{2}(E(a) + I - E(a')), \frac{1}{2}(E(b) + I - E(b'))] \hat{c}).$$

But if  $a \perp b$ , then all the projection operators appearing in the commutator bracket commute (cf. theorem 4.2) and thus the difference vanishes.

We can now extend the definition (5.2) to arbitrary members

$x = \sum_i \alpha_i a_i$  and  $y = \sum_j \beta_j b_j$  of  $[\mathcal{P}_f]$  by linearity to

$$x \circ y = \sum_i \sum_j \beta_i \beta_j a_i \circ b_j. \tag{5.4}$$

The distributive property of  $a \circ b$  is an immediate consequence of the definition. Uniqueness of the extension by linearity in the second member  $y$  follows from the definition. Commutativity then ensures uniqueness of linear extension in the first factor. Moreover, the form (5.2) is still valid if  $a$  and  $b$  are any finite propositions.

**Lemma 5.2.** *For any  $x \in [\mathcal{P}_f]$ , the Banach space  $[x]$ , when endowed with the product operation (5.3), is a real simple Jordan algebra.*

*Proof.* For any pair of atoms  $a, b \leq x$ , we conclude from theorem 5.1 that  $a \circ b \in [a \vee b]$ . Thus  $[x]$  is closed under the product operation. Furthermore, if  $[x]$  is finite dimensional, there exists only a finite number of atoms in any orthogonal set on  $[x]$ . Noting that  $y^2 + z^2 = 0 \Rightarrow y^2 = z^2 = 0$  (use the spectral resolution theorem 4.13 on  $y$  and  $z$ ) we see that all the conditions for a finite real Jordan algebra are satisfied [1].

In particular, the second condition in theorem 5.1 implies the more usual condition  $y^2 \circ (a \circ y) = (y^2 \circ a) \circ y$  for any  $a, y \in [x]$ . This is proved by JORDAN et al. [1]. The infinite dimensional case will not concern us in the sequel and we omit the proof.

**Lemma 5.3.** *The (infinite) Jordan algebra  $[\mathcal{P}_f]$  is simple.*

*Proof.* Suppose  $[\mathcal{P}_f]$  were not simple, then there would be a non-trivial idempotent element  $x$  in the centre of  $[\mathcal{P}_f]$ , i.e. one for which the left translation

$$L_x : y \rightarrow x \circ y \tag{5.5}$$

on  $[\mathcal{P}_f]$  commuted with all other left translations [1, Sec. 9]. But from (5.2) and the remarks following theorem 5.1, we have

$$L_x = \frac{1}{2}(E(x) - E(x') + I). \tag{5.5}$$

Hence  $[(E(x) - E(x')), (E(z) - E(z'))] = 0$  for all idempotent  $z \in [\mathcal{P}_f]$ . Using arguments similar to those appearing in the proof of theorem 4.2, we conclude that  $x \leftrightarrow z$  for all  $z \in [\mathcal{P}_f]$ , in contradiction with our initial assumption that  $\mathcal{P}$  is a simple logic (Sec. 2c).

**Lemma 5.4.** *Any Jordan subalgebra of  $[\mathcal{P}]$ , which is of the type  $[x]$ , is simple.*

*Proof.* This can be proved in the same way as Lemma 5.3, or by using a theorem due to ALBERT [34]. When  $x$  is a finite element and  $[x]$  finite dimensional, then  $[x]$  is a real simple finite Jordan algebra, which have been completely classified [1] into

- (i)  $\mathfrak{R}$  the usual algebra of the reals
- (ii)  $\mathfrak{S}_N$  for  $N = 3, 4, \dots$ , the algebra with basis given by a set of  $(N - 1)$  linearly independent matrices  $\{\gamma_\mu\}$  and a unit  $I$ , satisfying

$$\gamma_\mu \circ \gamma_\nu = \frac{1}{2} (\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu) = \delta_{\mu\nu} I; \gamma_\mu \circ I = \gamma_\mu I.$$

- (iii)  $\mathfrak{M}_\chi^\mathfrak{H}$   $\chi = 3, 4, 5, \dots$ ,  $\chi = 1, 2, 4$ . The algebra of  $(\gamma \times \gamma)$  Hermitian matrices in which the elements are real numbers ( $\chi = 1$ ) or complex numbers ( $\chi = 2$ ) or quaternions ( $\chi = 4$ ). The Jordan product is  $a \circ b = \frac{1}{2}(ab + ba)$  in terms of the usual matrix product.

- (iv)  $\mathfrak{M}_3^8$  the algebra of  $3 \times 3$  Hermitian matrices with Cayley number (octonion) elements.

If there exists a set of at least four orthogonal atoms in  $[\mathcal{P}]$ , then the only possibilities lie in class (iii). The Jordan product can be extended to  $[\mathcal{P}] = [\mathcal{P}_f] \oplus [\lambda I]$ , being the algebra obtained by formally adjoining an identity element. If  $[\mathcal{P}]$  is regarded as a countable union of finite dimensional algebras, the classification problem is then reduced essentially to that of the finite dimensional algebras. Using the notion of direct limit of algebras [35], we may represent  $[\mathcal{P}]$  as the direct limit (no

topology) of the direct limit system  $(\{[x] \oplus \{\lambda I\}\}, \{II_{xy}\})$  directed by inclusion. The maps  $II_{xy} : [x] \oplus \{\lambda I\} \rightarrow [y] \oplus \{\lambda I\}$  are the canonical injections. Corresponding to the three main types of Jordan algebra in class (iii), we find

**Theorem 5.5.** *There are three distinct simple infinite Jordan algebras  $\mathfrak{M}_\infty^\chi$  of the type  $[\mathcal{P}]$ , being the operator algebras in a separable infinite dimensional Hilbert space, formed from self-adjoint operators which differ from a multiple of the identity by an operator of finite rank. The Hilbert space is defined over the reals ( $\chi = 1$ ), the complexes ( $\chi = 2$ ) or the quaternions ( $\chi = 4$ ).*

*Proof.* In any direct limit system representing  $[\mathcal{P}]$ , the algebras  $[x]$  must all be of the type  $\mathfrak{M}_\gamma^\chi$ , where  $\gamma$  is determined by the dimension of  $x$  (or rather, a spectral resolution of  $x$ ). Note that  $\mathfrak{M}_2^\chi = \mathfrak{S}_{\chi+2}$  and  $\mathfrak{M}_1^\chi = \mathfrak{R}$ . Moreover, the value of  $\chi$  must be the same for all algebras in the system. Indeed, for any pair of comparable algebras in the system, the smaller may be obtained from the larger by decomposition relative to a suitable idempotent [1, theorem 8] (cf. lemma 4.7). Hence, in any common decomposition of the pair relative to a maximal set of pairwise orthogonal unresolvable idempotents, the value of  $\chi$  resulting [1, theorem 17] must be the same for both algebras. Any two algebras in the system are subalgebras of some common larger algebra and the result follows immediately. Thus the three essentially distinct direct limit systems (for  $\chi = 1, 2, 4$ ) determine, at least algebraically, the three infinite Jordan algebras  $\mathfrak{M}_\infty^\chi$  defined above.

#### b) Extended Classes of Observables

The three Jordan algebras of theorem 5.5 are the smallest infinite dimensional analogues of the algebras  $\mathfrak{M}_\gamma^\chi$  ( $\gamma \geq 4$ ), if we insist on the presence of an identity element. These are the algebras of primitive observables (Def. 3.1). However, the algebras used in ordinary quantum mechanics, especially those of all self-adjoint bounded operators on a Hilbert space, are much larger. By relaxing our finitistic requirements and allowing limiting operations in  $[\mathcal{P}]$ , we can include these cases. If we keep the same set of states, the largest possible space consists of all linear forms on  $\mathcal{L}$ , i.e., the algebraic dual  $\mathcal{L}^*$ . By embedding  $[\mathcal{P}]$  in  $\mathcal{L}^*$  and adjoining appropriate limits in the weak  $\sigma(\mathcal{L}^*, \mathcal{L})$  topology, we can construct various extended spaces. In particular, the Banach space  $\mathcal{L}'$  is given as the union of the  $\sigma(\mathcal{L}^*, \mathcal{L})$ -closures in  $\mathcal{L}^*$  of all bounded sets in  $[\mathcal{P}]$  [21, p. 143].

For each of the three Hilbert spaces  $\mathcal{H}_\chi$ ,  $\chi = 1, 2, 4$ , we obtain the following realisations of the various spaces defined in section 3.

(i)  $[\hat{\mathcal{P}}]$  is the completion of  $[\mathcal{P}]$  in norm, i.e. the usual operator bound, and so consists of all self-adjoint operators which differ from a multiple of the unit operator by a compact operator.

(ii)  $\mathcal{L}/\mathcal{R}_h$  is the Banach space of all self-adjoint trace-class operators, with trace norm [36]. The duality between  $[\mathcal{P}]$  and  $\mathcal{L}$  becomes

$$f(x) = \frac{1}{2} \text{tr}(O_f O_x + O_x O_f)$$

where the state  $f$  is represented by the trace class operator  $O_f$ .

(iii)  $(\mathcal{L}/\mathcal{R}_h)'$  is the Jordan Banach algebra of all bounded self-adjoint operators on  $\mathcal{H}$  [36].

Correspondingly, we have enlarged the original lattice of propositions. However, we regard any larger lattice, up to the lattice of all projection operators as being physically equivalent to the one obtained initially, i.e. the lattice  $\mathcal{P}$  formed from projection operators of finite ranks and their complements. The equivalence is understood in the same sense as the dual case discussed in section 3 b, i.e. for any given finite number of states in  $\mathcal{S}$  and corresponding finite errors, the weak neighbourhood of any proposition so determined always contains finite ones.

## 6. Notes and Comments

1. Axioms A.11 and A.12 can probably be replaced by many suitable alternatives. In the finite-dimensional case, it is known [28–33] that a necessary and sufficient condition for a self-adjoint convex cone to be the positive cone of a real Jordan algebra is that it possesses a transitive group of automorphisms. In the weakly infinite dimensional case considered here, a similar statement may be expected to hold. There are many conditions which ensure such homogeneity, but we have not discovered any physically satisfactory ones. For example, if the involution  $x \rightarrow -x^{-1}$  is order preserving, then this is a sufficient condition.

A natural requirement on the theory is that the order preserving one-to-one transformations of  $\mathcal{L}$  which leave invariant the set of states  $\mathcal{S}$  should be transitive on the pure states. It is this type of transformation which is required to describe the development in time of an isolated system according to some reversible equation of motion. Furthermore, it is known that such order preserving one-to-one maps are necessarily linear in all cases of interest [31] and hence are automorphisms of the order cone. Full transitivity may require somewhat more than this.

2. We do not know if axiom A.12 is independent of the earlier axioms. We suspect this to be the case, although we know of no supporting counterexample.

3. For the case of logics which are not simple, but can be expressed as direct sums of simple logics, the axioms do not require modification.

The resulting algebras are then the direct sums of the simple algebras corresponding to the simple logics.

4. The cases for and against the use of scalars from each of the real, complex and quaternion fields have been treated fully in several papers [38, 39, 40].

5. All the algebras we have considered are essentially of von Neumann type I, as we have explicitly postulated sufficient minimal idempotents. While we feel that no more physical content will be found in more general algebras, nevertheless it would seem to be of importance to cover the cases of types II and III as well. That the usual field commutation rules give rise to such algebras is well-known [41, 42, 43]. For this, the recent studies of Jordan  $C^*$ - and  $W^*$ -algebras will be useful [44].

6. The considerations of Sec. 3b can be developed further if we are given a realization of the abstract set of states  $\mathcal{S}$  in terms of a particular physical picture of the universe. With any experimenter  $E$ , we may associate a 'subjective state space'  $\mathcal{S}_E$  which depends on the maximum experimental resolution of which he is capable. This space may be defined as the pair  $(\mathcal{S}, T_E)$  in which  $T_E$  is a certain set of weak neighbourhoods in  $\mathcal{S}$ . For any real experimenter, there are not sufficient of these neighbourhoods to separate the points of  $\mathcal{S}$ . We cannot in general define it as a quotient state space in which indistinguishable elements are identified, as they do not necessarily fall into equivalence classes. However, pairs of elements of  $\mathcal{S}$  may be classified into 'distinguishable' if they satisfy the Hausdorff separation property of Sec. 3b, and indistinguishable otherwise. From this viewpoint, the purity or degree of mixedness of a state is a highly subjective characteristic. Indeed, we may remark that some pairs of states remain indistinguishable for all conceivable experimenters: It is unreasonable that experimental accuracies which require measuring tools greater than the size or mass of the universe and operating longer than the lifetime of the universe should be considered! Thus states of sufficient complexity which we may term 'macroscopic' may be expected to lose much of their quantum mechanical nature. We do not wish to pursue further the details of this classical limiting process here.

VON NEUMANN'S quantum mechanical theory of the measuring process [18] may be interpreted from this viewpoint, in that his controversial 'projection' postulate loses many of its apparently paradoxical properties if it is recognised that the projected and unprojected states are, to the relevant experimenter, indistinguishable. This point has been treated in detail by JAUCH [19].

The author would like to thank Prof. M. H. STONE for pointing out some errors in an earlier version and for useful comments.

## References

1. JORDAN, P., J. VON NEUMANN, and E. WIGNER: *Ann. Math.* **35**, 29 (1934).
2. VON NEUMANN, J.: *Mat. Sb.* **1**, 415 (1936).
3. BIRKHOFF, G., and J. VON NEUMANN: *Ann. Math.* **37**, 823 (1936).
4. SEGAL, I. E.: *Ann. Math.* **48**, 930 (1947).
5. LOWDENSLAGER, D. B.: *Proc. Am. Math. Soc.* **8**, 88 (1957).
6. SHERMAN, S.: *Ann. Math.* **64**, 593 (1956).
7. MACKEY, G. W.: *Mathematical foundations of quantum mechanics*. New York: Benjamin 1963.
8. ZIERLER, N.: *Proc. Am. Math. Soc.* **14**, 345 (1963).
9. — *Pacific J. Math.* **11**, 1151 (1961).
10. VARADARAJAN, V. S.: *Comm. Pure Appl. Math.* **15**, 189 (1962).
11. PIRON, C.: *Helv. Phys. Acta* **37**, 439 (1964).
12. JAUCH, J. M., and C. PIRON: *Helv. Phys. Acta* **36**, 827 (1963).
13. POOL, J. C. T.: *Simultaneous observability and the logic of quantum mechanics*. Univ. of Iowa Ph.D. thesis. Department of Physics Report No. 12.
14. LOOMIS, L. H.: *Mem. Am. Math. Soc. No. 18* (1955).
15. FOULIS, D. J.: *Portugal. Math.* **21**, 65 (1962).
16. RAMSAY, A.: *J. Math. Mech.* **15**, 227 (1966).
17. WICK, G. C., A. S. WIGHTMAN, and E. P. WIGNER: *Phys. Rev.* **88**, 101 (1952).
18. VON NEUMANN, J.: *Mathematical foundations of quantum mechanics*. Princeton: Princeton University Press 1956.
19. JAUCH, J. M.: *Helv. Phys. Acta* **37**, 293 (1964).
20. HAAG, R., and D. KASTLER: *J. Math. Phys.* **5**, 848 (1964).
21. SCHAEFER, H. H.: *Topological vector spaces*. New York: Macmillan 1966.
22. EDWARDS, D. A.: *Proc. Lond. Math. Soc.* **14**, 399 (1964).
23. RUSTON, A. F.: *Proc. Cambridge Phil. Soc.* **53**, 576 (1956).
24. GROTHENDIECK, A.: *Compt. Rend.* **130**, 605 (1950).
25. KLEE, V.: *Duke Math. J.* **22**, 263 (1955).
26. MILES, P. E.: *Trans. Am. Math. Soc.* **107**, 217 (1963).
27. BOURBAKI, N.: *Espaces Vectoriels Topologiques*. Ch. IV. *Actualités Scientifiques et Industrielles Paris: Hermann 1955*.
28. KOECHER, M.: *Am. J. Math.* **79**, 575 (1957).
29. ROTHBAUS, O.: *Abhandl. Math. Seminar Hamburg. Univ.* **24**, 199 (1960).
30. VINBERG, E. B.: *Soviet Math.* **1**, 787 (1960).
31. KOECHER, M.: *Math. Ann.* **148**, 244 (1962).
32. HERTNECKE, C.: *Math. Ann.* **146**, 433 (1962).
33. KOECHER, M.: *Bull. Am. Math. Soc.* **68**, 374 (1962).
34. ALBERT, A. A.: *Ann. Math.* **48**, 546 (1947). Theorem 14.
35. BOURBAKI, N.: *Algèbre Ch. 2. (3<sup>ème</sup> Ed<sup>n</sup>) § 6, No. 5. Actualités Scientifiques et Industrielles. Paris: Hermann 1962*.
36. SCHATTEN, R.: *Norm ideals of completely continuous operators*. Berlin-Göttingen-Heidelberg: Springer 1960.
37. LOEVMNER, C.: *Bull. Am. Math. Soc.* **70**, 1 (1964).
38. STUECKELBERG, E. C. C., and M. GUENIN: *Helv. Phys. Acta* **33**, 727 (1960).
39. FINKELSTEIN, D., J. M. JAUCH, S. SCHIMNOVICH, and D. SPEISER: *J. Math. Phys.* **3**, 207 (1962).
40. — *J. Math. Phys.* **4**, 788 (1963).
41. WIGHTMAN, A., and S. SCHWEBER: *Phys. Rev.* **98**, 812 (1955).
42. GLIMM, J.: *Ann. Math.* **73**, 572 (1961).
43. ARAKI, H.: *J. Math. Phys.* **5**, 1 (1964).
44. STORMER, E.: *Trans. Am. Math. Soc.* **120**, 438 (1965).