International Journal of Theoretical Physics, Vol. 24, No. 11, 1985

Fundamental Principles of Quantum Theory. II. From a Convexity Scheme to the DHB Theory

Pekka J. Lahti[1] **and Slawomir Bugajski**[2]

Received December 16, 1983

Some classical and quantum theories are characterized within the convexity approach to probabilistic physical theories. In particular, the structure of the so-called DHB quantum theory will be analyzed. It turns out that the natural generalization of the standard Hilbert space quantum mechanics, the operational one, is such a theory. The operational Hilbert space quantum theory will be reconstructed from the (weak) projection postulate and the complementarity principle. This is then used to argue that the DHB quantum theory is identical with the operational Hilbert space quantum theory.

ACKNOWLEDGMENTS

The preliminary version of this work was prepared during the first author's visit at the Institute of Physics of the Silesian University in Katowice, June 1980. The great hospitality of the hosts is warmly acknowledged. The rather long time which was needed to complete the work was partly caused by the dramatic developments in Poland. The second author is greatly indebted to those who supported him financially at the final stage of this work.

1. INTRODUCTION

This paper will continue the analysis of the foundations of the quantum theory started in (Bugajski and Lahti, 1980), hereafter referred to as FPI. The question posed therein was to what extent the standard Hilbert space quantum theory can be erected on its fundamental principles: the superposition principle, the uncertainty principle, and the complementarity principle. With respect to careful formalizations of these principles within the general

[1]Department of Physical Sciences, University of Turku, 20500 Turku, Finland.
[2]Katowice, Poland.

0020-7748/85/1100-1051$04.50/0

Mackey axiomatics it was demonstrated that these principles are not enough to determine the standard theory. Thus, either (i) the three principles do not exhaust the foundations of the quantum theory, or (ii) the usual formulation of the quantum theory is too specific to reflect properly the very ideas of the theory.

The standard quantum mechanics, though very useful, is highly idealized. This becomes most evident from the fact that observables on the system described by the theory are represented as spectral measures, or as self-adjoint operators, on a Hilbert space. Consequently, measurements on the system, or interactions of the system with its environment, described by the theory are restricted to those which give rise to observables as spectral measures, only. This is connected with the foundational status of the von Neumann–Lüders projection postulate in the standard formulation of the quantum theory. For a long time such an idealized scheme has been known to be too restrictive.

Thus the alternative (ii) above should be taken seriously. We follow here the proposal made in FPI that the very ideas of the quantum theory are most properly reflected in the more general DHB quantum theory, i.e., the quantum theory based directly on the superposition principle, the uncertainty principle, and the complementarity principle. It is the aim of the present investigation to analyze in detail the structure of such a theory.

The structure of the paper is the following. A sufficiently general but powerful theoretical framework is needed. Such a frame is provided by the so-called convexity approach to probabilistic physical theories, which is flexible enough to allow the formulations of the physical ideas important for the investigation at hand. This scheme, with its basic notions and structures, will be introduced in Section 2. In Section 3 we shall indicate how this general convexity scheme is further specified to provide physically interesting convex descriptions. Examples of some convex descriptions which are typically classical and typically quantal are there also given. In the following three Sections (4, 5 and 6) the three fundamental quantum principles are then formulated within the chosen framework and the basic properties, in particular the nonclassical features of the convex descriptions satisfying these principles are pointed out. The question on the interdependence of these principles is then considered in Section 7. In Section 8 we briefly return to the problem of characterizing classical convex descriptions with discussing some consequences of the requirement for the unique decomposability of mixed states. In Section 9 we make use of the great flexibility of the convex scheme in formalizing a measurement theoretical assumption called the projection postulate. This postulate, which is a remarkable weakening of the von Neumann–Lüders postulate, formalizes here the rather acceptable assumption that some observables of a physical system concerned may admit the so-called ideal first-kind measurements of their possible values. In Section 10 we are then finally in position to characterize quantum convex descriptions, and, in particular, the Dirac-Heisenberg-Bohr (DHB) quantum theories. We shall then argue that the DHB quantum theories are the operational Hilbertian descriptions which can essentially be based on the fundamental complementarity principle and on the measurement theoretical idealization here called the projection postulate.

2. THE GENERAL SCHEME

In this preliminary section we shall describe the basic notions and assumptions of the so-called operational or convexity approach to probabilistic (irreducibly or otherwise) physical theories. We prefer the term *convexity approach* as the adjective *operational* will subsequently be used to describe more specific structures. Moreover the term *convexity* is neutral with respect to some semantical or interpretative questions which are not relevant here. The general convexity scheme, CS for short, which will be sketched below, was formulated in the works of Davies and Lewis, Ludwig and Mielnik. This scheme provides a framework for convex descriptions for some physical theories. We do not repeat all the physical and mathematical arguments which, during its development, finally led to this scheme, but we refer the reader to the books of Davies (1976), Gudder (1978), and Ludwig (1982), where also references to original contributions can be found.

The basic notion of the scheme is *state* of a physical system, and the basic operation is forming *mixtures* of states. We do not adopt any particular operational interpretation of these basic notions here, in spite of the fact that some of the terminology used below might remind the reader of the so-called "beam semantics" or "ensemble interpretation" of the approach. The approach followed here is general and it is not, from the outset, engaged to any particular interpretation.

The set S of all states of the system will be equipped with an algebraic structure which allows the distinction between the pure and the mixed states of the system. Such a structure is that of a convex set. It is then a mathematical convenience to consider S as properly placed in a real vector space V. The set S defines a cone $V^+ = \{\lambda\alpha\colon \lambda \in \mathbf{R}^+,\ \alpha \in S\}$, which is a generating one: i.e., $V = V^+ - V^+$. S is a base for the cone V^+ of the real vector space V which is partially ordered by its cone: for any α and β in V, $\alpha \le \beta$ iff (if and only if) $\beta - \alpha \in V^+$. The set S of all states defines the strictly positive linear functional e on V such that $S = \{\alpha \in V^+\colon e(\alpha) = 1\}$. This functional will be called the *strength* or *intensity* functional. Owing to the convex structure of S the distinction between *pure states*, extreme elements of S, and *mixed states*, nonextreme elements of S, can now be made. $\text{Ex}(S)$ denotes the set of all pure states in S.

The physical interpretation attached to S can be extended to the positive cone V^+: any $\lambda\alpha \in V^+$ with $\lambda \in R^+$ and $\alpha \in S$ represents a new state of the system, obtained from α by changing its intensity or strength (or normalization). Thus original states (i.e., the elements of S) will be distinguished as normalized states, whereas the term "state" will be extended over all elements of V^+ including the "empty" state ω, the origin of V. The linear operations $(\alpha, \beta) \to \alpha + \beta$, $(\lambda, \alpha) \to \lambda\alpha$, $\lambda \geq 0$, α, $\beta \in V^+$ preserve their original interpretations as mixing and intensity changing, respectively. In particular, the term *pure state* can thus refer to an element of $\text{Ex}(S)$ or of $\text{Ed}(V^+) := \{\lambda\alpha: \lambda \in R^+, \alpha \in \text{Ex}(S)\}$, where Ed stands for "edge." This extension will be accepted only as a mathematically convenient one, and it has no physical implications. There are also some attempts to avoid the linear structures of V (see, e.g., Gudder, 1978), but it is not of any particular relevance for our considerations.

The base S defines also a natural seminorm $\|\cdot\|$ on V, $\|\alpha\| := \inf\{\lambda \in \mathbf{R}^+, \alpha \in \lambda \text{ conv } (S \cup -S)\}$, which is additive on V^+ and $\|\alpha\| = e(\alpha)$ for $\alpha \in V^+$. Assuming that also countable mixtures of states can be formed, the seminorm $\|\cdot\|$ is actually a norm, the base norm, with respect to which the base normed space (V, S) is complete. We assume also the closedness of V^+ as it is a rather harmless mathematical assumption. Thus we get the basic assumption of the convexity scheme:

(CS) The set of states of a physical system is represented by a norm closed generating cone V^+ of a base normed Banach space (V, S).

Other elements of the theory shall be defined on the basis of (CS) and will be specified according to the theory.

A distinguished role is played by linear positive contracting mappings of V into V, i.e., positive elements of the unit ball of $L(V)$. Some of such mappings are assumed to represent elementary physical *operations* performable on the physical system described by the theory. More complicated experimental arrangements are described by mappings from $\mathbf{B}(\mathbf{R})$—the Boolean lattice of Borel subsets of the real line $\mathbf{R}$, into the set of elementary operations. If $\mathbf{I}$: $\mathbf{B}(\mathbf{R}) \to L(V)$ denotes such a mapping, it is called an *instrument*, and assumed to represent an experimental procedure, iff: (i) $\mathbf{I}(X)$ is an operation for any $X \in \mathbf{B}(\mathbf{R})$, (ii) $e(\mathbf{I}(\mathbf{R})\alpha) = e(\alpha)$ for any $\alpha \in V$, (iii) $\mathbf{I}(\cup X_n) = \sum \mathbf{I}(X_n)$ for any countable family $\{X_n: n \in \mathbf{N}\}$ of pairwise disjoint Borel sets, where the sum converges in the strong operator topology. The notion of instrument is a straightforward generalization of quantum mechanical models of experimental arrangements as discussed, e.g., by von Neumann.

For any operation ϕ there corresponds a linear functional $e(\phi.)$ on V. Such functionals $V \to \mathbf{R}$ are called physical *effects*. It is easy to see that effects are elements of the order interval $[0, e] := \{a \in V^*: 0 \leq a \leq e\}$ of V^*—the topological dual space of V, where 0 is the zero functional. The order here is the one induced on V^* by the order on V, i.e., for any a, $b \in V^*$, $a \leq b$ iff $(b-a)(\alpha) \geq 0$ for any $\alpha \in V^*$. Similarly, if we apply the standard detector, symbolized by the strength functional e, after an instrument $\mathbf{I}$: $\mathbf{B}(\mathbf{R}) \to L(V)$ we get an effect-valued measure on the Borel space $(\mathbf{R}, \mathbf{B}(\mathbf{R}))$, called an *observable*. Actually we can take arbitrary Borel spaces instead of $(\mathbf{R}, \mathbf{B}(\mathbf{R}))$ as the value spaces of instruments and observables. Our restriction to $(\mathbf{R}, \mathbf{B}(\mathbf{R}))$ is not a serious one, since there are indications that the standard Borel spaces cover all the physically relevant cases (Davies, 1976).

The probabilistic character of the convexity scheme follows from the fact that each instrument-normalized state pair $(\mathbf{I}, \alpha)$ defines a probability measure $\mathbf{B}(\mathbf{R}) \to [0, 1]$, $X \to e(\mathbf{I}(X)\alpha)$. Consistent with the interpretation of the numbers $e(\alpha)$, $\alpha \in V^+$, the number $e(\mathbf{I}(X)\alpha)$ is taken to describe the (physical) probability that a measurement with the instrument $\mathbf{I}$ on the system in the state α leads to a result in X.

The mapping $\phi \to e(\phi.)$ of operations on effects defines an equivalence relation on operations, called *isotony*. Thus two operations ϕ_1 and ϕ_2 are isotonic iff the effects $e(\phi_1.)$ and $e(\phi_2.)$ they cause are the same. This isotony relation can easily be extended to instruments, as well: Two instruments $\mathbf{I}_1$ and $\mathbf{I}_2$ are isotonic whenever the observables $\mathbf{A}_1$ and $\mathbf{A}_2$ they define are the same.

The order interval $[0, e] \subset V^*$ is convex and weak* compact. The Krein–Milman theorem says now that it possesses a set $\mathrm{Ex}[0, e]$ of extreme points which is sufficiently rich to allow a weak* approximation of any element of $[0, e]$ by finite combinations of elements of $\mathrm{Ex}[0, e]$. Effects belonging to $\mathrm{Ex}[0, e]$ are called extreme effects (decision effects by Ludwig), whereas nonextreme effects will occasionally be called fuzzy effects.

The space V is ordered by the cone V^+ defined by the base S. If this order is a lattice order, S is a (Choquet) simplex (Alfsen, 1971; Asimow and Ellis, 1980)—a structure that has relevance for classical descriptions. If $V = V^+ - V^+$ is a vector lattice, then also its dual V^* is a vector lattice. In this case the interval $[0, e]$ is a lattice where, in particular, the greatest lower bound of any two elements a, $b \in [0, e]$ is given by $(a \wedge b)\ (\alpha) = \inf\{a(\alpha), b(\alpha): \alpha \in V\}$.

Thus in this special case we have the result

$$\text{if } a \wedge b = 0 \qquad \text{then } a \leq e - b$$

This fact will be useful later on. Let us note also that in this case the set Ex[0, e] is a Boolean lattice, with $a \to a^{\perp} := e - a$ as the orthocomplementation (Schaeffer, 1974).

3. EXAMPLES OF CONVEX DESCRIPTIONS

In its generality the *convexity scheme* CS provides us only with a prestructure of a physical theory. In particular, it is not specifically quantal nor classical, but the two types of theories can be formulated in terms of CS. The description of any physical system within this scheme requires more specification. What we have to specify are the set of states of the physical system, and the set of admissible (physically relevant) operations on the system. The *convex description* (CD) thus arises in specifying the couple (S_D, O_D) which consists of the sets S_D and O_D of all (normalized) states and of all admissible operations of the description D. CD denotes the family of all convex descriptions.

A convex description D will be called classical if S_D is a simplex, i.e., if the base normed Banach space (V_D, S_D) generated by S_D is a Banach lattice. The family of all classical convex descriptions will be denoted by CCD. Classical descriptions can be based on a measurement theoretical assumption, called classical ideal, according to which the mixing of states results from ignoring some relevant physical conditions in preparing the state of the system—a negligency which, in principle, could be eliminated.

Nonclassical descriptions are not necessarily quantal. Quantal descriptions are distinguished by their ability to account for the physical fact of the existence of the universal quantum of action. When basing quantal descriptions on this fact we face the problem of how to incorporate the quantum of action in the general scheme. The notion of a quantum description will be discussed in the following sections. The family of such descriptions will be denoted by QCD.

If the set of admissible operations of a convex description D is the set of all formally possible operations on V_D the description D is called "operational" or "fuzzy." The family of all such descriptions is denoted by OCD where O stands for "operational." In this case the set of effects equals to the whole order interval $[0, e]$ of V_D^*. As the maximally large set O_D contains all the operations on the state space V_D the physical idea behind OCD is simply that it allows one to study all kinds of physical influences on the considered system, which can be described within the approach. In particular OCD allows one to describe several kinds of measurements (ideal, nonideal, first kind, second kind, etc.) on the system and, e.g., its interactions with "reservoir," etc. (see, e.g., Davies, 1976; Busch and Lahti, 1985).

The other extreme of the convex descriptions are those with "maximally restricted" set of operations as the admissible operations. With this we mean convex descriptions with the set of all admissible operations consisting of the so-called filtering operations (to be discussed in Section 9) together with some other operations, which give rise to the trivial effect e. The set of effects of such a theory contains the extreme (nonfuzzy) elements of $[0, e]$ only. The family of such convex descriptions will be denoted as RCD, where R stands for "restricted."

3.1. Classical Convex Descriptions

Classical theories of physics are based on several idealizing measurement theoretical assumptions which might collectively be called *classical ideal.* The two best known of such assumptions are the *compatibility assumption* and *the assumption on the unique decomposability of mixed states.* Intuitively, the compatibility assumption claims the order independence of any two measurements on the system, whereas the unique decomposability of mixtures—assumption UDM says that in our description of physical systems we have to refer to mixed states only when we ignore some relevant physical conditions in preparing the state of the system, and that this ignorance is, in principle, avoidable. In the convexity approach one begins with states as the primary concepts and with forming mixtures of states as the primary operation of the theory. Thus it is the second of the two idealizing assumptions which now readily lends itself for a formalization. We shall comment the compatibility assumption and its relation to UDM further in Section 8.

The assumption UDM has a twofold consequence on the set S of all (normalized) states of the system: *Firstly,* the set $\mathrm{Ex}(S)$ of pure (extreme) states in S should be sufficiently rich so that each state α in S can be expressed as a convex combination (countable or otherwise) of the elements in $\mathrm{Ex}(S)$; *secondly,* this decomposition should be unique. This then allows one to say that if the system is in a mixed state $\alpha \in S$ it is actually in one of its pure components $\alpha_i \in \mathrm{Ex}(S)$, and the weights λ_i with which the pure states α_i participate the decomposition $\sum \lambda_i \alpha_i$ describe our knowledge on the actual state of the system. Formally UDM thus requires the simplicial shape for S. This suffices to justify the conception of classical convex description:

> A convex description (V, S) is a *classical convex description* iff V is a lattice in the natural ordering.

Thus for any $(V, S) \in \mathrm{CCD}$ S is a (Choquet) simplex.

There are at least two representations of this abstract structure which are of interest for physics: $M_{\mathbf{R}}(\Omega)$—the vector space of regular countable additive real Borel measures on a compact Hausdorff space Ω with the natural Banach structure of the topological dual $C_{\mathbf{R}}(\Omega)^*$ of the family $C_{\mathbf{R}}(\Omega)$ of all continuous real-valued functions on Ω; and $L_1(\Omega, \mu)$—the Lebesque space of measurable functions on a σ-finite measure space (Ω, μ). In both cases Ω represents the phase space (or its compactification) of the corresponding classical theory.

Any of the two representations of (V, S) could serve as a starting point for either operational or restricted classical descriptions. Thus two concrete elements of OCCD ($:=$CCD$\cap$OCD) and two of RCCD($:=$CCD$\cap$RCD) will be achieved. We shall now briefly comment on them.

The classical state space $M_{\mathbf{R}}(\Omega)$ is naturally ordered by the convex cone of positive measures on Ω. The positive cone $M_{\mathbf{R}}(\Omega)^+$ is generated by the set $M_{\mathbf{R}}(\Omega)_1^+$ of probability measures, which in the present approach represents the set of normalized states of the physical system concerned. The phase space Ω is here assumed to be compact and Hausdorff. The first assumption does not hold for usual phase space theories (where the phase space is taken to be the real Euclidean space $\mathbf{R}^{2n}$), so we would consider Ω rather as a compactification of the usual phase space. The concrete form of this compactification depends on the considered problems. The compactification which is natural for the operational approach is obtained as follows: We construct the set $M_{\mathbf{R}}(\Omega')$, where Ω' is the original locally compact phase space, and embed Ω' into $\mathrm{Ex}(M_{\mathbf{R}}(\Omega')^{**})_1^+$—the extreme boundary of the base of the base-normed Banach space $M_{\mathbf{R}}(\Omega')^{**}$. This is actually the Čech–Stone compactification of Ω'. We thus have a compact Hausdorff phase space Ω and the state space $M_{\mathbf{R}}(\Omega)$ of the convex description. The set $M_{\mathbf{R}}(\Omega)_1^+$, i.e., the base of $M_{\mathbf{R}}(\Omega)$, is a regular simplex and $\mathrm{Ex}(M_{\mathbf{R}}(\Omega)_1^+) \cong \Omega$ (Alfsen, 1971). The first property guarantees that UDM holds whereas the second property says that the pure states of the description can be identified with the points of the phase space. The order interval $[0, e]$ of $M_{\mathbf{R}}(\Omega)^*$ contains the set of all positive measurable functions on Ω less than the constantly one function, but is essentially larger than this set. All characteristic functions of Borel subsets of Ω can be naturally identified with extreme elements of $[0, e]$, but the latter set contains other elements as well. Thus the convex scheme provides a rather unconveniently general frame. Anyway, the set $[0, e]$ of effects of the description is a complete distributive lattice with the property: if $a \wedge b = 0$, then $a \leq b^\perp = e - b$ for any two a and b in $[0, e]$. The mapping $a \to a^\perp$ is however, not an orthocomplementation, so that $([0, e], \leq, ^\perp)$ fails to be Boolean, in general. In order to get the set of effects exactly equal to the set of all characteristic functions of Borel subsets of Ω we have to restrict the set of possible operations. The standard way of doing this is to take O_{RCl} as consisting of all mappings ϕ_X, X is a Borel subset of Ω, defined by

$a(\phi_X(\alpha)) := \int_X a\, d\alpha$ for any $\alpha \in M_{\mathbf{R}}(\Omega)_1^+$ and any bounded measurable function a on Ω. Such operations are just conditionalizations of the standard probability theory, so that the pair $(M_{\mathbf{R}}(\Omega_1^+, O_{\mathrm{RC1}})$ belonging to the class RCCD can be viewed as the convex description of classical probability theory. If we add to O_{RC1} operations which correspond to canonical transformations of Ω we get the set O_{RC2}, which together with $M_{\mathbf{R}}(\Omega)_1^+$ could be considered as the convex description of classical statistical mechanics. The description specified by the couple $(M_{\mathbf{R}}(\Omega)_1^+, O_{\mathrm{RC2}})$ belongs also to the family RCCD. In order to describe classical open systems we have to admit a larger set of operations than O_{RC2}. The natural choice is just the whole set O_{OC} of positive linear contracting transformations on $M_{\mathbf{R}}(\Omega)$. This leads to the pair $(M_{\mathbf{R}}(\Omega)_1^+, O_{\mathrm{OC}})$ generating a specific operational classical description, which is basic for classical mechanics of open systems. This kind of description is a typical sample of the class OCCD.

Other examples of classical convex descriptions arise if we take the Lebesque space $L_1(\Omega, \mu)$ as the state space. $L_1(\Omega, \mu)$ consists of all essentially bounded μ-integrable real functions on the phase space Ω defined up to set of μ-measure zero. It carries the natural structure of base normed Banach space with the standard strength functional $e(\alpha) := \int \alpha\, d\mu$ for any $\alpha \in L_1(\Omega, \mu)$. It is a lattice with respect to the pointwise ordering, so its base $L_1(\Omega, \mu)_1^+$ is a simplex. In the typical case of $\Omega = \mathbf{R}^{2n}$ and μ being the Lebesque measure λ, there are no pure states: $\mathrm{Ex}(L_1(\mathbf{R}^{2n}, \lambda)_1^+)$ is empty. Such a state space could be considered as the restriction of $M_{\mathbf{R}}(\Omega)$ to the measures absolutely continuous with respect to some fundamental measure μ. The model for classical statistical theory based on $L_1(\Omega, \mu)$ has frequently been advanced; see, e.g., Primas, 1981. An advantage of this approach is that $\mathrm{Ex}[0, e] \subset L_1(\Omega, \mu)^*[\cong L_\infty(\Omega, \mu)]$ consists only of characteristic functions of Borel subsets of Ω modulo the subsets of μ-measure zero. Thus the effects corresponding to O_{RC1} exhaust the set of all extreme elements of $[0, e]$, and the same holds for O_{RC2}. We have thus two more examples of descriptions in RCCD (defined by the pairs $(L_1(\Omega, \mu)_1^+, O_{\mathrm{RC1}})$ and $(L_1(\Omega, \mu)_1^+, O_{\mathrm{RC2}})$). The corresponding description in the family OCCD is defined by the pair $(L_1(\Omega, \mu)_1^+, O_{\mathrm{OC}})$. The description of classical systems based on $L_1(\Omega, \mu)$ is in fact most general, as any base normed Banach lattice (AL space) can be represented as $L_1(\Omega, \mu)$ for some locally compact measure space (Ω, μ) [the Kakutani theorem; see Schaeffer (1971) or Asimow and Ellis (1980)].

3.2. Quantum Convex Descriptions

Quantum theories of physics are based on the physical fact of the existence of the universal quantum of action, symbolized by the Planck constant h. Classical theories of physics are idealizations which can be unambiguously applied only in the limit where all actions involved are large compared with the quantum. Though $h>0$ or $h=0$, exclusively (physical actions cannot have negative values) there is no obvious way to characterize quantum convex descriptions. Such descriptions should not belong to the family CCD of classical convex descriptions, but this is not enough to provide a characterization of quantum descriptions. The problem of incorporating the physical fact "$h>0$" into the scheme is highly nontrivial, and it will be discussed subsequently. What we shall do now is to list some examples of theories which are surely of quantal nature so that they should belong to the family of QCD.

A large class of working quantum theories fits into the following scheme: Let $V(\mathbf{A})$ be the real part of the predual Banach space of a W^* algebra $\mathbf{A}$. $V(\mathbf{A})$ carries as a natural structure that of the base normed Banach space with the base $S(\mathbf{A})$ consisting of all ultraweakly continuous positive normalized functionals on $\mathbf{A}$. The Banach dual $V(\mathbf{A})^*$, being an order unit Banach space, can canonically be identified with the self-adjoint part of $\mathbf{A}$, with the unit of $\mathbf{A}$ as the strength functional e and the complete orthomodular, semimodular lattice of projection of $\mathbf{A}$ as the set of extreme effects $\mathrm{Ex}[0, e]$. It is worth noting that $(V(\mathbf{A})^*, e)$, which is an ordered linear space, is not a lattice except the case of commutative $\mathbf{A}$. Nevertheless, $\mathrm{Ex}[0, e]$ is a complete lattice with respect to the induced order (Primas, 1981). With a fixed $(V(\mathbf{A}), S(\mathbf{A}))$ we again meet the problem of deciding which elements of $L(V(\mathbf{A}))$ will be considered as describing physically admissible operations on the system. According to our division we get the operational and restricted variants of the scheme, or one of the intermediate cases.

In W^*-algebraic theories there is a distinguished class of operations which correspond to the ideal first-kind measurements. This class, which leads to a restricted description, can be defined as follows. Let $a\in \mathrm{Ex}[0, e]\subset V(\mathbf{A})^*$, so that a is a projection of $\mathbf{A}$. We define $\phi_a^*\in L(V(\mathbf{A})^*)$ by $b\to\phi_a^*b:=aba$, and $\phi_a\in L(V(\mathbf{A}))(\phi_a^*b)(\alpha):=b(\phi_a\alpha)$ for any $\alpha\in V(\mathbf{A})$. It is easy to see that ϕ_a is positive and of norm less than 1. Moreover ϕ_a belongs to the isotony class of a: $e(\phi_a\alpha)=(\phi_a^*e)(\alpha)=(aea)(\alpha)=a(\alpha)$. It is clear that the operations belonging to $O_{\mathrm{RW1}}^*:=\{\phi_a: a\in \mathrm{Ex}[0, e]\}$ are analogous to conditionalizations, so that the theories described by $(S(\mathbf{A}), O_{\mathrm{RW1}}^*)$ could be considered as noncommutative probability theories based on W^* algebras, cf., e.g., Gudder and Marchand, 1972. If we supplement the set of admissible operations by those corresponding to unitary transformations on the Hilbert space underlying the concrete representation of $\mathbf{A}$ as a von Neumann algebra, we get the pair $(S(\mathbf{A}), O_{\mathrm{RW2}}^*)$ which is fundamental for many physical theories. For a review and discussion of such W^* theories see Primas, 1981 and Emch, 1972.

An important property of any W^* theory is that there is a natural bijection between the set of all Ex[0, e]-valued measures on $\mathbf{R}^1$ with compact supports and elements of V^*. This is the content of the spectral theory for W^* algebras (see, e.g., Alfsen and Schultz, 1976). This spectral property justifies the name "algebra of observables" for **A**. In the case of generalized probability theory **A** is the algebra of random variables.

A W^* theory with commutativen **A** is classical. An example of such a theory is the one based on $(L_1(\Omega, \mu)_1^+, O_{RC1})$ discussed above. W^* theories with noncommutative **A** are considered as quantal, partly because **A** can be realized as a von Neumann algebra of operators acting on a complex Hilbert space, which makes it possible to construct the whole traditional machinery needed for concrete calculations. A paradigm of such a theory is the standard quantum mechanics, where **A** is just the von Neumann algebra $L(\mathbf{H})$ of all bounded linear operators acting on a complex separable (generally infinite dimensional) Hilbert space **H**. In this case $V(\mathbf{A})$ is the space of all self-adjoint trace class operators with the trace norm and $S(\mathbf{A})$ consists of the positive trace class operators of trace one. The operations ϕ_a, $a \in \mathrm{Ex}[0, e] \subset V(\mathbf{A})^*$, are exactly the filtering operations discussed by the standard measurement theory, and the existence of the unique ϕ_a for any a is more or less the content of the von Neumann-Lüders projection postulate, to be discussed subsequently.

Noncommutative probability theories $(S(\mathbf{A}), O^*_{RW1})$ as well as quantum W^* theories $(S(\mathbf{A}), O^*_{RW2})$ refer to isolated physical systems, or at most to systems coupled to environment in a very special way. Quantum open systems are described by theories $(S(\mathbf{A}), O^*_{OW})$ with noncommutativen **A**, where O^*_{OW} is the set of all linear positive contracting transformations on $V(\mathbf{A})$. Like the classical operational descriptions, the extension of the set of admissible elementary operations to O^*_{OW} introduce fuzzy effects represented by nonextreme elements in $[0, e] \subset V(\mathbf{A})^*$. The set of $\mathbf{R}^1$-based observables with compact support is now "larger" than **A** because of fuzzy-effect-valued (or semispectral) measures. The operational W^* theories with noncommutative **A** are quantal members of the family OCD. For $\mathbf{A} = L(\mathbf{H})$ we get the operational Hilbert space quantum mechanics (Davies, 1976; Gudder, 1978). Owing to their special importance we introduce the families CD_{OHQ} and CD_{SHQ} of operational Hilbert space quantum mechanics and its restricted variant, the standard Hilbert space quantum mechanics.

4. COMPLEMENTARITY

Experimental arrangements which permit unambiguous (operational) definitions of complementary observables are mutually exclusive. This intuitive idea of Pauli and Bohr was followed in (Lahti, 1980a) to define complementary observables in the Hilbert space quantum theory as well as in the more general quantum logic approach. However, in those approaches the notion of complementary observables has no explict reference to the mutual exclusiveness of the corresponding experimental arrangements. In the present scheme, where observables are defined through instruments, this connection can be made more explicit.

Each instrument $\mathbf{I}$: $\mathbf{B}(\mathbf{R}) \to O$, $X \to \mathbf{I}(X)$ defines an observable $\mathbf{A}$: $\mathbf{B}(\mathbf{R}) \to [0, e]$, $X \to \mathbf{A}(X)$ through the probabilistic relation $e(\mathbf{I}(X)\alpha) = \mathbf{A}(X)(\alpha)$ for any $\alpha \in V$, $X \in \mathbf{B}(\mathbf{R})$. Moreover, each observable is determined in such a way by at least one instrument. Through its operations such an instrument characterizes an experimental arrangement which can be used to measure all the possible values of the observables, or which serves to define unambiguously the observable. Actually, for each $\mathbf{A}$ there corresponds a unique family (an isotony class) of instruments $\mathbf{I}_i^{\mathbf{A}}$, $i \in I(\mathbf{A})$ = a suitable index set, which contains all the possible measurements of $\mathbf{A}$, i.e., $\mathbf{A}$-measurements which are describable within the scheme as operations.

Let $\mathbf{I}_i^{\mathbf{A}}$, $i \in I(\mathbf{A})$, and $\mathbf{I}_j^{\mathbf{B}}$, $j \in I(\mathbf{B})$, be any two instruments associated with the observables $\mathbf{A}$ and $\mathbf{B}$, respectively. The operations $\mathbf{I}_i^{\mathbf{A}}(X)$ and $\mathbf{I}_j^{\mathbf{B}}(Y)$, $X, Y \in \mathbf{B}(\mathbf{R})$, describe some particular measurements of $\mathbf{A}$ and $\mathbf{B}$; $e(\mathbf{I}_i^{\mathbf{A}}(X)\alpha)$ being the probability that a measurement of $\mathbf{A}$, with the instrument $\mathbf{I}_i^{\mathbf{A}}$ on the system in the state α yields a result in X. Assume now that there exists an operation ϕ such that $e(\phi\alpha) \leq e(\mathbf{I}_i^{\mathbf{A}}(X)\alpha)$ and $e(\phi\alpha) \leq e(\mathbf{I}_j^{\mathbf{B}}(Y)\alpha)$ for any state α. Such an operation, when applied, gives us (probabilistic) information which is potentially contained both in the operation $\mathbf{I}_i^{\mathbf{A}}(X)$ and $\mathbf{I}_j^{\mathbf{B}}(Y)$, too. It is a joint measurement of the observables $\mathbf{A}$ and $\mathbf{B}$, associated with the value sets X and Y. (For further analyses of the notion of joint measurement within the convexity scheme, see Busch and Lahti, 1984.) Intuitively, it is the lack of such measurements that the notion of complementarity aims at characterizing. This then leads to the following definition (cf. Lahti, 1980a):

(C) Observables $\mathbf{A}$ and $\mathbf{B}$ are *complementary* iff l.b. $\{\mathbf{A}(X), \mathbf{B}(Y)\} = \{0\}$ for any bounded X and Y in $\mathbf{B}(\mathbf{R})$ for which $\mathbf{A}(X) \neq e \neq \mathbf{B}(Y)$.

Here l.b.$\{.,.\}$ denotes the set of lower bounds of the elements in question in the relevant poset. The restriction to nonmaximal elements allows the possibility that also bounded observables, i.e., observables with bounded value sets, might be complementary. The restriction to bounded sets, on the other hand, can be motivated by considering operational definitions of observables. In fact, closed intervals would already suffice here. If $\mathbf{A}$ and $\mathbf{B}$ are complementary observables then any two instruments $\mathbf{I}_i^{\mathbf{A}}$, $i \in I(\mathbf{A})$, and $\mathbf{I}_j^{\mathbf{B}}$, $j \in I(\mathbf{B})$, are *mutually exclusive*, i.e., l.b. $\{\mathbf{I}_i^{\mathbf{A}}(X), \mathbf{I}_j^{\mathbf{B}}(Y)\} = \{0\}$ for any two bounded X and y in $\mathbf{B}(\mathbf{R})$ for which neither $\mathbf{I}_i^{\mathbf{A}}(X)$ nor $\mathbf{I}_j^{\mathbf{B}}(Y)$ is maximal. Also the converse holds true, i.e. if any two instruments $\mathbf{I}_i^{\mathbf{A}}$, $i \in I(\mathbf{A})$, and $\mathbf{I}_j^{\mathbf{B}}$, $j \in I(\mathbf{B})$ associated with $\mathbf{A}$ and $\mathbf{B}$ are mutually exclusive, then $\mathbf{A}$ and $\mathbf{B}$ are complementary. Indeed, if for given two effects a and b, like $\mathbf{A}(X)$ and $\mathbf{B}(Y)$, X, $Y \in \mathbf{B}(\mathbf{R})$, there exists an effect c which lies below them, i.e. $c \in \text{l.b.}\{a, b\}$ then, for any $\alpha \in \mathbf{B}$, the operation $\phi_\alpha^c : \beta \mapsto \phi_\alpha^c(\beta) := c(\beta)\alpha$ lies below the operations ϕ_α^c and ϕ_α^b, i.e. $\phi_\alpha^c \in \text{l.b.}\{\phi_\alpha^a, \phi_\alpha^b\}$. As $\phi_\alpha^c = 0$ only if $c = 0$ the above claim is hereby justified. This shows that within the present approach the notion of complementarity of observables can directly be based on the notion of mutual exclusiveness of the defining instruments.

(CP) A convex description (S, O) satisfies the *complementarity principle* if there exist, at least, two nonconstant observables which are complementary to each other.

CD_{CP} denotes the family of such descriptions. CD_{CP} is not empty as $CD_{SHQ} \subset CD_{CP}$. The two most characteristic properties of convex descriptions allowing complementary observables are given in the following two results.

Theorem 1. If $\mathbf{A}_1$ and $\mathbf{A}_2$ are complementary observables then $\mathbf{I}_1 \circ \mathbf{I}_2 \neq \mathbf{I}_2 \circ \mathbf{I}_1$ for any two instruments $\mathbf{I}_1$ and $\mathbf{I}_2$ which define these observables.

Proof. Recall first that two operations ϕ_1 and ϕ_2 commute weakly iff (i) $e \circ (\phi_1 \circ \phi_2) = e \circ (\phi_2 \circ \phi_1)$, and strongly if (ii) $\phi_1 \circ \phi_2 = \phi_2 \circ \phi_1$. Clearly, (ii) implies (i). Recall also that for any two operations ϕ_1 and ϕ_2, $e \circ (\phi_1 \circ \phi_2) \leq e \circ \phi_2$. Let $\mathbf{A}_1$ and $\mathbf{A}_2$ be complementary. Suppose that $\mathbf{I}_1$ and $\mathbf{I}_2$ commute: $\mathbf{I}_1 \circ \mathbf{I}_2 = \mathbf{I}_2 \circ \mathbf{I}_1$, i.e., $\mathbf{I}_1(X) \circ \mathbf{I}_2(Y) = \mathbf{I}_2(Y) \circ \mathbf{I}_1(X)$ for any two $X, Y \in \mathbf{B}(\mathbf{R})$. Thus for every $X, Y \in \mathbf{B}(\mathbf{R})$ we have $e \circ (\mathbf{I}_1(X) \circ \mathbf{I}_2(Y)) = e \circ (\mathbf{I}_2(Y) \circ \mathbf{I}_1(X))$, but also $e \circ (\mathbf{I}_1(X) \circ \mathbf{I}_2(Y)) \leq e \circ \mathbf{I}_2(Y)$ and $e \circ (\mathbf{I}_2(Y) \circ \mathbf{I}_1(X)) \leq e \circ \mathbf{I}_1(X)$. Thus for any bounded $X, Y \in \mathbf{B}(\mathbf{R})$ [for which $\mathbf{I}_1(X)$ and $\mathbf{I}_2(Y)$ are not maximal] $e \circ (\mathbf{I}_1(X) \circ \mathbf{I}_2(Y)) = e \circ (\mathbf{I}_2(Y) \circ \mathbf{I}_1(X)) = 0$. Let $\mathbf{I} = \mathbf{I}_1 \circ \mathbf{I}_2$ $(= \mathbf{I}_2 \circ \mathbf{I}_1)$ be the composed instrument (Davies and Lewis, 1970). Thus for every $X, Y \in \mathbf{B}(\mathbf{R})$, $\mathbf{I}(X \times Y) = \mathbf{I}_1(X) \circ \mathbf{I}_2(Y) = \mathbf{I}_2(Y) \circ \mathbf{I}_1(X)$. Let (I_{nm}) form a disjoint cover for $\mathbf{R} \times \mathbf{R}$ with $I_{nm} = X_n \times Y_m$, $X_n, Y_m \in \mathbf{B}(\mathbf{R})$. Now we have $e = e \circ \mathbf{I}(\mathbf{R} \times \mathbf{R}) = e \circ \mathbf{I}(\cup I_{nm}) = e \circ (\sum \mathbf{I}(I_{nm})) = \sum e \circ \mathbf{I}(I_{nm}) = \sum e \circ (\mathbf{I}_1(X_n) \circ \mathbf{I}_2(Y_m)) = 0$. But this is a contradiction, establishing the claim of the theorem.

Accordingly, the experimental arrangements which permit the unambiguous definitions of complementary observables are not only mutually exclusive but also noncommutative. Hence the result of an order measurement of complementary observables depends, in general, on the order in which the corresponding experimental arrangements are applied, a fact which seems to be widely accepted.

Theorem 2. If $(S, O) \in CD_{CP}$ then there exist at least two effects in $[0, e]$, say, a and b, such that $a \wedge b = 0$, $a \not\leq e - b$.

Proof. Assume that if $a \wedge b = 0$ then $a \leq e - b$ for any two effects a and b in $[0, e]$ and let $\mathbf{A}$ and $\mathbf{B}$ be complementary. Thus $\mathbf{A}(X) \leq e - \mathbf{B}(Y)$ for any two bounded $X, Y \in \mathbf{B}(\mathbf{R})$ [for which $\mathbf{A}(X) \neq e \neq \mathbf{B}(Y)$]. Let (I_n) form a disjoint bounded Borel cover for the real line $\mathbf{R}$. Now for any $n \in \mathbf{N}$, $\mathbf{A}(I_n) \leq e - \mathbf{B}(Y)$ for any bounded $Y \in \mathbf{B}(\mathbf{R})$ [for which $\mathbf{A}(I_n) \neq e \neq \mathbf{B}(Y)$] showing that $\mathbf{A}(\cup_1^N I_n) = \sum_1^N \mathbf{A}(I_n) \leq e - \mathbf{B}(Y)$ for any natural number $\mathbf{N}$. Hence $e = \lim \sum_1^N \mathbf{A}(I_n) \leq e - \mathbf{B}(Y)$ for any bounded $Y \in \mathbf{B}(\mathbf{R})$, for which $\mathbf{B}(Y) \neq e$. But this is a contradiction, closing the proof of the theorem.

Loosely speaking, this theorem shows that whenever we have complementary observables we also have disjoint effects which are not "orthogonal." When confronting this result with the fact that disjointness implies orthogonality in simplicial cases we also have the following:

Corollary 3. If $(S, O) \in CD_{CP}$ then S cannot be a simplex. Moreover $CD_{CP} \neq \varnothing$ and $CD_{CP} \cap CCD = \varnothing$.

5. SUPERPOSITION

States are the undefined axiomatic elements of the convexity scheme whereas its primary operation is the formation of mixtures of states. The state space is a linear space where the linearity reflects the idea of mixing states. It is, however, well known that there exists an important class of experimental conditions under which one observes (quantum) effects, e.g., interference effects, which cannot be understood on the above basis only. Some additional constituent is needed, and this is the superposition principle. After Dirac's monumental work (1930) this principle has been considered as one of the most important constituents of quantum descriptions. The Diracian tradition is very strong and it is closely followed in several well-known texts.

To underline the difference between the standard approach and the present one, we recall that in Dirac's approach the formulation of the principle of superposition led to the requirement of the linearity of the state space, too. An adequate analysis of some experiments which support this principle is given in Gerjuoy (1973); see also Badurek et al. (1983) and Summhammer et al. (1983).

The superposition principle as formulated in standard quantum mechanics assumes that pure states of an elementary quantum system are represented (up to a phase factor) by vectors of a complex Hilbert space $\mathbf{H}$, and it ascribes a fundamental physical meaning to the linearity of $\mathbf{H}$. In the corresponding operational description (see Section 3.2) the Hilbert space **H** does not appear explicitly. It is based on the real vector space of self-adjoint trace class operators on **H** equipped with the trace norm. This space shall be denoted as $V(\mathbf{H})$. The set $\mathrm{Ex}(V(\mathbf{H})_1^+)$ of extreme elements of the base $V(\mathbf{H})_1^+$ for the cone $V(\mathbf{H})^+$ is in a one-to-one correspondence with the factor set $\mathbf{H}/\sim$, where $\sim$ is the relation $\phi \sim \psi$ iff $\phi = \psi e^{i\lambda}$ for some real λ, with $\phi, \psi \in \mathbf{H}$. The relation $\sim$ is not congruent with the linear structure of **H** so that the operation of superposing states cannot be expressed as a mapping $\mathrm{Ex}(V(\mathbf{H})_1^+) \times \mathrm{Ex}(V(\mathbf{H})_1^+) \to \mathrm{Ex}(V(\mathbf{H})_1^+)$. The same difficulty plagues the quantum logic approach, as well as any other general scheme which does not start with an underlying Hilbert space housing the pure states of the description.

Following the quantum logic approach (see FPI and references therein) the notion of superposition could now be defined as follows: A pure state $\alpha \in \mathrm{Ex}(S)$ is a superposition of pure states $\alpha_1, \alpha_2 \in \mathrm{Ex}(S)$ iff $a(\alpha_1) = a(\alpha_2) = 0$ implies $a(\alpha) = 0$ for any extreme effect $a \in \mathrm{Ex}[0, e] \subset V^*$. This definition should be modified in a few aspects. Firstly, as any effect can be weak* approximated by convex combinations of extreme effects, we do not need to restrict ourselves to extreme effects. Without changing the notion of superposition we can refer to a broader class of effects. The notion of superposition of states does not depend on the "degree of fuzziness" of the description. Secondly, there is no need to consider normalized pure states only. In the convexity scheme a physical meaning can be attached to the whole positive cone of V, so that the assumption of normalization could simply be removed. Finally, the preliminary definition can easily be extended to describe superpositions of arbitrary sets of pure states. Thus we get the following:

(S) A pure state $\alpha \in V^+$ is a *superposition* of a set $W \subset \mathrm{Ed}(V^+)$ of pure states iff $a(\beta) = 0$ for all $\beta \in W$ implies $a(\alpha) = 0$ for any effect a of the description (V, S) (provided that the extreme elements of $[0, e] \subset V^*$ belong to the set of admissible effects).

A leading idea of the convexity approach is to construct the whole theory on the convex set of (normalized) states S. Hence it might be more elegant to formulate the notion of superposition of states in terms of geometry of S only, without any reference to effects. This can easily be done. Any $a \in [0, e]$ defines a (w-closed) hyperplane, $a^{-1}(0)$, and vice versa: any w-closed hyperplane which does not intersect the interior of V^+ can be considered as $a^{-1}(0)$ for some $a \in [0, e]$. A closed hyperplane $M \subset V$ will be called supporting (for V^+) if $M \cap V^+ = \varnothing$ or $V^+ \backslash M$ is still convex. The intersection of a supporting hyperplane M and the convex positive cone V^+ is a face of V^+, i.e., a convex subset F of V^+ such that $\lambda\alpha_1 + (1-\lambda)\alpha_2 \in F$ for $\alpha_1, \alpha_2 \in V^+$, $\lambda \in [0, 1]$ iff $\alpha_1, \alpha_2 \in F$. The face $M \cap V^+$ is called an exposed face of V^+, and an intersection of any family of exposed faces is called a semiexposed face. The smallest semiexposed face of V^+ containing a set $W \subset V^+$ will be denoted by $F(W)$. A physical interpretation of these notions is discussed by Mielnik and Rüttimann; faces are related to properties of the physical system in question (Mielnik, 1969), whereas exposed faces are related to detectable properties of the system (Rüttimann, 1981). The above discussion suggests the following definition:

(S′) A pure state $\alpha \in \mathrm{Ed}(V^+)$ is a *superposition* of a family $W \subset \mathrm{Ed}(V^+)$ of pure states iff α belongs to $F(W)$, the smallest semiexposed face of V^+ containing W, but does not belong to W.

Note, that all mixtures of a family $W \subset V^+$ are contained in the w closure of the set $\mathrm{Lin}_+(W)$ of all linear combinations with nonnegative coefficients of the elements of W. If W contains only pure states, i.e., $W \subset \mathrm{Ed}(V^+)$, then the w closure $\overline{\mathrm{Lin}_+(W)}^w$ of $\mathrm{Lin}_+(W)$ is also contained in the semiexposed face $F(W)$ generated by W. The elements of $F(W)$ which are both mixtures and superpositions of W will be called trivial superpositions. Obviously, if α is a trivial superposition of W, then there is an $\alpha_1 \in {}^w\bar{W}$ such that $\alpha = \lambda\alpha_1$ for some positive λ. Thus we can see $F(W)$ as containing the set of all superpositions as well as mixtures of elements of W. It is obvious that the superpositions produced by $W \subset \mathrm{Ed}(V^+)$ are determined by the set W. They do not depend on the order of elements of W nor on different clusterings of W. Thus the forming of superpositions is "symmetric" and "associative" so that the notion of superposition enjoys the properties (2c), (2d), (2e) considered in (Lahti, 1985). It does not need to satisfy, however, the property (2a), which in the present approach is equivalent to (2b), i.e., it is not guaranteed that the set $\{\alpha, \omega\}$ for arbitrary $\alpha \in \mathrm{Ed}(V^+)$ produces only trivial superpositions. This is due to the fact that the ray $\{\lambda\alpha: \lambda \in \mathbf{R}^+\}$ for $\alpha \in \mathrm{Ed}(V^+)$ is not necessarily a semiexposed face of V^+. Simple examples illustrating this fact can easily be invented (see, e.g., Alfsen and Schultz, 1976, Fig. 1.p.1.). On the other hand this "idempotency property" for quantal superpositions is of a fundamental nature, and it should hold in convex descriptions of quantum systems. We shall incorporate this property into the "convex" form of the principle of superposition of states.

As concerns the superposition principle, it seems that the principle should state, in its weakest form, simply the existence of superpositions of states. A stronger form, valid only for elementary systems (without superselection rules) should state that any pair of different pure states produces nontrivial superpositions (cf. FPI). Thus we have the following:

(SP) (i) For any pair of pure states α_1, $\alpha_2 \in \mathrm{Ed}(V^+)\backslash\{\omega\}$ there exists a pure state $\alpha \in \mathrm{Ed}(V^+)$ which is their (nontrivial) superposition, i.e., α does not lie in the rays generated by α_1 and α_2, but belongs to $F(\{\alpha_1, \alpha_2\})$.

(ii) No pair $\{\alpha, \omega\}$, $\alpha \in \text{Ed}(V^+)$, produces nontrivial superpositions, i.e., $F(\{\alpha\}) = \{\lambda\alpha: \lambda \in \mathbf{R}^+\}$ for any $\alpha \in \text{Ed}(V^+)$.

CD_{SP} denotes the family of convex descriptions which satisfy the superposition principle. Obviously, any Hilbertian description $(V(\mathbf{H}), V(\mathbf{H})_1^+)$ satisfies this principle so that the family CD_{SP} is not empty.

The given form of the superposition principle, which is merely a reformulation of the one considered in FPI, differs essentially from the variant advocated by Varadarajan (1968) and Gudder (1970). The latter would postulate, in our terms, an order isomorphism between the set of all semiexposed faces of V^+ ordered by the set inclusion and the set $\text{Ex}[0, e] \subset V^*$ with the induced order. It seems that the Varadarajan formulation of the principle of superposition of states imposes too strong, and in this connection not sufficiently justified, regularity assumptions on the description, forcing $\text{Ex}[0, e]$ to be a complete lattice (Gudder, 1970). Moreover, the Varadarajan principle holds also for classical systems with finite number of pure states (cf. Gudder, 1970), what is counterintuitive. The principle of superposition of states should express one of the distinguishing features of the quantum theories of physics. That this might be the case with our (SP) become evident below.

Let (V, S) be a classical convex description as defined in Section 3.1. Now S is a simplex, which means that for any two pure states $\alpha_1, \alpha_2 \in S$ $F(\{\alpha_1, \alpha_2\})$, the semiexposed face of S generated by $\{\alpha_1, \alpha_2\}$, is a simplex, too. Hence in $F(\{\alpha_1, \alpha_2\})$ there are no nontrivial superpositions of α_1 and α_2. We conclude this section with a simple but fundamental result.

Theorem 1. If $(S, O) \in \text{CD}_{\text{SP}}$ then S cannot be a simplex. Moreover, $\text{CD}_{\text{SP}} \neq \varnothing$ and $\text{CD}_{\text{SP}} \cap \text{CCD} = \varnothing$.

6. UNCERTAINTY

The convexity scheme followed here is probabilistic. The most typical probability distributions of the scheme are those defined by any instrument-(normalized) state pair $(\mathbf{I}, \alpha)$:

$$p(\mathbf{I}, \alpha): \mathbf{B}(\mathbf{R}) \to [0, 1], \qquad X \to p(\mathbf{I}, \alpha)(X) = e(\mathbf{I}(X)(\alpha))$$

Such distributions describe the idealized long-run results, recorded by the e detector, of the measurements $\mathbf{I}(X)$ $[X \in \mathbf{B}(\mathbf{R})]$ performed on the system in a given state α with a given experimental arrangement (instrument) $\mathbf{I}$. In the preceding two sections we have studied the two important consequences of the existence of the universal quantum of action h, namely, those expressed in the complementarity principle and the superposition principle. But they do not exhaust all the consequences of the finiteness of the h. If the phenomena involved are of quantum nature, i.e., the actions involved are comparable with the quantum of action, then it is also known that there may appear a strong correlation between the scattering of certain measurement results—a correlation which is expressed in the uncertainty relations.

Since the discovery of the uncertainty relations their interpretation has been one of the major problems in the foundation of quantum theory. These relations, which somehow arise from the inevitable interaction between measuring device and the system during the process of measurement, are said to indicate some "fundamental limitation": a limitation in the applicability of some classical concepts in quantum domain, a limitation in the definability of some concepts in quantum context, a limitation in the simultaneous measurability of certain observables, a limitation in the prepairing of states of a physical system, to mention only some of the best-known viewpoints. (For further discussion, see Busch and Lahti, 1984, 1985.)

In spite of the difficulties with the interpretation of the uncertainty relations their great importance in physics, and the foundational status of the uncertainty principle for quantum descriptions, is almost unanimously acknowledged. This is so even though the empirical support of the uncertainty principle is rather indirect. As Jammer wrote, "rarely in the history of physics has there been a principle of such universal importance with so few credentials of experimental tests" (Jammer, 1974, p. 81).

Accepting the general view on the foundational status of the uncertainty principle for quantum description we shall follow Heisenberg's intuitive idea that "... in many cases it is impossible to obtain an exact determination of the simultaneous values of two variables, but rather that there is a lower limit to the accuracy with which they can be known ... this lower limit to the accuracy with which certain variables can be known simultaneously may be postulated as a law of nature..." (Heisenberg, 1949, p. 3). Before formulating this "law of nature," the uncertainty principle, in the convexity scheme some further considerations are needed.

One of the advantages of the operational quantum mechanics over the standard quantum mechanics lies in the fact the former utilizes a far richer family of observables and of effects than the latter one. The set of measurements described by standard quantum mechanics is restricted to those giving rise to extreme effects and thus extreme-effect-valued observables, only. One of the consequences of introducing "fuzziness" into the description is explained by the following result (Davies, 1976): If $\mathbf{A}$ and $\mathbf{A}'$ are a $[0, e]$-valued and an $\mathrm{Ex}[0, e]$valued observables defining the same self-adjoint operator (in the Hilbertian approach), then $\mathrm{Var}(\mathbf{A}, \alpha) \geq \mathrm{Var}(\mathbf{A}', \alpha)$ for any state $\alpha \in S$. Thus, as it should, fuzziness increases "uncertainty" (or dispersion or variance). Physically, this increased uncertainty is caused by the possible outer disturbances on the system, which are excluded in the standard description (cf. Ingarden, 1974; Busch, 1982; Busch and Lahti, 1984). In addition to that, classical convex descriptions fulfill something like the uncertainty principle almost trivially. This is obvious, because if **A** is a simple fuzzy-valued observable defined through its range $\{0, a, e-a, e\}$ and spectrum $\{0, 1\}$ we have $\mathrm{Var}(A, \alpha) = \alpha(a)[1-\alpha(a)]$. But with the choice $\alpha(a) = \frac{1}{2} \forall \alpha \in S$, this gives $\mathrm{Var}(\mathbf{A}, \alpha) = \frac{1}{4}$ for any states $\alpha \in S$.

These results indicate that in the convexity scheme the uncertainty principle should be formulated in terms of optimal or extreme-effect-valued observables. We are finally in the position to formulate the uncertainty principle in the convexity scheme:

(UP) A convex description (S, O) satisfies the uncertainty principle if there exists at least one pair of extreme-effect-valued observables **A**, **B** and a positive number h, such that for any state $\alpha \in S$, for which the variances of **A** and **B** are well defined, the inequality $\mathrm{Var}(\mathbf{A}, \alpha) \cdot \mathrm{Var}(\mathbf{B}, \alpha) \geq h$ holds.

Let $\mathrm{CD}_{\mathrm{UP}}$ denote the family of those convex descriptions (S, O) which satisfy the uncertainty principle. Obviously, any Hilbertian description $(V(\mathbf{H})_1^+, O)$ satisfies the principle.

Let $(S, O) \in \mathrm{CCD}$ so that (V, S), and thus also (V^*, e) is a vector lattice. Identifying (V^*, e) with $(C_{\mathbf{R}}(X), 1_x)$ (Alfsen, 1971), with X denoting the (compact Hausdorff) space $\mathrm{Ex}(\beta S)$ of the extreme elements of the base βS of V^{**} we realize that the uncertainty principle cannot hold now as the extreme effects are exactly the characteristic functions of the Borel subsets of X. Note that S is weak* dense in βS. Thus we may again conclude the following:

Theorem 1. If $(S, O) \in \mathrm{CD}_{\mathrm{UP}}$ then S cannot be a simplex. Moreover, $\mathrm{CD}_{\mathrm{UP}} \neq \varnothing$ and $\mathrm{CD}_{\mathrm{UP}} \cap \mathrm{CCD} = \varnothing$.

Though there exists some disagreement on the interpretation of the uncertainty principle this principle is not so problematic from the present point of view. We can simply take the attitude that it is a law of Nature that only such states of the system can be prepared for which the product of the "uncertainties" of certain variables has a lower limit given by a positive constant h. Whether this holds or not is an empirical question.

7. INTERDEPENDENCE OF THE PRINCIPLES

We have now distinguished three important subfamilies $\mathrm{CD}_{\mathrm{CP}}$, $\mathrm{CD}_{\mathrm{UP}}$, and $\mathrm{CD}_{\mathrm{SP}}$ of the family CD of all convex descriptions through formalizing the principles of complementarity, uncertainty, and superposition within the convexity scheme. It turned out that any of these families is disjoint from the family CCD of classical convex descriptions, here characterized through the assumption UDM formalized as the requirement for the simplicial structure of the set of all normalized states of the description:

$$(\mathrm{CD}_{\mathrm{CP}} \cup \mathrm{CD}_{\mathrm{CP}} \cup \mathrm{CD}_{\mathrm{SP}}) \cap \mathrm{CCD} = \varnothing$$

Owing to the generality of the convex scheme there is, obviously, no reason to expect that the assumptions CP, UP, SP, and UDM, as formalized here, would cover all the descriptions in CD. This conjecture is confirmed, e.g., by the example due to Davies (1972). Thus

$$\mathrm{CCD} \cup \mathrm{CD}_{\mathrm{CP}} \cup \mathrm{CD}_{\mathrm{UP}} \cup \mathrm{CD}_{\mathrm{SP}} \neq \mathrm{CD}$$

We shall now briefly comment on the interdependence of the three fundamental quantum principles CP, UP, and SP. To show the mutual independence of these principles it suffices to provide examples of convex descriptions satisfying one or two of them, but not all. This method was used in Lahti (1981) to show the logical independence of the three principles within the quantum logic frame. Owing to the difference between the two approaches all the examples used therein are, however, not applicable in the present context.

To show that neither CP nor UP implies SP one may take the state space $V(\mathbf{H}_1 \oplus H_2)$ associated with the linear sum $\mathbf{H}_1 \oplus \mathbf{H}_2$ of two orthogonal Hilbert spaces $\mathbf{H}_1$ and $\mathbf{H}_2$. Such a description refers to a quantum system with superselection rules. Obviously CP and UP hold here, which is not the case with SP. The converse results that SP implies neither CP nor UP are demonstrated by the example 2 in Lahti (1981). That CP does not imply UP can be demonstrated by a sufficiently poor scheme, like, e.g., a convex scheme with a finite number of pure states. Then CP is satisfied owing to the lack of sufficiently many extreme points in $[0, e]$, whereas UP is not due to the compactness of S, which implies that any element in $[0, e]$ has a pure state in which it takes the value 0 and unbounded observables do not occur. The question whether UP implies CP or not is problematic. The example, which was used to demonstrate (Lahti, 1981) that UP does not imply CP, does not work in the present scheme as any change in the set S of states influences drastically the set $[0, e]$ of effects of the description. Thus we lack an example of a convex description satisfying UP but not CP. We leave it open whether such an example exists. However, as it becomes evident in Section 10, the results

$$\mathrm{CD}_{\mathrm{CP}} \cap \mathrm{CD}_{\mathrm{SP}}^{\perp} \neq \varnothing \quad \text{and} \quad \mathrm{CD}_{\mathrm{CP}} \cap \mathrm{CD}_{\mathrm{UP}}^{\perp} \neq \varnothing \quad \text{with} \quad \mathrm{CD}_{\alpha}^{\perp} := \mathrm{CD} \setminus \mathrm{CD}_{\alpha},$$
$$\alpha = \mathrm{SP}, \mathrm{UP}$$

showing that CP does imply neither SP nor UP, are sufficient to provide a fundamental reconstruction of the DHB quantum theories.

8. ON UNIQUE DECOMPOSABILITY OF MIXTURES

The lack of unicity in the decomposition of quantum mixtures into pure states has been emphasized as a crucial branching point between quantum and classical physics (see, e.g., Beltrametti and Cassinelli, 1981). In his convex scheme Mielnik (1974) underlined this property to the extent that he formulated it as one of the most general negative laws limiting the perception of quantum ensembles. This law, Mielnik's first *principle of impossibility*, is the following: "Having a mixed statistical ensemble of nonclassical objects one cannot determine uniquely its pure components and find out how the mixture has been prepared. Two mixtures created in two distinct ways by taking different collections of pure states may physically be indistinguishable" (Mielnik, 1974). In effect, this principle states that the set of all normalized states of the description cannot form a simplex.

In Sections 4, 5, and 6 we have considered the three principles of the quantum theory in which the very root of the quantum theory, the existence of the universal quantum of action, manifests itself most strikingly. It turned out that any of these three principles implies the above negative property, the nonsimplicial shape of the state space. Though the lack of unicity in the decomposing of mixed states into pure states is an important feature of quantum physics, it does not suffice to characterize quantum convex descriptions.

In Section 3.1 we based the notion of classical convex description on the requirement of unique decomposability of mixtures, UDM, which was formally expressed as the requirement of S being a simplex. Thus the family CD_{UDM} of those convex descriptions that satisfy UDM was identified, by definition, with the family CCD of all classical convex descriptions. This identification was justified by the fact that the convexity scheme for describing physical systems starts with the notion of state of the system and with the operation of forming mixtures of states. There are, however, at least two other characteristics of a description which are typically classical: The *Boolean* structure of the set of all properties of a system described by the theory, and the *nonprobabilistic*, or dispersion-free, character of the description. Both of these properties B and NPr, for short, can be made explicit in the present approach so that they can be used to single out the relevant subfamilies CD_B and CD_{NPr} of CD, which could form the starting point for classical convex descriptions, as well. Here we shall only show, with respect to the below given formalizations of the properties B and NPr, that

$CD_{UDM} \subset CD_B \cap CD_{NPr}$. In that we shall follow Bugajski (1981) where also the converse questions have been analyzed.

Let (V, S) be a convex description. The set $[0, e] \subset V^*$ of all formally possible effects is convex and weak* compact so that the weak* closure of the convex hull of the set $\mathrm{Ex}[0, e]$ of all extreme effects equals to $[0, e]$. The generality of the (V, S) scheme lies in the existence of a huge class of operations which give rise to nonextreme or fuzzy effects. Such operations can be interpreted to describe, e.g., some interactions of the system with its environment, or some possible outer disturbances on the system, or some inaccurate measurements on the system (cf., e.g., Ingarden, 1974; Davies, 1976; Busch and Lahti, 1984). Thus it is the set $\mathrm{Ex}[0, e]$ of extreme effects, or some of its relevant subset gained through specifying the admissible operations (cf. Section 3) that could describe the possible properties of the system (cf. Section 9). We then say the following:

(B) A convex description (S, O) is *Boolean* iff $\mathrm{Ex}[0, e]$ is Boolean.

We denote by CD_B the family of such descriptions. As already discussed in Sections 2 and 3.1, for any $(S, O) \in CCD$ $[0, e]$ is a complete distributive lattice whose Boolean sublattice $\mathrm{Ex}[0, e]$ is, with $a \mapsto a^{\perp} := e - a$ as the orthocomplementation. Hence $CCD := CD_{UDM} \subset CD_B$.

The convex scheme is probabilistic, and its basic probability assertions are of the form $e(\phi\alpha)$, $\phi \in O$, $\alpha \in S$, or simply $a(\alpha)$, with $a \in [0, e]$, $\alpha \in S$. The question then arises under which conditions the notion of probability can be eliminated. This question requires again further specification. Firstly, as conv $\mathrm{Ex}([0, e])$ is dense in $[0, e]$ it suffices to consider probability assertions of the type $a(\alpha)$ with $a \in \mathrm{Ex}([0, e])$. Secondly, only assertions $a(\alpha)$ associated with pure states should be considered. But this would require that each state α can be expressed as a convex combination (countable or otherwise) of pure states, i.e., $\mathrm{conv}(\mathrm{Ex}(S))$ should be dense in S. Such an assumption is physically well founded. However, if this is not the case we can always pass to the natural compactification βS of S in the second dual V^{**} of V, which has the required property. We then say the following:

(NPr) A convex description (S, O) is essentially *nonprobabilistic* iff $a(\alpha) \in \{0, 1\}$ for any extreme effect $a \in \mathrm{Ex}([0, e])$ and for any pure state $\alpha \in \mathrm{Ex}(\beta S)$.

If (S, O) is nonprobabilistic then any probability assertion $a(\alpha)$, $a \in [0, e]$, $\alpha \in S$, of the description can be expressed as a convex combination (countable or otherwise) of 0–1 probability statements. Customarily, one say that a state $\alpha \in S$ is *dispersion free* iff it takes only the value 0 or 1 on the set of extreme effects. Hence (S, O) is nonprobabilistic if any pure state in βS is dispersion free. The family of such descriptions will be denoted by CD_{NPr}.

Let $(S, O) \in CD_{UDM}$. Hence S, or βS, is a regular simplex which can be identified as $M_{\mathbf{R}}(\mathbf{X})_1^+$ for some compact Hausdorff space $\mathbf{X}$ (Alfsen, 1971). The pure states of S, or βS, are then identified with the Dirac measures δ_x, $x \in \mathbf{X}$, on $\mathbf{X}$ which obviously are dispersion free over the extreme effects. This then shows that $CD_{UDM} \subset CD_{NPr}$, and hence

$$CD_{UDM} \subset CD_B \cap CD_{NPr}$$

The above ideas have also been used to analyze the irreducibly probabilistic character of quantum theory (see Lahti, 1983).

9. PROJECTION POSTULATE

In Section 3 we defined the standard classical description $(M_{\mathbf{R}}(\mathbf{X})_1^+, O_{RC1})$ and the standard Hilbertian description $(V(\mathbf{H})_1^+, O_{RW1}^*)$ as rather peculiar restrictions of the corresponding operational descriptions. In the classical case the set O_{RC1} of admissible operations consists of the operations, called there conditionalizations, of the form

$$\phi_X: M_{\mathbf{R}}(\mathbf{X}) \to M_{\mathbf{R}}(\mathbf{X}), \quad \alpha \mapsto \phi_X(\alpha) \qquad \forall X \in \mathbf{B}(\mathbf{X})$$

$$\text{such that } a(\phi_X(\alpha)) = \int_X a\, d\alpha \qquad \forall a \in C_{\mathbf{R}}(\mathbf{X})$$

whereas in the Hilbertian case the set O_{RW1}^* of admissible operations contained the operations of the form

$$\phi_P: V(\mathbf{H}) \to V(\mathbf{H}), \quad \alpha \mapsto P\alpha P \qquad \forall P \in \mathbf{P}(\mathbf{H})$$

Such operations are of very special character: They are pure [mapping pure states onto pure states (or to ω) modulo normalization], ideal (satisfying the principle of minimal disturbance or least interference), and of the first kind (satisfying the repeatability hypothesis). In both cases the restrictions $O \to \{\phi_X: X \in \mathbf{B}(\mathbf{X})\}$ and $O \to \{\phi_P: P \in \mathbf{P}(\mathbf{H})\}$ lead to a natural one:one correspondences $\{\phi_X: X \in \mathbf{B}(\mathbf{X})\} \leftrightarrow \mathbf{B}(\mathbf{X}) \subset \mathrm{Ex}([0, e])$ and $\{\phi_P: P \in P(H)\} \leftrightarrow \mathbf{P}(\mathbf{H}) = \mathrm{Ex}([0, 1])$. This one:one correspondence between the distinguished class of operations, called filters, and the distinguished class of effects, called propositions, of the description formalizes the possibility of performing ideal first-kind measurements of some "properties" of the physical system concerned. This is the essence of the projection postulate; cf. FPI. To formulate this assumption in the convexity scheme the two intended classes of operations and effects should be defined.

Let (S, O) be a convex description. The set O_f of *filters* is defined as a sufficiently rich family of good operations in O. The qualities sufficiently rich and good, which grasps the qualities pure ideal and first kind, receive their exact meanings below:

Sufficiency. The set $O_f \subset O$ is *sufficiently rich* if

(S1) for any pure state α in $\mathrm{Ex}(S)$ there exists uniquely an operation ϕ_α in O_f such that $e(\phi_\alpha\beta) = e(\beta)$ implies $\beta = \alpha$ for any β in $\mathrm{Ex}(S)$;

(S2) for any operation ϕ in O_f there exists an operation ϕ' in O_f such that the resulting effects $e \circ \phi$ and $e \circ \phi'$ are orthogonal in the sense that $(e \circ \phi)^\perp = e \circ \phi'$.

Purity. An operation ϕ in O is *pure* if

(P1) $\phi\alpha \in [0, 1] \times \mathrm{Ex}(S)$ for any α in $\mathrm{Ex}(S)$.

Ideality. An operation ϕ in O is *ideal* if

(I1) $e(\phi\alpha) = e(\phi_{\alpha'}\alpha)$ for any α in $\mathrm{Ex}(S)$, with $\alpha' = e(\phi\alpha)^{-1}\ \phi\alpha$ and $\phi_{\alpha'}$ as in (S1).

First-Kindness. An operation ϕ in O is of the *first kind* if

(F1) $e(\phi\alpha) = e(\alpha)$ implies $\phi\alpha = \alpha$ for any α in $\mathrm{Ex}(S)$;

(F2) $e(\phi^2\alpha) = e(\phi\alpha)$ for any α in $\mathrm{Ex}(S)$.

The properties (S1) through (F2) which define the set O_f of filters has already been discussed in FPI, where also references to some other relevant works can be found. Owing to the differences between the present approach and the one employed in FPI some remarks are, however, called for.

As noted in FPI, the first sufficiency condition (S1) expresses the common belief that any pure state α can be produced by a particular selection or filtering process $\phi_{\alpha'}$ which under the assumptions (F1) and (F2) receives the form: $\phi_\alpha\beta \mathrel{\overline{\overline{\ \ }}} e(\phi_\alpha\beta)\alpha$ for any β in $\mathrm{Ex}(S)$. In the present scheme for any operation ϕ in O there exists an operation ϕ' in O such that the effects $e \circ \phi$ and $e \circ \phi'$ resulting from the two operations are orthogonal. With the second sufficiency condition (S2) one guarantees that whenever an effect a results from a "good" operation then also its "negation" $a^\perp$ results from a "good" operation.

The purity (P1) of an operation means simply that it takes a pure state onto a pure state with a possible loss in strength. As a pure state may be interpreted as a maximal-information state (see, e.g., Beltrametti and Casinelli, 1981), a pure operation leaves the system in a maximal-information state whenever it was in such a state.

With the so-called ideality assumptions one usually aims at minimalizing the influences on the states caused by an operation performed on the system. In addition to the purity condition (P1) and the first-kindness conditions (F1) and (F2), (I1) aims at that. It claims that an ideal ϕ maps any pure state α onto the closest to α eigenstate of ϕ, disturbing thus the system to a minimal extent.

Of the two first-kindness conditions (F1) and (F2), (F1) claims that if ϕ does not lead to a detectable effect when performed on the system in a pure state α then, provided that ϕ is "good enough," it does not alter the state of the system, either. According to (F2), a repeated application of a good operation does not lead to a new effect.

As an immediate consequence of the defining properties of filters, we note that they are not only weakly repeatable $[e(\phi^2\alpha)=e(\phi\alpha)$ for any $\alpha \in \text{Ex}(S)]$ but also repeatable $[\phi^2\alpha=\phi\alpha$ for any $\alpha \in \text{Ex}(S)]$ and even idempotent $[\phi^2=\phi]$ provided that any mixed state in S can be decomposed into its pure components in $\text{Ex}(S)$. Moreover filters satisfy the most usual ideality requirement: if a good operation ϕ_1 is performed on the system in a pure state α which is an eigenstate of a good operation ϕ_2 [i.e., $e(\phi_2\alpha)=e(\alpha)$] which commutes weakly with ϕ_1 (i.e., $\phi_1 \circ \phi_2$ and $\phi_2 \circ \phi_1$ lead to the same effect), then ϕ_1 leaves the system in a state which is still an eigenstate of ϕ_2.

The set L of *propositions* of a convex description (S, O) is defined as the set of all extreme effects a in $\text{Ex}[0, e]$ with nonempty certainly-yes-domain $a^1 := \{\alpha \in \text{Ex}(S)\colon\ a(\alpha)=1\}$ together with the null effect 0:

$$L=\{a \in \text{Ex}([0, e])\colon\ a=0 \quad \text{or} \quad a^1 \neq \varnothing\}$$

Thus propositions are exactly those extreme effects which, if they are possible [i.e., $a(\alpha)\neq 0$ for some α in $\text{Ex}(S)$], can also be actualized [i.e., there exists an α in $\text{Ex}(S)$ such that $a(\alpha)=1$]. As the "fuzziness" inherent in an (S, O) description may be interpreted as resulting from the possible outer disturbances on the system, the restriction to extreme effects guarantees that a proposition could describe a realizable property of the system.

For a given operational description (S, O) the set O_f of filters may or may not exist, and the set L of propositions may be trivial $\{0, e\}$. However, for any ϕ in O_f, $\phi \neq 0$, the resulting effect $e \circ \phi$ has a nonempty certainly-yes-domain $(e \circ \phi)^1$, and for any a in L, $a \neq 0$, one can associate through the Sasaki-projection-construction (see FPI) a filter ϕ_a such that the resulting effect $e \circ \phi_a$ equals to a. Following FPI the projection postulate is now expressed as a requirement for a natural one-to-one correspondence between the distinguished sets O_f and L of filters and propositions. However, in the present case, it appears to be reasonable to distinguish between the projection postulate and its strong variant.

(PP) An operational description (S, O) satisfies the *projection postulate* iff (1) the set O of operations admits a subset O_f of filters, and (2) there is a natural one-to-one correspondence I between the sets O_f and L with the following property: $a(\alpha)=e(I(a)\alpha)$ for every $a \in L$ and $\alpha \in \text{Ex}(S)$.

(SPP) An operational description (S, O) satisfies the *strong* form of the *projection postulate* iff it satisfies the projection postulate, and, in addition, (3) the set O_f of filters is the admissible set.

Let CD_{PP} and CD_{SPP} be the two corresponding families of convex descriptions.

Though the projection postulate PP is very restrictive from the general formal point of view, it is, however, far less restrictive than its strong variant SPP. We note that any operational classical and any operational Hilbertian description, as discussed above satisfies PP. But only the corresponding restricted standard descriptions satisfy SPP. Also from the physical point of view PP is rather plausible, which is not the case with SPP. The projection postulate guarantees the existence of the important class of operations associated with the pure, ideal, first-kind measurements, but it does not restrict the theory to deal with such measurements only, as does its strong form. We recall also that the major critique against the von Neumann–Lüders projection postulate is not so much against the existence of such measurements described by the postulate but rather against the apparently erroneous assumption that they exhaust all the physically relevant measurements.

10. THE DHB THEORY AND A MODEL

We shall now return to the problem of characterizing quantum convex descriptions. We accept the view that these descriptions are to be based more or less directly on the physical fact of the existence of the universal quantum of action symbolized by the Planck constant (cf. Lahti, 1980b). The important consequences of this fact are most properly reflected in the complementarity principle, the uncertainty principle, and in the superposition principle. This then leads us to the following definition of quantum convex descriptions:

$$QCD := CD_{CP} \cup CD_{SP} \cup CD_{UP}$$

Recall that $QCD \cap CCD = \varnothing$. Moreover, classical and quantum convex descriptions do not exhaust convex schemes in CD.

Owing to the foundational status of the three quantum principles we shall, following FPI, call the convex descriptions in

$$CD_{DHB} := CD_{CP} \cap CD_{SP} \cap CD_{UP}$$

as the *DHB quantum theories.* Clearly, any standard Hilbertian description $(V(\mathbf{H})_1^+, O_{RW1}^*)$ is such. More interesting is, however, the fact that also any operational Hilbertian description $(V(\mathbf{H})_1^+, O)$ is a DHB theory. It is clear that any $(V(\mathbf{H})_1^+, O)$ satisfies both the uncertainty principle and the superposition principle. That it satisfies also the complementarity principle can be seen from the following. If $P, Q \in \mathrm{Ex}([0, I]) = \mathbf{P}(\mathbf{H})$ are such that their meet in $\mathbf{P}(\mathbf{H})$ is zero then their meet in $[0, I]$ is zero, too. This is because the set $\{A \in [0, I]: A \leq P\}$ equals to the set $\{\lambda P: 0 \leq \lambda \leq 1\}$ for any P in $\mathrm{Ex}([0, I])$. Here $[0, I] := \{A \in L_s(\mathbf{H}): 0 \leq A \leq I\}$ and $\mathbf{P}(\mathbf{H}) := \{\mathrm{P} \in \mathrm{L_s}(\mathbf{H}): P^2 = P\}$ denote the set of all effects of the operational Hilbertian description and the standard Hilbertian description, respectively. As the observables of the standard description are $\mathrm{Ex}([0, I])$ valued, we see that complementary observables of the standard Hilbertian description, like position and momentum, are complementary also in the more general operational Hilbertian description. This shows that the operational Hilbertian descriptions are *relevant* (i.e., with actual applications; cf. Section 3.2) *nonstandard* (i.e., not standard Hilbertian) *models* for, or examples of, the DHB theories.

It has been argued elsewhere (Lahti, 1983, 1984) that the notion of complementary physical quantities presupposes the possibility of performing ideal first-kind measurements of such quantities. This is to say that whenever we have a convex description (S, O) which satisfies the complementarity principle, i.e., $(S, O) \in \mathrm{CD_{CP}}$, it should also satisfy, on physical grounds, the projection postulate, i.e., $(S, O) \in \mathrm{CD_{PP}}$. Though the projection postulate, as formalized in Section 9, is not very restrictive from the physical point of view it anyway implies very strong structural properties for a description (S, O) satisfying it. In effect, it allows one to infer that the set L of propositions of the description $(S, O) \in \mathrm{CD_{PP}}$ possesses the structure of a complete atomic lattice with the covering property (cf. Bugajska and Bugajski, 1973; FPI; Lahti, 1983). But if $(S, O) \in \mathrm{CD_{CP}} \cap \mathrm{CD_{PP}}$ L cannot be Boolean (cf. Section 4) so that the celebrated Piron–Maclaren representation theorem (Piron, 1976) allows one to identify (modulo the question of scalar field) the set L of propositions of the description (S, O) with the set $\mathbf{P}(\mathbf{H})$ of projections on a Hilbert space $\mathbf{H}$ with dimension $\dim(\mathbf{H}) \geq 3$. Reconstructing then the description (S, O) on the basis that $L \cong \mathbf{P}(\mathbf{H})$ we find that $(S, O) \cong (V(\mathbf{H})_1^+, O)$ for the Hilbert space $\mathbf{H}$. Thus these considerations strongly suggest that the important families of quantum descriptions $\mathrm{CD_{DHB}}$, $\mathrm{CD_{CP}} \cap \mathrm{CD_{PP}}$, and $\mathrm{CD_{OHQ}} := \{(S, O) \in \mathrm{CD}: S = V(\mathbf{H})_1^+, \mathbf{H}$ a Hilbert space with $\dim(\mathbf{H}) \geq 3\}$ are essentially the same:

> The Dirac–Heisenberg–Bohr quantum theories are the operational Hilbertian descriptions which can essentially be based on the fundamental complementarity principle and on the measurement theoretical idealization called the projection postulate.

Finally, we wish to emphasize that it is due to the great generality of the convexity scheme that a distinction between the projection postulate and its stronger variant became possible, and that it is only the projection postulate, in its weak form, which was applied above.

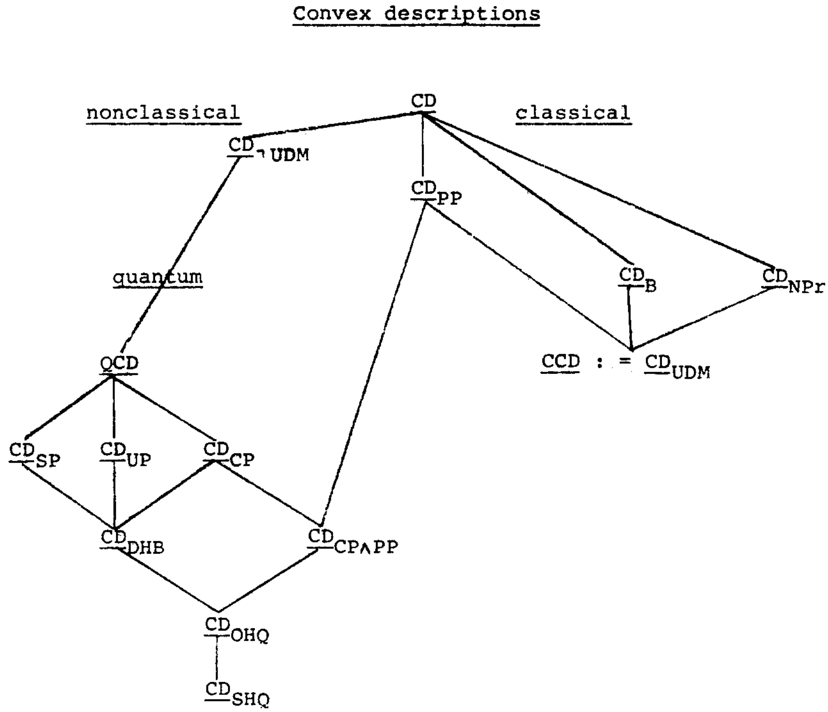

Fig. 1. A schematic representation of the interrelations of some convex descriptions.

Figure 1 above summarizes schematically the main results of this investigation in a set-theoretical form. The line connecting two families of descriptions indicate that the lower family is a subfamily, not necessarily a proper one, of the upper family. In that $CD_{\neg UDM}$ denotes the family $CD \setminus CD_{UDM}$. Other families of convex descriptions appearing in that figure have been defined throughout the text.

REFERENCES

Alfsen, E. M. (1971). *Compact Convex Sets and Boundary Integrals.* Springer-Verlag, Berlin.

Alfsen, E., and Schultz, F. (1976). Non-commutative spectral theory for affine function spaces on convex sets. *Memoirs of the American Mathematical Society 6,* No. 172.

Asimow, L., and Ellis, A. J. (1980). *Convexity Theory and Its Applications in Functional Analysis.* Academic Press, London.

Badurek, Gr., Rauch, H., Summhammer, J., Kischkol, U., and Zeilinger, A. (1983). Direct verification of the quantum spin-state superposition law. *Journal of Physics A: Math. Gen,* **16,** 1133–1139.

Beltrametti, E., and Cassinelli, G. (1981). *The Logic of Quantum Mechanics.* Addison-Wesley Publishing Company, Reading, Massachusetts.

Bugajska, K., and Bugajski, S. (1973). The projection postulate in quantum logic. *Bulletin of the Academy of Polish Science Series in Sciences Mathematics, Astronomy, and Physics*, **21**, 873-877.

Bugajski, S. (1981). Two questions on quantum mixture, in *14th Symposium on Mathematical Physics*, Torun, December 8-11.

Bugajski, S., and Lahti, P. (1980). Fundamental principles of quantum theory, *International Journal of Theoretical Physics*, **19**, 499-514.

Busch, P. (1982). *Unbestimmtheitsrelation und simultane Messung in der Quanten-Theoric*, doctoral dissertation, Köln.

Busch, P., and Lahti, P. (1984). On various joint measurements of position and momentum observables in quantum theory, *Physical Review D*, **29**, 1634-1646.

Busch, P., and Lahti, P. (1985). A note on quantum theory, complementarity, and uncertainty, *Philosophy of Science*, **52**, 64-77.

Davies, E. B. (1972). Example related to the foundations of quantum theory, *Journal of Mathematical Physics*, **13**, 39-41.

Davies, E. B. (1976). *Quantum Theory of Open Systems*. Academic Press, London.

Davies, E. B., and Lewis, J. T. (1970). An operational approach to quantum probability, *Communications on Mathematical Physics*, **17**, 239-260.

Dirac, P. A. N. (1930). *The Principle of Quantum Mechanics*. Oxford University Press, Amen House, London.

Emch, G. (1972). *Algebraic Methods in Statistical Mechanics and Quantum Field Theory*. Wiley-Interscience, New York.

Gerjuoy, E. (1973). Is the principle of superposition really necessary?, in *Contemporary Research in the Foundations and Philosophy of Quantum Theory*, C. A. Hooker, ed., D. Reidel Publ. Co., Dordrecht, Holland, pp. 114-142.

Gudder, S. (1970). A superposition principle in physics, *Journal of Mathematical Physics*, **11**, 1037-1040.

Gudder, S. (1978). *Stochastic Methods in Quantum Mechanics*. North-Holland, New York.

Gudder, S., and Marchand, J.-P. (1972). Noncommutative probability theory on von Neumann algebras, *Journal of Mathematical Physics*, **13**, 799-806.

Heisenberg, W. (1949). *The Physical Principles of the Quantum Theory*. Dover Publications, New York.

Ingarden, R. S. (1974). *Information Theory and Thermodynamics*, Part I, Chapter II, *Quantum Theory*. Reprint No 275, Torun, Poland.

Jammer, M. (1974). *The Philosophy of Quantum Mechanics*. John Wiley & Sons, New York.

Lahti, P. (1980a). Uncertainty and complementarity in axiomatic quantum mechanics, *International Journal of Theoretical Physics*, **19**, 789-842.

Lahti, P. (1980b). Characterization of quantum logics, *International Journal of Theoretical Physics*, **19**, 905-923.

Lahti, P. (1981). On the inter-relations of the three quantal principles, *Current Issues in Quantum Logic*, E. Beltrametti and B. van Fraassen, eds. Plenum Press, New York, pp. 447-454.

Lahti, P. (1985). A coherent superposition principle and the Hilbertian quantum theory. *Reports on Mathematical Physics*, **22**, 49-62.

Lahti, P. (1983). Hilbertian quantum theory as the theory of complementarity, *International Journal of Theoretical Physics*, **22**, 911-929.

Lahti, P. (1984). On the role of projection postulate in quantum theory, *Reports on Mathematical Physics*, **21**, 107-120.

Ludwig, G. (1982). *Foundations of Quantum Mechanics*, Vol I. Springer-Verlag, Berlin.

Mielnik, B. (1969). Theory of filters, *Communications on Mathematical Physics*, **15**, 1-46.

Mielnik, B. (1974). Generalized Quantum Mechanics. *Communications on Mathematical Physics*, **37**, 221-256.

Piron, C. (1976). *Foundations of Quantum Physics.* W. A. Benjamin, Reading, Massachusetts.

Primas, H. (1981). *Chemistry, Quantum Mechanics and Reductionism.* Springer-Verlag, Berlin.

Rüttimann, F. (1981). Expectation functionals of observables and counters, in *14th Symposium on Mathematical Physics, Torun,* December 8–11.

Summhammer, J., Badurek, G., and Rauch, H. (1983). Direct observation of fermion spin superposition by neutron interferometry, *Physical Review A,* **27**.

Schaefer, H. H. (1971). *Topological Vector Spaces.* Springer-Verlag, Berlin.

Schaefer, H. H. (1974). *Orderings of Vector Spaces. Lecture Notes in Physics,* No. 29, 4–10. Springer-Verlag, Berlin.

Varadaraja, V. S. (1968). *Geometry of Quantum Theory,* I. Van Nostrand Reinhold Company, New York.