Commun. math. Phys. 29, 249—264 (1973)

Convex Structures and Operational Quantum Mechanics

S. Gudder

Department of Mathematics, University of Denver, Denver, Colorado, USA

Received May 29, 1972; in revised form September 15, 1972

Abstract. A general mathematical framework called a convex structure is introduced. This framework generalizes the usual concept of a convex set in a real linear space. A metric is constructed on a convex structure and it is shown that mappings which preserve the structure are contractions. Convex structures which are isomomorphic to convex sets are characterized and for such convex structures it is shown that the metric is induced by a norm and that structure preserving mappings can be extended to bounded linear operators.

Convex structures are shown to give an axiomatization of the states of a physical system and the metric is physically motivated. We demonstrate how convex structures give a generalizing and unifying formalism for convex set and operational methods in axiomatic quantum mechanics.

1. Introduction

Until recently the main axiomatic frameworks for quantum mechanics have been the C^*-algebra approach [6, 14] and the "quantum logic" approach [4, 7, 15]. Recently a new and more general method of attack has been discussed, which might be called the "convex set" or "operational" approach. This method has been emphasized by Ludwig *et al.* [13] Gunson [5], Mielnik [9, 10], Davies and Lewis [2] and others [3, 11]. In this framework a basic role is played by the convex set of normalized states S_1 together with the geometric properties of the boundary of S_1. Although the particular terminology, physical interpretations, and certain details of these investigations may differ, they all use convex set methods.

In this paper we introduce a framework which is even more general but which we feel forms a unification of the convex set or operational approaches and at the same time provides a new and perhaps useful mathematical tool for further investigations. In this framework the only primitive axiomatic elements are the normalized states S_1 of a physical system and the only operation postulated on S_1 is that of forming mixtures. In this way the convex structure of S_1 is isolated and is, in fact, the only structure of S_1.

2. Convex Structures

Let $S_1 = \{p, q, r, \ldots\}$ be the set of normalized states for some physical system. For generality we do not specify any particular form for these states but take them to be undefined, primitive elements. In different axiomatic models for quantum mechanics the states take various forms. In the conventional model, for example, the states are positive trace-class operators with trace one [7, 15]; in the quantum logic approach, the states are probability measures on an orthomodular lattice [4, 7, 15]; in the algebraic approach, the states are positive, normal, linear functionals on a C^*-algebra [6, 14]; in the operational approach, the states are positive elements of an ordered Banach space [2, 10]. Although in these models the normalized states form a convex subset of a vector space we can formulate an axiomatic framework without any linear structure whatsoever. Thus our theory not only generalizes the usual models but leaves open the possibility of non-linear structures for quantum mechanics [9, 10].

We now assume there is a notion of mixtures of states. Thus if $\lambda_1, \lambda_2 \geqq 0$, $\lambda_1 + \lambda_2 = 1$, $p, q \in S_1$ we assume there is a state $T_2(\lambda_1, p; \lambda_2, q)$ which may be interpreted as a mixture with λ_1 parts p and λ_2 parts q. Depending upon ones interpretation $T_2(\lambda_1, p;\ \lambda_2, q)$ may be looked upon as a state in which the system is in state p with probability λ_1 or state q with probability λ_2; or $T_2(\lambda_1, p; \lambda_2. q)$ may describe a beam of non-interacting particles of two types in proportion λ_1 to λ_2; or some other interpretation. In any case it is clear that T_2 is a map $T_2\colon \{(\lambda_1, p; \lambda_2, q) : \lambda_1, \lambda_2 \geqq 0,\ \lambda_1 + \lambda_2 = 1,\ p, q \in S_1\} \to S_1$ satisfying:

(1) $T_2(\lambda_1, p; \lambda_2, q) = T_2(\lambda_2, q; \lambda_1, p)$,

(2) $T_2(1, p; 0, q) = p$.

Furthermore, we would like to form mixtures of three normalized states so we assume the existence of a map

$$T_3\colon \left\{(\lambda_1, p_1; \lambda_2, p_2; \lambda_3, p_3) : \lambda_i \geqq 0, \sum_1^3 \lambda_i = 1, p_1, p_2, p_3 \in S_1\right\} \to S_1$$

satisfying:

$$(3) \qquad T_3(\lambda_i, p_i; \lambda_j, p_j; \lambda_k, p_k) = T_3(\lambda_1, p_1; \lambda_2, p_2; \lambda_3, p_3)$$

whenever (i, j, k) is a permutation of (1, 2, 3),

(4) if $\lambda_1 \neq 1$ then

$$T_3(\lambda_1, p_1; \lambda_2, p_2; \lambda_3, p_3) = T_2\big(\lambda_1, p_1; 1 - \lambda_1, T_2(\lambda_2(1-\lambda_1)^{-1}, p_2; \lambda_3(1-\lambda_1)^{-1}, p_3)\big).$$

Condition (3) is clear, while condition (4) gives the connection between T_3 and T_2 and follows from the fact that the right-hand side is a mixture of λ_1 parts p_1 and $1-\lambda_1$ parts a state with $\lambda_2(1-\lambda_1)^{-1}$ parts p_2 and $\lambda_3(1-\lambda_1)^{-1}$ parts p_3 so the resulting state has λ_1 parts p_1, λ_2 parts p_2 and λ_3 parts p_3.

A set S_1 with two maps T_2, T_3 satisfying conditions (1)–(4) is called a *convex structure*. Let us define the map $T\colon [0,1]\times S_1\times S_1\to S_1$ by $T(\lambda,p,q)\equiv\langle\lambda,p,q\rangle=T_2(1-\lambda,p;\lambda,q)$. Of course, $\langle\lambda,p,q\rangle$ represents a mixture with $1-\lambda$ parts p and λ parts q. The reason for defining $\langle\lambda,p,q\rangle$ as $T_2(1-\lambda,p;\lambda,q)$ instead of $T_2(\lambda,p;1-\lambda,q)$ is merely for the convenience of making our later formulas simpler and also to agree with the standard notation in convex sets.

Theorem 2.1. *If* (S_1,T_2,T_3) *is a convex structure then* $T(\lambda,p,q)\equiv\langle\lambda,p,q\rangle=T_2(1-\lambda,p;\lambda,q)$ *satisfies:*

(a) $\langle\lambda,p,q\rangle=\langle 1-\lambda,q,p\rangle$ *(commutativity)*

(b) $\langle 0,p,q\rangle=p$ *(endpoint condition)*

(c) $\langle\lambda,p,\langle\mu,q,r\rangle\rangle=\langle\lambda\mu,\langle\lambda(1-\mu)(1-\lambda\mu)^{-1},p,q\rangle,r\rangle\quad(\lambda\mu\neq 1)$ *(associativity).*

Conversely, let $S_1=\{p,q,r,\ldots\}$ *be a set and let* $T\colon[0,1]\times S_1\times S_1\to S_1$, $T(\lambda,p,q)\equiv\langle\lambda,p,q\rangle$ *satisfy* (a), (b) *and* (c). *If we define*

$$T_2\colon\{(\lambda_1,p;\lambda_2,q)\colon\lambda_1,\lambda_2\geqq 0,\lambda_1+\lambda_2=1,p,q\in S_1\}\to S_1$$

and

$$T_3\colon\left\{(\lambda_1,p_1;\lambda_2,p_2;\lambda_3,p_3)\colon\lambda_i\geqq 0,\sum_1^3\lambda_i=1,p_1,p_2,p_3\in S_1\right\}\to S_1$$

by $T_2(\lambda_1,p;\lambda_2,q)=\langle 1-\lambda_1,p,q\rangle$ *and*

$$T_3(\lambda_1,p;\lambda_2,q;\lambda_3,r)=\langle 1-\lambda_1,p;\langle\lambda_3(1-\lambda_1)^{-1},q,r\rangle\rangle$$

if $\lambda_1\neq 1$, $=p$ *if* $\lambda_1=1$, *then* (S_1,T_2,T_3) *is a convex structure.*

Proof. (a) $\langle\lambda,p,q\rangle=T_2(1-\lambda,p;\lambda,q)=T_2(\lambda,q;1-\lambda,p)=\langle 1-\lambda,q,p\rangle$.

(b) $\langle 0,p,q\rangle=T_2(1,p;0,q)=p$.

(c)
$$\begin{aligned}\langle\lambda,p,\langle\mu,q,r\rangle\rangle&=T_2(1-\lambda,p;\lambda,\langle\mu,q,r\rangle)\\&=T_2(1-\lambda,p;\lambda,T_2(1-\mu,q;\mu,r))\\&=T_3(1-\lambda,p;(1-\mu)\lambda,q;\mu\lambda,r)\\&=T_3(\mu\lambda,r;1-\lambda,p;(1-\mu)\lambda,q)\\&=T_2(\mu\lambda,r;1-\mu\lambda,T_2((1-\lambda)(1-\mu\lambda)^{-1},p;\\&\quad\cdot(1-\mu)\lambda(1-\mu\lambda)^{-1},q))\\&=\langle 1-\mu\lambda,r,\langle\lambda(1-\mu)(1-\mu\lambda)^{-1},p,q\rangle\rangle\\&=\langle\mu\lambda,\langle\lambda(1-\mu)(1-\mu\lambda)^{-1},p,q\rangle,r\rangle\,.\end{aligned}$$

(1) $T_2(\lambda_1, p_1; \lambda_2, q) = \langle 1-\lambda_1, p, q\rangle = \langle \lambda_1, q, p\rangle = \langle 1-\lambda_2, q, p\rangle$
$= T_2(\lambda_2, q; \lambda_1, p)$.

(2) $T_2(1, p, 0, q) = \langle 0, p, q\rangle = p$.

(3) $T_3(\lambda_1, p_1; \lambda_2, p_2; \lambda_3, p_3) = \langle 1-\lambda_1, p_1, \langle \lambda_3(1-\lambda_1)^{-1}, p_2, p_3\rangle\rangle$
$= \langle \lambda_3, \langle \lambda_2(1-\lambda_3)^{-1}, p_1, p_2\rangle, p_3\rangle$
$= \langle 1-\lambda_3, p_3, \langle \lambda_2(1-\lambda_3)^{-1}, p_1, p_2\rangle\rangle$
$= T_3(\lambda_3, p_3; \lambda_1, p_1; \lambda_2, p_2)$

and the other permutations are similar.

(4) $\lambda_1 \neq 1$ then,

$$\begin{aligned} T_3(\lambda_1, p_1; \lambda_2, p_2; \lambda_3, p_3) &= \langle 1-\lambda_1, p_1, \langle \lambda_3(1-\lambda_1)^{-1}, p_2, p_3\rangle\rangle \\ &= T_2(\lambda_1, p_1; (1-\lambda_1), \langle \lambda_3(1-\lambda_1)^{-1}, p_2, p_3\rangle) \\ &= T_2(\lambda_1, p_1; (1-\lambda_1), T_2(\lambda_2(1-\lambda_1)^{-1}, p_2; \\ &\qquad \lambda_3(1-\lambda_1)^{-1}, p_3)). \end{aligned}$$

Thus an equivalent definition of a convex structure is a pair (S_1, T) where S_1 is a set and $T: [0,1] \times S_1 \times S_1 \to S_1$ satisfies (a), (b) and (c). Although our first definition is more physically motivated, the second is more convenient mathematically so we will use the second in the sequel. An example of a convex structure is a convex set S_0 in a real linear space where $T: [0,1] \times S_0 \times S_0 \to S_0$ is $T(\lambda, p, q) = (1-\lambda)p + \lambda q$. However one can construct convex structures that are not of this form. For instance, there are simple examples [1] of convex structures S_1 with finitely many elements and since a non-empty, non-singleton convex set S_0 must have infinitely many elements, S_1 cannot be isomorphic to an S_0. Our next result will characterize those convex structures that are isomorphic to a convex set.

A *convex prestructure* is a set $S = \{p, q, r, \ldots\}$ together with a map $T: [0,1] \times S \times S \to S$ denoted by $T(\lambda, p, q) = \langle \lambda, p, q\rangle$. Of course, any set is a convex prestructure since we have placed no requirements on T. If S_1, S_2 are convex prestructures, a map $A: S_1 \to S_2$ is *affine* if $A < \lambda, p, q\rangle_1 = \langle \lambda, Ap, Aq\rangle_2$ for every $\lambda \in [0,1]$, $p, q \in S_1$. We say S_1 and S_2 are *isomorphic* if there is an affine bijection from S_1 to S_2. If S_0 is a convex subset of a real vector space we always assume S_0 is equipped with its usual convex structure. An *affine functional* is an affine map f from a convex prestructure S to the real line R; that is, $f(\langle \lambda, p, q\rangle) = (1-\lambda) f(p) + \lambda f(q)$ for all $\lambda \in [0,1]$, $p, q \in S$. We let S^* denote the set of all affine functionals on S and say that S^* is *total* if for $p \neq q \in S$ there is an $f \in S^*$ such that $f(p) \neq f(q)$.

Theorem 2.2. *A convex prestructure S is isomorphic to a convex set if and only if S^* is total.*

Proof. Suppose S_0 is a convex set and $F: S \to S_0$ is an isomorphism. If S_0 is a convex subset of the vector space V, it is well-known that the algebraic dual V^* is total over V. Restricting the elements of V^* to S_0 we get a total set of affine functionals for S_0. Now if $f \in V^*$ then $f \circ F \in S^*$ so S^* is total. Conversely, suppose S^* is total. For $p \in S$ define $J(p): S^* \to R$ by $J(p)f = f(p)$. Clearly S^* is a vector space under pointwise operations and $J(p) \in S^{**}$ so that $J(S) \subseteq S^{**}$. Now $J(S)$ is a convex set since for $J(p), J(q) \in J(S)$ and $\lambda \in [0, 1]$ we have for $f \in S^*$,

$$[(1-\lambda)J(p)+\lambda J(q)]f = (1-\lambda)f(p)+\lambda f(q) = f(\langle\lambda,p,q\rangle) = J(\langle\lambda,p,q\rangle)f$$

so $(1-\lambda)J(p)+\lambda J(q) = J(\langle\lambda,p,q\rangle) \in J(S)$. Now $J: S \to S^{**}$ is injective if and only if S^* is total. Indeed, if S^* is total and $p \neq q \in S$ then there is an $f \in S^*$ such that $f(p) \neq f(q)$ so $J(p) \neq J(q)$ and conversely, if J is injective and $p \neq q \in S$ then $J(p) \neq J(q)$ so there is an $f \in S^*$ such that $f(p) = J(p)f \neq J(q)f = f(q)$. It follows that $J: S \to J(S)$ is an isomorphism.

We would now like to define a distance between normalized states p, q. The closeness of p to q can be measured by comparing mixtures $\langle\lambda, p, p_1\rangle, \langle\lambda, q, q_1\rangle$ of p and q with other normalizes states. If p and q are very close one would expect to find a mixture containing mostly p equal to a mixture containing mostly q; that is, $\langle\lambda, p, p_1\rangle = \langle\lambda, q, q_1\rangle$ in which λ is very small. Conversely, if $\langle\lambda, p, p_1\rangle = \langle\lambda, q, q_1\rangle$ and λ is small one expects that p and q are close. Thus the parameters λ such that $\langle\lambda, p, p_1\rangle = \langle\lambda, q, q_1\rangle$ give a measure of the closeness of p to q. We thus define

$$\sigma(p,q) = \inf\{0 \leqq \lambda \leqq 1 : \langle\lambda, p, p_1\rangle = \langle\lambda, q, q_1\rangle, p_1, q_1 \in S_1\},$$

where S_1 is a convex structure. Notice that since $\langle 2^{-1}, p, q\rangle = \langle 2^{-1}, q, p\rangle$ we have $0 \leqq \sigma(p,q) \leqq \frac{1}{2}$. It turns out to be more useful to make a change of scale and define the distance between p and q to be $\varrho(p,q) = \sigma(p,q) \cdot [1-\sigma(p,q)]^{-1}$ so that $0 \leqq \varrho(p,q) \leqq 1$. I am indebted to W. Cornette for help with the following proof.

Theorem 2.3. *On a convex structure S_1, σ and ϱ are semimetrics.*

Proof. It is clear that σ and ϱ are non-negative and symmetric. Since $\langle 0, p, q_1\rangle = \langle 0, p, q_1\rangle$ for every $p_1, q_1 \in S_1$ we have $\sigma(p,p) = \varrho(p,p) = 0$. We now prove the triangle inequality. Clearly, $\sigma(p,q) \leqq \sigma(p,s) + \sigma(s,q)$ if $p = s$ or $q = s$ so we can exclude these cases. Suppose

$$\lambda_1 \in \{0 < \lambda < 1 : \langle\lambda, p, p_1\rangle = \langle\lambda, s, s_1\rangle, p_1, s_1 \in S_1\}$$

and $\lambda_2 \in \{0 < \lambda < 1 : \langle\lambda, s, s_2\rangle = \langle\lambda, q, q_1\rangle, s_2, q_1 \in S_1\}$. Letting $\lambda_3 = \lambda_1 + \lambda_2 - 2\lambda_1\lambda_2$, $p_2 = \langle\lambda_2(1-\lambda_1)\lambda_3^{-1}, p_1, s_2\rangle$, $q_2 = \langle\lambda_2(1-\lambda_1)\lambda_3^{-1}, s_1 q_1\rangle$ and

$\lambda_0=\lambda_3(1-\lambda_1\lambda_2)^{-1}$ we have:

$$\begin{aligned}\langle\lambda_0,p,p_2\rangle &= \langle\lambda_3(1-\lambda_1\lambda_2)^{-1},p,\langle\lambda_2(1-\lambda_1)\lambda_3^{-1},p_1,s_2\rangle\rangle\\ &= \langle\lambda_2(1-\lambda_1)(1-\lambda_1\lambda_2)^{-1},\langle\lambda_1,p,p_1\rangle,s_2\rangle\\ &= \langle\lambda_2(1-\lambda_1)(1-\lambda_1\lambda_2)^{-1},\langle\lambda_1,s,s_1\rangle,s_2\rangle\\ &= \langle(1-\lambda_2)(1-\lambda_1\lambda_2)^{-1},s_2,\langle\lambda_1,s,s_1\rangle\rangle\\ &= \langle\lambda_1(1-\lambda_2)(1-\lambda_1\lambda_2)^{-1},\langle 1-\lambda_2,s_2,s\rangle,s_1\rangle\\ &= \langle(1-\lambda_1)(1-\lambda_1\lambda_2)^{-1},s_1,\langle\lambda_2,q,q_1\rangle\rangle\\ &= \langle(1-\lambda_1)(1-\lambda_2)(1-\lambda_1\lambda_2)^{-1},\langle\lambda_2(1-\lambda_1)\lambda_3^{-1},s_1,q_1\rangle,q\rangle\\ &= \langle\lambda_3(1-\lambda_1\lambda_2)^{-1},q,q_2\rangle=\langle\lambda_0,q,q_2\rangle\end{aligned}$$

so $\lambda_0\in\{0<\lambda<1\colon\langle\lambda,p,p_2\rangle=\langle\lambda,q,q_2\rangle,p_2,q_2\in S_1\}$. Now since $\lambda_0(1-\lambda_0)^{-1}=\lambda_1(1-\lambda_1)^{-1}+\lambda_2(1-\lambda_2)^{-1}$ we have

$$\begin{aligned}\varrho(p,q)=\sigma(p,q)\,[1-\sigma(p,q)]^{-1}&\leqq\sigma(p,s)\,[1-\sigma(p,s)]^{-1}\\ &+\sigma(s,q)\,[(1-\sigma(s,q)]^{-1}=\varrho(p,s)+\varrho(s,q)\,.\end{aligned}$$

The triangle inequality for σ follows in a similar way using the fact that $\lambda_0\leqq\lambda_1+\lambda_2$.

We call ϱ the *intrinsic semimetric* for S_1. As an example, let $p<q\in R$ and let us compute $\varrho(p,q)$ relative to some bounded convex subset S_0 of R. Since the only convex subsets of R are intervals we may assume that S_0 is a closed interval $[p_0,q_0]$, $p_0<q_0$, $p,q\in[p_0,q_0]$. It is easy to see that relative to S_0, $\sigma(p,q)=(p-q)\,[(p-q)+(p_0-q_0)]^{-1}$ and $\varrho(p,q)=(p-q)\,(p_0-q_0)^{-1}$. Thus ϱ is a metric on $[p_0,q_0]$ and is equivalent to the Euclidean metric. Notice however, that if ϱ is computed relative to an unbounded convex set, then all distances are zero and ϱ is the trivial semimetric $\varrho\equiv 0$.

Lemma 2.4. *A necessary and sufficient condition for σ,ϱ to be metrics is that whenever there are sequences $\lambda_i\in[0,1]$, $p_i,q_i\in S_1$ which satisfy $\lim_{i\to\infty}\lambda_i=0$, $\langle\lambda_i,p,p_i\rangle=\langle\lambda_i,q,q_i\rangle$ then $p=q$.*

Proof. Clearly ϱ is a metric if and only if σ is. Now if σ is a metric, since $\sigma(p,q)\leqq\lambda_i$ for all i we have $p=q$. Conversely, if $\sigma(p,q)=0$ then $V=\{0\leqq\lambda\leqq 1\colon\langle\lambda,p,p_1\rangle=\langle\lambda,q,q_1\rangle,p_1,q_1\in S_1\}$ either contains 0 or has 0 as a limit point. In the former case $p=\langle 0,p,p_1\rangle=\langle 0,q,q_1\rangle=q$. In the latter case there exist $\lambda_i\in V$ with $\lim\lambda_i=0$ so again $p=q$.

Corollary 2.5. *Let S_0 be a convex set in a real vector space X. If there is a topology on X that makes X a Hausdorff topological vector space in which S_0 is bounded, then ϱ is a metric.*

Proof. Suppose there are sequences $\lambda_i \in [0, 1]$, $\lim \lambda_i = 0$, $p_i, q_i \in S_0$ such that $(1-\lambda_i)\,p + \lambda_i p_i = (1-\lambda_i)\,q + \lambda_i q_i$. Then $p - q = \lambda_i(p-q) + \lambda_i(q_i - p_i)$. Let Λ be any neighborhood of 0. Then there is a neighborhood W of 0 such that $W + W + W \subseteq \Lambda$. Now for i sufficiently large $\lambda_i(p-q) \in W$. Since S_0 is bounded there is a $\mu > 0$ such that $\lambda S_0 \subseteq W$ for $|\lambda| \leqq \mu$. Thus for i sufficiently large $\lambda_i q_i - \lambda_i p_i \in W + W$. Hence for sufficiently large i, $p - q \in W + W + W \subseteq \Lambda$. Since X is Hausdorff, $p - q = 0$ and $p = q$.

The converse of the above corollary holds for finite dimensional spaces X. Indeed if S_0 is unbounded in X then one can show there is an infinite ray in S_0 and hence any two points on this ray will have distance zero. The converse of the corollary need not hold in infinite dimensional spaces.

If S_1, S_2 are convex structures we denote the affine maps $A: S_1 \to S_2$ by $Af(S_1, S_2)$. We use the notation $Af(S_1)$ for $Af(S_1, S_1)$ and the group of bijections in $Af(S_1)$ is denoted $\mathrm{Aut}(S_1)$.

Lemma 2.6. (1) *If* $A \in Af(S_1, S_2)$ *then* A *is a contraction (i.e.* $\varrho_2(Ap, Aq) \leqq \varrho_1(p, q)$*).* (2) *If* $A \in Af(S_1, S_2)$ *is bijective then* A *is an isometry.*

Proof. (1)
$$\begin{aligned}\sigma_2(Ap, Aq) &= \inf\{0 \leqq \lambda \leqq 1 : \langle \lambda, Ap, p_1\rangle_2 \\ &\qquad = \langle \lambda, Aq, q_1\rangle_2, p_1, q_1 \in S_2\} \\ &\leqq \inf\{0 \leqq \lambda \leqq 1 : A\langle \lambda, p, q_1\rangle_1 \\ &\qquad = A\langle \lambda, q, q_1\rangle_1, p_1, q_1 \in S_1\} \\ &\leqq \inf\{0 \leqq \lambda \leqq 1 : \langle \lambda, p, p_1\rangle_1 \\ &\qquad = \langle \lambda, q, q_1\rangle_1, p_1, q_1 \in S_1\} \\ &= \sigma_1(p, q)\,.\end{aligned}$$

Since $x(1-x)^{-1}$ is monotone increasing function we have $\varrho_2(Ap, Aq) \leqq \varrho_1(p, q)$. (2) follows from (1).

It follows, of course, that any $A \in Af(S_1, S_2)$ is continuous relative to the intrinsic metrics. As an application, if $f \in S_1^*$ and $0 \leqq f \leqq 1$ then $|f(p) - f(q)| \leqq \varrho(p, q)$.

Let S_1 be a convex structure. For applications it is important to consider the set $S = S_1^+ = \{(\alpha, p) : p \in S_1, \alpha \geqq 0\}$. We define $(\alpha, p) = (\beta, q)$ if $\alpha = \beta \neq 0$ and $p = q$, and $(0, p) = (0, q) \equiv 0$ for all $p, q \in S_1$. For convenience we write αp instead of (α, p). If S_1 corresponds to the normalized states of a physical system, then S corresponds to the states. For $\alpha p \in S$ and $\beta \geqq 0$ we define $\beta(\alpha p) = (\beta\alpha)\,p$, and clearly $\beta(\alpha p) = \alpha(\beta p)$. For $\alpha p, \beta q \in S$, $\lambda \in [0, 1]$, we define $\langle \lambda, \alpha p, \beta q\rangle = 0$ if $(1-\lambda)\,\alpha + \lambda\beta = 0$ and otherwise $\langle \lambda, \alpha p, \beta q\rangle = [(1-\lambda)\,\alpha + \lambda\beta]\,\langle \lambda\beta[(1-\lambda)\,\alpha + \lambda\beta]^{-1}, p, q\rangle$.

Lemma 2.7. *$(S, \langle\cdot,\cdot,\cdot\rangle)$ is a convex structure.*

Proof. (a) Let $\gamma=(1-\lambda)\alpha+\lambda\beta$. Then for $\gamma \neq 0$

$$\langle\lambda,\alpha p,\beta q\rangle=\gamma\langle\lambda\beta\gamma^{-1},p,q\rangle=\gamma\langle 1-\lambda\beta\gamma^{-1},q,p\rangle=\gamma\langle(1-\lambda)\alpha\gamma^{-1},q,p\rangle$$
$$=\langle 1-\lambda,\beta q,\alpha p\rangle\,.$$

If $\gamma=0$ the equality clearly holds. (b) $\langle 0,\alpha p,\beta q\rangle=\alpha\langle 0,p,q\rangle=\alpha p$. (c) Letting $\xi=(1-\mu)\beta+\mu\delta$, $\tau=(1-\lambda)\alpha+\lambda\xi$, and $\sigma=\alpha(1-\lambda)(1-\lambda\mu)^{-1}+\beta\lambda(1-\mu)(1-\lambda\mu)^{-1}$ we have

$$\begin{aligned}\langle\lambda,\alpha p,\langle\mu,\beta q,\delta r\rangle\rangle &= \langle\lambda,\alpha p,\xi\langle\mu\delta\xi^{-1},q,r\rangle\rangle\\ &=\tau\langle\lambda\xi\tau^{-1},p,\langle\mu\delta\xi^{-1},q,r\rangle\rangle\\ &=\tau\langle\lambda\mu\delta\tau^{-1},\langle\lambda\xi\tau^{-1}(1-\mu\delta\xi^{-1})(1-\lambda\mu\delta\tau^{-1})^{-1},\\ &\qquad\qquad p,q\rangle,r\rangle\\ &=\tau\langle\lambda\mu\delta\tau^{-1},\langle\lambda(1-\mu)\beta(1-\lambda\mu)^{-1}\sigma^{-1},p,q\rangle,r\rangle\\ &=\langle\lambda\mu,\sigma\langle\lambda(1-\mu)\beta(1-\lambda\mu)^{-1}\sigma^{-1},p,q\rangle,\delta r\rangle\\ &=\langle\lambda\mu,\langle\lambda(1-\mu)(1-\lambda\mu)^{-1},\alpha p,\beta q\rangle,\delta r\rangle\,.\end{aligned}$$

Thus all our results for convex structures can be applied to S. In particular S has an intrinsic metric

Let S_3 be a convex structure. We say S_2 is a *convex substructure* of S_3 if $S_2\subseteq S_3$ and $\langle\lambda,p,q\rangle\in S_2$ for all $\lambda\in[0,1]$, $p,q\in S_2$. If we identify elements of S_1 with elements of S_1^+ of the form $1\cdot p$, $p\in S_1$ then S_1 is a convex substructure of S_1^+. We now show how to extend elements of S_1^* to elements of S_1^{+*}.

Lemma 2.8. (a) *If $g\in S_1^{+*}$ then $g(\alpha p)=\alpha g(p)+(1-\alpha)g(0)$ for all $\alpha\geqq 0$, $p\in S_1$.* (b) *If $f\in S_1^*$ and $c\in R$ there exists a unique extension $g\in S_1^{+*}$ of f satisfying $g(0)=c$; in fact, $g(\alpha p)=\alpha f(p)+(1-\alpha)c$.*

Proof. (a) For any $\lambda\in[0,1]$, $\beta\geqq 0$, $p\in S_1$ we have $\langle\lambda,0p,\beta p\rangle=\lambda\beta\langle 1,p,p\rangle=\lambda\beta p$ and hence $g(\lambda\beta p)=(1-\lambda)g(0)+\lambda g(\beta p)$. If $0\leqq\alpha\leqq 1$ let $\beta=1$ and $\lambda=\alpha$ to obtain the result. If $\alpha>1$ let $\beta=\alpha$ and $\lambda=\alpha^{-1}$ to obtain $g(p)=(1-\alpha^{-1})g(0)+\alpha^{-1}g(\alpha p)$ which again gives the result. (b) Defining $g(\alpha p)=\alpha f(p)+(1-\alpha)c$ we need only show $g\in S_1^{+*}$ since clearly g is an extension of f, $g(0)=c$ and uniqueness follows from (a). Now letting $\gamma=(1-\lambda)\alpha+\lambda\beta$,

$$\begin{aligned}g(\langle\lambda,\alpha p,\beta q\rangle)&=g(\gamma<\lambda\beta\gamma^{-1},p,q\rangle)\\ &=\gamma f(\langle\lambda\beta\gamma^{-1},p,q\rangle)+(1-\gamma)c\\ &=\gamma[(1-\lambda\beta\gamma^{-1})f(p)+\lambda\beta\gamma^{-1}f(g)]+(1-\gamma)c\\ &=(1-\lambda)\alpha f(p)+\lambda\beta f(g)+(1-\gamma)c\\ &=(1-\lambda)[\alpha f(p)+(1-\alpha)c]+\lambda[\beta f(g)+(1-\beta)c]\\ &=(1-\lambda)g(\alpha p)+\lambda g(\beta q)\,.\end{aligned}$$

This completes the proof.

If $f_1 \in S_1^*$ is defined by $f_1(p) = 1$ for all $p \in S_1$ its extension $\tau \in S_1^{+*}$ given by $\tau(\alpha p) = \alpha$ will be quite important in the applications.

3. The Induced Seminorm

We now investigate the intrinsic semimetric in the special case of convex sets. In this section S_0 will denote a convex set in a real vector space V and ϱ the intrinsic semimetric on S_0. We first recall some definitions. S_0 is *absorbing* if for any $x \in V$ there is a $\delta(x) > 0$ such that $\lambda x \in S_0$ for any λ with $|\lambda| \leqq \delta(x)$. S_0 is *balanced* if $\lambda x \in S_0$ for any $x \in S_0$, $|\lambda| \leqq 1$. S_0 is *radial* if $x \in S_0$ implies $\lambda x \in S_0$ for any $0 \leqq \lambda \leqq 1$. S_0 is *normalized* if $x \in S_0$ implies $\alpha x \notin S_0$ for any $\alpha \neq 1$. If $K = \{\alpha S_0 : \alpha \geqq 0\}$ then K is *wedge* in V, i.e., $K + K \subseteq K$ and $\alpha K \subseteq K$ for any $\alpha \geqq 0$. We let $X = K - K \subseteq V$ be the subspace of V generated by K.

Let $D = \{cp - dq : 0 \leqq c, d \leqq 1; p, q \in S_0\}$. Then D is a convex, balanced, absorbing subset of X containing 0. Notice $D = S_0 - S_0$ if S_0 is radial. For $x \in X$ let $|x| = \inf\{\lambda > 0 : x \in \lambda D\}$. Then $|\cdot|$ is the Minkowski functional [12] for D in X. It is well-known that $|\cdot|$ is a seminorm and we call it the seminorm *induced* by the intrinsic semimetric ϱ. The reason for this terminology will become apparent from Theorem 3.2. We first give another way of computing $|\cdot|$ which is frequently more convenient.

Lemma 3.1. $|x| = \inf\{\max(c, d) : x = cp - dq; c, d \geqq 0; p, q \in S_0\}$.

Proof. By definition,

$$|x| = \inf\{\lambda > 0 : x = \lambda cp - \lambda dq; 0 \leqq c, d \leqq 1; p, q \in S_0\}.$$

Now if $x = \lambda cp - \lambda dq$, $0 \leqq c, d \leqq 1$, then $x = (\lambda c) p - (\lambda d) q$ where $\max(\lambda c, \lambda d) \leqq \lambda$ so $\inf\{\max(c, d) : x = cp - dq;\ c, d \geqq 0;\ p, q \in S_0\} \leqq |x|$. Conversely, if $x = cp - dq$, $c, d > 0$, $p, q \in S_0$ then

$$x = \max(c, d)\,[c \max(c, d)^{-1}]\,p - \max(c, d)\,[d \max(c, d)^{-1}]\,q$$

so the opposite inequality holds.

Theorem 3.2. *If S_0 is normalized or radial then $|p - q| = \varrho(p, q)$ for all $p, q \in S_0$, $|\cdot|$ is a norm if and only if ϱ is a metric.*

Proof. Let us first assume S_0 is normalized and $p, q \in S_0$. If $p - q = cp_1 - dq_1$, $c, d \geqq 0$, $p_1, q_1 \in S_0$ then $p + dq_1 = q + cp_1$ so

$$\begin{aligned}(1 + d)\,&[(1 + d)^{-1}\,p + d(1 + d)^{-1}\,q_1]\\ &= (1 + c)\,[1(1 + c)^{-1}\,q + c(1 + c)^{-1}\,p_1].\end{aligned}$$

Now since $q_2 \equiv (1 + c)^{-1}\,q + c(1 + c)^{-1}\,p_1 \in S_0$ and $(1 + c)\,(1 + d)^{-1}\,q_2 \in S_0$ we must have $(1 + c)\,(1 + d)^{-1} = 1$ so $c = d$. Thus all the representations of

$p-q$ are of the form $p-q=c(p_1-q_1)$ for $c \geqq 0$, $p_1, q_1 \in S_0$.

$$\begin{aligned}\sigma(p,q) &= \inf\{0 \leqq \lambda < 1 : (1-\lambda)p + \lambda p_1 = (1-\lambda)q + \lambda q_1, p_1, q_1 \in S_0\}\\ &= \inf\{0 \leqq \lambda < 1 : p-q = \lambda(1-\lambda)^{-1}(q_1 - p_1), q_1, p_1 \in S_0\}\\ &= \inf\{c(c+1)^{-1}, c \geqq 0 : p-q = c(q_1-p_1), q_1, p_1 \in S_0\}\\ &= \inf\{c \geqq 0 : p-q = c(q_1-p_1)\}\,[\inf\{c \geqq 0 : p-q=c(q_1-p_1)\}+1]^{-1}\\ &= |p-q|\,[|p-q|+1]^{-1}.\end{aligned}$$

Hence $\varrho(p,q) = \sigma(p,q)[1-\sigma(p,q)]^{-1} = |p-q|$. Now let us assume that S_0 is radial. Then suppose $p-q = cp_1 - dq_1$, $p, q, p_1, q_1 \in S_0$ where $c > d$. Then $p-q = cp_1 - c(d/c)q_1$ where $(d/c)q_1 \in S_0$. Thus for any representation of the form $p-q = cp_1 - dq_1$ we have another representation $p-q = b(p_2 - q_2)$ where $b = \max(c,d)$, $p_2, q_2 \in S_0$. We then obtain $|p-q| = \inf\{c \geqq 0 : p-q = c(p_1-q_1), p_1, q_1 \in S_0\}$ and just as in the normalized case we obtain $\varrho(p,q) = |p-q|$. Now it is clear that if $|\cdot|$ is a norm then ϱ is a metric. Conversely, suppose ϱ is a metric, $x, y \in X$ and $|x-y| = 0$. First assume that S_0 is radial. Since $|x-y|=0$ there exist $p, q \in S_0$, $0 \leqq c, d \leqq 1$ such that $x - y = cp - dq = p_1 - q_1$ where $p_1, q_1 \in S_0$. Then $\varrho(p_1, q_1) = |p_1 - q_1| = 0$ so $p_1 = q_1$ and hence $x = y$. Next assume that S_0 is normalized. We first show that in this case $|p| = 1$ for all $p \in S_0$. Indeed, if $p = cp_1 - dq_1$, $p_1, q_1 \in S_0$, $c, d \geqq 0$, then $(1+d)^{-1}p + d(1+d)^{-1}q_1 = c(1+d)^{-1}p_1$ and hence $c = 1 + \mathrm{d}$. Since $d \geqq 0$, $c \geqq 1$ and hence

$$|p| = \inf\{\max(c,d) : p = cp_1 - dq_1, c, d \geqq 0, p_1, q_1 \in S_0\} \geqq 1.$$

But $p = p - 0q$ so $|p| = 1$. Now suppose $x - y = cp - dq$, $c, d \geqq 0$, $p, q \in S_0$. Then

$$0 = |x-y| = |cp - dq| \geqq \big||cp| - |dq|\big| = \big|c|p| - d|q|\big| = |c-d|,$$

so $c = d$. Hence $0 = |x-y| = c|p-q| = c\varrho(p,q)$, if $c \neq 0$, then $\varrho(p,q) = 0$, giving $p = q$ and $x = y$, if $c = 0$ then again $x = y$.

In most examples S_0 is taken to be a *positive convex set*; that is if $p \in S_0$ then $\alpha p \notin S_0$ for all $\alpha < 0$. Notice if S_0 is normalized, it is positive. Positivity has the advantage of making K a cone. Indeed if $\alpha p = -\beta p_1$, $\alpha, \beta \geqq 0$, $p, p_1 \in S_0$ then if $\alpha > 0$ we have $p = -(\beta/\alpha)p_1$ so $p_1 \in S_0$ and $-(\beta/\alpha)p_1 \in S_0$ which is a contradiction. Hence $\alpha = 0$ and $K \cap (-K) = \{0\}$. The relation between positivity and normalized is given in the following:

Corollary 3.3. *S_0 is normalized if and only if it is positive and $|p| = 1$ for all $p \in S_0$.*

Proof. Necessity has been proved in the previous theorem. Conversely suppose S_0 is positive and $|p| = 1$ for all $p \in S_0$. Then if $p \in S_0$ and $\alpha p \in S_0$ we have $1 = |\alpha p| = |\alpha|$ so $\alpha = \pm 1$. Now $\alpha \neq -1$ by positivity.

We now consider linear extensions of affine maps from S_0 to X. If S_0 is radial and $T: S_0 \to S_0$ we say that T is *homogeneous* if $T(\lambda p) = \lambda Tp$ for all $p \in S_0$ and $0 < \lambda < 1$.

Theorem 3.4. *Let S_0 be a normalized or a positive and radial convex set in a real vector space V and let $|\cdot|$ be the induced seminorm on the subspace $X = K - K \subseteq V$. If S_0 is normalized (positive radial) and $T: S_0 \to S_0$ is an affine (and homogeneous) map then T has a unique linear extension $\hat{T}$ to X and $\|\hat{T}\| \leqq 1$ (i.e., $|Tx| \leqq |x|$ for all $x \in X$). If furthermore T is a bijection then $\hat{T}$ is an isometry on X (i.e., $|\hat{T}x| = |x|$ for all $x \in X$).*

Proof. Any $x \in X$ admits a representation $x = cp - dq$, $c, d \geqq 0$, $p, q \in S_0$. Define $\hat{T}x = cTp - dTq$. To show $\hat{T}$ is well-defined suppose also that $x = c_1 p_1 - d_1 q_1$, $c_1, d_1 \geqq 0$, $p_1, q_1 \in S_0$. First suppose S_0 is normalized and T is affine. Then since $cp - dq = c_1 p_1 - d_1 q_1$ we have (notice $c_1 + d, c + d_1 > 0$) $c(c + d_1)^{-1} p + d_1(c + d_1)^{-1} q_1 = (c_1 + d)(c + d_1)^{-1} \cdot (c_1(c_1 + d)^{-1} p_k + d(c_1 + d)^{-1} q)$ so $c_1 + d = c + d_1$ and hence

$$c(c + d_1)^{-1} Tp + d_1(c + d_1)^{-1} Tq_1 = c_1(c + d_1)^{-1} Tp_1 + d(c_1 + d)^{-1} q .$$

It follows that $cTp = dTq = c_1 Tp_1 - d_1 Tq_1$. Next suppose S_0 is positive, radial and T is affine and homogeneous. By positivity we have $c_1 + d$, $c + d_1 > 0$. Now either $c_1 + d \leqq c + d_1$ or $c_1 + d \geqq c + d$. For concreteness assume the former and apply the facts that S_0 is radial and T is affine, homogeneous to obtain

$$\begin{aligned} &c(c + d_1)^{-1} Tp + d_1(c + d_1)^{-1} Tq_1 \\ &\quad = (c_1 + d)(c + d_1)^{-1} (c_1(c_1 + d)^{-1} Tp_1 + d(c_1 + d)^{-1} Tq) . \end{aligned}$$

Again it follows that $cTp - dTq = c_1 Tp_1 - dTq_1$. Thus $\hat{T}$ is well-defined and it is easy to show that $\hat{T}$ is a linear operator on X. To show $\hat{T}$ is a contraction we have for $x \in X$

$$\begin{aligned} |\hat{T}x| &= \inf\{\max(c, d) \colon \hat{T}x = cp - dq, c, d \geqq 0, p, q \in S_0\} \\ &\leqq \inf\{\max(c, d) \colon x = cp - dq, c, d \geqq 0, p, q \in S_0\} = |x| . \end{aligned}$$

Corollary 3.5. *Let $S_0, X, |\cdot|$ be as in the previous theorem. If T is a linear operator on X that leaves S_0 invariant then T is a contraction (i.e., $|Tx| \leqq |x|$ for all $x \in X$). If furthermore T restricted to S_0 is a bijection then T is an isometry (i.e., $|Tx| = |x|$ for all $x \in X$.)*

Theorem 3.6. *Let S_0 be a normalized or radial convex set in a real vector space V and let X be the generated subspace. Let ϱ be the intrinsic semi-metric on S_0 and $|\cdot|$ the induced seminorm on X. If (S_0, ϱ) is complete then so is $(X, |\cdot|)$.*

Proof. First note that for $p \in S_0$, since $p = 1p$ we have $|p| \leqq 1$. Assuming (S_0, ϱ) is complete, let x_n be a Cauchy sequence in X. We may assume that $|x_{n+1} - x_n| < 2^{-n}$ for $n = 1, 2, \ldots$. Now we can write $x_{n+1} - x_n = c_n p_n - d_n q_n$ where $0 \leqq c_n, d_n < 2^{-n}, p_n, q_n \in S_0$. We can assume $c_1, d_1 > 0$. Let

$$a_n = \sum_{i=1}^{n} c_i\,, \qquad b_n = \sum_{i=1}^{n} d_i\,.$$

Now

$$\left\{\sum_{i=1}^{n} a_n^{-1} c_i p_i\,,\ i = 1, 2, \ldots\right\}$$

and

$$\left\{\sum_{i=1}^{n} b_n^{-1} d_i q_i,\ i = 1, 2, \ldots\right\}$$

are Cauchy sequences in S_0. Indeed, it is clear that $\{a_n\}$ is a Cauchy sequence and we have

$$\begin{aligned} &\varrho\left(\sum_{i=1}^{n+k} a_{n+k}^{-1} c_i p_i, \sum_{i=1}^{n} a_n^{-1} c_i p_i\right) \\ &\quad = \left|[a_{n+1}^{-1} - a_n^{-1}] \sum_{i=1}^{n} c_i p_i + \sum_{i=n+1}^{n+k} a_{n+k}^{-1} c_i p_i\right| \\ &\quad \leqq (a_n^{-1} - a_{n+1}^{-1}) \sum_{i=1}^{n} c_i + a_{n+k}^{-1} \sum_{i=n+1}^{n+k} c_i \\ &\quad = 1 - a_n a_{n+k}^{-1} (a_{n+k} - a_n) = 2(1 - a_n a_{n+k}^{-1})\,, \end{aligned}$$

where the last term approaches zero as $n, k \to \infty$. Thus there are elements $p, q \in S_0$ such that $\sum_{i=1}^{n} a_n^{-1} c_i p_i \to p$ and $\sum_{i=1}^{n} b_n^{-1} d_i q_i \to q$. Suppose $a_n \to a$ and $b_n \to b$; we claim that $x_n \to x_1 + ap - bq$. Indeed,

$$\begin{aligned} |x_{n+1} - x_1 - ap + bq| &= |x_{n+1} - x_n + x_n - x_{n-1} + \cdots + x_2 - x_1 - ap + bq| \\ &\leqq |a| \left|\sum_{i=1}^{n} a^{-1} c_i p_i - p\right| + |b| \left|\sum_{i=1}^{n} b^{-1} d_i q_i - q\right| \\ &\leqq |a| \left[\left|\sum_{i=1}^{n} (a^{-1} c_i - a_n^{-1} c_i)\, p_i\right| + \left|\sum_{i=1}^{n} a_n^{-1} c_i p_i - p\right|\right] \\ &\quad + |b| \left[|\Sigma (b^{-1} d_i - b_n^{-1} d_i)\, q_i| + \left|\sum_{i=1}^{n} b_n^{-1} d_i q_i - q\right|\right] \\ &\leqq |a| \left[1 - a^{-1} a_n + \left|\sum_{i=1}^{n} a_n^{-1} c_i p_i - p\right|\right] \\ &\quad + |b| \left[1 - b^{-1} b_n + \left|\sum_{i=1}^{n} b_n^{-1} d_i q_i - q\right|\right] \end{aligned}$$

$\to 0$ as $n \to \infty$.

We next consider the geometry of K relative to $|\cdot|$.

Lemma 3.7. *Let S_0 be a normalized or radial convex set in V, X the generated subspace and $|\cdot|$ the induced seminorm on X.* (1) *If $x, y \in X$ and $x-y \in K$, then $|x| \geqq |y|$.* (2) *Let $\gamma > 1$. For every $x \in X$ there exists $x_1, x_2 \in K$ such that $x = x_1 - x_2$ and $|x_i| \leqq \gamma |x|$, $i = 1, 2$.*

Proof. (1) Since $x-y \in K$ there is a $c_0 \geqq 0$ and $p_0 \in S_0$ such that $x - y = c_0 p_0$. We may assume $c_0 > 0$. If $y = cp - dq$, $c, d \geqq 0$, $p, q \in S_0$ then $x = y + c_0 p_0 = (c_0 + c)[c_0(c+c_0)^{-1} p_0 + c(c+c_0)^{-1} p] - dq$ and hence $|x| \geqq |y|$. (2) For $x \in X$, by definition of $|x|$ there exist $c, d \leqq \gamma |x|$, $p, q \in S_0$ such that $x = cp - dq$. Letting $x_0 = cp$, $x_2 = dq$ we have $x_1, x_2 \in K$, $x = x_1 - x_2$ and $|x_1| = |cp| \leqq c \leqq \gamma |x|$ and $|x_2| \leqq \gamma |x|$.

If K is a cone (e. g. if S_0 is positive) it follows from Lemma 3.7 that K is a normal strict b-cone in X [12].

Another norm on X that appears in the literature (we consider this norm in the next section) is the *natural seminorm*

$$|x|_1 = \inf\{c + d : x = cp - dq; c, d \geqq 0, p, q \in S_0\} .$$

It is clear that $|x| \leqq |x|_1 \leqq 2|x|$. The natural seminorm reduces to the *natural semimetric* $\varrho_1(p, q)$ which is twice the intrinsic semimetric. Indeed from the proof of Theorem 3.2 we have for $p, q \in S_0$

$$\begin{aligned}\varrho_1(p,q) &= |p-q|_1 = \inf\{c+d : p-q = cp_1 - dq_1, c, d \geqq 0, p_1, q_1 \in S_0\}\\ &= \inf\{2c : p - q = c(p_1 - q_1), c \geqq 0, p_1, q_1 \in S_0\}\\ &= 2|p-q| = 2\varrho(p,q) .\end{aligned}$$

Let us now compare our theory with the usual Hilbert space theory. In this case the normalized states $\mathscr{D}$ are the convex set of density operators; that is, the set of positive trace class operators with trace one. $\mathscr{D}$ is a base for the cone K of positive trace class operators which forms the set of states and K generates the real vector space $X = K - K$ of generalized states consisting of the self-adjoint trace class operators. In their latest study of scattering theory, Jauch, Misra and Gibson [8] use the trace norm $|\cdot|_1$ on X; that is, for $x \in X$, $|x|_1 = \Sigma |\lambda_i|$ where λ_i are the repeated eigenvalues of x. Now $|\cdot|_1$ reduces to the trace metric ϱ_1 on $\mathscr{D}$. We now compare the intrinsic metric ϱ and the induced norm $|\cdot|$ to ϱ_1 and $|\cdot|_1$ respectively.

Lemma 3.8. *Let $x \in X$ and let $\{\lambda_i\}, \{-\mu_i\}$ be the repeated positive eigenvalues and negative eigenvalues respectively of x. Then $x = x_1 - x_2$, where $x_1, x_2 \in K$, $tr x_1 = \Sigma \lambda_i$, $tr x_2 = \Sigma \mu_i$ and if $x = y_1 - y_2$, $y_1, y_2 \in K$ then $tr y_1 \geqq tr x_1$, $tr y_2 \geqq tr x_2$.*

Proof. Let $f^+(\lambda) = \lambda$ for $\lambda \geqq 0$ and $f^+(\lambda) = 0$ for $\lambda \leqq 0$. Let $f^-(\lambda) = -\lambda$ for $\lambda \leqq 0$ and $f^-(\lambda) = 0$ for $\lambda \geqq 0$. Then by the spectral theorem

$$\begin{aligned}x &= \int \lambda P^x(d\lambda) = \int_{\{\lambda > 0\}} \lambda P^x(d\lambda) + \int_{\{\lambda < 0\}} \lambda P^x(d\lambda) = f^+(x) - f^-(x)\\ &\equiv x_1 - x_2 .\end{aligned}$$

It is clear that the spectra $\sigma(x_1)=\{\lambda_i\}, \sigma(x_2)=\{\mu_i\}$. Now suppose $x=y_1=y_2, y_1, y_2 \in K$. If ϕ is an eigenvector of x corresponding to a positive eigenvalue then

$$\begin{aligned}\langle y_1\phi, \phi\rangle &= \langle x\phi, \phi\rangle + \langle y_2\phi, \phi\rangle = \langle x_1\phi, \phi\rangle + \langle y_2\phi, \phi\rangle \\ &\geqq \langle x_1\phi, \phi\rangle .\end{aligned}$$

If ϕ is an eigenvector of x corresponding to a negative eigenvalue then $\langle y_1\phi, \phi\rangle \geqq 0 = \langle x_1\phi, \phi\rangle$. Computing the traces using the eigenvectors of x we have $\operatorname{tr} y_1 \geqq \operatorname{tr} x_1$ and similarly $\operatorname{tr} y_2 \geqq \operatorname{tr} x_2$.

Corollary 3.9. *If $x \in X$ then*

$$|x|_1 = \inf\{c+d : x = cp_1 - dq_1, c, d \geqq 0, p_1, q_1 \in \mathscr{D}\} .$$

Proof. Applying Lemma 3.8,

$$x = x_1 - x_2 = (\operatorname{tr} x_1)\,[(\operatorname{tr} x_1)^{-1}\, x_1] - (\operatorname{tr} x_2)\,[(\operatorname{tr} x_2)^{-1}\, x_2]$$

so

$$|x|_1 = \operatorname{tr} x_1 + \operatorname{tr} x_2 = \inf\{c+d : x = cp_1 - dq_1, c, d \geqq 0, p_1, q_1 \in \mathscr{D}\} .$$

This completes the proof.

It follows from Lemma 3.8 that $|x| = \max(|x_1|_1, |x_2|_1) = \max(\Sigma\lambda_i, \Sigma\mu_i)$. Applying Corollary 3.9 we see that $|\cdot|_1$ is the natural norm considered in the previous section. It follows from the work of that section that $|x| \leqq |x|_1 \leqq 2|x|$ so $|\cdot|$ and $|\cdot|_1$ are equivalent and that $\varrho_1(p,q) = 2\varrho(p,q)$ for all $p, q \in \mathscr{D}$. In particular if $\{\lambda_i\}$ and $\{-\mu_i\}$ are the repeated positive and negative eigenvalues of $p-q$ we have $\Sigma\lambda_i + \Sigma\mu_i = 2\max(\Sigma\lambda_i, \Sigma\mu_i)$. Thus $\Sigma\lambda_i = \Sigma\mu_i$ which can also be easily derived from other means.

Jauch, Misra and Gibson [8] have shown that if $p, q \in \mathscr{D}$ are pure states corresponding to unit vectors ϕ, ψ then $\varrho_1(p,q) = 2(1 - |\langle\phi,\psi\rangle|^2)^{1/2}$. It follows that $\varrho(p,q)^2 = 1 - |\langle\phi,\psi\rangle|^2$.

4. Operational Quantum Mechanics

Returning to our general convex structure of states S, suppose there are enough observables to distinguish between states or equivalently that S^* is total. It then follows from Theorem 2.2 that S_1 is isomorphic to a convex set in a real linear space V. Then S is a cone with base S_1 and $X = S - S$ is a subspace of V generated by S. Let us make the reasonable assumption that the intrinsic semimetric ϱ is a metric and form the completion $\tilde{S}_1$ of S_1. Then $\tilde{S}_1$ is a base for a cone $\tilde{S}$ which generates the linear space $\tilde{X} = \tilde{S} - \tilde{S}$. If $|\cdot|$ is the intrinsic norm on $\tilde{X}$ then $\tilde{S}_1$ is normalized and hence by Theorem 3.6 $(\tilde{X}, |\cdot|)$ is a Banach space with closed generating cone $\tilde{S}$. Defining τ as above we see that $\tau(x) = |x|$ for every $x \in \tilde{S}$ and the triple $(\tilde{X}, \tilde{S}, \tau)$ becomes a complete base normed space (or state space), the basic framework for the operational quantum mechanics of Davies and Lewis. The norm $|\cdot|$ is equivalent to the natural norm $|\cdot|_1$ used by Davies and Lewis and applying Theorem 3.4 one can give definitions of observables, instruments, joint distributions, etc. on S which reduce to theirs. We thus see that in the case of a separating set of observables our framework reduces to that of Davies and Lewis and hence gives a simple, axiomatic motivation for their theory.

Mielnik gives convincing arguments for describing quantal situations in terms of a basic mathematical structure called a *quantum system* (e, D, T, B). In this structure, D is a linear subspace of the algebraic dual X^* of a real linear space X. Elements of D are called *detectors*. We give X the topology induced by D and assuming D separates points of X this makes X into a Hausdorff locally convex space. B is a generating closed cone for X whose elements are called *beams*. Denoting the set of continuous linear operators on X by $\mathscr{L}(X)$, T is a positive algebra in $\mathscr{L}(X)$ such that $T(B) \subseteqq B$. That is, T is closed under composition, addition, and multiplication by nonnegative scalars. T is called the *algebra of transmitters*. Finally $e \in D$ is a distinguished detector called the *standard quantum detector* or *quantum scale* and satisfies (1) $ex \geqq 0$ for all $x \in B$, (2) if $x \in B$ and $ex = 0$ then $x = 0$. The set $S = \{x \in B : ex = 1\}$ is the *figure of states* or the *statistical figure*. Although the Mielnik formalism is similar to that of Davies and Lewis, the former stresses convex set methods while the latter stresses the operational methods of probability theory.

It is clear that a Mielnik quantum system is a special case of our convex structure. Conversely, if $(S_1, \langle \cdot, \cdot, \cdot \rangle)$ is a convex structure then S_1 gives the statistical figure, $S = \{\alpha S_1 : \alpha \geqq 0\}$ is the set of beams, a subspace $D \subseteqq S^*$ gives the set of detectors, $Af(S)$ the set of transmitters, and τ the quantum scale. If it is assumed that S^* (or as Mielnik assumes, D) is total over S then applying Theorems 2.2 and 3.4 our structure reduces to a Mielnik quantum system. However, if S^* is not total over S then we get a non-linear generalization of Mielnik's theory.

References

1. Cornette, W., Ph. D.: Dissertation, University of Denver (in preparation).
2. Davies, E. B., Lewis, J. T.: An operational approach to quantum probability. Commun. math. Phys. **17**, 239—260 (1970).
3. Edwards, C. M.: Classes of operations in quantum theory. Commun. math. Phys. **20**, 26—56 (1971).
4. Gudder, S.: Axiomatic quantum mechanics and generalized probability theory in A. Bharucha-Reid (editor), Probabilistic Methods in Applied Mathematics, Vol. II. New York: Academic Press 1970.

5. Gunson. J.: Structure of quantum mechanics. Commun. math. Phys. **6**, 262—285 (1967).
6. Haag, R., Kastler, D.: An algebraic approach to quantum field theory. J. Math. Phys. **5**, 848—861 (1964).
7. Jauch, J.: Foundations of quantum mechanics. Reading, Mass.: Addison-Wesley 1968.
8. Jauch, J., Misra, B., Gibson, A.: On the asymptotic condition of scattering theory. Helv. Phys. Acta **41**, 513—527 (1968).
9. Mielnik, B.: Geometry of quantum states. Commun. math. Phys. **9**, 55—80 (1968).
10. Mielnik, B.: Theory of filters. Commun. math. Phys. **15**, 1—46 (1969).
11. Neumann, H.: Classical systems and observables in quantum mechanics. Commun. math. Phys. **23**, 100—116 (1971).
12. Peressini, A.: Ordered Topological Vector Spaces. New York: Harper and Row 1967.
13. Stolz, P.: Attempt of an axiomatic foundation of quantum mechanics and more general theories, VI. Commun. math. Phys. **23**, 117—126 (1971).
14. Segal, I.: Mathematical problems of relativistic physics. American Math. Soc. Lectures in Applied Mathematics. Providence, R.I. (1963).
15. Varadarajan, V.: Geometry of Quantum Theory, Vol. I. Princeton, N.J.: Van Nostrand 1968.

S. Gudder
Department of Mathematics
University of Denver
Denver Colorado 80210, USA

Erratum

Non-Existence of Axially Symmetric Massive Scalar Fields

A. R. Roy and J. R. Rao

Commun. math. Phys. **27**, 162—166 (1972)

Page 165: 4th line onwards from the top reads as follows:

Which in view of (3.4) implies

$$-\mu^2 V^2 + g^{11} g^{44} (F_{14})^2 - g^{22} g^{33} (F_{23})^2 = 0 . \tag{3.5}$$

Since g^{11}, g^{22}, g^{33} are all negative and g^{44} is positive, (3.5) will hold iff

$$\mu = 0, \quad F_{14} = 0, \quad F_{23} = 0 . \tag{3.6}$$

Hence, there cannot exist any solution for the coupled electromagnetic and massive scalar fields for the metric (2.1).

The rest of the calculations are unnecessary and may, therefore, be ignored.