

Université de Grenoble - Summer School of Theoretical Physics
Les Houches 1970
Supported by NATO
and the Commissariat à l'Energie Atomique

STATISTICAL MECHANICS
AND QUANTUM FIELD THEORY
MECANIQUE STATISTIQUE
ET THEORIE QUANTIQUE
DES CHAMPS

edited by C. DeWitt
Faculté des Sciences, Grenoble and
University of North Carolina, Chapel Hill
R. Stora
Centre d'Etudes Nucléaires de Saclay
Service de Physique théorique

1971

GORDON AND BREACH SCIENCE PUBLISHERS
New York London Paris

Selected Topics in Functional Analysis

Oscar E. Lanford III

Dept. of Mathematics

University of California, Berkeley, California

I Introduction

In this set of lectures I will attempt to present some of the mathematical tools which are currently of importance in the investigation of the statistical mechanics of infinite systems. It is my intention to give a reasonably self-contained *introduction* to the rudiments of the theory of integral representations on compact convex sets (Choquet theory), C^* -algebra theory, and von Neumann algebra theory. To do this I will need to make frequent use of the results of more "elementary" branches of mathematics—general topology, measure theory, Hilbert space theory, and the theory of topological vector spaces. While it is out of the question to try to make a systematic presentation of these preliminaries, I will attempt to summarize some of the most important points. With this summary, a little elementary knowledge of real analysis, and some good will, I hope it will be possible to make at least some sense out of the lectures.

A General Topology

We start with a very elementary definition: A *topological space* is a set X together with a collection \mathcal{G} of subsets of X , called open sets, such that:

- 1) X and ϕ (the null set) both belong to \mathcal{G} .
- 2) The intersection of two elements of \mathcal{G} again belongs to \mathcal{G} .
- 3) The union of any family of elements of \mathcal{G} belongs to \mathcal{G} .

Intuitively, one thinks of an open set as one which, if it contains x , also contains all other points "sufficiently near" to x . Thus, specifying a topology (a family of open sets) may be viewed as introducing a notion of approximation; from this point of view, the motivation for studying general topology, as opposed to the theory of metric spaces, may be seen as providing tools to investigate notions of approximation more subtle than those where one has a numerical measure of closeness. We will see plenty of examples later on.

If X is a topological space, a subset F of X is said to be *closed* if its complement $X \setminus F$ is open. In metric spaces, one has another way of describing closed sets—a set is closed if no sequence in it converges to a point outside it. We can give a corresponding characterization in general topology, at the expense of introducing the notion of net, which generalizes that of sequence. First, we recall that a sequence (x_n) in a set X is an indexed family of elements of X , the index set being $\{1, 2, 3, \dots\}$. We say that x_n converges to x if for every open set W containing x , x_n is eventually in W , i.e., there is a N such that, for $n \geq N$, $x_n \in W$. A net in X is, like a sequence, an indexed family of elements of X , but the index set is more general than $\{1, 2, 3, \dots\}$. We say

that a binary relation \geq on a set A is a *pre-order* if we have $\alpha \geq \alpha$ for all $\alpha \in A$, $\alpha \geq \beta$ and $\beta \geq \gamma$ implies $\alpha \geq \gamma$ for $\alpha, \beta, \gamma \in A$. A pre-ordered set A is said to be *directed* (*filtrant*) if, for any $\alpha, \beta \in A$, there exists $\gamma \in A$ such that $\gamma \geq \alpha$ and $\gamma \geq \beta$. A *net* in a set X is an indexed family (x_α) of elements of X , indexed by a directed set A (i.e., a net in X is a mapping $\alpha \mapsto x_\alpha$ from a directed set A to X). A net (x_α) *converges to* x if, for any open set W containing x , there exists an $\alpha_0 \in A$ such that $x_\alpha \in W$ for $\alpha \geq \alpha_0$. (We express this by saying that x is *eventually* in W .) We can now provide the promised characterization of closed set: A set F in a topological space is closed if, whenever a net (x_α) in F converges to a point x in X , the point x is actually in F .

It is perfectly possible (perhaps even desirable) to develop general topology without ever mentioning nets. The advantage of talking about nets is purely heuristic; it permits the use of intuition about sequences in Euclidean space to suggest arguments valid in general topological spaces. This tool must, however, be used cautiously; not *every* argument about sequences has a counterpart valid for nets. We shall soon see some examples.

It follows from the above characterization of closed set that, if one knows which nets converge to which limits, one knows what all the closed sets are; hence, what all the open sets are, i.e., one knows the topology. This suggests that one can specify a topology by specifying convergent nets. This can indeed be done, but of course not every collection of nets is precisely the set of convergent nets for some topology. (For example, a subnet (to be defined shortly) of a convergent net, like a subsequence of a convergent sequence, must converge.) For an axiomatization of topological spaces in terms of convergent nets, rather than open sets, see Kelley [1], p. 73. We will frequently define topologies by specifying what nets converge to what limits, leaving it as an exercise to verify that such a topology exists. (It is usually easier to construct the topology directly than to verify that the conditions in Kelley's book hold.) As an example, consider the product topology: Let $(X_i)_{i \in I}$ be an indexed family of topological spaces; we define a topology on the product set $\prod_i X_i$ by requiring that a net (x'_α) converges to (x') if and only if $\lim x'_\alpha = x'$ for all i . It is simple to verify that this description is equivalent to the usual construction of a product topology on $\prod_i X_i$.

In metric spaces, there are many equivalent characterizations of compact spaces: A space is compact if every open cover has a finite subcover, or if every sequence has a cluster point, or if every sequence has a convergent subsequence. These conditions cease to agree in general topological spaces. The first condition turns out to be the most useful one and is therefore taken as the general definition of compactness; examples can then be made of

subsequence. (It remains true that every sequence in a compact space has a cluster point.) However, by replacing "sequence" by "net", the alternative descriptions of compact sets can be saved. Thus, one says that a point x is a *cluster point* of a net (x_α) if, for every open set W containing x , and every $\alpha \in A$, there is an $\alpha' \geq \alpha$ such that $x_{\alpha'} \in W$ (We express this by saying that the net (x_α) is *frequently* in W). A topological space is then compact if and only if every net has a cluster point. A *subnet* of a net $(x_\alpha)_{\alpha \in A}$ is a net of the form $\beta \mapsto x_{\phi(\beta)}$, where ϕ is a mapping of the directed set B to the directed set A such that, for any $\alpha \in A$ there exists $\beta \in B$ with the property that $\phi(\beta') \geq \alpha$ if $\beta' \geq \beta$. (In other words, $\phi(\beta)$ becomes "large" in A as β becomes "large" in B .) Note that a subnet of $(x_\alpha)_{\alpha \in A}$ may have an index set which is not merely a subset of A . In particular, sequences may have subnets which are not sequences. Now it turns out that, with this definition, a point x is a cluster point of a net (x_α) if and only if the net has a subnet converging to x . Thus, a topological space X is compact if and only if every net has a convergent subnet.

There is yet another characterization of compact spaces, having no sequence analogue, corresponding to the fact that one can construct nets which cannot really be refined any further, and which hence converge if they have any cluster point at all. To be precise: A net (x_α) is said to be a *universal net* if, for any subset Y of X , (x_α) is either eventually in Y or eventually in $X \setminus Y$. It is nearly a tautology that, if (x_α) is a universal net and x is a cluster point of (x_α) , then $\lim x_\alpha = x$. It is also not hard to see that, if (x_α) is a universal net in X and if f is any mapping $X \rightarrow X'$, then $(f(x_\alpha))$ is a universal net in X' . What is not at all obvious, but true nevertheless, is that every net has a universal subnet (See Kelley [1], Chapter 2, Ex. J., p. 81.) Hence, a topological space is compact if and only if every universal net converges. This result leads immediately to Tychonoff's Theorem:

THEOREM Let $(X_i)_{i \in I}$ be an indexed family of compact spaces. Then $\prod_i X_i$ is compact.

Proof Let (x_i^α) be a universal net in $\prod_i X_i$. Then, for each i , x_i^α is a universal net in X_i ; hence, converges. By the definition of the product topology (x_i^α) converges in $\prod_i X_i$.

Let us look at another, related, example. Let \mathcal{X} be a Banach space, \mathcal{X}^* the dual space of \mathcal{X} (the space of continuous linear functionals on \mathcal{X}). The *weak-** topology on \mathcal{X}^* is the topology for which ϕ_α converges to ϕ means $\phi_\alpha(\xi)$ converges to $\phi(\xi)$ for all $\xi \in \mathcal{X}$. We now have the following theorem, which is usually derived as a corollary of the Tychonoff Theorem:

Theorem. The unit ball in the dual of
 114 Banach space X^* is compact in the weak- $*$ topology. E. LANFORD III

Proof Let ϕ_α be a universal net in the unit ball of X^* . For any $\xi \in X$, $\phi_\alpha(\xi)$ is a universal net of complex numbers of absolute value $\leq \|\xi\|$; since the set of such numbers is compact (ξ fixed!) $\phi_\alpha(\xi)$ converges for all ξ . Call the limit $\phi(\xi)$. Then $\xi \mapsto \phi(\xi)$ is linear, and $|\phi(\xi)| = \lim |\phi_\alpha(\xi)| \leq \|\xi\|$ for all ξ , so ϕ is an element of the unit ball of X^* , and the net ϕ_α converges in the weak- $*$ topology to ϕ . Thus, every universal net in the unit ball of X^* converges, so the unit ball of X^* is compact.

In many cases, we find ourselves forced to consider several different topologies on a single set. (For example, on the dual of a Banach space, one may consider the norm topology as well as the weak- $*$ topology.) The comparison of different topologies is usually confusing although in principle simple: We say that a topology \mathcal{T}_1 is stronger, or finer, than a topology \mathcal{T}_2 if every \mathcal{T}_2 -open set is also \mathcal{T}_1 -open (i.e., if \mathcal{T}_1 has more open sets than \mathcal{T}_2). Since a set is closed if and only if its complement is open, \mathcal{T}_1 also has more closed sets than \mathcal{T}_2 . Since the closure of a set $E \subset X$ is the intersection of all closed sets containing E , the closure in a stronger topology is smaller than in a weaker topology. If x_α is a net converging to x in a topology \mathcal{T} , it also converges to x in all weaker topologies, but not necessarily in stronger topologies. If one starts with a continuous map between two spaces, and weakens the topology on the range space or strengthens the topology on the domain space, the map remains continuous. If on the other hand, one strengthens the topology on the range space or weakens the topology on the domain space the map need not remain continuous. On the dual of a Banach space, the weak- $*$ topology is, as one would hope, weaker than the topology defined by the norm.

B Measure Theory

We continue with our impressionistic survey of the elements of analysis, by looking, first at abstract measure theory, then at measure theory on locally compact spaces. To start with: Let X be a set, Σ a collection of subsets of X . We say that Σ is a *ring* if $\phi \in \Sigma$ and if, for all $E, F \in \Sigma$, $E \cup F$, $E \cap F$, and $E \setminus F$ belong to Σ , and we say that Σ is a σ -*ring* if, in addition,

$\bigcup_{i=1}^{\infty} E_i \in \Sigma$ whenever E_1, E_2, \dots all belong to Σ . If the set X itself belongs to Σ we say that Σ is an *algebra* or σ -*algebra*. Given any collection of subsets, there is a unique smallest σ -ring containing all of them; we refer to this as the σ -ring generated by the collection of sets in question. Given a σ -ring Σ , we define a measure μ on Σ to be a mapping from Σ to the positive real numbers and $+\infty$ such that if E_1, E_2, \dots are in Σ with $E_i \cap E_j = \emptyset$ for

$i \neq j$ then $\mu(E_1 \cup E_2 \cup \dots) = \sum_i \mu(E_i)$ (countable additive) (to avoid

pathologies, we also require $\mu(\phi) = 0$; this eliminates the possibility $\mu(E) = \infty$ for all $E \in \Sigma$.)

With this general set-up (a set, a σ -ring, and a measure), one can define an integral, and one gets an integration theory with most of the nice formal properties of the Lebesgue-integral. In particular the monotone convergence theorem and the dominated convergence theorem hold. (It should be remarked that these theorems hold for *sequences*; the analogous statements for nets are false. For example, every positive real-valued function on the real line is the pointwise limit of an increasing net of functions each of which is zero except at a finite number of points.)

Now let X be a locally compact topological space (i.e., one in which each point belongs to some open set with a compact closure.) We want to investigate integration theory which will permit us to integrate continuous functions of compact support. As a minimal sort of condition to permit us to do this, we must have that, for any continuous non-negative function f of compact support, and every $\lambda > 0$, $\{x \in X : f(x) \geq \lambda\}$ belongs to the σ -ring Σ . This set is closed and contained in the support of f ; hence, is compact. However, it is a rather special kind of compact set: It can be written as the

intersection of countably many open sets by $\{x : f(x) \geq \lambda\} = \bigcap_{n=1}^{\infty} \{x : f(x) > \lambda - 1/n\}$. A set which can be obtained as a countable intersection of open sets is called a G_δ . (In a metric space, every closed set is a G_δ , so the distinction between general compact sets and compact G_δ 's becomes irrelevant.)

We consider, therefore, measures on the σ -ring generated by the compact G_δ 's. Elements of this σ -ring are called *Baire sets*. In order to ensure that the integral of every continuous function of compact support is finite, we also impose the condition that the measure of every compact G_δ is finite. Thus: A *Baire measure* on a locally compact space is a measure on the σ -ring of Baire sets which is finite on compact G_δ 's. It turns out that every continuous function of compact support is integrable with respect to every Baire measure μ . The mapping $f \mapsto \int f d\mu$ is linear and positive (positive means that $\int f d\mu \geq 0$ if $f(x) \geq 0$ everywhere). Thus each Baire measure determines a positive linear functional on the vector space of continuous functions of compact support. Conversely, every such functional comes from a measure, and the measure is uniquely determined by the functional. Thus we have:

THEOREM I. B. 1 (*Preliminary form of the Riesz Representation Theorem*): *The positive linear functionals on the space of continuous function of compact support on a locally compact space X are in one-one correspondence with the Baire measures on X the correspondence being defined by associating with a*

measure μ the functional $f \mapsto \int f d\mu$

This version of the Riesz Theorem is not fully satisfactory, since it does not permit us to speak of the measure of a general compact set. Hence, we want to extend Baire measures from the σ -ring of Baire sets to the larger σ -ring of Borel sets, defined as the σ -ring generated by the compact sets. It is always possible to do this, but the extension need not be unique. There is, however, a useful way of picking out a unique extension. To describe how this is done, we first mention a remarkable continuity property of Baire measures: Let μ be a Baire measure, E a Baire set. Then $\mu(E) = \sup \{ \mu(K) : K \text{ a compact } G_\delta \text{ contained in } E \} = \inf \{ \mu(O) : O \text{ an open Baire set containing } E \}$.

In other words, Baire sets can be approximated arbitrarily well from the inside by compact G_δ 's and from the outside by open Baire sets. We say that a Borel measure μ (measure on the σ -ring of Borel sets assigning finite values to compact sets) is inner regular if, for each Borel set E , $\mu(E) = \sup \{ \mu(K) : K \text{ compact; } K \subset E \}$; outer regular if, for each Borel set E , $\mu(E) = \inf \{ \mu(O) : O \text{ an open Borel set, } E \subset O \}$; regular if both inner regular and outer regular. It can then be shown that every Baire measure has a unique extension to a regular Borel measure. Hence:

THEOREM I, B. 2 (Riesz Representation Theorem): *Let X be a locally compact topological space. Then the set of positive linear functionals on the space of continuous functions of compact support on X is in one-one correspondence with the set of regular Borel measures on X , the correspondence being that associating with a measure μ the functional $f \mapsto \int f d\mu$.*

It should be made clear that the terminology we have adopted is far from universal. We have followed that used by Halmos [1], Chapter X. In particular the term "Borel set" is frequently taken to mean an element of the σ -algebra generated by the closed sets. A regular measure defined on the σ -ring generated by the compact sets may always be extended to a measure on the σ -algebra generated by the closed sets. The extension, however, is not in general unique. Furthermore, it cannot in general be taken to be both inner and outer regular. It is, however, always possible to extend a Borel measure to an inner regular measure on the σ -algebra generated by the closed sets, and the extension is evidently unique.

Once this extension has been made, we may formulate a slight modification of the monotone convergence theorem which will be useful later on. A real-valued function f is said to be lower semi-continuous if, for all real γ , $\{x : f(x) > \gamma\}$ is open. (Alternatively, f is lower semi-continuous if it is the supremum of a family of continuous functions.) f is upper semi-continuous if $-f$ is lower semi-continuous. If μ is a regular Borel measure extended as above to an inner-regular measure on the σ -algebra generated by

the closed sets and if f is a non-negative lower semi-continuous function, then $\int f d\mu$ makes sense (but may be $+\infty$). We now have the following result:

THEOREM I. B. 3 Let f_α be an increasing net of non-negative lower semi-continuous functions, and let $f = \lim f_\alpha$. Then f is lower semi-continuous and $\int f d\mu = \lim \int f_\alpha d\mu$. (In other words, the monotone convergence theorem is valid for increasing nets, rather than sequences, provided that the functions involved are lower semi-continuous.)

The proof of the theorem is not difficult; it uses the inner regularity of μ and the fact that, if K is a compact set and $f(x) > \gamma$ for $x \in K$ then $f_\alpha(x) > \gamma$ for $x \in K$ for all sufficiently large α .

C Topological Vector Spaces

A *semi-norm* on a vector space E is a function $\xi \mapsto \|\xi\|$ from E to the non-negative real numbers such that:

- a) $\|\lambda\xi\| = |\lambda| \cdot \|\xi\|$
- b) $\|\xi + \eta\| \leq \|\xi\| + \|\eta\|$.

The first condition evidently implies that $\|0\| = 0$; if in addition we have

- c) $\|\xi\| = 0$ implies $\xi = 0$ we say that $\|\cdot\|$ is a *norm* on E .

Given a family $(\|\cdot\|_i)_{i \in I}$ of semi-norms on a vector space E , we may define a topology on E by requiring that $\xi_\alpha \rightarrow \xi$ if and only if $\|\xi_\alpha - \xi\|_i \rightarrow 0$ for all i . This makes E into a topological vector space (i.e., a vector space, with a topology, such that the algebraic operations are continuous); a topological vector space whose topology obtained in this way from a family of semi-norms is called a *locally convex* topological vector space. We will normally want to consider only *Hausdorff* locally convex spaces, i.e., those for which, for any $\xi \neq 0$, there exists i such that $\|\xi\|_i \neq 0$. If the topology of a locally convex space is defined by a single norm, we speak of a *normed* vector space; a complete normed vector space is called a *Banach* space.

If E is any vector space, and ϕ is a linear functional on E , then $\ker(\phi) = \{\xi \in E: \phi(\xi) = 0\}$ is a linear subspace of E of codimension 1 (i.e., $E/\ker(\phi)$ is one-dimensional). (A linear subspace of codimension 1 in a vector space, or, more generally, a subset obtained by translating such a subset, is called a *hyperplane*.) Conversely, if E is a vector space, and F is a hyperplane in E passing through 0, then there is a linear functional on E with F as its kernel, and this functional is unique up to multiplication by a scalar.

If, now, E is a topological vector space, and ϕ is a continuous linear

functional, then $\ker \varphi$ is closed. Conversely if $\ker \varphi$ is closed then φ is

continuous. It is worth remarking that, since the closure of a linear subspace is again a linear subspace, the closure of a hyperplane is either the hyperplane itself or all of E , i.e., a hyperplane is either closed or dense. Hence, if we have a hyperplane and a non-empty open set which does not intersect it, then the hyperplane must be closed, i.e., must be a translate of the kernel of a continuous linear functional. If E is a locally convex space, with a topology defined by a family of semi-norms $(\|\cdot\|_i)$, then a linear functional ϕ is continuous if and only if there exists a positive real number λ and finitely many indices i_1, \dots, i_n such that

$$\|\phi(\xi)\| \leq \lambda(\|\xi\|_{i_1} + \dots + \|\xi\|_{i_n}).$$

This characterization generalizes the familiar fact that a linear functional on a normed vector space is continuous if and only if it is bounded.

Perhaps the most powerful tool in the study of locally convex spaces is the Hahn-Banach Theorem, in its many different forms. We will state an algebraic version of the theorem which has all the other versions as more or less straightforward consequences. First we need some terminology. A set K in a vector space is *convex* if, for $\xi_1, \xi_2 \in K$ and $0 \leq \alpha \leq 1$, $\alpha\xi_1 + (1 - \alpha)\xi_2 \in K$. Geometrically, this means that the line segment from ξ_1 to ξ_2 is contained in K . A point ξ of a convex set K in a vector space E is said to be an *algebraic interior point* of K if, for all $\eta \in E$, $\xi + \lambda\eta \in K$ for sufficiently small positive λ . In other words, ξ is an algebraic interior point of K if one can move from ξ a finite distance in any direction without getting outside of K (but the distance may depend upon the direction chosen.) If E is a locally convex topological vector space, then any topological interior point of K (i.e., a point contained in an open set contained in K) is an algebraic interior point, but the converse need not be true. (But in a finite dimensional space, an algebraic interior point is the same thing as a topological interior point...).

The key theorem is now the following:

THEOREM I. C. 1 *Let E be a vector space over the real numbers, X a convex subset of E with at least one algebraic interior point, η a point of E not belonging to X . Then there exists a hyperplane F in E separating η from X in the sense that η is on one side of the hyperplane or on it and X is on the other side of the hyperplane (but possibly intersecting it). In other words, there is a non-zero linear functional ϕ on E and a real number λ such that $\phi(\xi) \leq \lambda$ for $\xi \in X$ and $\phi(\eta) \geq \lambda$.*

Note that the condition $\phi(\xi) \leq \lambda$ for all $\xi \in X$ implies that $\phi(\xi) < \lambda$ for all algebraic interior points of X . Indeed, if ξ is an algebraic interior point and $\phi(\xi) = \lambda$, then, choosing $\zeta \in E$ such that $\phi(\zeta) > 0$, and choosing $\gamma > 0$ such that $\xi + \gamma\zeta \in X$ we get $\phi(\xi + \gamma\zeta) > \lambda$ contradiction.

dicting the requirement that $\rho \leq \lambda$ on X .

We can now read off a geometric form of the Hahn Banach Theorem:

THEOREM I. C. 2 (*Geometric Form of the Hahn Banach Theorem*): Let E be a locally convex topological vector space over the real numbers; if X and Y are disjoint convex subsets of E at least one of which has a non-empty interior, then there is a closed hyperplane F in E separating X and Y . In other words, there is a continuous linear functional ϕ and a real number λ such that $\phi(\cdot) \geq \lambda$ on X and $\phi(\cdot) \leq \lambda$ on Y .

If we drop the assumption that one of the sets X, Y has an interior point, but assume instead that there is an open set U containing 0 such that $(X + U) \cap Y = \emptyset$, then there is a closed hyperplane strictly separating X and Y in the sense that $\phi(\cdot) \geq \lambda + \epsilon$ on X , and $\phi(\cdot) \leq \lambda - \epsilon$ on Y , for some $\epsilon > 0$. This condition holds in particular if X is compact, Y is closed, and $X \cap Y = \emptyset$.

Proof For the first assertion, note that the set $X - Y = \{\xi - \eta : \xi \in X, \eta \in Y\}$ is convex, has a non-empty interior, and does not contain 0 . Hence, by the preceding theorem there is a non-zero linear functional ϕ such that $\phi(\cdot) \geq 0$ on $X - Y$. Let $\lambda = \inf \{\phi(\xi) : \xi \in X\}$. Then $\lambda \geq \sup \{\phi(\eta) : \eta \in Y\}$, so the hyperplane $\{\xi : \phi(\xi) = \lambda\}$ separates X and Y . It remains only to prove that ϕ is continuous, but this is immediate since, by the argument given just above, the hyperplane $\phi(\xi) = \lambda$ cannot intersect the interior of either X or Y ; hence, cannot be dense; hence, must be closed. To prove the second assertion: since E is locally convex, we may assume that U is convex; replacing U by $U \cap (-U)$, we may also assume that U is symmetric. Apply the first assertion to the sets $X + \frac{1}{2}U, Y + \frac{1}{2}U$. Then we get $\phi(\cdot) \geq \lambda$ on $X + \frac{1}{2}U$ and $\phi(\cdot) \leq \lambda$ on $Y + \frac{1}{2}U$. But $\phi(U)$ is a symmetric open interval; if we choose $\epsilon > 0$ so that $\phi(U) \supset (-2\epsilon, 2\epsilon)$, then $\phi(X + \frac{1}{2}U) \supseteq \phi(X) + (-\epsilon, \epsilon)$; since $\phi(\cdot) \geq \lambda$ on $X + \frac{1}{2}U$, $\phi(\cdot) \geq \lambda + \epsilon$ on X . Similarly, $\phi(\cdot) \leq \lambda - \epsilon$ on Y . To prove the final assertion, we must first show that $X - Y$ is closed. Let $\xi_n - \eta_n$ be a convergent net in $X - Y$. Since X is compact we can, by passing to a subnet, assume that ξ_n converges to $\xi \in X$. Since $\xi_n - \eta_n$ and ξ_n converge separately, η_n must also converge, and the limit η must belong to Y since Y is closed. Hence, $\xi_n - \eta_n$ converges to $\xi - \eta \in X - Y$, so $X - Y$ is closed. Since $0 \notin X - Y$, there is a neighborhood U of 0 , which may be taken to be convex and symmetric, such that $(X - Y) \cap U = \emptyset$. Then $(X + U) \cap Y = \emptyset$.

COROLLARY I. C.3 Let E be a Hausdorff locally convex topological vector space over the real numbers, and let $\xi \in E, \xi \neq 0$. Then there is a continuous linear functional ϕ on E with $\phi(\xi) \neq 0$.

Proof Let the topology on E be defined by the family of seminorms

(11.11) $t \in I$. Then $\{0\} = \cap \{\eta \in E : \|\eta\|_t = 0\}$
 is closed similarly $\{\xi\}$ is closed.

Clearly, $\{\xi\}$ is compact. Hence, there exists a continuous linear functional ϕ strictly separating $\{0\}$ from $\{\xi\}$. Since $\phi(0) = 0$, we must have $\phi(\xi) \neq 0$.

Next, we derive the so-called analytic form of the Hahn-Banach Theorem.

THEOREM I. C. 4 (*Analytic Form of the Hahn Banach Theorem*). Let E be a vector space over the real numbers, F a subspace of E , $\|\cdot\|$ a semi-norm on E , and ϕ a linear functional on F such that $|\phi(\xi)| \leq \|\xi\|$ for $\xi \in F$. Then there exists a linear functional $\tilde{\phi}$ on E which extends ϕ and satisfies $|\tilde{\phi}(\xi)| \leq \|\xi\|$ for all $\xi \in E$.

Proof Consider the vector space $E \times \mathbb{R}$ with the product topology and the two convex sets

$$X = \text{graph of } \phi, Y = \{(\zeta, \lambda) : \lambda > \|\zeta\|\}.$$

Then $X \cap Y = \emptyset$, and every point of Y is an algebraic interior point. Hence, we can find a hyperplane separating $X - Y$ from 0 , i.e., separating X from Y . Let the hyperplane come from a linear functional ψ . Then, since X is a linear subspace $\psi(X) = \{0\}$ or \mathbb{R} , and the second alternative is ruled out by the condition that ψ separate X from Y . Hence, $\psi(X) = \{0\}$. Since every point of Y is an algebraic interior point, we must have $\psi(Y) \cap \psi(X) = \emptyset$. Thus, if $\psi((\zeta, \mu)) = 0$, then $(\zeta, \mu) \notin Y$, so $\mu \leq \|\zeta\|$. Since $\ker(\psi)$ is a linear subspace, we get $-\mu \leq \|- \zeta\| = \|\zeta\|$, so $|\mu| \leq \|\zeta\|$. Thus, $\ker(\psi)$ is the graph of a linear functional $\tilde{\phi}$ satisfying $|\tilde{\phi}(\zeta)| \leq \|\zeta\|$. Since $\ker(\psi) \supset X = \text{graph of } \phi$, $\tilde{\phi}$ extends ϕ .

COROLLARY I. C. 5 Let E be a locally convex topological vector space over the real numbers, F a subspace of E , ϕ a continuous linear functional on F . Then ϕ may be extended to a continuous linear functional on E .

The analytic form of the Hahn-Banach Theorem has a generalization which is occasionally useful: A sublinear functional on a vector space E is a mapping $p: E \rightarrow \mathbb{R}$ such that

- $p(\lambda\xi) = \lambda p(\xi)$ for all $\lambda \geq 0$ and all $\xi \in E$ (p is positively homogeneous).
- $p(\xi + \eta) \leq p(\xi) + p(\eta)$ for all $\xi, \eta \in E$ (p is subadditive).

Any semi-norm is sublinear, as is any linear functional. Essentially the same argument as that used in proving the analytic form of the Hahn-Banach Theorem proves:

THEOREM I. C. 6 Let E be a vector space over the real numbers, p a sublinear functional on E , F a subspace of E , ϕ a linear functional on F such that $\phi(\xi) \leq p(\xi)$ for all $\xi \in F$. Then ϕ may be extended to a linear functional $\tilde{\phi}$ on E such that $\tilde{\phi}(\xi) \leq p(\xi)$ for all $\xi \in E$.

We will also have occasion to use the following:

THEOREM I. C. 7 (*Extension Theorem for Positive Functionals*): Let E be a vector space over the real numbers, K a convex cone in E (i.e., if $\xi, \eta \in K$, $\lambda > 0$, then $\xi + \eta$ and $\lambda\xi$ are in K). Let F be a linear subspace of E containing an algebraic interior point of K and let ϕ be a linear functional on E which is

non-negative on $E \cap K$. Then there is an extension $\tilde{\phi}$ of ϕ to E which is non-

negative on K . If E is a locally convex space and F contains a point of the interior of K , then $\tilde{\phi}$ is continuous.

Proof We can assume $\phi \neq 0$. Then we must have $\phi(\xi) > 0$ for ξ any algebraic interior point of K . The set of algebraic interior points of K is convex, and does not intersect $\ker(\phi)$. Hence, there exists a linear functional ψ on E separating K from $\ker(\phi)$. Since $\ker(\phi)$ is a linear subspace of E , $\psi(\ker(\phi)) = \{0\}$ or \mathcal{R} ; because of the separation property, $\psi(\ker(\phi)) = \{0\}$, i.e., $\ker(\psi) \supset \ker(\phi)$. Again by the separation property, ψ must take on only one sign on the set of algebraic interior points of K ; we can assume, then, that $\psi(\xi) \geq 0$ for any algebraic interior point ξ of K . As usual, this implies $\psi(\xi) > 0$ for any algebraic interior point of K . If ξ is any algebraic interior point of K which is actually in F , we let $\tilde{\phi}$ be a multiple of ψ such that $\tilde{\phi}(\xi) = \phi(\xi) (> 0)$. Now the restriction of $\tilde{\phi}$ to F is a linear functional on F whose kernel contains that of ϕ and which agrees with ϕ on a vector not in the kernel of ϕ ; this implies that $\tilde{\phi}$ is an extension of ϕ . It remains to check that $\tilde{\phi}(\cdot) \geq 0$ on K ; we know that this is true on the set of algebraic interior points of K . But if ξ' is an algebraic interior point of K , and ξ any point of K , then $\lambda\xi + (1 - \lambda)\xi'$ is an algebraic interior point of K for $0 \leq \lambda < 1$. Thus, $\lambda\tilde{\phi}(\xi) + (1 - \lambda)\tilde{\phi}(\xi') \geq 0$ for $0 \leq \lambda < 1$, so $\tilde{\phi}(\xi) \geq 0$.

For simplicity, we have limited the above discussion to real vector spaces. For complex vector spaces, the analytic form of the Hahn Banach Theorem remains valid as stated; the separation theorems have only to be changed by replacing "linear functional" by "real part of a complex linear functional" and "hyperplane" by "set of the form: $\{\xi: \operatorname{Re}\{\phi(\xi)\} = \lambda\}$, where ϕ is a non-zero complex-linear functional."

We can now describe a procedure for constructing new topologies for locally convex spaces. Let E be such a space, ϕ a continuous linear functional on E . Then $\xi \mapsto |\phi(\xi)|$ is easily verified to be a semi-norm on E . We consider the topology defined by all such semi-norms. In other words, we consider the topology such that $\xi_n \rightarrow \xi$ if and only if $\phi(\xi_n) \rightarrow \phi(\xi)$ for all continuous linear functionals ϕ . This topology is clearly weaker (i.e., not stronger) than the initial topology. (It may of course coincide with the initial topology.) It is, however, not too much weaker; every linear functional continuous for the initial topology is continuous for this new topology. In fact, the new topology can be uniquely described as the weakest topology with the same continuous linear functionals as the initial topology. It is called the *weak* (or, more properly, *weakened*) topology of E . The corollary to the geometric form of the Hahn-Banach Theorem ensures that the weak topology associated with a Hausdorff topology is again Hausdorff. One of the main reasons why weak topologies are useful is that they frequently have many more compact sets than the initial topology. For example, the unit ball of a Hilbert space is always compact for the weak topology, but is not compact

for the norm topology unless the space is
finite dimensional

As a prelude to Choquet theory, we will show how to derive the Krein-Milman theorem from the Hahn-Banach Theorem. The Krein-Milman Theorem says that a compact convex set is generated by its extremal points. A point of a convex set K is said to be an *extremal point* if it is not an internal point of any line segment contained in K , i.e., if it cannot be written as $\alpha\xi_1 + (1 - \alpha)\xi_2$ for $\xi_1 \neq \xi_2 \in K$ and $0 < \alpha < 1$. (A little thought shows that ξ is an extremal point of K if it cannot be written as $\frac{1}{2}(\xi_1 + \xi_2)$ with $\xi_1, \xi_2 \in K$ and $\xi_1 \neq \xi_2$.) The extremal points of a triangle (or of any convex polygon) are the corners; the extremal points of the circle $\{(x, y) : x^2 + y^2 \leq 1\}$ are all the boundary points. We let $\epsilon(K)$ denote the set of extremal points of the convex set K .

THEOREM I. C. 8 (Krein-Milman Theorem) *Any convex compact set in a locally convex topological vector space is the closed convex hull of its set of extremal points.*

In other words: We start with a convex compact set K . We form the set of its extremal points. We then form the set of all convex combinations of these extremal points, and finally take the closure of this set. This gives us a compact convex set contained in K ; the Krein-Milman Theorem asserts that it is all of K . Note that it is not at all obvious that a convex compact set has any extremal points at all; indeed, the proof of the Krein-Milman Theorem reduces essentially to proving that extremal points do exist.

The proof of the Krein-Milman Theorem uses as a technical device the notion of a support of a compact convex set. Let K be a compact convex set in a locally convex topological vector space, and let $A \subset K$. A is said to be a *support* of K if A is non-empty, convex, and closed (hence compact),

and if, whenever $\xi \in A$ can be written as $\frac{\xi_1 + \xi_2}{2}$, with $\xi_1, \xi_2 \in K$, then

ξ_1 and ξ_2 both belong to A . For example, if ϕ is a continuous linear functional on E , and if $\lambda = \sup \{\phi(\xi) : \xi \in K\}$, then $\{\xi \in K : \phi(\xi) = \lambda\}$ is a support of K . A set consisting of a single point is a support of K if and only if it is an extremal point of K . The key technical lemma in the proof of the Krein-Milman Theorem is the following:

LEMMA I. C. 9 *Every support of K contains an extremal point of K .*

Using this lemma, we can easily prove the Krein-Milman Theorem. Let K be a compact convex set, and let K_1 be the closed convex hull of the set of extremal points of K . Then K_1 is a compact convex set contained in K ; we want to show that it is all of K . Suppose not; let $\xi \in K \setminus K_1$. By the geometric form of the Hahn-Banach Theorem, there is a continuous linear functional ϕ on E such that $\phi(\xi) > \sup \{\phi(\eta) : \eta \in K_1\}$. Let $\lambda = \sup \{\phi(\zeta) : \zeta \in K\}$. Then $\{\zeta \in K : \phi(\zeta) = \lambda\}$ is a support of K ; hence, contains an extremal point of K . But by assumption K_1 contains all the extremal points of K , and

$\lambda > \sup \{ \phi(n) \mid n \in K \}$ This is a contradiction and prove

It remains to prove the lemma. This is an elementary exercise in the use of Zorn's Lemma.

ZORN'S LEMMA Let A be a partially ordered set. Suppose that every linearly ordered subset B of A has an upper bound (B is linearly ordered if, for all $\alpha, \beta \in B$, either $\alpha \geq \beta$ or $\beta \geq \alpha$). An upper bound for a subset B is an element γ of A such that $\gamma \geq \beta$ for all $\beta \in B$). Then if α is any element of A , there is a maximal element γ of A with $\gamma \geq \alpha$ (an element $\gamma \in A$ is maximal if $\gamma' \geq \gamma$ implies $\gamma' = \gamma$).

We consider the set of supports of K , ordered by $A \geq B$ if $A \subset B$. (Note that, somewhat confusingly, supports are "large" in the sense of the ordering if they are "small" sets, and that a maximal element of the partially ordered set is a minimal support, i.e., a support which contains no strictly smaller support.) We will use Zorn's Lemma to prove that every support contains a minimal support, and then argue that a minimal support must reduce to a single point. To prove that every support contains a minimal support, it suffices to show that any linearly ordered family of supports has an upper bound in the ordering considered. We do this by arguing that the intersection of a linearly ordered family of supports is again a support. Let $\{A_\alpha\}$ be such a family. Since the A_α 's are all compact, and since the intersection of any finite set of A_α 's contains some A_α ; hence, is non-empty, it follows that $\bigcap A_\alpha$ is non-empty. Since $\bigcap A_\alpha$ is the intersection of a family of closed sets, it is closed; hence, compact. Finally, if $\xi_1, \xi_2 \in K$, and if $(\frac{1}{2})\xi_1 + (\frac{1}{2})\xi_2 \in \bigcap A_\alpha$, then $(\frac{1}{2})\xi_1 + (\frac{1}{2})\xi_2 \in A_\alpha$ for all α , so ξ_1 and ξ_2 are in A_α for all α , so ξ_1 and ξ_2 are in $\bigcap A_\alpha$. Thus, $\bigcap A_\alpha$ is a support; it is evidently an upper bound for $\{A_\alpha\}$ in the ordering we have considered. Thus, applying Zorn's Lemma, every support contains a support which contains no strictly smaller support. Let B be such a minimal support. We want to show that B consists of a single point. Suppose not; then there is a continuous linear functional ϕ which is not constant on B (i.e., take two points of B and apply the geometric form of the Hahn-Banach Theorem to get a functional which separates them). Let $\lambda = \sup \{ \phi(\xi) : \xi \in B \}$ and let $B' = \{ \xi \in B : \phi(\xi) = \lambda \}$. Then $B' \subsetneq B$, and it is easy to check that B' is a support of K . This contradicts the assumed minimality of B and hence proves the lemma.

II Integral Representations on Compact Convex Sets

A Introduction

We will be concerned, in this chapter, with the problem of representing a general point of a compact convex set as a convex combination of extreme points.

general point of a compact convex set is a convex combination of extremal points. For orientation, let us consider convex polygons in the plane. It is

124

O. E. LANFORD III

then easy to see that the extremal points are just the corners of the polygon, and that, if ξ_1, \dots, ξ_n are the corners, then any point of the polygon may be written $\xi = \alpha_1 \xi_1 + \dots + \alpha_n \xi_n$, with $\alpha_1, \dots, \alpha_n \geq 0$ and $\sum_{i=1}^n \alpha_i = 1$. Indeed, since any point of a convex polygon is contained in at least one triangle with vertices at the corners of the polygon, we may choose the α 's so that only three of them are non-zero. Finally, if the polygon is a triangle, then the α 's are unique, and if the polygon is not a triangle, then at least some of the points of the polygon have more than one representation as convex combinations of corners.

The above remarks generalize easily to compact convex sets in finite dimensional spaces. A classical theorem, due to Minkowski, states that if X is a compact convex set in a p -dimensional vector space, and if $\xi \in X$, then ξ may be written as a convex combination of some set of $p + 1$ extremal points of X .

The generalization to infinitely many dimensions is not quite so simple. We are at least assured by the Krein-Milman Theorem that any compact convex set X in a locally convex topological vector space has enough extremal points so that every point in X can be approximated arbitrarily closely by convex combinations of extremal points. On the other hand, we should not be too surprised if, in passing to the infinite dimensional case, we had to consider integrals of extremal points rather than ordinary convex combinations. The concept needed is that of the *resultant*, or *barycenter*, of a probability measure (measure with total mass one) on a compact convex set X . The resultant of μ , denoted by $r(\mu)$, is just the integral $r(\mu) = \int \xi d\mu(\xi)$, where the integral is to be understood in the weak sense, i.e., $\int \xi d\mu(\xi)$ is an element of E such that $\phi\left(\int \xi d\mu(\xi)\right) = \int \phi(\xi) d\mu(\xi)$ for all continuous linear functionals ϕ on E . While it is easy to see (using the Hahn-Banach Theorem) that this condition uniquely specifies $\int \xi d\mu(\xi)$ if the integral exists, the existence is less obvious. We will prove it shortly. The equation $\xi = \sum_{i=1}^n \alpha_i \xi_i$ can be transcribed to $\xi = r\left(\sum_{i=1}^n \alpha_i \delta_{\xi_i}\right)$, where δ_{ξ_i} is the unit point-mass at ξ_i , i.e., the Dirac measure at ξ_i . We now want to investigate generalizations of Minkowski's Theorem asserting that every point of a compact convex set X may be represented as the resultant of a probability measure concentrated on the set of extremal points of X . From another point of view, the results we will discuss are a more precise version of the Krein-Milman Theorem: The Krein-Milman Theorem asserts that every point of X may be approximated by resultants of measures with finite support concentrated on the set of extremal points of X ; the results we will discuss say that every point of X is equal to the resultant of a measure (not necessarily with finite support) concentrated on the set of extremal points of X .

The prototype of such results in the following:

THEOREM (Choquet) *Let X be a compact convex set in a Hausdorff locally convex topological vector space. Let ϵ denote the set of extremal points of X . Assume that X is metrizable, i.e., that its topology has a countable base. Then:*

- i) ϵ is a Borel subset of X and
- ii) Every point ξ of X is the resultant of a Borel probability measure μ on X which is concentrated on ϵ in the sense that $\mu(X \setminus \epsilon) = 0$.

This theorem has been refined in two directions. In the first place, the requirement that X be metrizable can almost be eliminated. The difficulty in doing this lies in the fact that the set of extremal points of X may not be a Borel set, i.e., may be pathological from the point of view of measure theory, and the sense in which the measure μ should be concentrated of the set of extremal points is therefore somewhat complicated. In the second place, an algebraic condition on X is given which is necessary and sufficient for every point of X to be the resultant of a unique measure concentrated on the set of extremal points of X .

The following notational conventions will be used throughout this chapter.

- 1) All vector spaces will be vector spaces over the real numbers, and all numerical functions will be real-valued.
- 2) X will denote a compact convex set in a locally convex topological vector space E .
- 3) ϵ will denote the set of extremal points of X .
- 4) $C(X)$ will denote the set of continuous, real valued functions on X .
- 5) S will denote the set of continuous convex functions on X ; $S = \{f \in C(X) : f(\alpha\xi + (1 - \alpha)\eta) \leq \alpha f(\xi) + (1 - \alpha)f(\eta) \text{ for all } \xi, \eta \in X, 0 \leq \alpha \leq 1\}$.
- 6) $A = S \cap (-S)$ will denote the set of continuous affine functions on X . If ϕ is a continuous linear functional on E , and if α is a real number, then $\phi(\cdot) + \alpha$ is an element of A (but not every element of A is necessarily of this form).
- 7) M^+ will denote the set of positive Borel measures on X , and M_1 the set of Borel probability measures on X . M^+ is contained in the dual of $C(X)$; we equip it with the weak-* topology, thus making M_1 compact.

We now have to prove a few preliminary results.

PROPOSITION II. A.1 *Let $\mu \in M^+$. Then there exists exactly one element $r(\mu)$ of E such that*

$$\phi(r(\mu)) = \int \phi(\xi) d\mu(\xi)$$

for all continuous linear functionals ϕ on E . If $\mu \in M_1$, $r(\mu) \in X$.

Proof We first prove uniqueness: If

$$\phi(\xi_1) = \int \phi(\xi) d\mu(\xi) \quad \text{and} \quad \phi(\xi_2) = \int \phi(\xi) d\mu(\xi)$$

for all continuous linear functionals ϕ on E , then

$$\phi(\xi_1 - \xi_2) = 0$$

for all continuous linear functionals ϕ on E , so $\xi_1 - \xi_2 = 0$. Also, if $\mu \in M_1$, and if $\phi(\xi) \leq \alpha$ for all $\xi \in X$, then

$$\int \phi(\xi) d\mu(\xi) \leq \alpha;$$

by the geometric form of the Hahn-Banach Theorem, this implies $r(\mu) \in X$. It remains to prove the existence of $r(\mu)$. It suffices to consider $\mu \in M_1$.

If μ is a measure with finite support $\left(\mu = \sum_{i=1}^n \alpha_i \delta_{\xi_i}\right)$, then $r(\mu)$ exists and is equal to $\sum \alpha_i \xi_i$. Now let μ be a general element of M_1 . Then there exists

a net $\mu^{(\alpha)}$ of measures with finite support converging in the weak-* topology to μ . (This is just the approximability of integrals of continuous functions by Riemann sums.) Each $r(\mu^{(\alpha)})$ exists and belongs to X . Since X is compact, we can suppose (passing to a subnet if necessary) that the net $r(\mu^{(\alpha)})$ converges. Then for all continuous linear functionals ϕ on E

$$\begin{aligned} \phi\left(\lim_{\alpha} r(\mu^{(\alpha)})\right) &= \lim_{\alpha} \phi(r(\mu^{(\alpha)})) = \lim_{\alpha} \int \phi(\xi) d\mu^{(\alpha)}(\xi) \\ &= \int \phi(\xi) d\mu(\xi) \quad (\text{by the definition of weak-* convergence}), \end{aligned}$$

so $\phi\left(\lim_{\alpha} r(\mu^{(\alpha)})\right) = \int \phi(\xi) d\mu(\xi)$, so we can take $\lim_{\alpha} r(\mu^{(\alpha)}) = r(\mu)$.

PROPOSITION II. A. 2 $S - S$ is dense in $C(X)$, i.e., any continuous function on X may be approximated uniformly by differences of continuous convex functions.

Proof By the Hahn-Banach Theorem, the continuous affine functions A on X separate points. Therefore, by the Stone-Weierstrass Theorem, polynomials in elements of A are dense in $C(X)$. Thus, it will suffice to prove that, if $P(Z_1, \dots, Z_n)$ is a polynomial in n variables, and if $f_1, \dots, f_n \in A$, then $\xi \mapsto P(f_1(\xi), \dots, f_n(\xi))$ is the difference of two convex functions. Let $Y = \{(f_1(\xi), \dots, f_n(\xi)) : \xi \in X\} \subset \mathbb{R}^n$. Then Y is convex and compact. If we can show

$$P(Z_1, \dots, Z_n) = P_1(Z_1, \dots, Z_n) - P_2(Z_1, \dots, Z_n),$$

where P_1 and P_2 are polynomials convex on Y , we will be done. Since a polynomial $Q(Z_1, \dots, Z_n)$ is convex on Y if the matrix $\frac{\partial^2 Q}{\partial Z_i \partial Z_j}$

is positive semi-definite on Y , we get a decomposition of the desired form by writing

$$P(Z_1, \dots, Z_n) = (P(Z_1, \dots, Z_n) + \lambda(Z_1^2 + \dots + Z_n^2)) - \lambda(Z_1^2 + \dots + Z_n^2),$$

with λ sufficiently large.

PROPOSITION II. A. 3 Any $\mu \in M_1$ can be approximated arbitrarily well in the weak-* topology by measures with finite support having the same resultant as μ .

Proof Let $f_1, \dots, f_m \in C(X)$ and let $\varepsilon > 0$. We have to find $\lambda_1, \dots, \lambda_m \geq 0$, with $\sum_{i=1}^m \lambda_i = 1$, and $\xi_1, \dots, \xi_m \in X$, such that $\sum_{i=1}^m \lambda_i \xi_i = r(\mu)$ and

$$\left| \mu(f_j) - \sum_{i=1}^m \lambda_i f_j(\xi_i) \right| \leq \varepsilon \quad \text{for } 1 \leq j \leq m.$$

Since X is compact, there exists a finite set U_1, \dots, U_n of open convex sets in X , with $\bigcup_{i=1}^n U_i = X$, such that each f_j varies by less than ε on each U_i . Next, we can write $\mu = \sum_{i=1}^n \lambda_i \mu_i$, where each $\mu_i \in M_1$ and has support in U_i . (For example, if ϕ_i is the characteristic function of U_i , we can take

$$\lambda_i \mu_i = \frac{\phi_i \mu}{\sum_{i=1}^n \phi_i}.)$$

Let $\xi_i = r(\mu_i)$; then evidently $r(\mu) = \sum_i \lambda_i \xi_i$. Since U_i is compact and convex, $r(\mu_i) \in U_i$, so

$$\left| \int d\mu_i f_j - f_j(\xi_i) \right| \leq \varepsilon \quad \text{for all } i, j.$$

Therefore, for $1 \leq j \leq m$,

$$\left| \int d\mu f_j - \sum_{i=1}^n \lambda_i f_j(\xi_i) \right| \leq \sum_{i=1}^n \lambda_i \left| \int d\mu_i f_j - f_j(\xi_i) \right| \leq \varepsilon.$$

B The Existence Theorem

We will first outline the strategy for proving that every $\xi \in X$ is the resultant of a measure concentrated on the set of extremal points. Given ξ , we want to find a measure with resultant ξ which is pushed out as much as possible to the "corners" of X . One way to tell how much a measure μ is pushed toward the "corners" is to evaluate $\mu(f)$ for convex f ; the larger $\mu(f)$ is for fixed convex f and fixed $r(\mu)$, the more we expect μ to be concentrated near the extremal points of X .

With this in mind, we define an order on M^+ by $\mu \succ \nu$ if $\mu(f) \geq \nu(f)$ for

all continuous convex f . It is evident that this defines a preorder, i.e. that

128

O. E. LANFORD III

\succ is reflexive and transitive. The relation \succ is also anti-symmetric and hence an order; if $\mu \succ \nu$ and $\nu \succ \mu$, then $\mu(f) = \nu(f)$ for all $f \in S$, therefore, for all $f \in S - S$; therefore, since $S - S$ is dense in $C(X)$, for all $f \in C(X)$.

Now let $\mu \succ \nu$ and let f be a continuous affine function. Since both f and $-f$ are convex, we have

$$\mu(f) \geq \nu(f) \quad \text{and} \quad \mu(-f) \geq \nu(-f), \quad \text{so} \quad \mu(f) = \nu(f).$$

In particular, $\|\mu\| = \mu(1) = \nu(1) = \|\nu\|$, and $r(\mu) = r(\nu)$, so two measures which are comparable in the sense of \succ have the same total mass and the same resultant. It is thus easy to see that $\mu \succ \delta_x$ if and only if μ is a probability measure and $r(\mu) = x$.

We now have good candidates for measures with resultant ξ which are concentrated on the set of extremal points of X ; they are the measures μ satisfying $\mu \succ \delta_x$ which are maximal in the sense of \succ . An easy argument using Zorn's Lemma shows that such maximal measures exist. It remains, however, to determine in what sense maximal measures are concentrated on the set of extremal points. For this purpose, we consider upper envelopes of functions: If f is a bounded function on X , we define the upper envelope of f , denoted \hat{f} , to be the smallest concave function which is everywhere greater than or equal to f or, more precisely, as the infimum of all continuous concave functions $\geq f$. (Recall that the infimum of any family of concave functions is concave.)

Unfortunately, \hat{f} need not be continuous even if f is. If we forget this fact for the moment, but remember that \hat{f} is concave, we see that, to make μ maximal, we want to make $\mu(\hat{f})$ as small as possible (keeping $r(\mu)$ fixed). On the other hand, $\hat{f} \geq f$, so $\mu(\hat{f}) \geq \mu(f)$ for all $\mu \in M^+$. Thus, we might guess that, in order to have μ maximal we should have $\mu(f) = \mu(\hat{f})$ and it turns out, in fact, that μ is maximal if and only if $\mu(f) = \mu(\hat{f})$ for all continuous convex functions f . Since $\hat{f} \geq f$, this means that μ is maximal if and only if

$$\mu(\{\xi : \hat{f}(\xi) > f(\xi)\}) = 0$$

for all continuous convex f , i.e., if and only if μ is carried by $\{\xi : \hat{f}(\xi) = f(\xi)\}$ for all continuous convex f . Finally, it turns out that the set of extremal points of X is precisely the set of points on which every continuous convex function is equal to its upper envelope, i.e.,

$$\varepsilon = \bigcap_{f \in S} \{\xi : \hat{f}(\xi) = f(\xi)\}.$$

The intersection is, in general, over a non-denumerable set of f 's, so we cannot conclude that $\mu(X \setminus \varepsilon) = 0$ for μ maximal, or even that $\mu(X \setminus \varepsilon)$ is defined. Thus, it is reasonable to say that a maximal measure is concentrated on ε , but the sense in which the measure is concentrated on ε is a bit subtle. If X is metrizable all these difficulties disappear: There is a single continuous convex function f such that $\varepsilon = \{\xi : \hat{f}(\xi) = f(\xi)\}$

so in this case every maximal measure is concentrated on ϵ in the sense that $\mu(X \setminus \epsilon) = 0$. Combining all these results gives the theorem of Choquet quoted in the introduction.

To give the proof of the existence theorem, then, we must prove four things:

- a) Every measure on X is majorized in the sense of \succ by a maximal measure.
- b) A measure μ is maximal if and only if $\mu(f) = \mu(\hat{f})$ for all continuous convex f .
- c) $\epsilon = \bigcap_{f \in S} \{\xi : \hat{f}(\xi) = f(\xi)\}$.
- d) If X is metrizable, there exists $f \in S$ such that

$$\epsilon = \{\xi : \hat{f}(\xi) = f(\xi)\}.$$

Points a), c), and d) are more or less routine; the subtle part of the argument comes in the proof of b).

PROPOSITION II. B. 1. *Every $\mu \in M^+$ is majorized, in the sense of the ordering \succ , by a maximal measure.*

Proof By Zorn's Lemma, it suffices to show that any family $(\mu_i)_{i \in I} \subset M^+$ which is totally ordered by \succ admits an upper bound. For $f \in S$, $(\mu_i(f))$ is monotonic, by the definition of \succ , and bounded since $\|\mu_i\|$ is independent of i . Therefore, $\mu_i(f)$ converges to something for all $f \in S$; hence, for all $f \in S - S$; hence, since $S - S$ is dense in $C(X)$, for all $f \in C(X)$. The limit defines a positive measure (i.e., a positive linear functional on $C(X)$) which evidently majorizes all the μ_i 's.

As we indicated above, we define the upper envelope \hat{f} of a bounded real-valued function f by $\hat{f}(\xi) = \inf \{g(\xi) : g \text{ concave and continuous, } g \geq f \text{ everywhere}\}$. In other words, we define \hat{f} to be the infimum of the set of all concave continuous functions $\geq f$. Since the infimum of an arbitrary family of concave functions is concave, \hat{f} is concave. Although \hat{f} need not be continuous, it is the infimum of a family of continuous functions and is therefore upper semi-continuous. The mapping $f \mapsto \hat{f}$ is:

i) increasing (If $f_1 \geq f_2$, then $\hat{f}_1 \geq \hat{f}_2$) and

ii) sublinear ($\widehat{\lambda f} = \lambda \cdot \hat{f}$ if $\lambda \geq 0$; $\widehat{f_1 + f_2} \leq \hat{f}_1 + \hat{f}_2$)

For any positive measure μ , we define $\hat{\mu}(f) = \mu(\hat{f})$. Then $\hat{\mu}$ is an increasing sublinear functional on $C(X)$, and $\hat{\mu} \geq \mu$, i.e., $\hat{\mu}(f) \geq \mu(f)$ for all $f \in C(X)$. Furthermore, the functional $\hat{\mu}$ is continuous. First to prove that $\hat{\mu}$ is continuous, note that

where $\|g\| = \sup_{t \in X} |g(t)|$. This is true since the constant function $\|g\|$ is concave and $\geq g$. Thus,

$$|\hat{\mu}(g)| = |\mu(\hat{g})| \leq \|g\| \mu(1).$$

The continuity of $\hat{\mu}$ at a general point of $C(X)$ follows easily from continuity at zero and subadditivity: For $f, g \in C(X)$,

$$\hat{\mu}(f + g) \leq \hat{\mu}(f) + \hat{\mu}(g),$$

$$\hat{\mu}(f) \leq \hat{\mu}(f + g) + \hat{\mu}(-g),$$

so

$$-\hat{\mu}(-g) \leq \hat{\mu}(f + g) - \hat{\mu}(f) \leq \hat{\mu}(g),$$

so

$$|\hat{\mu}(f + g) - \hat{\mu}(f)| \leq \|g\| \mu(1).$$

Recall now that we are trying to show that μ is maximal if and only if $\mu(f) = \hat{\mu}(f)$ for all convex continuous f , i.e., if and only if $\hat{\mu} = \mu$ on S .

PROPOSITION II. B. 2 If $\mu \in M^+$, the following are equivalent

- i) $\hat{\mu} = \mu$ on S ,
- ii) $\hat{\mu} = \mu$ on $C(X)$,
- iii) $\hat{\mu}$ is linear on $C(X)$.

Proof It is immediate that ii) implies iii) and that ii) implies i). To see that iii) implies ii), recall that $\hat{\mu}(f) \geq \mu(f)$ for all f in $C(X)$. But we can equally well apply this inequality with f replaced by $-f$; then, since both μ and $\hat{\mu}$ are linear, $-\hat{\mu}(f) = \hat{\mu}(-f) \geq \mu(-f) = -\mu(f)$.

So $\hat{\mu}(f) \leq \mu(f)$, so $\hat{\mu}(f) = \mu(f)$ for all f in $C(X)$. We have therefore only to prove that i) implies ii). Note first that, if $g \in S$, $\widehat{-g} = -g$ (since $-g$ is already concave), so $\hat{\mu}(-g) = \mu(-g)$.

Now assume i), and let $f, g \in S$. By subadditivity,

$$\hat{\mu}(f) \leq \hat{\mu}(f - g) + \hat{\mu}(+g) \quad \text{and} \quad \hat{\mu}(f - g) \leq \hat{\mu}(f) + \hat{\mu}(-g),$$

so

$$|\hat{\mu}(f) - \hat{\mu}(g)| \leq \hat{\mu}(f - g) \leq \hat{\mu}(f) + \hat{\mu}(-g).$$

Using

$$\hat{\mu}(f) = \mu(f); \quad \hat{\mu}(g) = \mu(g); \quad \hat{\mu}(-g) = \mu(-g);$$

and the linearity of μ , we get

$$\mu(f - g) \leq \hat{\mu}(f - g) \leq \mu(f - g),$$

so $\hat{\mu} = \mu$ on $S - S$. Since $S - S$ is dense in $C(X)$ and since μ and $\hat{\mu}$ are both continuous, $\hat{\mu} = \mu$ on all of $C(X)$.

The key remark is the following lemma which will allow us to use the

LEMMA II. B. 3 Let $\mu \in M^+$ and let ν be a linear functional on $C(X)$. Then the following are equivalent:

i) $\nu(f) \leq \hat{\mu}(f)$ for all $f \in C(X)$.

ii) $\nu \in M^+$ and $\nu \succ \mu$.

Proof Let us assume i) and prove ii). Note that, since $\hat{f} = f$ for f concave, $\nu(f) \leq \hat{\mu}(f)$ means $\nu(f) \leq \mu(f)$ for f concave, so $\nu(f) \leq \mu(f)$ for f convex, so $\nu \succ \mu$. It still has to be verified that, if $f \geq 0$, $\nu(f) \geq 0$. It is simpler to verify that $f \leq 0$ implies $\nu(f) \leq 0$; indeed, this is immediate since $f \leq 0$ implies $\hat{f} \leq 0$ (the constant function 0 is concave, continuous, and $\geq f$), so $\hat{\mu}(f) = \mu(\hat{f}) \leq 0$, so $\nu(f) \leq \hat{\mu}(f) \leq 0$.

Now assume ii). The fact we want to exploit is that $\nu(g) \leq \mu(g)$ for all concave continuous g . We will use this fact by remarking that $\hat{f} = \inf \{g : g \text{ concave, continuous, and } g \geq f\}$. The set of concave, continuous functions $g \geq f$ is a decreasing net of continuous functions converging pointwise to \hat{f} ; hence, by the generalized monotone convergence theorem for semi-continuous functions (Theorem I. B. 3);

$$\mu(\hat{f}) = \inf \{ \mu(g) : g \text{ concave, continuous, and } g \geq f \},$$

$$\nu(\hat{f}) = \inf \{ \nu(g) : g \text{ concave, continuous, and } g \geq f \}, \text{ so}$$

$$\nu(\hat{f}) \leq \nu(f) \leq \mu(\hat{f}) = \hat{\mu}(f) \text{ for all } f \text{ in } C(X).$$

We are now ready to prove:

THEOREM II. B. 4 Let $\mu \in M^+$. Then μ is maximal if and only if $\mu(f) = \mu(\hat{f})$ for all $f \in S$.

Proof Suppose first that $\mu(f) = \mu(\hat{f})$ for all $f \in S$. Then $\hat{\mu} = \mu$ on $C(X)$ by Proposition II. B. 2. If $\nu \succ \mu$, then $\nu(f) \leq \hat{\mu}(f) = \mu(f)$ for all $f \in C(X)$; since μ and ν are both linear, this implies $\nu(f) = \mu(f)$ for all $f \in C(X)$, i.e., $\mu = \nu$. Therefore, μ is maximal. (It is majorized only by itself.)

Now assume $\mu(g) \neq \mu(\hat{g})$ for some $g \in S$; we will prove that μ is not maximal by producing $\nu \succ \mu$ such that $\nu(g) = \mu(\hat{g}) > \mu(g)$. Thus, we have to construct a measure, i.e., a positive linear functional on $C(X)$, so we will use the Hahn-Banach Theorem. By Lemma II. B. 3, we have only to construct a linear functional ν such that $\nu(f) \leq \hat{\mu}(f)$ for all f . We start by defining $\nu(\lambda g) = \lambda \hat{\mu}(g)$; this defines ν on the one-dimensional subspace generated by g , and $\nu(\lambda g) \leq \hat{\mu}(\lambda g)$ since $\hat{\mu}(\lambda g) \geq \lambda \hat{\mu}(g)$ by the sublinearity of $\hat{\mu}$. Now, applying the generalization of the analytic form of the Hahn-Banach Theorem (Theorem I. C. 6), we extend ν to all of $C(X)$, preserving the relation $\nu(f) \leq \hat{\mu}(f)$ for all f . This is possible since $\hat{\mu}$ is sublinear. We are now finished; since $\nu(f) \leq \hat{\mu}(f)$ for all f , $\nu \succ \mu$ by Lemma II. B. 3, but $\nu \neq \mu$

Since $\nu(g) = \hat{\mu}(g) \neq \mu(g)$

132

O. E. LANFORD III

Now let f be a continuous convex function on X . We define $B(f) = \{\xi \in X: \hat{f}(\xi) = f(\xi)\}$, and any set of the form $B(f)$, $f \in \mathcal{S}$, is called a *boundary set* of X . Since \hat{f} is upper semi-continuous and f is continuous, $\hat{f} - f$ is upper semi-continuous so, for all α ,

$$\{\xi \in X: \hat{f}(\xi) - f(\xi) < \alpha\}$$

is open. Hence

$$B(f) = \bigcap_{n=1}^{\infty} \{\xi \in X: \hat{f}(\xi) - f(\xi) < 1/n\}$$

is a countable intersection of open sets and is therefore a Borel set. Since $\hat{f} - f \geq 0$, we have $\mu(\hat{f}) = \mu(f)$ if and only if

$$\mu(\{\xi \in X: \hat{f}(\xi) > f(\xi)\}) = 0.$$

If Y is a Borel subset of X , and if $\mu \in M^+$, we will say that μ is *carried by* Y if $\mu(X \setminus Y) = 0$. Thus we have:

PROPOSITION II. B. 5 *A measure $\mu \in M^+$ is maximal in the sense of the order $>$ if and only if it is carried by every boundary set of X .*

COROLLARY II. B. 6 *Every point $\xi \in X$ is the resultant of a probability measure μ on X carried by every boundary set of X .*

COROLLARY II. B. 7 *Any sum, or, more generally, any integral, of maximal measures is maximal. The greatest lower bound or least upper bound of a finite family of maximal measures is maximal.*

We now must investigate the relation between boundary sets and extremal points.

PROPOSITION II. B. 8 *The set of extremal points of X is the intersection of all the boundary sets of X .*

Proof We have to prove:

i) Any extremal point of X belongs to every boundary set.

ii) Every non-extremal point of X lies outside some boundary set.

i) Let ξ be an extremal point of X . Then the only probability measure with finite support having ξ as resultant is δ_ξ . But any probability measure having ξ as resultant can be approximated in the weak-* topology by measures with finite support having ξ as resultant (Proposition II. A. 3), so the only probability measure having ξ as resultant is δ_ξ . In other words, if $\mu > \delta_\xi$, then $\mu = \delta_\xi$. Hence, the measure δ_ξ is maximal, i.e., is carried by every boundary set of X , so ξ belongs to every boundary set of X .

ii) Let ξ be a non-extremal point of X . Then

$$\xi = 1/2 \xi_1 + 1/2 \xi_2, \text{ with } \xi_1 \neq \xi_2$$

Choose a continuous affine f such that $f(\xi_1) \neq f(\xi_2)$. Then f^2 is convex, and

$$\begin{aligned} & (f((1/2)\xi_1 + (1/2)\xi_2))^2 \\ &= (1/4)(f(\xi_1))^2 + (1/4)(f(\xi_2))^2 + (1/2)f(\xi_1)f(\xi_2) \\ &= (1/2)(f(\xi_1))^2 + (1/2)(f(\xi_2))^2 - (1/4)(f(\xi_1) - f(\xi_2))^2. \end{aligned}$$

Since $\widehat{f^2}$ is concave,

$$\begin{aligned} \widehat{f^2}(\xi) &= \widehat{f^2}((1/2)\xi_1 + (1/2)\xi_2) \geq (1/2)\widehat{f^2}(\xi_1) + (1/2)\widehat{f^2}(\xi_2) \\ &\geq (1/2)f^2(\xi_1) + (1/2)f^2(\xi_2) > f^2(\xi); \end{aligned}$$

$$\text{so } \widehat{f^2}(\xi) > f^2(\xi), \text{ so } \xi \notin B(f^2).$$

We can prove a somewhat more precise result if X is metrizable.

PROPOSITION II. B. 9 *If X is metrizable, then ϵ is a boundary set of X .*

Proof We will use only the following consequence of the metrizability of X : There exists a sequence f_n of continuous affine functions separating the points of X (This means that, if $\xi_1 \neq \xi_2$, then $f_n(\xi_1) \neq f_n(\xi_2)$ for some n). It is left as an exercise in general topology to verify that the existence of such a sequence is equivalent to the metrizability of X . By scaling the f_n 's, we can assume $|f_n| \leq 1/n$ for each n . Since each f_n is affine, f_n^2 is convex for each n . Let $f = \sum_{n=1}^{\infty} f_n^2$;

then f is continuous and convex. We claim that f is strictly convex, i.e., that if

$$\xi_1 \neq \xi_2, \text{ then } f((1/2)\xi_1 + (1/2)\xi_2) < (1/2)f(\xi_1) + (1/2)f(\xi_2).$$

Since

$$\widehat{f}((1/2)\xi_1 + (1/2)\xi_2) \geq (1/2)\widehat{f}(\xi_1) + (1/2)\widehat{f}(\xi_2) \geq (1/2)f(\xi_1) + (1/2)f(\xi_2),$$

this will imply

$$\widehat{f}((1/2)\xi_1 + (1/2)\xi_2) > f((1/2)\xi_1 + (1/2)\xi_2)$$

for all $\xi_1 \neq \xi_2$, i.e., that $\widehat{f}(\xi) > f(\xi)$ if ξ is not an extremal point of X . Thus, $B(f) \subset \epsilon$, and since we know that, in general, $B(f) \supset \epsilon$ for all continuous convex f , we must have $B(f) = \epsilon$.

It remains to show that f is strictly convex. Let $\xi_1 \neq \xi_2$, and choose n so that $f_n(\xi_1) \neq f_n(\xi_2)$. By the calculation in the proof of the preceding proposition,

$$f_n^2((1/2)\xi_1 + (1/2)\xi_2) < (1/2)f_n^2(\xi_1) + (1/2)f_n^2(\xi_2).$$

Since $f = f_n^2 + \sum_{j \neq n} f_j^2$ and since $\sum_{j \neq n} f_j^2$ is convex, f is strictly convex.

Putting together the preceding results, we get the theorem quoted in the introduction:

THEOREM II. B. 10 *Let X be a metrizable compact convex set in a Hausdorff locally convex topological vector space. Then*

- i) ϵ is a Borel subset of X and
- ii) Each $\xi \in X$ is the resultant of a probability measure on X carried by ϵ .

If X is not metrizable, the situation is a little more cloudy. We have shown that every point of X is the resultant of a maximal measure, and that a measure is maximal if and only if it is carried by every boundary set of X . Since the intersection of all boundary sets is the set of extremal points of X , a maximal measure is in some weak sense concentrated on the set of extremal points of X . The sense in which a maximal measure is concentrated on the set of extremal points may be clarified a bit; we quote by way of illustration the following result, which we will not prove. (It is Lemma 18 of Choquet and Mayer⁽¹⁾.)

PROPOSITION II. B. 11. *Let $K \subset X$ be a countable union of compact sets and contain the set of extremal points of X . Then every maximal measure on X is carried by K .*

C The Uniqueness Theorem

We now turn to an investigation of the conditions under which each $\xi \in X$ is the resultant of a unique maximal measure. We have first to prove several preliminary results. The first group of results concerns the upper envelope of a function f and provides some new ways to compute it.

LEMMA II. C. 1 *Let f be a bounded real-valued function on X . Then*

$$\Gamma = \{(\xi, t) \in E \times \mathbb{R} : \xi \in X, t \leq \hat{f}(\xi)\}$$

is the closed convex hull of

$$\{(\xi, t) \in E \times \mathbb{R} : \xi \in X; t \leq f(\xi)\}.$$

(i.e., the region below the graph of \hat{f} is the closed convex hull of the region below the graph of f .)

Proof It follows easily from the concavity of \hat{f} that Γ is convex and from the upper semi-continuity of \hat{f} that Γ is closed. Thus, Γ is a closed convex set containing

$$\{(\xi, t) \in E \times \mathbb{R} : \xi \in X; t \leq f(\xi)\}.$$

If we let Γ' denote the closed convex hull of the latter set, then $\Gamma' \subset \Gamma$, and what we have to prove is that $\Gamma' = \Gamma$. Thus, assume there is a $(\xi_0, t_0) \in \Gamma \setminus \Gamma'$

~~By the geometric form of the Borel Theorem...~~

hyperplane in $E \times \mathbb{R}$ strictly separating (ξ_0, t_0) from Γ . In other words, there is a linear functional ϕ on $E \times \mathbb{R}$ and a constant c such that

$$\phi(\xi_0, t_0) < c; \quad \phi(\xi, t) \geq c \quad \text{for } t \leq f(\xi).$$

We claim that the hyperplane $\phi(\xi, t) = c$ is the graph of a continuous affine functional on X , which is greater than f everywhere, but less than t_0 at ξ_0 , thus contradicting the assumption that $t_0 \leq \hat{f}(\xi_0)$. To see this, note that $\phi(\xi, t)$ must have the form $\phi(\xi) - \alpha t$, where ϕ is a continuous linear functional on E and $\alpha \in \mathbb{R}$. Since

$$\phi(\xi_0) - \alpha t_0 < c; \quad \phi(\xi_0) - \alpha t > c \quad \text{for } t \leq f(\xi_0),$$

we must have $\alpha > 0$. Letting

$$\psi(\xi) = \frac{\phi(\xi) - c}{\alpha}$$

we have $\psi(\xi_0) < t_0$ but $\psi(\xi) > t$ for all $t \leq f(\xi)$. Thus, ψ is a continuous affine function everywhere greater than f ; it is therefore impossible to have $\psi(\xi_0) < t_0 \leq \hat{f}(\xi_0)$, so we have a contradiction and $\Gamma' = \Gamma$ as desired.

In the course of the proof, we have also proved the following:

LEMMA II. C. 2 If h is any concave upper semi-continuous function on X , then $h = \inf \{g \in A : g \geq h\}$.

LEMMA II. C. 3 Let $f \in C(X)$ and let $\xi \in X$. Then

$$\begin{aligned} \hat{f}(\xi) &= \sup \{ \mu(f) : \mu \in M_1; r(\mu) = \xi \} \\ &= \sup \{ \mu(f) : \mu \in M_1; r(\mu) = \xi, \mu \text{ has finite support} \}. \end{aligned}$$

Proof Since \hat{f} is concave,

$$\hat{f}(\xi) \geq \sum_{i=1}^n \lambda_i \hat{f}(\xi_i) \geq \sum_{i=1}^n \lambda_i f(\xi_i) \quad \text{if } \sum_{i=1}^n \lambda_i \xi_i = \xi.$$

Hence, $\hat{f}(\xi)$ is not smaller than the second supremum. On the other hand, Lemma II. C. 1, translated into the language of measures with finite supports, says that there exists a net of measures with finite supports

$$\mu^{(a)} = \sum_{i=1}^{n^{(a)}} \lambda_i^{(a)} \delta_{\xi_i^{(a)}}$$

with $r(\mu^{(a)}) \rightarrow \xi$ and $\liminf \mu^{(a)}(f) \geq \hat{f}(\xi)$. If μ is any limit point of this net, then $r(\mu) = \xi$ and $\mu(f) \geq \hat{f}(\xi)$, so the first supremum is not less than $\hat{f}(\xi)$. But since any measure with resultant ξ may be approximated arbitrarily well in the weak-* topology by measures with finite supports and resultant μ (Proposition II. A. 3) the two suprema are equal and so $\hat{f}(\xi)$ is equal to ~~the second~~ *the first* ~~value~~ *value*.

PROPOSITION II. C. 4 Let f be an affine upper semi-continuous function on X , and let $\mu \in M_1$. Then

$$\int f(\xi) d\mu(\xi) = f(r(\mu)).$$

Proof In more picturesque form, we want to show that, if f is affine and upper semi-continuous, then

$$\int f(\xi) d\mu(\xi) = f\left(\int \xi d\mu(\xi)\right).$$

This is immediate if f is continuous, or if μ has finite support. We get the general case by a two-fold approximation. We first approximate μ by measures with resultant $r(\mu)$ and finite support. If ν is such a measure, and if g is continuous, concave and nowhere less than f , we have

$$f(r(\mu)) = \nu(f) \leq \nu(g) \leq g(r(\mu)).$$

Since μ is a weak limit of measures with finite support and resultant $r(\mu)$, we have

$$f(r(\mu)) \leq \mu(g) \leq g(r(\mu)) \quad (*)$$

for any continuous concave function $g \geq f$. Now f is concave and upper semi-continuous, so by Lemma II. C. 2, $f = \inf \{g : g \text{ concave and continuous, } g \geq f\}$. Thus, $\{g : g \text{ concave and continuous, } g \geq f\}$ is a decreasing family of continuous functions with infimum f ; by the generalized Monotone Convergence Theorem for semi-continuous functions (Theorem I. B. 3),

$$\int f(\xi) d\mu(\xi) = \inf \{\mu(g) : g \text{ concave and continuous; } g \geq f\}.$$

By (*), then,

$$f(r(\mu)) \leq \int f(\xi) d\mu(\xi) \leq f(r(\mu)),$$

so the proposition is proved.

We are now ready to formulate a condition on X which is equivalent to the statement that every point of X is majorized by a unique maximal measure. To do this, it is convenient to suppose that X is the base of a convex cone K . By this we mean that there is a closed hyperplane in E which intersects K in X , which does not pass through 0, and which intersects each ray in K exactly once. We can always arrange this by replacing E by $E \times \mathbb{R}$, identifying X with $X \times \{1\}$, and taking for K the cone generated by $X \times \{1\}$. In general, there may be many ways of realizing X as the base of a cone, but any two cones obtained in this way are linearly isomorphic, since any point of such a cone is uniquely specified as $\lambda\xi$, where λ is a positive real number and $\xi \in X$. Thus, the algebraic properties of the cone K depend only on the algebraic properties of X , and not on how X is imbedded in a topological vector space. It is convenient to extend every numerical function on X to a positively homogeneous function on the cone K denoted by $\phi \circ$

same symbol. Thus, for example, the condition that f be convex (on X) translates into the condition

$$f(\xi + \eta) \leq f(\xi) + f(\eta) \quad (\text{on } K).$$

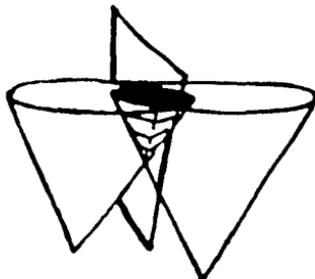
By definition, the *natural ordering* of a cone K in a vector space is the order in which $\xi \geq \eta$ means $\xi - \eta \in K$. Also, a partially ordered set Y is a *lattice* if, for any pair ξ, η of elements of Y , there exists a least upper bound $\xi \vee \eta$ and a greatest lower bound $\xi \wedge \eta$ for ξ and η in Y . By definition, $\xi \vee \eta$ is an element of Y such that

$$\xi \vee \eta \geq \xi, \quad \xi \vee \eta \geq \eta, \quad \text{and } \zeta \geq \xi \text{ and } \zeta \geq \eta \text{ implies } \zeta \geq \xi \vee \eta.$$

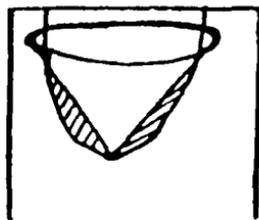
The element $\xi \vee \eta$ is uniquely determined if it exists. The definition of greatest lower bound is obtained by replacing " \geq " by " \leq " in the definition of least upper bound.

We now say that a compact convex set X is a *simplex* if it is the base of a cone which is a lattice with its natural ordering.

It does not seem to be altogether evident that this definition reduces to the usual one when X is finite-dimensional. To clarify the situation a little, let us see why a circle is not a simplex. Let ξ and η belong to the cone generated by a circle, and suppose that they are not comparable in the ordering defined by the cone. We will show that no least upper bound of ξ and η exists.



$\zeta \geq \xi$ and $\zeta \geq \eta$ if and only if ζ belongs to the intersections of the two cones with vertices ξ and η . A section of the figure by a vertical plane, perpendicular to the plane containing the axes of the cones, and passing through the lowest intersection point, is a hyperbola.



The only possible candidate for the least upper bound of ξ and η is the vertex of this hyperbola. A cone drawn at this vertex, however, does not contain all of the hyperbola, so there are points which are greater than ξ and greater than η , but not greater than the vertex of the hyperbola. Therefore, ξ and η have no least upper bound.

On the other hand, the usual n -dimensional simplex $\{(\lambda_1, \dots, \lambda_{n+1}) : \sum_{i=1}^n \lambda_i = 1\}$ is a base of the cone $\{(\lambda_1, \dots, \lambda_{n+1}) : \lambda_i \geq 0 \text{ for all } i\}$ which is certainly a lattice with its natural ordering. The real justification for the definition, however, is contained in the following theorem:

THEOREM II. C. 5 *Let X be a compact convex set. Then the following are equivalent:*

- i) X is a simplex
- ii) For any $f \in S$, f is affine
- iii) Every $\xi \in X$ is the resultant of a unique maximal measure on X .

Proof We will show that ii) \Rightarrow iii) \Rightarrow i) \Rightarrow ii). Everything except i) \Rightarrow ii) is easy.

ii) \Rightarrow iii). Let ii) hold, let μ_1, μ_2 be two maximal measures on X both with resultant ξ , and let $f \in S$. Since μ_1 and μ_2 are maximal,

$$\mu_1(f) = \mu_1(f^\wedge); \quad \mu_2(f) = \mu_2(f^\wedge).$$

By hypothesis, f^\wedge is affine, and f^\wedge is always upper semi-continuous so by Proposition II. C. 3

$$\mu_1(f^\wedge) = f^\wedge(r(\mu_1)) = f^\wedge(\xi) = \mu_2(f^\wedge),$$

so

$$\mu_1(f) = \mu_1(f^\wedge) = \mu_2(f^\wedge) = \mu_2(f),$$

so

$$\mu_1 = \mu_2 \text{ on } S, \text{ so } \mu_1 = \mu_2 \text{ on } S-S$$

so $\mu_1 = \mu_2$ on $C(X)$ (since $S-S$ is dense in $C(X)$ by Proposition II. A. 2). Thus, there is only one maximal measure with resultant ξ . iii) \Rightarrow i). By Corollary II. B. 7, the cone of maximal measures is a lattice; the mapping $\mu \mapsto r(\mu)$ is a linear isomorphism from this cone to the cone generated by X .

i) \Rightarrow ii). Let f be convex. We want to show that f^\wedge is affine. Regarding f as defined and homogeneous on K , we have

$$f^\wedge(\xi + \eta) = \sup \left\{ \sum_{i=1}^n f(\zeta_i) : \zeta_i \in K; \sum_{i=1}^n \zeta_i = \xi + \eta \right\}$$

by Lemma III. C. 3. We will prove shortly that whenever we have $\xi + \eta = \zeta_1 + \dots + \zeta_n$, with the ζ_i in K , we can decompose

$$\zeta_i = \xi_i + \eta_i, \quad \text{where } \xi_i, \eta_i \in K; \quad \sum \xi_i = \xi; \quad \sum \eta_i = \eta.$$

Assuming this for the moment, we have:

$$\begin{aligned} f(\xi + \eta) &= \sup \left\{ \sum_i f(\zeta_i) : \zeta_i \in K; \sum_i \zeta_i = \xi + \eta \right\} \\ &= \sup \left\{ \sum_i f(\xi_i + \eta_i) : \xi_i, \eta_i \in K; \sum_i \xi_i = \xi; \sum_i \eta_i = \eta \right\}. \end{aligned}$$

(This is just the decomposition lemma stated above.)

$$\leq \sup \left\{ \sum_i f(\xi_i) : \sum_i \xi_i = \xi \right\} + \sup \left\{ \sum_i f(\eta_i) : \sum_i \eta_i = \eta \right\} = \hat{f}(\xi) + \hat{f}(\eta).$$

(Since f is convex.)

Thus,

$$\hat{f}(\xi + \eta) \leq \hat{f}(\xi) + \hat{f}(\eta).$$

On the other hand,

$$\hat{f}(\xi + \eta) \geq \hat{f}(\xi) + \hat{f}(\eta)$$

since \hat{f} is concave. Thus, for $\xi, \eta \in K$,

$$\hat{f}(\xi + \eta) = \hat{f}(\xi) + \hat{f}(\eta),$$

i.e., \hat{f} is affine.

We still have to prove the following lemma.

LEMMA II. C. 6 Let K be a cone which is a lattice with its natural order, let $\xi, \eta \in K$, and suppose $\xi + \eta = \zeta_1 + \dots + \zeta_n$ where $\zeta_1, \dots, \zeta_n \in K$. Then there exist $\xi_1, \dots, \xi_n, \eta_1, \dots, \eta_n \in K$ such that $\zeta_i = \xi_i + \eta_i$ for $1 \leq i \leq n$, and

$$\xi_1 + \dots + \xi_n = \xi; \quad \eta_1 + \dots + \eta_n = \eta.$$

Proof It is easy to prove by induction that the lemma with $n = 2$ implies the lemma for general n (Write $\xi + \eta = \zeta_1 + (\zeta_2 + \dots + \zeta_n)$; use the lemma with $n = 2$ to decompose

$$\zeta_1 = \xi_1 + \eta_1; \quad (\zeta_2 + \dots + \zeta_n) = \xi' + \eta',$$

where

$$\xi_1 + \xi' = \xi; \quad \eta_1 + \eta' = \eta.$$

Then iterate.) For $n = 2$, it is trivial to verify that

$$\begin{aligned} \xi_1 &= \zeta_1 \wedge \xi; & \eta_1 &= \zeta_1 - \zeta_1 \wedge \xi \\ \xi_2 &= \xi - \zeta_1 \wedge \xi; & \eta_2 &= \zeta_2 - \xi + \zeta_1 \wedge \xi \end{aligned}$$

satisfy all the conditions except, possibly, $\eta_2 \in K$. But we can rewrite:

$$\begin{aligned} \eta_2 &= (\zeta_2 - \xi) + (\zeta_1 \wedge \xi) \\ &= (\zeta_1 + (\zeta_2 - \xi)) \wedge (\xi + (\zeta_2 - \xi)) \\ &= \eta \wedge \xi_1 \in K \end{aligned}$$

In many applications to statistical mechanics, the following theorem which bypasses the above characterization of simplices can be used.

THEOREM II. C. 7 *Let X be a compact convex set. Suppose there exists an affine mapping $\xi \mapsto \mu_\xi$ from X to M_1 , such that $r(\mu_\xi) = \xi$ for all. Then each ξ on X is the resultant of a unique maximal measure, and μ_ξ is that maximal measure.*

Proof We will show that, if $r(\nu) = \xi$, then $\nu < \mu_\xi$, i.e., $\nu(f) \leq \mu_\xi(f)$ for all continuous convex f . Since ν may be approximated by measures with finite support and resultant ξ it is enough to show that $\nu(f) \leq \mu_\xi(f)$ for all ν with finite support and resultant ξ and all $f \in S$. But:

$$\nu(f) = \sum_i \lambda_i f(\xi_i) \leq \sum_i \lambda_i \mu_{\xi_i}(f)$$

$$\text{(Since } f \text{ is convex.)} \quad = \mu_{\sum_i \lambda_i \xi_i}(f)$$

$$\text{(Since } \xi \mapsto \mu_\xi \text{ is affine.)} \quad = \mu_\xi(f).$$

III C* Algebras

A Definitions and Algebraic Preliminaries

An *algebra* is a vector space equipped with a law of composition (multiplication), usually written $A \cdot B$ or AB , which is:

Associative: $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ for all $A, B, C \in \mathfrak{A}$, and

Distributive: $(\alpha A + \beta B) \cdot C = \alpha(A \cdot C) + \beta(B \cdot C)$ for all $A, B, C \in \mathfrak{A}$

$C \cdot (\alpha A + \beta B) = \alpha(C \cdot A) + \beta(C \cdot B)$ and all scalars α, β .

We will consider only algebras for which the field of scalars is the complex numbers \mathbb{C} . \mathfrak{A} is *commutative* if $A \cdot B = B \cdot A$ for all $A, B \in \mathfrak{A}$. An element $\mathbf{1}$ of \mathfrak{A} is an *identity* if $\mathbf{1} \cdot A = A \cdot \mathbf{1} = A$ for all $A \in \mathfrak{A}$. \mathfrak{A} has at most one identity element; if $\mathbf{1}, \mathbf{1}'$ are both identities, then $\mathbf{1} = \mathbf{1} \cdot \mathbf{1}' = \mathbf{1}'$.

A vector subspace \mathfrak{B} of \mathfrak{A} is said to be a *subalgebra* if $A \cdot B \in \mathfrak{B}$ for all $A, B \in \mathfrak{B}$,

a *left ideal* if $A \cdot B \in \mathfrak{B}$ for all $A \in \mathfrak{A}, B \in \mathfrak{B}$ and if $\mathfrak{B} \neq \mathfrak{A}$,

a *right ideal* if $B \cdot A \in \mathfrak{B}$ for all $A \in \mathfrak{A}, B \in \mathfrak{B}$ and if $\mathfrak{B} \neq \mathfrak{A}$,

a *two-sided ideal* if it is both a left-ideal and a right ideal.

If \mathfrak{C} is a two-sided ideal in \mathfrak{A} , then the quotient vector space $\mathfrak{A}/\mathfrak{C}$ may be made into an algebra by defining $\hat{A} \cdot \hat{B} = (\hat{A} \cdot B)$, where A is any element of \hat{A} and B any element of \hat{B} (Here, we are regarding elements of $\mathfrak{A}/\mathfrak{C}$ as subsets of \mathfrak{A} of the form $A + \mathfrak{C}$; such a subset is denoted by \hat{A}). The requirement

that \mathcal{C} be a two-sided ideal is just what is required to guarantee that the product be well-defined (i.e., not depend on the choice of representatives A, B of \bar{A}, \bar{B}).

An *involution* of an algebra \mathfrak{A} is a mapping $A \mapsto A^*$ of \mathfrak{A} into itself such that:

- 1) $*$ is conjugate linear: $(\alpha A + \beta B)^* = \bar{\alpha}A^* + \bar{\beta}B^*$,
- 2) reverses the order of products: $(A \cdot B)^* = B^* \cdot A^*$,
- 3) $(A^*)^* = A$.

If $\mathfrak{A}, \mathfrak{B}$ are algebras, a *morphism* $\phi: \mathfrak{A} \rightarrow \mathfrak{B}$ is a linear mapping from \mathfrak{A} to \mathfrak{B} such that

$$\phi(AB) = \phi(A)\phi(B)$$

for all A, B in \mathfrak{A} . If \mathfrak{A} and \mathfrak{B} are algebras with involution, we will assume that any morphisms we consider preserve the involution, i.e.,

$$\phi(A^*) = \phi(A)^*$$

for all A in \mathfrak{A} .

A *normed algebra* is an algebra \mathfrak{A} equipped with a norm $\| \cdot \|$ such that

$$\|A \cdot B\| \leq \|A\| \cdot \|B\|.$$

If the algebra \mathfrak{A} is complete in the norm, we speak of a *Banach algebra*. If \mathfrak{A} is equipped with an involution, and if $\|A^*\| = \|A\|$ for all $A \in \mathfrak{A}$, we speak of a *normed algebra with involution* or (if \mathfrak{A} is complete) *Banach algebra with involution*. If \mathfrak{A} has an identity, we will always assume $\|1\| = 1$. This is not really necessary, but it simplifies things, and, if it is not true, \mathfrak{A} can always be given an equivalent norm which makes it true.

We can now define: A *C^* algebra* is a Banach algebra with involution such that $\|A^*A\| = \|A\|^2$ for all $A \in \mathfrak{A}$. We shall see that this rather innocent-looking condition is very restrictive.

Examples

1. Let X be a locally compact space; consider the set $C_0(X)$ of continuous functions vanishing at ∞ on X . (A continuous function is said to vanish at ∞ if, for every $\epsilon > 0$, there is a compact set K such that $|f(x)| < \epsilon$ for $x \in X \setminus K$.) Define:

$$\|f\| = \sup_{x \in X} |f(x)|$$

$$(f + g)(x) = f(x) + g(x)$$

$$(xf)(x) = \alpha f(x)$$

$$(f \cdot g)(x) = f(x)g(x)$$

$$f^*(x) = \overline{f(x)}.$$

Then $C_0(X)$ is a commutative C^* algebra. The equality: $\|f^*f\| = \|f\|^2$ holds since $\|f^*f\| = \sup_{x \in X} |f(x)|^2 = \|f\|^2$. $C_0(X)$ has an identity if and

only if X is compact: We will see shortly that every commutative C^* algebra is isomorphic to $C_0(X)$ for some locally compact space X , and that, moreover, X is determined up to homeomorphism by the C^* algebra $C_0(X)$.

2. Let $D = \{z \in \mathbb{C} : |z| < 1\}$, and let $\mathcal{H}^\infty(D)$ denote the set of bounded analytic functions on D . Define addition and multiplication on $\mathcal{H}^\infty(D)$ pointwise (i.e., as for $C_0(X)$); define

$$\|f\| = \sup_{z \in D} |f(z)|$$

$$f^*(z) = \overline{f(\bar{z})}.$$

Then $\mathcal{H}^\infty(D)$ is a commutative Banach algebra with involution ($\mathcal{H}^\infty(D)$ is complete since a uniform limit of analytic functions is analytic), but is not a C^* algebra; for example, if $f(z) = e^z$, then $\|f\| = e$ but

$$f(z)f^*(z) = e^z e^{-\bar{z}} = 1, \quad \text{so} \quad \|f^*f\| = 1 \neq \|f\|^2.$$

3. Let \mathcal{H} be a Hilbert space, and let $\mathcal{L}(\mathcal{H})$ be the set of all bounded operators on \mathcal{H} . Define sums and products in the usual way, and let the norm be the usual operator norm:

$$\|A\| = \sup \{\|A\xi\| : \xi \in \mathcal{H}, \|\xi\| \leq 1\}.$$

Let $*$ be the ordinary adjoint operation, i.e.,

$$(\xi | A^*\eta) = (A\xi | \eta)$$

for all $\xi, \eta \in \mathcal{H}$. Then $\mathcal{L}(\mathcal{H})$ is a C^* algebra:

$$\begin{aligned} \|A^*A\| &= \sup \{\|A^*A\xi\| : \xi \in \mathcal{H}, \|\xi\| \leq 1\} \\ &= \sup \{(|\eta| A^*A\xi) : \xi, \eta \in \mathcal{H}, \|\xi\| \leq 1; \|\eta\| \leq 1\} \\ &= \sup \{|(A\xi | A\xi)| : \xi \in \mathcal{H}, \|\xi\| \leq 1\} = \|A\|^2. \end{aligned}$$

Evidently, any subalgebra of $\mathcal{L}(\mathcal{H})$ which is closed (and hence complete) in the operator norm, and which is self-adjoint (i.e., contains A^* if it contains A) is also a C^* algebra. It turns out that this example gives all C^* algebras, i.e., that every C^* algebra is isomorphic to a norm-closed self-adjoint algebra of bounded operators on a Hilbert space.

It is worth making a remark about terminology at this point: The term " C^* algebra" is sometimes (e.g., in Ruelle [1]) defined to mean a norm-closed self-adjoint algebra of operators on a Hilbert space, and the term " B^* algebra" is used for the more abstract object we have called C^* algebra. The distinction is not so very important since:

a) Every norm-closed self-adjoint algebra of operators on a Hilbert space is a C^* algebra.

b) Every C^* algebra is isomorphic to a norm-closed self-adjoint algebra of operators on a Hilbert space.

When we have occasion to distinguish between the two notions, we will refer to a norm-closed self-adjoint algebra of operators on a Hilbert space as a *concrete C^* algebra*.

The presence of an identity in an algebra is useful for many technical purposes. Therefore, if we have to deal with an algebra with no identity, it is useful to be able to imbed it in a larger algebra which has one. This we can do as follows: Let \mathfrak{A} be an algebra; and let $\tilde{\mathfrak{A}}$ be the set of pairs (λ, A) , with $\lambda \in \mathbb{C}$ and $A \in \mathfrak{A}$. Instead of writing pairs (λ, A) , we will use the more suggestive notation $\lambda \mathbf{1} + A$. Define:

$$(\lambda \mathbf{1} + A) + (\mu \mathbf{1} + B) = (\lambda + \mu) \mathbf{1} + (A + B)$$

$$(\lambda \mathbf{1} + A)(\mu \mathbf{1} + B) = (\lambda\mu) \mathbf{1} + (\mu A + \lambda B + AB).$$

It is easy to see that, with this structure, $\tilde{\mathfrak{A}}$ is an algebra with identity $\mathbf{1} \mathbf{1} + 0$, containing \mathfrak{A} (the set of elements of the form $0 \mathbf{1} + A$) as a two-sided ideal. (Note that the construction works perfectly well even if \mathfrak{A} has an identity, but that, in this case, the identity of \mathfrak{A} is no longer an identity for $\tilde{\mathfrak{A}}$.) If \mathfrak{A} has an involution, we may extend it to $\tilde{\mathfrak{A}}$ by defining $(\lambda \mathbf{1} + A)^* = \bar{\lambda} \mathbf{1} + A^*$. If \mathfrak{A} is a Banach algebra, or a C^* algebra, we would like to extend the norm of \mathfrak{A} to $\tilde{\mathfrak{A}}$ in such a way that $\tilde{\mathfrak{A}}$ becomes an object of the same sort. Defining

$$\|\lambda \mathbf{1} + A\| = |\lambda| + \|A\|$$

can easily be seen to make $\tilde{\mathfrak{A}}$ into a Banach algebra if \mathfrak{A} is a Banach algebra, but need not make $\tilde{\mathfrak{A}}$ into a C^* algebra if \mathfrak{A} is a C^* algebra. We will give a construction that does this, under the assumption that \mathfrak{A} does not already have an identity. (There is another construction which works in the other case; but we do not need it. See Dixmier C^*A , 1.3.8.) The algebra $\tilde{\mathfrak{A}}$ acts by left multiplication on \mathfrak{A} (since \mathfrak{A} is a left ideal in $\tilde{\mathfrak{A}}$); we will define $\|\lambda \mathbf{1} + A\|$ to be the norm of this multiplication operator (i.e., $\|\lambda \mathbf{1} + A\| = \sup \{\|\lambda B + AB\| : B \in \mathfrak{A}; \|B\| \leq 1\}$.) It is immediate from the above description that this expression defines a semi-norm, and that

$$\|(\lambda \mathbf{1} + A)(\mu \mathbf{1} + B)\| \leq \|(\lambda \mathbf{1} + A)\| \cdot \|(\mu \mathbf{1} + B)\|.$$

For any $A \in \mathfrak{A}$, $\|0 \mathbf{1} + A\| \leq \|A\|$, but also

$$\|(0 \mathbf{1} + A)\| \geq \frac{\|AA^*\|}{\|A^*\|} = \|A\|, \text{ so}$$

$\|0 \mathbf{1} + A\| = \|A\|$. We have still to check

1) $\|\lambda \mathbf{1} + A\| = 0$ implies $\lambda = 0$ and $A = 0$.

2) $\|\bar{\lambda} \mathbf{1} + A^*\| = \|\lambda \mathbf{1} + A\|$.

3) $\|(\lambda \mathbf{1} + A)(\bar{\lambda} \mathbf{1} + A^*)\| = \|(\lambda \mathbf{1} + A)\|^2$.

~~$\lambda \mathbf{1} + A$ is a left ideal in $\tilde{\mathfrak{A}}$ and $\tilde{\mathfrak{A}}$ acts on it by left multiplication. The norm of this operator is $\|(\lambda \mathbf{1} + A)\|$.~~

4) $\tilde{\mathfrak{A}}$ is complete in the norm we defined

To prove 1: Since $\|0\mathbb{1} + A\| = \|A\|$, $\|0\mathbb{1} + A\| = 0$ only if $A = 0$. Hence, we have to show that $\|\lambda\mathbb{1} + A\| \neq 0$ if $\lambda \neq 0$. By homogeneity, it suffices to consider the case $\lambda = 1$. Thus, suppose $\|\mathbb{1} - I\| = 0$. Then $(\mathbb{1} - I)B = B - IB = 0$ for all $B \in \mathfrak{A}$. In other words, I is a left identity for \mathfrak{A} . But then, for any $B \in \mathfrak{A}$, $B \cdot (I)^* = ((I) \cdot B)^* = B$, so $(I)^*$ is a right identity for \mathfrak{A} . Now

$$I = I \cdot I^* = I^*,$$

so I is a two-sided identity for \mathfrak{A} , contradicting the assumption that \mathfrak{A} has no identity. This proves 1.

To prove 2 and 3, consider:

$$\begin{aligned} \|\lambda\mathbb{1} + A\|^2 &= \sup \{ \|\lambda B + AB\|^2 : \|B\| \leq 1 \} \\ &= \sup \{ \|B^*(\lambda\mathbb{1} + A)^*(\lambda\mathbb{1} + A)B\| : \|B\| \leq 1 \} \\ &\leq \|(\lambda\mathbb{1} + A)^*(\lambda\mathbb{1} + A)\|. \end{aligned}$$

Now in particular:

$$\|\lambda\mathbb{1} + A\|^2 \leq \|(\lambda\mathbb{1} + A)^*(\lambda\mathbb{1} + A)\| \leq \|(\lambda\mathbb{1} + A)^*\| \|\lambda\mathbb{1} + A\|,$$

so $\|\lambda\mathbb{1} + A\| \leq \|(\lambda\mathbb{1} + A)^*\|$. The opposite inequality follows by replacing $\lambda\mathbb{1} + A$ by $(\lambda\mathbb{1} + A)^*$, so 2. is proved. Also

$$\|\lambda\mathbb{1} + A\|^2 \leq \|(\lambda\mathbb{1} + A)^*(\lambda\mathbb{1} + A)\| \leq \|(\lambda\mathbb{1} + A)^*\| \|\lambda\mathbb{1} + A\| = \|\lambda\mathbb{1} + A\|^2,$$

which proves 3. To prove that $\tilde{\mathfrak{A}}$ is complete, note first that \mathfrak{A} is complete and hence closed in $\tilde{\mathfrak{A}}$. Since \mathfrak{A} is the kernel of the linear functional $\lambda\mathbb{1} + A \mapsto \lambda$, this functional is continuous. Thus, if $\lambda_n\mathbb{1} + A_n$ is a Cauchy sequence in $\tilde{\mathfrak{A}}$, λ_n is a Cauchy sequence in \mathbb{C} , and hence A_n is a Cauchy sequence in \mathfrak{A} . Therefore $\lambda_n \rightarrow \lambda$ and $A_n \rightarrow A$, so $\lambda_n\mathbb{1} + A_n \rightarrow \lambda\mathbb{1} + A$.

The above construction provides a powerful tool for investigating C^* algebras without identity. Unfortunately, it does not solve all problems, and keeping track of what happens in algebras without identity introduces substantial complications into the theory of C^* algebras. By and large, these complications seem to be irrelevant for the application of C^* algebras to physics. We will therefore; in what follows, concentrate on algebras with identity, and make only occasional comments about the general case.

B Spectrum and Resolvent

Let \mathfrak{A} be an algebra with identity. An element A of \mathfrak{A} is said to be *invertible* if there exists an element A^{-1} of \mathfrak{A} such that

$$AA^{-1} = A^{-1}A = \mathbb{1}$$

The element A^{-1} is uniquely determined if it exists; it is called the *inverse* of A . The *resolvent set* of $A(R(A))$ is the set of all complex numbers λ such that $\lambda I - A$ is invertible; the inverse $(\lambda I - A)^{-1}$ is called the *resolvent* of A . The *spectrum* of $A(\sigma(A))$ is the complement of the resolvent set. In general, the spectrum or the resolvent set may be quite arbitrary, but we will show that, in a Banach algebra, the spectrum of any element is a non-empty compact set.

PROPOSITION III. B. 1 *Let A be an element of a normed algebra \mathfrak{A} . Then $\lim_{n \rightarrow \infty} \|A^n\|^{1/n}$ exists and is equal to $\inf_n \|A^n\|^{1/n}$.*

Proof Choose any $m \geq 1$. For any n , write $n = j_n m + i_n$, with $0 \leq i_n < m$. Then

$$\|A^n\|^{1/n} \leq \|A^{j_n m}\|^{1/n} \|A^{i_n}\|^{1/n} \leq \|A^m\|^{j_n/n} \|A^{i_n}\|^{1/n}.$$

Since $\lim_{n \rightarrow \infty} j_n/n = 1/m$ and $\limsup_{n \rightarrow \infty} \|A^{i_n}\|^{1/n} = 1$ (unless $A^n = 0$ for some n , in which case the proposition is trivial), we get

$$\limsup_{n \rightarrow \infty} \|A^n\|^{1/n} \leq \|A^m\|^{1/m}.$$

This is true for all m , so

$$\limsup_{n \rightarrow \infty} \|A^n\|^{1/n} \leq \inf_n \|A^n\|^{1/n}.$$

But, trivially,

$$\inf_n \|A^n\|^{1/n} \leq \liminf_{n \rightarrow \infty} \|A^n\|^{1/n},$$

so the proposition is proved.

The limit whose existence is proved in this proposition is called the *spectral radius* of A and is denoted by $\rho(A)$. The reason for this terminology is the fact, to be proved shortly, that, if the algebra \mathfrak{A} is complete, $\rho(A) = \sup \{|\lambda| : \lambda \in \sigma(A)\}$.

PROPOSITION III. B. 2 *Let \mathfrak{A} be a Banach algebra with identity, and let $A \in \mathfrak{A}$ be invertible. If $B \in \mathfrak{A}$ is such that $\rho(BA^{-1}) < 1$ (in particular, if $\|B\| < \frac{1}{\|A^{-1}\|}$), then $A + B$ is invertible and*

$$(A + B)^{-1} = A^{-1} \sum_{n=0}^{\infty} (-1)^n (BA^{-1})^n;$$

the series converges absolutely in the norm.

Proof The fact that the series converges absolutely in the norm follows at once from the fact that the spectral radius of BA^{-1} is < 1 ; the fact that the sum of the series is an inverse for $A + B$ is a straightforward computation.

COROLLARY III. B. 3 Let \mathfrak{A} be a Banach algebra with identity; let $A \in \mathfrak{A}$. If λ is a complex number with $|\lambda| > \rho(A)$, then λ is in the resolvent set of A , and

$$(\lambda \mathbb{1} - A)^{-1} = \lambda^{-1} \sum_{n=0}^{\infty} \left(\frac{A}{\lambda} \right)^n.$$

As $\lambda \rightarrow \infty$, $\|(\lambda \mathbb{1} - A)^{-1}\| \rightarrow 0$.

COROLLARY III. B. 4 Let \mathfrak{A} be a Banach algebra with identity; let $A \in \mathfrak{A}$; and let λ be in the resolvent set of A . If $|\mu| < \frac{1}{\|(\lambda \mathbb{1} - A)^{-1}\|}$, then $\lambda + \mu$ is in the resolvent set of A and

$$((\lambda + \mu) \mathbb{1} - A)^{-1} = (\lambda \mathbb{1} - A)^{-1} \sum_{n=0}^{\infty} [-\mu(\lambda \mathbb{1} - A)^{-1}]^n.$$

In particular, the resolvent set of A is open (so the spectrum of A is closed), and $\lambda \mapsto (\lambda \mathbb{1} - A)^{-1}$ is an analytic \mathfrak{A} -valued function on the resolvent set of A .

So far, we have not ruled out the possibility that the spectrum is empty.

PROPOSITION III. B. 5 Let \mathfrak{A} be a Banach algebra with identity, and let $A \in \mathfrak{A}$. Then the spectrum of A is non-empty.

Proof Suppose not. Then, for any continuous linear functional ϕ on \mathfrak{A} $\phi((\lambda \mathbb{1} - A)^{-1})$ is an entire function of λ going to zero at infinity. Hence, by Liouville's Theorem, $\phi((\lambda \mathbb{1} - A)^{-1}) \equiv 0$. This is true for all continuous linear functionals, so $(\lambda \mathbb{1} - A)^{-1} = 0$, which is impossible.

Remark The above proposition is valid without the requirement that \mathfrak{A} is complete (but it is necessary that \mathfrak{A} be normed). In the general case, use the above argument to show that the resolvent set of A in the completion $\bar{\mathfrak{A}}$ of \mathfrak{A} cannot be all of \mathbb{C} , and then remark that the resolvent set of A as an element of \mathfrak{A} is contained in the resolvent set of A as an element of $\bar{\mathfrak{A}}$.

THEOREM III. B. 6 (Gelfand) Let \mathfrak{A} be a Banach algebra with identity in which every non-zero element is invertible. Then $\mathfrak{A} = \{\lambda \mathbb{1} : \lambda \in \mathbb{C}\}$, i.e., \mathfrak{A} is isomorphic to \mathbb{C} .

Proof Let $A \in \mathfrak{A}$. Then $\sigma(A) \neq \emptyset$, so $A - \lambda \mathbb{1}$ is not invertible for some λ . This implies $A - \lambda \mathbb{1} = 0$, i.e., $A = \lambda \mathbb{1}$.

PROPOSITION III. B. 7 Let A be a Banach algebra with identity, and let $A \in \mathfrak{A}$. Then $\rho(A) = \sup \{|\lambda| : \lambda \in \sigma(A)\}$.

Proof We know from Proposition III. B. 3 that $\rho(A) \geq \sup \{|\lambda| : \lambda \in \sigma(A)\}$. Hence, assume that $\rho(A) > \sup \{|\lambda| : \lambda \in \sigma(A)\}$. Note that $\mu \mapsto (\mathbb{1} - \mu A)^{-1}$ is analytic on $|\mu| < \frac{1}{\sup \{|\lambda| : \lambda \in \sigma(A)\}}$. Choose r so that $\rho(A) > r > \sup \{|\lambda| : \lambda \in \sigma(A)\}$. Then for any continuous linear functional ϕ on \mathfrak{A}

$\phi((\mathbb{I} - \mu A)^{-1})$ has a power series expansion in μ convergent for $|\mu| \leq 1/r$. But we know that the expansion coefficients must be $\phi(A^n)$ (since $(\mathbb{I} - \mu A)^{-1} = \sum \mu^n A^n$ for $|\mu| < (1/\|A\|)$). Hence, $\phi(A^n)/r^n$ is bounded with respect to n for all continuous linear functionals ϕ . By the uniform boundedness principle, this implies that $\|(A/r)^n\|$ is bounded. This violates the assumption that $r < \rho(A)$ and proves the proposition.

C Commutative Banach Algebras

In this section, we make a preliminary investigation of the structure of commutative Banach algebras. Since the proofs are somewhat technical, we will give a heuristic summary of the approach to be taken. The method of attack is to try to realize, in one sense or another, such algebras as algebras of complex-valued continuous functions. To see how to do this, we first suppose we have a function algebra \mathfrak{A} and ask how we can recover the points of the space on which the functions are defined. One fact about the points of the space is clear: If x is a point, then evaluation at that point $f \mapsto f(x)$ is a morphism from \mathfrak{A} to the complex numbers. Such a morphism (or, at least, one which is not identically zero) we will call a character of \mathfrak{A} . Thus, given an abstract commutative algebra \mathfrak{A} , we will try to realize it as an algebra of functions on the set of its characters, suitably topologized. (Of course, if we start from an algebra of functions, there may be characters which do not come from points of the space on which the functions are defined. We may, however, regard all the characters as points of some "completion" of the original domain.)

The first technical problem which arises in this program is that of proving the existence of characters. There seems to be no very good direct way of constructing them in general. One therefore makes the remark that a character, as a morphism from \mathfrak{A} to the complex numbers, is uniquely determined by its kernel, which is a maximal ideal in \mathfrak{A} . Thus, one wants to find maximal ideals. Now one can make an argument using Zorn's lemma to show that every non-invertible element of \mathfrak{A} is in at least one maximal ideal (Proposition III. C. 2). Next, one has to argue that every maximal ideal is the kernel of a character. This argument has several parts: First one shows that a maximal ideal must be closed by showing that its closure must either be an ideal or all of \mathfrak{A} and ruling out the latter possibility by showing that all elements sufficiently near the identity are invertible so the identity cannot be in the closure of any ideal. (Proposition III. C. 4.) Next one shows that the quotient of a Banach algebra by a closed ideal is again a Banach algebra. Finally, one argues that, in the quotient of an algebra by a maximal ideal, every non-zero element is invertible. (Proposition III. C. 3.) Combining these two remarks with the theorem that the only Banach algebra in which

~~every non-zero element is invertible is isomorphic to \mathbb{C} it follows that the only~~

every non-zero element is invertible is \mathbb{C} itself
 148 concludes that the quo-
 O. E. LANFORD III

tient of \mathfrak{A} by any maximal ideal is isomorphic to \mathbb{C} , i.e., that every maximal ideal of \mathfrak{A} is the kernel of a character. (Proposition III. C. 5.)

Now one topologizes the set of characters conveniently and associates with each element A of \mathfrak{A} the function \hat{A} on the set of characters of \mathfrak{A} which, at the character χ , takes $\hat{A}(\chi) = \chi(A)$. This gives a morphism of \mathfrak{A} into the algebra of continuous functions on the set of characters of \mathfrak{A} . This morphism is not very satisfactory in general: It need not be norm preserving (or even injective) and it need not take the involution on \mathfrak{A} (if one exists) to complex conjugation. In the next section, we will see that the situation is much better if \mathfrak{A} is a commutative C^* algebra.

We now turn to the technical details, starting with some remarks valid in general normed algebras: Let \mathfrak{A} be a normed algebra, \mathcal{C} a closed two-sided ideal in \mathfrak{A} . For $A \in \mathfrak{A}/\mathcal{C}$, define $\|A\| = \inf \{\|A\| : A \in \hat{A}\}$. It is not hard to verify that this defines a norm on \mathfrak{A}/\mathcal{C} (If $\|A\| = 0$, i.e., if A contains elements of arbitrarily small norm, then, since A is closed in \mathfrak{A} , $0 \in A$, i.e. $\hat{A} = 0$ in \mathfrak{A}/\mathcal{C}) making \mathfrak{A}/\mathcal{C} into a normed algebra. If \mathfrak{A} is complete, so is \mathfrak{A}/\mathcal{C} . (This is a general fact about Banach spaces: Let (A_n) be a Cauchy sequence in \mathfrak{A}/\mathcal{C} . It suffices to show that a subsequence of (A_n) converges. Choose a subsequence A_{n_j} so that $\|A_{n_j} - A_{n_m}\| \leq 1/2^j$ for $m \geq n_j$. Then $\|A_{n_j} - A_{n_{j+1}}\| \leq 1/2^j$. Choose representatives A_{n_j} of A_{n_j} such that $\|A_{n_j} - A_{n_{j+1}}\| \leq 1/2^{j-1}$. Then $\sum_{j=1}^{\infty} \|A_{n_j} - A_{n_{j+1}}\| < \infty$, so (A_{n_j}) converges to, say, A ; hence, A_{n_j} converges to A).

If \mathfrak{A} is an algebra with involution, and if $\mathcal{C}^* = \mathcal{C}$, then we can make \mathfrak{A}/\mathcal{C} an algebra with involution by defining $(A + \mathcal{C})^* = A^* + \mathcal{C}$. The involution is compatible with the quotient norm on \mathfrak{A}/\mathcal{C} . (If \mathfrak{A} is a C^* algebra, and if \mathcal{C} is a closed, two-sided ideal, then \mathcal{C} is automatically self-adjoint and \mathfrak{A}/\mathcal{C} is a C^* algebra. We will not prove these facts; they are contained in Proposition I. 8.2 of Dixmier C^*A .)

Next, we investigate ideals in algebras:

PROPOSITION III. C.1 *Let \mathfrak{A} be an algebra with identity and let \mathcal{C} be a (left, right, two-sided) ideal in \mathfrak{A} . Then \mathcal{C} is contained in a maximal (left, right, two-sided) ideal.*

Proof Order the set of ideals by inclusion. Let (\mathcal{C}_α) be an increasing family of ideals; then $\bigcup_{\alpha} \mathcal{C}_\alpha$ is an ideal containing all the \mathcal{C}_α . ($\bigcup_{\alpha} \mathcal{C}_\alpha$ cannot be all of \mathfrak{A} , since $\mathbf{1}$ cannot belong to any \mathcal{C}_α .) The proposition follows by Zorn's Lemma.

PROPOSITION III. C.2 *Let \mathfrak{A} be a commutative algebra with identity, and let $A \in \mathfrak{A}$. Then A belongs to some maximal ideal of \mathfrak{A} if and only if A*

is not invertible.

Proof If A is invertible, then it cannot belong to any ideal of \mathfrak{A} . If A is not invertible, then $\mathfrak{A} \cdot A$ is not all of \mathfrak{A} (it cannot contain 1); hence, is an ideal of \mathfrak{A} ; hence, is contained in a maximal ideal of \mathfrak{A} .

PROPOSITION III. C. 3 *Let \mathfrak{A} be a commutative algebra with identity, \mathcal{C} an ideal of \mathfrak{A} . Then \mathcal{C} is maximal if and only if \mathfrak{A}/\mathcal{C} is a field (i.e., if and only if every non-zero element of \mathfrak{A}/\mathcal{C} is invertible).*

Proof Suppose \mathfrak{A}/\mathcal{C} contains a non-invertible element different from 0 . Then \mathfrak{A}/\mathcal{C} contains an ideal different from $\{0\}$. The inverse image of this ideal under the canonical morphism $\mathfrak{A} \rightarrow \mathfrak{A}/\mathcal{C}$ is an ideal of \mathfrak{A} properly containing \mathcal{C} . Thus, \mathcal{C} is not maximal. Conversely, if \mathcal{C} is not maximal, there is an ideal \mathcal{J} in \mathfrak{A} properly containing \mathcal{C} ; the image of \mathcal{J} in \mathfrak{A}/\mathcal{C} is an ideal which is not equal to $\{0\}$; any element of this ideal must be non-invertible.

So far, we have been doing pure algebra. We now look specifically at Banach algebras.

PROPOSITION III. C. 4 *Let \mathfrak{A} be a Banach algebra with identity, and let \mathcal{C} be a (left, right, two-sided) ideal in \mathfrak{A} . Then $\bar{\mathcal{C}}$ (the closure of \mathcal{C}) is also (left, right, two-sided) ideal in \mathfrak{A} . Every maximal (left, right, two-sided) ideal in \mathcal{C} is closed.*

Proof We consider the case of \mathcal{C} a left ideal. Let $A \in \bar{\mathcal{C}}$ and let $B \in \mathfrak{A}$; we want to show that $B \cdot A \in \bar{\mathcal{C}}$. Let A_n be a sequence in \mathcal{C} converging to A ; then $B \cdot A_n$ is a sequence in \mathcal{C} converging to $B \cdot A$. Hence, $B \cdot A \in \bar{\mathcal{C}}$. Thus, all that remains to be shown is that $\bar{\mathcal{C}}$ is not all of \mathfrak{A} . By Proposition III. B. 2, if $\|A - 1\| < 1$, then A is invertible and hence cannot belong to $\bar{\mathcal{C}}$. Thus $\{A: \|A - 1\| \geq 1\}$ is a closed set containing \mathcal{C} and hence $\bar{\mathcal{C}}$, so $1 \notin \bar{\mathcal{C}}$. A maximal ideal is closed since, if it were not, its closure would be a strictly larger ideal.

PROPOSITION III. C. 5 *Let \mathfrak{A} be a commutative Banach algebra with identity. If \mathcal{C} is a maximal ideal, then \mathfrak{A}/\mathcal{C} is isomorphic to \mathbb{C} .*

Proof If \mathcal{C} is a maximal ideal, then \mathfrak{A}/\mathcal{C} is a Banach algebra in which every non-zero element is invertible; hence, by Theorem III. B. 6, is isomorphic to \mathbb{C} .

A character of a commutative algebra \mathfrak{A} is a non-zero morphism of \mathfrak{A} into the complex numbers. If \mathfrak{A} has an identity, and if χ is a character of \mathfrak{A} , then $\chi(1) = 1$. The kernel of a character of a commutative algebra with identity is a maximal ideal of the algebra. Conversely, by Proposition III. C. 5, every maximal ideal of a commutative Banach algebra with identity is the kernel of a character. Thus, characters are, in this case, in one-one

PROPOSITION III. C. 6 Let \mathfrak{A} be a commutative Banach algebra with identity, and let $A \in \mathfrak{A}$. There is a character χ of \mathfrak{A} such that $\chi(A) = \lambda$ if and only if $\lambda \in \sigma(A)$. In particular, for any character χ of \mathfrak{A} ,

$$|\chi(A)| \leq \rho(A) \leq \|A\|.$$

Proof There is a character χ such that $\chi(A) = \lambda$ if and only if $A - \lambda I$ is sent to zero by some character, which is true if and only if $A - \lambda I$ is not invertible (Proposition III. C. 2), i.e., if and only if $\lambda \in \sigma(A)$.

A character χ of \mathfrak{A} is in particular a linear functional on \mathfrak{A} ; the above proposition says that χ is bounded and has norm not greater than 1. (The norm of a character is in fact equal to 1, since $\chi(I) = 1$.) Thus, the set of characters of \mathfrak{A} is a subset of the unit ball of the dual \mathfrak{A}^* of \mathfrak{A} . We claim that it is in fact a closed subset in the weak- $*$ topology. Let χ_α be a net of characters converging in the weak topology to the linear functional χ ; since $\chi_\alpha(I) = 1$ for all α , $\chi(I) = 1$; if $A, B \in \mathfrak{A}$, then

$$\chi(AB) = \lim \chi_\alpha(AB) = \lim \chi_\alpha(A) \cdot \lim \chi_\alpha(B) = \chi(A) \cdot \chi(B),$$

so χ is a character. Hence we have:

PROPOSITION III. C. 7 Let \mathfrak{A} be a commutative Banach algebra with identity. Then the separate set of characters of \mathfrak{A} is compact in the weak- $*$ topology on the dual of \mathfrak{A} .

The set of characters of \mathfrak{A} , as a topological space, is called the *spectrum*, or *maximal ideal space*, of \mathfrak{A} ; we will denote it by $S(\mathfrak{A})$. If $A \in \mathfrak{A}$, we may define a function \hat{A} on $S(\mathfrak{A})$ by $\hat{A}(\chi) = \chi(A)$. We collect in the following proposition the elementary properties of the map $A \mapsto \hat{A}$, which is called the *Gelfand transform*.

PROPOSITION III. C. 8 Let \mathfrak{A} be a commutative Banach algebra with identity. The mapping $A \mapsto \hat{A}$ is a morphism of the algebra \mathfrak{A} into the algebra of continuous complex-valued functions on $S(\mathfrak{A})$ with the pointwise operations. The range of \hat{A} is precisely the spectrum of A ; in particular, $\|\hat{A}\| = \sup_{\chi \in S(\mathfrak{A})} |\chi(A)| = \rho(A) \leq \|A\|$.

The functions \hat{A} separate points of $S(\mathfrak{A})$, i.e., given $\chi_1, \chi_2 \in S(\mathfrak{A})$, $\chi_1 \neq \chi_2$, we may find $A \in \mathfrak{A}$ such that $\hat{A}(\chi_1) \neq \hat{A}(\chi_2)$.

Proof The fact that \hat{A} is continuous for all $A \in \mathfrak{A}$ follows tautologically from the choice of the topology on $S(\mathfrak{A})$. The mapping $A \mapsto \hat{A}$ is certainly linear; also $\widehat{A \cdot B}(\chi) = \chi(AB) = \chi(A)\chi(B) = \hat{A}(\chi)\hat{B}(\chi)$, so $A \mapsto \hat{A}$ is a morphism. The fact that the range of \hat{A} is the spectrum of A is Proposition III. C. 6. The statement that $\chi_1 \neq \chi_2$ means $\chi_1(A) \neq \chi_2(A)$ for some A , i.e. $\hat{A}(\chi_1) \neq \hat{A}(\chi_2)$ for some A .

The Gelfand transform generalizes, in a certain sense, the Fourier transform. Let $l^1(\mathbb{I})$ denote the space of two-sided sequences of complex numbers $(\dots, a_{-2}, a_{-1}, a_0, a_1, \dots)$ such that $\sum_{n=-\infty}^{\infty} |a_n| < \infty$. If $A = (a_n)$, $B = (b_n)$, define $(A \cdot B)_n = \sum_j a_j \cdot b_{n-j}$, i.e., define products by convolution. Then it may be verified that $l^1(\mathbb{I})$ is a commutative Banach algebra with identity. It turns out that the spectrum of $l^1(\mathbb{I})$ may be identified with the complex numbers of absolute value 1 with the usual topology; the character identified with $e^{i\theta}$ sends $A = (a_n)$ to $\sum_n a_n e^{in\theta}$. Thus, the map $A \mapsto \hat{A}$ is just the classical Fourier transform. In this case, the range of the Gelfand transform is not all continuous functions but the set of continuous functions with absolutely convergent Fourier series, and the supremum norm of \hat{A} is generally strictly smaller than the norm of A .

D Commutative C^* Algebras

At the end of the last section, we saw an example of a commutative Banach algebra with identity in which the Gelfand transform sends \mathfrak{A} to a dense subalgebra of $C(S(\mathfrak{A}))$. The point of the present section is to show that for C^* algebras the situation is better. Specifically, we will prove:

THEOREM III. D.1 *Let \mathfrak{A} be a commutative C^* algebra with identity. Then the Gelfand transform is a norm-preserving isomorphism of \mathfrak{A} onto the algebra of continuous functions on $S(\mathfrak{A})$.*

Before proving this theorem, we will establish some terminology, make a few remarks, and establish a subsidiary result. An element A in an algebra \mathfrak{A} with involution is said to be *self-adjoint* if $A^* = A$ and *normal* if A commutes with A^* . For any A , A^*A is self-adjoint. Any element A of an algebra with involution may be written uniquely as $A = A_1 + iA_2$ with A_1 and A_2 self-adjoint; indeed $A = \frac{1}{2}(A + A^*) + i\left(\frac{A - A^*}{2i}\right)$. A is normal if and only if A_1 and A_2 commute. If A, B are self-adjoint elements of \mathfrak{A} , then $A \cdot B$ is self-adjoint if and only if A and B commute.

PROPOSITION III. D.2 *Let \mathfrak{A} be a C^* algebra, A a normal element of \mathfrak{A} . Then $\rho(A) = \|A\|$.*

Proof

$$\|A^{2^n}\|^2 = \|A^* A^{2^n} A^{2^n}\| = \|(A^* A)^{2^n}\| = \|(A^* A)^{2^{n-1}}\|^2 = \dots = \|A^* A\|^{2^n} = \|A\|^{2^{n+1}}.$$

Thus,

~~$$\|A\|^{2^{n+1}} = \|A^* A\|^{2^n} = \|A\|^{2^{n+1}}$$~~

$$\|A^n\|^{1/2^n} = \|A\| \text{ so } \rho(A) = \lim_{n \rightarrow \infty} \|A^{2^n}\|^{1/2^n} = \|A\|$$

Proof of Theorem III. D. 1 Since \mathfrak{A} is commutative, every element of \mathfrak{A} is normal, so $\|\hat{A}\| = \rho(A) = \|A\|$ for all $A \in \mathfrak{A}$, i.e., the Gelfand transform is norm-preserving as a map from \mathfrak{A} to $C(S(\mathfrak{A}))$. We will show that, if A is self-adjoint, \hat{A} is real. This will imply that $\{\hat{A} : A^* = A\}$ is a norm-complete (hence, norm-closed) algebra of continuous real-valued functions on $S(\mathfrak{A})$ containing the constants and separating points. By the Stone-Weierstrass Theorem, then, every continuous real-valued function is the image under the Gelfand transform of a self-adjoint element of \mathfrak{A} , so the Gelfand transform is surjective and thus an isomorphism.

Let $A \in \mathfrak{A}$ be self-adjoint. Define $\exp(\pm iA) = \sum_{n=0}^{\infty} \frac{(\pm iA)^n}{n!}$. It is straightforward to verify that $[\exp(+iA)]^* = \exp(-iA)$ and that $\exp(+iA) \times \exp(-iA) = \mathbf{1}$. Hence, $\|\exp(\pm iA)\|^2 = \|\exp(\pm iA) \exp(\mp iA)\| = 1$. Thus; for any character χ of \mathfrak{A} ,

$$\begin{aligned} 1 &\geq |\chi(\exp(\pm iA))| = \left| \chi \left(\sum_{n=0}^{\infty} 1/n! (\pm iA)^n \right) \right| = \left| \sum_{n=0}^{\infty} 1/n! (\pm i\chi(A))^n \right| \\ &= |\exp(\pm i\chi(A))|, \end{aligned}$$

so $\chi(A)$ is real.

Let us look briefly at what happens if \mathfrak{A} does not have an identity. Then we can imbed \mathfrak{A} in $\tilde{\mathfrak{A}}$. The Gelfand transform for $\tilde{\mathfrak{A}}$ then maps \mathfrak{A} isometrically onto a subalgebra (in fact, a maximal ideal) of the algebra of continuous functions on $S(\tilde{\mathfrak{A}})$. Now every multiplicative linear functional χ on \mathfrak{A} extends uniquely to a character $\tilde{\chi}$ on $\tilde{\mathfrak{A}}$ by $\tilde{\chi}(\lambda \mathbf{1} + A) = \lambda + \chi(A)$. Conversely, every character of $\tilde{\mathfrak{A}}$ restricts to a character of \mathfrak{A} except for the character $\tilde{\chi}_\infty$ defined by $\tilde{\chi}_\infty(\lambda \mathbf{1} + A) = \lambda$ (which restricts to the zero functional). Thus, the set of characters of \mathfrak{A} , which we will again denote by $S(\mathfrak{A})$ and equip with the weak-* topology, may be identified with $S(\tilde{\mathfrak{A}}) \setminus \{\tilde{\chi}_\infty\}$ (which is a locally compact space since it is obtained by deleting one point from a compact space). Moreover, an element A of $\tilde{\mathfrak{A}}$ belongs to \mathfrak{A} if and only if $\tilde{\chi}_\infty(A) = 0$, so the Gelfand transform for $\tilde{\mathfrak{A}}$ sends \mathfrak{A} to the algebra of all continuous functions on $S(\tilde{\mathfrak{A}})$ vanishing at $\tilde{\chi}_\infty$, i.e., the Gelfand transform for \mathfrak{A} sends \mathfrak{A} isomorphically and isometrically onto the algebra of all continuous functions vanishing at infinity on the locally compact space $S(\mathfrak{A})$.

With the detailed information we have about commutative C^* algebras, we can easily analyze morphisms of such objects. Let $\mathfrak{A}, \mathfrak{B}$ be commutative C^* algebras with identity, and let $\varphi: \mathfrak{A} \rightarrow \mathfrak{B}$ be a morphism such that $\varphi(\mathbf{1}_\mathfrak{A}) = \mathbf{1}_\mathfrak{B}$. If χ is a character of \mathfrak{B} , then $\chi \circ \varphi$ is a character of \mathfrak{A} . Thus, we may define a mapping $\varphi^*: S(\mathfrak{B}) \rightarrow S(\mathfrak{A})$ by $\varphi^*(\chi) = \chi \circ \varphi$. It follows directly from the definition of the topologies on $S(\mathfrak{A})$ and $S(\mathfrak{B})$ that φ^* is continuous. Conversely, if φ^* is any continuous mapping of $S(\mathfrak{B})$ into $S(\mathfrak{A})$, then $\hat{A} \mapsto \hat{A} \circ \varphi^*$

defines a morphism from the continuous functions on $S(\mathcal{B})$ into the continuous functions on $S(\mathcal{A})$, i.e., a morphism from \mathfrak{A} to \mathfrak{B} . Thus, morphisms from \mathfrak{A} to \mathfrak{B} sending $\mathbf{1}_{\mathfrak{A}}$ to $\mathbf{1}_{\mathfrak{B}}$ are in one-one correspondence with continuous mappings of $S(\mathfrak{B})$ into $S(\mathfrak{A})$.

ous functions on $S(\mathfrak{B})$, i.e., a morphism from \mathfrak{A} to \mathfrak{B} . Thus, morphisms from \mathfrak{A} to \mathfrak{B} sending $\mathbf{1}_{\mathfrak{A}}$ to $\mathbf{1}_{\mathfrak{B}}$ are in one-one correspondence with continuous mappings of $S(\mathfrak{B})$ into $S(\mathfrak{A})$.

PROPOSITION III. D. 3 Let $\mathfrak{A}, \mathfrak{B}$ be commutative C^* algebras with identity, φ a morphism of \mathfrak{A} into \mathfrak{B} sending $\mathbf{1}_{\mathfrak{A}}$ to $\mathbf{1}_{\mathfrak{B}}$, φ^* the associated mapping from $S(\mathfrak{B})$ to $S(\mathfrak{A})$. Then φ is injective if and only if φ^* is surjective, and φ is surjective if and only if φ^* is injective. If φ is injective, then $\|\varphi(A)\| = \|A\|$ for all $A \in \mathfrak{A}$.

Proof Since φ^* is continuous and $S(\mathfrak{B})$ is compact, $\varphi^*(S(\mathfrak{B}))$ is compact and hence closed in $S(\mathfrak{A})$. Now $\varphi(A) = 0$ if and only if $\chi(\varphi(A)) = 0$ for all $\chi \in S(\mathfrak{A})$, i.e., if and only if $\hat{A} = 0$ on $\varphi^*(S(\mathfrak{B}))$. This shows that φ is injective if and only if zero is the only continuous function on $S(\mathfrak{A})$ vanishing on $\varphi^*(S(\mathfrak{B}))$. Since $\varphi^*(S(\mathfrak{B}))$ is closed, this is equivalent to $\varphi^*(S(\mathfrak{B})) = S(\mathfrak{A})$. If φ is surjective, then $\|\varphi(A)\| = \sup \{|\chi(\varphi(A))| : \chi \in S(\mathfrak{A})\} = \sup \{|\varphi^*(\chi)(A)| : \chi \in S(\mathfrak{B})\} = \sup \{|\chi(A)| : \chi \in S(\mathfrak{A})\} = \|A\|$.

The morphism φ is surjective if and only if, for all $B \in \mathfrak{B}$, there is an $A \in \mathfrak{A}$ such that $\hat{A} \circ \varphi^* = \hat{B}$. If φ^* is not injective this is clearly not possible since, if $\varphi^*(\chi_1) = \varphi^*(\chi_2)$, no \hat{B} separating χ_1 from χ_2 can be so obtained. On the other hand, if φ^* is injective, then it is a continuous one-one mapping from the compact space $S(\mathfrak{B})$ to the compact set $\varphi^*(S(\mathfrak{B})) \subset S(\mathfrak{A})$. But such a mapping has a continuous inverse (it sends closed sets, i.e., compact sets, to sets which are compact; hence closed, so the inverse image under its inverse of a closed set is closed.) Thus, all we have to do is to extend the continuous function $\hat{B} \circ (\varphi^*)^{-1}$, defined on $\varphi^*(S(\mathfrak{B}))$, to a continuous function \hat{A} on all of $S(\mathfrak{A})$. The Tietze Extension Theorem asserts the existence of such an extension, so the proposition is proved.

E The Spectral Theorem for Bounded Normal Operators on Hilbert Space

As an example of the power of the methods we have been discussing, we will show that they lead to a very quick proof of the spectral theorem for bounded normal operators. Let \mathcal{H} be a Hilbert space, and let A be a bounded operator on \mathcal{H} which is normal, i.e., which commutes with its adjoint. (In particular, A can be self-adjoint or unitary.) Let \mathfrak{A} be the C^* algebra generated by A and $\mathbf{1}$, i.e., the norm closure of the set of all polynomials in A and A^* . \mathfrak{A} is a commutative C^* algebra, and hence is isomorphic to the algebra of all continuous functions on a compact space. (It is worth remarking that in this case, $S(\mathfrak{A})$ may be identified with the spectrum of A , regarded as a subset of \mathbb{C} . The proof is as follows: Consider the mapping $S(\mathfrak{A}) \rightarrow \mathbb{C}$ defined by $\chi \mapsto \chi(A)$. This is continuous by the definition of the topology on $S(\mathfrak{A})$, and its image is exactly the spectrum of A . Since $\chi(A^*) = \overline{\chi(A)}$ for any character χ , the mapping $\chi \mapsto \chi(A)$ is a homeomorphism of $S(\mathfrak{A})$ onto the spectrum of A .)

for any character of \mathfrak{A} and since a character is a multiplicative linear

functional, two characters which agree on A agree on all polynomials in A and A^* and, hence, by continuity, are equal on \mathfrak{A} . Thus, the mapping $\chi \mapsto \chi(A)$ is continuous and one-one from $S(\mathfrak{A})$ to $\sigma(A)$. But we have already observed that a continuous one-one mapping between compact spaces is a homeomorphism, so $S(\mathfrak{A})$ is homeomorphic to $\sigma(A)$.

Now let us assume, temporarily, that \mathfrak{A} has a cyclic vector, i.e., that there is a vector $\xi \in \mathcal{H}$, which we may take to have norm one, such that $\{B\xi; B \in \mathfrak{A}\}$ is dense in \mathcal{H} . We will define a linear functional μ on $C(S(\mathfrak{A}))$ by $\mu(\hat{B}) = (\xi | B\xi)$. We claim that μ is positive and hence defines a measure on $S(\mathfrak{A})$. To see this, suppose that $\hat{B} \geq 0$ everywhere. Then $\sqrt{\hat{B}}$ ($\sqrt{\quad}$ denotes the positive square root) is again a continuous function on $S(\mathfrak{A})$ and hence is the Gelfand transform of a self-adjoint element of \mathfrak{A} , which we will denote by \sqrt{B} , and which evidently has the property that $(\sqrt{B})^2 = B$. Then

$$\mu(\hat{B}) = (\xi | B\xi) = (\xi | \sqrt{B} \cdot \sqrt{B} \xi) = (\sqrt{B} \xi | \sqrt{B} \xi) \geq 0.$$

This shows that μ defines a measure on $S(\mathfrak{A})$; we will also denote the measure by μ . We now claim that we can define a unitary mapping of \mathcal{H} into $\mathcal{L}^2(\mu)$ by

$$B\xi \mapsto \hat{B}.$$

It is not clear at the moment that this mapping is well defined, but we have, for $B, C \in \mathfrak{A}$,

$$(B\xi | C\xi) = (\xi | B^*C\xi) = \int d\mu \widehat{B^*C} = \int d\mu \hat{B}^* \hat{C}.$$

so the proposed mapping preserves scalar products, so it is well defined and length-preserving; since it is defined on the dense subset $\mathfrak{A}\xi$ of \mathcal{H} and has range the dense subset $C(S(\mathfrak{A})) \subset \mathcal{L}^2(\mu)$, it extends uniquely to a unitary operator U from \mathcal{H} onto $\mathcal{L}^2(\mu)$. Now let $B \in \mathfrak{A}$ be arbitrary; we have

$$\hat{AB} = UAB\xi = UAU^{-1}UB\xi = UAU^{-1}\hat{B},$$

so UAU^{-1} is equal to the operator M_A of multiplication by A on the dense subset $C(S(\mathfrak{A}))$ of $\mathcal{L}^2(\mu)$; since both operators are bounded, they must agree on all of $\mathcal{L}^2(\mu)$. Stripping away the details of the construction, we see that we have proved the following: If \mathfrak{A} has a cyclic vector, then A is unitarily equivalent to the operator of multiplication by a continuous function on $\mathcal{L}^2(\mu)$, where μ is a Borel measure on a compact space (Moreover, the space may be taken to be $\sigma(A)$ and the function to be the co-ordinate function z .)

We now must remove the requirement that \mathfrak{A} have a cyclic vector. To do this, we use the fact, to be proved later, that if \mathfrak{A} is any concrete C^* algebra containing the identity operator on a Hilbert space \mathcal{H} , then \mathcal{H} may be decomposed into a direct sum of subspaces $\mathcal{H} = \bigoplus_{i \in I} \mathcal{H}_i$, such that each

\mathcal{H}_i is mapped into itself by each element of \mathfrak{A} and such that the restriction of \mathfrak{A} to each \mathcal{H}_i has a cyclic vector (see Proposition 11.1, F. 2). Then we can

use the above argument to show that A , restricted to any \mathcal{X}_i , is unitarily equivalent to the operator of multiplication by \tilde{A} on $\mathcal{L}^2(\mu_i)$ for some measure μ_i on $S(\mathfrak{A})$. We may now construct a (possibly very large) space \mathcal{X} by taking the disjoint union of one copy of $S(\mathfrak{A})$ for each $i \in I$ and defining a Borel measure μ on this union by requiring that its restriction to the i^{th} copy be just μ_i . Then $\mathcal{L}^2(\mu)$ may be identified with $\bigoplus_{i \in I} \mathcal{L}^2(\mu_i)$, and this identification makes A unitarily equivalent to the operator on $\mathcal{L}^2(\mu)$ of multiplication by the function whose restriction to each copy of $S(\mathfrak{A})$ is just \tilde{A} . Thus, we can formulate the following theorem.

THEOREM III. E. 1 *Let A be a bounded normal operator on a Hilbert space. Then A is unitarily equivalent to the operator of multiplication by a bounded continuous function on \mathcal{L}^2 of a Borel measure on a locally compact space.*

The realization of A as a multiplication operator may be thought of as "diagonalizing" A . We may easily deduce from this statement the version of the spectral theorem expressing A in terms of a projection-valued measure: Let $A = UM_fU^{-1}$, where M_f is multiplication by f on $\mathcal{L}^2(\mu)$ and let S be a Borel subset of \mathbb{C} . Let $U^{-1}P(S)U$ be the operator of multiplication by the characteristic function of $\{x : f(x) \in S\}$ on $\mathcal{L}^2(\mu)$. It is easy to check that $S \mapsto P(S)$ is a projection-valued measure on \mathbb{C} and that, for any $\xi \in \mathcal{X}$,

$$(\xi | M_f \xi) = \int_{\mathbb{C}} \lambda(\xi | dP(\lambda) \xi).$$

If A is self-adjoint, then f is real so the projection-valued measure P is concentrated on the real axis, and we get the usual representation of a bounded self-adjoint operator in terms of its spectral resolution.

In the classical versions of the spectral theorem, it is shown that, if A is self-adjoint and if B commutes with A , then B commutes with the spectral projections of A . We can prove this as follows: Let F be any closed set in the spectrum of A . Then there is a non-negative continuous function $C(\lambda)$ on $\sigma(A)$ equal to one on F , and strictly less than one on the complement of F . C is the Gelfand transform of an element C of \mathfrak{A} . Since, by assumption, B commutes with A , it commutes with all polynomials in A , and hence with all elements of \mathfrak{A} ; in particular, B commutes with C . By realizing A as a multiplication operator, and applying the monotone convergence theorem, we see that $C^n \xi$ converges to $P(F) \xi$ for all $\xi \in \mathcal{X}$. Thus

$$B P(F) \xi = \lim_{n \rightarrow \infty} B C^n \xi = \lim_{n \rightarrow \infty} C^n B \xi = P(F) B \xi$$

for all $\xi \in \mathcal{X}$, i.e., B commutes with $P(F)$ for all closed $F \subset \sigma(A)$. It is easy to see that the collection of subsets G of $\sigma(A)$ such that $P(G)$ commutes with B is a σ -algebra of sets.

with B is a σ -field; since it contains
~~closed~~ closed sets by the above argument,

it contains all Borel sets, i.e., $P(E)$ commutes with B for all Borel subsets E of $\sigma(A)$.

One advantage of the derivation of the spectral theorem that we have given over the more classical ones is that it shows that any collection, finite or infinite, of commuting self-adjoint operators can be simultaneously diagonalized, i.e., realized as multiplication operators on an \mathcal{L}^2 space. The proof is the same; one has only to replace \mathfrak{A} by the C^* algebra generated by the family of operators in question.

F Generalities on Representations

Let \mathfrak{A} be an algebra with involution. A *representation* of \mathfrak{A} on a Hilbert space \mathcal{X} is a morphism of \mathfrak{A} into the C^* algebra $\mathcal{L}(\mathcal{X})$ of bounded operators on \mathcal{X} . (Thus, we are defining "representation" to mean "representation by bounded operators". We will see later that, if \mathfrak{A} is a Banach algebra with involution and identity and if π is any morphism of \mathfrak{A} into the set of (possibly unbounded) operators on a dense domain in a Hilbert space with adjoint defined in the obvious way, then $\pi(A)$ is bounded for all $A \in \mathfrak{A}$ and, indeed,

$$\|\pi(A)\| \leq \|A\|.$$

Two representations π and π' on Hilbert spaces \mathcal{X} and \mathcal{X}' are said to be *unitarily equivalent* ($\pi \cong \pi'$) if there is a unitary operator U from \mathcal{X} to \mathcal{X}' such that

$$U^{-1} \pi'(\cdot) U = \pi(\cdot).$$

Let $(\pi_i)_{i \in I}$ be an indexed set of representations of \mathfrak{A} on Hilbert spaces (\mathcal{X}_i) . If, for each $A \in \mathfrak{A}$, $\sup_i \|\pi_i(A)\| < \infty$ (In particular, if \mathfrak{A} is a Banach algebra with identity, by the remark made above), we can form the direct sum representation $\bigoplus_i \pi_i$ on the direct sum Hilbert space $\bigoplus_i \mathcal{X}_i$, as follows: $\bigoplus_i \mathcal{X}_i$ is the set of indexed families (ξ_i) with $\xi_i \in \mathcal{X}_i$, such that $\sum_i \|\xi_i\|^2 < \infty$. For $A \in \mathfrak{A}$, the operator $\bigoplus_i \pi_i(A)$ is defined to take (ξ_i) to $(\pi_i(A) \xi_i)$.

If π is a representation of \mathfrak{A} on \mathcal{X} , and if \mathcal{X}^1 is a linear subspace of \mathcal{X} , we say that \mathcal{X}^1 is an *invariant subspace* for π if $\pi(A) \mathcal{X}^1 \subset \mathcal{X}^1$ for all $A \in \mathfrak{A}$. If \mathcal{X}^1 is invariant, so is its orthogonal complement $\mathcal{X}^{1\perp}$, since if $\xi \in \mathcal{X}^1$, $\eta \in \mathcal{X}^{1\perp}$, and $A \in \mathfrak{A}$,

$$(\xi | \pi(A) \eta) = (\pi(A^*) \xi | \eta) = 0.$$

Now if \mathcal{X}^1 is *closed* and invariant, and if $P_{\mathcal{X}^1}$ is the projection onto \mathcal{X}^1 , we have

$$\begin{aligned} P_{\mathcal{X}^1} \pi(A) \xi &= P_{\mathcal{X}^1} \pi(A) P_{\mathcal{X}^1} \xi + P_{\mathcal{X}^1} \pi(A) (\mathbb{I} - P_{\mathcal{X}^1}) \xi \\ &= \pi(A) P_{\mathcal{X}^1} \xi \end{aligned}$$

since $\pi(A) P_{\mathcal{X}^1} \xi \in \mathcal{X}^1$ and $\pi(A) (\mathbb{I} - P_{\mathcal{X}^1}) \xi \in \mathcal{X}^{1\perp}$,

Thus, $P_{\mathcal{X}_1}$ commutes with $\pi(A)$ for all $A \in \mathfrak{A}$. Conversely, if $P_{\mathcal{X}_1}$ commutes with $\pi(A)$ for all $A \in \mathfrak{A}$, then \mathcal{X}_1 is invariant. Thus: A closed subspace \mathcal{X}_1 of \mathcal{X} is invariant for π if and only if $P_{\mathcal{X}_1}$ commutes with $\pi(A)$ for all $A \in \mathfrak{A}$.

There is a particularly trivial kind of representation for any algebra \mathfrak{A} , that in which $\pi(A) = 0$ for all $A \in \mathfrak{A}$. We want to split any representation into a part of this kind and a part in which such pathology is entirely absent.

PROPOSITION III. F. 1 Let \mathfrak{A} be an algebra with involution, π a representation of \mathfrak{A} on a Hilbert space \mathcal{X} . Then π splits into the direct sum of two orthogonal invariant subspaces \mathcal{X}_1 and \mathcal{X}_2 ; \mathcal{X}_1 is the closed linear span of

$$\{\pi(A)\xi : A \in \mathfrak{A}; \xi \in \mathcal{X}\},$$

and

$$\mathcal{X}_2 = \{\xi \in \mathcal{X} : \pi(A)\xi = 0 \text{ for all } A \in \mathfrak{A}\}.$$

If \mathfrak{A} has an identity, then $\pi(\mathbf{1})$ is the projection onto \mathcal{X}_1 .

Proof Let \mathcal{X}_1 and \mathcal{X}_2 be defined as in the statement of the proposition. We have to show that $\mathcal{X}_2 = \mathcal{X}_1^\perp$. First note that, for a fixed A , the orthogonal complement of the range of $\pi(A)$ is the null space of $\pi(A^*)$. This is true since η is orthogonal to the range of $\pi(A)$ if and only if $(\eta | \pi(A)\xi) = 0$ for all $\xi \in \mathcal{X}$, which is true if and only if $(\pi(A^*)\eta | \xi) = 0$ for all $\xi \in \mathcal{X}$, which is true if and only if $\pi(A^*)\eta = 0$. Thus, the orthogonal complement of \mathcal{X}_1 is the intersection over $A \in \mathfrak{A}$ of the null space of $\pi(A^*)$, and this is what we have called \mathcal{X}_2 . It is clear that \mathcal{X}_2 is invariant, so \mathcal{X}_1 is also. If \mathfrak{A} has an identity, then $\pi(\mathbf{1})$ is the identity on \mathcal{X}_1 and zero on \mathcal{X}_2 , so $\pi(\mathbf{1})$ is the projection onto \mathcal{X}_1 .

With the notation of the above proposition, we say that π is *non-degenerate* if $\mathcal{X}_2 = \{0\}$. A vector ξ is said to be a *cyclic vector* for π if $\{\pi(A)\xi : A \in \mathfrak{A}\}$ is dense in \mathcal{X} and π is said to be a *cyclic representation* if it admits a cyclic vector. From the above proposition it is clear that a cyclic representation is non-degenerate. The converse is not true, but we have:

PROPOSITION III. F. 2 Let \mathfrak{A} be an algebra with involution. Then every non-degenerate representation of \mathfrak{A} is a direct sum of cyclic representations.

To prove this, we need the following:

PROPOSITION III. F. 3 Let \mathfrak{A} be an algebra with involution, π a non-degenerate representation of \mathfrak{A} on \mathcal{X} , and $\xi \in \mathcal{X}$. Then $\overline{\pi(\mathfrak{A})\xi}$ is a closed invariant subspace of \mathcal{X} containing ξ .

Proof of III. F. 3 It is clear that $\overline{\pi(\mathfrak{A})\xi}$ is invariant; let P be the projection onto this subspace. Then P commutes with $\pi(A)$ for all $A \in \mathfrak{A}$, so $0 = (\mathbf{1} - P)\pi(A)\xi = \pi(A)(\mathbf{1} - P)\xi$, so $(\mathbf{1} - P)\xi = 0$ by the non-degeneracy

~~of π~~ of π , so $\xi \in \overline{\pi(\mathfrak{A})\xi}$

Proof of III. F. 2 We use Zorn's Lemma. Consider the collection of all sets $\{\mathcal{H}_i\}$ of pairwise orthogonal closed invariant non-zero subspaces of \mathcal{H} such that the restriction of π to each \mathcal{H}_i has a cyclic vector. Order this collection by inclusion. It is nearly obvious that the hypotheses of Zorn's Lemma are satisfied. Thus, there exists a maximal collection $\{\mathcal{H}_i\}_{i \in I}$. The maximality implies that there is no non-zero cyclic subspace orthogonal to all the \mathcal{H}_i . By III. F. 3, this means that $\mathcal{H} = \bigoplus_{i \in I} \mathcal{H}_i$. (Otherwise, there would exist $\xi \neq 0$ orthogonal to all the \mathcal{H}_i 's; then $\overline{\pi(\mathbb{N})\xi}$ would be a cyclic subspace orthogonal to all the \mathcal{H}_i 's.)

If \mathfrak{A} is an algebra with involution and π is a representation of \mathfrak{A} on a Hilbert space \mathcal{H} , we say that π is *irreducible* if the only closed invariant subspaces for π are $\{0\}$ and \mathcal{H} . The term "topologically irreducible" is sometimes used for this concept, the term "irreducible" being defined to mean that the only invariant linear subspaces, closed or not, are $\{0\}$ and \mathcal{H} . We will not have occasion to use the stronger notion; anyway, it turns out that, for \mathfrak{A} a C^* algebra, they are equivalent. (See Dixmier, C^* A, Corollaire 2.8.4, p. 45).

PROPOSITION III. F. 4 *Let \mathfrak{A} be an algebra with involution, π a representation of \mathfrak{A} on a Hilbert space \mathcal{H} . Then the following are equivalent.*

- i) π is irreducible.
- ii) The only orthogonal projections on \mathcal{H} commuting with $\pi(A)$ for all A are 0 and $\mathbf{1}$.
- iii) The only bounded operators on \mathcal{H} commuting with $\pi(A)$ for all A are scalar multiples of $\mathbf{1}$.

If π is non-degenerate, these are all equivalent to

- iv) Every non-zero vector $\xi \in \mathcal{H}$ is a cyclic vector for π .

Proof i) and ii) are equivalent, since we have already seen that a closed subspace of \mathcal{H} is invariant for π if and only if its orthogonal projection commutes with $\pi(A)$ for all A . Clearly, iii) implies ii); we now show that ii) implies iii). Thus, let ii) hold, and let T be a bounded operator commuting with all $\pi(A)$'s. Then $[\pi(A), T^*] = [T, \pi(A^*)]^* = 0$ for all A , so T^* also commutes with all $\pi(A)$'s. Hence, $\frac{T + T^*}{2}$ and $\frac{T - T^*}{2i}$ also commute with all $\pi(A)$'s so if we can show that, if T is a self-adjoint operator commuting with all $\pi(A)$'s, then T is a scalar multiple of $\mathbf{1}$, we are through. But if T is self-adjoint and commutes with all $\pi(A)$'s, then the spectral projections of T must commute with all $\pi(A)$'s; hence, by ii), must all be zero or $\mathbf{1}$. This implies that T is a scalar multiple of $\mathbf{1}$.

Now let π be non-degenerate and irreducible, and let $\xi \in \mathcal{H}$, $\xi \neq 0$. By Proposition III. F. 3. $\overline{\pi(\mathbb{N})\xi}$ is a non-zero closed invariant subspace.

hence, must be all of \mathcal{H} , so ξ is a cyclic vector for π . Thus, i) implies iv). Conversely, suppose π is not irreducible, and let ξ be a non-zero element of a proper closed invariant subspace, \mathcal{H}_1 . Then $\overline{\pi(\mathfrak{A})\xi} \subset \mathcal{H}_1 \neq \mathcal{H}$, so ξ cannot be a cyclic vector for π . Thus, iv) implies i) and the proof is complete.

G Positive Linear Functionals and the Gelfand-Segal Construction

We saw in the last section that every non-degenerate representation of an algebra \mathfrak{A} with involution may be written as a direct sum of cyclic representations. We are now going to show how to describe a cyclic representation of \mathfrak{A} by a linear functional on \mathfrak{A} of a special kind. Thus, the study of representations is reduced, in a certain sense, to the study of the so-called positive linear functionals on \mathfrak{A} . The definition of positive linear functionals is motivated by the following remark: Let \mathfrak{A} be an algebra with involution, π a representation of \mathfrak{A} on a Hilbert space \mathcal{H} , ξ a vector in \mathcal{H} . Define a linear functional ϕ on \mathfrak{A} by $\phi(A) = (\xi | \pi(A)\xi)$. Then, for any $A \in \mathfrak{A}$, $\phi(A^*A) = (\xi | \pi(A^*A)\xi) = (\pi(A)\xi | \pi(A)\xi) \geq 0$. We are thus led to define: A *positive linear functional* on an algebra \mathfrak{A} with involution is a linear functional ϕ on \mathfrak{A} such that $\phi(A^*A) \geq 0$ for all $A \in \mathfrak{A}$.

The study of positive linear functionals is largely based on the remark that, if ϕ is a positive linear functional on \mathfrak{A} , then we can define something which is almost an inner product on \mathfrak{A} by

$$\langle A | B \rangle = \phi(A^*B).$$

We have therefore to review some elementary facts about spaces with inner products.

Let E be a vector space over the complex numbers. A *sesquilinear form* on E is a mapping $(\xi, \eta) \mapsto \langle \xi | \eta \rangle$ from $E \times E$ to \mathbb{C} , such that

$$\langle \alpha\xi + \beta\eta | \zeta \rangle = \bar{\alpha}\langle \xi | \zeta \rangle + \bar{\beta}\langle \eta | \zeta \rangle,$$

$$\langle \zeta | \alpha\xi + \beta\eta \rangle = \alpha\langle \zeta | \xi \rangle + \beta\langle \zeta | \eta \rangle$$

for $\alpha, \beta \in \mathbb{C}$, $\xi, \eta, \zeta \in E$. A sesquilinear form is *positive semi-definite* if $\langle \xi | \xi \rangle \geq 0$ for all $\xi \in E$, and *positive definite* if $\langle \xi | \xi \rangle > 0$ for $\xi \in E$, $\xi \neq 0$. A positive-definite sesquilinear form is also called an *inner product*. A vector space equipped with a positive semi-definite sesquilinear form is called a *pre-Hilbert space*; if the sesquilinear form is positive definite, we speak of a *strict pre-Hilbert space*. Elementary arguments show that a positive semi-definite sesquilinear form is *hermitian*, i.e.

$$\langle \xi | \eta \rangle = \overline{\langle \eta | \xi \rangle}$$

and satisfies the Schwarz inequality

$$\langle \xi, \eta \rangle^2 \leq \langle \xi | \xi \rangle \langle \eta | \eta \rangle.$$

If we define $\|\xi\| = \sqrt{\langle \xi | \xi \rangle}$ then $\|\cdot\|$ is a semi-norm on E ; the semi-norm is a norm if E is a strict pre-Hilbert space. Any strict pre-Hilbert space may be completed, i.e., linearly imbedded as a dense subspace of a Hilbert space in an inner-product preserving way.

There is also a standard way of constructing from a pre-Hilbert space E , $\langle \cdot | \cdot \rangle$, a strict pre-Hilbert space: Let

$$I = \{\xi \in E : \langle \xi | \xi \rangle = 0\}.$$

Since $\|\cdot\|$ is a seminorm, I is a linear subspace. If $\xi - \xi^1$ and $\eta - \eta^1$ belong to I ,

$$\begin{aligned} |\langle \xi | \eta \rangle - \langle \xi^1 | \eta^1 \rangle| &\leq |\langle \xi - \xi^1 | \eta \rangle| + |\langle \xi^1 | \eta - \eta^1 \rangle| \leq \|\eta\| \|\xi - \xi^1\| \\ &\quad + \|\xi^1\| \|\eta - \eta^1\| = 0. \end{aligned}$$

Thus, $\langle \xi | \eta \rangle$ only depends on $\xi + I, \eta + I$. In other words, the sesquilinear form $\langle \cdot | \cdot \rangle$ may be regarded as mapping $E/I \times E/I$ to \mathbb{C} ; this mapping is easily seen to be a scalar product on E/I . We shall speak of E/I with the scalar product constructed in this way as the strict pre-Hilbert space associated with $E, \langle \cdot | \cdot \rangle$.

Returning to positive linear functionals, we have:

PROPOSITION III. G. 1 *Let \mathfrak{A} be an algebra with involution, ϕ a positive linear functional on \mathfrak{A} . Then, for $\xi, \eta \in \mathfrak{A}$,*

$$\phi(\xi^* \eta) = \overline{\phi(\eta^* \xi)};$$

$$|\phi(\xi^* \eta)|^2 \leq \phi(\xi^* \xi) \cdot \phi(\eta^* \eta).$$

If \mathfrak{A} has an identity, then

$$\phi(\xi^*) = \overline{\phi(\xi)};$$

$$|\phi(\xi)|^2 \leq \phi(\xi^* \xi) \cdot \phi(\mathbf{1}).$$

Proof Everything follows from the fact that $\langle \xi | \eta \rangle = \phi(\xi^* \eta)$ is a positive semi-definite sesquilinear form.

We now come to a less trivial result, that positive linear functionals on Banach algebras with identity and involution are automatically continuous.

PROPOSITION III. G. 2 *Let \mathfrak{A} be a Banach algebra with identity and involution, and let ϕ be a positive linear functional on \mathfrak{A} . Then ϕ is continuous and $\|\phi\| = \phi(\mathbf{1})$.*

Proof Let $\xi \in \mathfrak{A}$, and assume $\|\xi\| < 1$. We will prove that

$$\phi(\xi^* \xi) \leq \phi(\mathbf{1}).$$

Then, by the Schwarz inequality

$$|\phi(\xi)|^2 \leq \phi(\xi^*\xi) \cdot \phi(\mathbf{1}) \leq \phi(\mathbf{1})^2$$

if $\|\xi\| < 1$; thus, ϕ is continuous and $\|\phi\| \leq \phi(\mathbf{1})$. The other inequality is immediate; $\|\phi\| \geq \phi(\mathbf{1})$ since $\|\mathbf{1}\| = 1$.

To show that $\phi(\xi^*\xi) \leq \phi(\mathbf{1})$, we prove that $\mathbf{1} - \xi^*\xi$ has a positive square-root. Consider the series

$$\eta = \mathbf{1} - \frac{1}{2} \cdot (\xi^*\xi) - \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{(\xi^*\xi)^2}{2!} - \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{3}{2} \cdot \frac{(\xi^*\xi)^3}{3!} - \dots$$

(We have inserted $\xi^*\xi$ for z in the Taylor series for $\sqrt{1-z}$ about $z=0$). The series converges since $\sqrt{1-z}$ is analytic for $|z| < 1$ and since $\|\xi^*\xi\| < 1$. Also, η is self-adjoint since it is a sum of self-adjoint terms. Exactly the same calculation as that required to show that

$$\left(1 - \frac{1}{2}z - \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{z^2}{2!} - \dots\right)^2 = 1 - z$$

for z a complex number of modulus less than 1 shows that

$$\eta^2 = \mathbf{1} - \xi^*\xi.$$

Thus

$$0 \leq \phi(\eta^2) = \phi(\mathbf{1}) - \phi(\xi^*\xi), \quad \text{so} \quad \phi(\mathbf{1}) \geq \phi(\xi^*\xi).$$

Remark For C^* algebras, one can prove by a different argument that every positive linear functional is continuous without the assumption that \mathfrak{A} has an identity. See Dixmier, C^*A , 2.1.8.

We define as *state* of a Banach algebra with involution to be a continuous positive linear functional ϕ of norm one. If \mathfrak{A} has an identity, we may remove the requirement that ϕ be continuous, and replace the requirement $\|\phi\| = 1$ by $\phi(\mathbf{1}) = 1$. Notice a peculiar property of the norms of positive linear functionals: If \mathfrak{X} is an arbitrary Banach space, and if ϕ, ψ are linear functionals on \mathfrak{X} , then $\|\phi + \psi\|$ is normally strictly less than $\|\phi\| + \|\psi\|$. For example, if \mathfrak{X} is a Hilbert space, then

$$\|\phi + \psi\| = \|\phi\| + \|\psi\|$$

only if ϕ and ψ are linearly dependent. However, if ϕ and ψ are positive linear functionals on a Banach algebra with identity and involution, then $\phi + \psi$ is also positive, and

$$\|\phi + \psi\| = (\phi + \psi)(\mathbf{1}) = \phi(\mathbf{1}) + \psi(\mathbf{1}) = \|\phi\| + \|\psi\|.$$

By this remark, the set of states of a Banach algebra with identity is convex.

Using the above proposition we can prove a result about the continuity of representations.

PROPOSITION III G. 3 Let \mathfrak{A} be a Banach algebra with involution, π a morphism of \mathfrak{A} into the algebra of all linear operators on a pre-Hilbert space \mathfrak{X} , such that, for $\xi, \eta \in \mathfrak{X}$

$$\langle \xi | \pi(A) \eta \rangle = \langle \pi(A^*) \xi | \eta \rangle.$$

Then $\|\pi(A) \xi\| \leq \|A\| \|\xi\|$, for all $\xi \in \mathfrak{X}$.

Proof If \mathfrak{A} has no identity, extend π to $\bar{\mathfrak{A}}$ by defining $\pi(\mathbf{1})$ to be the identity operator on \mathfrak{X} . Thus, we may assume that \mathfrak{A} has an identity. Now, for any $\xi \in \mathfrak{X}$, define a linear functional ϕ_ξ on \mathfrak{A} by $\phi_\xi(A) = \langle \xi | \pi(A) \xi \rangle$. ϕ_ξ is positive, so, for any $A \in \mathfrak{A}$,

$$\begin{aligned} \|\pi(A) \xi\|^2 &= \langle \xi | \pi(A^*A) \xi \rangle = \phi_\xi(A^*A) \leq \|A\|^2 (\phi_\xi(\mathbf{1})) \\ &= \|A\|^2 \|\xi\|^2. \end{aligned}$$

We are now ready to construct a representation associated with a positive functional. The idea is as follows: We begin with a positive linear functional ϕ on an algebra \mathfrak{A} with involution and identity. We make \mathfrak{A} itself into a pre-Hilbert space by $\langle A | B \rangle = \phi(A^*B)$. Let $I = \{A \in \mathfrak{A} : \phi(A^*A) = 0\}$. For each $A \in \mathfrak{A}$, we define a linear mapping L_A of \mathfrak{A} into itself by $L_A B = A \cdot B$.

Evidently, $L_A L_B = L_{AB}$ and

$$\langle L_A B | C \rangle = \langle B | L_A^* C \rangle.$$

We want to show that each L_A induces a linear mapping on the quotient space \mathfrak{A}/I . To do this, we have to show that

$$\phi(B^*B) = 0 \text{ implies } \phi(B^*A^*AB) = 0.$$

To do this, we note that $\phi_B(\cdot) = \phi(B^*(\cdot)B)$ is a positive linear functional, and hence, by the Schwarz inequality, we have:

$$\begin{aligned} |\phi(B^*A^*AB)|^2 &= |\phi_B(A^*A)|^2 \leq \phi_B(\mathbf{1}) \cdot \phi_B((A^*A)^2) \\ &= \phi(B^*B) \phi_B((A^*A)^2) = 0. \end{aligned}$$

Now let $\hat{\mathfrak{X}}_\phi$ denote \mathfrak{A}/I , $\hat{\pi}_\phi(A)$ denote the linear operator induced on $\hat{\mathfrak{X}}_\phi$ by L_A , and $\hat{\xi}_\phi$ denote $\mathbf{1} + I \in \hat{\mathfrak{X}}_\phi$. Then we have:

THEOREM III. G. 4 Let ϕ be a positive linear functional on an algebra \mathfrak{A} with involution and identity. Then there exist a strict pre-Hilbert space $\hat{\mathfrak{X}}_\phi$, a morphism $\hat{\pi}_\phi$ from \mathfrak{A} to the algebra of all linear mappings of $\hat{\mathfrak{X}}_\phi$ into itself, and a vector $\hat{\xi}_\phi \in \hat{\mathfrak{X}}_\phi$ such that:

- i) $\phi(A) = (\hat{\xi}_\phi | \hat{\pi}_\phi(A) \hat{\xi}_\phi)$ for all $A \in \mathfrak{A}$.
- ii) $(\hat{\pi}_\phi(A) \xi | \eta) = (\xi | \hat{\pi}_\phi(A^*) \eta)$ for all $A \in \mathfrak{A}$, $\xi, \eta \in \hat{\mathfrak{X}}_\phi$.
- iii) $\hat{\mathfrak{X}}_\phi = \hat{\pi}_\phi(\mathfrak{A}) \hat{\xi}_\phi$.

Finally, these objects are unique up to unitary equivalence, i.e., if \mathcal{X}' , π' and ξ' also satisfy i), ii) and iii), then there exists a mapping U of \mathcal{X}_ϕ onto \mathcal{X}' , preserving inner products, such that

$$U^{-1} \pi'(\cdot) U = \hat{\pi}_\phi(\cdot)$$

$$U \xi_\phi = \xi'.$$

Proof Conditions i), ii), iii) are straightforward verifications. To prove the uniqueness statement, show, again by a simple computation, that

$$\hat{\pi}_\phi(A) \xi_\phi \mapsto \pi'(A) \xi'$$

preserves inner products (and is therefore well-defined) and has the desired algebraic properties.

COROLLARY III. G. 5 Let α be an automorphism of \mathfrak{A} such that $\phi(\alpha A) = \phi(A)$ for all $A \in \mathfrak{A}$. Then there is a uniquely determined unitary operator $U_\phi(\alpha)$ on \mathcal{X}_ϕ such that

$$U_\phi(\alpha) \xi_\phi = \xi_\phi; U_\phi(\alpha) \hat{\pi}_\phi(A) U_\phi(\alpha)^{-1} = \hat{\pi}_\phi(\alpha A) \quad \text{for all } A \in \mathfrak{A}.$$

Proof Apply the uniqueness part of the preceding theorem with $\mathcal{X}'_\phi = \mathcal{X}_\phi$, $\xi_\phi = \xi'$, and $\pi'_\phi \cdot \alpha = \pi'$.

We may regard \mathcal{X}'_ϕ as a dense linear subspace of its completion \mathcal{X}_ϕ , and, if \mathfrak{A} is a Banach algebra, Proposition III. G. 3, assures us that each $\hat{\pi}_\phi(A)$ may be extended to a bounded linear operator $\pi_\phi(A)$ on \mathcal{X}_ϕ . It is worth proving the theorem in this form, allowing for representations by unbounded operators, since it gives the Wightman Reconstruction Theorem in field theory. Let $\mathcal{S}(\mathcal{S})$ be the set of all sequences $(f^{(n)})$, $n = 0, 1, 2, \dots$ such that $f^{(0)}$ is a complex number, $f^{(n)}$ is a smooth, rapidly decreasing complex-valued function on $(\mathbb{R}^4)^n$ and $f^{(n)} = 0$ for all but finitely many n 's. $\mathcal{S}(\mathcal{S})$ is a vector space in an obvious way, and we define

$$(f \cdot g)^{(n)}(x_1, \dots, x_n) = \sum_{j=0}^n f^{(j)}(x_1, \dots, x_j) g^{(n-j)}(x_{j+1}, \dots, x_n)$$

$$(f^*)^{(n)}(x_1, \dots, x_n) = \overline{f(x_n, \dots, x_1)}.$$

These operations make $\mathcal{S}(\mathcal{S})$ into an algebra with involution and identity (but not a Banach algebra). Representations π of $\mathcal{S}(\mathcal{S})$ with reasonable continuity properties may be thought of as specified by an operator-valued distribution $\phi(x)$ by

$$\pi((f)) = \sum_j \int \dots \int dx_1 \dots dx_j f^{(j)}(x_1, \dots, x_n) \phi(x_1) \dots \phi(x_n).$$

Thus, a scalar Wightman field defines a representation of $\mathcal{S}(\mathcal{S})$. Conversely, a set of Wightman fields $\{W_n(x_1, \dots, x_n)\}$ defines a positive linear criteria functions

tional on $\mathcal{S}(\mathcal{S})$ by

$$\phi(f) = \sum_j \int \cdots \int dx_1 \cdots dx_j f^{(j)}(x_1, \dots, x_n) \mathcal{W}_n(x_1, \dots, x_n),$$

so the above theorem enables us to reconstruct a Wightman field given a set of vacuum expectation values.

We now restate the above theorem specialized to the case in which \mathfrak{A} is a Banach algebra with involution and identity, taking advantage of the fact that the $\pi_\phi(A)$'s are automatically continuous.

THEOREM III. G. 5 *Let \mathfrak{A} be a Banach algebra with involution and identity, and let ϕ be a positive linear functional on \mathfrak{A} . Then there exist a Hilbert space \mathcal{X}_ϕ , a representation π_ϕ of \mathfrak{A} on \mathcal{X}_ϕ , and a cyclic vector ξ_ϕ for π_ϕ such that*

$$\phi(A) = (\xi_\phi | \pi_\phi(A) \xi_\phi)$$

for all $A \in \mathfrak{A}$. If \mathcal{X}' , π' , ξ' are another triple of objects satisfying these conditions, there exists a unique unitary operator U mapping \mathcal{X} to \mathcal{X}' such that

$$U\xi_\phi = \xi'; \quad U\pi_\phi(A)U^{-1} = \pi'(A) \quad \text{for all } A \in \mathfrak{A}.$$

If α is an automorphism of \mathfrak{A} such that $\phi(\alpha A) = \phi(A)$ for all $A \in \mathfrak{A}$, there is a unique unitary operator $U_\phi(\alpha)$ on \mathcal{X}_ϕ such that

$$U_\phi(\alpha)\xi_\phi = \xi_\phi; \quad U_\phi(\alpha)\pi_\phi(A)U_\phi(\alpha)^{-1} = \pi_\phi(\alpha A).$$

We will speak of $(\mathcal{X}_\phi, \pi_\phi, \xi_\phi)$ as the *canonical cyclic representation of \mathfrak{A} associated with ϕ* ; the construction of this representation is called the *Gelfand-Segal construction*.

H Pure States and Irreducible Representations

We continue our investigation of the relation between positive linear functionals and representations by determining when the cyclic representation associated with a positive functional is irreducible. Suppose we start with a functional ϕ and construct the cyclic representation $(\mathcal{X}_\phi, \pi_\phi, \xi_\phi)$, and suppose that this representation is not irreducible. Let P be a non-trivial projection commuting with $\pi_\phi(A)$ for all $A \in \mathfrak{A}$. Consider the functional

$$\varrho(A) = (P\xi_\phi | \pi_\phi(A) P\xi_\phi).$$

This is certainly a positive, and

$$\begin{aligned} \phi(A) - \varrho(A) &= (\xi_\phi | \pi_\phi(A) \xi_\phi) - (P\xi_\phi | P\pi_\phi(A) \xi_\phi) = (\xi_\phi | \pi_\phi(A) \xi_\phi) \\ &\quad - (P\xi_\phi | \pi_\phi(A) \xi_\phi) - ((\mathbf{1} - P)\xi_\phi | \pi_\phi(A) \xi_\phi) \\ &= ((\mathbf{1} - P)\xi_\phi | \pi_\phi(A) (\mathbf{1} - P)\xi_\phi), \end{aligned}$$

so $\phi(A) - \varrho(A)$ is again a positive linear functional on \mathfrak{A} . Next, we claim that ϱ is not simply a numerical multiple of ϕ . Suppose the contrary is true, i.e., $\varrho(A) = (\lambda)\phi(A)$ for all $A \in \mathfrak{A}$. Using the fact that P commutes with $\pi_\phi(A)$ for all A we compute that

$$(\pi_\phi(B) \xi_\phi | P\pi_\phi(A) \xi_\phi) = \varrho(B^*A) = \lambda\phi(B^*A) = \lambda \cdot (\pi_\phi(B) \xi_\phi | \pi_\phi(A) \xi_\phi).$$

Since vectors of the form $\pi_\phi(B) \xi_\phi (B \in \mathfrak{A})$ are dense in \mathfrak{H} , this equality implies that $P = \lambda\mathbf{I}$, which contradicts the assumption that P is a non-trivial projection.

We now make two definitions: If ϕ and ϱ are positive linear functionals on an algebra \mathfrak{A} with involution and identity, we say that ϕ *majorizes* ϱ ($\phi \succeq \varrho$) if $\phi - \varrho$ is a positive linear functional. We say that a positive linear functional ϕ is *pure* if the only positive functionals majorized by ϕ are scalar multiples of ϕ . What the above argument shows is that if the representation π_ϕ is not irreducible, then ϕ is not pure.

Let us next try to prove the converse: Let ϱ be a positive linear functional majorized by ϕ . We note that

$$\begin{aligned} |\varrho(B^*A)|^2 &\leq \varrho(B^*B)\varrho(A^*A) \leq \phi(B^*B)\phi(A^*A) \\ &= \|\pi_\phi(B) \xi_\phi\|^2 \cdot \|\pi_\phi(A) \xi_\phi\|^2. \end{aligned}$$

Thus, there exists a unique linear operator T such that

$$(\pi_\phi(B) \xi_\phi | T\pi_\phi(A) \xi_\phi) = \varrho(B^*A).$$

If ϱ is not a scalar multiple of ϕ , then T is not a scalar multiple of \mathbf{I} . Moreover, T is positive

$$((\pi_\phi(A) \xi_\phi | T\pi_\phi(A) \xi_\phi) = \varrho(A^*A) \geq 0),$$

and

$$\|T\| \leq 1. \text{ For } A, B, C \in \mathfrak{A},$$

$$\begin{aligned} (\pi_\phi(B) \xi_\phi | T\pi_\phi(C)\pi_\phi(A) \xi_\phi) &= \varrho(B^*CA) \\ &= (\pi_\phi(C^*)\pi_\phi(B) \xi_\phi | T\pi_\phi(A) \xi_\phi), \end{aligned}$$

so $[T, \pi_\phi(C)] = 0$. Thus, T commutes with $\pi_\phi(\mathfrak{A})$. But if any operator which is not a scalar multiple of \mathbf{I} commutes with $\pi_\phi(\mathfrak{A})$, π_ϕ is not irreducible. We have thus shown that, if the functional ϕ is not pure, the representation π_ϕ is not irreducible. We have, in fact, proved the following proposition.

PROPOSITION III. H. 1 *Let \mathfrak{A} be a Banach algebra with involution and identity, let ϕ be a positive linear functional on \mathfrak{A} , and let $(\mathfrak{H}_\phi, \pi_\phi, \xi_\phi)$ be the associated cyclic representation. Then π_ϕ is irreducible if and only if ϕ is a pure positive functional. Moreover, there is a one-one correspondence between positive functionals on \mathfrak{A} majorized by ϕ and positive operators on \mathfrak{H}_ϕ of norm not greater than one commuting with $\pi_\phi(A)$ for all $A \in \mathfrak{A}$, the correspondence being defined to associate with the operator T the functional*

$$A \mapsto (T\xi_\phi | \pi_\phi(A) \xi_\phi).$$

Suppose, now, we consider states of \mathfrak{A} (i.e., positive linear functionals of norm one) and ask when a state ϕ is pure. We claim that this is the case if and only if ϕ is an extremal point of the set of states of \mathfrak{A} . Suppose first that ϕ is not extremal; then we can write

$$\phi = \frac{1}{2}\phi_1 + \frac{1}{2}\phi_2,$$

where ϕ_1 is a state different from ϕ . Then $\frac{1}{2}\phi_1$ is a positive functional on \mathfrak{A} majorized by ϕ but not a scalar multiple of ϕ , so ϕ is not a pure positive functional. Conversely, suppose ϕ is not a pure positive functional, and let ϱ be a positive functional majorized by ϕ but not a scalar multiple of ϕ . Then:

$$\phi(\cdot) = \varrho(\cdot) + (\phi - \varrho)(\cdot) = \varrho(\mathbf{1}) \cdot \frac{\varrho(\cdot)}{\varrho(\mathbf{1})} + (\phi(\mathbf{1}) - \varrho(\mathbf{1})) \frac{(\phi - \varrho)(\cdot)}{\phi(\mathbf{1}) - \varrho(\mathbf{1})},$$

and since $\frac{\varrho(\cdot)}{\varrho(\mathbf{1})}$ and $\frac{(\phi - \varrho)(\cdot)}{(\phi - \varrho)(\mathbf{1})}$ are states of \mathfrak{A} , we have displayed ϕ as a non-trivial convex combination of states of \mathfrak{A} and hence have shown that ϕ is not an extremal point of the set of states of \mathfrak{A} . We thus have:

COROLLARY III. H. 2 *Let \mathfrak{A} be a Banach algebra with involution and identity, and let ϕ be a state of \mathfrak{A} . Then π_ϕ is irreducible if and only if ϕ is an extremal point of the set of states of \mathfrak{A} .*

Remark The set of states of \mathfrak{A} is a weak- $*$ closed subset of the unit ball of the dual of \mathfrak{A} ; hence, is weak- $*$ compact. If \mathfrak{A} is separable, then the set of states of \mathfrak{A} is metrizable in the weak- $*$ topology. We may therefore use Choquet theory to decompose general states of \mathfrak{A} into integrals of pure states. His decomposition is closely connected with the decomposition of the corresponding cyclic representation as a direct integral of irreducible representations. The decomposition of states into pure states is, however, usually non-unique. It may be shown that the set of states of a C^* algebra \mathfrak{A} with identity is a simplex only if \mathfrak{A} is commutative. (Dixmier, C^*A , Exercise 2.12.17, p. 57).

I Morphisms of C^* Algebras

In this section, we will prove some technical results about morphisms of C^* algebras. Specifically, we will show that any morphism of C^* algebras is norm-decreasing (in particular, continuous), and that any injective morphism is norm preserving. These two results imply that the norm on a C^* algebra is unique.

PROPOSITION III. I. 1 *Let \mathfrak{A} be a Banach algebra with involution, \mathfrak{B} a C^* algebra, and π a morphism from \mathfrak{A} to \mathfrak{B} . Then*

$$\|\pi(A)\| \leq \|A\|$$

for all $A \in \mathfrak{A}$.

Proof We can assume that \mathfrak{A} has an identity (If not, adjoin one). If \mathfrak{A} does not have an identity, or if it does have an identity but it is not sent by π to the identity of \mathfrak{B} , adjoin an identity and extend π to send $\mathbf{1}_{\mathfrak{A}}$ to $\mathbf{1}_{\mathfrak{B}}$. Thus, we can assume that \mathfrak{A} and \mathfrak{B} both have identities and that π sends $\mathbf{1}_{\mathfrak{A}}$ to $\mathbf{1}_{\mathfrak{B}}$. Let $A \in \mathfrak{A}$; we claim that the spectral radius of $\pi(A)$ is not greater than the spectral radius of A :

$$\rho(\pi(A)) \leq \rho(A).$$

Indeed, if $|\lambda| > \rho(A)$, then $\lambda\mathbf{1}_{\mathfrak{A}} - A$ is invertible in \mathfrak{A} ; since π sends identity to identity, $\pi((\lambda\mathbf{1}_{\mathfrak{A}} - A)^{-1})$ is an inverse for $\lambda\mathbf{1}_{\mathfrak{B}} - \pi(A)$, so $\lambda \notin \sigma(\pi(A))$. Now for $A \in \mathfrak{A}$,

$$\|A\|^2 \geq \|(A^*A)\| \geq \rho(A^*A) \geq \rho(\pi(A^*A)).$$

But $\pi(A^*A)$ is a self-adjoint, and hence normal, element of the C^* algebra \mathfrak{B} , so its norm is equal to its spectral radius, i.e.,

$$\rho(\pi(A^*A)) = \|\pi(A^*A)\| = \|\pi(A)^* \pi(A)\| = \|\pi(A)\|^2.$$

Combining these relations gives

$$\|A\|^2 \geq \|\pi(A)\|^2.$$

PROPOSITION III. 1. 2 *Let $\mathfrak{A}, \mathfrak{B}$ be C^* algebras, and let π be an injective morphism of \mathfrak{A} into \mathfrak{B} . Then $\|\pi(A)\| = \|A\|$ for all $A \in \mathfrak{A}$.*

Proof We have already proved this result if $\mathfrak{A}, \mathfrak{B}$ are both commutative. (C^* algebras with identity and if π sends identity to identity (Proposition III. 1. 3) We will reduce the general statement to the commutative case by using relation $\|A^*A\| = \|A\|^2$. We first have to fix the identities. If \mathfrak{A} has an identity, then $\pi(\mathbf{1}_{\mathfrak{A}})$ is an identity for $\overline{\pi(\mathfrak{A})}$; replacing \mathfrak{B} by $\overline{\pi(\mathfrak{A})}$ if necessary, we can assume that \mathfrak{B} has an identity and that π sends identity to identity. Suppose \mathfrak{A} has no identity; then, since π is injective, $\mathbf{1}_{\mathfrak{B}}$ (which it may be necessary to adjoin) cannot belong to $\overline{\pi(\mathfrak{A})}$; then we can extend π to $\tilde{\mathfrak{A}}$ by $\pi(\tilde{\mathbf{1}}_{\mathfrak{A}}) = \mathbf{1}_{\mathfrak{B}}$ without spoiling the injectiveness of π . Thus, we may assume that \mathfrak{A} and \mathfrak{B} have identities and that π sends identity to identity.

Now let A be a self-adjoint element of \mathfrak{A} ; then π defines an injective homomorphism from the commutative C^* sub-algebra of \mathfrak{A} generated by A and $\mathbf{1}_{\mathfrak{A}}$ to the commutative C^* sub-algebra of \mathfrak{B} generated by $\pi(A)$ and $\mathbf{1}_{\mathfrak{B}}$. Since we already know the result for commutative C algebras,

$$\|\pi(A)\| = \|A\|$$

for A self-adjoint. Now let A be as general element of \mathfrak{A} ; then $\|A\|^2 = \|A^*A\| = \|\pi(A^*A)\| = \|\pi(A)\|^2$, so the result is proved.

We can now see that the norm on a C^* algebra is uniquely determined. The crudest form of this assertion is the remark that, if \mathfrak{A} is an algebra with

involution and if $\|\cdot\|, \|\cdot\|'$ are two norms making
 168 \mathfrak{A} into a C^* -algebra, then

O. E. LANFORD III

the identity mapping is a morphism from \mathfrak{A} with $\|\cdot\|$ to \mathfrak{A} with $\|\cdot\|'$ and hence, by the first proposition, $\|A\|' \leq \|A\|$. Interchanging the roles of $\|\cdot\|$ and $\|\cdot\|'$ gives the opposite inequality and hence proves that $\|\cdot\| = \|\cdot\|'$. The second proposition enables us to prove a somewhat more subtle form of the uniqueness of the norm. Let \mathfrak{A} be a C^* algebra with a norm $\|\cdot\|$, and let $\|\cdot\|'$ be another norm making \mathfrak{A} into a normed algebra with involution and satisfying $\|A^*A\|' = (\|A\|')^2$, but with respect to which \mathfrak{A} is not necessarily complete. Then the completion \mathfrak{A}' of \mathfrak{A} with respect to the prime norm is a C^* algebra; the identity mapping is an injective morphism from \mathfrak{A} to \mathfrak{A}' , and is hence norm-preserving by the second proposition. Thus, $\|\cdot\|' = \|\cdot\|$. In other words, if \mathfrak{A} is an algebra with involution, and if it admits one norm making it into a C^* algebra, then it admits no other norm with the right algebraic properties to make it a C^* algebra, i.e., the topology of \mathfrak{A} is built into its algebraic properties.

There is another consequence of the above propositions which is sometimes useful. Let $\pi: \mathfrak{A} \rightarrow \mathfrak{B}$ be a morphism of C^* algebras. Then, since π is continuous (first proposition), $\ker(\pi)$ is closed in \mathfrak{A} , and, by general algebra, $\ker(\pi)$ is a self-adjoint two-sided ideal of \mathfrak{A} . We have already quoted the fact that the quotient of a C^* algebra by a closed, self-adjoint, two-sided ideal is a C^* algebra. π induces an injective morphism $\bar{\pi}: \mathfrak{A}/\ker(\pi) \rightarrow \mathfrak{B}$. By the second proposition, $\bar{\pi}$ is norm-preserving, so its image is complete and hence closed in \mathfrak{B} . But the image of $\bar{\pi}$ is the same as the image of π , i.e., the image of any morphism of C^* algebras is closed.

J Positive Elements of C^* Algebras

In this section, we will show that every C^* algebra is isomorphic to a norm-closed self-adjoint algebra of operators on a Hilbert space. By Proposition III. J. 2, it suffices to show that every C^* algebra admits an injective (i.e., faithful) representation. The main ingredient of the proof is the fact that, if A is any element of a C^* algebra, then A^*A has spectrum contained in the positive real axis. We need first:

LEMMA III. J. 1 Let \mathfrak{A} be an algebra with identity, $A, B \in \mathfrak{A}$. If $\lambda \neq 0$ is in the resolvent set of $A \cdot B$, then λ is in the resolvent set of $B \cdot A$ (i.e., the spectrum of $A \cdot B$ is the same as the spectrum of $B \cdot A$ except possibly for 0).

Proof We can assume $\lambda = 1$ by scaling, say, A . Thus, we have to show that $\mathbf{1} - B \cdot A$ is invertible if $\mathbf{1} - A \cdot B$ is. This we do by exhibiting an explicit formula:

$$(\mathbf{1} - B \cdot A)^{-1} = \mathbf{1} + B(\mathbf{1} - A \cdot B)^{-1} A.$$

It is trivial to verify that this formula is correct; the formula may be remembered by noting that, if $\|A\| < 1, \|B\| < 1$, then

$$\begin{aligned} (\mathbf{1} - B \cdot A)^{-1} &= \sum_{n=0}^{\infty} (B \cdot A)^n = \mathbf{1} + B(\mathbf{1} + A \cdot B + (A \cdot B)^2 + \dots) \cdot A \\ &= \mathbf{1} + B(\mathbf{1} - A \cdot B)^{-1} A. \end{aligned}$$

Now if \mathfrak{A} is a C^* algebra, we will say that an element A of \mathfrak{A} is *positive* ($A \geq 0$) if A is self-adjoint and if the spectrum of A is contained in $[0, \infty)$. The set of positive elements of \mathfrak{A} will be denoted by \mathfrak{A}_+ .

Remark Let A be a self-adjoint element of \mathfrak{A} , and let \mathfrak{B} be the sub C^* algebra of \mathfrak{A} generated by A and $\mathbf{1}$. We claim that A is positive as an element of \mathfrak{A} if and only if it is positive as an element of \mathfrak{B} , and, more generally, that the spectrum of A as an element of \mathfrak{A} is the same as its spectrum as an element of \mathfrak{B} . Certainly, the spectrum of A as an element of \mathfrak{A} is contained in the spectrum of A as an element of \mathfrak{B} , since an inverse for $(\lambda \mathbf{1} - A)$ in \mathfrak{B} is an inverse in \mathfrak{A} . In particular, the spectrum in \mathfrak{A} is contained in the real axis. Let λ be a point of the spectrum of A as an element of \mathfrak{B} . Realizing \mathfrak{B} as an algebra of continuous functions, we see that

$$\lim_{\epsilon \rightarrow 0} \|(\lambda + i\epsilon) \mathbf{1} - A\|^{-1} = \infty.$$

But if λ were not in the spectrum of A as an element of \mathfrak{A} , we would have to have $\lim_{\epsilon \rightarrow 0} \|(\lambda + i\epsilon) \mathbf{1} - A\|^{-1} = (\lambda \mathbf{1} - A)^{-1}$; this is impossible by the above, so λ is in the spectrum of A as an element of \mathfrak{A} , i.e., the two spectra are identical.

PROPOSITION III. J. 2 *Let \mathfrak{A} be a C^* algebra with identity. The set of positive elements of \mathfrak{A} is a convex cone in the set of self-adjoint elements of \mathfrak{A} with $\mathbf{1}$ as an interior point, and $\mathfrak{A}_+ \cap (-\mathfrak{A}_+) = \{0\}$.*

Proof It is clear that, if $\lambda > 0$ and $A \in \mathfrak{A}_+$, then $\lambda A \in \mathfrak{A}_+$. Thus, to prove that \mathfrak{A}_+ is a convex cone, it suffices to show that

$$\{A \in \mathfrak{A}_+ : \|A\| \leq 1\} \text{ is convex.}$$

Now let A be self-adjoint. We can identify the C^* algebra generated by $\mathbf{1}$ and A as the algebra of continuous functions on a compact space. It is then clear that, if $\|A\| \leq 1$, then A is positive if and only if $\|A - \mathbf{1}\| \leq 1$. The set of such A 's is clearly convex, so $\{A \in \mathfrak{A}_+ : \|A\| \leq 1\}$ is convex. Furthermore $\mathbf{1}$ is an interior point of \mathfrak{A}_+ in the set of self-adjoint elements of \mathfrak{A} . Finally, if $A \in \mathfrak{A}_+ \cap (-\mathfrak{A}_+)$, then since $\sigma(A) = \{0\}$ and A is self-adjoint, $A = 0$.

The key technical result is now the following.

PROPOSITION III. J. 3 *Let \mathfrak{A} be a C^* algebra with identity. An element of \mathfrak{A} is positive if and only if it can be written as A^*A , with $A \in \mathfrak{A}$.*

Proof The "only if" is trivial; any positive element of \mathfrak{A} has a positive square root. Thus, what we have to show is that $A^*A \in \mathfrak{A}_+$ for any $A \in \mathfrak{A}$. This is clear if A is self-adjoint. In the general case, by realizing the sub-algebra generated by A^*A and $\mathbf{1}$ as an algebra of continuous functions, we see that we can write

$$A^*A = B + C, \quad \text{with } B, C \in \mathfrak{A}_+, B \cdot C = C \cdot B = 0;$$

we want to show that $C = 0$. We have:

$$(AC)^* AC = C(B - C)C = -C^3 \in -\mathfrak{A}_+.$$

We can also write

$$AC = S + iT, S \text{ and } T \text{ self-adjoint, so}$$

$$(AC)^* (AC) = (S - iT)(S + iT) = S^2 + T^2 + i[S, T].$$

$$(AC)(AC)^* = S^2 + T^2 - i[S, T], \text{ so}$$

$$(AC)(AC)^* - (AC)^*(AC) = 2i[S, T] \in \mathfrak{A}_+, \text{ since}$$

$$(AC)^*(AC) \in -\mathfrak{A}_+; S^2 \in \mathfrak{A}_+, T^2 \in \mathfrak{A}_+, \text{ and } \mathfrak{A}_+ \text{ is a convex cone.}$$

Thus, the spectrum of $(AC)(AC)^*$ is contained in $[0, \infty)$. By Lemma III. J. 1., this implies that the spectrum of $(AC)^*(AC)$ is contained in $\{0, \infty\}$ i.e., that $(AC)^*(AC) \in \mathfrak{A}_+$. Since we already know that $(AC)^*(AC) = -C^3 \in -\mathfrak{A}_+$; we have $-C^3 \in \mathfrak{A}_+ \cap -\mathfrak{A}_+$, so $C^3 = 0$, so $C = 0$.

Once we have this result, we know that a linear functional on \mathfrak{A} is positive if and only if it is positive on \mathfrak{A}_+ . We can therefore prove:

PROPOSITION III. J. 4 *Let \mathfrak{A} be a C^* algebra with identity, \mathfrak{B} a sub- C^* algebra containing $\mathbf{1}$, ϕ a positive linear functional on \mathfrak{B} . Then ϕ may be extended to a positive linear functional $\check{\phi}$ on \mathfrak{A} , and we have $\|\check{\phi}\| = \|\phi\|$. If ϕ is a pure state of \mathfrak{B} , then $\check{\phi}$ may be taken to be a pure state of \mathfrak{A} .*

Proof Consider the set \mathfrak{A}_+ of self-adjoint elements of \mathfrak{A} as a real vector space; \mathfrak{A}_+ is a convex cone in \mathfrak{A}_+ with $\mathbf{1}$ as an interior point. The extension theorem for positive functionals (Sec. I. C.) tells us that the real-linear functional ϕ on $\mathfrak{B} \cap \mathfrak{A}_+$ positive on $\mathfrak{B} \cap \mathfrak{A}_+$ may be extended to a real-linear functional $\check{\phi}$ on \mathfrak{A}_+ positive on \mathfrak{A}_+ . Extend $\check{\phi}$ to a complex-linear functional on \mathfrak{A} by $\check{\phi}(A + iB) = \check{\phi}(A) + i\check{\phi}(B)$. Then $\check{\phi}$ is positive and extends ϕ . Also $\|\check{\phi}\| = \check{\phi}(\mathbf{1}) = \phi(\mathbf{1}) = \|\phi\|$; the middle equality holds because $\mathbf{1} \in \mathfrak{B}$. It remains to prove the last assertion. Let ϕ be a pure state of \mathfrak{B} (i.e., an extremal point of the set of states of \mathfrak{B}), and let $\check{\mathcal{F}}$ be the set of all positive linear functionals on \mathfrak{A} extending ϕ . Then $\check{\mathcal{F}}$ is a non-empty convex subset of the unit ball of the dual of \mathfrak{A} ; moreover, $\check{\mathcal{F}}$ is weak- $*$ closed and therefore weak- $*$ compact. By the Krein-Milman Theorem, $\check{\mathcal{F}}$ has at least one extremal point $\check{\phi}$. We claim that $\check{\phi}$ must be a pure state of \mathfrak{A} . To see this, let $\check{\phi} = \frac{1}{2}(\check{\phi}_1 + \check{\phi}_2)$, where $\check{\phi}_1$ and $\check{\phi}_2$ are states of \mathfrak{A} . Then $\phi_1 = \check{\phi}_1|_{\mathfrak{B}}$ and $\phi_2 = \check{\phi}_2|_{\mathfrak{B}}$ are states of \mathfrak{B} , and $\frac{1}{2}(\phi_1 + \phi_2) = \phi$. But by assumption ϕ is a pure state of \mathfrak{B} , so $\phi_1 = \phi_2 = \phi$. Hence, $\check{\phi}_1$ and $\check{\phi}_2$ are both extensions of ϕ , i.e., $\check{\phi}_1$ and $\check{\phi}_2$ belong to $\check{\mathcal{F}}$. Since $\check{\phi}$ is an extremal point of $\check{\mathcal{F}}$, and since $\frac{1}{2}(\check{\phi}_1 + \check{\phi}_2) = \check{\phi}$, we have $\check{\phi}_1 = \check{\phi}_2 = \check{\phi}$, so $\check{\phi}$ is an extremal point of the set of states of \mathfrak{A} , i.e., a pure state of \mathfrak{A} .

From the above extension theorem, we get:

PROPOSITION III. J. 5 *Let \mathfrak{A} be a C^* algebra with identity, and let $A \in \mathfrak{A}$, $A \neq 0$. Then there is a state ϕ of \mathfrak{A} such that $\phi(A^*A) > 0$.*

Proof Let \mathfrak{B} be the sub- C^* algebra of \mathfrak{A} generated by 1 and A^*A . Since $A^*A \neq 0$; there is a character χ of \mathfrak{B} such that $\chi(A^*A) > 0$. Then χ is a state of \mathfrak{B} ; we may therefore extend χ to a state ϕ of \mathfrak{A} .

Now, finally:

THEOREM III. J. 6 *Let \mathfrak{A} be a C^* algebra. Then \mathfrak{A} is isomorphic to a norm-closed self-adjoint algebra of operators on a Hilbert space.*

Proof By adjoining an identity of necessary, we can assume that \mathfrak{A} has an identity. Since an injective homomorphism of C^* algebras is norm-preserving, we have only to produce a representation π of \mathfrak{A} such that, if $A \neq 0$, $\pi(A) \neq 0$. For each state ϕ of \mathfrak{A} , construct the associated cyclic representation $(\mathcal{H}_\phi, \pi_\phi, \xi_\phi)$, and let $\pi = \bigoplus \pi_\phi$. We claim that π is injective.

Let $A \in \mathfrak{A}$, $A \neq 0$; then by the preceding proposition, $\phi(A^*A) > 0$ for some state ϕ . This implies that $\|\pi_\phi(A)\xi_\phi\|^2 = (\xi_\phi | \pi_\phi(A^*A)\xi_\phi) = \phi(A^*A) > 0$, so $\pi_\phi(A) \neq 0$, so $\pi(A) \neq 0$.

IV Von Neumann Algebras

A Introduction and Preliminaries

We have seen that C^* algebras may be regarded as algebras with involution which are isomorphic to norm closed algebras of bounded operators on Hilbert space. We now want to investigate von Neumann algebras, i.e., self-adjoint algebras of operators on Hilbert space which are closed in the weak operator topology. Unlike C^* algebras, which have many important properties which can be investigated abstractly, i.e., without realizing the algebras concretely on Hilbert space, von Neumann algebras are very closely tied to the Hilbert space on which they act, and their study is based largely on Hilbert-space techniques.

Let \mathcal{H} be a Hilbert space; we are going to define various topologies on the set $\mathcal{L}(\mathcal{H})$ of all bounded operators on \mathcal{H} .

1. The *strong operator topology* on $\mathcal{L}(\mathcal{H})$ is defined by requiring that a net A_α converges to A if and only if, for all $\xi \in \mathcal{H}$, $A_\alpha \xi \rightarrow A\xi$ in the Hilbert space \mathcal{H} . Alternatively, we say that a set $G \subset \mathcal{L}(\mathcal{H})$ is open in the strong operator topology if, for all $A \in G$, there exists a finite set ξ_1, \dots, ξ_n of elements of \mathcal{H} and an $\epsilon > 0$ such that G contains $\{B \in \mathcal{L}(\mathcal{H}) : \|B\xi_i - A\xi_i\| \leq \epsilon \text{ for } 1 \leq i \leq n\}$. Note that this condition imposes no constraints at all on what B does on the orthogonal complement of the subspace generated by ξ_1, \dots, ξ_n , so no non empty strongly open set in $\mathcal{L}(\mathcal{H})$ is bounded in the norm.

2. The *weak operator topology* is defined by requiring that a net A_α converges to A if and only if, for all $\xi, \eta \in \mathcal{X}$, $(\eta | A_\alpha \xi) \rightarrow (\eta | A \xi)$. As before, we can also describe explicitly the open sets: A set $G \subset \mathcal{L}(\mathcal{X})$ is open for the weak operator topology if, for every $A \in G$ there exist $\xi_1, \xi_n, \eta_1, \dots, \eta_n \in \mathcal{X}$ and $\epsilon > 0$ such that $|(\eta_i | B \xi_i) - (\eta_i | A \xi_i)| \leq \epsilon$ for $1 \leq i \leq n$ implies $B \in G$.

It is clear that a net which converges in the norm, or *uniform*, topology on $\mathcal{L}(\mathcal{X})$ converges to the same limit in the strong operator topology, and that a net which converges in the strong operator topology converges to the same limit in the weak operator topology. In other words, the uniform topology is stronger (finer; has more open sets) than the strong operator topology, which in turn is stronger than the weak operator topology.

We can now make the essential definition: A *von Neumann algebra* is a self-adjoint algebra of operators on a Hilbert space, which contains the identity operator and is closed in the weak operator topology. We will see shortly that it is equivalent to require that it be closed in the strong operator topology. Requiring that the algebra be strongly closed is, however, much more restrictive than requiring that it be norm closed. For example: Consider the algebra of all continuous functions on $[0, 1]$, regarded as multiplication operators on \mathcal{L}^2 of Lebesgue measure. This is a norm-closed algebra of operators, but it is not closed in the weak operator topology. Its weak closure is the algebra of all bounded Borel functions modulo functions which are zero almost everywhere, these functions again being regarded as defining multiplication operators on \mathcal{L}^2 .

There are two other topologies on $\mathcal{L}(\mathcal{X})$ whose usefulness is less immediately evident, but which turn out to be very important for technical purposes:

3. The *ultrastrong operator topology* is defined by requiring that a net A_α converges to A if and only if, whenever (ξ_i) is a sequence of elements of \mathcal{X} such that

$$\sum_i \|\xi_i\|^2 < \infty, \quad \sum_i \|A_\alpha \xi_i - A \xi_i\|^2 \rightarrow 0.$$

Roughly speaking, to approximate an operator A in the strong operator topology, one must be able to approximate it on any finite set of vectors simultaneously while to approximate it in the ultrastrong topology one must be able to approximate it on a countable set of vectors simultaneously but the approximation on most of the vectors need not be very good.

4. The *ultraweak operator topology* is similarly defined by requiring that a net A_α converges to A if and only if, whenever (ξ_i) and (η_i) are two sequences of vectors in \mathcal{X} , such that

$$\sum_i \|\xi_i\| \cdot \|\eta_i\| < \infty, \quad \sum_i (\eta_i | A_\alpha \xi_i) \rightarrow \sum_i (\eta_i | A \xi_i).$$

It is nearly obvious that the ultrastrong operator topology is weaker than the uniform topology, but stronger than the strong operator topology,

and also stronger than the ultraweak operator topology, and that the ultraweak operator topology is stronger (i) than the weak operator topology. Note, however, that a *bounded* net A_α which converges to A in the strong operator topology also converges in the ultrastrong operator topology: Let (ξ_i) be any sequence of vectors such that $\sum_i \|\xi_i\|^2 < \infty$, and assume $\|A_\alpha\| \leq M$ for all α . We want to show

$$\lim_n \sum_i \|A_n \xi_i - A \xi_i\|^2 = 0,$$

given that

$$\lim_n \|A_n \xi_i - A \xi_i\| = 0 \quad \text{for each } i.$$

Now

$$\begin{aligned} \lim_n \sup_i \sum_1^n \|A_n \xi_i - A \xi_i\|^2 &\leq \lim_n \sup_i \sum_{i=1}^n \|A_n \xi_i - A \xi_i\|^2 \\ &\quad + \sum_{i=n+1}^\infty 4M^2 \|\xi_i\|^2 = \sum_{i=n+1}^\infty 4M^2 \|\xi_i\|^2. \end{aligned}$$

This is true for all n , so

$$\lim_n \sup_i \sum_1^n \|A_n \xi_i - A \xi_i\|^2 = 0.$$

We may re-express this remark by saying that the strong and ultrastrong topologies on $\mathcal{L}(\mathcal{X})$ agree on bounded sets. A similar argument shows that the weak and ultraweak topologies agree on bounded sets.

There is another way of looking at the ultrastrong and ultraweak topologies which is sometimes useful: We form the direct sum $\bigoplus_{i=1}^\infty \mathcal{X}$ of countably many copies of \mathcal{X} , and we represent $\mathcal{L}(\mathcal{X})$ on $\bigoplus_{i=1}^\infty \mathcal{X}$ by the direct sum of countably many copies of the identity representation. More concretely, given $A \in \mathcal{L}(\mathcal{X})$, we define an operator \tilde{A} on $\bigoplus_{i=1}^\infty \mathcal{X}$ by $\tilde{A}(\xi_i) = (A\xi_i)$. It is nearly trivial to verify that a net A_α converges ultrastrongly (ultra-weakly) to A if and only if \tilde{A}_α converges strongly (weakly) to \tilde{A} . It is also sometimes useful to identify $\bigoplus_{i=1}^\infty \mathcal{X}$ as $\mathcal{X} \otimes l^2$, where l^2 is the space of all sequences $(a_n)_{n=1,2,\dots}$ with $\sum_n |a_n|^2 < \infty$. The identification may be carried out by mapping (ξ_i) to $\sum_{i=1}^\infty \xi_i \otimes l_i$ where $l_i \in l^2$ is the sequence with zeros everywhere except in the i^{th} place where there is a one. With this identification \tilde{A} corresponds to the operator $A \otimes \mathbf{1}$.

The ultrastrong, ultraweak, strong, and weak operator topologies all have some unpleasant features with respect to the algebraic operations. Consider, for example, the mapping $A \mapsto A^*$. This is continuous in the weak operator topology: A net A_α converges to A if and only if $(\eta | A_\alpha \xi) \rightarrow (\eta | A \xi)$ for all $\xi, \eta \in \mathcal{X}$; taking complex conjugates gives $(\xi | A_\alpha^* \eta) \rightarrow (\xi | A^* \eta)$, which

implies that A_n^* converges weakly to A^* . A similar argument shows that $A \mapsto A^*$ is ultraweakly continuous. On the other hand, this mapping is not continuous with respect to the strong or ultrastrong operator topologies: Assume for simplicity that \mathcal{X} is separable and has a complete orthonormal set (ϕ_i) . Define an operator S on \mathcal{X} by $S\phi_1 = 0$; $S\phi_{i+1} = \phi_i$ for $i = 1, 2, 3 \dots$. For any

$$\xi = \sum_i \lambda_i \phi_i \in \mathcal{X},$$

$$S^n \xi = \sum_{i=n+1}^{\infty} \lambda_i \phi_{i-n}, \quad \text{so} \quad \|S^n \xi\|^2 = \sum_{i=n+1}^{\infty} |\lambda_i|^2$$

which goes to zero as n goes to infinity. Thus S^n converges strongly to 0 as n goes to infinity. On the other hand, it is easy to check that S^* is given by $S^* \phi_i = \phi_{i+1}$, $i = 1, 2, 3 \dots$ and hence that

$$\|(S^n)^* \left(\sum_i \lambda_i \phi_i \right)\|^2 = \left\| \sum_i \lambda_i \phi_{i+n} \right\|^2 = \sum_i |\lambda_i|^2,$$

so $(S^n)^*$ does not converge strongly to zero as n goes to infinity. Since the sequence (S^n) is bounded, and since the strong and ultrastrong topologies agree on bounded sets, this same example shows that $A \mapsto A^*$ is not continuous in the ultrastrong topology.

Furthermore, the mapping $(A, B) \mapsto A \cdot B$ is not continuous in any one of the four topologies we are considering (i.e., multiplication is not jointly continuous). This is easy to see for the weak topology: The sequence S^n constructed in the preceding paragraph converges strongly to zero and hence also converges weakly to zero; since taking adjoints is continuous in the weak operator topology, $(S^n)^*$ also converges weakly to zero. But $S^n (S^n)^* = \mathbf{I}$ which certainly does not converge weakly to zero. The argument for the strong operator topology is a bit more subtle. We note first that, if a sequence A_n converges strongly, then $\|A_n \xi\|$ is bounded in n for each ξ , and hence, by the uniform boundedness principle, $\|A_n\|$ is bounded in n . Now let A_n be sequence converging strongly to A , and let B_n be a sequence converging strongly to B ; we will show that $A_n B_n$ converges strongly to $A \cdot B$. Thus, let $\xi \in \mathcal{X}$; then

$$\begin{aligned} \|A_n B_n \xi - A B \xi\| &\leq \|A_n B_n \xi - A_n B \xi\| + \|A_n B \xi - A B \xi\| \\ &\leq \|A_n\| \cdot \|(B_n - B) \xi\| + \|(A_n - A)(B \xi)\| \rightarrow 0. \end{aligned}$$

Nevertheless, the product is not jointly continuous in the strong topology. (And this shows, among other things, that one can make mistakes by arguing about sequences rather than general nets.) Indeed, if \mathcal{W} is any strong neighborhood of zero in $\mathcal{L}(\mathcal{X})$, we will show that there exist $A, B \in \mathcal{W}$ such that $A \cdot B = \mathbf{I}$. By the definition of the strong topology, \mathcal{W} contains a set of the form

$$\{C \in \mathcal{L}(\mathcal{X}) : \|C \zeta_i\| \leq \epsilon \text{ for } 1 \leq i \leq j\}.$$

Here, ζ_1, \dots, ζ_j are elements of \mathcal{X} and $\varepsilon > 0$. Let

$$M = \sup_i \frac{\|\zeta_i\|}{\varepsilon}$$

We again assume for simplicity that \mathcal{X} is separable, and we construct the operator S as above. Since (S^n) converges strongly to zero, we may choose n so that

$$\|S^n \zeta_i\| \leq \frac{\varepsilon}{M} \quad \text{for } i = 1, 2, \dots, j.$$

Let $A = M \cdot S^n$; $B = \frac{1}{M} (S^n)^*$. Then

$$\|A \zeta_i\| = M \cdot \|S^n \zeta_i\| \leq \frac{M \cdot \varepsilon}{M} = \varepsilon \quad \text{for } i = 1, 2, \dots, j,$$

so $A \in W$, and

$$\|B \zeta_i\| = \frac{1}{M} \|(S^n)^* \zeta_i\| = \frac{1}{M} \|\zeta_i\| \leq \varepsilon \quad \text{for } i = 1, 2, \dots, j$$

(by the choice of M), so $B \in W$. But $A \cdot B = S^n (S^n)^* = \mathbf{I}$ as asserted. A similar argument shows that the product is not jointly continuous in the ultrastrong topology.

The pathologies of the operator topologies are somewhat mitigated by the following remarks:

1. The product is continuous in each variable separately in any one of the operator topologies. We will prove this for the weak operator topology: Let A_α be a net converging weakly to A , and let $B \in \mathcal{L}(\mathcal{X})$. Then, for any

$$\xi, \eta \in \mathcal{X}, (\eta | A_\alpha B \xi) \rightarrow (\eta | AB \xi),$$

so $A_\alpha B$ converges weakly to AB , and

$$(\eta | BA_\alpha \xi) = (B^* \eta | A_\alpha \xi) \rightarrow (B^* \eta | A \xi) = (\eta | BA \xi),$$

so BA_α converges weakly to BA .

2. The mapping $(A, B) \mapsto A \cdot B$ is jointly strongly continuous on bounded sets in $\mathcal{L}(\mathcal{X})$. Let A_α, B_α be bounded nets converging strongly to A, B respectively. Then, for any $\xi \in \mathcal{X}$

$$\begin{aligned} \|A_\alpha B_\alpha \xi - AB \xi\| &\leq \|A_\alpha B_\alpha \xi - A_\alpha B \xi\| + \|A_\alpha B \xi - AB \xi\| \\ &\leq \|A_\alpha\| \|B_\alpha \xi - B \xi\| + \|A_\alpha(B \xi) - A(B \xi)\| \rightarrow 0 \end{aligned}$$

so $A_\alpha \cdot B_\alpha$ converges strongly to $A \cdot B$. (Incidentally, we needed only the boundedness of A_α to make the above argument work). Since the strong and ultrastrong topologies agree on bounded sets, multiplication is also jointly continuous on bounded sets. We have, however, already given an example which shows that, even on bounded sets, multiplication is not jointly continuous in the weak operator topology.

We summarize the state of affairs in a table:

Topology	$A \leftrightarrow A^*$	Multiplication
Uniform	Continuous	Jointly continuous
Ultrastrong strong	Not continuous	Separately continuous Not jointly continuous
Ultraweak weak	Continuous	Jointly continuous on bounded sets Separately continuous Not jointly continuous even on bounded sets

It should be remarked that the negative statement made above are correct only if the Hilbert space on which the operators are acting is infinite-dimensional. If the Hilbert space is finite-dimensional, all the operator topologies coincide. We also add a remark about terminology. One commonly says "weak topology", or "strong topology", rather than "weak operator topology" or "strong operator topology". This language is somewhat unfortunate, since, for example, the term "weak topology" ought to be applied to the weak topology of $\mathcal{L}(\mathcal{X})$ as a Banach space (which is not at all the same as the weak operator topology). Nevertheless, it is frequently convenient to use expressions like "weakly continuous" instead of saying "continuous in the weak operator topology". We will use the imprecise terminology freely, except in places where it seems to lead to serious ambiguities.

We will review here a few ideas from operator theory which will be needed later. Let \mathcal{X} be a Hilbert space, \mathcal{X}_1 and \mathcal{X}_2 closed subspaces of \mathcal{X} . An operator U on \mathcal{X} is said to be a *partial isometry with initial subspace \mathcal{X}_1 and terminal subspace \mathcal{X}_2* if:

i) $U\xi = 0$ if ξ is orthogonal to \mathcal{X}_1 .

ii) U , restricted to \mathcal{X}_1 , is an isometry from \mathcal{X}_1 onto \mathcal{X}_2 (i.e., U is unitary from \mathcal{X}_1 to \mathcal{X}_2). If U is a partial isometry with initial subspace \mathcal{X}_1 and terminal subspace \mathcal{X}_2 , then it may be verified that U^* is a partial isometry with initial subspace \mathcal{X}_2 and terminal subspace \mathcal{X}_1 , and that

$$U^*U = P_{\mathcal{X}_1}; \quad UU^* = P_{\mathcal{X}_2}.$$

Conversely, if U is any operator such that $U^*U = P_{\mathcal{X}_1}$, then U is a partial isometry with initial subspace \mathcal{X}_1 (and terminal subspace $U\mathcal{X}_1$).

Now let $A \in \mathcal{L}(\mathcal{X})$; then A^*A is a positive self-adjoint operator, and therefore has a unique positive square-root, which we will denote by $|A|$. For any $\xi \in \mathcal{X}$,

$$\|A\xi\|^2 = (A\xi | A\xi) = (\xi | A^*A\xi) = (\xi | |A|^2 \xi) = \||A| \xi\|^2.$$

Thus, there is a unique unitary operator U_A from the closure of the range of $|A|$ to the closure of the range of A such that

$$U_A |A| = A.$$

We extend U_A to a partial isometry by making it zero on the orthogonal complement of the range of $|A|$. Note, incidentally, that the orthogonal complement of the range of $|A|$ is just the null space of $|A|$, which is the same as the null-space of A . Thus, any bounded operator A may be written uniquely as

$$A = U_A |A|,$$

where $|A|$ is a positive operator vanishing on the null-space of A and U_A is a partial isometry with initial subspace the orthogonal complement of the null-space of A and terminal subspace the closure of the range of A . This way of writing A is called the *polar decomposition* of A . From the uniqueness statement, it follows that, if W is any unitary operator such that $WAW^{-1} = A$, then $W|A|W^{-1} = |A|$ and $WU_AW^{-1} = U_A$, (since $A = (WU_AW^{-1}) \times (W|A|W^{-1})$ is another polar decomposition of A). Thus, any unitary operator commuting with A also commutes with $|A|$ and U_A . It is not necessary for the construction that A map a Hilbert space into itself. If A is a bounded operator from \mathcal{X} to \mathcal{X}' , then A can be written uniquely as $A = U_A |A|$; where $|A|$ is a positive operator on \mathcal{X} and U_A is a partial isometry with initial subspace the orthogonal complement of the null-space of A (in \mathcal{X}) and terminal subspace the closure of the range of A (in \mathcal{X}').

We also need some facts about increasing nets of operators: A net A_α of self-adjoint operators is *increasing* if $A_\alpha \geq A_\beta$ whenever $\alpha \geq \beta$. If $\|A_\alpha\|$ is bounded, then for each $\xi \in \mathcal{X}$, $(\xi | A_\alpha \xi)$ is a bounded increasing net of real numbers and hence converges. The polarization identity gives:

$$(\eta | A_\alpha \xi) = \frac{1}{4} [(\xi + \eta | A_\alpha (\xi + \eta)) - (\xi - \eta | A_\alpha (\xi - \eta))]$$

so $(\eta | A_\alpha \xi)$ converges for all $\xi, \eta \in \mathcal{X}$. The limit is a bounded bilinear form; hence, defines a bounded operator A , and we have shown that A_α converges weakly to A . It may also be shown that A_α converges strongly to A : Let $\xi \in \mathcal{X}$, and consider $\|(A - A_\alpha) \xi\|^2 = (\xi | (A - A_\alpha)^2 \xi)$.

The sesquilinear form $\langle \xi | \eta \rangle = (\xi | (A - A_\alpha) \eta)$ is positive semi-definite, so the Schwarz inequality holds. Inserting $(A - A_\alpha) \xi$ for η , we get

$$\begin{aligned} (\xi | (A - A_\alpha)^2 \xi) &= \langle \xi | (A - A_\alpha) \xi \rangle \overline{\langle (A - A_\alpha) \xi | (A - A_\alpha) \xi \rangle} \\ &= \sqrt{(\xi | (A - A_\alpha) \xi)} \sqrt{(\xi | (A - A_\alpha)^3 \xi)}. \end{aligned}$$

The first term goes to zero and the second term remains bounded, so $\lim \|(A - A_\alpha) \xi\| = 0$, i.e., A_α converges strongly to A .

In particular, let $(P_i)_{i \in I}$ be a family of mutually orthogonal projections, and consider the net of finite partial sums of this family. This is an increasing net of projections, and hence converges strongly to a limit, which may easily be seen to be the projection onto the subspace generated by the union of the ranges of the P_i . We will write this limiting projection as $\sum_{i \in I} P_i$.

More generally, if $(P_i)_{i \in I}$ is any family of projections, not necessarily mutually orthogonal, we consider the net, labelled by the finite subsets

(i_1, \dots, i_n) of I , defined by $P_{(i_1, \dots, i_n)} = P_{i_1} \vee \dots \vee P_{i_n}$, i.e. the projection onto the subspace generated by the union of the ranges of P_{i_1}, \dots, P_{i_n} . This is again an increasing net which converges strongly to the projection onto the subspace generated by the union of the ranges of all the P_i 's; we denote this limiting projection by $\bigvee_{i \in I} P_i$.

B Linear Functionals on Operator Algebras

Before we can begin the study of von Neumann algebras proper, we need a few more technicalities having to do with linear functionals on operator algebras. Since the various operator topologies we have introduced are all weaker than the uniform topology, we do not expect, in general, that all norm-continuous linear functionals will be, say, strongly continuous. We want to see what can be said about special properties of functionals which are continuous in one or another of the operator topologies. The basic fact is that such functionals can always be written as sums (possibly infinite) of functionals of the form $A \mapsto (\xi | A\eta)$ with $\xi, \eta \in \mathcal{X}$.

PROPOSITION IV. B. 1 *Let $(\xi_i), (\eta_i)$ be two sequences of elements of \mathcal{X} such that $\sum_i \|\xi_i\| \cdot \|\eta_i\| < \infty$. Then the mapping $A \mapsto \sum_i (\eta_i | A\xi_i)$ is ultraweakly, and hence ultrastrongly, continuous. Conversely, if \mathfrak{A} is a linear subspace of $\mathcal{L}(\mathcal{X})$, and if ϕ is an ultrastrongly continuous linear functional on \mathfrak{A} , then there exist two sequences $(\xi_i), (\eta_i)$ as above such that*

$$\phi(A) := \sum_i (\eta_i | A\xi_i).$$

In particular, ϕ is ultraweakly continuous, i.e., a linear functional is ultraweakly continuous if and only if it is ultrastrongly continuous. The above statements remain true with "ultrastrong" replaced by "strong", "ultra-weak" replaced by "weak", and infinite sequences replaced by finite ones.

Proof The first assertion is immediate: By the definition of the ultraweak topology, if a net A_α converges ultraweakly to A , then

$$\sum_i (\eta_i | A_\alpha \xi_i) \rightarrow \sum_i (\eta_i | A\xi_i),$$

i.e., the linear functional $\sum_i (\eta_i | A\xi_i)$ is ultraweakly continuous.

Now let ϕ be an ultrastrongly continuous functional. By the definition of the ultrastrong topology, there exists a sequence (ξ_i) in \mathcal{X} such that $\sum_i \|\xi_i\|^2 < \infty$ and such that $\sum_i \|A\xi_i\|^2 < \varepsilon$ implies $|\phi(A)| < 1$. Regard (ξ_i) as an element of $\bigoplus_{i=1}^{\infty} \mathcal{X}$, and consider the linear subspace $\{(A\xi_i) : A \in \mathfrak{A}\}$ of $\bigoplus_{i=1}^{\infty} \mathcal{X}$. On this subspace, the correspondence $(A\xi_i) \mapsto \phi(A)$ is well-defined and continuous, since $\sum_i \|A\xi_i\|^2 < \varepsilon$ implies $|\phi(A)| < 1$.

Extending by continuity, we get a continuous linear functional on the closed linear subspace $\overline{\{(A\xi_i) : A \in \mathfrak{A}\}}$ of $\bigoplus_{i=1}^{\infty} \mathcal{X}$. By elementary Hilbert space theory, a continuous linear functional on a Hilbert space is always given by taking the scalar product with an element of the space, so there exists an element of $\bigoplus_{i=1}^{\infty} \mathcal{X}$, i.e., a sequence (η_i) with $\sum_i \|\eta_i\|^2 < \infty$, such that $\phi(A) = \sum_i (\eta_i | A\xi_i)$ for all $A \in \mathfrak{A}$. This, then, proves the assertions about the ultrastrong-ultraweak topologies; the proofs for the strong-weak topologies are similar.

Remark It follows from the above proposition that the ultraweak topology could have been defined as follows: A net A_α in $\mathcal{L}(\mathcal{X})$ converges to A in the ultraweak topology if and only if $\phi(A_\alpha)$ converges to $\phi(A)$ for all ultrastrongly continuous linear functionals ϕ . In other words, the ultraweak topology is the weakened topology associated with the ultrastrong topology. Similarly, the weak operator topology is the weakened topology associated with the strong operator topology. It is a general fact about topological vector spaces that a locally convex topology and its associated weakened topology have the same closed convex sets (although the initial topology may have many non-convex closed sets which are not closed in the weakened topology). We will prove this result for the case at hand; the proof is valid in general.

PROPOSITION IV. B. 2 *If K is a convex set in $\mathcal{L}(\mathcal{X})$, the strong operator closure of K is the same as its weak operator closure, and the ultrastrong closure of K is the same as its ultraweak closure.*

Proof We will prove only that the strong operator closure of K is the same as the weak operator closure. The weak operator closure of K is a strongly closed set containing K , hence, contains the strong operator closure of K . If we show that the strong operator closure of K is weakly closed, we get the opposite inclusion and thus finish the proof. Replacing K by its strong operator closure, which is again convex, we assume that K is closed in the strong operator topology but not in the weak operator topology, and attempt to derive a contradiction. Let A be in the weak operator closure of K but not in K . Since K is convex, and closed in the strong operator topology, we may apply the Hahn-Banach Theorem to show that A may be separated from K by a strongly continuous linear functional, i.e. there is a strongly continuous linear functional ϕ such that

$$\operatorname{Re} \{ \phi(A) \} > \sup \{ \operatorname{Re} \{ \phi(B) \} : B \in K \}.$$

But since ϕ is strongly continuous, it is also weakly continuous, and the above inequality contradicts the assumption that there is a net A_α in K converging weakly to A .

Remark There do exist convex sets which are ultrastrongly closed but not strongly closed, e.g., the kernel of a linear functional which is ultrastrongly continuous but not strongly continuous.

We next want to see what happens when the functional ϕ is assumed to be positive:

PROPOSITION IV B.3 *Let \mathfrak{A} be a self-adjoint subalgebra of $\mathcal{L}(\mathcal{X})$, containing the identity operator, and let ϕ be an ultrastrongly continuous positive linear functional on \mathfrak{A} . Then there exists a sequence (ζ_i) of vectors of \mathcal{X} such that $\sum_i \|\zeta_i\|^2 < \infty$ and such that*

$$\phi(A) = \sum_i (\zeta_i | A \zeta_i)$$

or all $A \in \mathfrak{A}$. If ϕ is strongly continuous, the sequence of ζ_i 's may be taken to be finite.

Proof We will consider only the ultrastrong topology; the proof of the assertion for the strong operator topology is best obtained by repeating the argument with appropriate modifications. Furthermore, we may replace \mathfrak{A} by its norm closure, i.e., we may assume that \mathfrak{A} is a C^* algebra.

We know two things already:

1. Since ϕ is ultrastrongly continuous, we may write $\phi(A) = (\eta | \bar{A}\xi)$, where $\xi, \eta \in \bigoplus_{i=1}^{\infty} \mathcal{X}$. (We have used the fact that, if $\xi = (\xi_i), \eta = (\eta_i)$, $(\eta | \bar{A}\xi) = \sum_i (\eta_i | A\xi_i)$.)

2. $\phi(A) \geq 0$ for $A \geq 0$.

We want to combine these to show that $\phi(A) = (\zeta | \bar{A}\zeta)$ for some $\zeta \in \bigoplus_{i=1}^{\infty} \mathcal{X}$.

One might hope to prove this simply by being careful in the choice of η , but it turns out to be easier to use a trick. Consider the positive linear functional $\varphi(A) = ((\xi + \eta) | \bar{A}(\xi + \eta))$. Let A be a positive operator in \mathfrak{A} ; then

$$\varphi(A) = (\xi | \bar{A}\xi) + (\eta | \bar{A}\eta) + 2(\eta | \bar{A}\xi) \geq 2\phi(A).$$

Hence, the positive functional φ majorizes ϕ . Consider the cyclic subspace of $\xi + \eta$ in $\bigoplus_{i=1}^{\infty} \mathcal{X}$, i.e., the closure of $\{\bar{A}(\xi + \eta) : A \in \mathfrak{A}\}$, and the representation of \mathfrak{A} defined by restricting each \bar{A} to this cyclic subspace. By the uniqueness of the Gelfand-Segal construction, this representation is unitarily equivalent to the canonical cyclic representation associated with the positive linear functional ϕ . Since φ majorizes ϕ , we know by Proposition III.H.1 that there is a positive operator T on the cyclic subspace, commuting with \bar{A} for each $A \in \mathfrak{A}$, such that

$$\phi(A) = (T(\xi + \eta) | \bar{A}(\xi + \eta)) = (\sqrt{T}(\xi + \eta) | \bar{A} \sqrt{T}(\xi + \eta)).$$

Letting $\zeta = (\zeta_i) = \sqrt{T}(\xi + \eta)$, we have the desired result.

Now let ϕ be a ultrastrongly continuous positive linear functional on a self-adjoint sub-algebra \mathfrak{A} of $\mathcal{L}(\mathcal{X})$ containing the identity operator. We know that we can write:

$$\phi(A) = \sum (\zeta_i | A\zeta_i), \quad \text{where} \quad \sum_i \|\zeta_i\|^2 = \phi(\mathbf{1}) < \infty.$$

Define a linear operator ρ on \mathcal{X} by

$$\rho\xi = \sum_i \zeta_i (\zeta_i | \xi).$$

Then ρ is positive as $(\xi | \rho\xi) = \sum_i |(\zeta_i | \xi)|^2$. Also, if (ξ_1, \dots, ξ_n) is any finite orthonormal set in \mathcal{X} ,

$$\sum_{j=1}^n (\xi_j | \rho\xi_j) = \sum_{j=1}^n \sum_i |(\zeta_i | \xi_j)|^2 \leq \sum_i \|\zeta_i\|^2 = \phi(\mathbf{1}),$$

The bound on $\sum_{j=1}^n (\xi_j | \rho\xi_j)$ which is independent of n implies that the operator ρ is of trace class.

By the properties of (ζ_i) , we have for any complete orthonormal set (ξ_α) and any $A \in \mathfrak{A}$,

$$\begin{aligned} \phi(A) &= \sum_i (\zeta_i | A\zeta_i) = \sum_{\alpha, i} (\xi_\alpha | A\zeta_i) (\zeta_i | \xi_\alpha) \\ &= \sum_{\alpha} (\xi_\alpha | A\rho\xi_\alpha) = \text{Tr}(A\rho). \end{aligned}$$

i.e., $\phi(A) = \text{Tr}(A\rho)$. Thus we have:

PROPOSITION IV.B.4 *Let \mathfrak{A} be a self-adjoint subalgebra of $\mathcal{L}(\mathcal{X})$ containing the identity operator, and let ϕ be a positive ultrastrongly continuous linear functional on \mathfrak{A} . Then there exists a positive linear operator ρ of trace class such that*

$$\phi(A) = \text{Tr}(A\rho)$$

for all $A \in \mathfrak{A}$. Conversely, if ρ is a positive linear operator of trace class, $A \mapsto \text{Tr}(A\rho)$ is an ultrastrongly continuous positive linear functional on $\mathcal{L}(\mathcal{X})$.

It remains only to prove the converse statement. By the spectral theorem, we may write $\rho\xi = \sum \zeta_i (\zeta_i | \xi)$, where the ζ_i are eigenvectors of ρ and the corresponding eigenvalue is $\|\zeta_i\|^2$.

By the finiteness of the trace of ρ , $\sum_i \|\zeta_i\|^2 < \infty$. We have

$$\begin{aligned} \text{Tr}(\rho A) &= \sum_{\alpha} (\xi_\alpha | \rho A\xi_\alpha) = \sum_{\alpha, i} (\xi_\alpha | \zeta_i) (\zeta_i | A\xi_\alpha) \\ &= \sum_i (\zeta_i | A\zeta_i), \end{aligned}$$

(Here (ξ_α) is any complete orthonormal set in \mathcal{X}).

Thus $A \mapsto \text{Tr}(\rho A)$ is a positive ultraweakly continuous linear functional on $\mathcal{L}(\mathcal{X})$. Note incidentally that the ζ_i are mutually orthogonal, so

Proposition IV.B.3 can be refined by adding the requirement that the ζ_i be mutually orthogonal.

We will refer to ϱ as a *density matrix* determining ϕ . It is not, in general, uniquely determined by ϕ (but is uniquely determined if $\mathfrak{A} = \mathcal{L}(\mathcal{X})$).

Note A positive linear functional ϕ on a von Neumann algebra \mathfrak{A} is said to be *normal* if, whenever P_α is an increasing net of projections in \mathfrak{A} (this means $P_\alpha \geq P_\beta$ when $\alpha \geq \beta$) converging to P (i.e., P is the projection onto the closure of $U P_\alpha \mathcal{X}$), we have

$$\phi(P) = \lim_{\alpha} \phi(P_\alpha).$$

Since P_α converges to P in the ultrastrong topology, any ultrastrongly continuous positive linear functional is normal. It turns out that the converse is true: A positive linear functional on a von Neumann algebra is normal if and only if it is ultrastrongly continuous. (We will not give the proof that normality implies ultrastrong continuity; it is sketched in exercise 9, p. 65 of Dixmier AvN.). The usefulness of this result lies in the fact that it is frequently easier to verify that a functional is normal than that it is ultrastrongly continuous.

To summarize: A positive linear functional on a von Neumann algebra is normal if and only if it is ultrastrongly continuous, which is true if and only if it is given by a density matrix.

C. The von Neumann and Kaplansky Density Theorems

We have defined a von Neumann algebra to be a weakly closed self-adjoint algebra of bounded operators on a Hilbert space containing the identity operator. In this section, we give a more algebraic characterization of von Neumann algebras.

Let M be any subset of $\mathcal{L}(\mathcal{X})$; we define the *commutant* of M , written M' , to be the set of elements of $\mathcal{L}(\mathcal{X})$ commuting with every element of M .

PROPOSITION IV.C.1 *For any subset M of $\mathcal{L}(\mathcal{X})$, $(M \cup M^*)'$ is a von Neumann algebra.*

Proof It is clear that $(M \cup M^*)'$ is an algebra of operators, and that the identity operator belongs to $(M \cup M^*)'$. If $A \in (M \cup M^*)'$, then

$$[A^*, B] = ([B^*, A])^* = 0 \quad \text{for all } B \in (M \cup M^*), \quad \text{so } (M \cup M^*)'$$

is self-adjoint. Finally, $(M \cup M^*)'$ is weakly closed. If A_α is a net in $(M \cup M^*)'$ converging weakly to A , and if $B \in M \cup M^*$, then

$$[A, B] = AB - BA = \lim_{\alpha} (A_\alpha B - BA_\alpha) = \lim_{\alpha} 0 = 0,$$

by the separate continuity of multiplication, so $A \in (M \cup M^*)'$.

PROPOSITION IV.C.2 *If M, N are any subsets of $\mathcal{L}(\mathcal{X})$, and if $M \subset N$, then $M' \supset N'$, and $M \subset M''$.*

Proof If $M \subset N$ and $A \in N'$, then A commutes with every element of N ; hence, with every element of M , so $A \in M'$. If $A \in M$, and $B \in M'$, then A commutes with B . Hence, A commutes with every element of M' , i.e., $A \in M''$.

We now come to the crucial:

THEOREM IV.C.3 (*von Neumann Density Theorem, Bicommutant Theorem*). *Let \mathfrak{A} be a self-adjoint subalgebra of $\mathcal{L}(\mathcal{X})$ containing the identity operator. Then the ultrastrong, strong, ultraweak, and weak closures of \mathfrak{A} are all the same, and are equal to \mathfrak{A}'' .*

Thus, a von Neumann algebra could alternatively have been defined as a self-adjoint subalgebra of $\mathcal{L}(\mathcal{X})$ equal to its bicommutant.

Proof We know that \mathfrak{A}'' is a von Neumann algebra, and that $\mathfrak{A}'' \supset \mathfrak{A}$. If we can show that \mathfrak{A} is ultrastrongly dense in \mathfrak{A}'' , then, since \mathfrak{A}'' is ultrastrongly closed, it is equal to the ultrastrong closure of \mathfrak{A} . But \mathfrak{A}'' is also strongly, weakly, and ultraweakly closed, so the ultrastrong closure of \mathfrak{A} coincides with the strong, weak and ultraweak closure, and with \mathfrak{A}'' . The problem, then, is to prove that \mathfrak{A} is ultrastrongly dense in \mathfrak{A}'' , i.e., that, given any $B \in \mathfrak{A}''$, any sequence (ξ_i) in \mathcal{X} such that $\sum_i \|\xi_i\|^2 < \infty$, and any $\epsilon > 0$, there exists $A \in \mathfrak{A}$ such that $\sum_i \|B\xi_i - A\xi_i\|^2 < \epsilon$. We now use the trick of thinking of (ξ_i) as an element ξ of $\bigoplus_{i=1}^{\infty} \mathcal{X}$ and note that what we have to prove is precisely that $\bar{B}\xi$ is in the closed linear span of $\{\bar{A}\xi : A \in \mathfrak{A}\}$. Let P denote the projection from $\bigoplus_{i=1}^{\infty} \mathcal{X}$ onto this closed linear span; then what we have to show is that $P\bar{B}\xi = \bar{B}\xi$. Regard $A \mapsto \bar{A}$ as a representation of \mathfrak{A} ; then the projection P onto the cyclic subspace $\overline{\mathfrak{A}\xi}$ commutes with \bar{A} for every $A \in \mathfrak{A}$ (since the cyclic subspace is invariant). Thus, $P \in \mathfrak{A}$. Let us admit, without proof for the moment, that $\bar{B} \in \mathfrak{A}''$. Then

$$\begin{aligned} P\bar{B}\xi &= \bar{B}P\xi \quad (\text{since } P \in \mathfrak{A}' \text{ and } \bar{B} \in \mathfrak{A}'') \\ &= \bar{B}\xi, \quad (\text{since } \xi \text{ belongs to the cyclic subspace } \overline{\mathfrak{A}\xi}), \end{aligned}$$

so we are through, except for the verification that $\bar{B} \in \mathfrak{A}''$. To carry out this verification, we note that any operator C on $\bigoplus_{i=1}^{\infty} \mathcal{X}_i$ is defined by a matrix (C_{ij}) of bounded operators on \mathcal{X} , so that

$$(C\xi)_i = \sum_j C_{ij}\xi_j.$$

(Warning: The conditions on the matrix (C_{ij}) required to make it the matrix of a bounded operator are very complicated; we do not need to investigate them since we are starting from a bounded operator C). Now $C \in \mathfrak{U}'$ if and only if $\sum_j AC_{ij}\xi_j = \sum_j C_{ij}A\xi_j$ for all i and all (ξ_j) , i.e. if and only if $C_{ij} \in \mathfrak{U}'$ for all i, j . Reversing the above calculation shows that, if $B \in \mathfrak{U}'$, and if $C_{ij} \in \mathfrak{U}'$, for all i, j , then \bar{B} commutes with C i.e., that $\bar{B} \in \mathfrak{U}''$.

We can now read out some consequences. Let A be any element of a von Neumann algebra \mathfrak{U} . We may write $A = A_1 + iA_2$, with A_1 and A_2 self-adjoint. In the proof of the spectral theorem we showed that the spectral projections of any self-adjoint operator are strong limits of polynomials in that operator. In particular, the spectral projections of A_1 and A_2 belong to \mathfrak{U} . But on the other hand, any self-adjoint operator can be approximated arbitrarily well in norm by finite linear combinations of its spectral projections. Conclusion: Let \mathfrak{U} be a von Neumann algebra; then the set of finite linear combinations of projections in \mathfrak{U} is dense in \mathfrak{U} . Thus, an operator A' is in \mathfrak{U}' if and only if it commutes with all the projections in \mathfrak{U} . If P is a projection, then $2P - \mathbf{1}$ is a unitary operator, so we can similarly conclude: An operator A' is in \mathfrak{U}' if and only if it commutes with all the unitary operators in \mathfrak{U} . So far, we have not used the bicommutant theorem. We use it by remarking that, since \mathfrak{U} is a von Neumann algebra, $\mathfrak{U} = \mathfrak{U}''$, so we may interchange the roles of \mathfrak{U} and \mathfrak{U}' , to get:

PROPOSITION IV.C.4 *Let \mathfrak{U} be a von Neumann algebra; then if $B \in \mathcal{L}(\mathcal{X})$ commutes with all unitary operators commuting with all operators in \mathfrak{U} , then $B \in \mathfrak{U}$. In particular, if $A \in \mathfrak{U}$, and if we write the polar decomposition*

$$A = U_A |A|,$$

then $|A|$ and U_A belong to \mathfrak{U} .

It is worth knowing what happens when we take closures of algebras without identity. Thus, let \mathfrak{U} be a self-adjoint subalgebra of $\mathcal{L}(\mathcal{X})$. Regarding \mathfrak{U} as a representation of itself, we may use the general decomposition of a representation of an algebra with identity into a non-degenerate part and a trivial part (Proposition III.F.1) to decompose $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$, where

$$\mathcal{X}_2 = \{\xi \in \mathcal{X} : A\xi = 0 \text{ for all } A \in \mathfrak{U}\}$$

and $\mathcal{X}_1 =$ closed linear span of the union of the ranges of the operators in \mathfrak{U} . Let P be the projection on \mathcal{X}_1 ; then, for all $A \in \mathfrak{U}$, $PA = AP = A$, so $P \in \mathfrak{U}'$. It then turns out that the strong, weak, ultrastrong, and ultraweak closures of \mathfrak{U} all coincide; the closure contains P , and is obtained by restricting all operators in \mathfrak{U} to \mathcal{X}_1 , taking the bicommutant (in $\mathcal{L}(\mathcal{X}_1)$), and then extending by zero. Thus, an arbitrary weakly (or strongly, etc.) closed self-adjoint subalgebra of $\mathcal{L}(\mathcal{X})$ splits into the direct sum of a von Neumann algebra and the zero algebra.

We sketch the proof of the above assertions. It is clear that the closure of \mathfrak{A} in any of the topologies is obtained by restricting to \mathcal{X}_1 , taking the closure in the corresponding topology on $\mathcal{L}(\mathcal{X}_1)$, then extending by zero. Since P restricted to \mathcal{X}_1 is the identity, we have only to prove that P is in the ultrastrong closure of \mathfrak{A} and then to apply the bicommutant theorem to the ultrastrong closure of \mathfrak{A} restricted to \mathcal{X}_1 .

To prove that P is in the ultrastrong closure of \mathfrak{A} , we note that P is the supremum of the projections onto the ranges of the self-adjoint elements of \mathfrak{A} . First, let A be a single self-adjoint operator in \mathfrak{A} . It follows easily from the spectral theorem and the dominated convergence theorem that, if $P_n(z)$ is a sequence of polynomials which is uniformly bounded on the spectrum of A and which converges pointwise to the function equal to 1 for $z \neq 0$ and to zero for $z = 0$, then the sequence of operators $P_n(A)$ is uniformly bounded and converges strongly (hence, ultrastrongly) to the projection onto the range of A . It is clear from the Weierstrass Approximation Theorem that such a sequence of polynomials can be found and that, moreover, they can all be taken to have constant term zero. Then $P_n(A) \in \mathfrak{A}$ for all n , so the projection onto the range of A is in the ultrastrong closure of \mathfrak{A} . Now suppose A_1, A_2 are two self-adjoint elements of \mathfrak{A} and let R_1, R_2 be the projections on to their ranges. Then $R_1 \vee R_2$ is the projection onto the range of $R_1 + R_2$ which, by the above argument, is a strong limit of a bounded sequence of operators which are obtained as polynomials in $R_1 + R_2$ with zero constant term. Thus, $R_1 \vee R_2$ is in the ultrastrong closure of \mathfrak{A} , and it is clear how to extend the argument to show that, if A_1, \dots, A_n is any finite set of elements of \mathfrak{A} , then the projection onto the closed linear span of union of the ranges of A_i is in the ultrastrong closure of \mathfrak{A} . Thus, P is a limit of an increasing net of projections in the ultrastrong closure of \mathfrak{A} and therefore belongs to the ultrastrong closure of \mathfrak{A} .

We next prove a general principle asserting that, to approximate a self-adjoint operator, we can take the approximants to be self-adjoint.

PROPOSITION IV.C.5 *Let E be a vector subspace of $\mathcal{L}(\mathcal{X})$ which contains A^* if it contains A , and let B be a self-adjoint operator in the strong closure of E . Then there is a net of self-adjoint operators in E converging strongly to B .*

Proof Let A_α be a net in E converging strongly to B . Since taking adjoints is continuous in the weak operator topology, A_α^* converges weakly to $B^* = B$. Hence, $\frac{1}{2}(A_\alpha + A_\alpha^*)$ converges weakly to B . In other words, B is in the weak closure of the set of self-adjoint elements of E , and we want to show that it is in the strong closure. We do this by invoking the result that, for a convex set in $\mathcal{L}(\mathcal{X})$, the strong closure is the same as the weak closure (Proposition IV.B.2).

If \mathfrak{A} is a self-adjoint algebra of operators containing $\mathbf{1}$, the von Neumann Density Theorem asserts that, given any $A \in \mathfrak{A}''$, there exists a net A_α in \mathfrak{A} converging strongly to A . It does not, however, rule out the possibility that,

to approximate a given element A of \mathfrak{A}' , it might be necessary to use an unbounded net A_α . This is unfortunate since, if the net A_α is unbounded and converges to A , we cannot be sure that, for example, A_α^2 converges strongly to A^2 . We will now prove a refinement of the von Neumann Density Theorem which eliminates all such problems:

THEOREM IV.C.5 (Kaplansky Density Theorem) *Let \mathfrak{A} be a self-adjoint algebra of operators on a Hilbert space \mathcal{H} , and let $\overline{\mathfrak{A}}$ denote the strong closure of \mathfrak{A} . For any element A of $\overline{\mathfrak{A}}$, there exists a net A_α in \mathfrak{A} such that:*

- i) $\|A_\alpha\| \leq \|A\|$ for all α .
- ii) A_α converges strongly to A .
- iii) A_α^* converges strongly to A^* .

If A is self-adjoint, A_α may be taken to be self-adjoint.

Proof We will first assume that A is self-adjoint. By scaling, we can assume that $\|A\| = 1$. Also, since any self-adjoint element of the norm closure of \mathfrak{A} may be approximated arbitrarily well in norm by self-adjoint elements of \mathfrak{A} of smaller norm, we may assume that \mathfrak{A} is norm-closed, i.e., that \mathfrak{A} is a C^* algebra (but we are not assuming that \mathfrak{A} has an identity).

Consider the function

$$f(t) = \frac{2t}{1+t^2} = \frac{2}{\frac{1}{t} + t}.$$

Its absolute value is nowhere greater than one; it sends zero to zero; and it maps the unit interval monotonically onto itself. Hence there exists a continuous function g on the unit interval such that

$$f(g(t)) = t \quad \text{for } |t| \leq 1.$$

Realizing the C^* algebra generated by A as an algebra of functions, we may construct $B = g(A)$ which is self-adjoint and which satisfies:

$$A = 2B(1 + B^2)^{-1}.$$

Since B is in the C^* algebra generated by A , B is in $\overline{\mathfrak{A}}$, so there is a net B_α of self-adjoint elements of \mathfrak{A} converging strongly to B . The plan is to show that the net $A_\alpha = 2B_\alpha(1 + B_\alpha^2)^{-1}$ of self-adjoint elements of \mathfrak{A} of norm not greater than one converges strongly to A . Thus, we look at:

$$\begin{aligned} A - A_\alpha &= 2B(1 + B^2)^{-1} - 2B_\alpha(1 + B_\alpha^2)^{-1} \\ &= 2(1 + B_\alpha^2)^{-1} [(1 + B_\alpha^2)B - B_\alpha(1 + B^2)](1 + B^2)^{-1} \\ &= 2(1 + B_\alpha^2)^{-1} [B - B_\alpha + B_\alpha^2B - B_\alpha B^2](1 + B^2)^{-1} \\ &= 2(1 + B_\alpha^2)^{-1} [(B - B_\alpha) + B_\alpha(B_\alpha - B) \cdot B](1 + B^2)^{-1} \\ &= 2(1 + B_\alpha^2)^{-1} (B - B_\alpha)(1 + B^2)^{-1} \\ &\quad + 2B_\alpha(1 + B_\alpha^2)^{-1} (B_\alpha - B)B(1 + B^2)^{-1} \end{aligned}$$

Now $(B_n - B)(1 + B^2)^{-1}$ converges strongly to zero, and $\|(1 + B_n^2)^{-1}\| \leq 1$, $\|2B_n(1 + B_n^2)^{-1}\| \leq 1$, so $A - A_n$ converges strongly to zero, as asserted.

That, then, completes the proof if A is self-adjoint. To deal with the general case, we use a trick: Let $\mathfrak{A}_2(\overline{\mathfrak{A}}_2)$ be the self-adjoint algebra of operators on $\mathcal{X} \oplus \mathcal{X}$ given by matrices of the form $\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, with $A_{11}, A_{12}, A_{21}, A_{22}$ in $\mathfrak{A}(\overline{\mathfrak{A}})$. It is easy to see that $(\overline{\mathfrak{A}}_2) = (\overline{\mathfrak{A}}_2)$. Let A be an element of $\overline{\mathfrak{A}}$; then $\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}$ is a self-adjoint element of $\overline{\mathfrak{A}}_2$ with norm equal to $\|A\|$. Hence, there exists a net $\begin{pmatrix} A_{11}^{(\alpha)} & A_{12}^{(\alpha)} \\ A_{21}^{(\alpha)} & A_{22}^{(\alpha)} \end{pmatrix}$ of self-adjoint elements of $\overline{\mathfrak{A}}_2$, with norm not greater than $\|A\|$ converging strongly to $\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}$. But then $\|A_{12}^{(\alpha)}\| \leq \|A\|$ for all α , $(A_{12}^{(\alpha)})^* = A_{21}^{(\alpha)}$, $A_{12}^{(\alpha)}$ converges strongly to A , and $A_{21}^{(\alpha)} = (A_{12}^{(\alpha)})^*$ converges strongly to A^* , proving all the assertions of the theorem.

It is worth remarking that, if \mathcal{X} is separable, we may replace "net" in the above theorem by "sequence". We will prove something slightly stronger: we will show that bounded sets in $\mathcal{L}(\mathcal{X})$ are metrizable in the strong topology if \mathcal{X} is separable.

PROPOSITION IV.C.6 *Let \mathcal{X} be a separable Hilbert space, \mathfrak{B} a bounded set in $\mathcal{L}(\mathcal{X})$. Then the topology induced on \mathfrak{B} by the strong topology on $\mathcal{L}(\mathcal{X})$ is metrizable. More precisely, there exists a norm $||| \cdot |||$ on $\mathcal{L}(\mathcal{X})$ such that a bounded net A_α converges strongly to A if and only if $\lim |||A_\alpha - A||| = 0$.*

Proof Let (ξ_α) be a countable dense set in \mathcal{X} , such that no ξ_α is zero, and define

$$|||A||| = \sum_{\alpha} \frac{1}{2^\alpha} \cdot \frac{\|A\xi_\alpha\|}{\|\xi_\alpha\|}.$$

$||| \cdot |||$ is clearly a norm on $\mathcal{L}(\mathcal{X})$, and $|||A||| \leq \|A\|$. We claim that, if A_α is a bounded net in $\mathcal{L}(\mathcal{X})$, then A_α converges strongly to A if and only if $|||A_\alpha - A||| \rightarrow 0$. (Warning: It is definitely *not* true that a general net A_α converges strongly to A if $|||A_\alpha - A||| \rightarrow 0$; the strong operator topology is not given by a single norm.) By passing to the net $A_\alpha - A$, we can assume $A = 0$. Now let A_α be a net converging strongly to zero, and assume $\|A_\alpha\| \leq M$ for all α . Then

$$\limsup_{\alpha} |||A_\alpha||| \leq \limsup_{\alpha} \left(\sum_{n=1}^N \frac{1}{2^n} \cdot \frac{\|A_\alpha \xi_n\|}{\|\xi_n\|} \right) + \limsup_{\alpha} \sum_{n=N+1}^{\infty} \frac{1}{2^n} \frac{\|A_\alpha \xi_n\|}{\|\xi_n\|}.$$

The first term on the right is zero since $\|A_\alpha \xi_n\| \rightarrow 0$ for all n ; the second term on the right is $\leq \frac{M}{2^N}$ since $\frac{\|A_\alpha \xi_n\|}{\|\xi_n\|} \leq M$ for all n, α .

Thus,

$$\limsup_n \|A_n\| \leq \frac{M}{2^n} \quad \text{for all } N, \text{ so } \limsup_n \|A_n\| = 0.$$

Conversely, let A_n be a net satisfying $\|A_n\| \leq M$ for all n , and suppose $\|A_n\| \rightarrow 0$. We will show that $A_n \rightarrow 0$ strongly. Thus, let $\xi \in \mathcal{H}$ and let $\varepsilon > 0$. Choose n such that

$$\|\xi_n - \xi\| < \varepsilon,$$

and choose α_0 such that

$$\|A_n\| \leq \frac{\varepsilon}{2^n \cdot \|\xi_n\|} \quad \text{if } \alpha \geq \alpha_0.$$

Then, for

$$\alpha \geq \alpha_0, \quad \frac{1}{2^n} \cdot \frac{\|A_n \xi_n\|}{\|\xi_n\|} \leq \|A_n\| \leq \frac{\varepsilon}{2^n \cdot \|\xi_n\|},$$

so

$$\|A_n \xi_n\| \leq \varepsilon.$$

Hence,

$$\|A_n \xi\| \leq \|A_n \xi - A_n \xi_n\| + \|A_n \xi_n\| \leq M \cdot \|\xi - \xi_n\| + \varepsilon = (M + 1)\varepsilon.$$

Since $(M + 1)\varepsilon$ may be made as small as desired, we have:

$$\lim_n \|A_n \xi\| = 0,$$

i.e., A_n converges strongly to zero.

As a corollary, we have:

COROLLARY IV. C.7 *Let \mathfrak{A} be a self-adjoint algebra of operators on a separable Hilbert space, and let A be in the strong closure of \mathfrak{A} . Then there exists a sequence A_n in \mathfrak{A} such that*

- 1) $\|A_n\| \leq \|A\|$ for all n .
- 2) A_n converges strongly to A .
- 3) A_n^* converges strongly to A^* .

Proof By the Kaplansky Density Theorem, we can find for each n an operator A_n in \mathfrak{A} such that $\|A_n\| \leq \|A\|$ and such that

$$\|A_n - A\| \leq \frac{1}{n}; \quad \|A_n^* - A^*\| \leq \frac{1}{n}.$$

Then A_n converges strongly to A and A_n^* converges strongly to A^* , by the preceding proposition.

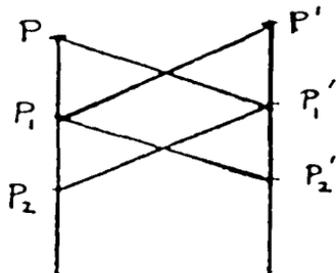
D Comparison of Projections in a von Neumann Algebra

We will begin with some motivation: Let \mathfrak{A} be an algebra with involution, π a representation of \mathfrak{A} on a Hilbert space \mathcal{H} , and let \mathfrak{B} be the von Neumann algebra $\pi(\mathfrak{A})'$. We will study the multiplicity of the representation π . To

begin with the most elementary of considerations, we note that a projection P is in \mathfrak{B} if and only if the subspace $P\mathcal{X}$ is invariant for π , i.e., that the closed invariant subspaces for π are parametrized by the projections in \mathfrak{B} . Let P and P' be two projections in \mathfrak{B} . The representation $\pi(\cdot)|_{P\mathcal{X}}$ is unitarily equivalent to $\pi(\cdot)|_{P'\mathcal{X}}$ if and only if there exists a unitary operator W from $P\mathcal{X}$ to $P'\mathcal{X}$, such that $\pi(\cdot)W = W\pi(\cdot)$ on $P\mathcal{X}$. Extending W to be zero on $(1 - P)\mathcal{X}$, we get a partial isometry with initial subspace $P\mathcal{X}$ and terminal subspace $P'\mathcal{X}$, and this partial isometry commutes with $\pi(A)$ for all A , i.e., belongs to \mathfrak{B} . To summarize: The representation $\pi(\cdot)|_{P\mathcal{X}}$ is unitarily equivalent to $\pi(\cdot)|_{P'\mathcal{X}}$ if and only if there is a partial isometry W in \mathfrak{B} with initial subspace $P\mathcal{X}$ and terminal subspace $P'\mathcal{X}$. We thus make the following definition: Let \mathfrak{B} be a von Neumann algebra, P and P' two projections in \mathfrak{B} . We say that P is *equivalent* to P' ($P \simeq P'$) if there is a partial isometry W in \mathfrak{B} with initial subspace $P\mathcal{X}$ and terminal subspace $P'\mathcal{X}$, i.e., if there is an operator W in \mathfrak{B} such that $W^*W = P$; $WW^* = P'$. (If $\mathfrak{B} = \mathcal{L}(\mathcal{X})$, two projections are equivalent if and only if their ranges have the same dimension.) We extend this definition slightly by defining $P < P'$ or $P' > P$ to mean that P is equivalent to a subprojection of P' (i.e., to a projection in \mathfrak{B} whose range is contained in the range of P'). It is important to distinguish between $P' \geq P$ (P is a subprojection of P') and $P' > P$ (P is equivalent to a subprojection of P'). The relation $P' < P$ clearly defines a pre-order on the set of equivalence classes of projections; the following proposition asserts that it actually defines an order:

PROPOSITION IV. D.1 *Let \mathfrak{B} be a Neumann algebra, P and P' two projections in \mathfrak{B} such that $P > P'$ and $P' > P$. Then $P \simeq P'$.*

Proof We have $P \simeq P_1 \leq P'$ and $P' \simeq P_1 \leq P$. Composing the partial isometry W taking P to P_1 with the partial isometry W' taking P' to P_1 gives partial isometry taking P to a subprojection P_2 of P_1 , i.e., $P \simeq P_2$; Similarly, we get $P' \simeq P_2'$. The partial isometry taking P' to P_1 takes $P' - P_1'$ to $P_1 - P_2$ (i.e., the range of $W'(P' - P_1')$ is $P_1 - P_2$) so $P' - P_1' \simeq P_1 - P_2$, and similarly, $P - P_1 \simeq P_1' - P_2'$. We may represent the situation schematically by a diagram:



Lines sloping downward to the right indicate the action of the partial isometry W taking P to P'_1 ; lines sloping downward to the left indicate the action of the partial isometry W' taking P' to P_1 , and opposite vertical ends of parallelograms represent equivalent projections. We may continue the diagram downward by defining P_3 as the range of $W' P'_2$, P'_3 as the range of WP_2 , etc. We get thus two decreasing sequences of projections:

$$P \geq P_1 \geq P_2 \geq \dots$$

$$P' \geq P'_1 \geq P'_2 \geq \dots$$

such that the range of $WP_i = P'_{i+1}$ and the range of $W'P'_i = P_{i+1}$. Then $P_i - P_{i+1} \approx P'_{i+1} - P'_{i+1} - P_{i+2} - P_{i+2} \approx P'_i - P'_{i+1}$ for $i = 0, 1, 2, \dots$ (We let $P_0 = P$; $P'_0 = P'$). Furthermore, let

$$P_\infty = iP_i \quad \text{and} \quad P'_\infty = iP'_i = iP'_{i+1}.$$

Since the range of WP_i is P'_{i+1} , the range of $W \cdot P_\infty$ is exactly P'_∞ , i.e., P_∞ is equivalent to P'_∞ . We may now write:

$$\begin{aligned} P &= P_\infty + (P - P_1) + (P_1 - P_2) + (P_2 - P_3) + (P_3 - P_4) + \dots \\ &\approx \quad \quad \approx \quad \quad \approx \quad \quad \approx \quad \quad \approx \\ P' &= P'_\infty + (P'_1 - P'_2) + (P'_2 - P'_3) + (P'_3 - P'_4) + (P'_4 - P'_5) + \dots \end{aligned}$$

A simple argument (which we omit) shows that P is equivalent to P' .

E Disjointness of Projections and Central Projections

We have indicated that our investigation of the structure of von Neumann algebras will be motivated by regarding the von Neumann algebra \mathcal{B} in question as the commutant of a representation π of some algebra \mathfrak{A} , and translating questions about subrepresentations of π into questions about the von Neumann algebra \mathcal{B} . Let us elaborate on this approach a little by asking what the situation is if \mathfrak{A} has the property that every representation may be written as a direct sum of irreducible representations. (This will be true, for example, if \mathfrak{A} is finite-dimensional). In this case, every projection P in \mathcal{B} can be written as a sum of mutually orthogonal projections onto irreducible subspaces, and two projections P and P' are equivalent if and only if each irreducible representation of \mathfrak{A} appears the same number of times in $\pi(\cdot) | P\mathcal{X}$ as in $\pi(\cdot) | P'\mathcal{X}$. Thus, the structure of the set of equivalence classes of projections is entirely transparent. If we drop the requirement that every representation of \mathfrak{A} decompose into a direct sum of irreducible representations, the situation becomes much more intricate. For example, the algebra of continuous functions on $[0, 1]$ represented

multiplication operators on \mathcal{L}^2 of Lebesgue measure on $[0, 1]$, has no irreducible subspaces. Much more peculiar things can also happen, e.g., there exist C^* algebras \mathfrak{A} and representations π of \mathfrak{A} which are not irreducible, but which have the property that any two non-zero subrepresentations of \mathfrak{A} are unitarily equivalent. (In fact, this is what occurs in quantum statistical mechanics for the cyclic representation of the algebra of quasilocal observables associated with an equilibrium state representing a pure phase at non-zero temperature.)

We will approach the general situation in the following way: We look for properties of subrepresentations of π which have a straightforward interpretation if π can be decomposed into a direct sum of irreducible representations, but which can be formulated without mentioning irreducible representations. We then translate these properties into properties of projections in the von Neumann algebra $\mathfrak{B} = \pi(\mathfrak{A})'$, and make them into definitions which no longer depend on regarding \mathfrak{B} as the commutant of a representation. Of course, there is no way to see at the outset which properties of subrepresentations of π can fruitfully be translated into definitions in general von Neumann algebras...

We have already seen one example of this procedure: Unitary equivalence of subrepresentations of π translates into the definition of equivalence of projections in a von Neumann algebra. As a second example: If every representation of \mathfrak{A} decomposes into a direct sum of irreducible representations, we will say that two subrepresentations of π are disjoint if no irreducible representation appearing in one of them also appears in the other. This we translate into: Two projections P and P' in a von Neumann algebra \mathfrak{B} are disjoint if no non-zero subprojection of P is equivalent to a subprojection of P' . Conversely, we say that P covers Q ($P \succ \succ Q$ or $Q \prec \prec P$) if no non-zero subprojection of Q is disjoint from P (Intuitively, if every irreducible contained in $\pi(\cdot) | Q\mathcal{H}$ is contained in $\pi(\cdot) | P\mathcal{H}$) and that P is quasi-equivalent to Q ($P \sim Q$) if P covers Q and Q covers P . One can thus think of quasi-equivalent projections as corresponding to subrepresentations which are the same except for multiplicity.

The key to analyzing the notion of disjointness is contained in the following proposition, which is really just a variant of Schur's Lemma.

PROPOSITION IV. E.1 *Let P and Q be projections in a von Neumann algebra \mathfrak{B} . In order that P be disjoint from Q it is necessary and sufficient that $PBQ = 0$ for all $B \in \mathfrak{B}$. In particular, if P is disjoint from Q , then $PQ = 0$, i.e., P is orthogonal to Q .*

Proof Assume first that P is not disjoint from Q , and let W be a non-zero partial isometry in \mathfrak{B} with initial domain contained in $Q\mathcal{H}$ and terminal domain contained in $P\mathcal{H}$. Then $W = PIWQ$, so $PIWQ \neq 0$, i.e., W is an element of \mathfrak{B} such that $PIWQ \neq 0$. Next, assume that, for some $B \in \mathfrak{B}$, $PBQ \neq 0$. Then $PBQ \in \mathfrak{B}$ and we may use the polar decomposition to write

$PBQ = W \cdot |PBQ|$ where W is a partial isometry with initial subspace the orthogonal complement of the null-space of PBQ (which is contained in $Q\mathcal{X}$ since $PBQ = 0$ on $(1 - Q)\mathcal{X}$) and terminal subspace the closure of the range of PBQ (which is contained in $P\mathcal{X}$). Moreover, W is in \mathcal{A} since it is the partial isometric part of PBQ which is in \mathcal{A} . Hence, Q is not disjoint from P (since the initial subspace of W is equivalent to the terminal subspace of W).

In any algebra the center of \mathcal{A} , denoted by $\mathfrak{Z}(\mathcal{A})$, is the set of elements of \mathcal{A} commuting with all elements of \mathcal{A} . If \mathcal{A} is a von Neumann algebra, we may equivalently define $\mathfrak{Z}(\mathcal{A}) = \mathcal{A} \cap \mathcal{A}'$. Thus $\mathfrak{Z}(\mathcal{A})$ is a commutative von Neumann algebra contained in \mathcal{A} .

COROLLARY IV. E.2 *Let \mathcal{A} be a Neumann algebra, P a projection in \mathcal{A} . Then P is disjoint from $1 - P$ if and only if $P \in \mathfrak{Z}(\mathcal{A})$.*

Proof Suppose first that $P \in \mathfrak{Z}(\mathcal{A})$. Then, for all $B \in \mathcal{A}$, $PB(1 - P) = B \cdot P \cdot (1 - P) = 0$, so P is disjoint from $1 - P$. Conversely, if P is disjoint from $1 - P$, then $0 = P \cdot B \cdot (1 - P) = PB - PBP$ i.e., $PB = PBP$ for all $B \in \mathcal{A}$. Taking adjoints gives

$$B^*P = PB^*P = PB^*$$

for all $B \in \mathcal{A}$, so $B^*P = PB^*$ for all $B \in \mathcal{A}$, so $P \in \mathfrak{Z}(\mathcal{A})$.

COROLLARY IV. E.3 *Let \mathcal{A} be a von Neumann algebra, P a projection in \mathcal{A} . Let $(Q_i)_{i \in I}$ be a family of projections in \mathcal{A} each of which is disjoint from P . Then $\bigvee_i Q_i$ is disjoint from P .*

Proof Since $Q_iBP = 0$ for all $B \in \mathcal{A}$, $(\bigvee_i Q_i)BP = 0$ for all $B \in \mathcal{A}$, so $\bigvee_i Q_i$ is disjoint from P .

PROPOSITION IV. E.4 *Let P be a projection in a von Neumann algebra \mathcal{A} . Then there exists a unique projection $Q \in \mathcal{A}$, $Q \geq P$, such that P is quasi-equivalent to Q and disjoint from $1 - Q$. Furthermore, $Q \in \mathfrak{Z}(\mathcal{A})$.*

Proof Intuitively, we may think of Q as the projection onto the subspace generated by all irreducible subspaces equivalent to an irreducible subspace contained in P , and $1 - Q$ as the projection onto the subspace generated by all irreducible subspaces which are not equivalent to an irreducible subspace contained in P . The latter description gives the key to constructing Q ; we define $1 - Q$ to be the supremum of all projections in \mathcal{A} disjoint from P . Then $1 - Q$ is disjoint from P , and thus in particular orthogonal to P , by the preceding proposition. By the construction of Q , no non-zero subprojection of Q can be disjoint from P (otherwise, this projection would be in $1 - Q$), so P covers Q . On the other hand $Q \geq P$, so Q covers P and thus Q is quasi-equivalent to P . This proves the existence of Q . Moreover, Q is

disjoint from $\mathbf{1} - Q$, so $Q \in \mathfrak{B}(\mathfrak{A})$. To prove the uniqueness, let Q' be another projection such that P is quasi-equivalent to Q' and disjoint from $\mathbf{1} - Q'$. Then Q' must also belong to $\mathfrak{X}(\mathfrak{A})$, so $Q'(\mathbf{1} - Q)$ is a projection. Now $Q'(\mathbf{1} - Q)$ is covered by P (since it is a subprojection of Q') and disjoint from P (since it is a subprojection of $\mathbf{1} - Q$.) The only possibility, then, is that $Q'(\mathbf{1} - Q) = 0$, i.e., $Q' = Q'Q$. Interchanging the roles of Q and Q' we see that $Q = QQ'$, so $Q = Q'$ and uniqueness is proved.

The projection Q constructed in the preceding proposition is called the *central support* of P . There are some other ways to characterize it:

PROPOSITION IV. E.5 *Let P be a projection in the von Neumann algebra \mathfrak{A} , and let Q be its central support. Then if Q' is any projection in $\mathfrak{X}(\mathfrak{A})$ containing P , $Q' \geq Q$ (i.e., Q is the smallest projection in $\mathfrak{X}(\mathfrak{A})$ containing P). Q is also equal to the projection onto the closed linear span of the set of vectors obtained by applying elements of \mathfrak{A} to elements of $P\mathfrak{X}$ (we denote this closed linear span by $[\mathfrak{A}P\mathfrak{X}]$).*

Proof. Let Q' be a projection in $\mathfrak{X}(\mathfrak{A})$ which contains P . Then QQ' is a central projection containing P and contained in Q ; hence it is quasi-equivalent to Q (we have $Q \succ \succ QQ' \geq P \succ \succ Q$). On the other hand $Q - QQ' \leq \mathbf{1} - QQ'$ and is therefore disjoint from QQ' . Thus we have $Q - QQ'$ is covered by QQ' and on the other hand $Q - QQ'$ is disjoint from QQ' , so we must have $Q - QQ' = 0$, i.e. $Q = QQ'$, i.e. $Q' \geq Q$.

Now let P' denote the projection onto $[\mathfrak{A}P\mathfrak{X}]$. Since $[\mathfrak{A}P\mathfrak{X}]$ is invariant for \mathfrak{A} , $P' \in \mathfrak{B}'$. If Q is the central support of P , if $B \in \mathfrak{B}$, and if $\xi \in \mathfrak{X}$, then $QB P \xi = BQP \xi$ (since $Q \in \mathfrak{X}(\mathfrak{A}) = BP \xi$ (since $Q \geq P$); hence $Q \geq P'$. To prove that $Q = P'$, we have thus only to show that $P' \in \mathfrak{X}(\mathfrak{A})$, and we already know that $P' \in \mathfrak{B}'$. Thus, what we have to show is that $[\mathfrak{A}P\mathfrak{X}]$ is invariant for \mathfrak{B}' (since this will imply $P' \in (\mathfrak{B}')' = \mathfrak{A}$.) But if $B' \in \mathfrak{B}'$, $B \in \mathfrak{B}$, and $\xi \in P\mathfrak{X}$, $B'B\xi = BB'\xi \in [\mathfrak{A}P\mathfrak{X}]$ since $B'P\mathfrak{X} \subset P\mathfrak{X}$. Thus, B' maps $[\mathfrak{A}P\mathfrak{X}]$ into itself, so $P' \in \mathfrak{A}$, so $P' = Q$.

COROLLARY IV. E.6 *Let P be a projection in a von Neumann algebra \mathfrak{A} , and let Q be the central support of P . Then, for $B' \in \mathfrak{B}'$, $B' \upharpoonright P\mathfrak{X} = 0$ implies $B' \upharpoonright Q\mathfrak{X} = 0$; i.e., the mapping $B' \upharpoonright Q\mathfrak{X} \leftrightarrow B' \upharpoonright P\mathfrak{X}$ defines an isomorphism of the algebra $B' \upharpoonright Q\mathfrak{X}$ to the algebra $B' \upharpoonright P\mathfrak{X}$.*

Proof. If $B' \upharpoonright P\mathfrak{X} = 0$, then, for any $B \in \mathfrak{A}$, $B'BP\mathfrak{X} = BB'P\mathfrak{X} = 0$, so $B' \upharpoonright Q\mathfrak{X} = 0$.

We can now analyze more or less completely the structure of the set of quasi-equivalence classes of projections.

- 1) Every projection is quasi-equivalent to a central projection, its central support.

- 2) If Q, Q' are central projections, then Q covers Q' if and only if $Q \geq Q'$ (and hence Q is quasi-equivalent to Q' if and only if $Q = Q'$).

Proof Suppose Q covers Q' ; then since $(\mathbf{1} - Q) \cdot Q'$ is contained in Q' and disjoint from Q , it must be zero, so $Q' = QQ'$, i.e., $Q \geq Q'$. Conversely, if $Q \geq Q'$, the Q surely covers Q' .

3) Combining 1) and 2) we see that every quasi-equivalence class of projections contains exactly one central projection, i.e., the quasi-equivalence classes of projections are parametrized by the central projections and the relation $\succ \succ$ just translates into the relation \geq .

4) If P_1, P_2 are projections in \mathcal{B} , then P_1 covers P_2 if and only if, for $B' \in \mathcal{B}'$, $B' | P_1 \mathcal{X} = 0$ implies $B' | P_2 \mathcal{X} = 0$, and P_1 is quasi-equivalent to P_2 if and only if the mapping $B' | P_1 \mathcal{X} \leftrightarrow B' | P_2 \mathcal{X}$ defines an isomorphism from $\mathcal{B}' P_1 | \mathcal{X}$ to $\mathcal{B}' P_2 | \mathcal{X}$.

From remark 4), it is a short step, but not a trivial one, to the following characterization of quasi-equivalence: Let \mathfrak{A} be an algebra, π a representation of \mathfrak{A} , $\mathcal{B} = \pi(\mathfrak{A})$, P, P' two projections in \mathcal{B} . Then P is quasi-equivalent to P' if and only if there exists an isomorphism ϕ from the von Neumann algebra on $P\mathcal{X}$ generated by $\pi(\mathfrak{A}) | P\mathcal{X}$ to the von Neumann algebra on $P'\mathcal{X}$ generated by $\pi(\mathfrak{A}) | P'\mathcal{X}$, such that $\phi(\pi(A) | P\mathcal{X}) = \pi(A) | P'\mathcal{X}$ for all $A \in \mathfrak{A}$. To prove the "only if" statement from remark 4), we have only to show that the von Neumann algebra generated by $\pi(\mathfrak{A}) | P\mathcal{X}$ is just $\mathcal{B}' | P\mathcal{X}$, and this follows easily from the bicommutant theorem. To prove the "if" statement, we have to show that if ϕ is an isomorphism from $\mathcal{B}' | P\mathcal{X}$ to $\mathcal{B}' | P'\mathcal{X}$ extending the mapping $\pi(A) | P\mathcal{X} \leftrightarrow \pi(A) | P'\mathcal{X}$, then ϕ must be given the same formula on all of $\mathcal{B}' | P\mathcal{X}$. This follows easily from the fact that an isomorphism of von Neumann algebras is ultraweakly continuous, which we will not prove (See Dixmier, *Av N.*, p. 57, Corollaire 1.)

There are some remarks to be made about the significance of the center $\mathcal{Z}(\mathcal{B})$ of a von Neumann algebra \mathcal{B} . First let P be a projection in \mathcal{B}' . Then, since $P\mathcal{X}$ and $(\mathbf{1} - P)\mathcal{X}$ are both invariant for \mathcal{B} , we may decompose $\mathcal{X} = P\mathcal{X} \oplus (\mathbf{1} - P)\mathcal{X}$ and get an analogous decomposition for every element of \mathcal{B} : $B = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}$. We can now ask what will be special about this decomposition if P is in $\mathcal{Z}(\mathcal{B})$, i.e., if it is in \mathcal{B} as well as in \mathcal{B}' . It is not hard to see that P is in $\mathcal{Z}(\mathcal{B})$ if and only if the entries B_1 and B_2 in matrices $\begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}$ representing elements of \mathcal{B} are independent. By "independent" we mean that, given any two elements B', B'' in \mathcal{B} , we can find an element B of \mathcal{B} such that

$$B | P\mathcal{X} = B' | P\mathcal{X}$$

$$B | (\mathbf{1} - P)\mathcal{X} = B'' | (\mathbf{1} - P)\mathcal{X}.$$

In other words, if P is in the center of \mathcal{B} , then \mathcal{B} is isomorphic to $\mathcal{B} \mid P\mathcal{X} \oplus \mathcal{B} \mid (\mathbb{I} - P)\mathcal{X}$. Thus, we can study the von Neumann algebras $\mathcal{B} \mid P\mathcal{X}$ and $\mathcal{B} \mid (\mathbb{I} - P)\mathcal{X}$ separately, then reconstruct \mathcal{B} from them. The algebras $\mathcal{B} \mid P\mathcal{X}$ and $\mathcal{B} \mid (\mathbb{I} - P)\mathcal{X}$ are smaller than \mathcal{B} , and ought to be less complicated. In particular, if it happens that P and $\mathbb{I} - P$ are the only non-trivial projections in $\mathcal{L}(\mathcal{B})$, then it is easy to see that the centers of $\mathcal{B} \mid P\mathcal{X}$ and $\mathcal{B} \mid (\mathbb{I} - P)\mathcal{X}$ are both trivial, i.e., consist only of the scalars. A von Neumann algebra \mathcal{B} such that $\mathcal{L}(\mathcal{B}) = \{\lambda\mathbb{I}\}$ is called a *factor*, so what we have just shown is that, if \mathcal{B} is a von Neumann algebra whose center is generated by a single non-trivial projection (i.e., whose center is two-dimensional), then \mathcal{B} may be decomposed into a direct sum of two factors. Evidently, this argument extends trivially to any von Neumann algebra with finite-dimensional center (if $\mathcal{L}(\mathcal{B})$ is n dimensional, \mathcal{B} may be written as a direct sum of n factors) and even to a Neumann algebra whose center is generated by a countable set of mutually orthogonal projections. The center of \mathcal{B} need not have this property, however, and it is therefore necessary to pass from direct sums to direct integrals. Very crudely speaking, the general situation (subject to certain technical restrictions) is the following: Starting from a Hilbert space \mathcal{X} and a von Neumann algebra \mathcal{B} on \mathcal{X} , one can find a measure space (X, μ) , and, for each $x \in X$, a Hilbert space \mathcal{X}_x and a factor \mathcal{B}_x on \mathcal{X}_x such that:

1) The Hilbert space \mathcal{X} may be realized as the space of "all" functions $x \mapsto \xi_x \in \mathcal{X}_x$ such that $\int d\mu(x) \|\xi_x\|^2 < \infty$. The word "all" is in quotation marks since we really only want to consider only functions which are in some sense measurable; the way in which "measurable" should be defined is not at all obvious since the Hilbert spaces \mathcal{X}_x for different x are to be thought of as different spaces.

2) In this realization, the elements of \mathcal{B} are in one-one correspondence with functions $x \mapsto B_x \in \mathcal{B}_x$, where these functions are subjected to the condition $\sup_x \|B_x\| < \infty$ and to some measurability condition; the correspondence is defined by $(B\xi)_x = B_x\xi_x$.

The above theory, known as reduction theory, in some sense reduces the theory of general von Neumann algebras to the theory of factors. It has however two disadvantages: It is valid only under some restrictions on the size of the von Neumann algebra (these restrictions are not too serious for physical applications since they are always satisfied if the space on which \mathcal{B} acts is separable), and applying it frequently leads to unpleasant problems of measurability. For many purposes, therefore, it is both simpler and clearer to make arguments directly about \mathcal{B} itself rather than invoking the decomposition of \mathcal{B} into factors. Nevertheless, thinking of a general von Neumann algebra as some kind of generalized direct sum of factors is frequently an invaluable heuristic device.

F The Comparability Theorem

We now want to get some idea of what the collection of equivalence classes of projections in \mathcal{B} looks like as a partially ordered set. For orientation, we go back to the pictures of \mathcal{B} as the commutant of a representation π of a C^* algebra \mathfrak{A} and we assume, for heuristic purposes, that every representation of \mathfrak{A} may be written as a direct sum of irreducible representations. Let I denote the set of equivalence classes of irreducible representations of \mathfrak{A} . For each projection $P \in \mathcal{B} = \pi(\mathfrak{A})'$, we associate an indexed set $(n_i)_{i \in I}$ of cardinal numbers, n_i being the number of times the representation i appears in a direct sum decomposition of $\pi(\cdot) | P\mathcal{X}$. If P' is another projection, and (n'_i) is its set of multiplicities, then $P' \succ P$ if and only if $n'_i \geq n_i$ for each i . Hence, the relation \succ does not linearly order the set of projections in \mathcal{B} , unless π is a direct sum of copies of a single irreducible representation. Moreover, in the simplified case we are considering, π is a direct sum of copies of a single irreducible representation if and only if no two non-zero projections in \mathcal{B} are disjoint, i.e., if and only if the only central projections in \mathcal{B} are 0 and 1, i.e. (since $\mathcal{Z}(\mathcal{B})$ is generated by its projections) if and only if \mathcal{B} is a factor. We are thus led to the general conjecture that, if \mathcal{B} is a factor, the set of projections in \mathcal{B} is linearly ordered by \succ . In this conjecture of course, all mention of irreducible representations has disappeared. To prove the conjecture, we will prove the following slightly stronger result, which is also easy to interpret in the simplified case considered above:

THEOREM IV. F.1 *Let \mathcal{B} be a von Neumann algebra, P and Q two projections in \mathcal{B} . There exists a central projection R such that $RP \succ RQ$; $(\mathbf{1} - R)Q \succ (\mathbf{1} - R)P$. In particular, if \mathcal{B} is a factor (so either $R = \mathbf{1}$ or $\mathbf{1} - R = \mathbf{1}$), we have either $P \succ Q$ or $Q \succ P$. In other words, any two projections in a factor are comparable.*

Proof The idea of the proof is simple: We first find projections $S \leq P$, $T \leq Q$ which are equivalent and as large as possible; we then see what can be done with what is left over. To carry out the first step, we form the set whose elements are sets $\{(S_i, T_i)\}$ of pairs of non-zero projections in \mathcal{B} such that

- 1) $S_i \leq P$, $T_i \leq Q$ for all i .
- 2) $S_i \simeq T_i$ for all i .
- 3) If $i \neq j$, S_i is orthogonal to S_j and T_i is orthogonal to T_j .

Since the elements of this set are sets (of pairs of projections), we can order the set by inclusion: $\{(S_i, T_i)\} \leq \{(S'_k, T'_k)\}$ means that each (S_i, T_i) is equal to (S'_k, T'_k) for some k . It follows at once from the definitions that the union of a linearly ordered family of such sets of pairs of projections is again such a set of pairs of projections; hence by Zorn's Lemma, there exists a maximal

such set $\{(S_i, T_i)\}$. Let $S = \sum_i S_i$; $T = \sum_i T_i$. Then $S \leq P$, $T \leq Q$, and $S \simeq T$. Furthermore, $P - S$ is disjoint from $Q - T$; (If this were not true, there would exist a non-zero projection $S_0 \leq P - S$ which is equivalent to $T_0 \leq Q - T$; then $\{(S_i, T_i)\} \cup \{(S_0, T_0)\}$ properly contains $\{(S_i, T_i)\}_{i \in I}$ and therefore violates maximality.) In a factor, no two non-zero projections can be disjoint, and this implies that either $P - S = 0$ (i.e., $P = S \simeq T \leq Q$, so $P < Q$) or $T - Q = 0$ (i.e., $Q = T \simeq S \leq P$, so $P > Q$), so if \mathscr{A} is a factor, either $P < Q$ or $Q < P$. To deal with the general case, let R be the central support of $P - S$. Then $R(P - S) = P - S$, and $R(Q - T) = 0$ since $Q - T$ is disjoint from $P - S$. Thus:

$$RP = R(P - S) + RS = P - S + RS,$$

while

$$RQ = R(Q - T) + RT = RT.$$

But since $R \in \mathcal{Z}(\mathscr{A})$ and $S \simeq T$, $RS \simeq RT$ (Proof. Let W be a partial isometry in \mathscr{A} with initial subspace $S\mathscr{X}$ and terminal subspace $T\mathscr{X}$. Then RWR is a partial isometry in \mathscr{A} with initial subspace $RS\mathscr{X}$ and terminal subspace $RT\mathscr{X}$). Thus, $RQ < RP$. Similarly,

$$(\mathbf{1} - R)P = (\mathbf{1} - R)(P - S) + (\mathbf{1} - R)S = (\mathbf{1} - R) \cdot S,$$

$$(\mathbf{1} - R)Q = (\mathbf{1} - R)(Q - T) + (\mathbf{1} - R)T = Q - T + (\mathbf{1} - R)T,$$

so

$$(\mathbf{1} - R)P < (\mathbf{1} - R)Q.$$

G Finite and Infinite Projections

Let π be a representation of an algebra \mathfrak{A} with involution, and assume that every representation of \mathfrak{A} can be written as a sum of irreducibles. We will say that π is finite if it contains at most finitely many copies of each irreducible representation of \mathfrak{A} ; otherwise, we say π is infinite. It is easy to see that a representation is infinite if and only if it is unitarily equivalent to a proper subrepresentation of itself. Translating, in the usual way, statements about representations into statements about projections in von Neumann algebras, we say that a projection P in a von Neumann algebra \mathscr{B} is *infinite* if there exists a projection Q in \mathscr{B} ; $Q \leq P$, such that $Q \simeq P$, and we say that a projection P is *finite* if it is not infinite. Note that, if P is an infinite projection and if $P' \geq P$, then P' is also infinite (If Q is a proper subprojection of P equivalent to P , then $P' = (P' - P) + P \simeq (P' - P) + Q$.) We say that the von Neumann algebra \mathscr{B} is *finite (infinite)* if the projection $\mathbf{1}$ is finite (infinite). The algebra $\mathcal{L}(\mathscr{X})$ is finite if and only if \mathscr{X} is finite dimensional.

For some purposes, it is useful to have a strengthening of the notion of an infinite projection. In the simple case where \mathfrak{A} is an algebra all representations

of which can be written as direct sums of irreducible representations, we have defined a representation π of \mathfrak{A} to be infinite if it contains infinitely many copies of at least one irreducible representation. In the same context, we will say that a representation is properly infinite if every irreducible representation which appears once in it appears there infinitely often. We thus make the following definition: a projection P in a von Neumann algebra is *properly infinite* if, for every central projection R of \mathfrak{A} , either $RP = 0$ or RP is infinite.

There is another distinction, related to finiteness, which has no analogue in the irreducible representations example: We will say that a projection P in a von Neumann algebra \mathfrak{A} is *purely infinite* if every non-zero subprojection of P is infinite and *semi-finite* if every non-zero subprojection of P contains a finite non-zero subprojection. We will say that a von Neumann algebra is *properly infinite* (*purely infinite*, *semifinite*) if the projection $\mathbf{1}$ is properly infinite (purely infinite, semifinite). It is not at all evident at this point that purely infinite von Neumann algebras exist.

We next show that every von Neumann algebra can be split uniquely into a finite part, a properly infinite but semifinite part, and a purely infinite part.

PROPOSITION IV. G.1 *Let \mathfrak{A} be a von Neumann algebra. There exist three central projection R_f , $R_{s.f.}$ and $R_{p.i.}$ such that $\mathbf{1} = R_f + R_{s.f.} + R_{p.i.}$ and such that R_f is finite, $R_{s.f.}$ is semi-finite, and $R_{p.i.}$ is purely infinite. Furthermore, these projections are uniquely determined.*

Proof If $(R_i)_{i \in I}$ is a family of mutually orthogonal finite central projections, $\sum_{i \in I} R_i$ is finite. (Proof If $S \not\leq T \subseteq \sum_i R_i$, and if $S \simeq T$, then $SR_i \not\leq TR_i$ for some i , so TR_i is an infinite projection contained in R_i , contradicting the assumed finiteness of R_i .) Hence, Zorn's Lemma gives the existence of a maximal finite central projection R_f . Also, if $(R_i)_{i \in I}$ is a mutually orthogonal family of purely infinite central projections, then $\sum_{i \in I} R_i$ is purely infinite. (If $Q \leq \sum_{i \in I} R_i$, then $QR_i \neq 0$ for some i , so QR_i is a non-zero subprojection of the purely infinite projection R_i , so QR_i is infinite, but $QR_i \leq Q$, so Q is infinite.) Again by Zorn's Lemma, there exists a maximal purely infinite central projection $R_{p.i.}$, and clearly $R_{p.i.}R_f = 0$, being both finite and purely infinite, must be zero, so $R_{p.i.} \leq \mathbf{1} - R_f$. Define $R_{s.f.} = \mathbf{1} - R_f - R_{p.i.}$. Then $R_{s.f.}$ is properly infinite (It is a central subprojection of the properly infinite central projection $\mathbf{1} - R_f$), but it contains no purely infinite central projection. Now if Q is any purely infinite projection, the central support of Q is also purely infinite. (Any subprojection contained in the central support of Q has a non-zero sub-

projection equivalent to a subprojection of Q ; hence, must be infinite.) Thus, $R_{s,r}$ contains no purely infinite subprojection, i.e., is semifinite. The uniqueness is proved by remarking that R_r must contain every finite central projection and that $R_{r,t}$ must contain every purely infinite central projection.

LEMMA IV. G.2 A projection P in a von Neumann algebra \mathfrak{A} is infinite if and only if there exists a mutually orthogonal sequence of non-zero projections (Q_n) such that $Q_n \leq P$ for all n and $Q_n \simeq Q_m$ for all n, m .

Proof If projections (Q_n) exist, then $\sum_{n=1}^{\infty} Q_n \simeq \sum_{n=2}^{\infty} Q_n$, so $\sum_{n=1}^{\infty} Q_n$ is infinite, so $P \geq \sum_{n=1}^{\infty} Q_n$ is infinite.

Now let P be infinite, and let W be a partial isometry in \mathfrak{A} with initial subspace $P\mathcal{H}$ and terminal subspace $P_1\mathcal{H}$, where $P_1 \leq P$. Define a decreasing sequence P_i of projections by $P_{i+1}\mathcal{H} = WP_i\mathcal{H}$. Then $W(P_i - P_{i+1})$ is a partial isometry with initial subspace $(P_i - P_{i+1})\mathcal{H}$ and terminal subspace $(P_{i+1} - P_{i+2})\mathcal{H}$; hence $Q_1 = P - P_1 \simeq Q_2 = P_1 - P_2 \simeq Q_3 = P_2 - P_3 \dots$

We now come to what seems to be the most delicate point on the theory of comparison of projections: The proof that the supremum of a finite number of finite projections is finite. The key to the proof is contained in the following lemma:

LEMMA IV. G.3 Let (P_1, P_2) and (Q_1, Q_2) be two pairs of projections in the von Neumann algebra \mathfrak{A} . Assume that P_1 is orthogonal to P_2 , that Q_1 is orthogonal to Q_2 , and that $P_1 + P_2 = Q_1 + Q_2$.

Then there is a central projection R in \mathfrak{A} such that

$$RQ_1 < RP_1, (I - R)Q_2 < (I - R)P_2.$$

(If \mathfrak{A} is a factor, so $R = 0$ or I , this implies that either $Q_1 < P_1$ or $Q_2 < P_2$.)

Proof We define projections P'_1, P'_2, Q'_1, Q'_2 , by:

$$P_1 = P_1 \wedge Q_1 + P_1 \wedge Q_2 + P'_1,$$

$$P_2 = P_2 \wedge Q_1 + P_2 \wedge Q_2 + P'_2,$$

$$Q_1 = Q_1 \wedge P_1 + Q_1 \wedge P_2 + Q'_1,$$

$$Q_2 = Q_2 \wedge P_1 + Q_2 \wedge P_2 + Q'_2.$$

We will show that P'_1, P'_2, Q'_1 , and Q'_2 are all equivalent. Let us first show how this implies the lemma. Assume first that \mathfrak{A} is a factor; then either $P_1 \wedge Q_2 < P_2 \wedge Q_1$ or $P_2 \wedge Q_1 < P_1 \wedge Q_2$. If the first relation holds, then

$$P_1 = P_1 \wedge Q_1 + P_1 \wedge Q_2 + P'_1 < P_1 \wedge Q_1 + P_2 \wedge Q_1 + Q'_1 = Q_1;$$

if the second relation holds, a similar argument shows $Q_2 < P_2$. If \mathfrak{A} is not a factor, then $P_1 \wedge Q_2$ and $P_2 \wedge Q_1$ may not be comparable, but Theorem IV. F.1 shows that there is a central projection R such that

$$R(P_2 \wedge Q_1) < R(P_1 \wedge Q_2), (\mathbb{1} - R)(P_2 \wedge Q_1) > (\mathbb{1} - R)(P_1 \wedge Q_2),$$

and then a straightforward modification of the above argument completes the proof. It remains to prove the equivalence of $P'_1, P'_2, Q'_1,$ and Q'_2 . Consider first the operator $P'_1 Q'_1$. We will show that its null space is exactly the null space of Q'_1 ; since its range is clearly contained in $P'_1 \mathcal{A}$, this will imply that its polar decomposition gives a partial isometry in \mathfrak{A} with initial subspace $Q'_1 \mathcal{A}$ and terminal subspace contained in $P'_1 \mathcal{A}$, i.e., it will show that $Q'_1 < P'_1$. Now if the null space of $P'_1 Q'_1$ is not contained in $(\mathbb{1} - Q'_1) \mathcal{A}$, it must contain a non-zero vector in $Q'_1 \mathcal{A}$. Thus, suppose $\xi \in Q'_1 \mathcal{A}$; and suppose that $P'_1 Q'_1 \xi = P'_1 \xi = 0$. Now the projection Q'_1 is $\leq Q_1$; hence, $Q_2 \xi = 0$, so $(Q_2 \wedge P_1) \xi = 0$. Similarly, since Q'_1 is orthogonal to $P_1 \wedge Q_1$, $(P_1 \wedge Q_1) \xi = 0$. Since $P_1 = P_1 \wedge Q_1 + P_2 \wedge Q_2 + P'_1$, $P_1 \xi = 0$. Since $Q'_1 \leq Q_1 + Q_2 = P_1 + P_2$, we must have $(P_1 + P_2) \xi = \xi$, i.e., $P_2 \xi = \xi$, so $\xi \in P_2 \mathcal{A} \wedge Q_1 \mathcal{A}$. But Q'_1 is orthogonal to $P_2 \wedge Q_1$, so $\xi = 0$. This then, shows $Q'_1 < P'_1$. Interchanging the roles of P_1 and P_2 gives $Q'_1 < P'_2$; interchanging the roles of Q_1 and Q_2 gives $Q'_2 < P'_1, Q'_2 < P'_2$; interchanging the roles of P 's and the Q 's gives $P'_1 \approx P'_2 \approx Q'_1 \approx Q'_2$ as desired.

THEOREM IV. G.4 Let P_1, \dots, P_n be a finite set of finite projections in the von Neumann algebra \mathfrak{A} . Then $\bigvee_{i=1}^n P_i$ is finite.

Proof By induction, it suffices to consider $n = 2$. Next, we reduce to the case P_1 and P_2 orthogonal. Let $P'_2 = P_1 \vee P_2 - P_1$. Then $P_2 P'_2$ is injective from $P'_2 \mathcal{A}$ into $P_2 \mathcal{A}$, so its partial isometric part has initial subspace $P'_2 \mathcal{A}$ and terminal subspace contained in $P_2 \mathcal{A}$, so $P_2 > P'_2$. Since P_2 is finite, P_2 is finite, and $P_1 \vee P'_2 = P_1 + P'_2 = P_1 \vee P_2$, so we may replace P_2 by P'_2 , i.e., we may assume P_2 is orthogonal to P_1 . Thus, let P_1, P_2 be mutually orthogonal finite projections in \mathfrak{A} , and suppose $P_1 + P_2$ is not finite. Then, by Lemma IV. G.2 there exists a mutually orthogonal sequence S_1, S_2, \dots of pairwise equivalent non-zero projections all contained in $P_1 + P_2$. Let $Q_1 = S_1 + S_3 \dots$; $Q_2 = P_1 + P_2 - Q_1 \geq S_2 + S_4 + \dots$. Evidently, Q_1 and Q_2 are mutually orthogonal, and $Q_1 + Q_2 = P_1 + P_2$. By Lemma IV. G.3, there is a central projection R such that

$$RQ_1 < RP_1 \leq P_1 \\ (\mathbb{1} - R)Q_2 < (\mathbb{1} - R)P_2 \leq P_2.$$

Since P_1 is finite and $RQ_1 = RS_1 + RS_2 + \dots$ is the sum of infinitely many mutually orthogonal, mutually equivalent projections, we must have

$RS_i = 0$ for all i . But then

$$\begin{aligned} P_2 &\geq (\mathbf{I} - R)P_2 > (\mathbf{I} - R)Q_2 \geq (\mathbf{I} - R)S_2 + (\mathbf{I} - R)S_4 + \dots \\ &= S_2 + S_4 + \dots \end{aligned}$$

contradicting the finiteness of P_2 and completing the proof.

II Tensor Product Decompositions

We have already seen how to decompose a von Neumann algebra into smaller pieces by using central projections. In this section, we will discuss another technique for breaking up von Neumann algebras into smaller pieces. The starting point here is a family $(P_i)_{i \in I}$ of projections in \mathfrak{B} ; (the index set may be finite or infinite); we assume that these projections are mutually orthogonal, pairwise equivalent, and add up to the identity operator. Let P be some fixed projection equivalent to each P_i . We will show that the Hilbert space \mathcal{H} on which \mathfrak{B} acts may be decomposed as a tensor product $\mathcal{H} \cong \tilde{\mathcal{H}} \otimes \mathcal{H}_1$, where $\tilde{\mathcal{H}}$ has Hilbert space dimension equal to the cardinality of the index set I , $\mathcal{H}_1 \cong P\mathcal{H}$, and \mathfrak{B} goes over into the von Neumann algebra generated by $\mathcal{L}(\tilde{\mathcal{H}}) \otimes \mathbf{1}$ and $\mathbf{1} \otimes P\mathfrak{B}P$: Here $P\mathfrak{B}P$ denotes the collection of all operators on $P\mathcal{H}$ obtained by restricting to $P\mathcal{H}$ operators $B \in \mathfrak{B}$ satisfying $B = PBP$, i.e., operators in \mathfrak{B} which are zero on $(\mathbf{I} - P)\mathcal{H}$ and whose range is contained in $P\mathcal{H}$. It is easy to verify from this description that $P\mathfrak{B}P$ is a weakly closed self-adjoint algebra of operators on $P\mathcal{H}$, containing the identity operator, i.e., that $P\mathfrak{B}P$ is a von Neumann algebra on $P\mathcal{H}$. This decomposition is particularly interesting when P is a *minimal projection* in \mathfrak{B} , i.e., when there is no non-zero projection of \mathfrak{B} contained in P except P itself. In this case, the von Neumann algebra $P\mathfrak{B}P$ contains no non-trivial projections; hence, consists only of the scalars, so $\mathfrak{B} \cong \mathcal{L}(\tilde{\mathcal{H}}) \otimes \mathbf{1}$, i.e., \mathfrak{B} is isomorphic to $\mathcal{L}(\tilde{\mathcal{H}})$. Let us expand a bit on this point before proceeding to the proof of the existence of a tensor product decomposition. Let \mathfrak{B} be a factor and suppose \mathfrak{B} contains a minimal projection P . (We say that a factor which contains a minimal projection is of *type I*.) By Zorn's Lemma, there exists a maximal collection $(P_i)_{i \in I}$ of pairwise orthogonal projections, such that each P_i is equivalent to P . We claim that $\mathbf{1} = \sum_i P_i$. To see this, let $P' = \mathbf{1} - \sum_i P_i$. Then either $P' > P$ or $P > P'$, $P' \not\sim P$. The first alternative is ruled out by the maximality of the family $(P_i)_{i \in I}$, and the second alternative implies $P' = 0$ by the minimality of P . Combining the above remarks, we get:

PROPOSITION IV. H.1 *Let \mathfrak{B} be a type I factor on a Hilbert space \mathcal{H} . Then \mathcal{H} may be decomposed as a tensor product $\mathcal{H} \cong \tilde{\mathcal{H}} \otimes \mathcal{H}_1$, in such a way that \mathfrak{B} corresponds to the algebra of operators of the form $A \otimes \mathbf{1}$, $A \in \mathcal{L}(\tilde{\mathcal{H}})$.*

We now proceed to the construction of the tensor product decomposition. We first state the result precisely:

PROPOSITION IV. H.2 *Let \mathfrak{A} be a von Neumann algebra on a Hilbert space \mathcal{X} . Let P be a projection in \mathfrak{A} , and suppose that we can decompose $\mathbf{1} = \sum_{i \in I} P_i$, where the P_i are mutually orthogonal projections in \mathfrak{A} each of which is equivalent to P . Then there exists a Hilbert space $\tilde{\mathcal{X}}$, with Hilbert-space dimension equal to the cardinality of I , and a unitary operator W from \mathcal{X} to $\tilde{\mathcal{X}} \otimes (P\mathcal{X})$, such that $W\mathfrak{A}W^{-1}$ is the von Neumann algebra generated by $\mathcal{L}(\tilde{\mathcal{X}}) \otimes \mathbf{1}$ and $\mathbf{1} \otimes P\mathfrak{A}P$. (In other words, \mathfrak{A} is unitarily equivalent to the tensor product of $\mathcal{L}(\tilde{\mathcal{X}})$ and $P\mathfrak{A}P$.)*

Proof We will construct the tensor product decomposition in two steps. First, we note that the decomposition $\mathbf{1} = \sum_i P_i$ gives a decomposition $\mathcal{X} = \bigoplus_i P_i\mathcal{X}$. The subspaces $P_i\mathcal{X}$ are all unitarily equivalent to $P\mathcal{X}$, and the unitary operator may be taken to be in \mathfrak{A} . More precisely, we can choose for each i a partial isometry $U_i \in \mathfrak{A}$ with initial subspace $P_i\mathcal{X}$ and terminal subspace $P\mathcal{X}$. Using this family of partial isometries, we can identify each $P_i\mathcal{X}$ with $P\mathcal{X}$ and thus transform the decomposition $\bigoplus_i P_i\mathcal{X}$ into a realization $\mathcal{X} \cong \bigoplus_{i \in I} P\mathcal{X}$. To be completely explicit, we can define a unitary operator $U: \mathcal{X} \rightarrow \bigoplus_{i \in I} P\mathcal{X}$ by $U\xi = (U_i\xi)$ (Note that, since U_i has initial subspace $P_i\mathcal{X}$, $\|U_i\xi\| = \|P_i\xi\|$, so $\sum_i \|U_i\xi\|^2 = \sum_i \|P_i\xi\|^2 = \|\xi\|^2$.)

The second step is to transform the direct sum representation $\mathcal{X} \cong \bigoplus_{i \in I} P\mathcal{X}$ into a tensor product representation. This is a standard construction; we let $\tilde{\mathcal{X}}$ be a Hilbert space with an orthonormal basis (ϕ_i) labelled by I , and we define a unitary operator $V: \bigoplus_{i \in I} P\mathcal{X} \rightarrow \tilde{\mathcal{X}} \otimes P\mathcal{X}$ by $V: (\xi_i) \mapsto \sum_i \phi_i \otimes \xi_i$. Now, of course, we define $W = V \cdot U$.

The next step is to compute what happens to various operators on \mathcal{X} when they are transformed by W . Tracing through the definitions, it is easy to see that

$$WU_i^*U_jW^{-1} = \hat{U}_{ij} \otimes \mathbf{1},$$

where \hat{U}_{ij} is the operator on $\tilde{\mathcal{X}}$ defined by

$$\hat{U}_{ij}\phi_j = \phi_i; \quad \hat{U}_{ij}\phi_k = 0 \quad \text{if } k \neq j.$$

It is also easily verified that the only operators on $\tilde{\mathcal{X}} \otimes P\mathcal{X}$ commuting with $\hat{U}_{ij} \otimes \mathbf{1}$ for all i, j are those of the form

$$\mathbf{1} \otimes A, \quad \text{with } A \in \mathcal{L}(P\mathcal{X}).$$

Next, suppose $B' \in \mathcal{B}'$. Then it may be seen that

$$WB'W^{-1} = \mathbf{1} \otimes (B' | P\mathcal{X}).$$

Finally, if $B \in P\mathcal{B}P$, then $\mathbf{1} \otimes B = W\tilde{B}W^{-1}$, where

$$\tilde{B} = \sum_i U_i^* B U_i \in \mathcal{B}.$$

Now using the fact that the commutant of $P\mathcal{B}P$ is exactly $\mathcal{B}' | P\mathcal{X}$, we see that the operators on $\tilde{\mathcal{X}} \otimes P\mathcal{X}$ commuting with $\tilde{U}_i, \otimes \mathbf{1}$ for all i, j and with $\mathbf{1} \otimes B$ for all $B \in P\mathcal{B}P$ are precisely the operators in $W\mathcal{B}W^{-1}$. Thus:

1) For any $A \in \mathcal{L}(\tilde{\mathcal{X}})$, $A \otimes \mathbf{1} \in (W\mathcal{B}W^{-1})' = W\mathcal{B}'W^{-1}$.

2) The bicommutant of $(\mathcal{L}(\tilde{\mathcal{X}}) \otimes \mathbf{1}) \cup (\mathbf{1} \otimes P\mathcal{B}P)$ is the commutant of $W\mathcal{B}W^{-1}$, i.e., is $W\mathcal{B}'W^{-1}$, so $W\mathcal{B}W^{-1}$ is the von Neumann algebra generated by $\mathcal{L}(\tilde{\mathcal{X}}) \otimes \mathbf{1}$ and $\mathbf{1} \otimes P\mathcal{B}P$.

I Classification of Factors

In the preceding section, we defined a factor to be of type I if it contains a minimal projection. We then showed that, if \mathcal{B} is such a factor, there is a tensor product decomposition of the Hilbert space \mathcal{X} as $\tilde{\mathcal{X}} \otimes \mathcal{X}_1$, such that $\mathcal{B} = \mathcal{L}(\tilde{\mathcal{X}}) \otimes \mathbf{1}$. In particular, \mathcal{B} is isomorphic to $\mathcal{L}(\tilde{\mathcal{X}})$. Conversely, if a von Neumann algebra \mathcal{B} is isomorphic to $\mathcal{L}(\tilde{\mathcal{X}})$ for some Hilbert space $\tilde{\mathcal{X}}$, then \mathcal{B} is a factor and contains a minimal projection (any one-dimensional projection is minimal in $\mathcal{L}(\tilde{\mathcal{X}})$); hence, is a factor of type I. Therefore, a von Neumann algebra is a factor of type I if and only if it is isomorphic to the algebra of all bounded operators on some Hilbert space. If that Hilbert space is finite-dimensional and of dimension n , we say that the algebra is of type I_n ; if the Hilbert space is infinite-dimensional, we say that the algebra is of type I_∞ .

At the opposite extreme from factors of type I are factors of type III; a factor \mathcal{B} is said to be of type III if every non-zero projection in it is infinite. Finally, a factor \mathcal{B} is said to be of type II if it is neither of type I nor of type III, i.e., if it contains no minimal projection but does contain a non-zero finite projection. A finite factor of type II is called a factor of type II_1 ; an infinite factor of type II is called a factor of type II_∞ .

The decomposition theory of the preceding section provides a convenient way of passing from factors of type II_∞ to factors of type II_1 . Let \mathcal{B} be a factor of type II_∞ , and let P be a finite projection in \mathcal{B} . We will prove shortly (Proposition IV. I.1) that there exists a mutually orthogonal family (P_i) of projections in \mathcal{B} equivalent to P , such that $\mathbf{1} = \sum_i P_i$. Thus, we may write

$\mathcal{H} \cong \hat{\mathcal{H}} \otimes P\mathcal{H}$, and \mathcal{B} is the von Neumann algebra generated by $\mathcal{L}(\hat{\mathcal{H}}) \otimes \mathbf{1}$ and $\mathbf{1} \otimes P\mathcal{B}P$. It is nearly immediate that $P\mathcal{B}P$ is a factor (if $A \in \mathcal{L}(P\mathcal{B}P)$, then $\mathbf{1} \otimes A \in \mathcal{L}(\mathcal{B})$, so $\mathbf{1} \otimes A$ is a scalar, so A is a scalar) and that $P\mathcal{B}P$ is finite (If W is a partial isometry in $P\mathcal{B}P$ with initial domain $P\mathcal{H}$ and terminal domain strictly contained in $P\mathcal{H}$, then since W is the restriction to $P\mathcal{H}$ of a partial isometry in \mathcal{B} which is zero on $(\mathbf{1} - P)\mathcal{H}$, P is equivalent (in \mathcal{B}) to a strictly smaller projection, and this violates the assumed finiteness of P). Furthermore, $P\mathcal{B}P$ certainly cannot contain minimal projections (since it can be regarded as a subalgebra of \mathcal{B}), so $P\mathcal{B}P$ is a factor of type II_1 . Thus, a type II_∞ factor can be decomposed into a tensor product (in a sense which should be clear from the above) of a factor of type II_1 , and a factor of type I_∞ . Conversely, starting with a factor of type II_1 , one can form a factor of type II_∞ by taking the tensor product with a factor of type I_∞ .

In the above discussion, we make use of the fact that the identity in a type II_∞ factor can be written as a sum of infinitely many mutually orthogonal, pairwise equivalent projections. The following proposition generalizes this assertion:

PROPOSITION IV. 1.1 *Let \mathcal{B} be a factor, P and Q projections in \mathcal{B} with P finite and non-zero and Q infinite. Then there exists a family $(P_i)_{i \in I}$ of mutually orthogonal projections in \mathcal{B} such that $Q = \sum_i P_i$ and such that $P_i \simeq P$ for all i .*

To prove this proposition, we first prove the following lemma:

LEMMA IV. 1.2 *Let \mathcal{B} be a factor, Q a projection in \mathcal{B} , and $(S_i)_{i \in I}$ an infinite mutually orthogonal family of pairwise equivalent non-zero projections contained in Q . Then we can write $Q = \sum_{j \in J} S_j$ where each S_j is equivalent to each S_i .*

Proof of Lemma IV. 1.2 By using Zorn's Lemma, we may extend $(S_i)_{i \in I}$ to a maximal mutually orthogonal family $(S'_j)_{j \in J}$ of pairwise equivalent non-zero projections contained in Q (i.e., we make $(S'_j)_{j \in J}$ by adjoining to $(S_i)_{i \in I}$ as many projections orthogonal to each other and to each S_i , but also equivalent to each S_i , as possible.) Let $Q' = Q - \sum_{j \in J} S'_j$. If j_0 is some element of J , then, since \mathcal{B} is a factor, either $Q' < S'_{j_0}$ or $Q' > S'_{j_0}$. If the second relation were true, we could adjoin to $(S'_j)_{j \in J}$ a projection $S'' \leq Q'$ which is equivalent to S'_{j_0} and thus contradict the maximality of the family $(S'_j)_{j \in J}$. Thus, we must have $Q' < S'_{j_0}$. Now:

$$\sum_{i \in I} S'_i = S'_{j_0} + \sum_{j \neq j_0} S'_j > Q' + \sum_{j \neq j_0} S'_j$$

(we have replaced S'_{j_0} by the smaller projection Q')

$$\simeq Q' + \sum_j S'_j$$

(since $J \setminus \{j_0\}$ and J have the same number of elements)

$$= Q,$$

so $\sum_{j \in J} S_j \succ Q$, but, on the other hand,

$$\sum_{j \in J} S_j \leq Q, \quad \text{so} \quad \sum_{j \in J} S_j \prec Q.$$

Thus, $\sum_{j \in J} S_j \approx Q$, and this means precisely that Q can be written as a sum of a mutually orthogonal family of subprojections labelled by J each of which is equivalent to each S_j .

Proof of Proposition IV. 1.1 Let $(P'_i)_{i \in I}$ be a maximal mutually orthogonal family of subprojections of Q equivalent to P , and let $Q' = Q - \sum_I P'_i$.

The maximality of $(P'_i)_{i \in I}$ implies, as in the above lemma, that $Q' \prec P$. In particular, Q' is finite. If the family I were finite, then we could write $Q = P_{i_1} + P_{i_2} + \dots + P_{i_n} + Q'$, so Q , as a finite sum of finite projections, would be finite. Since we have assumed Q to be infinite, this cannot be the case, so $(P'_i)_{i \in I}$ is an infinite family of subprojections of Q equivalent to P , and the proposition follows from the lemma.

We can also use Lemma IV. 1.2 to prove

PROPOSITION IV. 1.3 *Let \mathfrak{A} be a factor on a separable Hilbert space. Then any two infinite projections in \mathfrak{A} are equivalent.*

Proof Let P and Q be two infinite projections in \mathfrak{A} . Then there are sequences $S'_n(T'_n)$ of mutually orthogonal pairwise equivalent non-zero subprojections of $P(Q)$. Since \mathfrak{A} is a factor, either $S'_n \prec T'_m$ or $S'_n \succ T'_m$ for all m, n . Assume for definiteness that the first alternative holds. Then P and Q each contain an infinite mutually orthogonal family of projections equivalent to, say, S'_1 . Thus, by Lemma IV, 1.2, we can write

$$P = \sum_{i \in I} P_i; \quad Q = \sum_{j \in J} Q_j,$$

where each P_i and each Q_j is equivalent to S'_1 , and where the index sets I and J are both infinite. Now, on a separable Hilbert space, any infinite mutually orthogonal family of non-zero projections must be countable, so the index sets I and J can both be taken to be $\{1, 2, 3, \dots\}$. Thus, we have $P = \sum_{n=1}^{\infty} P_n; Q = \sum_{n=1}^{\infty} Q_n$, and $P_n \approx Q_n$ for all n , so $P \approx Q$.

COROLLARY IV. 1.4 *Let \mathfrak{A} be a factor of type III on a separable Hilbert space. Then any two non-zero projections in \mathfrak{A} are equivalent.*

J Dimension Theory for Projections in Type II Factors

Let \mathcal{H} be a separable Hilbert space, and consider the algebra $\mathcal{L}(\mathcal{H})$. For any projection $P \in \mathcal{L}(\mathcal{H})$, we define the dimension of P , $\dim(P)$, to be equal to the dimension of the range of P if this dimension is finite and to be ∞ if the dimension is infinite. Then the dimension function has two important properties:

- a) $P \simeq Q$ if and only if $\dim(P) = \dim(Q)$.
- b) If P is orthogonal to Q , $\dim(P + Q) = \dim(P) + \dim(Q)$.

It is easy to see that these two properties determine the dimension function uniquely up to multiplication by an overall constant factor. Since the equivalence of projections is preserved under isomorphisms of von Neumann algebras ($P \simeq Q$ in \mathcal{B} if and only if there exists $W \in \mathcal{B}$ such that $W^*W = P$; $WW^* = Q$), and since every type I factor on a separable Hilbert space is isomorphic to $\mathcal{L}(\mathcal{H})$, where \mathcal{H} is separable (but may be finite dimensional), we see that every type I factor \mathcal{B} admits a dimension function satisfying a) and b) which may be normalized to take on the values $\{0, 1, 2, \dots, n\}$ if \mathcal{B} is of type I_n , or $\{0, 1, 2, \dots, \infty\}$, if \mathcal{B} is of type I_∞ . The point of the present section is to prove that, if \mathcal{B} is a type II factor on a separable Hilbert space, \mathcal{B} again admits a dimension function (a positive real-valued function on the set of projections of \mathcal{B} satisfying a) and b)), that this dimension function is again uniquely determined up to multiplication by an overall constant but that this time the range of values is a closed interval $[0, a]$, where a is a finite number if \mathcal{B} is of type II_1 and is $+\infty$ if \mathcal{B} is of type II_∞ . Thus, from the present point of view, the difference between a factor of type II and a factor of type I is that the dimension function for a factor of type II takes on a continuous range of values, while the dimension function for a factor of type I takes on a discrete set of values. (A factor of type III on a separable Hilbert space also admits a dimension function, but it is not very interesting. Since any two non-zero projections in such a factor are equivalent (Corollary IV. I.4), condition a) implies that a dimension function must take on exactly one value other than zero, and condition b) forces this value to be $+\infty$.) It should be remarked that the condition that \mathcal{H} be separable is not essential to the discussion; it can be eliminated by allowing the dimension function to take on values which are infinite cardinal numbers and thus to distinguish between different infinite projections.

Before constructing the dimension function on a factor of type II, we prove some results about the comparison of projections in general von Neumann algebras. These results enable us to do simple "arithmetic operations" on equivalence classes of finite projections. By this, we mean the following: Suppose we have two equivalence classes of projections $\{P\}$ and $\{Q\}$. (We will use $\{P\}$ to denote the set of all projections equivalent to P .) If we can find $P \in \{P\}$ and $Q \in \{Q\}$ such that P is orthogonal to Q , we

define $[P] + [Q]$ to be the equivalence class of the projection $P + Q$. This definition is unambiguous since, if $P_1 \in [P]$, $Q_1 \in [Q]$ and if P_1 is orthogonal to Q_1 , then $P_1 + Q_1 \approx P + Q$. (Note, however, that $[P] + [Q]$ may not be defined; this is the case for $[P] = [Q] = [1]$ in a finite factor.) The next thing we want to show is that, if $[Q]$ is a finite equivalence class and $[P] < [Q]$, then $[Q] - [P]$ is well-defined. Thus, we want to prove:

PROPOSITION IV. J.1 *Let \mathfrak{A} be a von Neumann algebra, Q_1 and Q_2 equivalent finite projections in \mathfrak{A} , P_1 and P_2 equivalent projections in \mathfrak{A} such that $P_1 \leq Q_1$; $P_2 \leq Q_2$. Then $Q_1 - P_1 \approx Q_2 - P_2$.*

Proof Let us first assume that \mathfrak{A} is a factor. Then, by Theorem IV. F.1, either $Q_1 - P_1 > Q_2 - P_2$ or $Q_2 - P_2 > Q_1 - P_1$. Assume for definiteness that the first relation holds but that $Q_1 - P_1$ is not equivalent to $Q_2 - P_2$. Then $Q_2 - P_2$ is equivalent to a projection S which is strictly smaller than $Q_1 - P_1$. But then $Q_1 \approx Q_2 = (Q_2 - P_2) + P_2 \approx S + P_2 \leq Q_1$ and this violates the assumed finiteness of Q_1 . If \mathfrak{A} is not a factor, then, again by Theorem IV. F.1, there is a central projection R such that

$$R(Q_1 - P_1) > R(Q_2 - P_2); (1 - R)(Q_2 - P_2) > (1 - R)(Q_1 - P_1).$$

Straightforward modification of the above argument then yields:

$$R(Q_1 - P_1) \approx R(Q_2 - P_2); (1 - R)(Q_2 - P_2) \approx (1 - R)(Q_1 - P_1),$$

and hence $Q_1 - P_1 \approx Q_2 - P_2$.

Secondly, we want to prove that, for a finite equivalence class $[P]$ in \mathfrak{A} ,

(†) $[P]$ is uniquely defined if it makes sense at all. Specifically, we want to prove:

PROPOSITION IV. J.2 *Let P_1, P_2, Q_1, Q_2 be finite projections in a von Neumann algebra \mathfrak{A} . Assume that P_1 and P_2 are mutually orthogonal and equivalent; that Q_1 and Q_2 are mutually orthogonal and equivalent; and that $P_1 + P_2$ is equivalent to $Q_1 + Q_2$. Then $P_1 \approx Q_1$.*

Proof Assume first that \mathfrak{A} is a factor. Then we have either $P_1 < Q_1$ or $Q_1 < P_1$. Assume for definiteness $P_1 < Q_1$. Then, if P_1 is not equivalent to Q_1 , we have $P_1 \approx Q_1' \leq Q_1$. Since $P_1 \approx P_2$ and $Q_1 \approx Q_2$, we also have $P_2 \approx Q_2' \leq Q_2$. Thus, $Q_1 + Q_2 \approx P_1 + P_2 \approx Q_1' + Q_2' \leq Q_1 + Q_2$. But by assumption Q_1 and Q_2 are both finite, so $Q_1 + Q_2$ is finite, so $Q_1 + Q_2$ cannot be equivalent to a proper subprojection of itself; hence, $P_1 \approx Q_1$. Now, if \mathfrak{A} is not a factor, there is a central projection R such that $RQ_1 < RP_1$; $(1 - R)P_1 < (1 - R)Q_1$ (Theorem IV. F.1). The above argument can easily be reworked to show that $RQ_1 \approx RP_1$; $(1 - R)Q_1 \approx (1 - R)P_1$, so $Q_1 \approx P_1$.

We now restrict ourselves to factors of type II, and show that dividing a finite equivalence class by two always makes sense.

PROPOSITION IV. J.3 Let P be a projection in a factor \mathfrak{A} of type II. Then we can write $P = P_1 + P_2$, where P_1 and P_2 are mutually orthogonal and equivalent.

Proof Consider the set of all collections of pairs of non-zero subprojections of P $\{(Q_i, Q'_i)\}$ in \mathfrak{A} , where

- 1) $Q_i \simeq Q'_i$ for all i .
- 2) Q_i is orthogonal to Q_j for $i \neq j$.
 Q'_i is orthogonal to Q'_j for $i \neq j$.
- 3) Q_i is orthogonal to Q'_j for all i, j .

By Zorn's Lemma, there is a maximal such family $\{(Q_i, Q'_i)\}$. Let $P_1 = \sum_i Q_i$; $P_2 = \sum_i Q'_i$. Then P_1 and P_2 are equivalent mutually orthogonal subprojections of P . If we can show $P_1 + P_2 = P$, we are through. Suppose not. Then $P - P_1 - P_2$ is a non-zero projection in \mathfrak{A} . Since \mathfrak{A} is of type II, $P - P_1 - P_2$ cannot be a minimal projection, so we can write $P - P_1 - P_2 = Q_0 + Q'_0$, where neither Q_0 nor Q'_0 is zero. Either $Q_0 < Q_0 + Q'_0$ or $Q'_0 < Q_0 + Q'_0$, and we can assume the former is the case. Then if Q'_0 is a subprojection of Q_0 equivalent to Q_0 , $\{(Q_i, Q'_i)\} \cup \{Q_0, Q'_0\}$ satisfies 1), 2), 3), and hence violates the maximality of $\{(Q_i, Q'_i)\}$. Thus, $P - P_1 - P_2 = 0$, so we are through.

By induction, then, if $[P]$ is a finite equivalence class of projections in a type II factor, there is a uniquely defined equivalence class $\frac{1}{2^n} [P]$ for each n , and hence a uniquely defined equivalence class $\frac{m}{2^n} [P]$ for any integer m between 0 and 2^n . If \mathfrak{A} is of type II_∞ , then we may write $\mathbf{1}$ as a sum of a sequence of mutually orthogonal projections P_i such that each P_i is equivalent to P (Proposition IV. I.1). Thus, there is a uniquely defined equivalence class $\frac{m}{2^n} [P]$ for any integer m . Specifically, the statement $Q \in \frac{m}{2^n} [P]$ means that we can write $Q = \sum_{i=1}^m Q_i$; $P = \sum_{j=1}^{2^n} P_j$ with $Q_i \simeq P_j$ for all i, j . We define separately $0[P] = [0]$ for all P . It is important to recognize that the statement $Q \in \frac{m}{2^n} [P]$ has nothing to do with any assertion relating the projection Q to the linear operator obtained by multiplying the projection P by the real number $\frac{m}{2^n}$. We can now define the dimension function on projections in \mathfrak{A} .

a) If \mathfrak{A} is of type II_1 , and if P is a projection in \mathfrak{A} , we define

$$\dim(P) = \sup \left\{ \frac{m}{2^n} : \frac{m}{2^n} [\mathbf{1}] < [P] \right\}.$$

Then $\dim(P)$ is a real number in $[0, 1]$.

b) If \mathfrak{A} is of type II_∞ , we choose an arbitrary finite non-zero projection $P_0 \in \mathfrak{A}$, and we define for an arbitrary projection $P \in \mathfrak{A}$

$$\dim(P) = \sup \left\{ \frac{m}{2^n} : \frac{m}{2^n} [P_0] < [P] \right\}.$$

This time, $\dim(P)$ is a real number or $+\infty$; it is $+\infty$ if and only if P is infinite (Proof: If P is infinite, then by Proposition IV. 1.1, P may be written as a sum of infinitely many mutually orthogonal projections each of which is equivalent to P_0 . Thus, $[P] > m[P_0]$ for all integers m , so $\dim(P) = +\infty$. Conversely, if P is finite, then a maximal mutually orthogonal family of sub-projections of P equivalent to P_0 must be finite, so $P = P_{i_1} + \dots + P_{i_m} + P'$ where $P' \succ P_0$. Thus,

$$[P] < (m+1)[P_0], \quad \text{so } \dim(P) \leq m+1 < \infty).$$

The choice of P_0 in this construction amounts to a choice of a normalization for the dimension function. Another choice of P_0 would have given a dimension function differing at most by multiplication by a constant. Similarly, in the construction of a dimension function for a type II_1 factor, there is an arbitrariness in the normalization, but we have made the natural convention of defining the dimension of 1 to be one. In order to have a unified notation for the two cases, we will sometimes write P_0 for 1 in the II_1 case.

PROPOSITION IV. J.2 *Let \mathfrak{A} be a type II factor, P a finite projection in \mathfrak{A} . Then $[P] > \frac{m}{2^n} [P_0]$ if and only if $\dim(P) \geq \frac{m}{2^n}$. In particular, $\dim\left(\frac{m}{2^n} [P_0]\right) = \frac{m}{2^n}$.*

Proof From the definition of the dimension function, it follows that $\dim(P) \geq \frac{m}{2^n}$ if $[P] > \frac{m}{2^n} [P_0]$. Conversely, suppose it is not true that $[P] > \frac{m}{2^n} [P_0]$. Then, if $\frac{m'}{2^{n'}} \geq \frac{m}{2^n}$, it is not true that $[P] > \frac{m'}{2^{n'}} [P_0]$, so $\dim(P) \leq \frac{m}{2^n}$. To prove the final assertion, note that $\frac{m}{2^n} \leq \dim\left(\frac{m}{2^n} [P_0]\right)$ but that, if $\frac{m'}{2^{n'}} > \frac{m}{2^n}$, then it is not true that $\frac{m'}{2^{n'}} [P_0] < \frac{m}{2^n} [P_0]$ (by the finiteness of $\frac{m'}{2^{n'}} [P_0]$), so $\frac{m'}{2^{n'}} \geq \dim\left(\frac{m}{2^n} [P_0]\right)$. Thus, we have also $\frac{m}{2^n} \geq \dim\left(\frac{m}{2^n} [P_0]\right)$, so equality holds.

PROPOSITION IV. J.3 Let P and Q be mutually orthogonal projections in a type II factor \mathcal{A} . Then $\dim(P + Q) = \dim(P) + \dim(Q)$.

Proof If either $\dim(P)$ or $\dim(Q)$ is infinite, then $P + Q$ is infinite, $\dim(P + Q) = \infty = \dim(P) + \dim(Q)$. Thus, we have only to consider the case where $\dim(P)$ and $\dim(Q)$ are both finite. For each integer n , there exist integers m, m' such that

$$\frac{m}{2^n} [P_0] < [P] < \frac{m+1}{2^n} [P_0]; \quad \frac{m'}{2^n} [P_0] < [Q] < \frac{m'+1}{2^n} [P_0].$$

In other words, we can write

$$P = P_n + P'_n, \quad \text{with } P_n \in \frac{m}{2^n} [P_0] \text{ and } [P'_n] < \left(\frac{1}{2^n}\right) [P_0],$$

$$Q = Q_n + Q'_n, \quad \text{with } Q_n \in \frac{m'}{2^n} [P_0] \text{ and } [Q'_n] < \left(\frac{1}{2^n}\right) [P_0].$$

Then, since

$$P + Q = P_n + Q_n + P'_n + Q'_n,$$

since

$$P_n + Q_n \in \frac{m+m'}{2^n} [P_0]$$

and since

$$[P'_n + Q'_n] < \left(\frac{1}{2^{n-1}}\right) [P_0],$$

we have:

$$\frac{m+m'}{2^n} \leq \dim(P + Q) \leq \frac{m+m'+2}{2^n}.$$

As $n \rightarrow \infty$, $\frac{m}{2^n} \rightarrow \dim(P)$; $\frac{m'}{2^n} \rightarrow \dim(Q)$, so we get

$$\dim(P) + \dim(Q) = \dim(P + Q).$$

PROPOSITION IV. J.4 Let P be a projection in a type II factor \mathcal{A} , and suppose $\dim(P) = 0$. Then $P = 0$.

Proof By induction, we can find a sequence P_1, P_2, \dots of mutually orthogonal subprojections of P_0 with

$$P_1 \in \left(\frac{1}{2}\right) [P_0]; \quad P_2 \in \left(\frac{1}{4}\right) [P_0], \dots, P_i \in \left(\frac{1}{2^i}\right) [P_0] \dots$$

Since $\dim(P) = 0$, we cannot have $[P] > \left(\frac{1}{2^i}\right) [P_0]$ for any i , and, since any two equivalence classes of projections in a factor are comparable, we must have $[P] < \left(\frac{1}{2^i}\right) [P_0]$ for each i . Thus, for each i , we can find a subprojection S_i of P_i which is equivalent to P . The finite projection P_0 therefore contains a mutually orthogonal sequence of projections all of which are equivalent to P . This is possible only if $P = 0$. (Lemma IV. G.2).

PROPOSITION IV. J.5 *Let \mathcal{A} be a type II factor on a separable Hilbert space. Then two projections P and Q in \mathcal{A} are equivalent if and only if $\dim(P) = \dim(Q)$.*

Proof The definition of $\dim(P)$ depends only on the equivalence class of P , so two equivalent projections have the same dimension. Conversely, suppose $\dim(P) = \dim(Q)$. If the dimensions are infinite, then P and Q are both infinite, and hence are equivalent (Proposition IV. I.3). (This is the only use we make of the assumption that the Hilbert space on which \mathcal{A} acts is separable.) We can therefore assume that the dimensions are finite. Since \mathcal{A} is a factor, we have either $P > Q$ or $Q > P$, and we can assume that the first alternative holds. Thus, there is a projection Q' equivalent to Q with $Q' \leq P$. Now $\dim(P) = \dim(Q') + \dim(P - Q')$ by Proposition IV. J.3, so $\dim(P - Q') = 0$. Hence, by Proposition IV. J.4, $P - Q' = 0$, so P is equivalent to Q .

Combining the above results, we get:

THEOREM IV. J.6 *Let \mathcal{A} be a type II factor on a separable Hilbert space. Then there exists a function $P \mapsto \dim(P)$ on the set of projections in \mathcal{A} , with values in $[0, \infty]$, such that:*

i) $P \simeq Q$ if and only if $\dim(P) = \dim(Q)$.

ii) If P and Q are orthogonal, $\dim(P + Q) = \dim(P) + \dim(Q)$.

Moreover, the dimension function is uniquely specified up to multiplication by an overall constant factor by conditions a) and b).

Proof We have already proved the existence statement. To prove the uniqueness statement, let \dim' be another function satisfying a) and b). We first claim that P is finite if and only if $\dim'(P)$ is finite. Suppose P is any projection; then we can write $P = P_1 + P_2$ where P_1 is equivalent to P_2 (Proposition IV. 5.3). Thus, $\dim'(P) = \dim'(P_1) + \dim'(P_2) = 2 \dim'(P_1)$. Now if P is finite, P is not equivalent to P_1 , so $\dim'(P) \neq 2 \dim'(P_1)$, so $\dim'(P_1) < \infty$, so $\dim'(P) < \infty$. Conversely, if P is infinite, then P_1 must also be infinite (otherwise P would be the sum of two finite projections). Thus, $P \simeq P_1$ (Two infinite projections in a factor on a separable Hilbert space are equivalent.) This implies $\dim'(P_1) = \dim'(P) = 2 \dim'(P_1)$. This implies $\dim'(P_1) = \infty$, so $\dim'(P) = \infty$.

Now consider the projection P_0 which we have chosen to have dimension one. Since P_0 is finite, $\dim'(P_0) \neq \infty$, and since $P_0 \neq 0$, $\dim'(P_0) \neq 0$. Hence, by multiplying \dim' by a suitably chosen constant, we can arrange $\dim'(P_0) = 1$. Having done this, we no longer have a normalization free, so we must prove $\dim'(P) = \dim(P)$ for all finite P . We can of course regard \dim' as a function on the set of equivalence classes of projections. Repeated

application of properties a) and b) of the function \dim' gives:

$$\dim\left(\frac{m}{2^n} [P_0]\right) = \frac{m}{2^n} = \dim'\left(\frac{m}{2^n} [P_0]\right).$$

Now if P is any finite projection in \mathcal{A} , and if n is any positive integer, we can find a non-negative integer m such that

$$\frac{m}{2^n} [P_0] < [P] < \frac{m+1}{2^n} [P_0].$$

Condition b) implies that \dim' is an increasing function on equivalence classes of projections, so

$$\frac{m}{2^n} = \dim'\left(\frac{m}{2^n} [P_0]\right) \leq \dim'(P) \leq \dim'\left(\frac{m+1}{2^n} [P_0]\right) = \frac{m+1}{2^n}.$$

Since, as $n \rightarrow \infty$, $\frac{m}{2^n}$ approaches $\dim(P)$, we get

$$\dim'(P) = \lim_{n \rightarrow \infty} \frac{m}{2^n} = \dim(P).$$

We have not yet shown that the dimension function takes on all values in $[0, 1]$ if \mathcal{A} is of type II_1 and all values in $[0, \infty]$ if \mathcal{A} is of type II_∞ . This fact is easily obtained from the following proposition, which is of interest in its own right since it gives a continuity property of the dimension function.

PROPOSITION IV. J.7 *Let (P_n) be a mutually orthogonal sequence of projections in a type II factor. Then*

$$\dim\left(\sum_{n=1}^{\infty} P_n\right) = \sum_{n=1}^{\infty} \dim(P_n).$$

(In other words, the dimension function is countably additive.)

Proof $\sum_{n=1}^{\infty} P_n \geq \sum_{n=1}^N P_n$ for all N , so

$$\dim\left(\sum_{n=1}^{\infty} P_n\right) \geq \dim\left(\sum_{n=1}^N P_n\right) = \sum_{n=1}^N \dim(P_n)$$

for all N , so

$$\dim\left(\sum_{n=1}^{\infty} P_n\right) \geq \sum_{n=1}^{\infty} \dim(P_n).$$

Let us assume that

$$\sum_{n=1}^{\infty} \dim(P_n) < \dim\left(\sum_{n=1}^{\infty} P_n\right)$$

and derive a contradiction. Choose Q so that

$$\sum_{n=1}^{\infty} \dim(P_n) < \dim(Q) < \dim\left(\sum_{n=1}^{\infty} P_n\right).$$

We may take, for example, $Q \in r[P_0]$ with r an appropriately chosen dyadic rational). We now define inductively a mutually orthogonal sequence (Q_1, Q_2, \dots) of subprojections of Q such that $Q_n \simeq P_n$ for each n . This we do as follows: Since $\dim(Q) > \dim(P_1)$, we can find $Q_1 \leq Q$ with $Q_1 \simeq P_1$. Then

$$\dim(Q) = \dim(Q_1) + \dim(Q - Q_1) = \dim(P_1) + \dim(Q - Q_1)$$

But $\dim(Q) > \dim(P_1) + \dim(P_2)$, so

$\dim(Q - Q_1) > \dim(P_2)$, so there exists $Q_2 \leq Q - Q_1$ with $Q_2 \simeq P_2$.

Continuing in this way we get the desired sequence (Q_n) . Then, on the one hand,

$$\sum_n Q_n \leq Q, \text{ so } \dim\left(\sum_n Q_n\right) \leq \dim(Q);$$

on the other hand,

$$\sum_n Q_n \simeq \sum_n P_n, \text{ so } \dim\left(\sum_n Q_n\right) = \dim\left(\sum_n P_n\right) > \dim(Q).$$

We have therefore obtained a contradiction and hence proved the proposition.

COROLLARY IV. J.8 *Let Q_n be a decreasing sequence of finite projections in a type II factor, and let Q_∞ be the infimum of this sequence. Then $\dim(Q_\infty) = \lim_{n \rightarrow \infty} \dim(Q_n)$.*

Proof We can write $Q_1 = Q_\infty + (Q_1 - Q_2) + (Q_2 - Q_3) + \dots$. Thus,

$$\begin{aligned} \dim(Q_1) &= \dim(Q_\infty) + \dim(Q_1 - Q_2) + \dim(Q_2 - Q_3) + \dots \\ &= \lim_{n \rightarrow \infty} [\dim(Q_\infty) + \dim(Q_1) - \dim(Q_2) + \dim(Q_2) - \dots - \dim(Q_n)] \\ &= \lim_{n \rightarrow \infty} [\dim(Q_\infty) + \dim(Q_1) - \dim(Q_n)] \end{aligned}$$

so $\dim(Q_\infty) = \lim_{n \rightarrow \infty} \dim(Q_n)$ as asserted.

Bibliographic Note

In this section I will attempt to provide some indication of when to look for more information about the subjects treated in these notes. The listing is incomplete, unsystematic, and subjective; the works mentioned are some of those I have found useful in my own study.

Chapter I There are numerous excellent textbooks on general topology and measure theory; I have referred to Kelley [1] on topology and Halmos [1] on measure theory. The summaries of these subjects in the books of

Loomis [1] and Naimark [1] are also useful in providing a unencumbered exposition of the essential points. For more complete discussions of the theory of topological vector spaces, see Kelley and Namioka [1] or Köthe [1].

Chapter II The version of Choquet theory presented here is largely based on the review article of Choquet and Mayer [1], which is comprehensive but rather condensed. A more leisurely presentation of the subject may be found in the book of Phelps [1].

Chapter III The most useful single source is the treatise of Dixmier (Dixmier [2], referred to in the text as C^*A). In particular, the first two sections of this work give in sixty pages an excellent and complete exposition of the general theory of C^* algebras. The books of Loomis [1], Naimark [1], and Rickart [1] are also useful.

Chapter IV Here again, the standard reference is the treatise of Dixmier (Dixmier [1], referred to in the text as AvN). A more traditional exposition is given in Schwartz [1], and an excellent presentation of the elementary theory of von Neumann algebras, on about the same level as the present set of notes, may be found in Chapter VII of Naimark [1]. The exposition given in these notes of the comparison theory of projections and its relation to multiplicity theory of representations is based largely on Chapter I of Mackey [1]. For a brief survey of the theory of von Neumann algebras, see Kadison [1].

Bibliography

- Choquet, G. and P. A. Mayer, [1] Existence et unicité des représentations intégrales dans les convexes compacts quelconques, *Ann. Inst. Fourier* (Grenoble), 13, 139-154 (1963);
Dixmier, J., [1] *Les algèbres d'opérateurs dans l'espace hilbertien (Algèbres de von Neumann)*, Paris, Gauthier-Villars (1957).
[2] *Les C^* algèbres et leurs représentations*, Paris, Gauthier-Villars (1964).
Halmos, P. R., [1] *Measure Theory*, New York, van Nostrand (1950).
Kadison, R. V., [1] Theory of Operators, Part II. Operator Algebras, *B.A.M.S.*, 64, No. 3, part 2, 61-85 (1958).
Kelley, J. L., [1] *General Topology*, New York, van Nostrand (1955).
Kelley, J. L. and I. Namioka, [1] *Linear Topological Spaces*, New York, van Nostrand (1963).
Köthe, G., [1] *Topologische Lineare Räume I*, Berlin/Göttingen/Heidelberg, Springer (1960).
Loomis, L. H., [1] *An Introduction to Abstract Harmonic Analysis*, New York, van Nostrand (1953).
Mackey, G. W., [1] *The Theory of Group Representations*, Mimeographed notes from lectures delivered at the University of Chicago (1955).
Naimark, M. A., [1] *Normed Rings*, trans. L. F. Boron, Groningen, P. Noordhoff (1959).
Phelps, R. R., [1] *Lectures on Choquet's Theorem*, New York, van Nostrand (1966).
Rickart, C., [1] *General Theory of Banach Algebras*, New York, van Nostrand (1960).
Ruelle, D., [1] *Statistical Mechanics: Rigorous Results*, New York, W. A. Benjamin (1969).
Schwartz, J. T., [1] *W^* Algebras*, New York, Gordon and Breach (1967).