Pacific Journal of Mathematics

IN THIS ISSUE—

Vol. 39, No. 1 **May, 1971**

PACIFIC JOURNAL OF MATHEMATICS

Printed in Japan by International Academic Printing Co., Ltd., Tokyo, Japan

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

A GELFAND REPRESENTATION THEORY FOR C^*-ALGEBRAS

CHARLES A. AKEMANN

Recent work by the author which was independently duplicated in part by Giles and Kummer has made it possible to generalize the Gelfand representation theorem for abelian C-*algebras to the non-abelian case. Let A be a C-algebra with unit. If A is abelian, it can be identified with the algebra of all continuous complex-valued functions on its maximal ideal space (with the hull-kernel topology). A less precise way of looking at this result would be to say that an abelian A is completely recoverable from the set of maximal ideals and a certain structure thereon (in this case, a topology). If we use the latter description as the basis for a theory applicable to non-abelian A, we find immediately that two changes are necessary. The set of maximal ideals is replaced by the set of maximal left ideals, and secondly, the structure defined thereon will not be a topology, though it will have many similar properties when viewed correctly. This paper shows how the C^*-algebra is recovered from the maximal left ideals (with structure).

I. Preliminaries. Consider the W^*-algebra A^{**}, the second Banach space dual of A [9, p. 236]. There exists a central projection $z \in A^{**}$ which is the supremum of all the minimal projections in A^{**} [3, p. 278]. Set $M = zA^{**}$. The minimal projections of M are in one to one correspondence with the maximal left ideals of A [3, p. 280 and 9, p. 48], so that we can define a structure on this set of minimal projections instead of directly on the maximal left ideals. Naturally the first thing we "build" is the algebra M. We then single out a class L of projections in M as the q-open projections as follows. First note that we can consider $A \subset M$ since $A \subset A^{**}$ and $A \rightarrow zA$ is a *-isomorphism [9, p. 39]. (Also we can view M as the direct sum of irreducible representations of A, one from each equivalence class.) A projection p in M is q-open if there exists a closed left ideal I of A such that the weak* closure $\bar{I}$ of I in M is of the form Mp. The q-open projections are analogous to the open sets of a topology.

If A were abelian, M would be the algebra of all bounded complex function on its maximal ideal space K. The q-open projections would be characteristic functions of open sets of K for the hull-kernel topology. A self-adjoint operator b in M actually lies in $A(\subset M)$ if and only if the spectral projections of b corresponding to open sets of real numbers are q-open projections in the above sense.

This is a restatement of Gelfand's theorem since a function is continuous if and only if its inverse images of open sets are open.

We may now state an identical theorem for the non-abelian case. The proof follows immediately from the addendum to [4] and Theorem II. 17 of [3].

THEOREM I.1. *A self-adjoint operator $b \in M$ lies in $A(\subset M)$ if and only if each spectral projection of b which corresponds to an open subset of the real numbers is also a q-open projection.*

This theorem says that we may reconstruct A from its set of maximal left ideals together with the above defined structure. As a corollary we note that if two algebras A_1 and A_2 have "isomorphic structures" then they are isomorphic.

COROLLARY I.2. *Let A_1 and A_2 be C^*-algebras with $M_i = z_i A_i^{**}$ and L_i the q-open projections in M_i $(i = 1, 2)$. If there exists a $*$-isomorphism $\varphi: M_1 \rightarrow M_2$ which maps L_1 onto L_2, then $\varphi \mid A_1$ is an isomorphism of A_1 onto A_2.*

This paper extends these results to C^*-algebras without unit with appropriate modifications suggested by the abelian case. A number of other "topological" results are proved, and counter-examples are given to close off several tempting avenues of approach.

To complete our terminology, we shall assume from now on that A is a C^*-algebra which may not have a unit. The above discussion still applies to get $z \in A^{**}$ and we set $M = zA^{**}$. Identify A and $zA \subset M$ and call M the pure state q-space of A. (The terminology is lifted from [11].) We have already defined q-open projections in M, and their complements (in M) are called q-closed. $\tilde{A}$ will denote the algebra A with unit adjoined as in [9, p. 7]. Note that A is a closed two-sided ideal in $\tilde{A}$ of co-dimension one. Thus $\tilde{A}^* \cong A^* \oplus \{\lambda f_\infty\}$, where f_∞ is the unique pure state of $\tilde{A}$ which vanishes on A. Also the pure state q-space $\tilde{M}$ of $\tilde{A}$ is: $M \cong M \oplus \{\lambda 1_\infty\}$, with $f_\infty(1_\infty) = 1$. In view of Theorem II.17 of [3] all the properties of open or closed projections in A^{**} (as considered in [3 and 4]) carry over immediately to corresponding properties of q-open or q-closed projections in M.

II. The problem of compactness. Although the notion of compactness is vaguely introduced in [3], it is clear that a theory which claims to generalize locally compact Hausdorff spaces should generalize the notion of a compact set.

DEFINITION II.1. A projection $p \in M$ is *q-compact* if p is q-closed and there exists $b \in A^+$ $(= \{a \in A: a \geqq 0\})$ with $bp = p$.

There are a number of conditions equivalent to compactness for a set in a locally compact Hausdorff space. It would be desirable to show that many of them can be extended to equivalent conditions for q-compactness. The most desirable such condition would be:

Conjecture II.2. A regular [10, p. 408] projection $p \in M$ is compact if for every family $\{p_\alpha\}$ of q-closed projections such that the family $\{p_\alpha \wedge p\}$ has the finite intersection property, then $p \wedge \bigwedge_\alpha p_\alpha \neq 0$.

We shall prove this for certain p in Theorem II.6. The conjecture is false without the assumption of regularity (see Example IV.5).

LEMMA II.3. *Suppose B is a C^*-algebra, $b \in B^+$, $p \in B^{**}$ a projection and $\{a_\alpha\} \subset B$ an increasing net of positive elements with $\|b^{1/2} - b^{1/2}a_\alpha\| \underset{\alpha}{\rightarrow} 0$. If $b \geqq p$ (considering $B \subset B^{**}$), then $\|p - a_\alpha p\| \underset{\alpha}{\rightarrow} 0$.*

Proof. Since $\|b^{1/2} - b^{1/2}a_\alpha\| \underset{\alpha}{\rightarrow} 0$, clearly $\|(1 - a_\alpha)\, b\, (1 - a_\alpha)\| \underset{\alpha}{\rightarrow} 0$. Since $(1 - a_\alpha)\, b\, (1 - a_\alpha) \geqq (1 - a_\alpha)\, p\, (1 - a_\alpha)$, we get

$$\|(1 - a_\alpha)\, p\, (1 - a_\alpha)\| = \|(1 - a_\alpha) p\|^2 = \|p - a_\alpha p\|^2 \longrightarrow 0 \ .$$

LEMMA II.4. *If p is q-closed for A and we consider $\tilde{A}$ and $\tilde{M}$ as above with $M \subset \tilde{M}$ (hence $p \in \tilde{M}$) and there exists $b \in A^+$ with $b \geqq p$, then p is q-closed in $\tilde{M}$.*

Proof. Let $K = (p\ A^*\ p)^+$. Then K is $\sigma(A^*, A)$ closed by [3, II.2]. If K is not $\sigma(\tilde{A}^*, \tilde{A})$ closed, then there is a net $\{f_\alpha\} \subset K$ with $\|f_\alpha\| \subset K$ with $\|f_\alpha\| = 1$ and $f_\alpha \underset{\alpha}{\rightarrow} f$, $\sigma(\tilde{A}^*, \tilde{A})$, for some $f \in \tilde{A}^*$ with $\|f\| = f(1) = 1$. Since $\tilde{A}^* = A^* \bigoplus \{\lambda f_\infty\}$, we get $f = f_0 + \lambda f_\infty$ where $f_0 \in A^{*+}$ and $\lambda \geqq 0$. For any $c \in A$ with $c \geqq p$,

$$f_0(c) = f(c) = \lim_\alpha f_\alpha(c) \geqq \overline{\lim_\alpha}\, f_\alpha(p) = 1$$

since each $f_\alpha \in K$.

Now if $1 \in A$, then A^* is $\sigma(\tilde{A}^*, \tilde{A})$ closed in $\tilde{A}^*$, so the conclusion of this lemma is immediate. If $1 \notin A$, let $\{a_\gamma\} \subset A^+$ be an increasing approximate unit. Then, by Lemma II.3, $\{a_\gamma\}$ is an approximate unit for p also. Thus given $\varepsilon > 0$ there exists $c \in A$ with $c \geqq p$ and $\|c\| \leqq 1 + \varepsilon$ by Theorem 1.2 of [2]. Hence $f_0(c) \geqq 1$ by the above. Since $\varepsilon > 0$ was arbitrary, $\|f_0\| = 1$, so $\lambda = 0$, since

$$\|f\| = \|f_0\| + |\lambda| = 1 \ .$$

Thus $f \in A^*$. Since $\{f_\alpha\} \subset K$, K is $\sigma(A^*, A)$ closed, and $f_\alpha \underset{\alpha}{\rightarrow} f$ in the $\sigma(\tilde{A}^*, \tilde{A})$ topology, we see that $f \in K$, so K is $\sigma(\tilde{A}^*, \tilde{A})$ closed.

THEOREM II.5. *If p is q-closed and there exists $b \in A$ with $b \geqq p$, then p is q-compact.*

Proof. Since p is q-closed for $\tilde{A}$ by Lemma II.4, there exist $\{b_\alpha\} \subset \tilde{A}$, $b_\alpha = a_\alpha + \lambda_\alpha 1$ with $a_\alpha \in A$, $1 \geqq b_\alpha \geqq p$ and $b_\alpha \downarrow p$ in M [5, proof of Prop. 1]. Thus each b_α (and hence a_α) commutes with p. Since $f_\infty(b_\alpha) \underset{\alpha}{\rightarrow} 0$, there exists α_0 with $f_\infty(b_{\alpha_0}) < 1/2$. Thus $\lambda_{\alpha_0} < 1/2$ since $f_\infty(a_{\alpha_0}) = 0$. Let $g(t)$ be a continuous function which has $g(t) = 1$ for $t \geqq 1/2$, $g(0) = 0$, $0 \leqq g(t) \leqq 1$ for all t. Then $g(a_{\alpha_0}) \geqq p$. (Since a_{α_0}, b_{α_0} and p all commute, we may view them as functions on a common locally compact space; this makes the assertion clear.) Since $g(a_{\alpha_0}) \in A$, the theorem follows.

The construction in the proof of last theorem will not work for all projections p in M having only the property that $p \leqq b \in A$, even though it easily works whenever p is central.

THEOREM II.6. *Suppose $1 \in A$ and A is separable. Then Conjecture* II. 2 *holds for central projections $p \in M$.*

Proof. Suppose p satisfies the intersection condition of Conjecture II.2. We need only show p is q-closed since $1 \in A$. If it is not q-closed, let $\bar{p}$ be its closure [3, II. 11] and let $q \leqq \bar{p} - p$ be a minimal projection. As in [1] there exists a strictly positive element a_0 in $\{a \in A: aq = qa = 0\} = I$, so we let p_n be the spectral projection of a_0 corresponding to the interval $[0, 1/n]$. Since $\bigwedge_n p_n \wedge p = 0$, there is some n_0 with $p_{n_0} p = 0$ by hypothesis. Since p is central, the spectral projection x of a_0 corresponding to $[1/n_0, \infty)$ is q-closed and $x \geqq p$. This contradicts $xq = 0$ and $q \leqq \bar{p}$.

THEOREM II.7. *If p is q-compact, then p satisfies the intersection condition of Conjecture* II.2.

Proof. Since p is also q-closed in $\tilde{M}$ by Lemma II.4, the theorem follows from [3, II.10] for if $\{p_\alpha\}$ are q-closed in M, then their q-closures $\{\bar{p}_\alpha\}$ in $\tilde{M}$ have no larger M component. (Recall that $\tilde{M} = M \oplus \{\lambda 1_\infty\}$ with $1_\infty M = \{0\}$.) Thus if $p \wedge \bigwedge_{\alpha \in J} \bar{p} \neq 0$ for all finite sets J, $p \wedge \bigwedge_\alpha \bar{p}_\alpha \neq 0$, so $p \wedge \bigwedge_\alpha p_\alpha \neq 0$, since $p \wedge \bar{p}_\alpha = p \wedge p_\alpha$.

Next we move in a different direction for a characterization of

M. If A were an abelian C^*-algebra of functions containing the constants and separating the points of the topological space Ω, then A consists of all continuous functions on Ω if and only if Ω is compact. Following [11] we define a q-space to be an atomic W^*-algebra. If M_1 is a q-space and $A \subset M_1$, is a weak* dense C^*-subalgebra with $1 \in A$, we can define a q-open projection in M_1 as a sup of range projections of elements of A. Naturally q-closed projections are complements of q-open projections. If $M_1 = M$, the two definitions coincide.

THEOREM II.8. *If A is separable and $A \subset M_1$ as above, then there is an A-preserving *-isomorphism between M_1 and M if and only if the q-closed projections of M_1 satisfy the intersection condition of Conjecture* II.2.

Proof. If M_1 is *-isomorphic to M under an A-preserving map the verification is routine. Now suppose the q-closed projections of M_1 satisfy the intersection condition. If every pure state of A extends to a normal state of M_1, there is a natural isomorphism between M_1 and M which preserves A because of the definition of M as a subset of A^{**}. Thus let f be a pure state of A with no normal extension to M_1. Let $\{a_j\} \subset A$ be an increasing positive abelian [1] approximate unit for $\{a \in A\colon f(a^*a + aa^*) = 0\}$. Then let p_{jn} be the spectral projection of a_α corresponding to the interval $(1/n, \infty)$. Cleary $\bigvee_{j,n} P_{jn} = 1$ in M_1, for if not, then $(1 - \bigvee_{j,n} P_{jn})$ would be one-dimensional, hence f could be extended to a normal functional on M_1 with support $(1 - \bigvee_{j,n} p_{jn})$. But $\{(1 - p_{jn}\}\}$ is a decreasing net of closed projections in M_1 with $\bigwedge_{j,n}(1 - \vee p_{jn}) = 0$. Thus $(1 - p_{jn}) = 0$ for some j and n. Hence a_j is invertible, so $f = 0$, a contradiction.

III. The Gelfand representation.

LEMMA III.1. *If p is q-closed, p_1 is q-compact, and $p_1 p = 0$, then there exists $a \in A^+$ with $\|a\| = 1$, $ap = 0$, and $ap_1 = p_1$.*

Proof. Set $A_1 = \{a \in A\colon ap = pa = 0\}$. Consider $\widetilde{A}_1 \subset \widetilde{A}$. By Lemma II.4, p_1 is q-closed for $\widetilde{A}$. Thus the unit ball of $p_1 A^* p_1 = p_1 \widetilde{A}_1^* p_1$ $= p_1 \widetilde{A}_1^* p_1$ is compact for the $\sigma(\widetilde{A}^*, \widetilde{A})$ topology, hence also for the weaker $\sigma(\widetilde{A}_1^*, \widetilde{A}_1)$. Thus $p_1 \widetilde{A}_1^* p_1$ is $\sigma(\widetilde{A}_1^*, \widetilde{A})$ closed, so p_1 is q-closed for $\widetilde{A}_1$ [3, II.2]. Now by [4, I.1] there exists $a \in \widetilde{A}_1^+$ with $\|a\| = 1$, $ap_1 = p_1$ and $ap_2 = 0$, where p_2 is the one dimensional projection in $\widetilde{M}_1$ which supports the pure state f_∞ which vanishes on A_1. Since $ap_2 = 0$, $a \in A_1$, so $ap = 0$.

This last Lemma generalizes Urysohn's Lemma. We now define an analog for a continuous function.

DEFINITION III.2. A self-adjoint operator $b \in M$ is *q-continuous* if each spectral projection of b corresponding to an open subset of the spectrum of b is also q-open.

Now we can state our best Gelfand representation theorem.

THEOREM III.3. *The self-adjoint elements of A are exactly those q-continuous elements b of M such that the spectral projections of b corresponding to closed subsets of the spectrum of b which don't contain 0 are q-compact* (i.e., b "vanishes at ∞").

Proof. Consider $A \subset \tilde{A}$, $M \subset \tilde{M}$. If $b \in \tilde{A}$, then $b \in A$, since $b \subset M$. But if p is the spectral projection of b corresponding to an open subset U of the spectrum of b, we consider two cases. First if $0 \notin U$, then $p \in M$, hence p is q-open since it is q-open for A by hypothesis. Secondly if $0 \in U$, then the complement of U is closed and doesn't contain 0, thus the spectral projection corresponding to it is q-compact for A, hence q-closed for $\tilde{A}$ by Lemma II.4. Thus b is q-continuous for $\tilde{A}$ and Theorem I.1 applies.

For the abelian case it is well-known that if B is a C^*-algebra of continuous bounded functiohs on a locally compact Hausdorff space Ω such that the smallest topology on Ω making all $b \in B$ continuous agrees with the given topology, then B contains all continuous functions vanishing at ∞ on Ω. A similar result is true in general.

THEOREM III.4. *Let A_1 be a C^*-subalgebra of M such that the q-open projections for A_1 in M are the same as the q-open projections for A. Then $A_1 \supset A$ and $A_1 = A$ if $1 \in A$.*

Proof. Let $A_2 = A \cap A_1$. If p is q-open for A, then $p = \bigvee_\alpha p_\alpha$ where p_α is q-open with q-compact closure. For each α, p_α is also A_1 open, so there exists a net $\{a_\alpha^\gamma\} \subset A_1$ with $0 \leqq a_\alpha^\gamma \underset{\gamma}{\uparrow} p_\alpha$. By hypothesis each $a \in A_1$ is q-continuous, and since p_α has compact closure, Theorem III.3 applies to give $\{a_\alpha^\gamma\} \subset A$, hence in A_2. Thus p is A_2 open. We now apply Theorem III.3 of [3] and get $A_2 = A$. (Theorem III.3 of [3] is stated for algebras with unit, but considering $\tilde{A}_2$ and $\tilde{A}$ we get the result.)

Now if $1 \in A$, Theorem I.1 gives that $A_1 \subset A$, so $A_1 = A$.

Recall that one way of constructing the double centralizer $M(A)$ of A is to let $M(A)$ be the idealizer of A in A^{**}, i.e.,

$$M(A) = \{b \in A^{**}: bA + Ab \subset A\} .$$

We first prove a lemma bringing $M(A)$ into M.

LEMMA III.5. *The mapping $b \to bz$ is a $*$-isomorphism of $M(A)$ into M.*

Proof. Suppose $b \geqq 0$ in $M(A)$ and $zb = 0$. Then let $a \in A$ with $0 < a \leqq b$. Then $za = 0$ since $za \leqq zb = 0$. This means $a = 0$, a contradiction.

From now on consider $M(A)$ as a subalgebra of M. A tempting conjecture would be;

Conjecture III.6. The self-adjoint elements of $M(A)$ are exactly the q-continuous elements of M.

Our next result is one half of the conjecture.

THEOREM III.7. *Every self-adjoint element of $M(A)$ is q-continuous.*

Proof. Let $\{a_\alpha\} \subset A$ be a positive increasing approximate unit for A. Let $b \in M(A)$ be self-adjoint and let U be an open subset of the spectrum of b with p the spectral projection of b corresponding to U. Let $\{b_n\}$ be a sequence of continuous functions of b with $0 \leqq b_n \uparrow p$. Then $\{b_n^{1/2} a_\alpha b_n^{1/2}\}$ is a net in A which is $\leqq p$ and converges to p. Thus p is q-open for A.

In [7] Dixmier introduces the ideal center of a C^*-algebra which is a C^*-subalgebra of $M(A)$ containing A. Dixmier constructs it in A^{**} but Lemma III.5 assures us the idea carries over to M as well. We can characterize it in the obvious way.

COROLLARY III.8. *The ideal center of A consists of exactly those central elements of M which are q-continuous.*

Proof. We need to show that if d is central in M and p-continuous and $a \in A$, then $da \in A$. Clearly we need only consider $d, a \geqq 0$ and $\|d\| = \|a\| = 1$. For $\lambda > 0$, the spectral projection p of (da) corresponding to the interval $[\lambda, \infty)$ is less than or equal to the spectral projection of a corresponding to $[\lambda, \infty)$ which is q-compact since $a \in A$. By III.3 we need only show ad is q-continuous.

To show that (ad) is q-continuous, let (α, β) be an open interval and consider a and d as real functions on $\sigma(ad)$ (the spectrum of ad). Then let $t_0 \in K = \{t: a(t)d(t) \in (\alpha, \beta)\}$. For sufficiently small ε and δ we have $U \cap V \subset K$, where $U = \{t: a(t_0) - \varepsilon < t < a(t_0) + \varepsilon\}$ and $V = \{t; d(t_0) - \delta < t < d(t_0) + \delta\}$. Since K is a union of open sets of the form $U \cap V$, the spectral projection p of ad in M corresponding to K is a union of projections corresponding to sets of the form $U \cap V$. But for any U and V as above, the spectral projections of (ad) corresponding to U and V are both q-open and they commute. Hence their intersection corresponds to $U \cap V$ and it is q-open [3, II.7]. Thus p is a union of q-open projections, hence it is q-open [3, II.5].

IV. **Assorted results and examples.** One interesting question is: What are all the different C^*-algebras which have a factor for their pure state p-space? If M is countably decomposable, then the question was answered in [13] where it was shown that the C^*-algebra must consist of exactly the compact operators in M (i.e., the C^*-algebra generated by the minimal projections). We can slightly extend this result.

THEOREM IV.1. *Suppose M is a factor. Then A consists of exactly the compact operators in M if any q-open projection p is countably decomposable.*

Proof. Let $A_0 = \{a \in A: ap = pa = a\}$. Then the pure state q-space M_0 of A_0 is pMp. By [13] A_0 consists of the compact operators in pMp. Thus A contains all the compact operators in M by [9, p.85]. But if A is strictly larger than the compact operators, then they form an ideal in A, so A has at least two inequivalent irreducible representations. This contradicts the assumption that M is a factor.

Next is a theorem of the Stone-Weierstrass type.

THEOREM IV.2. *Let $B \subset A$ be a C^*-subalgebra which separates the pure states of A and 0. If pBp is norm closed in M for each q-closed projection p for A, then $B = A$.*

Proof. By [3, III.2] M is also the pure state q-space for B. Let p_1 be the B-closure of p in M (i.e., the smallest projection $\geqq p$ which is q-closed for B). If $p_1 > p$, then there is a minimal projection p_2 in M with $p_2 \leqq p_1 - p$. Let $\{b_\alpha\} \subset B$ with $1 \geqq b_\alpha \downarrow_\alpha p_2$ in M. Then

$\| p_1 b_\alpha p_1 \| = 1$ for all α, but $\| p b_\alpha p \| \underset{\alpha}{\rightarrow} 0$ since p is q-closed. By [3, II.12] the map $B \rightarrow p_1 B p_1$ has closed range, and by hypothesis the map $p_1 B p_1 \xrightarrow{\varphi} pBp$ has closed range also. But since p_1 is the q-closure of p for B, the map φ is $1-1$. Thus φ^{-1} is continuous by the closed graph theorem, and this contradicts $\| p_1 b_\alpha p_1 \| = 1$, $\| p b_\alpha p \| \underset{\alpha}{\rightarrow} 0$.

The most difficult aspect of the q-theory is the existence of non-regular projections, even in the best of circumstances [4, I.2]. The next result shows that some interesting projections are regular.

PROPOSITION IV.3. *If p' is finite-dimensional, then p is regular.*

Proof. Let p_1 be the q-closure of p. Then p_1' is finite dimensional, so p_1' is q-closed [3, II.8]. Hence p_1 is q-open and q-closed, so $p_1' \in A$ by [3, II.18]. By considering $p_1 A p_1$, we can assume $p_1 = 1$. Let $b \in A$ with $\| b \| = 1$ and suppose $\| bp \| < 1$. This would be the case if p were not regular. Since $\| b^* b \| = 1$ and $\| b^* b p \| < 1$, we can assume $b > 0$. Let p_2 be the spectral projection of b corresponding to the open interval (δ, ∞), where $\| bp \| < \delta < 1$. Then p_2 is q-open and $p_2 \neq 0$, so $p_2 \wedge p \neq 0$ as follows. If $p_2 \wedge p = 0$, then $p_2' \vee p' = 1$. Since p' is finite dimensional, this implies that p_2 is finite dimensional. But then $p_2 \in A$, so we can get a minimal projection $p_3 \in A$ with $p_3 \leqq p'$. This contradicts $\bar{p} = 1$. Now if g is a pure state of A with $g(p_2 \wedge p) = 1$, then

$$g(bp) = g(b) = g(p_2 b p_2) \geqq g(\delta p_2) = \delta \, .$$

This contradicts the definition of δ.

The next proposition and example show how badly behaved non-regular projections can be and how reasonable regular projections are.

PROPOSITION IV.4. *If $p \in M$ is regular, f a pure state of A, $b \in A$ with $b \geqq p$ and $f(b) = 0$, then $f(\bar{p}) = 0$ ($\bar{p}$ =* closure of *p).*

Proof. Let $\{a_\alpha\}$ be an increasing positive approximate unit for $\{a \in A\colon f(a^*a + aa^*) = 0\}$. By Lemma II.3 and by [2, I.2] we can get $\{b_n\} \subset A$ with $b_n \geqq p$, $\| b_n \| \leqq 1 + 1/n$, $f(b_n) = 0$. Let p_1 be the support projection of f. If $f(\bar{p}) \neq 0$, then there exists a pure state g of A with $g(\bar{p}) = 1$ and $g(p_1) \neq 0$. By regularity and [10, 6.1] there exists a net $\{g_\gamma\}$ of states of A with $g_\gamma \underset{\gamma}{\rightarrow} g$, $\sigma(A^*, A)$, and $g_\gamma(p) = 1$ for all γ. Let b_0 be a limit point of $\{b_n\}$ for the weak* topology of M, clearly $\| b_0 \| \leqq 1$. Since $g_\gamma(b_n) \geqq g_\gamma(p) = 1$ for all γ

and all n, then $g(b_n) \geqq 1$ for all n. Hence $g(b_0) \geqq 1$. But $\|b_0 + p_1\| = 1$ since $b_n p_1 = 0$ for all n implies $b_0 p_1 = 0$ (and similarly $p_1 b_0 = 0$). Hence $g(b_0 + p_1) \geqq 1 + g(p_1) > 1$, contradicting the asssumption that $\|g\| = 1$.

EXAMPLE IV.5. Let us work in the direct sum $\sum_{n=1}^{\infty} \oplus B(H_n)$ of matrix algebras where dimension $H_n = 2$ for all n. Set

$$a = \sum_{n=1}^{\infty} \begin{pmatrix} 1/n & 0 \\ 0 & 0 \end{pmatrix}, \quad p = \sum_{n=1}^{\infty} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad q = \sum_{n=1}^{\infty} \begin{pmatrix} 1 - \gamma_n & (\gamma_n - \gamma_n^2)^{1/2} \\ (\gamma_n - \gamma_n^2)^{1/2} & \gamma_n \end{pmatrix}$$

where $\{\gamma_n\}_{n-1}^{\infty}$ is an enumeration of the rationals between 0 and 1 which contains each rational an infinite number of times. Set $b = p + q$ and let A be the C^*-algebra generated by a and b. Let p_1 be the range projection of a in M.

Conclusions from the example. (1) $b \geqq p_1$ but there is no $d \in A^+$ with $dp_1 = p_1$ (c.*f.*, [12] page 11, line 11). (2) If f is the pure state at ∞ for A, then $f(b) = 0$ but $f(\bar{p}_1) \neq 0$, so p_1 is nonregular by Proposition IV.4. (3) Let p_2 be the support projection of f. Then $p_1 + p_2$ satisfies the intersection condition of Conjecture II.2, but $p_1 + p_2$ is not q-closed.

If $\varphi\colon A_1 \to A_2$ is a $*$-homomorphism of A_1 onto A_2, we may easily extend it to a normal $*$-homomorphism of M_1 onto M_2. However if φ is not onto, this extension may not be possible. The natural representation of the continuous function on the interval $[0, 1]$ into the algebra of all bounded operators on $L^2\,[0, 1]$ by $\varphi(f)h = fh$ has no such extension (the proof was communicated to me by R. Giles). In order to place q-theory into a category theory setting, one must restrict the class of allowable "*morphisms*" between two C^*-algebras. The following restriction is empty in the abelian case.

PROPOSITION IV.6. *A $*$-homomorphism φ taking the C^*-algebra A_1 into the C^*-algebra A_2 has a normal extension $\tilde{\varphi}\colon M_1 \to M_2$* (necessarily unique) *if and only if φ is continuous for the topologies generated by the seminorms $\|a\|_f = f(a^*a)$ for all pure states f of A_1* (or A_2 for the topology on A_2).

Proof. It $\tilde{\varphi}$ exists, the continuity is automatic for $\tilde{\varphi}$, hence for φ. The converse follows immediately from [14, p. 3 of appendix].

REFERENCES

1. J. F. Aarnes, and R. V. Kadison, *Pure states and approximate identities*, Proc. Amer. Math. Soc., **21** (1969), 749–752.

2. C. A. Akemann, *Approximate units and maximal abelian C^*-subalgebras*, Pacific, J. Math., **33** (3) (1970), 543-550.
3. ———, *The general Ston-Weierstrass problem*, J. Functional Analysis, **4**(2) (1969), 277-294.
4. ———, *Left ideal structure of C^*-algebras*, J. Functional Analysis, **6** (2) (1970), 305-317.
5. ———, *Sequential convergence in the dual of a W-*algebra*, Comm. Math. Phys. **7** (1968), 222-224.
6. R. C. Busby, *Double centralizers and extensions of C^*-algebras*, Trans. Amer. Math. Soc., **132** (1968), 79-99.
7. J. Dixmier, *Ideal center of a C^*-algebra*, Duke Math. J., **35** (1968), 375-382.
8. ———, *Les algebrès d'opérateurs dans l'espace hilbertsen* (Algèbres von Neumann), Gauthier-Villars, Paris, 1957.
9. ———, *Les C^*-algebres et leurs représentations*, Gauthier-Villars, Paris, 1964.
10. E. G. Effros, *Order ideals in a C^*-algebra and its dual*, Duke Math. J. **30** (1963), 391-412.
11. R. Giles and H. Kummer, *A non-commutative generalization of topology*, Queens Mathematical Reprint No. 1970-12, Kingston.
12. G. K. Pedersen, *C^*-integrals*, Kobenhavns Universitet Matematisk Institut Preprint Series 1970 No. 7.
13. A. Rosenberg, *The number of irreducible representations of simple rings with no minimal ideals*, Amer. J. Math. **75** (1953), 523-530.
14. S. Sakai, *The theory of W^*-algebras*, Yale Notes, 1963.

Received January 7, 1971. This research was supported by National Science Foundation Grant GP11238 and GP19101.

UNIVERSITY OF CALIFORNIA SANTA BARBARA.

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

SPECTRAL THEORY FOR A FIRST-ORDER SYMMETRIC SYSTEM OF ORDINARY DIFFERENTIAL OPERATORS

SORRELL BERMAN

For a symmetric differential expression associated with a first order system

$$A_0(t)x' + A(t)x, \quad a < t < b$$

where A_0 and A are $n \times n$ matrices and x is an $n \times 1$ vector, a spectral decomposition will be developed. That is, if S is a closed symmetric differential operator determined by the differential system, the explicit nature of the generalized resolutions of the identity for all the self-adjoint extensions of S in any Hilbert space will be determined in terms of a fundamental matrix and spectral matrices associated with these extensions. An important aspect is that these self-adjoint extensions may be defined in Hilbert spaces larger than the natural one $\mathscr{H}$ in which the operator S is defined.

The development proceeds as in Coddington [5]; however, the consideration of systems of differential equations introduces matrix techniques and notation. It is hoped that this formulation will have application to such problems as open end (infinite time) control theory problems, and facilitate the canonical formulation of the associated spectral analysis.

Preliminary definitions. Let $\mathscr{H}$ be a Hilbert space with an inner product $(\ ,\ )$.

(1) *Generalized Resolution of the Identity.* Let $F = \{F(\lambda)\}$ be a family of bounded self-adjoint operators in $\mathscr{H}$, depending on real λ, such that:

(i) $F(\lambda) \geqq F(\mu)$, $\lambda > \mu$,

(ii) $F(\lambda + 0) = F(\lambda)$,

(iii) $F(\lambda) \to I$, as $\lambda \to +\infty$,
$F(\lambda) \to 0$, as $\lambda \to -\infty$,

then F is a generalized resolution of the identity.

The family F is said to be associated with a symmetric operator Z (or F is a "spectral function" for Z, Naimark [7]) if

$$(Zu, v) = \int \lambda d(F(\lambda)u, v),$$

and

$$||Zu||^2 = \int \lambda^2 d(F(\lambda)u, u),$$

for all $u \in \mathscr{D}(Z)$ and $v \in \mathscr{H}$.

(2) *Generalized Resolvent.* Let Z be a symmetric operator and $F = \{F(\lambda)\}$ be an associated generalized resolution of the identity. For $\operatorname{Im} l \neq 0$, let $\mathscr{R} = \{\mathscr{R}(l)\}$ be a family of operators such that

$$(\mathscr{R}(l)u, v) = \int \frac{d(F(\lambda)u, v)}{\lambda - l}.$$

Then $\mathscr{R}$ is a generalized resolvent of Z associated with F. The development for symmetric operators will include the case for self-adjoint operators.

1. Basic vector and matrix definitions. In addition to the usual definitions and notation for the absolute magnitude of a vector, the inner product of two vectors, the norm of a vector and the absolute magnitude of a matrix, the norm of a matrix is defined as

$$||A|| = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \int |A_{ij}(t)|^2 dt\right)^{1/2} = \left(\int \operatorname{trace}(A^*(t)A(t))dt\right)^{1/2};$$

and a matrix "inner product" is introduced,

$$(A, B) = \int B^*(t)A(t)dt,$$

which is a matrix whose (i, j)th element is

$$\sum_{l=1}^{n} \int \bar{B}_{li}(t)A_{lj}(t)dt.$$

This "inner product" makes sense for any two matrices for which $B^*(t)A(t)$ exists and is integrable.

An inner product of a matrix and a vector can be defined in some situations; it is a special case of the matrix "inner product." For example, if f is an $n \times 1$ vector and G an $n \times n$ matrix,

$$(f, G) = \int G^*(t)f(t)dt.$$

2. The "basic operators" T_0 and T. Let (a, b) be an open interval on the real line (a may be $-\infty$ and/or b may be $+\infty$). A differential operator L is defined by

$$Lx = A_0(t)x' + A(t)x.,$$

where: x is an $n \times 1$ vector, A_0 and A are n by n suitably regular matrix-valued functions (Brauer [2]) and $'$ denotes d/dt. The Lagrange adjoint, L^+, associated with L is defined by

$$\begin{aligned} L^+y &= -(A_0^*(t)y)' + (A^*(t)y) \\ &= -A_0^*(t)y' + (-A_0^{*\prime}(t) + A^*(t)y). \end{aligned}$$

The operator L is formally self-adjoint if $L = L^+$, that is when

$$A_0 = -A_0^* \text{ and } A = -A_0^{*\prime} + A^* = A_0' + A^*.$$

Throughout the remainder of this paper L will be assumed to be formally self-adjoint.

Using the definitions for the inner product of two vectors, and for the norm, a Hilbert space, $\mathscr{H}$, can be defined,

$$\mathscr{H} = \mathscr{L}^2(a, b) = \{u: \|u\| < \infty\}.$$

Defining a domain $\mathscr{D}$ in $\mathscr{H}$ by $\mathscr{D} = \{u \in \mathscr{H}$: (i) u is absolutely continuous on every compact subinterval of (a, b), (ii) $Lu \in \mathscr{H}\}$, an operator T, having domain $\mathscr{D}$, can be defined by

$$Tu = Lu, \quad u \in \mathscr{D}.$$

Let, for $u, v \in \mathscr{D}$,

$$\langle uv \rangle = (Lu, v) - (u, L^+v) = (Lu, v) - (u, Lv).$$

Then, similarly for a domain $\mathscr{D}_0$,

$$\mathscr{D}_0 = \{u \in \mathscr{D}: \langle uv \rangle = 0 \text{ for all } v \in \mathscr{D}\},$$

an operator T_0 can be defined by

$$T_0u = Lu, \quad u \in \mathscr{D}_0.$$

The development of the operators T_0 and T is motivated by the fact that T_0 is the smallest closed symmetric operator in $\mathscr{H}$ (associated with the differential operator L) having a domain which contains all vectors which are infinitely differentiable on (a, b) and vanish outside closed bounded subintervals of (a, b). Further, if F_1 is any generalized resolution of the identity for a closed symmetric operator T_1, where $T_0 \subset T_1 \subset T$, then F_1 is a generalized resolution of the identity for T_0, also. Thus, by considering T_0, a maximal set of generalized resolutions of the identity, which are naturally associated with L, can be obtained.

The following theorem provides an important relation between T_0 and T.

THEOREM 2.1. *The operator T_0 is closed, symmetric, and $T_0^* =$*

T, $T^* = T_0$.

Proof. Let

$$K(t, \tau) = \begin{cases} \Phi(t)\Phi^{-1}(\tau)A_0^{-1}(\tau), & t \geqq \tau, \\ 0, & t < \tau, \end{cases}$$

where Φ is a fundamental matrix, that is, a matrix whose columns are independent solutions of $Lx = 0$. Thus, Φ is a nonsingular $n \times n$ matrix such that $L\Phi = 0$. As a function of t, $LK(t, \tau) = 0$, and

$$K(t+,t) - K(t-,t) = \Phi(t+)\Phi^{-1}(t)A_1^{-0}(t) - 0 = A_0^{-1}(t).$$

The representation for K can be simplified

$$(\Phi^* A_0 \Phi)' = 0, \text{ or } \Phi^* A_0 \Phi = D^{-1},$$

where D is a skew-Hermitian constant matrix, and hence

$$\Phi^{-1}A_0^{-1} = D\Phi^*.$$

The matrix K can now be written as

$$K(t, \tau) = \begin{cases} \Phi(t)D\Phi^*(\tau), & t \geqq \tau, \\ 0, & t < \tau. \end{cases}$$

Let Δ be a closed bounded subinterval $[\tilde{a}, \tilde{b}]$ of (a, b). The Hilbert space $\mathscr{L}^2(\Delta)$ is defined by

$$\mathscr{L}^2(\Delta) = \{u: \|u\|_\Delta < \infty\}.$$

For $t \in \Delta$, the vector x defined by

$$\begin{aligned} x(t) &= \int_{\tilde{a}}^{\tilde{b}} K(t, \tau)y(\tau)d\tau \\ &= \int_{\tilde{a}}^{t} K(t, \tau)y(\tau)d\tau, \end{aligned}$$

where $y \in \mathscr{L}^2(\Delta)$, is such that

(0) $x \in \mathscr{L}^2(\Delta)$,

(i) x is absolutely continuous on Δ,

(ii) $Lx \in \mathscr{L}^2(\Delta)$.

Having verified that for $t \in \Delta$ and $y \in \mathscr{L}^2(\Delta)$ the vector x satisfies conditions (0), (i), and (ii), the proof follows exactly as in Theorem 1 of reference 3.

3. The Green's function G_Δ. In § 5. the generalized resolvents associated with T_0 will be constructed. The generalized resolvent will be developed starting from the Green's function G_Δ associated with certain self-adjoint boundary-value problems on finite subintervals

Δ. The purpose of this section is to derive such Green's functions G_Δ.

Once again, let Δ be a closed bounded subinterval of (a, b), denoted by $[\tilde{a}, \tilde{b}]$. Analogous to previous definitions, a domain $\mathscr{D}_\Delta$ is defined and the associated operator T_Δ, having domain $\mathscr{D}_\Delta$,

$$T_\Delta u = Lu, \quad u \in \mathscr{D}_\Delta.$$

Similarly, for $\mathscr{D}_{0\Delta}$, an operator $T_{0\Delta}$, having domain $\mathscr{D}_{0\Delta}$, is defined by

$$T_{0\Delta} u = Lu, \quad u \in \mathscr{D}_{0\Delta}.$$

[NOTE: The conditions and relations of Theorem 2.1 hold for $T_{0\Delta}$ and T_Δ.]

It will now be shown that abstract self-adjoint boundary conditions can be constructed by considering the self-adjoint extensions of $T_{0\Delta}$. Let

$$\mathscr{E}_\Delta(\pm i) = \{v \in \mathscr{D}_\Delta : T_\Delta v = \pm iv\};$$

It is clear that dim $\mathscr{E}_\Delta(i) =$ dim $\mathscr{E}_\Delta(-i) = n$. The domain $\mathscr{D}_\Delta$ can be written as a direct sum

$$\mathscr{D}_\Delta = \mathscr{D}_{0\Delta} + \mathscr{E}_\Delta(i) + \mathscr{E}_\Delta(-i).$$

From the theory of the Cayley transform (see Riesz-Nagy [8], for example) every self-adjoint extension, $T_{\Delta U}$, of $T_{0\Delta}$ has a domain

$$\mathscr{D}_{\Delta U} = \mathscr{D}_{0\Delta} + (I - U)\mathscr{E}_\Delta(-i),$$

where U is a unitary mapping from $\mathscr{E}_\Delta(-i)$ onto $\mathscr{E}_\Delta(i)$; and

$$T_{\nabla U} u = Lu, \quad u \in \mathscr{E}_{\Delta U}.$$

Let $\{\varphi_{\Delta i}\}$ $i = 1, \cdots, n$, be an orthonormal basis for $\mathscr{E}_\Delta(i)$; also let $\{\psi_{\Delta i}\}$ $i = 1, \cdots, n$, be an orthonormal basis for $\mathscr{E}_\Delta(-i)$; finally let

$$v_{\Delta j} = \psi_{\Delta j} - U\psi_{\Delta j},$$

and

$$v_{\Delta j*} = \varphi_{\Delta j} - U^*\varphi_{\Delta j}, \quad j = 1, \cdots, n.$$

The following theorem describes the abstract self-adjoint boundary conditions induced by the domain $\mathscr{D}_{\Delta U}$.

THEOREM 3.1. *The domain $\mathscr{D}_{\Delta U}$ of $T_{\Delta U}$ has the following representation:*

$$\mathscr{D}_{\Delta U} = \{u \in \mathscr{D} : \langle uv_{\Delta j*}\rangle = 0, \mathrm{j} = 1, \cdots, n\}$$

where $\{\langle uv_{\Delta j*}\rangle = 0, j = 1, \cdots, n\}$ form a self-adjoint set of boundary *conditions.*

Proof. This follows by direct analogy from the proof of Theorem 3 in Coddington [3].

The set $\{\langle uv_{\Delta j*}\rangle = 0, j = 1, \cdots, n\}$ forms a self-adjoint set of boundary conditions since the $v_{\Delta j*}$ are linearly independent and $\langle v_{\Delta j*}v_{\Delta k*}\rangle = 0$ for all j, k.

The set of self-adjoint boundary conditions $\{\langle uv_{\Delta j*}\rangle = 0, j = 1, \cdots, n\}$ can be represented in matrix form by

$$V_{\Delta*}^*(\tilde{b})A_0(\tilde{b})u(\tilde{b}) - V_{\Delta*}^*(\tilde{a})A_0(\tilde{a})u(\tilde{a}) = 0,$$

where $V_{\Delta*}$ is the matrix whose ith column is the vector $v_{\Delta i*}$. Letting

$$M_\Delta = - V_{\Delta*}^*(\tilde{a})A_0(\tilde{a}),$$

and

$$N_\Delta = V_{\Delta*}^*(\tilde{b})A_0(\tilde{b}),$$

the self-adjoint boundary conditions can be written in standard form

$$U_\Delta u = M_\Delta u(\tilde{a}) + N_\Delta u(\tilde{b}) = 0.$$

The self-adjoint boundary-value problem (on Δ)

$$Lu = lu, \quad U_\Delta u = 0 \quad (bv)$$

will now be considered. The Green's function G_Δ associated with the problem (bv) is a unique function $G_\Delta(t, \tau, l)$ (l not an eigenvalue of (bv)) satisfying the following conditions:

(i) $G_\Delta(t, \tau, l)$ and $\partial/\partial t G_\Delta(t, \tau, l)$ are continuous on $\tilde{a} \leqq t \leqq \tau \leqq \tilde{b}$ and $\tilde{a} \leqq \tau \leqq t \leqq b$, and for each fixed (t, τ) are analytic in l,

(ii) $G_\Delta(t+, \tau, l) - G_\Delta(t-, t, l) = A_0^{-1}(t)$, $\tilde{a} < t < \tilde{b}$,

(iii) G_Δ satisfies $LG_\Delta = lG_\Delta$ (as a function of t),

(iv) G_Δ satisfies $U_\Delta G_\Delta = 0$ (as a function of t),

(v) $G_\Delta(t, \tau, l) = G_\Delta^*(\tau, t, \bar{l})$,

(vi) if $f \in \mathscr{L}^2(\Delta)$ and $Lu = lu + f$,

then,

$$u(t) = \int_\Delta G_\Delta(t, \tau, l)f(\tau)d\tau, \quad U_\Delta u = 0$$

and if

$$\mathscr{G}_\Delta(l)f(t) = \int_\Delta G_\Delta(t, \tau, l)f(\tau)d\tau,$$

then

$$(L - l)\mathscr{G}_\Delta(l)f(t) = f(t),$$

and

$$\|\mathscr{G}_\Delta(l)\|_\Delta \leqq |\text{Im } l|^{-1}.$$

The Green's function will now be constructed starting from the kernel

$$K_\Delta(t, \tau, l) = \begin{cases} \Phi(t, l)D\Phi^*(\tau, l), & t \geqq \tau, \\ 0, & t < \tau, \end{cases}$$

for $t, \tau \in \Delta$; where Φ is a fundamental matrix for $(L - l)u = 0$, having the property that for some c, $\tilde{a} < c < \tilde{b}$, $\Phi(c, l) = I$. The matrix $D(= A_0^{-1}(c))$ is a constant, skew-Hermitian matrix. From Theorem 8.4 Coddington and Levinson [6], Φ is continuous as a function of (t, l), and for fixed t is an analytic function of $l(\text{Im } l \neq 0)$. Let

$$G_\Delta(t, \tau, l) = K_\Delta(t, \tau, l) + \Phi(t, l)J(\tau, l).$$

Introducing the notation

$$U_\Delta\Phi(l) = M_\Delta\Phi(\tilde{a}, l) + N_\Delta\Phi(\tilde{b}, l),$$

G_Δ can be written as

$$G_\Delta(t, \tau, l) = \begin{cases} \Phi(t, l)(U_\Delta\Phi(l))^{-1}M_\Delta\Phi(\tilde{a}, l)D\Phi^*(\tau, \bar{l}), & t \geqq \tau, \\ -\Phi(t, l)(U_\Delta\Phi(l))^{-1}N_\Delta\Phi(\tilde{b}, l)D\Phi^*(\tau, \bar{l}), & t < \tau. \end{cases}$$

It now follows by direct verification that G_Δ as constructed satisfies the remaining five conditions.

4. **The limit function G.** In this section it will be shown that a type of limit function G exists for the set $\{G_\Delta\}$, as Δ approaches (a, b).

Let Δ_0, Δ_1, and Δ be closed bounded subintervals of (a, b) such that Δ_0 is properly contained in Δ_1, and Δ_1 is properly contained in Δ; these will be denoted by

$$\Delta_0 = [a_0, b_0],\ \Delta_1 = [a_1, b_1],\ \Delta = [\tilde{a}, \tilde{b}].$$

Let μ be a function, having a continuous first derivative, such that for some open interval Δ_2, $\Delta_0 \subset \Delta_2 \subset \Delta_1$

$$\mu(t) = \begin{cases} 1, & t \in \Delta_2 \\ 0, & t \text{ outside } \Delta_1. \end{cases}$$

Let

$$W_\Delta(t, \tau, l) = G_\Delta(t, \tau, l) - \mu(t)G_{\Delta_1}(t, \tau, l).$$

Then for $t, \tau \in \Delta_0$, $W_\Delta(t, \tau, l)$ is continuous; as a function of t, W_Δ satisfies

$$U_\Delta W_\Delta = M_\Delta W_\Delta(\tilde{a}) + N_\Delta W(\tilde{b}) = 0;$$

and also

$$(L_t - l)W_\Delta(t, \tau, l) = -A_0(t)\mu'(t)G_{\Delta_1}(t, \tau, l), \quad t \neq \tau.$$

Since $\mu'(t) = 0$ for t outside of Δ_1, W_Δ can be written as

$$W_\Delta(t, \tau, l) = -\int_{\Delta 1} G_\Delta(t, s, l)A_0(s)\mu'(s)G_{\Delta_1}(s, \tau, l)ds;$$

(Note: The integral over Δ_1 actually represents the sum of integrals over $[a, t-]$, $[t+, \tau-]$, $[\tau+, b]$ for $\tau > t$), or

$$G_\Delta(t, \tau, l) = \mu(t)G_{\Delta_1}(t, \tau, l) - \int_{\Delta_1} G_\Delta(t, s, l)A_0(s)\mu'(s)G_{\Delta_1}(s, \tau, l)ds.$$

It can be shown that the set $\{W_\Delta\}$ is uniformly bounded and equicontinuous on any compact (t, τ, l) – region, $\operatorname{Im} l \neq 0, t \neq \tau$. Thus, by Ascolis' theorem a uniform limit W exists and from this a limit function G, where

$$G = \mu G_{\Delta_1} + W,$$

and G is a limit function for the set $\{G_\Delta\}$.

THEOREM 4.1. *The function G satisfies the following conditions*:

(i) *$G(t, \tau, l)$ and $\partial/\partial t G(t, \tau, l)$ are continuous on $a < t \leqq \tau < b$ and $a < \tau \leqq t < b$, and for* $\operatorname{Im} l \neq 0$ *G is analytic in l*,

(ii) $G(t+, t, l) - G(t-, t, l) = A_0^{-1}(t), a < t < b$,

(iii) $L_t G = lG, t \neq \tau$,

(iv) $G(t, \tau, l) = G^*(\tau, t, \bar{l})$,

(v) $G(t, , l) \in \mathscr{L}^2(a, b), a < t < b$,

(vi) *If $f \in \mathscr{L}^2(a, b)$, then the vector v defined by*

$$v(t) = \int_a^b G(t, \tau, l)f(\tau)d\tau, \quad \operatorname{Im} l \neq 0,$$

is such that $v \in \mathscr{D}$ and

$$Lv(t) = lv(t) + f(t).$$

Proof. Again, this follows by direct verification.

It is thus seen that G satisfies all the conditions of a Green's function except for satisfying boundary conditions. Further, from property (vi), if

$$\mathscr{G}(l)f(t) = \int_a^b G(t, \tau, l)f(\tau)d\tau,$$

then,

$$(L - l)\mathscr{G}(l)f(t) = f(t)$$

and $\mathscr{G}(l)$ is a right inverse for $L - l$.

5. **The generalized resolvent.** Having constructed the closed symmetric operator T_0, all its self-adjoint extensions will now be considered. In § 3. the self-adjoint extensions for an operator in $\mathscr{H}$ having equal deficiency indices were considered and these self-adjoint extensions were also in the space $\mathscr{H}$. A spectral analysis of those self-adjoint extensions occurring in $\mathscr{H}$ was carried out, by quite different methods, by Brauer in [2]. The problem to be considered next is for unequal deficiency indices or equivalently, singular problems with equal deficiency indices such that the self-adjoint extensions are outside the original space.

Naimark [7] and others have defined extensions of T_0 for this case in larger Hilbert spaces. Theorem 7. in Straus [12] provides a means for an explicit construction in $\mathscr{H}$ itself. Let $A(l)$ map $\mathscr{E}(-i)$ into $\mathscr{E}(i)$, where $A(l)$ is analytic and $\|A(l)\| \leqq 1$ for $\operatorname{Im} l > 0$. Analogously to the case of equal deficiency indices, a domain $\mathscr{D}(l) \subset \mathscr{D}$ is defined by

$$\mathscr{D}(l) = \mathscr{D}_0 + (I - A(l))\mathscr{E}(-i),$$

and an operator $T_{A(l)}$, having domain $\mathscr{D}(l)$ is defined by

$$T_{A(l)}u = Tu, \quad u \in \mathscr{D}(l).$$

Then, $T_0 \subset T_{A(l)} \subset T$, and the generalized resolvent $\mathscr{R}$ can be represented as

$$\mathscr{R}(l) = (T_{A(l)} - lI)^{-1}, \mathscr{R}(\bar{l}) = \mathscr{R}^*(l), \operatorname{Im} l > 0.$$

Further, from Straus [12] every generalized resolvent is generated by such $A(l)$.

Again, analogously to the case for equal deficiency indices, the domain $\mathscr{D}(l)$ can be characterized in an alternate manner which leads to an explicit formulation for the generalized resolvent $\mathscr{R}(l)$. The domain $\mathscr{D}$ can be represented as a direct sum

$$\mathscr{D} = \mathscr{D}_0 + \mathscr{E}(i) + \mathscr{E}(-i);$$

let ω^+ be the dimension of $\mathscr{E}(i)$, and ω^- be the dimension of $\mathscr{E}(-i)$, where $0 \leqq \omega^+, \omega^- \leqq n$; let $\{\varphi_j(i)\}$, $j = 1, \cdots, \omega^+$, be an orthonormal basis for $\mathscr{E}(i)$, and let $\{\psi_k(-i)\}, k = 1, \cdots, \omega^-$, be an orthonormal basis for $\mathscr{E}(-i)$; finally, let

$$v_j(l) = \psi_j(-i) - A(l)\psi_j(-i),$$

and

$$v_{j*}(l) = \varphi_j(i) = A^*(l)\varphi_j(i).$$

THEOREM 5.1. *For* $\operatorname{Im} l > 0$, *the domain* $\mathscr{D}(l)$ *of* $T_{A(l)}$ *can be represented as*

$$\mathscr{D}(l) = \{u \in \mathscr{D}: \langle uv_{j*}(l)\rangle = 0, \quad j = 1, \cdots, \omega^+\};$$

and the domain of $T^*_{A(l)}$ *is*

$$\mathscr{D}^*(l) = \{w \in \mathscr{D}: \langle wv_k(l)\rangle = 0, \quad k = 1, \cdots, \omega^-\}.$$

Proof. The proof is the same as the proof of Theorem 3.1 with the operator $A(l)$ in place of U.

It will now be shown, again analogous to § 3, that the domain $\mathscr{D}(l)$ induces limiting abstract boundary conditions. For $u \in \mathscr{D}$ and any closed bounded subinterval $[c, d]$ of (a, b)

$$[uv_{j*}](d) - [uv_{j*}](c) = v^*_{j*}(d)A_0(d)u(d) - v^*_{j*}(c)A_0(c)u(c).$$

Since u, v_{j*}, Lu, and Lv_{j*} are each in $\mathscr{L}^2(a, b)$, then $\lim_{d\to b} v^*_{j*}(d)A_0(d)u(d)$ exists, and $\lim_{c\to a} v^*_{j*}(c)A_0(c)u(c)$ exists; these limits will be denoted by $v^*_{j*}(b)A_0(b)u(b)$ and $v^*_{j*}(a)A_0(a)u(a)$. The conditions $\{\langle uv_{j*}\rangle = 0, j = 1, \cdots, \omega^+\}$ can then be represented in matrix form as

$$0 = \langle uV_*\rangle = V^*_*(b, l)A_0(b)u(b) - V^*_*(a, l)A_0(a)u(a)(LB),$$

where V_* is the matrix with v_{j*} in the jth column; these are limiting abstract boundary conditions.

Having obtained the limiting abstract boundary conditions, the following theorem describes a method for the construction of the generalized resolvent $\mathscr{R}(l)$ starting from the integral operator $\mathscr{G}(l)$ developed in § 4.

THEOREM 5.2. *Each generalized resolvent* $\mathscr{R}(l)$ *of* T_0 *is an integral operator of Carleman type, having a kernel* $R(t, \tau, l)$, *which is continuous in* (t, τ, l) *and analytic in* l *in any region for which* $\operatorname{Im} l \neq 0$, *and* $t \neq \tau$.

Proof. The integral operator $\mathscr{G}(l)$ obtained in § 4. is of Carleman type; further $\mathscr{G}(l)$ satisfies the conditions of the theorem except that $\mathscr{G}(l)$ is only a right inverse of $(T - l)$. It will now be shown that a matrix G_1 can be constructed such that the kernel of $\mathscr{R}(l)$ is

$$R(t, \tau, l) = G(t, \tau, l) + G_1(t, \tau, l).$$

For fixed l, $\operatorname{Im} l > 0$, let $\{\theta_i(l)\}$, $i = 1, \cdots, \omega^+$, be an orthonormal basis for $\mathscr{E}(l)$, let $\Theta_+(l)$ be the $n \times \omega^+$ matrix having $\theta_j(l)$ in the jth column, similarly, let $\{\chi_k(\bar{l})\}$, $k = 1, \cdots, \omega^-$, be an orthonormal basis for $\mathscr{E}(\bar{l})$, and $\chi_-(\bar{l})$ be the $n \times \omega^-$ matrix having $\chi_k(\bar{l})$ in the kth column. From the orthonormal property of the $\theta_i(l)$ and the $\chi_k(\bar{l})$,

$$\begin{aligned}(\Theta_+(l), \Theta_+(l)) &= \int_a^b \Theta_+^*(t, l)\Theta_+(t, l)dt \\ &= I_{\omega^+ \times \omega^+},\end{aligned}$$

where $I_{\omega^+ \times \omega^+}$ is the identity matrix of rank ω^+; similarly,

$$(\chi_-(\bar{l}), \chi_-(\bar{l})) = I_{\omega^- \times \omega^-}.$$

For any vector f in $\mathscr{L}^2(a, b)$, $(T - l)(\mathscr{R}(l) - \mathscr{G}(l))f) = f - f = 0$, and thus $(\mathscr{R}(l) - \mathscr{G}(l))f$ is in $\mathscr{E}(l)$. Thus for some $\omega^+ \times 1$ vector $a(f, l)$

$$(\mathscr{R}(l) - \mathscr{G}(l))f = \Theta_+(l)a(f, l).$$

Also

$$((\mathscr{R}(l) - \mathscr{G}(l)f, \Theta_+(l)) = (f, (\mathscr{R}(\bar{l}) - \mathscr{G}(\bar{l}))\Theta_+(l));$$

and, for each column $\theta_k(l)$ of $\Phi_+(l)$, $(T - \bar{l})(\mathscr{R}(\bar{l}) - \mathscr{G}(\bar{l}))\theta_k(l) = \theta_k(l) - \theta_k(l) = 0$. Thus, for some $\omega^+ \times \omega^-$ matrix $B(l)$

$$(\mathscr{R}(\bar{l}) - \mathscr{G}(\bar{l}))\Phi_+(l) = \chi_-(\bar{l})B^*(l).$$

Combining the preceding calculations yields

$$(\mathscr{R}(l) - \mathscr{G}(l))f(t) = \Phi_+(t, l)B(l)(f, \chi_-(\bar{l})) = (f, \chi(\bar{l})B^*(l)\Phi_+^*(t, l));$$

thus,

$$R(t, \tau, l) - G(t, \tau, l) = \Phi_+(t, l)B(l)\chi_-^*(\tau, \bar{l}).$$

Similarly, for some $\omega^+ \times \omega^-$ matrix $H(l)$,

$$R(t, \tau, \bar{l}) - G(t, \tau, \bar{l}) = \chi_-(t, \bar{l})H^*(l)\Phi_+^*(t, \tau)$$

and

$$R^*(\tau, t, \bar{l}) - G^*(\tau, t, \bar{l}) = \Phi_+(t, l)H(l)\chi_-^*(\tau, \bar{l}).$$

Further,

$$\begin{aligned} B(l) &= (\chi_-(\bar{l}), \chi_-(\bar{l})B^*(l)) \\ &= (\chi_-(\bar{l}), (\mathscr{R}(\bar{l}) - \mathscr{G}(\bar{l}))\Phi_+(l)) \\ &= ((\mathscr{R}(l) - \mathscr{G}(l))\chi_-(\bar{l}), \Phi_+(l)) \\ &= (\Theta_+(l)H(l), \Theta_+(l)) \\ &= H(l), \end{aligned}$$

so that

$$R^*(\tau, t, \bar{l}) - G^*(\tau, t, \bar{l}) = R(t, \tau, l) - G(t, \tau, l).$$

Since $\theta_j(l) \in \mathscr{E}(l) \subset \mathscr{D}$, for $j = 1, \cdots, \omega^+$, and $\chi_k(l) \in \mathscr{E}(\bar{l}) \subset \mathscr{D}$, for $k = 1, \cdots, \omega^-$, then $\Theta_+(l)B(l)\chi_-^*(\tau, \bar{l}) \in \mathscr{L}^2(a, b)$ for $a < \tau < b$, $\operatorname{Im} l \neq 0$. Thus $\mathscr{R}(l)$ is an integral operator of Carleman type.

The operator $\mathscr{R}(l)$ will completely satisfy the condition of the theorem when it is shown that $R(t, \tau, l)$ is analytic in l, $\operatorname{Im} l \neq 0$, and $t \neq \tau$. To facilitate the proof of the analyticity of $R(t, \tau, l)$, analytic bases for $\mathscr{E}(l)$ and $\mathscr{E}(\bar{l})$ will be introduced, as in Coddington [5], related to an arbitrary l_0, $\operatorname{Im} l_0 > 0$.

Matrices $\Psi_-(\bar{l})$and $\Phi_+(l)$ are defined by this process such that the columns of $\Phi_+(l)$ form a basis for $\mathscr{E}(l)$ and the columns of $\Psi_-(\bar{l})$ form a basis for $\mathscr{E}(\bar{l})$; thus for some nonsingular matrix $T(\bar{l})$,

$$\Psi_-(\bar{l}) = \chi_-(\bar{l})T(\bar{l}),$$

and for some nonsingular matrix $S(l)$

$$\Phi_+(l) = \Theta_+(l)S(l).$$

Thus,

$$\begin{aligned} \Theta_+(t, l)B(l)\chi_-^*(\tau, \bar{l}) &= \Phi_+(t, l)S^{-1}(l)B(l)(T^*(\bar{l}))^{-1}\chi_-^*(\tau, \iota), \\ &= \Phi_+(t, l)C(l)\Psi_-(\tau, \bar{l}), \end{aligned}$$

where $C(l) = S^{-1}(l)B(l)(T^*(\bar{l}))^{-1}$. The matrix $\Phi_+(l)$ is analytic in l and $\Psi_-^*(\bar{l})$ is analytic in l for any compact subset of $\operatorname{Im} l \neq 0$. Thus it remains to show that $C(l)$ satisfies the same conditions of analyticity.

Let Z be an $n \times r$ matrix each of whose columns z_k is in $\mathscr{L}^2(a, b)$, $k = 1, \cdots, r$. Then $\mathscr{R}(l)z_k$ is in $\mathscr{D}(l)$ and thus satisfies the boundary condition (LB),

$$0 = \langle(\mathscr{R}(l)z_k)V_*(l)\rangle;$$

the set $\{0 = \langle(\mathscr{R}(l)z_k)V_*(l)\rangle,\ k = 1, \cdots, r\}$ can be written in matrix form as

$$0 = \langle(\mathscr{R}(l)Z)V_*(l)\rangle.$$

Expanding $\mathscr{R}(l)$, yields

$$0 = \langle(\mathscr{G}(l)Z)V_*(l)\rangle + \langle(\Phi_+(l)C(l)(Z, \Psi_-(\bar{l}))) V_*(l)\rangle,$$

also

$$\langle(\Phi_+(l)C(l)(Z, \Psi_-(\bar{l})))V_*(l)\rangle = \langle\Phi_+(l)V_*(l)\rangle C(l)(Z, \Psi_-(\bar{l})).$$

Thus,

$$-\langle(\mathscr{G}(l)Z)V_*(l)\rangle = \langle\Phi_+(l)V_*(l)\rangle C(l)(Z, \Psi_-(\bar{l})).$$

The matrix $C(l)$ will be analytic if $\langle(\mathscr{G}(l)Z)V_*(l)\rangle$ is analytic, and $\langle\Phi_+(l)V_*(l)\rangle$ and $(Z, \Psi_-(\bar{l}))$ are each nonsingular and analytic.

First it can be shown that $\langle\Phi_+(l)V_*(l)\rangle$ is nonsingular and analytic. Next, for $(Z, \Psi_-(\bar{l}))$ to be nonsingular Z must be an $n \times \omega^-$ matrix, it can be verified that for $Z = \Psi_-(-i)$, $(Z, \Psi_-(\bar{l}))$ is nonsingular and analytic. Finally, $\langle(\mathscr{G}(l)\Psi_-(-i)V_*(l)\rangle$ is analytic in l for $|l - l_0| < \operatorname{Im} l_0/2$. Thus

$$C(l) = -\langle\Phi_+(l)V_*(l)\rangle^{-1}\langle(\mathscr{G}(l)\Psi_-(-i)V_*(l)\rangle(\Psi_-(-i), \Psi_-(l))^{-1}$$

is analytic and

$$\Phi_+(t, l)C(l)\Phi_-^*(\tau, \bar{l}) = \Theta_+(t, l)B(l)\chi_-^*(\tau, \bar{l})$$

is analytic in l in a compact subset of $\operatorname{Im} l \neq 0$, $|l - l_0| < \operatorname{Im} l_0/2$. Theorem 5.2 is now proved, the generalized resolvent $\mathscr{R}(l)$ with kernel $\mathscr{R}(t, \tau, l)$ has been constructed.

6. The spectral matrix.

DEFINITION. A matrix ρ, (associated with an eigenvalue problem) is a spectral matrix if it satisfies:

(i) ρ is Hermitian,

(ii) $\rho(\Delta) = \rho(\lambda) - \rho(\mu) \geqq 0$ if $\lambda > \mu$, (where $\Delta = [\mu, \lambda]$),

(iii) ρ is of bounded variation on every finite λ interval.

To develop the spectral matrix associated with the problem $(L - l)u = 0$ with the boundary conditions (LB), and thus associated with the generalized resolvent $\mathscr{R}$ and the generalized resolution of the identity F, the kernel of $\mathscr{R}(l)$ will be split into two parts,

$$R(t, \tau, l) = R_0(t, \tau, l) + R_1(t, \tau, l)$$

where $R_0(t, \tau, l)$ is a certain fundamental matrix for $(L - l)u = 0$. Once again, let Φ be a fundamental matrix for $(L - l)u = 0$, satisfying $\Phi(c, l) = I$, for some c, $a < c < b$. Then, as shown in § 3,

$$[\Phi(t, l)\Phi(t, \bar{l})] = \Phi^*(t, \bar{l})A_0(t)\Phi(t, l) = D^{-1},$$

where D is a nonsingular, constant, skew-Hermitian matrix. Defining

$R_0(t, \tau, l)$ by

$$R_0(t, \tau, l) = \begin{cases} \dfrac{1}{2}\Phi(t, l)D\Phi^*(\tau, \bar{l}), & t \geqq \tau, \\ \dfrac{1}{2}\Phi(t, l)D^*\varphi^*(t, \bar{l}), & t < \tau , \end{cases}$$

then,

$$R_0^*(\tau, t, \bar{l}) = R_0(t, \tau, l).$$

Also,

$$R_0(t+, t, l) - R_0(t-, t, l) = A_0^{-1}(t),$$

which is the same jump that $R(t, \tau, l)$ has at $t = \tau$.

Now, let

$$R_1(t, \tau, l) = R(t, \tau, l) - R_0(t, \tau, l).$$

Then as a function of t, R_1 has a continuous first derivative, and $(L_t - l)R_1(t, \tau, l) = 0$.

From the symmetry property $R_1^*(\tau, t, \bar{l}) = R_1(t, \tau, l)$ it follows that for some matrix $\Psi(l)$

$$R_1(t, \tau, l) = \Phi(t, l)\Psi(l)\Phi^*(\tau, \bar{l}).$$

THEOREM 6.1. *The matrix* Ψ *is analytic for* $\operatorname{Im} l > 0$, $\Psi^*(l) = \Psi(\bar{l})$, *and* $\operatorname{Im}\Psi(l)/\operatorname{Im} l > 0$, *where* $\operatorname{Im}\Psi = (\Psi - \Psi^*)/2i$.

Proof. The analyticity of Ψ follows from the choice of $\Phi(c, l) = I$.

Next,

$$R_1(t, \tau, l) = R_1^*(\tau, t, \bar{l}),$$

implying

$$\Phi(t, l)\Psi(l)\Phi^*(\tau, \bar{l}) = \Phi(t, l)\Psi^*(\bar{l})\Phi^*(\tau, \bar{l}),$$

and, since Φ^{-1} exists, $\Psi(l) = \Psi^*(\bar{l})$, or $\Psi^*(l) = \Psi(\bar{l})$.

Let

$$H(t, \tau, l) = \frac{R(t, \tau, l) - R(t, \tau, \bar{l})}{2i},$$

direct computation yields

$$H(c, c, l) = \operatorname{Im}\Psi(l).$$

The proof for $\operatorname{Im}\Psi(l)/\operatorname{Im} l \geqq 0$ now follows as in the proof of Theorem 3 of Coddington [5].

THEOREM 6.2. *The matrix* ρ *defined by*

$$\rho(\lambda) = \lim_{\varepsilon \to +0} \frac{1}{\pi} \int_0^\lambda \operatorname{Im} \Psi(\nu + i\varepsilon) d\nu$$

exists, is nondecreasing and is of bounded variation on any finite interval.

Proof. This follows directly from Theorem 4 of Coddington [5]. The matrix ρ is the spectral matrix associated with the generalized resolvent $\mathscr{R}$ and the generalized resolution of the identity F.

7. **The generalized resolution of the identity.** Let ρ be the spectral matrix derived in § 6, let $\Delta = (\mu, \lambda]$ be a finite interval, and let $F(\Delta) = F(\lambda) - F(\mu)$.

THEOREM 7.1. *Let* $f \in \mathscr{H}$ *and vanish outside a closed bounded subinterval* $[c, d]$ *of* (a, b). *If* μ *and* λ *are continuity points of* F, *then*

$$F(\Delta) f(t) = \int_\Delta \Phi(t, \nu) d\rho(\nu) (f, \Phi(\nu)).$$

Proof. It follows from the relationship

$$(\mathscr{R}(l) f, f) = \int_{-\infty}^{\infty} \frac{d(F(\gamma) f, f)}{\gamma - l}$$

that

$$(F(\Delta) f, f) = \lim_{\varepsilon \to +0} \frac{1}{\pi} \int_\Delta (\operatorname{Im} \mathscr{R}(\nu + i\varepsilon) f, f) d\nu$$

at continuity points μ, λ of F. The generalized resolvent $\mathscr{R}(l)$ can be written as $\mathscr{R}(l) = \mathscr{R}_0(l) + \mathscr{R}_1(l)$, where $\mathscr{R}_0(l)$ has kernel $R_0(t, \tau, l)$, and $\mathscr{R}_1(l)$ has kernel $R_1(t, \tau, l)$. Then

$$\mathscr{R}_0(l) f(t) = \int_a^b R_0(t, \tau, l) f(\tau) d\tau = \int_c^d R_0(t, \tau, l) f(\tau) d\tau.$$

However, $(\operatorname{Im} \mathscr{R}_0(\nu + i\varepsilon) f, f)$ tends to zero $\varepsilon \longrightarrow +0$, uniformly in Δ. Consequently, it follows that

$$(F(\Delta) f, f) = \lim_{\varepsilon \to +0} \frac{1}{\pi} \int_\Delta (\operatorname{Im} \mathscr{R}_1(\nu + i\varepsilon) f, f) d\nu,$$

where

$$\begin{aligned}
&(\operatorname{Im} \mathscr{R}_1(\nu + i\varepsilon)f, f) \\
&\quad = (\Phi(\nu), f)\left[\frac{\Psi(\nu + i\varepsilon) - \Psi(\nu - i\varepsilon)}{2i}\right](f, \Phi(\nu)) \\
&\qquad + \frac{1}{2i}\{\Phi(\nu + i\varepsilon), f)\Psi(\nu + i\varepsilon)(f, (\Psi(\nu - i\varepsilon) - \Phi(\nu))) \\
&\qquad + ((\Phi(\nu + i\varepsilon) - \Phi(\nu)), f)\Psi(\nu + i\varepsilon)(f, \Psi(\nu))\} \\
&\qquad + \frac{1}{2i}\{(\Phi(\nu), f)\Psi(\nu - i\varepsilon)(f, (\Phi(\nu) - \Phi(\nu + i\varepsilon))) \\
&\qquad + ((\Phi(\nu) - \Phi(\nu - i\varepsilon)), f)\Psi(\nu - i\varepsilon)(f, \Phi(\nu + i\varepsilon))\} \\
&\quad = T_1 + T_2 + T_3,
\end{aligned}$$

where $T_i = T_i(\nu, \varepsilon, f)$.

In Lemma 3 of Straus [13], it is shown that

$$\lim_{\varepsilon \to +0} \frac{1}{\pi}\int_\Delta T_2(\nu)d\nu = \lim_{\varepsilon \to +0} \frac{1}{\pi}\int_\Delta T_3(\nu)d\nu = 0.$$

Finally, for T_1,

$$\begin{aligned}
&\lim_{\varepsilon + 0}\frac{1}{\pi}\int_\Delta (\Phi(\nu), f) \operatorname{Im} \Psi(\nu + i\varepsilon)(f, \Phi(\nu))d\nu \\
&\qquad = \int_\Delta (\Phi(\nu), f)d\rho(\nu)(f, \Phi(\nu)) \\
&\qquad = \int_\Delta (f, \Phi(\nu))^* d\rho(\nu)(f, \Phi(\nu)),
\end{aligned}$$

and therefore,

$$(F(\Delta)f, f) = \int_a^b f^*(t)\left[\int_\Delta \Phi(t, \nu)d\rho(\nu)(f, \Phi(\nu))\right]dt.$$

Since this representation must hold for all $f \in \mathscr{H}$ which vanish outside closed finite subintervals of (a, b),

$$F(\Delta)f(t) = \int_\Delta \Phi(t, \nu)d\rho(\nu)f, \Phi(\nu))$$

for all such f.

Thus the generalized resolutions of the identity associated with the first order system of differential operators $Lx(t) = A_0(t)x'(t) + A(t)x(t)$ can be represented explicitly in terms of a certain fundamental matrix Φ and an associated spectral matrix ρ.

8. The expansion and completeness relations. Expansion and completeness relations can be defined in terms of the spectral matrix

ρ and the fundamental matrix Φ. For two vectors $\hat{\alpha}$, $\hat{\beta}$, an inner product is defined in terms of ρ by

$$(\hat{\alpha}, \hat{\beta})_\rho = \int_{-\infty}^{\infty} \hat{\beta}^*(\nu) d\rho(\nu) \hat{\alpha}(\nu).$$

Thus a norm can be defined by

$$||\hat{\alpha}||_\rho = (\hat{\alpha}, \hat{\alpha})_\rho^{1\backslash 2}.$$

The Hilbert space $\mathscr{L}^2(\rho)$ is defined by

$$\mathscr{L}^2(\rho) = \{\hat{\alpha}: ||\hat{\alpha}||_\rho < \infty\}.$$

Defining a mapping from $\mathscr{L}^2(a, b)$ into $\mathscr{L}^2(\rho)$ by

$$\hat{f}(\nu) = (f, \Phi(\nu)) = \int_a^b \Phi^*(t, \nu) f(t) dt,$$

the expansion and completeness relations have the following form:

$$f(t) = (\hat{f}, \Phi^*(t))_\rho = \int_{-\infty}^{\infty} \Phi(t, \nu) d\rho(\nu) \hat{f}(\nu) \quad \text{(expansion)}$$

and

$$||f|| = ||\hat{f}||_\rho \quad \text{(completeness)}.$$

References

1. N. I. Akhiezer and I. M. Glazman, *Theory of Linear Operators in Hilbert Space* (English Translation) F. Ungar Publishing Co., N. Y., Vol. 1 and Vol. 2, 1961, 1963.
2. F. Brauer, *Spectral theory for linear systems of differential equations*, Pacific J. Math., V. **10** (1960), 17-34.
3. E. A. Coddington, *The spectral representation of ordinary self-adjoint differential operators*, Annals of Math., **60** (July 1954), 192-211.
4. _______, *The spectral matrix and Green's function for singular self-adjoint boundary value problems*, Canad. Math., **6** (1954), 169-185.
5. _______, *Generalized resolutions of the identity for symmetric ordinary differential operators*, Annals of Math., **68** (Sept. 1958) 378-392.
6. _______, and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill Book Co., Inc., 1955.
7. M. A. Naimark, *Spectral functions of a symmetric operator*, Izvest. Akad. Nauk. S. S. S. R. Ser. Mat., **4** (1940), 277-318.
8. F. Riesz and B. Sz-Nagy, *Functional Analysis*, (English Translation) F. Ungar Publishing Co., N. Y., 1955.
9. M. H. Stone, *Linear Transformations in Hilbert Space*, Amer. Math. Soc. Colloqu. Publ., 1932.
10. A. V. Straus, *On generalized resolvents of a symmetric operator*, Dok. Akad. Nauk. S. S. S. R., **71** (1950), 241-244.
11. _______, *On the theory of generalized resolvents of a symmetric operator*, Dok. Akad. Nauk. S. S. S. R., **78** (1951), 217-220.

12. _______, *Generalized resolvents of symmetric operators,* Izvest. Akad. Nauk. S. S. S. R., Ser. Mat., **18** (1954), 51-86.
13. _______, *On spectral functions of differential operators,* Izvest. Akad. Nauk S. S. S. R., Ser. Mat., **19** (1955), 201-220.

Received September 4, 1970 and in revised form June 19, 1971.

PACIFIC MISSILE RANGE, POINT MAGU, CALIFORNIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

ON SPLITTING IN HEREDITARY TORSION THEORIES

ROBERT L. BERNHARDT

Let $(\mathscr{T}, \mathscr{F})$ denote a hereditary torsion theory for the category of modules over a ring R. In this paper the splitting of projective modules is studied, and it is shown that this is not equivalent to the splitting of quasi-projective modules. In addition, situations arising from the class of torsion modules $\mathscr{T}$ (or the class of torsionfree modules $\mathscr{F}$) being contained in the injective or in the projective modules are considered, and several conditions sufficient for an especially strong form of splitting are given. Finally when $\mathscr{T}$ is closed under injective envelopes the following is shown: every module splits if R is an artinian generalized uniserial ring, and projective modules split if R is a QF-2 ring.

The term "ring" will mean an associative ring with unity 1, and all modules are assumed to be unitary left modules. We denote the category of all modules over a ring R by ${}_R\mathscr{M}$. Dickson [6] defined a torsion theory for ${}_R\mathscr{M}$ to be a pair $(\mathscr{T}, \mathscr{F})$ of classes of modules satisfying the following:

(a) $\mathscr{T} \cap \mathscr{F} = 0$;

(b) $\mathscr{T}$ is closed under homomorphic images and $\mathscr{F}$ is closed under submodules;

(c) For each module M there exists a (unique) submodule $M_t \in \mathscr{T}$ such that $M/M_t \in \mathscr{F}$.

A torsion theory $(\mathscr{T}, \mathscr{F})$ is said to be *hereditary* if $\mathscr{T}$ is closed under submodules, and *stable* if $\mathscr{T}$ is closed under injective envelopes. We remark that from (b) above it is clear that $\operatorname{Hom}(T, F) = 0$ for all $T \in \mathscr{T}$ and all $F \in \mathscr{F}$; also Dickson has shown that $\mathscr{T}$ is closed under submodules if and onlf if $\mathscr{F}$ is closed under injective envelopes. In this paper we shall always be concerned with hereditary torsion theories.

If $\mathscr{T}$ is a hereditary torsion class, then Gabriel [8] has shown that $\mathscr{T}$ is uniquely associated with an (topologizing and) idempotent *filter*

$F(\mathscr{T}) = \{L \subseteqq R \mid L$ is a left ideal of R and $R/L \in \mathscr{T}\}$. Moreover, $\mathscr{T}$ is a torsionfree class for some torsion class $\mathscr{C}$ if and only if $F(\mathscr{T})$ contains a unique minimal left ideal (see [9]); in this case Jans has called $\mathscr{T}$ a *torsion-torsionfree* (TTF) *class*, and we shall call $(\mathscr{T}, \mathscr{F})$ and $(\mathscr{C}, \mathscr{T})$ the torsion theories *associated* with $\mathscr{T}$. If R is a right perfect ring, Alin [1] has shown that every hereditary torsion class for ${}_R\mathscr{M}$ is a TTF class.

If $(\mathscr{T}, \mathscr{F})$ is a hereditary torsion theory for ${}_R\mathscr{M}$ and if $M \in {}_R\mathscr{M}$, we say that M *splits* provided $M = M_t \oplus M'$; we shall call $(\mathscr{T}, \mathscr{F})$ *splitting* if every module in ${}_R\mathscr{M}$ splits. We say that $(\mathscr{T}, \mathscr{F})$ is *centrally splitting* provided $\mathscr{T}$ is a TTF class with associated torsion theories $(\mathscr{T}, \mathscr{F})$ and $(\mathscr{C}, \mathscr{T})$, and $M = M_t \oplus M_c$ (i.e., M is the direct sum of its two torsion submodules) for every $M \in {}_R\mathscr{M}$. Centrally splitting is clearly a strong form of splitting; the interested reader may see [5] for more information on this topic.

1. Splitting in projective modules. In this section we shall study the dual for projective modules to the following result of Armendariz [3] on the splitting of injective modules. We denote the injective envelope of a module M by $E(M)$.

THEOREM A (Armendariz). *If $(\mathscr{T}, \mathscr{F})$ is a hereditary torsion theory, then the following are equivalent:*

(1) *$\mathscr{T}$ is stable;*

(2) *Every injective module splits;*

(3) *Every quasi-injective module splits;*

(4) *$E(M_t) = E(M)_t$ for every $M \in {}_R\mathscr{M}$.*

If N is a submodule of the module M, we call N *invariant* in M provided that $f(N) \subseteq N$ for every endomorphism f of M. We call N *small* in M provided that if K is a submodule of M and if $K + N = M$, then $K = M$. We shall say that a class $\mathscr{C}$ of modules is *closed under projective covers* provided that whenever $M \in \mathscr{C}$ has a projective cover $P(M)$, then $P(M) \in \mathscr{C}$.

THEOREM 1.1. *Let $(\mathscr{T}, \mathscr{F})$ be a hereditary torsion theory for ${}_R\mathscr{M}$ such that every torsionfree module has a projective cover. Then the following are equivalent:*

(1) *$\mathscr{F}$ is closed under projective covers;*

(2) *Every projective module splits.*

Proof. (1) $\rightarrow$ (2): Let Q be a projective module, and let $\pi: P(Q/Q_t) \rightarrow Q/Q_t$ be the projective cover of Q/Q_t. Let n be the natural epimorphism of Q onto Q/Q_t. By [4, Lemma 2.3] there exists a monomorphism $h: P(Q/Q_t) \rightarrow Q$ such that $nh = \pi$ and such that $Q = Imh + Q'$, where $Q' \subseteq \operatorname{Ker} n = Q_t$. But Imh is torsionfree; so that $Imh \cap Q_t = 0$ and $Q = Imh \oplus Q_t$.

(2) $\rightarrow$ (1): Choose $M \in \mathscr{F}$, and let $\pi: P(M) \rightarrow M$ be the projective cover of M. Then $P(M)_t \subseteq \operatorname{Ker} \pi$, so that $P(M)_t$ is small in $P(M)$. But $P(M)$ splits by hypothesis; thus $P(M) \in \mathscr{F}$.

EXAMPLE 1.2. The splitting of projective modules does not imply the splitting of quasi-projective modules in left artinian generalized uniserial rings.

Let K be a field, and let R be the ring of 4×4 upper triangular matrices over K. Let

$$I = \left\{ \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & 0 \end{vmatrix} | a_{ij} \in K \right\} ;$$

then I is an idempotent, two-sided ideal of R. Thus by a result of Jans [9], $\mathscr{T} = \{M \in {}_R\mathscr{M} | IM = 0\}$ is a TTF class. Further $R \in \mathscr{F}$, so that every free module is torsionfree. Hence every projective module is torsionfree, and thus splits. Now let e_{ij} denote the matrix with 1 in the i^{th} row and j^{th} column and 0 elsewhere, and let $J = Re_{14}$. Then J is a two-sided ideal of R, and hence $M = Re_{44}/Je_{44} = Re_{44}/J$ is quasi-projective and indecomposable. But $M \notin \mathscr{T}$, and $Re_{24}/J \subseteqq M_t$. Thus M_t is a nontrivial submodule of M.

We next turn our attention to the quasi-projective cover; this was introduced in [12], and there it was shown that a sufficient (but not necessary) condition for the quasi-projective cover of a module M to exist is that the projective cover of M exist.

PROPOSITION 1.3. *Let M be a quasi-projective module which has a projective cover. If N is an invariant submodule of M, then the module M/N is quasi-projective.*

Proof. Let $\pi: P(M) \to M$ be the projective cover of M, and choose an endomorphism f of $P(M)$. By [12, Proposition 2.2], f induces an endomorphism g of M such that $g\pi = \pi f$. Let $K = \pi^{-1}(N)$; then $\pi f(K) = g\pi(K) = g(N) \subseteqq N$, and hence $f(K) \subset \pi^{-1}(N) = K$. We have shown that K is invariant in $P(M)$; thus by [12, Proposition 2.1] we have $P(M)/K \cong M/N$ is quasi-projective.

THEOREM 1.4. *Let M be a module with a projective cover, let $\pi': QP(M) \to M$ denote the quasi-projective cover of M, and let $(\mathscr{T}, \mathscr{F})$ be a hereditary torsion theory for ${}_R\mathscr{M}$. If $M \in \mathscr{F}$, then $QP(M) \in \mathscr{F}$.*

Proof. Let $\pi: P(M) \to M$ denote the projective cover of M; by [12, Propositions 2.6, 2.1 and 2.2] we have that $QP(M) \cong P(M)/X$, where X is the unique maximal invariant submodule of $P(M)$ contained in Ker π. Let n denote the natural epimorphism of $P(M)$ onto

$QP(M)$. Since $\operatorname{Ker} n \subseteqq \operatorname{Ker} \pi$, we have that $\operatorname{Ker} n$ is small in $P(M)$, and thus $n: P(M) \to QP(M)$ is the projective cover of $QP(M)$. Further $QP(M)_t \subseteqq \operatorname{Ker} \pi'$ since $M \in \mathscr{F}$, and also $QP(M)_t$ is invariant in $QP(M)$. Hence $QP(M)/QP(M)_t$ is quasi-projective by Proposition 1.3; thus $QP(M)_t = 0$ by condition (3) of the definition of the quasi-projective cover in [12].

2. Classes of projective and injective modules. Let $(\mathscr{T}, \mathscr{F})$ be a hereditary torsion theory for ${}_R\mathscr{M}$. In this section we investigate the following condition:

$\mathscr{T}$ is stable and all torsionfree modules are injective. This has been studied previously in [3] (also see [2] for the special case that $\mathscr{T}$ is the Goldie torsion class), where it was shown to imply that $(\mathscr{T}, \mathscr{F})$ is splitting. In Theorem 2.2 we shall give a statement equivalent to this condition, and, in addition, we shall show that it implies the much stronger result: $(\mathscr{T}, \mathscr{F})$ is centrally splitting. Finally we shall obtain a dual to Theorem 2.2.

LEMMA 2.1. *Let $(\mathscr{T}, \mathscr{F})$ be a hereditary torsion theory for ${}_R\mathscr{M}$, let $R = R_t \oplus K$, and let $\mathscr{F}$ be closed under homomorphic images. Then $R = R_t \dotplus K$(ring direct sum), $\mathscr{T}$ is a* TTF *class, and $(\mathscr{T}, \mathscr{F})$ is centrally splitting.*

Proof. Since right multiplication by an element of R is a left R-homomorphism on K, and since $\mathscr{F}$ is closed under homomorphic images, K is a two-sided ideal of R and $R = R_t \dotplus K$.

By [5, Theorem 1] it now suffices to see that $\mathscr{T}$ is a TTF class. Choose $L \in F(\mathscr{T})$; then $K \cap L \in F(\mathscr{T})$, and hence $R/K \cap L \in \mathscr{T}$. Thus $K/K \cap L \in \mathscr{T}$. But $K \to K/K \cap L \to 0$ is exact and $K \in \mathscr{F}$; thus $K/K \cap L \in \mathscr{T} \cap \mathscr{F} = 0$ and $K = K \cap L \subseteqq L$. We have shown that K is the unique minimal ideal in $F(\mathscr{T})$; thus $\mathscr{T}$ is a TTF class.

THEOREM 2.2. *If $(\mathscr{T}, \mathscr{F})$ is a hereditary torsion theory for ${}_R\mathscr{M}$, then the following are equivalent:*

(1) *$\mathscr{T}$ is stable, and all torsionfree modules are injective;*

(2) *$\mathscr{F}$ is closed under homomorphic images, and all torsionfree modules are projective;*

(3) *$\mathscr{F}$ is clossed under homomorphic images, $R = R_t \dotplus K$ (ring direct sum), and K is a semi-simple ring with minimum condition.*

In addition, whenever (1), (2), *and* (3) *are true, then $\mathscr{T}$ is a* TTF *class and $(\mathscr{T}, \mathscr{F})$ is centrally spilitting.*

Proof. (1) $\to$ (3) follows from [3, Theorem 3.2].

(3) $\to$ (2): If M is a torsionfree module, then $R_t M = 0$ and hence

M is a projective K-module. But now M is a direct summand of a free K-module, and hence M is a direct summand of a free R-module. Thus M is projective as an R-module.

$(2) \to (1)$: Choose $M \in \mathscr{T}$, and let n be the natural epimorphism of $E(M)$ onto $E(M)/E(M)_t$. Since this torsionfree module is projective, there exists a monomorphism f from $E(M)/E(M)_t$ into $E(M)$ such that $E(M) = \operatorname{Ker} n \oplus \operatorname{Im} f$. But $M \subseteq \operatorname{Ker} n$ and M is large in $E(M)$; hence $\operatorname{Im} f = 0$ and $E(M) = E(M)_t \in \mathscr{T}$. Thus $\mathscr{T}$ is stable. Now choose $M \in \mathscr{F}$; then $E(M) \in \mathscr{F}$ and so the module $E(M)/M$ is torsionfree, and hence projective. Thus $E(M) \cong M \oplus E(M)/M$. This proves that M is injective.

The final statement follows from Lemma 2.1.

THEOREM 2.3. *Let* $(\mathscr{T}, \mathscr{F})$ *be a hereditary torsion theory for* ${}_R\mathscr{M}$ *for which cyclic torsionfree modules have projective covers; the following are equivalent:*

(1) $\mathscr{F}$ *is closed under projective covers, and every torsion module is projective;*

(2) $\mathscr{F}$ *is closed under homomorphic images, and every torsion module is injective;*

(3) $\mathscr{F}$ *is closed under homomorphic images,* $R = R_t \dotplus K$ *(ring direct sum), and* R_t *is a semi-simple ring with minimum condition.*

In addition, whenever (1), (2), *and* (3) *are true, then* $\mathscr{T}$ *is a* TTF *class and* $(\mathscr{T}, \mathscr{F})$ *is centrally splitting.*

Proof. $(1) \to (3)$: Choose $N \in \mathscr{F}$, and let L be a homomorphic image of N. Since L_t is projective, there exists a monomorphism f from L_t to N. But $\operatorname{Hom}(L_t, N) = 0$; thus $L_t = 0$ and $L \in \mathscr{F}$. Thus $\mathscr{F}$ is closed under homomorphic images.

Since R/R_t is a cyclic module, it has a projective cover $\pi: P(R/R_t) \to R/R_t$, and $P(R/R_t) \in \mathscr{F}$ by hypothesis. If n denotes the natural epimorphism from R onto R/R_t, then there exists a homomorphism $f: P(R/R_t) \to R$ such that $R = \operatorname{Im} f + \operatorname{Ker} n = \operatorname{Im} f + R_t$. But $\operatorname{Im} f \in \mathscr{F}$, so that $R_t \cap \operatorname{Im} f = 0$ and $R = R_t \oplus \operatorname{Im} f$. Thus $R = R_t \dotplus K$ — and we also get the final statement of the theorem — by Lemma 2.1.

Finally, it is easy to see that R_t is a completely reducible ring since every torsion module is projective; this is equivalent to saying that R_t is a semi-simple ring with minimum condition.

$(3) \to (2)$: If $M \in \mathscr{T}$, then $KM = 0$ since $\mathscr{F}$ is closed under homomorphic images. Hence M is an injective R_t-module, and, by Baer's Lemma, it is easy to see that M is an injective R-module.

$(2) \to (1)$: Let $M \in \mathscr{F}$ have a projective cover $\pi: P(M) \to M$; then

$P(M)_t$ is injective and $P(M) = P(M)_t \oplus P'$. Further, $\mathrm{Hom}(P(M)_t, M) = 0$ and thus $P(M)_t \subseteq \mathrm{Ker}\,\pi$. Hence $P(M)_t$ is small in $P(M)$, and $P(M) = P' \in \mathscr{F}$. Thus $\mathscr{F}$ is closed under projective covers.

Since R_t is injective, we have $R = R_t \oplus K$. Thus, by Lemma 2.1, we have that $R = R_t \dotplus K$. Since $\mathscr{F}$ is closed under homomorphic images, one can easily see that $M \in \mathscr{T}$ if and only if $KM = 0$. But if every R_t-module is injective, then every R_t-module is projective. Thus every torsion R-module is projective.

3. Stable torsion theories. In [5] the following result is given; its proof depends strongly upon the dualities present in quasi-Frobenius rings.

THEOREM B. *Let R be a quasi-Frobenius ring and let $(\mathscr{T}, \mathscr{F})$ be a hereditary torsion theory for ${}_R\mathscr{M}$. The following are equivalent:*

(1) *$\mathscr{T}$ is stable;*

(2) *$(\mathscr{T}, \mathscr{F})$ is splitting;*

(3) *$(\mathscr{T}, \mathscr{F})$ is centrally splitting.*

It is easily seen that the implications $(3) \rightarrow (2) \rightarrow (1)$ are always true, regardless of the type of ring involved. We are motivated to examine the remaining implications in types of left artinian rings more general that the quasi-Frobenius ones, especially since Fuller [7] has shown that QF-3 rings possess dualities somewhat similar to those in quasi-Frobenius rings.

THEOREM 3.1. *Let R be a left artinian generalized uniserial ring, and let $(\mathscr{T}, \mathscr{F})$ be a hereditary torsion theory for ${}_R\mathscr{M}$. Then $\mathscr{T}$ is stable if and only if $(\mathscr{T}, \mathscr{F})$ is splitting.*

Proof. We need only consider the case where $\mathscr{T}$ is stable. Now every module M is a direct sum of indecomposable cyclic submodules, and each of these submodules is a homomorphic image of a left ideal Re where e is a primitive idempotent of R [10]. But each such Re has a lattice of submodules which is a finite chain, and thus every homomorphic image of an Re has a lattice of submodules which is a finite chain.

If L is an indecomposable cyclic submodule of M, then by the preceding its socle, denoted $\mathrm{soc}(L)$, is simple. Thus either $\mathrm{soc}(L) \in \mathscr{T}$ or $\mathrm{soc}(L) \in \mathscr{F}$. But $\mathrm{soc}(L)$ is large in L, so that L is contained in the injective envelope of $\mathrm{soc}(L)$. By hypothesis either $E(\mathrm{soc}(L)) \in \mathscr{T}$ or $E(\mathrm{soc}(L)) \in \mathscr{F}$; thus either $L \in \mathscr{T}$ or $L \in \mathscr{F}$. Hence M splits.

EXAMPLE 3.2. Splitting does not imply centrally splitting in left artinian generalized uniserial rings.

Let K be a field, and let R be the ring of two by two upper triangular matrices over K. Let

$$I = \left\{ \begin{vmatrix} a & b \\ 0 & 0 \end{vmatrix} \mid a, b \in K \right\};$$

then I is an idempotent, two-sided ideal of R. Thus by Jans [9],

$$\mathscr{T} = \{M \in {}_R\mathscr{M} \mid IM = 0\}$$

is a TTF class with associated torsion theories $(\mathscr{T}, \mathscr{F})$ and $(\mathscr{C}, \mathscr{T})$, where

$$\begin{aligned} \mathscr{F} &= \{N \in {}_R\mathscr{M} \mid \mathrm{Hom}(T, N) = 0 \quad \text{for all } T \in \mathscr{T}\} \text{ and} \\ \mathscr{C} &= \{L \in {}_R\mathscr{M} \mid \mathrm{Hom}(L, T) = 0 \quad \text{for all } T \in \mathscr{T}\} \\ &= \{L \in {}_R\mathscr{M} \mid IL = L\}. \end{aligned}$$

Clearly $(\mathscr{C}, \mathscr{T})$ does not split, since $R_c = I$ is not a direct summand of R. Hence $\mathscr{T}$ is not centrally splitting.

Note that $F(\mathscr{T}) = \{I, R\}$; thus for $M \in {}_R\mathscr{M}$, $M_t = \{x \in M \mid (0: x) \in F(\mathscr{T})\} = \{x \in M \mid I \subseteqq (0: x)\}$, where $(0: x) = \{r \in R \mid rx = 0\}$. Since I is the only large proper left ideal of R, we see that M_t is the singular submodule $Z(M)$ of M. Also $Z(R) = 0$, so that $\mathscr{T}$ is the Goldie — and $E(R)$ — torsion class (see [1] and [9] for an explanation of these). It is well-known that the Goldie torsion class is stable; thus $(\mathscr{T}, \mathscr{F})$ splits by Theorem 3.1.

As an aside, we note that the class $\mathscr{C}$ above is hereditary but is not stable. Also we remark that Teply [11, Propositions 4.5 and 4.7] gives several necessary and sufficient conditions for splitting to imply centrally splitting.

PROPOSITION 3.3. *Let R be a QF-2 ring, and let $(\mathscr{T}, \mathscr{F})$ be a hereditary torsion theory for ${}_R\mathscr{M}$. If $\mathscr{T}$ is stable, then every projective module splits.*

Proof. If e is a primitive idempotent in R, then $\mathrm{soc}(Re)$ is both a simple module and is large in Re. Hence Re is contained in the injective envelope of $\mathrm{soc}(Re)$, and thus either $Re \in \mathscr{T}$ or $Re \in \mathscr{F}$. But any projective module P over a left artinian ring R is isomorphic to a direct sum of modules Re_α, where each e_α is a primitive idempotent of R. Thus every projective module splits.

If $\mathscr{T}$ is a stable hereditary torsion class for a QF-2 ring, then, by Theorem A and Proposition 3.3, every quasi-injective and every projective module splits. It seems reasonable to conjecture that every module will split, and in fact we have been unable to find examples to the contrary.

REFERENCES

1. J. S. Alin and E. P. Armendariz, *TTF-classes and $E(R)$-torsion modules*, preprint.
2. J. S. Alin and S. E. Dickson, *Goldie's torsion theory and its derived functor*, Pacific J. Math., **24** (1968), 195-203.
3. E. P. Armendariz, *Quasi-injective modules and stable torsion classes*, Pacific J. Math., **31** (1969), 277-280.
4. H. Bass, *Finitistic dimension and a homological generalization of semi-primary rings*, Trans. Amer. Math. Soc., **95** (1960), 466-488.
5. R. L. Bernhardt, *Splitting hereditary torsion theories over semiperfect rings*, Proc. Amer. Math. Soc., **22** (1969), 681-687.
6. S. E. Dickson, *A torsion theory for abelian categories*, Trans. Amer. Math. Soc., **121** (1966), 223-235.
7. K. R. Fuller, *The structure of QF-3 rings*, Trans. Amer. Math. Soc., **134** (1968), 343-354.
8. P. Gabriel, *Des categories abeliennes*, Bull. Soc. Math. France, **90** (1962), 323-449.
9. J. P. Jans, *Some aspects of torsion*, Pacific J. Math., **15** (1965), 1249-1259.
10. T. Nakayama, *On Frobeniusean algebras* II, Ann. of Math., **42** (1941), 1-22.
11. M. L. Teply, *Homological dimension and splitting torsion theories*, Pacific J. Math., **34** (1970), 193-206.
12. L. E. T. Wu and J. P. Jans, *On quasi-projectives*, Ill. J. Math., **11** (1967), 439-448.

Received October 13, 1970 and in revised form April 29, 1971.

THE UNIVERSITY OF NORTH CAROLINA AT GREENSBORO

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

GERŠGORIN THEOREMS, REGULARITY THEOREMS, AND BOUNDS FOR DETERMINANTS OF PARTITIONED MATRICES II SOME DETERMINANTAL IDENTITIES

J. L. BRENNER

A square matrix $A = [a_{ij}]_1^n$ has dominant diagonal if $\forall_i\{|a_{ii}| > R_i = \sum_{j \neq i} |a_{ij}|\}$. A more complicated type of dominance is the following. Suppose for each i, there is assigned a set $I(i)$ (subset of $\{1, \cdots, n\}$), $i \in I(i)$: Define B_{ij} as the $I(i) \times I(i)$ submatrix of A that uses columns $I(i)$, and rows $\{I(i)\backslash i, j\}$, i.e., the set obtained from $I(i)$ by replacing the ith row by the jth row. Set $b_{ij} = \det B_{ij}$. Then $[b_{ij}]_1^n$ is a matrix, the elements of which are determinants of minor matrices of A. In an earlier paper, bounds for det A were derived in case $[b_{ij}]$ has dominant diagonal in the special case that $\{I(i)\}_i$ represents a partitioning of the indices into disjoint subsets.

In this article the general case is treated; $I(i)$ can be any subset of $\{1, \cdots, n\}$ that contains i. An identity is derived connecting $\det [b_{ij}]_1^n$ with $\det A$.

To establish the identity, a general multinomial identity is first derived, connecting determinants of certain submatrices of an $r \times 2r$ matrix of indeterminates. This result, reminiscent of Sylvester's determinantal identity, is used to bound $\det A$.

1. Application of a characterization of the determinant function.

LEMMA 1.01. *Let $A = [a_{ij}]_1^n$ be a matrix of complex numbers* [or indeterminates]; *let a function ϕ: $A \to C$*[or ϕ: $A \to C[a_{11}, \cdots, a_{nn}]$] *have the following properties for all $n \times n$ matrices A.*

(1.02) [1.03] *If any row* [*column*] *of A is replaced by the sum of that row* [*column*] *and a multiple of another row* [*column*], *$\phi(A)$ is unaltered.*

(1.04) *If any row of A is multiplied* (*throughout*) *by a constant α, $\phi(A)$ is multiplied α^r.*

Then $\phi(A)$ is a constant c_0 (independent of a_{ij}) multiplied by the rth power of det A.

Proof. The hypotheses (1.02, 1.03) guarantee that $\phi(A)$ is the same as $\phi(B)$, where B is any matrix obtainable from A by means of elementary transformations. It is known that $B = \operatorname{diag}[\det A, 1, \cdots, 1]$ can be so obtained; see for example [1]. Thus $\phi(A)$ is some function of $\det A$; the conclusion of lemma 1.01 follows on applying hypothesis 1.04 to the matrix B: If $\phi(\alpha x) = \alpha^r \phi(x)$, then $\phi(x) \equiv c_0 x^r$, since $\phi(x)/x^r$ is constant.

An application of this result was made in [2], to which the reader should refer. In slightly changed notation, this application is as follows.

LEMMA 1.05. *Let $A = [a_{ij}]_{i=1,}^{r}{}_{j=1}^{2r}$ be an $r \times 2r$ matrix of indeterminates, let $b_{ij} = \det A\begin{pmatrix} 1 \cdots r \\ 1 \cdots i-1, i+1, \cdots, r, j \end{pmatrix}$ be the determinant of the $r \times r$ submatrix of A that uses columns $\{1, \cdots, r\} \backslash i, j$. This is the almost-principal submatrix of A in which the ith column is replaced by the jth column.* (For $j = i$, this is $A\begin{pmatrix} 1 \cdots r \\ 1 \cdots r \end{pmatrix}$. For $1 \leqq j \neq i \leqq r$, this submatrix has determinant 0.)

Then

$$X = \det [b_{ij}]_{i=1, j=r+1}^{r \quad 2r} = G_1^{r-1} \det [a_{ij}]_{i=1, j=r+1}^{r \quad 2r} , \tag{1.06}$$

where

$$G_1 = \det [a_{ij}]_{11}^{rr} .$$

Note that in 1.06, the column indices are $r + 1, \cdots, 2r$.

To prove this Lemma, it is only necessary to observe that it is a multinomial identity, and that the hypotheses of Lemma 1.01 concerning the function X are satisfied.

1° if X is regarded as a function of $\{a_{ij}, 1 \leqq i, j \leqq r\}$;

2° if X is regarded as a function of $\{a_{ij}, 1 \leqq i \leqq r, r < j \leqq 2r\}$.

COROLLARY 1.07. *With the same hypothesis, the conclusion*

$$Y = \det [b_{ij}]_{i=1, j \in S}^{r} = G_1^{r-1} \det [a_{ij}]_{i=1, j \in S}^{r} \tag{1.08}$$

is valid, where S is any set of r distinct positive integers not exceeding $2r$.

Proof. Since 1.06 is a multinomial identity, the r^2 indeterminates $a_{ij}\,(j > r)$ on the right can simply be replaced by the r^2 indeteminates $a_{ij}\,(j \in S)$. But this replacement changes not only the range of j in the set variables $\{a_{ij}\}$, but also the range of j in the set $\{b_{ij}\}$, as the definition of b_{ij} shows.

LEMMA 1.09. *Suppose*

$$I(1) = \{1\}, I(2) = \{1, 2\}, \cdots, I(r) = \{1, 2, \cdots, r\} .$$

Let $B = \{b_{ij}\}_{i=1,j=1}^{r,\ 2r}$ *be defined as in* 1.05. *Then*

$$\det B\binom{1, \cdots, r}{r+1, \cdots, 2r}$$

$$(1.10) \qquad = a_{11} \det A\binom{12}{12} \det A\binom{123}{123} \cdots \det A\binom{12 \cdots r-1}{12 \cdots r-1}$$

$$\times \det A\binom{1, 2, \cdots, r}{r+1, \cdots, 2r} .$$

REMARK. This is again a multinomial identity in the $2r^2$ indeterminates a_{ij}. Therefore 1.09 has the Corollary

$$(1.11) \quad \det B\binom{1 \cdots r}{j_1 \cdots j_r} = a_{11} \det A\binom{12}{12} \det A\binom{123}{123} \cdots \det A\binom{1\,2 \cdots r}{j_1 \cdots j_r}$$

in view of the definition of b_{ij}.

Proof of Lemma 1.09. To show that a_{11} is a factor in (1.10), as shown, a_{21} times the first row is added to the second row. The second row becomes

$$(1.13) \qquad a_{11}a_{2,r+1},\ a_{11}a_{2,r+2},\ \cdots,\ a_{11}a_{2,r+j},\ \cdots$$

which obviously has a_{11} as a factor.

It is a little more complicated to show $\det\begin{pmatrix}a_{11}a_{12}\\a_{21}a_{22}\end{pmatrix}$ is also a factor, as is asserted in relation (1.10). The trick is to add to the third row $-\det\begin{pmatrix}a_{21}a_{22}\\a_{31}a_{32}\end{pmatrix}$ times the first row as well as $a_{11}^{-1}\det\begin{pmatrix}a_{11}a_{12}\\a_{31}a_{32}\end{pmatrix}$ times the second row (1.13). The new third row is

$$(1.14) \qquad \det\begin{pmatrix}a_{11}a_{12}\\a_{21}a_{22}\end{pmatrix}[a_{3,r+1},\ a_{3,r+2},\ \cdots,\ a_{3,r+j},\ \cdots] ,$$

i.e., every element of that row has the common prefactor indicated.

The formal proof of (1.10) is inductive, as follows. As an induction hypothesis, assume that the left member of (1.10) can be written in the form

$$(1.15) \qquad a_{11} \det A\binom{12}{12} \cdots \det A\binom{12 \cdots k-1}{12 \cdots k-1} \det C_k ,$$

where C_k is the $r \times r$ matrix, the jth column of which is

$$\begin{array}{c} a_{1,r+j} \\ a_{2,r+j} \\ \vdots \\ a_{k,r+j} \end{array}$$

$$\det\begin{bmatrix} a_{11} \cdots a_{1k} a_{1,r+j} \\ a_{k1} \cdots a_{kk} a_{k,r+j} \\ a_{k+1,1} \cdots a_{k+1,r+j} \end{bmatrix}$$
$$\vdots \qquad .$$

This has already been established for $k = 1, 2$. The inductive assertion is: the factor $\det A\binom{12 \cdots k}{12 \cdots k}$ splits off from $\det C_k$. To prove this, subtract from the $k+1$st row of the matrix C_k appropriate multiples of the preceding rows. The multiple of $a_{i,r+j}$ needed is precisely the cofactor of $a_{i,r+j}$ in C_k itself.

This completes inductive proof. To establish (1.10) in its entirety, a final visual check is needed of the circumstance that for $k = r$, the matrix C_r is indeded the matrix $A\binom{1 \cdots r}{r+1 \cdots 2r}$. See (1.05).

2. Some special factorizations.

THEOREM 2.01. *Let $A = [a_{ij}]$ be a matrix with r rows: $i = 1(1)r$, and $2r$ columns: $j = 1(1)r < j_1 < \cdots < j_r$. Suppose, for $i = 1, 2, \cdots, r-1$, $I(i) = \{1, 2, \cdots, r-1\}$; $I(r) = \{1, 2, \cdots, r\}$. For $j = j_1, j_2, \cdots, j_r$ set $B_{ij} = A\binom{I(i)}{I(i)\backslash i, j}$, $i = 1(1)r$.*

Denote $\det B_{ij}$ *by* b_{ij}; $B = [b_{ij}]$. *Then*

$$\det B = \pm\, C^{r-1} \det A\binom{1, 2, \cdots, r}{j_1 j_2 \cdots j_r};\ C = \det A\binom{I(1)}{I(1)} . \tag{2.02}$$

Proof. Consider the last row of B. The element b_{rj} in column "j" of this row is the determinant of the $r \times r$ matrix B_{rj}. If this determinant is expanded by minors of the elements $a_{rj}, a_{r1}, a_{r2}, \cdots a_{r,r-1}$ of the last row of B_{rj}, the result is

$$b_{rj} = \pm a_{rj} C \pm a_{r1} b_{1j} \pm a_{r2} b_{2j} \pm \cdots \pm a_{r,r-1} b_{r-1,j} . \tag{2.03}$$

Relation (2.03) shows that $\det B$ is not altered if every element b_{rj} of the last row of B is replaced by $\pm a_{rj} C$. (This replacement would merely omit from the last row of B a linear combination of the preceding rows.)

At this point it is clear that C is a factor of $\det B$, and that the other factor has the same first $r-1$ rows does B, and has last row a_{rj}. The conclusion of the theorem now follows by expanding $\det B$ by its last row and applying Corollary 1.07. See Lemmas 4.3, 4.4 of [2].

COROLLARY 2.04. *Suppose*

$$I(i) = \{1, 2, \cdots, r - k\} \textit{ for } i = 1, 2, \cdots, r - k\ ;$$

and $I(i) = \{1, 2, \cdots, r - k, i\}$ *for* $i = r - k + 1, \cdots, r$. *Then* (2.02) *holds; where* C *now means* $\det A\binom{1, 2, \cdots, r - k}{1, 2, \cdots, r - k}$.

2.05. Another special case is the case $I(1) = \{1, 2\}, I(2) = \{2, 3\}, I(3) = \{3, 1\}$. The formula

$$\det B = G \det A, \qquad G = \det \begin{bmatrix} a_{11} & -a_{12} & 0 \\ 0 & a_{22} & -a_{23} \\ -a_{31} & 0 & a_{23} \end{bmatrix} \tag{2.06}$$

can be verified by appropriate devices. A generalization of (2.06) is the formula

$$\det B\binom{1\ 2\ 3}{j_1 j_2 j_2} = G \cdot \det A\binom{1\ 2\ 3}{j_1 j_2 j_3}\ , \tag{2.07}$$

valid for any 3×6 matrix A, with $I(i)$ defined as above. Among several valid proofs of this formula, the following is presented. It proves (2.07) as a special case of a still more general result.

THEOREM 2.08. *Let* $A = [a_{ij}]$ *be an* $r \times 2r$ *matrix,* $i = 1(1)r$, $j = 1(1)2r$. *Let* B *be the* $r \times r$ *matrix with* (i, j) *element* $b_{ij} = \det B_{ij}$, *where* $B_{ij} = A\binom{i\ \ i + 1}{j\ \ i + 1}$, $i = 1(1)r - 1$, $B_{rj} = A\binom{r1}{j1}$; $j = r + 1(1)2r$. *Then the relation*

$$\det B = G \det A\binom{1 \cdots r}{r + 1 \cdots 2r}, \quad G = \det \begin{bmatrix} a_{11} & -a_{12} & & \\ & \ddots & \ddots & \\ & & \ddots & -a_{r-1,r} \\ -a_{r,1} & & & a_{rr} \end{bmatrix} \tag{2.09}$$

holds; G *is a bidiagonal matrix with* $2r$ *nonzero elements.*

REMARK. This is the case $I(1) = \{1, 2\}, I(2) = \{2, 3\}, \cdots, I(r) = \{r, 1\}$.

Proof. Subtract a multiple of the first row of B from the second, then a multiple of the second from the third, $\cdots$, a multiple of the

$r-1st$ from the last. The resulting matrix has the same determinant as B, and the multiples mentioned can be chosen so that this resulting matrix is, row by row,

$$\begin{array}{rr} a_{22}[a_{1j}] - a_{12}[a_{2j}] & \cdots 1 \\ (a_{22}a_{33}/a_{12})[a_{1j}] - a_{23}[a_{3j}] & \cdots 2 \\ (a_{22}a_{33}a_{44}/a_{12}a_{23})[a_{1j}] - a_{34}[a_{4j}] & \cdots 3 \\ \vdots & \vdots \\ (a_{11}a_{22}\cdots a_{rr}/a_{12}\cdots a_{r-1r} - a_{r1})[a_{1j}] & \cdots r\,. \end{array}$$

Now subtract a multiple of the new last row from each of the preceding rows; the first $r-1$ rows of the new matrix are $-a_{12}[a_{2j}]$, $-a_{23}[a_{3j}]$, $\cdots$. This matrix obviously has determinant (2.09). ||

3. General factorization of det B. The function $i \mapsto I(i)$ induces a (weak) separation of the indices $\{1, \cdots, n\}$ into agglomerated mutually exclusive sets $S(k)$, as follows.

DEFINITION 3.01. Let $i \mapsto I(i)$ be a function from the integers $\{1, \cdots, n\}$ to sets of these same integers, with the further property $i \in I(i)$ for all i. In the usual way, the sets $I(i)$ are now agglomerated into the smallest possible (minimal) *mutually exclusive* sets $S(k)$ so that: *Every $I(i)$ is in one or another of the sets $S(k)$.* Then $S(k)$ are the *mutually separated sets defined by the function I.* For example, the function

$$1 \longmapsto \{1\},\ 2 \longmapsto \{1, 2\},\ 3 \longmapsto \{1, 2, 3\},\ 4 \longmapsto \{4, 5\},\ 5 \longmapsto \{5, 6\},$$
$$6 \longmapsto \{6, 7\},\ 7 \longmapsto \{7\}$$

defines a separation of the indices $\{1, 2, 3, 4, 5, 6, 7\}$ into the mutually exclusive sets $S(1) = \{1, 2, 3\}$, $S(2) = \{4, 5, 6, 7\}$.

Parallel to the separation of Definition 3.01, there is a factorization of det B into a product of factors, one for each set $S(k)$. The kth factor is the determinant of a matrix; in general the elements of this matrix are again determinants of matrices: the elements of these matrices are elements a_{ij} of the matrix A, where $i, j \in S(k)$. The point is that the polynomial function det B of the elements of A factors into the product of multinomial factors; the kth factor is a polynomial in the indeterminates a_{ij}, where i, j belong only to the kth set $S(k)$ of indices. Besides these factors, det A also appears as a factor.

It there are two or more sets $S(k)$ in the separation, then det A, but not $(\det A)^2$, is thus a factor of det B. Even when the entire set $\{1, 2, \cdots, n\}$ of indices are connected through the sets I (there is but a single set S), the factor det A appears only to first power "in general." The exact meaning of "in general" is explained below.

The above remarks are summarized in the following theorem. Its proof, together with a more detailed atatement, unfold in § 4.

THEOREM 3.02. *Let $A = [a_{ij}]$ be an $n \times n$ matrix of indeterminates; for $i = 1(1)n$ let $I(i)$ be a subset of the first n integers with $i \in I(i)$. Denote by B_{ij} the minor $A\binom{I(i)}{I(i)\backslash i, j}$ on rows $I(i)$; and on columns $I(i)$, but with index i replaced by j. Set $b_{ij} = \det B_{ij}$; $B = [b_{ij}]$. Thus B is an $n \times n$ matrix. Let the function $I(i)$ induce a separation of the indices $\{1, \cdots, n\}$ into $s \geqq 1$ mutually exclusive sets $S_1, S_2, \cdots, S_s$. Then $\det B$, which is obviously a polynomial function of the n^2 indeterminates a_{ij} with integer coefficients, can be factored in the form*

$$\det B = G \det A ,$$

where $G = M_1 M_2 \cdots M_s$, and where each M_k is a multinomial in those indeterminates a_{ij} for which both indices i, j belong to the set S_k. In particular, $\det A$ is always a factor of $\det B$.

The details of the proof depend on the following lemma.

LEMMA 3.03. *Let $A = [a_{ij}]$ be an $r \times 2r$ matrix of indeterminates, $i = 1(1)r$, $j = 1(1)2r$. For each i, let $I(i)$ be a subset of the first r integers. Let B_{ij}, b_{ij} be defined formally as in Theorem 3.02. B_1 is the $r \times r$ matrix $[b_{ij}]$, $1 \leqq i \leqq r < j \leqq 2r$. A_1 is the $r \times r$ matrix $[a_{ij}]_{1 \leqq i \leqq r < j \leqq 2r}$.* (Note the range for j.)

Then the polynomial identity

$$\det B_1 = F \cdot \det A_1 \tag{3.04}$$

holds, where F is a multinomial with integer coefficients in the indeterminates $\{a_{ij}, 1 \leqq i, j \leqq r\}$.

REMARK 3.05. This lemma is more general than any of previous ones, since the sets $I(i)$ are more general.

COROLLARY 3.06. *Det A_1 is, but $(\det A_1)^2$ is not a factor of det B_1.*

Proof. The variables that figure in F are disjoint from those in B_1.

REMARK 3.07. This is the meaning of the phrase "in general" above.

COROLLARY 3.08. *Let A_1, B_1 redefined conformally. That is,* *without changing the sets $I(i)$, let the range for j in the definitions of A_1, B_1 be replaced by any range of r distinct integers, including some or all of the first r integers. Then* (3.04) *still holds.*

Proof. If some of the indices j in the polynomial $\det A_1$ are changed, the definition of b_{ij} shows that a conformal change is concurrently made in the polynomial $\det B_1$. In other words, the change amounts solely to a change of the names of the variables in (3.04). But (3.04) is a polynomial identity.

Under the change $a_{i,j} \to a_{i,j-r}$, $b_{i,j} \to b_{i,j-r}$ in (3.04), the factor $\det A_1$ could appear as a factor in F for suitable choice of $I(i)$. For example, if $I(i) \equiv \{1, 2, \cdots, r\}$, and if j runs through the range $1 \leqq j \leqq r$, then (3.04) becomes $\det B_1 = (\det A_1)^r$.

Proof of Lemma 3.03. To avoid difficulties with an algebraic sign, the columns of $B_{ij} \equiv A\binom{I(i)}{(I(i)\backslash i, j)}$ are to be thought of as written in a definite order: the jth column a_{ij} first, followed by the other columns in natural order. For example, if $I(1) = \{1, 2, 3\}$ then B_{1j} is the matrix

$$\begin{bmatrix} a_{1j} & a_{12} & a_{13} \\ a_{2j} & a_{22} & a_{23} \\ a_{3j} & a_{22} & a_{33} \end{bmatrix} .$$

Without this convention, the formula to be obtained for F would be determined only up to sign.

It will be instructive to carry through the proof in a special case, since a rather simple special case already embodies all the points of difficulty and interest. The case $I(1) = \{1, 2\}$, $I(2) = \{1, 2, 3\}$, $I(3) = \{1, 2, 3\}$ will serve as an illustration. The matrix B_1 has as jth column B_{1j}, where

$$B_{1j} = \begin{bmatrix} \det \begin{bmatrix} a_{1j} & a_{12} \\ a_{2j} & a_{22} \end{bmatrix} \\ \det \begin{bmatrix} a_{1j} & a_{11} & a_{13} \\ a_{2j} & a_{21} & a_{23} \\ a_{3j} & a_{31} & a_{33} \end{bmatrix} \\ \det \begin{bmatrix} a_{1j} & a_{11} & a_{12} \\ a_{2j} & a_{21} & a_{22} \\ a_{3j} & a_{31} & a_{32} \end{bmatrix} \end{bmatrix} \quad j = 4, 5, 6 . \tag{3.09}$$

The first step in the proof is to border the 3×3 matrix B_1 with 3 rows and columns as shown below. The enlarged matrix B_2 clearly has the same determinant as B_1, except for the factor $(-1)^r$. Only the subscripts are printed; thus $1j$ is an abbreviation for a_{1j}. The reader must also supply the symbol det throughout: [ ] is an abbreviation for det [ ].

$$\left[\begin{array}{cccccc}
14, & 15, & 16, & 1, & 0, & 0 \\
24, & 25, & 26, & 0, & 1, & 0 \\
34, & 35, & 36, & 0, & 0, & 1 \\
\begin{bmatrix} 14 & 12 \\ 24 & 22 \end{bmatrix}, & \begin{bmatrix} 15 & 12 \\ 25 & 22 \end{bmatrix}, & \begin{bmatrix} 16 & 12 \\ 26 & 22 \end{bmatrix}, & 0, & 0, & 0 \\
\begin{bmatrix} 14 & 11 & 12 \\ 24 & 21 & 23 \\ 34 & 31 & 33 \end{bmatrix}, & \begin{bmatrix} 15 & 11 & 13 \\ 25 & 21 & 23 \\ 35 & 31 & 33 \end{bmatrix}, & \begin{bmatrix} 16 & 11 & 13 \\ 26 & 21 & 23 \\ 36 & 31 & 33 \end{bmatrix}, & 0, & 0, & 0 \\
\begin{bmatrix} 14 & 11 & 12 \\ 24 & 21 & 22 \\ 34 & 31 & 32 \end{bmatrix}, & \begin{bmatrix} 15 & 11 & 12 \\ 25 & 21 & 22 \\ 35 & 31 & 32 \end{bmatrix}, & \begin{bmatrix} 16 & 11 & 12 \\ 26 & 21 & 22 \\ 36 & 31 & 32 \end{bmatrix}, & 0, & 0, & 0
\end{array}\right].$$

To show that the factor $\det A_1$ splits off from the determinant of this 6×6 matrix, it need only be noted that the matrix can be reduced to the form $\begin{bmatrix} A_1 & I \\ 0 & F_1 \end{bmatrix}$ by adding appropriate linear combinations of the first three rows to each of the last three. This argument is an alternative to a general argument of Loewy [3], who proved by another method that if $\det A_1 = 0$, then necessarily $\det B_1 = 0$. In the special case being expounded, $\det B_2 = -(\det F_1)(\det A_1)$, where F_1 is the 3×3 matrix

$$\left[\begin{array}{ccc}
a_{22}, & -a_{12}, & 0 \\
\begin{bmatrix} 21 & 23 \\ 31 & 33 \end{bmatrix}, & -\begin{bmatrix} 11 & 13 \\ 31 & 33 \end{bmatrix}, & \begin{bmatrix} 11 & 13 \\ 21 & 23 \end{bmatrix} \\
\begin{bmatrix} 21 & 22 \\ 31 & 32 \end{bmatrix}, & -\begin{bmatrix} 11 & 12 \\ 31 & 32 \end{bmatrix}, & \begin{bmatrix} 11 & 12 \\ 21 & 22 \end{bmatrix}
\end{array}\right].$$

The argument given above has general applicability. Formula (3.04) is established. The multinomial F is in fact the determinant of an $r \times r$ matrix. The (k, l) element of this matrix is the negative of the cofactor of $a_{l\,r+l}$ in $b_{k,r+l} = \det A\begin{pmatrix} I(k) \\ I(k)\backslash k,\ r+l \end{pmatrix}$, and is thus

$$f_{kl} = -(-1)^{1+\operatorname{pos} l} \det A\begin{pmatrix} I(k)\backslash l \\ I(k)\backslash k \end{pmatrix},$$

where $\operatorname{pos} l$ is the position of l in the set $I(k)$. If $l \notin I(k)$, then $f_{kl} = 0$, and conversely. For consistency, f_{kk} must be defined as 1 when $I(k) = \{k\}$.

COROLLARIES.

(3.09) $$\det B_1 = (-1)^r(\det F_1)(\det A_1)$$

3.10 [3] If $\det A_1 = 0$, then $\det B_1 = 0$.

3.11. If F_1 is a triangular matrix, then

(3.12) $$\det B_1 = -(-1)^r \Pi(\det G_1^{(i)})\cdot(\det A_1), \text{ where}$$

(3.13) $$G_1^{(i)} = A\binom{I(i)\backslash i}{I(i)\backslash i}\,.$$

In particular, relation (1.10) follows; this proof differs from the first proof.

(3.14) In case $I(1)=\{1,2\}$, $I(2)=\{2,3\}$, $\cdots$, $I(i)=\{i, i+1\}$, $\cdots$, $I(n)=\{n, 1\}$, then formula

(3.15) $\det B_1 = G\cdot\det A_1$ holds, where $G = \det\begin{bmatrix} a_{11} & -a_{12} & \\ & a_{22} & -a_{23} \\ & \cdot\ \cdot\ \cdot & \end{bmatrix}$ is the determinant of the bidiagonal matrix shown. This proof is again different from the earlier proof of (2.09).

3.16. Note that the case $I(1) = \{1, 2, 3\}$, $I(2) = \{2, 3, 4\}$, $\cdots$ is considerably more complicated than the case (3.14); indeed while the first type of proof is more direct for the hypothesis (3.14), an attempt to generalize this proof to the case (3.16) is unrewarding.

3.17. *Relation* (1.06) *holds.*

The following proof of 1.06 is somewhat less direct than the original proof. The matrix F_1 is not triangular, so that the determinant $\det F_1$ does not factor for this simple reason. However F_1 is seen on inspection to be the $r-1st$ compound of the matrix $A\binom{I(1)}{I(1)}$; thus $\det F_1 = \det A\binom{I(1}{I(1}^{r-1}$. This proof requires a knowledge of the formula

(3.18) $\det C^{(t)} = (\det C)^e$, $e = \binom{r-1}{t-1}$, where $C^{(t)}$ is the tth compound of the $r \times r$ matrix C.

4. General factorization of $\det B$ (continued). In this section, Corollary 3.08 is applied to obtain a general formula for the determinant of the $n \times n$ matrix $B = [b_{ij}]$ defined in Theorem 3.02.

Since Theorem 3.02 holds for a matrix A of indeterminates, it holds in particular for a matrix A of complex numbers.

Proof of Theorem 3.02. The function $i \mapsto I(i)$ induces a separation of the indices $\{1, 2, \cdots, n\}$ into $s \geqq 1$ *mutually exclusive sets* $S(k)$ such that *every set* $I(i)$ *is in exactly one of the sets* $S(k)$, and the sets $S(k)$ cannot be further decomposed without destroying these properties.

In following the details of the proof, the reader may prefer to think of the indices of the sets $S(1), S(2), \cdots$ as occuring in natural order.

To continue the proof, the rows of B are partitioned into (mutually exclusive) sets $S(1), S(2), \cdots$ and $\det B$ is expanded according to the generalized Laplace expansion on these rows. Corollary 3.08 asserts that the determinants of all the $S(1) \times S(1)$ minor matrices on the set of rows with indices in $S(1)$ have a common factor M_1. The corollary asserts further that this common factor is a multinomial in the particular variables a_{ij} $(i, j \in S(1))$. Similarly for $S(2), \cdots$. Thus $M_1 M_2 \cdots M_s$ is a factor of $\det B$.

Besides the factor common to the determinants of all the $S(1) \times S(1)$ matrices, there is a factor, see (3.04), peculiar to the particular minor matrix. This peculiar factor is just what is needed, in the Laplace expansion of $\det B$, to produce $\det A$. The proof of Theorem 3.02 is complete.

Let A be a matrix of indeterminates. If there is more than one set $S(k)$, then $\det A$ is, but $(\det A)^2$ is not, a factor of $\det B$.

5. **Applications.** Theorem 3.02 can be used to obtain bounds for $\det A$ in case the matrix B has dominant diagonal. The details and results are similar to those of [2]. These results have one remarkable feature: This is the first occasion on which such bounds have been obtained for a "partitioning" of a matrix, in which the sets of rows in the "partitioning" overlap one another.

The results of this paper will be needed in any attempt to obtain minimal Geršgorin sets related to the Hoffman-Brenner theorem. If it can be accomplished, this will be an interesting generalization of the results of [5].

REFERENCES

1. J. L. Brenner, *Applications of the Dieudonné Determinant,* Linear Algebra and its Applications **1** (1968), 511-536.

2. ———, *Geršgorin theorems, regularity theorems, and bounds for determinants of partitioned matrices,* SIAM J. Appl. Math., **18** (1970), 443-450.

3. F. R. Gantmacher, *Applications of the theory of matrices,* N. Y.: Interscience (Wiley) 1959.

4. R. Loewy, *Gerschgorin-type theorems for partitioned matrices*, SIAM J. Num. Anal., **7** (1970), 125-128.
5. R. S. Varga, *Minimal Gerschgorin sets*, Pacific J. Math., **15** (1965), 719-729.

Received September 10, 1970, and in revised form April 13, 1971. Supported by NSF-GP9483.

UNIVERSITY OF ARIZONA
AND
UNIVERSITY OF VICTORIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

ON REPRESENTING F^*-ALGEBRAS

R. M. BROOKS

The purpose of this paper is to obtain a concrete representation for F^*-algebras with identity: a Frechet algebra with involution for which there exists a determining sequence of B^*-seminorms. The main result is Theorem 3.4 which is described here. Let A be an F^*-algebra with identity. Let $\{(\pi_\lambda, H_\lambda): \lambda \in \Lambda\}$ be a complete family of irreducible Hilbert space representations of A. Let $H = \Sigma_\lambda \oplus H_\lambda$, define $E \subseteqq \Lambda$ to be equicontinuous provided $\sup_{\lambda \in E} ||\pi_\lambda(a)|| < \infty$ $(a \in A)$, and let $X = \{x \in H: \mathrm{Supp}\,(x)$ is equicontinuous$\}$. The linear space X is given the final topology τ_f determined by the family $\{H_E = [x \in H: \mathrm{Supp}\,(x) \subseteqq E]: E$ equicontinuous$\}$ of subspaces of X. Let X_f be (X, τ_f) and let $\mathscr{L}^*(X_f)$ be all operators on X which have an adjoint relative to the inner product inherited from H such that both the operator and its adjoint are τ_f-continuous. This algebra will be endowed with the topology $\mathscr{T}_b$ of bounded convergence. Let $\mathscr{L}^+(X)$ be all operators which have adjoints. It has a natural topology $\mathscr{T}_+$ described in § 2. Define $\pi: A \to \mathscr{L}_a(X)$ by $\pi(a)\{x_\lambda\} = \{\pi_\lambda(a)x_\lambda\}$ for $a \in A$ and $x = \{x_\lambda\} \in X$. Then $\pi(A) \subseteqq \mathscr{L}^+(X) = \mathscr{L}^*(X_f)$, and (1) $\pi: A \to (\mathscr{L}^*(X_f), \mathscr{T}_b)$ is a topological *-isomorphism (into) and (2) $\pi: A \to (\mathscr{L}^+(X), \mathscr{T}_+)$ is a topological *-isomorphism (into).

In § 1 we recall some results about Fréchet *-algebras with identity, their positive functionals and Hilbert space representations, and set the notation for the remainder of the paper.

In § 2 we obtain the results about algebras of operators on certain inner product spaces necessary to prove the main representation theorem.

In § 4 we define the concept of an enveloping algebra $E(A)$ for a Fréchet *-algebra with identity, A, and show that $E(A)$ can be realized either as the inverse limit of the enveloping algebras of the Banach *-algebras in an inverse limit decomposition of A or as an algebra of operators naturally constructed from the irreducible Hilbert space representations of A. Also we show that $E(A)$ has the property that every Hilbert space representation of A factors through $E(A)$, but that there are representations of A in algebras $\mathscr{L}^+(X)$ which fail to factor through $E(A)$.

1. Preliminaries. A *Fréchet algebra* is a complete metrizable topological algebra whose topology is determined by a (countable) family of seminorms (submultiplicative, convex, symmetric functionals). We may assume that such a family $\{\|\cdot\|_n\}_{n=1}^\infty$ for A is *ascending*: $\|a\|_n \leqq \|a\|_{n+1} (a \in A, n \in N)$, and that $\|e\|_n = 1 (n \in N)$ if A has an identity e. A *Fréchet *-algebra* is a Fréchet algebra with a continuous involution. If A is a Fréchet *-alebra with identity we can choose a sequence $\{\|\cdot\|_n\}_{n=1}^\infty$ of seminorms for A such that (1) $\{\|\cdot\|_n\}$ determines the topology of A, (2) $\{\|\cdot\|_n\}$ is ascending, (3) $\|e\|_n = 1 (n \in N)$, and (4) $\|a^*\|_n = \|a\|_n (a \in A, n \in N)$. Such a sequence we shall call a **-sequence of seminorms for A*. An *F^*-algebra* is a Fréchet *-algebra, for which there is an ascending determining sequence $\{\|\cdot\|_n\}$ of seminorms for A each of which has the B^*-property: $\|a^*a\|_n = \|a\|_n^2 (a \in A, n \in N)$. Such a sequence we shall call an *F^*-sequence of seminorms*. The usual constructions (see [5]) show that every Fréchet *-algebra (resp., F^*algebra) is an inverse limit of Banach *-algebra (resp., B^*-algebras).

Let $(A, \{\|\cdot\|_n\})$ be a Fréchet *-algebra with identity e. We denote by $P(A)$ the set of all positive functionals on A and by $K(A)$ those $f \in P(A)$ for which $f(e) = 1$. For each $n \in N$ we let $P_n(A)$ (resp., $K_n(A)$) be the set of all $f \in P(A)$ (resp., $K(A)$) such that $|f(a)| \leqq f(e)\|a\|_n (a \in A)$. If $\{A_n, \rho^n, N\}$ is the inverse system generated by $\{\|\cdot\|_n\}$ with $\rho_n: A \rightarrow A_n$ natural map of A onto the nth member A_n, then for each n the dense homomorphism ρ_n induces a one-to-one map ρ_n^* of $P(A_n)$ onto $P_n(A)$. $(K(A_n))$ onto $K_n(A))$. Moreover, ρ_n^* preserves indecomposability. A theorem of Do-Shing [2] states every positive functional on A is continuous so we have $P(A) = \bigcup_{n=1}^\infty P_n(A)$ and $K(A) = \bigcup_{n=1}^\infty K_n(A)$. Also, $K(A)$ is a weak*-closed, convex subset of A^* and is the closed convex hull of its extreme points $\operatorname{ext}(K(A))$ which is exactly $\bigcup_{n=1}^\infty \operatorname{ext}(K_n(A))$.

A *Hilbert space representation* of a Fréchet *-algebra A with identity is a *-homomorphism $\mu: A \rightarrow \mathfrak{B}(H)$ for a Hilbert space H. A consequence of Do-Shing's theorem (see Lemma 3.1 below) is that every such representation is continuous. Moreover, there is a one-to-one correspondence between the members of $K(A)$ and the equivalence classes of cyclic Hilbert space representations of A (with unit cyclic vectors). This correspondence matches elements of $K_n(A)$ with those representations which can be factored through A_n. Also, the indecomposable positive functionals on A correspond to classes of irreducible Hilbert space representations of A.

The **-radical*, $R^*(A)$, of A is the set $\{a \in A: f(a^*a) = 0\ (f \in P(A))\} = \{a \in A: f(a^*a) = 0\ (f \in \operatorname{ext}(K(A))\} = \bigcap \{\ker \pi: \pi$ is an irreducible Hilbert space representation of $A\}$. If A is an F^*-algebra with identity, then $R^*(A) = (0)$. Hence, if we let Λ be all equivalence classes of irreducible Hilbert space representations of A and for each $\lambda \in \Lambda$ we choose $\pi_\lambda \in \lambda$ with representation space H_λ, then $\{(\pi_\lambda, H_\lambda): \lambda \in \Lambda\}$ is a complete family of irreducible Hilbert space representations of A. A family constructed in this manner for a Fréchet *-algebra A will be called a *standard family of irreducible Hilbert space representations of* A. If $\{\pi_\lambda, H_\lambda): \lambda \in \Lambda\}$ is a standard family for A and we let $E_n = \{\lambda: \pi_\lambda$ factors through $A_n\}$, then for each $\lambda \in E_n$ there exists a unique irreducible representation σ_λ of A_n on H_λ so that $\sigma_\lambda \circ \rho_n = \pi_\lambda$. The family $\{\sigma_\lambda: \lambda \in E_n\}$ is a complete family of irreducible representations for A_n (in case A is F^*) and the direct sum $\sum_{\lambda \in E_n} \sigma_\lambda$ on $\sum_{\lambda \in E_n} \oplus H_\lambda$ is an isometry and *-isomorphism of A_n into $\mathfrak{B}(\sum_{\lambda \in E_n} \oplus H_\lambda)$.

We have included no proofs of the facts quoted above since those concerning Fréchet *-algebras are proved for the more general class of locally m-convex *-algebras in [1], and those relating to Banach *-algebras can be found in [6].

2. Certain operator algebras. In this section we obtain the results about special algebras of operators on direct sums and inductive limits of Hilbert spaces which we need in the proof of the main representation theorem in § 3. The concepts considered in the first part of this section are discussed in detail in G. Lassner's work [4].

We first establish our notation. If X is a complex vector space we denote by $\mathscr{L}_a(X)$ the algebra of all linear transformations on X. If X has a locally convex topology τ we denote by $\mathscr{L}(X)$, or by $\mathscr{L}(X_\tau)$ if there are several topologies on X in the discussion, the subalgebra of $\mathscr{L}_a(X)$ consisting of all τ-continuous operators. For a locally convex TVS (X, τ) we denote by $\mathscr{S}$ the family of all τ-bounded subsets of X (with an appropriate subscript on $\mathscr{S}$ if there are several topologies on X). The topology of bounded convergence $\mathscr{T}_b$ is the topology on $\mathscr{L}(X)$ with base at 0 $\{Nbd(M, U): M \in \mathscr{S}, U$ a τ-neighborhood of 0 in $X\}$, where $Nbd(M, U) = \{T \in \mathscr{L}(X): T(M) \subseteqq U\}$.

DEFINITION. Let X be an inner product space with inner product $(\cdot, \cdot)$, and let H be the completion of X. $\mathscr{L}^+(X)$ is the subset of $\mathscr{L}_a(X)$ which consists of all $T \in \mathscr{L}_a(X)$ which have an adjoint in $\mathscr{L}_a(X)$: there exists $S \in \mathscr{L}_a(X)$ such that $(Tx, y) = (x, Sy) (x, y \in X)$.

LEMMA 2.1. (Lassner) *$\mathscr{L}^+(X)$ is a *-algebra with involution $T \to T^*$. Also,* (1) *each $T \in \mathscr{L}^+(X)$ is closed,* (2) *if $X = H$, then $\mathscr{L}^+(X) = \mathfrak{B}(H)$, and* (3) *if there is a closed operator in $\mathscr{L}^+(X)$, then $X = H$.*

Proof. This is merely a compilation of Lemmas 2.1 and 2.2 of [4].

DEFINITION. An *Op^*-algebra* on X is a *-subalgebra $\mathfrak{A}$ with identity I of $\mathscr{L}^+(X)$, (i.e., the identify of $\mathfrak{A}$ is the identity operator on X).

DEFINITION. Let $\mathfrak{A}$ be an Op^*-algebra on X. We define a locally convex topology $\tau_{\mathfrak{A}}$ by taking as a sub-basic family of seminorms $\{\|\cdot\|_T: T \in \mathfrak{A}\}$, where $\|x\|_T = \|Tx\|$ $(x \in X)$. This is the coarsest topology on X with respect to which each operator in $\mathfrak{A}$ is a continuous map into H.

Lassner shows [Lemma 3.1, 4] that each $T \in \mathfrak{A}$ is a continuous linear transformation on $(X, \tau_{\mathfrak{A}})$. Since $I \in \mathfrak{A}$ it follows that $\tau_{\mathfrak{A}}$ is finer than the norm topology of H restricted to X.

DEFINITION. Let $\mathfrak{A}$ be an Op^*-algebra on X. We define two topologies $\mathscr{T}_{\mathfrak{A}}$ and $\mathscr{T}^{\mathfrak{A}}$ on $\mathfrak{A}$ by:

(1) $\mathscr{T}_{\mathfrak{A}}$ is defined the family $\{\|\cdot\|_M: M \in \mathscr{S}_{\mathfrak{A}}\}$ of seminorms where, (a) $\mathscr{S}_{\mathfrak{A}}$ is the family of $\tau_{\mathfrak{A}}$-bounded subset, of X and (b) $\|T\|_M = \sup\{|(Tx, y)|: x, y \in M\}$.

(2) $\mathscr{T}^{\mathfrak{A}}$ is the restriction to $\mathfrak{A}$ of the topology $\mathscr{T}_b$ on $\mathscr{L}(X, \tau_{\mathfrak{A}})$.

LEMMA 2.2. (Lassner) *If $\mathfrak{A}$ is an Op^*-algebra on X, then,*

(1) $(\mathfrak{A}, \mathscr{T}^{\mathfrak{A}})$ *is a locally convex algebra* (separate continuity of multiplication), *but the involution is not in general continuous.*

(2) $(\mathfrak{A}, \mathscr{T}_{\mathfrak{A}})$ *is a locally convex algebra with continuous involution.*

(3) $\mathscr{T}_{\mathfrak{A}} \leqq \mathscr{T}^{\mathfrak{A}}$, *and* $\mathscr{T}_{\mathfrak{A}} = \mathscr{T}^{\mathfrak{A}}$ *if, and only if, the multiplication in* $(\mathfrak{A}, \mathscr{T}_{\mathfrak{A}})$ *is* (jointly) *continuous.*

Proof. This is a compilation of Theorems 4.1 and 4.2 and Example 5.1 of [4].

NOTATION. For the maximal Op^*-algebra $\mathscr{L}^+(X)$ on X we shall let τ_+ and $\mathscr{T}_+$ replace the clumsier notation $(\tau_{\mathscr{L}^+(X)}, \mathscr{T}_{\mathscr{L}^+(X)})$ of the definitions above.

We now specialize to a particular class of inner product spaces. Let $\{H_\beta: \beta \in B\}$ be a family of Hilbert spaces, let $H = \Sigma_\beta \oplus H_\beta$ and let $X = \Sigma_\beta H_\beta$ (the algebraic direct sum). For $\beta \in B$ we let $p_\beta: X \to H_\beta$ be the natural projection and let $q_\beta: H_\beta \to X$ be the natural injection. For $x \in X$ we define $\mathrm{Supp}\,(x) = \{\beta: p_\beta(x) \neq 0\}$.

The locally convex direct sum topology, τ_f, on X is the final topology determined by the family $\{q_\beta: \beta \in B\}$. We shall abbreviate (X, τ_f) by X_f.

LEMMA 2.3. *If $X = \Sigma_\beta H_\gamma$, then $\mathscr{L}^+(X)$ is isomorphic to the algebra of all B^2-matrices $(T_{\alpha\beta})_{\alpha,\beta \in B}$ such that*

(1) $T_{\alpha\beta} \in \mathscr{L}(H_\beta, H_\alpha)(\alpha, \beta \in B)$,
(2) *for each* $\alpha \in B$ *the set* $B_\alpha = \{\beta\colon T_{\alpha\beta} \neq 0\}$ *is finite, and*
(3) *for each* $\beta \in B$ *the set* $B^\beta = \{\alpha\colon T_{\alpha\beta} \neq 0\}$ *is finite.*

Proof. *If* $(T_{\alpha\beta})$ is a matrix satisfying (1)—(3) we define $T\colon X \to X$ by $(T\{X_\beta\})_\alpha = \sum_\beta T_{\alpha\beta}x_\beta$ for each α. Since Supp (x) is finite, $\sum_\beta T_{\alpha\beta}x_\beta$ converges in H_α for each $\alpha \in B$ and it is easily seen that the set of α for which $\Sigma_\beta T_{\alpha\beta}x_\beta$ is nonzero is contained in $\cup\{B^\beta\colon \beta \in \mathrm{Supp}\,(x)\}$, a finite set. Thus, $T \in \mathscr{L}_a(X)$ and by considering the matrix $(T^*_{\alpha\beta})(T^*_{\alpha\beta}\colon H_\alpha \to H_\beta)$ we obtain an adjoint for T in $\mathscr{L}_a(X)$; hence, $T \in \mathscr{L}^+(X)$.

Fix $T \in \mathscr{L}^+(X)$. For $\alpha, \beta \in B$ define $T_{\alpha\beta} = p_\alpha T q_\beta\colon H_\beta \to H_\alpha$. Clearly, $T_{\alpha\beta}$ is a linear transformation. We show that it has an everywhere defined adjoint, hence is bounded. Set $S_{\beta\alpha} = p_\beta T^* q_\alpha\colon H_\alpha \to H_\beta$. Fix $x_\alpha \in H_\alpha$, $x_\beta \in H_\beta$, then;

$$\begin{aligned}(T_{\alpha\beta}x_\beta, x_\alpha) &= (p_\alpha T q_\beta x_\beta, x_\alpha)\\ &= (p_\alpha T q_\beta x_\beta, p_\alpha q_\alpha x_\alpha)\\ &= \Sigma_\gamma(p_\gamma T q_\beta x_\beta, p_\gamma q_\alpha x_\alpha)\\ &= (Tq_\beta x_\beta, q_\alpha x_\alpha) = (q_\beta x_\beta, T^* q_\alpha x_\alpha)\\ &= (x_\beta, S_{\beta\alpha}x_\alpha) \text{ (by reversing the steps above)}\,.\end{aligned}$$

Fix $\beta \in B$. If B^β is not finite, then there exists a sequence $\{\alpha_j\}$ it B so that $T_{\alpha_j\beta} \neq 0$ (j = 1, 2, $\cdots$). For each $x_\beta \in H_\beta$ there exists $n(x_\beta)$ so that $T_{\alpha_j\beta}(x_\beta) = 0$ for $j > n(x_\beta)$. We choose sequences $\{n_j\}$ in $\boldsymbol{N}$ and $\{x_j\}$ in H_β by the following procedure. Let $n_1 = 1$ and choose $x_1 \in H_\beta$ so that $||x_1|| < 2^{-1}$ and $T_{n_1}x_1 \neq 0$ (hereafter T_j will be used for $T_{\alpha_j\beta}$). There exists $n_2 > n_1$ such that $T_j x_1 = 0$ for $j \geqq n_2$. Choose $x_2 \in H_\beta$ such that $T_{n_2}x_2 \neq 0$ and $||x_2|| < \min\,(2^{-2}, 2^{-2}||T_{n_1}x_1||/||T_{n_1}||)$. Continuing inductively we obtain sequences $\{n_j\}$ and $\{x_j\}$ so that:

(1) $1 = n_1 < n_2 < \cdots$.
(2) $||x_j|| < \min\,(2^{-j}, 2^{-j}||T_{n_1}x_1||/||T_{n_1}||, \cdots, 2^{-j}||T_{n_{j-1}}x_{j-1}||/||T_{n_{j-1}}||)$
(3) $T_{n_j}x_i \neq 0$
(4) $T_{n_i}x_j = 0$ $(i > j)$.

We let $x = \sum_{j=1}^\infty x_j \in H_\beta$. We claim that $T_{n_k}x \neq 0$ $(k \geqq 1)$. Fix $k \in \boldsymbol{N}$, then $T_{n_k}x = \sum_{j-1}^{k-1} T_{n_k}x_j + T_{n_k}x_k + \sum_{j=k+1}^\infty T_{n_k}x_j$. For $j \leqq k-1$ we have $T_{n_k}x_j = 0$ and for $j > k+1$ we have $||T_{n_k}x_j|| \leqq ||T_{n_k}||\,||x_j|| < 2^{-j}||T_{n_k}x_k||$. If $T_{n_k}x = 0$, then, $||T_{n_k}x_k|| = ||\sum_{j=k+1}^\infty T_{n_k}x_j|| \leqq \sum_{j=k+1}^\infty 2^{-j}$ $||T_{n_k}x_k|| < ||T_{n_k}x_k||$, a contradiction.

That B_α is finite for each $\alpha \in B$ follows by applying the same argument to T^*.

LEMMA 2.4. *Let* $\{X_\beta\}$ *be a family of Banach spaces and let* $X = \sum_\beta X_\beta$. *For* $c \in \boldsymbol{R}_+^B$ *define* $p_c\colon X \to \boldsymbol{R}_+$ *by* $p_c(x) = \sum_\beta c_\beta ||x_\beta||$. *Then*

$\{p_c: c \in \boldsymbol{R}_+^B\}$ *defines the locally convex direct sum topology* τ_f *on* X.

Proof. Clearly, $\{p_c\}$ defines a separated locally convex topology τ' on X. Since τ_f is the final topology generated by the injections q_β: $X_\beta \to X(\beta \in B)$, it suffices to show (1) if p is a τ_f-continuous seminorm on X, then there exists $c \in \boldsymbol{R}_+^B$ so that $p \leqq p_c$ (hence, $\tau_f \leqq \tau'$), and (2), for each $\beta \in B$ the map q_β: $X_\beta \to (X, \tau')$ is continuous.

(1) Fix a τ_f-continuous seminorm p on X. For each $\beta \in B$ the map $p \circ q_\beta$: $X_\beta \to \boldsymbol{R}$ is a continuous seminorm on X_β. Hence, there exists $c_\beta \in \boldsymbol{R}_+$ so that $p \circ q_\beta(x_\beta) \leqq c_\beta \| x_\beta \| \, (x_\beta \in X_\beta)$. This defines the function $c \in \boldsymbol{R}_+^B$. If $x = \{x_\beta\} \in X$, then,

$$\begin{aligned} p(x) &= p(\textstyle\sum_\beta q_\beta(x_\beta)) \leqq \textstyle\sum_\beta p \circ q_\beta(x_\beta) \\ &\leqq \textstyle\sum_\beta c_\beta \| x_\beta \| = p_c(x) \ . \end{aligned}$$

(2) Fix $\beta \in B, c \in \boldsymbol{R}_+^B$. Then $p_c(q_\beta x_\beta) = c_\beta \| x_\beta \|$ and q_β: $X_\beta \to (X, \tau')$ is continuous.

LEMMA 2.5. *Let* $X = \sum_\beta H_\beta$. *For each* $c \in \boldsymbol{R}_+^B$ *we define* $\| \ \|_c$: $X \to \boldsymbol{R}$ *by* $\| x \|_c = [\sum_\beta c_\beta^2 \| x_\beta \|^2]^{1/2}$ *for* $x = \{x_\beta\} \in X$. *If* B *is countable then* τ_f *is defined by the seminorms* $\{\| \cdot \|_c: c \in \boldsymbol{R}_+^B\}$.

Proof. Suppose $B = \boldsymbol{N}$. We have for each $c \in \boldsymbol{R}_+^N$ that $\| \cdot \|_c \leqq p_c$, so $\tau_{\{\| \cdot \|_c\}} \leqq \tau_f$. We fix $c \in \boldsymbol{R}_+^N$.

For $x \in X$ we have:

$$\begin{aligned} p_c(x) &= \textstyle\sum_n c_n \| x_n \| = \textstyle\sum_n n^{-1}(nc_n \| x_n \|) \\ &\leqq (\textstyle\sum_n n^{-2})^{1/2} (\textstyle\sum_n (nc_n)^2 \| x_n \|^2)^{1/2} \\ &= (\text{constant}). \quad \| x \|_{\{nc_n\}} \ . \end{aligned}$$

THEOREM 2.6. *If* $X = \sum_\beta H_\beta$, *then* $\mathscr{L}^+(X) \subseteqq \mathscr{L}(X_f)$; *hence,* $\mathscr{L}^+(X) = \mathscr{L}^*(X_f)$, *the algebra of continuous operators on* X_f *with continuous adjoints.*

Proof. It suffices to show that for each $T \in \mathscr{L}^+(X)$ and $\beta \in B$ the operator $T \circ q_\beta$; $H_\beta \to X_f$ is continuous (see [Prop. 6.1, p. 54, 7]). Fix $T \in \mathscr{L}^+(X)$, $\beta \in B$ and a seminorm p_c, $c \in \boldsymbol{R}_+^B$, for τ_f. The set $B^\beta = \{\alpha: T_{\alpha\beta} \neq 0\}$ is finite, so for $x_\beta \in H_\beta$ we have:

$$\begin{aligned} p_c(T \circ q_\beta(x_\beta)) &= \textstyle\sum_\alpha c_\alpha \| (T \circ q_\beta(x_\beta))_\alpha \| \\ &= \textstyle\sum_{\alpha \in B^\beta} c_\alpha \| T_{\alpha\beta} x_\beta \| \ . \end{aligned}$$

Hence, $T \circ q_\beta$ is continuous.

We now turn from direct sums to inductive limits. Let $\{H_\lambda: \lambda \in \Lambda\}$ be a family of Hilbert spaces. Let $H = \sum_\lambda \oplus H_\lambda$. We fix a family $\mathscr{E}$ of subsets of Λ which satisfies (a)$\mathscr{E}$ is closed under finite unions, and (b) all subsets of a member of $\mathscr{E}$ belong to $\mathscr{E}$ (i.e., $\mathscr{E}$ is an ideal in the lattice 2^Λ). We let $X = \{x \in H: \operatorname{Supp}(x) \in \mathscr{E}\}$, and for $E \in \mathscr{E}$ we let $H_E = \{x \in H: \operatorname{Supp}(x) \subseteqq E\}$, a Hilbert space, and let i_E be the identity injection of H_E into X. Finally, let τ_f be the final topology on X determined by the family $\{i_E: E \in \mathscr{E}\}$ of injections.

Since X has an inner product the Op^*-algebra $\mathscr{L}^+(X)$ is defined. A subalgebra of $\mathscr{L}^+(X)$ of importance here is $\mathscr{L}_r(X) = \{T \in \mathscr{L}^+(X)$: for each $E \in \mathscr{E}$, $T(H_E) \subseteqq H_E$ and $T^*(H_E) \subseteqq H_E\}$; i.e., $\mathscr{L}_r(X)$ consists of all elements of $\mathscr{L}^+(X)$ which are reduced by each $H_E (E \in \mathscr{E})$.

We denote the topologies on X determined by $\mathscr{L}^+(X)$ and $\mathscr{L}_r(X)$ by τ_+ and τ_r (respectively) and the corresponding families of bounded sets by $\mathscr{S}_+$ and $\mathscr{S}_r$.

ASSUMPTION. Throughout the remainder of this section we assume the existence of an ascending *countable* cofinal (with respect to the partial order $\subseteqq$ on $\mathscr{E}$) subfamily $\mathscr{E}_0 = \{E_n\}_{n=1}^\infty$. We let the corresponding Hilbert spaces H_{E_n} and injections i_{E_n} be denoted H_n and $i_n (n \in N)$.

LEMMA 2.7. *The final topology on X generated by the family $\{i_n\}_{n=1}^\infty$ is τ_f. Hence, (X, τ_f) is a strict inductive limit of the sequence of Hilbert spaces $\{H_n\}$ and;*

(1) *$\tau_f \mid H_n$ is the norm topology on H_n.*

(2) *$M \subseteqq X$ is τ_f-bounded if, and only if, there exists $n \in N$ such that M is a* (norm) *bounded subset of H_n.*

Proof. That the final topologies are the same can either be easily proved directly or deduced from Proposition 3, p. 159 of [3]. That we have a strict inductive limit and (1) follow from the fact that $\tau_{n+1} \mid H_n = \tau_n$ (trivial if one writes out the norm of an element of H_n considered as an element of H_{n+1}) and a theorem of Dieudonné-Schwartz (see [pp. 159–160, 3]). Claim (2) is another theorem of Dieudonné-Schwartz (see [p. 161, 3]).

THEOREM 2.8. $\mathscr{S}_r = \mathscr{S}_f$

Proof. We show first that $\mathscr{L}_r(X)$ is a subalgebra of $\mathscr{L}(X_f)$. Fix $T \in \mathscr{L}_r(X)$, $n \in N$. We must show that $T \circ i_n = T \mid H_n: H_n \to X_f$ is continuous. But $T(H_n) \subseteqq H_n$ and $\tau_f \mid T_n = \tau_n$. Thus, we must show that $T \mid H_n: H_n \to H_n$ is continuous. Since $T \in \mathscr{L}_r(X)$ we have that $T^*(H_n) \subseteqq H_n$, so $(T \mid H_n)^* = T^* \mid H_n$ and $T \mid H_n$ has an everywhere defined adjoint on H_n, hence is continuous.

Let $M \in \mathscr{S}_f$. Then M is a bounded subset of H_n for some $n \in N$.

For each $T \in \mathscr{L}_r(X)$, $T(M)$ is a bounded subset of H_n since $T \mid H_n$ is continuous on H_n. Hence, $T(M)$ is bounded in H. Since T was arbitrary, $M \in \mathscr{S}_r$.

Suppose $M \notin \mathscr{S}_f$. Case (a). There exist sequences $\{n_j\}$ in N and $\{x_j\}$ in M so that:

(1) $1 = n_1 < n_2 < \cdots$.

(2) $x_j \in H_{n_{j+1}} \backslash H_{n_j} (j = 1, 2, \cdots)$.

Choose $x_1 \in M \backslash H_{n_1} (n_1 = 1)$ let n_2 be sufficiently large that $\operatorname{Supp}(x_1) \subseteqq E_{n_2}$, choose $x_2 \in M \backslash H_{n_2}$, etc. Let $D_j = E_{n_j} \backslash E_{n_{j-1}} (n_0 = 0, E_0 = \varnothing)$, and set $C_j = (\sum_{\lambda \in D_j} \| x_{j,\lambda} \|)^{1/2}$. Define $T: X \to X$ by $(Tx)_\lambda = jC_j^{-1} x_\lambda$ if $\lambda \in D_j$. It is easily verified that $T \in \mathscr{L}_a(X)$, $T^* = T$, and $\operatorname{Supp}(Tx) \subseteqq \operatorname{Supp}(x)$ for $x \in X$. Hence, $T \in \mathscr{L}_r(X)$. Also,

$$\begin{aligned} \| Tx_n \|^2 &= \textstyle\sum_{j=1}^{\infty} \sum_{\lambda \in D_j} (jC_j^{-1})^2 \| x_{n,\lambda} \|^2 \\ &\geqq (nC_n^{-1})^2 \textstyle\sum_{\lambda \in D_n} \| x_{n,\lambda} \|^2 = n^2 . \end{aligned}$$

Hence, $\sup_{x \in M} \| Tx \| = \infty$, and $M \notin \mathscr{S}_r$. Case (b). $M \subseteqq H_n$, for some n, but is unbounded. Easy to show that $M \notin \mathscr{S}_r$.

THEOREM 2.9. *$\mathscr{T}_r = \mathscr{T}_b \mid \mathscr{L}_r(X)$, where $\mathscr{T}_b$ is the topology of bounded convergence on the algebra $\mathscr{L}(X_f)$.*

Proof. (1) $\mathscr{T}_r \leqq \mathscr{T}_b$ on $\mathscr{L}_r(X)$: Fix a $\mathscr{T}_r$-neighborhood of 0 in $\mathscr{L}_r(X)$, $\operatorname{Nbd}(M, \varepsilon) = \{T: \| T \|_M < \varepsilon\}$, where $M \in \mathscr{S}_r$, $\varepsilon > 0$. Since $\mathscr{S}_r = \mathscr{S}_f$ there exists $n \in N$ so that $M \subseteqq H_n$ and $\| M \| = \sup_{x \in M} \| x \| < \infty$. Let U be a τ_f-neighborhood of 0 in X so that $U \cap H_n \subseteqq S_n(0, (2 \| M \|)^{-1} \varepsilon)$, the $(2 \| M \|)^{-1}$ ε-ball about 0 in H_n. If $T \in \mathscr{L}_r(X) \cap N\text{bd}(M, U) = \{S \in \mathscr{L}_r(X): S(M) \subseteqq U\}$, then $T(M) \subseteqq U \cap H_n$ and $\| Tx \| < (2 \| M \|)^{-1} \varepsilon$ for $x \in M$. Now

$$\begin{aligned} \| T \|_M &= \sup \{| (Tx, y) |: x, y \in M\} \\ &\leqq \sup_{x \in M} \{\sup_{\| y \| \leqq \| M \|} | (Tx, y) |\} \\ &\leqq \sup \| M \| \cdot \| Tx \| \leqq \varepsilon/2 < \varepsilon . \end{aligned}$$

This shows that $\mathscr{T}_r \leqq \mathscr{T}_b$.

(2) $\mathscr{T}_b \leqq \mathscr{T}_r$ on $\mathscr{L}_r(X)$: Fix a $\mathscr{T}_b$-neighborhood of 0 in $\mathscr{L}_r(X)$, $N\text{bd}(M, U)$ where $M \in \mathscr{T}_f$, U is a τ_f-neighborhood of 0 in X. There exists $n \in N$ so that M is a bounded subset of H_n. Let $M_1 = M \cup S_n(0, 1)$, bounded subset of H_n; hence, a τ_f-bounded subset of X. Choose $\varepsilon > 0$ so that $S_n(0, \varepsilon) \subseteqq U \cap H_n$. Suppose $T \in N\text{bd}(M_1, \varepsilon)$. If $x \in M$, then

$$\| Tx \| = \sup_{\| y \| \leqq 1} | (Tx, y) | \leqq \sup \{| (Tz, y) |: z, y \in M_1\} = \| T \|_{M_1} < \varepsilon .$$

So $T(M) \subseteqq S_n(0, \varepsilon) \subseteqq U$, and $T \in N\text{bd}(M, U)$. Hence $\mathscr{T}_b \leqq \mathscr{T}_r$.

We now show that $\mathscr{L}^+(X)$ is a subalgebra of $\mathscr{L}(X_f)$. The problem here is that we do not have an obvious charactrization of elements of $\mathscr{L}^+(X)$. We have a fixed cofinal sequence $\{E_n\}$ in $\mathscr{S}$. We let $D_1 = E_1$ and for $n > 1$ we set $D_n = E_n \backslash E_{n-1}$. Let $K_n = \sum_{y \in D_n} \bigoplus H_\lambda$ and let $Y = \sum_n K_n$.

LEMMA 2.10. *The map $u\colon X \to Y$ defined by $u(\{x_\lambda\}_{\lambda \in \Lambda}) = \{\{x_\lambda\}_{\lambda \in D_n}\}_{n=1}^\infty$ has the following properties*:

(1) *u is a linear isomorphism* (onto)

(2) *u is unitary*: $(u(x), u(x')) = (x, x')$ *for $x, x' \in X$.*

(3) *$u\colon (X, \tau_f) \to (Y, \tau_f)$ is topological.*

Proof. It is easily verified that u is a linear isomorphism (into). If $y \in Y$, then there exists $n \in N$ so that $y_j = 0$ for $j > n$. Then $y_{j\lambda} = 0$ if $\lambda \in D_j$, $j > n$. Set $x = \{y_{k\lambda}\colon \lambda \in D_k, k = 1, 2, \cdots\}$. Then $x \in H$ and $\operatorname{Supp}(x) \subseteqq \cup_{j=1}^n D_j = E_n$. Clearly $u(x) = y$, and (1) is proved. That u is unitary depends only on the fact that the series obtained by taking inner products is absolutely convergent so can be rearranged at will.

(3) We show first that u is continuous. We recall Lemma 2.5 and fix a seminorm $\|\cdot\|_c$, $c \in \boldsymbol{R}_+^N$. If $x \in H_n$, then:

$$\begin{aligned}\|u(x)\|_c^2 &= \textstyle\sum_{j=1}^\infty c_j^2 \|u(x)_j\|^2 \\ &= \textstyle\sum_{j=1}^\infty c_j^2 \{\sum_{\lambda \in D_j} \|x_\lambda\|^2\} \\ &= \textstyle\sum_{j=1}^n c_j^2 \{\sum_{\lambda \in D_j} \|x_\lambda\|^2\} \\ &\leqq (\max_{1 \leqq j \leqq n} c_j^2) \|x\|^2 .\end{aligned}$$

Thus, $\|u(x)\|_c \leqq C(n) \|x\|$ for $x \in H_n$ and $u \circ i_n\colon H_n \to Y_f$ is continuous for arbitrary $n \in N$. Hence, u is continuous $(X_f \to Y_f)$.

Since X_f and Y_f are (LF) spaces (each is a strict inductive limit of Hilbert spaces) and u is a continuous surjection of X_f to Y_f it follows that u is an open map (see [Prop 2.2, p. 78, 7]).

THEOREM 2.11. *$\mathscr{L}^+(X) \subseteqq \mathscr{L}(X_f)$; hence, $\mathscr{L}^+(X) = \mathscr{L}^*(X_f)$, the subalgebra of $L(X_f)$ which consists of all operators $T \in \mathscr{L}(X_f)$ whose adjoint T^* exists and belongs to $\mathscr{L}(X_f)$.*

Proof. Let u and Y be as in Lemma 2.10. Since u is a unitary linear isomorphism the induced map $u^*\colon \mathscr{L}_a(X) \to \mathscr{L}_a(Y)$ defined by $u^*(T) = u \circ T \circ u^{-1}$, which maps $\mathscr{L}_a(X)$ isomorphically onto $\mathscr{L}_a(Y)$, maps $\mathscr{L}^+(X)$ onto $\mathscr{L}^+(Y)$. Also, since $u\colon X_f \to Y_f$ is topological the same map u^* maps $\mathscr{L}(X_f)$ onto $\mathscr{L}(Y_f)$. Since $\mathscr{L}^+(Y) \subseteqq \mathscr{L}(Y_f)$ (Theorem 2.6) we must have that $\mathscr{L}^+(X) \subseteqq \mathscr{L}(X_f)$.

THEOREM 2.12. $\mathscr{T}_+ \mid \mathscr{L}_r(X) = \mathscr{T}_b \mid \mathscr{L}_r(X)$.

Proof. We show first that $\tau_+ \leqq \tau_f$. It suffices to show that for each $n \in N$ the injection $i_n\colon H_n \to (X, \tau_+)$ is continuous. Fix a seminorm $\|\cdot\|_T$ for τ_+, where $T \in \mathscr{L}^+(X)$. Since $T \in \mathscr{L}(X_f)$ we have that $T \circ i_n = T \mid H_n\colon H_n \to X_f$ is continuous; hence, $T \mid H_n\colon H_n \to H$ is continuous ($\tau_{\text{norm}} \leqq \tau_f$). But then there exists $C_T > 0$ so that $\| T \mid H_n(x) \| \leqq C_T \|x\|$ for $x \in H_n$; i.e., $\|x\|_T \leqq C_T \|x\|$ $(x \in H_n)$.

We have $\mathscr{S}_r = \mathscr{S}_f \subseteqq \mathscr{S}_+ (\tau_+ \leqq \tau_f)$. But since $\mathscr{L}_r(X) \subseteqq \mathscr{L}^+(X)$ we also have $\mathscr{S}_+ \subseteqq \mathscr{S}_r$. Hence, $\mathscr{S}_r = \mathscr{S}_+$ and $\mathscr{T}_r = \mathscr{T}_+$ on $\mathscr{L}_r(X)$. But Theorem 2.9 says that $\mathscr{T}_r = \mathscr{T}_b$ on $\mathscr{L}_r(X)$.

THEOREM 2.13. *$(\mathscr{L}_r(X), \mathscr{T}_r)$ is complete.*

Proof. Let $\{T_\alpha\}$ be a $\mathscr{T}_r$-Cauchy set in $\mathscr{L}_r(X)$. For each $E \in \mathscr{E}$ we let M_E be the unit ball in H_E. Then, $\{M_E\} \subseteqq \mathscr{S}_r$ and if we fix $E \in \mathscr{E}$, then $\{T_\alpha \mid H_E\}$ is a Cauchy net in $\mathfrak{B}(H_E)$:

$$\begin{aligned} \| T_\alpha \mid H_E - T_\beta \mid H_E \| &= \sup \{\| T_\alpha - T_\beta)x \| : x \in H_E, \|x\| \leqq 1\} \\ &= \sup \{| ((T_\alpha - T_\beta)x, y) | : x, y \in M_E\} \\ &= \| T_\alpha - T_\beta \|_{M_E} . \end{aligned}$$

Note also $\| T_\alpha^* \mid H_E - T_\beta^* \mid H_E \| = \|T_\alpha - T_\beta\|_{M_E}$. Thus, $T_\alpha \mid H_E \to T_E \in \mathfrak{B}(H_E)$ and $T_\alpha^* \mid H_E \to S_E \in \mathfrak{B}(H_E)$. We define T and S on X by $Tx = T_E x$ if $x \in H_E$ and $Sx = S_E x$ if $x \in H_E$. If $E \subseteqq F$, then $T_F \mid H_E = T_E$ $(S_F \mid H_E = S_E)$. Hence, T and S are well-defined linear transformations on X. Clearly, both T and S leave each H_E invariant and $S = T^*$. Thus, $T \in \mathscr{L}_r(X)$. That $\mathscr{T}_r$-$\lim_\alpha T_\alpha = T$ is easily checked.

THEOREM 2.14. *$(\mathscr{L}_r(X), \mathscr{T}_r)$ is an F^*-algebra with identity. In fact, $(\mathscr{L}_r(X), \mathscr{T}_r) \cong \lim_n \operatorname{inv} \mathfrak{B}(H_n)$.*

Proof. For each $n \in N$ we let M_n be the unit ball in H_n. Then $\{M_n\} \subseteqq \mathscr{S}_r$ and is "essentially" cofinal: if $M \in \mathscr{S}_r$, then M is a bounded subset of some H_n, hence there exists $k \in \boldsymbol{R}_+$ so that $M \subseteqq kM_n$. But then $\| T\|_M \leqq k^2 \| T\|_{M_n} (T \in \mathscr{L}_r^*(X))$. Thus, the topology $\mathscr{T}_r$ is determined by th ascending family $\{\|\cdot\|_{M_n}\}$ of (linear) seminorms, and $(\mathscr{L}_r(X), \mathscr{T}_r)$ is a complete metrizable algebla. As we saw in Theorem 2.13 for $T \in \mathscr{L}_r(X)$ and $n \in N$ we have $\| T\|_{M_n} = \| T^* \|_{M_n} = \| T | H_n \|$. Thus, each $\|\cdot\|_{M_n}$ is a B^*-seminorm, and $(\mathscr{L}_r(X), \mathscr{T}_r)$ is an F^*-algebra with identity. The last part of the conclusion was essentially proved in Theorem 2.13. The map $\rho_n\colon \mathscr{L}_r(X) \to \mathfrak{B}(H_n)$ is just the restriction map; as is the bonding map $\rho^n\colon \mathfrak{B}(H_n) \to \mathfrak{B}(H_{n-1}) (n \in N)$.

3. A representation theorem for F^*-algebras. In this section we give three concrete faithful (topologically and algebraically) representations for an abstract F^*-algebra with identity as an algebra of operators on a vector space formed from the irreducible Hilber space representations of the algebra.

We let A be an F^*-algebra with identity and let $\{(\pi_\lambda H_\lambda): \lambda \in \Lambda\}$ be a standard family of irreducible Hilbert space representations of A. Since A is an F^*-algebra the family is complete.

If A is a B^*-algebra and $\{(\pi_\lambda, H_\lambda)$ is a standard family for A, then $\pi: A \to \mathfrak{B}(\sum_\lambda \oplus H_\lambda)$ defined by $\pi(a)\{x_\lambda\} = \{\pi_\lambda(a)x_\lambda)$ is an isometry and $*$-isomorphism. It is easily seen that for non-B^*-algebras (but still F^*-) this is impossible. In fact, one cannot even define $\pi(a)$ on $\sum_\lambda \oplus H_\lambda$ for all $a \in A$, unless every $a \in A$ has bounded norm: $\sup_n ||a||_n < \infty$ for some determining family of seminorms. If one moves to the other extreme and defines $\pi(a)$ by the same formula on $X = \sum_\lambda H_\lambda$ (algebraic sum), then $\pi(a)$ makes sense and $\pi: A \to (\mathscr{L}(X_f), \mathscr{T}_b)$ is a continuous $*$-isomorphism but fails to be topological. This is the case because the final topology on X, hence the topology $\mathscr{T}_b$ on $\mathscr{L}(X_f)$, depends on finite subsets of Λ whereas that of A depends on much larger subsets of Λ. Thus, we must seek a middle ground in order to achieve a faithful representation of A in this manner. Before we introduce the basic concept we first prove a crucial fact about Hilbert space representations of Fréchet $*$-algebras.

LEMMA 3.1. *Let A be a Fréchet $*$-algebra with identity, and let $\mu: A \to \mathfrak{B}(H)$ be a representation of A on the Hilbert space H. Then μ is continuous.*

Proof. Fix $\varepsilon > 0$. Let $V = \{a \in A: ||\mu(a)|| \leqq \varepsilon\} = \cap \{V_{x\,y}: ||x||, ||y|| \leqq 1\}$, where $V_{x\,y} = \{x \in A: |(\mu(a)x, y)| \leqq \varepsilon\}$. For each pair $x, y \in H$ such that $||x||, ||y|| \leqq 1$ the set $V_{x, y}$ is convex and balanced. Since for each $z \in H$ the map $a \to (\mu(a)z, z)$ is continuous (Do-Shing's Theorem [2]), we have that $a \to (\mu(a)x, y)$ is continuous (polarization formula). Thus, each $V_{x\,y}$ is closed. So V is closed, convex, and balanced. It is easily verified that V is absorbing; hence is a neighborhood of 0 in A.

DEFINITION. Let A be a Fréchet $*$-algebra with identity and let $\{(\pi_\lambda, H_\lambda): \lambda \in \Lambda\}$ be a stadard family of irreducible Hilbert space representations of A. A subset E of Λ will be called *equicontinuous* if, and only if $\sup_{\lambda \in E} ||\pi_\lambda(a)|| < \infty$ for each $a \in A$. The family of all equicontinuous subsets of Λ will be denoted $\mathscr{E}(\Lambda)$.

LEMMA 3.2. *If A is a Fréchet $*$-algebra with identity and* $\{(\pi_\lambda,$

H_λ): $\lambda \in \Lambda\}$ is a standard family for A, then $E \subseteqq \Lambda$ is equicontinuous if, and only if, $\sum_{\lambda \in E} \pi_\lambda$ defines a continuous representation of A on $\sum_{\lambda \in E} \oplus H_\lambda$.

Proof. Suppose $E \subseteqq \Lambda$ is equicontinuous. For $a \in A$ we let $C_a = \sup_{\lambda \in E} \|\pi(a)\|$ and define $\pi: A \rightarrow \mathfrak{B}(\sum_{\lambda \in E} \oplus H_\lambda)$ by $\pi(a)\{x_\lambda\} = \{\pi_\lambda(a)x_\lambda\}$. Now $\|\{\pi_\lambda(a)x_\lambda\}\|^2 = \sum_\lambda \|\pi_\lambda(a)x_\lambda\|^2 \leqq C_a^2 \|(x)\|^2$. So $\pi(a)$ maps $\sum_{\lambda \in E} \oplus H_\lambda$ into itself, and π is a representation of A on $\sum_{\lambda \in E} \oplus H_\lambda$.

Conversely, suppose we can define a representation of A on $\sum_{\lambda \in E} \oplus H_\lambda$ by the direct sum formula. It is clear that $\|\pi_\lambda(a)\| \leqq \|\pi(a)\|$ for each $\lambda \in E$ and $a \in A$.

LEMMA 3.3. *Let A, $\{(\pi_\lambda, H_\lambda): \lambda \in \Lambda\}$ be as above. Let $\{\|\cdot\|_n\}$ be a *-sequence of seminorms for A. For $n \in N$ we set $E_n = \{\lambda \in \Lambda: \|\pi_\lambda(a)\| \leqq \|a\|_n\ (a \in A)\}$. Then $E \subseteqq \Lambda$ is equicontinuous if, and only if, E is contained in some E_n. In particular, the increasing sequence $\{E_n\}$ is cofinal in $\mathscr{E}(\Lambda)$.*

Proof. If $E \subseteqq E_n$ for some n, then clearly $E \in \mathscr{E}(\Lambda)$. Conversely, if $E \in \mathscr{E}(\Lambda)$ then $\pi: A \rightarrow \mathfrak{B}(\sum_{\lambda \in E} \oplus H_\lambda)$ defined as in Lemma 3.2 is a continuous representation of A. Hence, there exists $C > 0$, $n \in N$ so that $\|\pi(a)\| \leqq C\|a\|_n\ (a \in A)$. It is easily verified that we can take $C = 1$, and the condition is satisfied.

We set $H = \sum_{\lambda \in \Lambda} \oplus H_\lambda$ and let $X = \{x \in H: \mathrm{Supp}\,(x) \in \mathscr{E}(\Lambda)\}$. We are now in the situation of the second part of § 2 with $H_E = \{x: \mathrm{Supp}\,(x) \subseteqq E\}$, $i_E: H_E \rightarrow X$ the natural injection, τ_f the final topology determined by the family $\{i_E: E \in \mathscr{E}(\Lambda)\}$. If $\{\|\cdot\|_n\}$ is any F^*sequence of seminorms for our F^*-algebra A with identity, then we let $H_n = H_{E_n}$ and $\pi_n: A \rightarrow \mathfrak{B}(H_n)$ the induced representation of A on H_n.

LEMMA 3.4. *With the definitions given immediately above for each $n \in N$ and $a \in A$ it is the case that $\|a\|_n = \|\pi_n(a)\|$.*

Proof. In § 1 we indicated that $\{\pi_\lambda: \lambda \in E_n\}$ induces a complete standard family $\{\sigma_\lambda: \lambda \in E_n\}$ of irreducible Hilbert space representations of the B^*-algebra A_n, the completion of $A/\{a: \|a\|_n = 0\}$ with respect to the induced norm, and if ρ_n is the natural projection of A into A_n we have $\sigma_\lambda \circ \rho_n = \pi_\lambda (\lambda \in E_n)$. If we let $\sigma_n = \sum_{\lambda \in E_n} \sigma_\lambda: A_n \rightarrow \mathfrak{B}(H_n)$, then $\|\sigma_n(a_n)\| = \|a_n\|$ for each $a_n \in A_n$. But $\|\sigma_n(a_n)\| = \sup_{\lambda \in E_n} \|\sigma_\lambda(a_n)\|$. Thus, if $a \in A$, then $\|a\|_n = \|\rho_n a\| = \sup_{\lambda \in E_n} \|\sigma_\lambda(\rho_n a)\| = \sup_{\lambda \in E_n} \|\pi_\lambda(a)\| = \|\pi_n(a)\|$.

From the above construction we can infer more. For each $a \in A$ there exists $\lambda_0 \in E_n$ such that $\|a\|_n = \|\pi_{\lambda_0}(a)\|$. This can be proved for A_n by reducing the problem to that for a hermitian elements, then showing that it holds on the algebra generated by the element and extending to the full algebra.

THEOREM 3.5. *Let A be an F^*-algebra with identity, $\{(\pi_\lambda, H_\lambda)\colon \lambda \in \Lambda\}$ a standard family of irreducible Hilbert representations of A, $H = \sum_{\lambda\in\Lambda} \oplus H_\lambda$, and $X = \{x \in H\colon \mathrm{Supp}\,(x)$ is equicontinuous$\}$. Let $\pi\colon A \to \mathscr{L}_a(X)$ be defined by $\pi(a)\{x_\lambda\} = \{\pi_\lambda(a)x_\lambda\}(a \in A, x = \{x_\lambda\} \in X)$. Then*

(1) *For each $a \in A$ the function $\pi(a)$ defined above is indeed in $\mathscr{L}_a(X)$; in fact, $\pi(a) \in \mathscr{L}_r(X)$.*

(2) *$\pi\colon A \to (\mathscr{L}^*(X_f), \mathscr{T}_b)$ is a topological $*$-isomorphism* (into).

(3) *$\pi\colon A \to (\mathscr{L}^+(X), \mathscr{T}_+)$ is a topological $*$-isomorphism* (into).

(4) *$\pi\colon A \to (\pi(A), \mathscr{T}_{\pi(A)})$ is a topological $*$-isomorphism.*

Proof. (1) Fix $a \in A, x \in X$. Then

$$\sum_\lambda ||\pi_\lambda(a)x_\lambda||^2 \leqq \sum \{||\pi_\lambda(a)||^2\, ||x_\lambda||^2\colon \lambda \in \mathrm{Supp}\,(x)\}$$
$$\leqq \sup \{||\pi_\lambda(a)||^2\colon \lambda \in \mathrm{Supp}\,(x)\} \cdot ||x||^2 .$$

Thus, $\pi(a)x \in H$ and $\mathrm{Supp}\,(\pi(a)x) \subseteqq \mathrm{Supp}\,(x) \in \mathscr{E}(\Lambda)$, so $\pi(a)$ maps X into itself. Moreover, if $a \in A$ and $x, y \in X$ we have $(\pi(a)x, y) = (x, \pi(a^*)y)$; so $\pi(a) \in \mathscr{L}^+(X)$. It is clear that $\pi(a) \in \mathscr{L}_r(X)$.

(2) and (3) It is clear that π is a $*$-isomorphism. Since $\mathscr{T}_+ | \mathscr{L}_r(X) = \mathscr{T}_b | \mathscr{L}_r(X) = \mathscr{T}_r$ (Theorem 2.12) and $\pi(A) \subseteqq \mathscr{L}_r(X)$ it is necessary and sufficient that we show $\pi\colon A \to (\mathscr{L}_r(X), \mathscr{T}_r)$ is topological. We fix an F^*-sequence $\{||\cdot||_n\}$ of seminorms for A, let $\{E_n\}$, $\{H_n\}$, and $\{\pi_n\}$ be the corresponding cofinal sequence in $\mathscr{E}(A)$, Hilbert space sequence in X, and sequence of representations of A (respectively). We note that $\pi_n(a) = \pi(a) | H_n$ for $n \in N, a \in A$. We recall from Theorem 2.14 that $(\mathscr{L}_r(X), \mathscr{T}_r)$ is an F-$*$algebra with identity and that $\{||\cdot||_{M_n}\}$ is an F^*-sequence of seminorms for $\mathscr{L}_r(X)$, where M_n is the closed unit ball in $H_n (n \in N)$. Moreover, for each $n \in N$ and $T \in \mathscr{L}_r(X)$ we have $||T||_{M_n} = ||T | H_n||$, the norm of the (bounded) restriction of T to H_n. If $a \in A$, then $||a||_n = ||\pi_n(a)||$ (Lemma 3.4) and the latter is $||\pi(a) | H_n|| = ||\pi(a)||_{M_n}$.

We again fix an F^*-sequence $\{||\cdot||_n\}$ of seminorms for A. Let $\{E_n\}$, etc. be as in "(2) and (3)" above. We let $\mathfrak{A} = \pi(A)$, an Op^*-algebra on X with corresponding family $\mathscr{S}_\mathfrak{A}$ of bounded subsets of X. The topology $\mathscr{T}_\mathfrak{A}$ on $\mathfrak{A}$ is defined by the seminorms $\{||\cdot||_M\colon M \in \mathscr{S}_\mathfrak{A}\}$, where $M \subseteqq X$ belongs to $\mathscr{S}_\mathfrak{A}$ if, and only if, $\sup_{x\in M} ||Tx|| < \infty$ for each $T \in \mathfrak{A}$. Lassner's Lemma 5.2 [4] says that $\pi\colon A \to (\mathfrak{A}, \mathscr{T}_\mathfrak{A})$ is continuous. Fix $n \in N$ and let M_n be the closed unit ball in H_n as above. Since $\mathfrak{A} \subseteqq \mathscr{L}^+(X)$, $\mathscr{S}_+ \subseteqq \mathscr{S}_\mathfrak{A}$ and $\{M_n\} \subseteqq \mathscr{S}_+$; so $\{M_n\} \subseteqq \mathscr{S}_\mathfrak{A}$. We know from above that $||a||_n = ||\pi(a)||_{M_n} (a \in A, n \in N)$. This establishes the openness of $\pi: A \to (\mathfrak{A}, \mathscr{T}_{\mathfrak{a}})$.

REMARKS. (1) Do-Shing [2] obtains a representation theorem for LMC^*-algebras (the same as F^*-algebra without the metrizability restriction) which uses essentially the same Hilbert space, but he maps A onto an algebra of unbounded operators with special properties. Also, he does not consider topological properties of the map.

(2) The main problem in studying non-commutative Fréchet *-algebras is the lack of models against which to compare the abstract algebras. A corollary to Do-Shing's theorem on positive functionals on Fréchet *-algebras is that every one induces a cyclic Hilbert space representation, but as we have seen we cannot represent these algebras faithfully on Hilbert spaces. The examples discussed above, the algebras $\mathscr{L}_r(X)$, are quite similar to those considered by E. A. Michael in Appendix A of his memoir [5], where in our case the underlying locally convex space is an inductive limit of Banach (Hilbert) spaces. It seems that the class he defined in [5] might include most examples of noncommutative F-algebras, except those built from a commutative F-algebra and a noncommutative Banach algebra by tensor products, e.g., $C(X, B)$ where X is an appropriate topological space and B is a Banach algebra.

4. Enveloping algebras. In this section we define the enveloping algebra of a Fréchet *-algebra with identity, relate it to inverse limit decompositions of the algebra, and realize it as an algebra of operators naturally constructed from A.

We fix a Fréchet *-algebra with identity, A, and also fix a standard family $\{(\pi_\lambda, H_\lambda): \lambda \in \Lambda\}$ of irreducible Hilbert space representations of A. We recall that $K(A) = \{f : f$ is a positive functional on A, $f(e) = 1\}$.

LEMMA 4.1. *If* $E \subseteqq K(A)$ *and* $\{\|\cdot\|_n\}$ *is a *-sequence of seminorms for* A, *then the following statements are equivalent.*

(1) *E is equicontinuous.*

(2) $\sup_{f \in E} f(a^*a) < \infty \, (a \in A)$.

(3) *There $n \in N$ such that $E \subseteqq K_n(A)$.*

Proof. (1) and (2) are clearly equivalent by the uniform boundedness principle for Fréchet spaces: if $E \subseteqq A^*$ and E is pointwise bounded ($\sigma(A^*, A)$-bounded), then E is equicontinuous (see [Theorem 4.2, p. 83, 7]). It is also clear that (1) and (3) are equivalent, since $K_n(A)$ is the intersection with $K(A)$ of the polar of the neighborhood $\{a \in A: \|a\|_n \leqq 1\}$.

DEFINITION. Let $\mathscr{E}(K)$ be all equicontinuous subsets of $K(A)$. For $E \in \mathscr{E}(K)$ we define

$$|a|_E = [\sup\{f(a^*a): f \in E\}]^{1/2} \qquad (a \in A) .$$

THEOREM 4.2. *If A is a Fréchet $*$-algebra, $\{|\cdot|_E: E \in \mathscr{E}(K)\}$ is the family of maps defined above, then*

(1) *Each $|\cdot|_E$ is a linear seminorm on A.*

(2) *$R^*(A) = \{a \in A: |a|_E = 0$ for each $E \in \mathscr{E}(K)\}$.*

(3) *If $\{\|\cdot\|_n\}$ is any $*$-sequence of seminorms for A, then the topology of $(A/R^*(A), \{|\cdot|_E\})$ is determined by the B^*-seminorms $\{|\cdot|_n\}$, where $|\cdot|_n = |\cdot|_{K_n(A)}$. Hence,*

(4) *The completion $E(A)$ of $(A/R^*(A), \{|\cdot|_E\})$ is an F^*-algebra with identity.*

Proof. (1) and (2) are trivial to verify and (4) follows from (3), which we now prove. Fix a $*$-sequence $\{\|\cdot\|_n\}$ of seminorms for A. For each $n \in N$ we set $E_n = \{\lambda \in \Lambda: \|\pi_\lambda(a)\| \leqq \|a\|_n \ (a \in A)\}$ and define $\pi_n: A \to \mathfrak{B}(\sum_{\lambda \in E_n} \oplus H_\lambda)$ by $\pi_n(a)(\{\xi\}_\lambda\}_{\lambda \in E_n}) = \{\pi_\lambda(a)\xi_\lambda\}_{\lambda \in E_n}$ for each $a \in A$. We shall show that for each $n \in N$ and $a \in A$ we have $\|\pi_n(a)\| = |a|_n$. Fix $n \in N$. For $\lambda \in \Lambda$ we choose a unit vector $\xi_\lambda \in H_\lambda$, define $f_\lambda: A \to C$ by $f_\lambda(a) = (\pi_\lambda(a)\xi_\lambda, \xi_\lambda)$, let K_λ be the completion of $A/\{a: f_\lambda(a^*a) = 0\}$ with respect to the induced inner product $([a]_\lambda, [b]_\lambda) = f_\lambda(b^*a)$, where $[a]_\lambda$ is the coset containing a. Finally, define $\psi_\lambda: A \to \mathfrak{B}(K_\lambda)$ by $\psi_\lambda(a)([b]_\lambda) = [ab]_\lambda$ on $A/\{a: f_\lambda(a^*a) = 0\}$, and extending these norm-continuous operators to K_λ. There exists an isomorphism $U: H_\lambda K_\lambda$ so that $U\pi_\lambda = \psi_\lambda U$. Hence, for $a \in A$ we have $\|\pi_\lambda(a)\|^2 = \|\psi_\lambda(a)\|^2 = \sup\{f_\lambda(b^*a^*ab): f_\lambda(b^*b) = 1\} \geqq f_\lambda(a^*a)$. If $f_\lambda(b^*b) = 1$, then $f_{\lambda,b}: c \to f_\lambda(b^*cb)$ also belongs to $K_n(A)$ (that f_λ does is clear) and $f_{\lambda,b}(a^*a) \leqq |a|_n$. Hence, $\|\pi_\lambda(a)\| \leqq |a|_n$ for each $\lambda \in E_n$, and $\|\pi_n(a)\| = \sup\|\pi_\lambda(a)\|: \lambda \in E_n\} \leqq |a|_n$. Then reverse inequality follows from the fact that $|a|_n = \sup\{f(a^*a): f \in K_n(A), f$ is extreme$\}$.

DEFINITION. We shall call the algebra $E(A)$ in Theorem 4.2 (4) the *enveloping algebra* of A.

THEOREM 4.3. *If $(A, \{\|\cdot\|_n\})$ is a Fréchet $*$-algebra with identity and if $\{A_n\}$ is the corresponding inverse limit system of Banach $*$-algebras with identity, then $E(A) = \lim_n \operatorname{inv}\{E(A_n)\}$.*

Proof. We let ρ_n be the natural map of A into A_n and $\rho^n: A_n \to A_{n-1} (n \geqq 2)$ the induced bonding map. For $n \in N$ we let E_n be the enveloping algebra of A_n and let Ψ_n be the natural map of A_n into E_n. Finally, we let φ be the map of A into $E(A)$.

For $n \in N$ we have the diagram

$$\begin{tikzcd} A_{n-1} \arrow[d, "\Psi_{n-1}"'] & A_n \arrow[l, "\rho^n"'] \arrow[d, "\Psi_n"] \\ E_{n-1} & E_n \end{tikzcd} .$$

Now $\ker(\Psi_n) = R^*(A_n)$ and $\rho^n(R^*(A_n)) \subseteqq R^*(A_{n-1})$. Thus, we have an induced map $\sigma^n: E_n \to E_{n-1}$. It is easily verified that $\{E_n, \sigma^n, N\}$ is a dense inverse limit system of B^*-algebras (i.e., the bonding maps have dense range and are norm-decreasing). We let $E = \lim_n \operatorname{inv} \{E_n, \sigma^n, N\}$ and consider E a subset of $\Pi_n E_n$ with the relative product topology.

We define $\tau: E(A) \to E$ by first defining τ on $A/R^*(A)$ by the formula $\tau(\varphi a) = \{\Psi_n \rho_n a\}$. If $a \in A$, then $\varphi(a) = 0$ if, and only if, $a \in R^*(A)$ if, and only if, $\rho_n(a) \in R^*(A_n) (n \in N)$ if, and only if, $\Psi_n \rho_n(A) = 0$ $(n \in N)$. Thus, τ is well-defined and one-to-one $A/R^*(A)$. Also, since all the maps involved have dense range it follows that $\tau(A/R^*(A))$ is dense in E. Finally,

$$\begin{aligned} |\tau(\varphi a)|_n^2 &= |\Psi_n \sigma_n(a)|_n^2 = |\rho_n(a)|_n^2 \\ &= \sup\{f_n(\rho_n(a^*a)): f \in K(A_n)\} \\ &= \sup\{f(a^*a): f \in K_n(A)\} \\ &= |a|_n^2 \ (n \in N, a \in A) . \end{aligned}$$

Thus, τ is an isometry in each seminorm; hence, extends to a topological map of E onto $E(A)$. It is clear that the map is a $*$-isomorphism.

We now realize $E(A)$ as an algebra of operators on $X = \{x \in \sum_{\lambda \in \Lambda} \oplus H_\lambda: \operatorname{Supp}(x)$ is equicontinuous$\}$. We use the same notation as in § 3.

THEOREM 4.4. *Let $(A, \{\|\cdot\|_n\})$ be a Fréchet $*$-algebra with identity and let $(E(A), \{|\cdot|_n\})$ be its enveloping algebra, with natural map φ: $A \to E(A)$. For $a \in A$ we define $\pi(a)$ on X by $\pi(a)\{x_\lambda\} = \{\pi_\lambda(a)x_\lambda\}$. Then $\pi: A \to \mathscr{L}_r(X)$ induces a topological $*$-isomorphism $\bar{\sigma}$ of $E(A)$ onto $\overline{\pi(A)}$, where "topological" refers to any of the (equal on $\mathscr{L}_r(X)$) topologies $\mathscr{T}_r$, $\mathscr{T}_+$, or $\mathscr{T}_b$ on $\mathscr{L}_r(X)$ and the closure of $\pi(A)$ is with respect to these topologies.*

Proof. Since for each $a \in A$ and $n \in N$ we have $|a|_n = \|\pi_n(a)\|$ we have that $\ker \pi = R^*(A)$, so there is an induced map $\sigma: A/R^*(A) \to \mathscr{L}_r(X)$ so that the following diagram commutes:

$$\begin{tikzcd} A \arrow[r, "\pi"] \arrow[d, "\varphi"'] & \mathscr{L}_r(X) \\ A/R^*(A) \arrow[ur, "\sigma"'] & \\ E(A) \arrow[u, phantom, "\cap\!\!\parallel"] & \end{tikzcd}$$

We have shown in Theorem 2.13 that the topologies $\mathscr{T}_r = \mathscr{T}_b = \mathscr{T}_+$ on $\mathscr{L}_r(X)$ are defined by the sequence $\{\|\cdot\|_{M_n}\}$ of seminorms and that $\|T\|_{M_n} = \|T|H_n\|$. Also, we know that $\pi_n(a)$ is $\pi(a)|H_n$. So from Theorem 4.2 we have $|a|_n = \|\pi_n(a)\|$, and hence, $|a|_n = \|\pi(a)\|_{M_n}$. It follows that $\sigma\colon A/R^*(A) \to \mathscr{L}_r(X)$ is topological, and since $(\mathscr{L}_r(X), \mathscr{T}_r)$ is complete σ extends to a topological $*$-isomorphism of $E(A)$ into $\mathscr{L}_r(X)$.

If A is a Banach $*$-algebra with identity and $E(A)$ its enveloping algebra, then every Hilbert space representation of A factors through $E(A)$. We conclude our discussion of enveloping algebras by examing this problem for Fréchet $*$-algebras. We consider only representations in $\mathscr{L}^+(X)$, since this is enough to illustrate the problems involved.

LEMMA 4.5. *If $(A, \{\|\cdot\|_n\})$ is a Fréchet $*$-algebra with identity, $\{|\cdot|_n\}$ the corresponding sequence of B^*-seminorms on A used to define the topology of $E(A)$, and if $\mu\colon A \to \mathscr{L}^+(X)$ is an essential representation of A on X ($\mu(e) = I$) then, for each $M \in \mathscr{S}_+$ there exists $n \in N$ and $C > 0$ such that $\|\mu(a)\|_M \leqq C|a|_n (a \in A)$.*

Proof. Fix $M \in \mathscr{S}_+$. We let $\|M\| = \sup\{\|x\|\colon x \in M\}$ ($\|M\| < \infty$, since M is bounded in the Hilbert space completion of X). Since μ is continuous there exist $n \in N$ and $C > 0$ such that $\|\mu(a)\|_M \leqq C\|a\|_n (a \in A)$.

Fix $x \in M$. Then $f_x\colon a \to (\mu(a)x, x)$ is a positive functional on A. Also, $|f_x(a)| = |(\mu(a)x, x)| \leqq \|\mu(a)\|_M \leqq C\|a\|_n$. Hence, $f_x \in P_n(A)$ for each $x \in M$. Therefore, if $x \in M$ and $x \neq 0$, the positive functional $f_x(e)^{-1}f_x$ belongs to $K_n(A)$ and $f_x(e)^{-1}f_x(a^*a) \leqq |a|_n^2 (a \in A)$. So we have $f_x(a^*a) \leqq f_x(e)|a|_n^2 (x \in M, a \in A)$. But $f_x(e) = (\mu(e)x, x) = \|x\|^2 \leqq \|M\|^2$. Hence, $f_x(a^*a) \leqq \|M\|^2 |a|_n^2 (x \in M, a \in A)$.

For $x, y \in M, a \in A$ we have

$$\begin{aligned} |(\mu(a)x, y)| &= \|\mu(a)x\| \cdot \|y\| \\ &\leqq \|M\| \cdot (\|\mu(a)x\|^2)^{1/2} \\ &\leqq \|M\| \cdot (\mu(a^*a)x, x)^{1/2} \\ &\leqq \|M\|^2 |a|_n . \end{aligned}$$

Thus, $\|\mu(a)\|_M \leqq \|M\|^2 |a|_n (a \in A)$.

THEOREM 4.6. *If $(A, \{\|\cdot\|_n\})$ is a Fréchet *-algebra with identity, $(E(A), (|\cdot|_n\})$ its enveloping algebra, φ the natural map of A into $E(A)$, and if $\mu: A \to \mathscr{L}^+(X)$ is an essential representation of A on X, then there exists a continuous representation σ of $A/R^*(A)$ on X so that $\sigma\varphi = \mu$. If $\pi(A)$ is contained in a $\mathscr{T}_+$-complete subalgebra of $\mathscr{L}^+(X)$, then σ extends to a representation of $E(A)$ on X. In particular, this is the case if X is Hilbert space. Hence, all Hilbert space representations of A factor through $E(A)$.*

Proof. We need only show that σ can be defined on $A/R^*(A)$ so that $\sigma\varphi = \mu$. The other claims follow from Lemma 4.5. It is sufficient to show that $\ker\varphi \subset \ker\mu$. *If $a \in \ker\varphi = R^*(A)$ and if $x \in X$, then $b \to (\mu(b)x, x)$ is a positive functional on A; hence, $(\mu(a^*a)x, x) = \|\mu(a)x\|^2 = 0$. Thus, $\mu(a) = 0$ and $a \in \ker\mu$.*

EXAMPLE 4.7. We show here that some representations $\mu: A \to \mathscr{L}^+(X)$ fail to factor through $E(A)$.

Let $A = C^\infty(\boldsymbol{R})$ with the topology determined by the seminorms $\|a\|_n = \sum_{k=0}^n (k!)^{-1} \|a^{(k)}\|_{n,\infty}$, where $a^{(k)}$ is the kth derivative of a and $\|\cdot\|_{n,\infty}$ is the supremum on $[-n, n]$. Then A is a commutative Fréchet *-algebra with identity (involution is conjugation), and

(1) $|a|_n = \|a\|_{n,\infty} (a \in A, n \in \boldsymbol{N})$

(2) $R^*(A) = 0$, hence

(3) $A/R^*(A) = (A, \{|\cdot|_n\})$ and $E(A)$ is just $C(\boldsymbol{R})$ with the compact-open topology. We use hereafter $|\cdot|_n$ for $\|\cdot\|_{n,\infty}$.

Let $X = C_0^\infty(\boldsymbol{R})$, the compactly-supported C^∞ functions on $\boldsymbol{R}$, considered as a dense subspace of $L^2(\boldsymbol{R})$. We note that if $a \in C(\boldsymbol{R})$ and $f \in C_0^\infty(\boldsymbol{R})$, then there exists $n \in \boldsymbol{N}$ such that $\|af\| \leqq |a|_n \|f\|$ (n depends on f, n is any positive integer so that $\operatorname{Supp}(f) \subseteqq [-n, n]$), and $\|\cdot\|$ is the norm in $L^2(\boldsymbol{R})$.

Define $\mu: A \to \mathscr{L}_a(X)$ by $\mu(a)f = af$. It is clear that this formula actually does define a linear transformation on X and that (1) $\mu(a)^* = \mu(\bar{a})$, (2) μ is a representation of A in $\mathscr{L}^+(X)$, hence, (3) μ; $(A, \{|\cdot|_n\}) \to \mathscr{L}^+(X)$ is continuous (by Theorem 4.6). We now show that μ cannot be extended continuously to $C(\boldsymbol{R})$. We prove (4): if $\tilde{\mu}$ is the extension to $C(\boldsymbol{R})$ of μ and if $f \in X$, then, we must have $\tilde{\mu}(a)f = af$ $(a \in C(\boldsymbol{R}))$. We know that there exists $n \in \boldsymbol{N}$ and $C > 0$ so that $\|\tilde{\mu}(a)f\| \leqq C|a|_n$ for each $a \in C(\boldsymbol{R})$. Fix $a \in C(\boldsymbol{R})$ and choose $\{a_j\} \subseteqq C^\infty(\boldsymbol{R})$ so that $C(\boldsymbol{R}) - \lim_j a_j = a$. Choose $C > 0$ and $n \in \boldsymbol{N}$ as above (for $f \in X$) and such that $\operatorname{Supp}(f) \subseteqq [-n, n]$. Then $\|\tilde{\mu}(a)f - \tilde{\mu}(a_j)f\| \leqq C|a - a_j|_n$. Hence, $\{\tilde{\mu}(a_j)f\}$ converges in $L^2(\boldsymbol{R})$ to $\tilde{\mu}(a)f$. But $\tilde{\mu}(a_j)f = a_jf$ and by our earlier estimate $\|a_jf - af\| \leqq \|f\| |a_j - a|_n$. So

$\tilde{\mu}(a_j)f = a_j f$ converges to $\tilde{\mu}(a)f$ and to af. Thus $\tilde{\mu}(a)f = af$ for each $f \in X$. But $C(\boldsymbol{R}) \cdot C_0^\infty(\boldsymbol{R}) \nsubseteq C_0^\infty(\boldsymbol{R})$, so μ fails to extend to $C(\boldsymbol{R})$.

REMARKS. In the last example we could have considered μ a representation of $C^\infty(\boldsymbol{R})$ in $\mathscr{L}(C_0^\infty(\boldsymbol{R}), \mathscr{T}_b)$. It is not too difficult to show that μ is continuous when thought of this way. It clearly still fails to extend.

Fainally, we do not know whether representations of A in $(\mathscr{L}^*(X), \mathscr{T}_b)$ where X is a locally convex TVS with a continuous inner product are necessarily continuous, in contrast to representations in $\mathscr{L}^+(X)$. It probably is possible to find an example of a discontinuous representation, since the topology $\mathscr{T}_b$ need not be related to the inner product.

REFERENCES

1. R. M. Brooks, *On locally m-convex *-algebras*, Pacific J. Math., **23** (1967), 5-23.
2. Siu Do-Shing, С. *Дошпн, О по-лунормированных кольцах с инволюцией*, Изв. АН СССР, Сер. Мат., **23** (1959), p. 509-528.
3. J. Horváth, *Topological vector spaces and distributions*, vol. I, Addison-Wesley, Reading, Mass., 1966.
4. G. Lasser, *Topological Algebras of Operators*, (Preprint, Joint Institute for Nuclear Research, Dubna, 1969).
5. E. A. Michael, *Locally multiplicatively-convex topological algebras*, Mem. Amer. Math. Soc., **11**, 1952.
6. C. E. Rickart, *General Theory of Banach algebras*. van Nostrand, Princeton, 1960.
7. H. H. Schaefer, *Topological vector spaces*, Macmillan, New York, 1966.

Received September 10, 1970. Research for this paper was supported in part by NSF Grants GP-8346 and GP-18729.

UNIVERSITY OF UTAH

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

EXTENSIONS OF TOPOLOGICAL GROUPS

LAWRENCE G. BROWN

In this paper, we will be concerned with topological group extensions of a polonais group A by a polonais group G. When A is abelian, we will consider two cohomology groups: $H^2(G, A)$ and $H^2_V(G, A)$. $H^2(G, A)$ is based on Borel cochains and was studied by Moore. $H^2_V(G, A)$ is obtained by identifying cochains which differ only on a first category set, an idea suggested to us by D. Wigner. We will show that each of the groups classifies the extensions and that the hypothesis that A be abelian can be eliminated.

By a polonais group, we mean a separable metrizable topological group which is complete in its two-sided uniformity. The completeness requirement is equivalent to topological completeness (see [2], Exercise $Q(d)$, p. 212). If E is a topological group having a polonais normal subgroup A such that E/A is polonais, then it is elementary to prove that E must be polonais also.

If A is abelian, then the cohomology groups are defined in terms of an a priori action of G on A. If A is non-abelian, then we will define sets (not groups) $H^2(G, A)$ without given action of G on A. For brevity, we proceed directly to the general case.

Let $\mathscr{A}$ be the group of topological automorphisms of A. We will write ${}^{\theta}a$ for the action of $\theta \in \mathscr{A}$ on $a \in A$ and I_a for the inner automorphism $b \rightarrow aba^{-1}$. Let e denote the identity of any group. Then $H^2(G, A)$ is defined by means of cocycles (σ, ρ) where:

$$\sigma\colon G \times G \rightarrow A,\ \rho\colon G \rightarrow \mathscr{A}, \tag{1}$$

σ is a Borel function on $G \times G$ and $(x, a) \rightarrow {}^{\rho(x)}a$ is a Borel function on $G \times A$.

$$\begin{aligned} \sigma(x, y)\cdot\sigma(xy, z) &= {}^{\rho(x)}\sigma(y, z)\cdot\sigma(x, yz) \ , \\ \rho(x)\cdot\rho(y) &= I_{\sigma(x,y)}\rho(x, y) \ , \\ \sigma(x, e) &= \sigma(e, y) = e, \text{ and} \\ \rho(e) &= e \ . \end{aligned} \tag{2}$$

(σ, ρ) and (σ', ρ') are identified in $H^2(G, A)$ if there is a Borel function $\lambda\colon G \rightarrow A$ such that:

$$\begin{aligned} \sigma'(x, y) &= \lambda(x)\cdot{}^{\rho(x)}\lambda(y)\cdot\sigma(x, y)\cdot\lambda(xy)^{-1} \ , \\ \rho'(x) &= I_{\lambda(x)}\cdot\rho(x) \ . \end{aligned} \tag{3}$$

$H^2_V(G, A)$ can be defined by simply stipulating that (2) and (3) hold only on the complement of a first category set. However, for purposes of motivation, we prefer a slightly different description.

If X is a complete metric space and Y a metric space, then (following Wigner) by a *virus* from X to Y, we mean an equivalence class of continuous functions $f: R \to Y$ where R is a dense G_δ set in X. f and f' are equivalent if they agree on $R \cap R'$. (We could eliminate the equivalence relation by using maximal continuous extensions.) It is clear that if Y is separable, the viruses amount to Borel functions modulo first category sets. The concept of virus will be used later. In the abelian case, it could be incorporated directly into the definition of $H^2_V(G, A)$.

$H^2_V(G, A)$ is defined by means of cocycles (σ, ρ) where:

(1′) $\sigma: R_1 \to A$, $\rho: R_2 \to \mathscr{A}$ where R_1 and R_2 are dense G_δ's in $G \times G$ and G. σ is continuous on R_1, and $(x, a) \to {}^{\rho(x)}a$ is continuous on $R_2 \times A$.

(2′) $\sigma(x, y)\sigma(xy, z) = {}^{\rho(x)}\sigma(y, z)\sigma(x, yz)$ and $\rho(x)\rho(y) = I_{\sigma(x,y)}\rho(xy)$ whenever everything is defined. (σ, ρ) and (σ', ρ') are identified in $H^2_V(G, A)$ if there is a continuous function $\lambda: R \to A$ (R a dense G_δ in G) such that:

(3′) (3) holds whenever everything is defined. (In particular, (σ, ρ) is identified with (σ', ρ') if they agree off a first category set.)

We need two technical results:

LEMMA 1 (Dixmier [1]). *If E is a polonais group, A a closed normal subgroup, and $p: E \to E/A$ the projection, then there is a Borel function $f: E/A \to E$ such that $p(f(x)) = x$ for all $x \in E/A$.*

LEMMA 2 (Wigner). *If E is a metrizable group complete in its two-sided uniformity, A a closed normal subgroup, and $p: E \to E/A$ the projection, then there is a dense G_δ, $R \subset E/A$ and a continuous function $f: R \to E$ such that $p(f(x)) = x$ for all $x \in R$.*

Proof. (One may note first that Lemma 2 follows from Lemma 1 in the polonais case.) Let $U_1, U_2 \cdots$ be a fundamental system of symmetric neighborhoods of e in G. Define recursively subsets O_n of G such that O_n is maximal with respect to:

(4) O_n is open; $O_n \subset O_{n-1}$; and if $x, y \in O_n$ and $p(x) = p(y)$, then $x^{-1}y, yx^{-1} \in U_n$. ($O_0 = E$.)

It is not hard to see that $p(O_n)$ is dense in E/A:

If $V \subset O_{n-1}$ is open, VV^{-1} and $V^{-1}V \subset U_n$, and $p(V)$ is disjoint from $p(O_n)$, then $O_n \cup V$ satisfies (4). If $R = \cap p(O_n)$, then it is not hard to prove that R is dense (i.e., that E/A is second category in itself. Actually, E/A is even complete in its two-sided uniformity.) For $x \in R$, let $y_n \in O_n$ be such that $p(y_n) = x$. Then (y_n) is Cauchy and hence $y_n \to y$, for some $y \in G$. y can easily be seen to be independent of the choices of y_n, and if we define $f(x) = y$, it is straightforward to prove f is continuous.

Now we can define a map $\pi\colon \mathrm{Ext}\,(G, A) \to H^2(G, A)$. ($\mathrm{Ext}\,(G, A)$ is the set of equivalence classes of topological group extensions of A by G.) If E is a given extension, let f be as in Lemma 1 such that $f(e) = e$.

Define (σ, ρ) by:

$$f(x)\cdot f(y) = \sigma(x, y)f(xy);\ {}^{\rho(x)}a = f(x)\cdot a\cdot f(x)^{-1}\,. \tag{5}$$

It is easy to see that $\pi(E) = [(\sigma, \rho)]$ gives a well defined function π where $[(\sigma, \rho)]$ denotes the class of (σ, ρ).

There is a natural map $j\colon H^2(G, A) \to H^2_V(G, A)$. If (σ, ρ) satisfies (1) and (2), then there are dense G_δ's $R_1 \subset G \times G$ and $R_3 \subset G \times A$ such that the restriction of σ to R_1 and the restriction of $(x, a) \to {}^{\rho(x)}a$ to R_3 are continuous.

LEMMA 3. *If $R_2 = \{x \in G\colon \{a\colon (x, a) \in R_3\}$ is a dense G_δ in $A\}$, then R_2 is a dense G_δ in G and $(x, a) \to {}^{\rho(x)}a$ is continuous on $R_2 \times A$.*

Proof. That R_2 is a dense G_δ is clear. Now suppose $x_n \to x$ and $a_n \to a (x_n, x \in R_2)$. We can find $b \in A$ such that:

$$(x_n, b) \in R_3,\ (x, b) \in R_3,\ (x_n, b^{-1}a_n) \in R_3,\ (x, b^{-1}a) \in R_3\,.$$

Then ${}^{\rho(x_n)}a_n = {}^{\rho(x_n)}b\cdot{}^{\rho(x_n)}b^{-1}a_n \to {}^{\rho(x)}b\cdot{}^{\rho(x)}b^{-1}a = {}^{\rho(x)}a$.

Now let $\bar{\sigma}$ and $\bar{\rho}$ be the restrictions of σ and ρ to R_1 and R_2, and define $j[(\sigma, \rho)] = [(\bar{\sigma}, \bar{\rho})]$. It should be clear that j is a well-defined map on $H^2(G, A)$.

The map $j\pi\colon \mathrm{Ext}\,(G, A) \to H^2_V(G, A)$ can be described a little more simply. If f and R satisfy the conclusion of Lemma 2, then we can use (5) to define σ, ρ, R_1 and R_2. Note that $R_2 = R$ and $R_1 = \widetilde{R} = \{(x, y) \in G \times G\colon x, y, xy \in R\}$.

We can now state:

THEOREM. *If G and A are polonais groups, then π and $j\pi$ are bijections.*

Proof. The main part of the proof is that $j\pi$ is surjective, and we prove that first. Let σ, ρ, R_1 and R_2 be given. We first reduce to the case where R_1 and R_2 are as in the previous paragraph. Let $R = \{x \in R_2\colon \{y\colon (x, y) \in R_1\}$ is a dense G_δ in $G\}$. Then for $(x, y) \in \widetilde{R}$, define:

(6) $\Sigma(x, y)(z) = {}^{\rho(x)}\sigma(y, z) \cdot \sigma(x, yz) \cdot \sigma(xy, z)^{-1}$ for all z such that this makes sense.

Then Σ induces a function from $\tilde{R}$ into the set of viruses from G to A. For $(x, y) \in \tilde{R} \cap R_1$, it is easy to see from (2′) that $\Sigma(x, y)$ is the constant virus $\sigma(x, y)$. Since $\tilde{R} \cap R_1$ is dense in $\tilde{R}$ and Σ is continuous, it is now clear that $\Sigma(x, y)$ is a constant virus, $\tilde{\sigma}(x, y)$, for all $(x, y) \in \tilde{R}$, and $\tilde{\sigma}$ is continuous. We now replace σ by $\tilde{\sigma}$ (and R_2 by R) to obtain the desired situation. We will also assume that $e \notin R$.

We next extend (σ, ρ) to a cocycle defined everywhere, following Moore [7]. To this end, let V be the group of all viruses from G to A. For $x \in R, a \in A$, define $\theta(x)\colon V \to V$ and $J_a\colon V \to V$ by:

$$(7) \qquad \begin{aligned} [\theta(x)g](y) &= \sigma(y, x)g(yx) \text{ and} \\ (J_a g)(x) &= {}^{\rho(x)}a \cdot g(x) \text{ for } g \in V . \end{aligned}$$

It is a straightforward computation to verify:

$$(8) \qquad \begin{aligned} &\theta(x) \cdot \theta(y) = J_{\sigma(x\ y)} \cdot \theta(xy),\ (x, y, xy \in R),\ J_{ab} = J_a \cdot J_b, \text{ and} \\ &\theta(x)^{-1} \text{ exists and } \theta(x) J_a \theta(x)^{-1} = J_{\rho(x)_a} . \end{aligned}$$

Now let E' be the group generated by $J(A)$ and $\theta(R)$. We can see that if A is identified with $J(A)$, then A is normal in E', and each element of E' induces a continuous automorphism of A. Let $p\colon E' \to E'/A$ be the projection. Clearly, for $(x, y) \in \tilde{R}$, $p\theta(xy) = p\theta(x) \cdot p\theta(y)$. It is now easy to see that $p\theta$ extends to a homomorphism $\bar{\theta}\colon G \to E'/A$. Let $f\colon G \to E'$ be an extension of θ such that $f(e) = e$ and $pf = \bar{\theta}$. Then we can extend (σ, ρ) by:

$$(9) \qquad \begin{aligned} f(x) \cdot f(y) &= J_{\sigma(x\ y)} \cdot f(xy), \text{ and} \\ J_{\rho(x)_a} &= f(x) \cdot J_a \cdot f(x)^{-1} . \end{aligned}$$

It is clear that the extended (σ, ρ) satisfies the cocycle relations (2) (though not (1)).

We will later have occasion to use a uniqueness result for the extended (σ, ρ). Thus let (σ', ρ') satisfy (2) everywhere, suppose σ' agrees with σ on a dense G_δ, R_1, and ρ' agrees with ρ on a dense G_δ, R_2. Define $R' = \{y \in R \cap R_2\colon \{x\colon (x, y) \in R_1\}$ is a dense G_δ in $G\}$. We can see that for any $v \in G$, there is a dense G_δ, R_v, such that $\sigma'(u, v)$ is continuous in u on R_v. This follows from the following consequence of (2):

$$(10) \qquad \sigma'(u, yz) = {}^{\rho'(u)}\sigma'(y, z)^{-1} \cdot \sigma'(u, y) \cdot \sigma'(uy, z)$$

where we choose $y, z \in R'$ such that $yz = v$.

Thus we can use (7) to define θ' on all of G, agreeing with θ on R', and θ' will satisfy (8) (with (σ', ρ') instead of (σ, ρ)). If we define λ by $\theta'(x) = \lambda(x)f(x)$, then we see that (3) is satisfied and λ vanishes on R'. This is the desired uniqueness result.

Using the extended (σ, ρ), we can now construct a (non-topological) extension E of A by G and a function $f: G \to E$ such that if $p: E \to G$ is the projection, then $pf(x) = x, f(e) = e$, and (5) is satisfied. We must topologize E, and we first define sequential convergence to e. If $\alpha_n \in E$, we say that $\alpha_n \to e$ if there exists $\beta = b \cdot f(y) (y \in R)$ such that if $\beta\alpha_n = b_n \cdot f(y_n)$, then $y_n \in R$, $b_n \to b$, and $y_n \to y$. If $\alpha_n = a_n \cdot f(x_n)$, then it is readily seen that this condition is equivalent to:

$$yx_n \in R,\ x_n \to e, \text{ and } {}^{\rho(y)}a_n \cdot \sigma(y, x_n) \to e\ . \tag{11}$$

The condition is thus independent of b, and we now show it is independent of y (subject to $y, yx_n \in R$). We need:

LEMMA 4. (a) *For* $x \in G, \sigma(x, y)$ *is continuous in* y *on* $R_x = \{y: y, xy \in R\}$.

(b) $\sigma(x, y)$ *is continuous in* x *on* $R_y = \{x: x, xy \in R\}$.

Proof. (a) Let $y_n \to y$ in R_x. Consider:

$$\sigma(x, y_n) = \sigma(v, w)^{-1} \cdot {}^{\rho(v)}\sigma(w, y_n) \cdot \sigma(v, wy_n) \tag{12}$$

where $v = xw^{-1}$ and $w, xw^{-1}, wy_n, wy \in R$.

Then from the continuity of σ on $\tilde{R}$, we see that $\sigma(x, y_n) \to \sigma(x, y)$. (b) is proved similarly.

Now if $y = uy'$ where $y, yx_n, y', y'x_n \in R$, we find:

$${}^{\rho(y)}a_n \cdot \sigma(y, x_n) = \sigma(u, y')^{-1\rho(u)}[{}^{\rho(y')}a_n \cdot \sigma(y', x_n)] \cdot \sigma(u, y'x_n)\ . \tag{13}$$

Using Lemma 4(a), we see that (11) is satisfied for y if it is for y'.

LEMMA 5. *If* $\alpha_n \to e$, *then* $\beta\alpha_n\beta^{-1} \to e$ *for any* $\beta \in E$.

Proof. First, let $\beta = b \in A$ where $\alpha_n = a_n f(x_n)$. If $y, yx_n \in R$ and $f(y)\beta\alpha_n\beta^{-1} = c_n \cdot f(x_n)$, then:

$$c_n = {}^{\rho(y)}b \cdot [{}^{\rho(y)}a_n \cdot \sigma(y, x_n)] \cdot {}^{\rho(yx_n)}b^{-1}\ . \tag{14}$$

Hence (11) is satisfied for $\beta\alpha_n\beta^{-1}$ if satisfied for α_n. Now let $\beta = f(z)$. Then if $y, yzx_nz^{-1}, yz, yzx_n \in R$, and $f(y)\beta\alpha_n\beta^{-1} = d_n f(yzx_nz^{-1})$,

$$d_n = \sigma(y, z) \cdot [\rho^{(yz)}a_n \cdot \sigma(yz, x_n)] \cdot \sigma(yzx_nz^{-1}, z)^{-1}\ . \tag{15}$$

Hence (by Lemma 4 (b)), (11) is satisfied for $\beta\alpha_n\beta^{-1}$ if satisfied for α_n.

From this lemma, it is easy to see that the definition of $\alpha_n \to e$ could equivalently have been given in terms of $\alpha_n\beta$ instead of $\beta\alpha_n$. The criterion would then have been:

(11′) If $w, x_n w \in R$, then $a_n \cdot \sigma(x_n, w) \to e$ and $x_n \to e$. $(\alpha_n = a_n \cdot f(x_n))$.

LEMMA 6. *If $\alpha_n \to e$ and $\beta_n \to e$, then $\alpha_n \cdot \beta_n \to e$.*

Proof. Let $\alpha_n = a_n \cdot f(x_n)$, $\beta_n = b_n \cdot f(y_n)$. First, choose z such that z, zx_n, and $zx_n y_n \in R$. Then choose w such that $w, y_n w$, and $zx_n y_n w \in R$. Then if $f(z) \cdot \alpha_n \cdot \beta_n = c_n f(zx_n y_n)$, we find:

$$c_n = [{}^{\rho(z)}a_n \cdot \sigma(z, x_n)] \cdot {}^{\rho(zx_n)}[b_n \cdot \sigma(y_n, w)] \cdot \sigma(zx_n, y_n w) \cdot \sigma(zx_n, y_n, w)^{-1} . \tag{16}$$

Hence (11), for $\alpha_n\beta_n$ follows from (11) for α_n and (11′) for β_n.

LEMMA 7. *If $\alpha_n \to e$, $\alpha_n^{-1} \to e$.*

Proof. Let $\alpha_n = a_n \cdot f(x_n)$. Choose z such that $z, zx_n^{-1} \in R$. Then choose w such that $w, zw, x_n w \in R$. Then if $f(z) \cdot \alpha_n^{-1} = c_n f(zx_n^{-1})$, we find

$$c_n^{-1} = {}^{\rho(zx_n^{-1})}[a_n \cdot \sigma(x_n, w)] \cdot \sigma(zx_n^{-1}, x_n w) \cdot \sigma(z, w)^{-1} . \tag{17}$$

Thus (11), for α_n^{-1} follows from (11′) for α_n.

We now must show that there is a metrizable group-topology on E such that convergence as defined above is convergence in the topology. To do this, it is sufficient to find a sequence W_m such that $e \in W_m \subset E$, and $\alpha_n \to e$ as defined above if and only if α_n is eventually in W_m, for each m. Let U_m be a fundamental system of neighborhoods of e in A and V_m a fundamental system of neighborhoods relative to R of $u_0 \in R$. Then we define:

$$W_m = [U_m f(V_m)]^{-1} \cdot U_m f(V_m) . \tag{18}$$

Suppose that α_n is eventually in each W_m. Then $\alpha_n = \beta_n^{-1}\gamma_n$ where β_n, γ_n are eventually in each $U_m f(V_m)$. Hence $f(u_0)^{-1}\beta_n \to e$ and $f(u_0)^{-1}\gamma_n \to e$. Thus $\alpha_n = [f(u_0)^{-1}\beta_n]^{-1} \cdot [f(u_0)^{-1}\gamma_n] \to e$. Now assume $\alpha_n = a_n f(x_n) \to e$. Since $V_m^{-1} V_m$ is a neighborhood of e in G, we can find $y_n, z_n \in R$ such that $y_n \to u_0$, $z_n \to u_0$ and $y_n^{-1} z_n = x_n$. Now define $\beta_n = f(y_n)$ and $\gamma_n = c_n f(z_n)$ where c_n is chosen such that $\beta_n^{-1}\gamma_n = \alpha_n$. Then calculation shows:

$$c_n = {}^{\rho(y_n)}[a_n \cdot \sigma(x_n, w)] \cdot \sigma(y_n, x_n w) \cdot \sigma(z_n, w)^{-1} \tag{19}$$

where we assume w chosen so that $w, x_n w, z_n w$ and $u_0 w \in R$. Then (11′) shows that $c_n \to e$. Hence β_n, γ_n are eventually in each $U_m f(V_m)$, as desired.

Now that we have a metrizable topology on E, it is easy to see that the subspace topology on A is the correct one. It is also not hard to see that the quotient topology on G is the correct one. Indeed, all we must show is that if $x_n \to e$ in G, then we can find $\alpha_n \to e$ in E such that $p(\alpha_n) = x_n$. To do this, choose y_n and z_n as in the preceding paragraph. Then $\alpha_n = f(y_n)^{-1}f(z_n)$. From the fact that A and G are complete and separable, it now follows easily that E is complete and separable and hence polonais. It is clear that f is continuous on R. Hence $j\pi(E) = [(\sigma, \rho)]$, and we have proved that $j\pi$ is surjective.

To show that $j\pi$ is $1-1$, assume $j\pi(E) = j\pi(E')$. Then we can find a dense $G_\delta R \subset G$ and continuous $f: R \to E$, $f': R \to E'$ such that $pf(x) = x = p'f'(x)$, and $(\sigma, \rho) = (\sigma', \rho')$ where σ, ρ, σ', and ρ' are defined by (5). Then we can define a function $\varphi: p^{-1}(R) \to p'^{-1}(R)$ by:

$$\varphi(a \cdot f(x)) = a \cdot f'(x) \;, \tag{20}$$

and clearly $\varphi(\alpha\beta) = \varphi(\alpha)\cdot\varphi(\beta)$ for $\alpha, \beta, \alpha\beta \in p^{-1}(R)$. From this, it is not hard to prove that φ can be extended to a homomorphism of E onto E' and that this homomorphism is a topological isomorphism.

It is now clear that π is $1-1$. To show that π is surjective, let (σ, ρ) satisfy (1) and (2). Let E be such that $j\pi(E) = j[(\sigma, \rho)]$. Then we can find a Borel function $f: G \to E$ such that $pf(x) = x$, $f(e) = e$, and if (σ', ρ') is defined by (5) (with (σ', ρ') substituted for (σ, ρ)), then (σ, ρ) agrees with (σ', ρ') except on a first category set. Now the uniqueness result proved above shows that there is a function $\lambda: G \to A$, which vanishes on a dense G_δ, R', such that (3) holds. We must show that λ is Borel. (3) implies:

$$\lambda(xy) = \sigma'(x, y)^{-1}\cdot\lambda(x)\cdot{}^{\rho(x)}\lambda(y)\cdot\sigma(x, y) = \sigma'(x, y)^{-1}\cdot\sigma(x, y) \;, \tag{21}$$

for $(x, y) \in R' \times R'$. Thus if $m: R' \times R' \to G$ is the group operation, then $\lambda \circ m$ is Borel. Since m is surjective, it follows from well-known results on Borel sets (see Kuratowski [3] or the first few pages of Mackey [4]) that λ is Borel. The theorem is now proved.

We make two final remarks:

1. If A is abelian and we are given an action of G on A, then it is easy to see that the bijections π, j, and $j\pi$ preserve the group operations (where the group operation on Ext (G, A) is the usual Baer product).

2. The hypothesis that A be separable can be dropped by considering A as a direct limit of closed separable subgroups (G-subgroups

in the abelian case).

REFERENCES

1. J. Dixmier, *Dual et quasi-dual d'une algebre de Banach involutive,* Trans. Amer. Math. Soc., **104** (1962), 278-283.

2. J. Kelley, *General Topology,* Van Nostrand, Princeton, 1955.

3. K. Kuratowski, *Topology Vol.* I (translated from the French by J. Jawarowski), Academic Press, New York, 1966.

4. G. Mackey, *Borel structures in groups and their duals,* Trans. Amer. Math. Soc., **85** (1957), 134-165,

5. C. Moore, *Extensions and low dimensional cohomology theory of locally compact groups I,* Trans. Amer. Math. Soc., **113** (1964), 40-63.

6. ———, *Extensions and low dimensional cohomology theory of locally compact groups II,* Trans. Amer. Math. Soc., **113** (1964), 64-86.

7. ———, *Group extensions of p-adic and adelic linear groups,* I.H.E.S. Publ. Math., **35** (1968), 5-70.

Received November 2, 1970.

STANFORD UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

REVERSIBLE HOMEOMORPHISMS OF THE REAL LINE

ARNOLD B. CALICA

Let G be the group of germs of C^k local homeomorphisms of the real line which fix the origin and have nonzero derivative there. In this paper the possibility of factoring an element of G which is conjugate to its inverse into the product of two involutions is investigated. It is shown that it is always possible to do this in the analytic case and not always possible in the continuous case. In the intermediate cases several necessary and sufficient conditions are developed for determining whether or not such a factorization is possible. Included is a construction which allows one to determine an explicit factorization. Indication is given of the application of this material to the same problem in higher dimensions. This work is related to some material in Dynamics.

1. Introduction. If G is an abstract group an element $g \in G$ is called *reversible* in G if there exists an element $h \in G$ such that $hgh^{-1} = g^{-1}$. The product of two involutions is always reversible by an obvious argument. There arises the following question:

Question $\sharp$. If g is reversible in G can g be factored into the product of two involutions in G?

D. C. Lewis has decided this issue in the case $G = GL(n, \boldsymbol{C})$ affirmatively (Lewis [4]).

This paper concerns itself with the investigation of this question in the case where G is the group of germs of continuous or differentiable homeomorphisms of the real line.

Reversible transformations play a role in Dynamics. For further information on this connection see the references in Lewis [4].

2. Definitions and Notation. $C^k = \{F\colon F$ is a local homeomorphism of a neighborhood of 0 in $\boldsymbol{R}$ to another such neighborhood which fixes the origin and is of class C^k on some neighborhood of 0, $F'(0) \neq 0$ if $k > 0$, for $0 \leqq k \leqq \infty$ or analytic for $k = \omega\}$. $T^k =$ {germs of elements of $\boldsymbol{C}^k$}

Let $\phi_k\colon \boldsymbol{C}^k \rightarrow \boldsymbol{T}^k$ be the map which assigns to each element of $\boldsymbol{C}^k$ its germ in $\boldsymbol{T}^k$. The binary operation of (local) composition of mappings in $\boldsymbol{C}^k$ induces a group multiplication in $\boldsymbol{T}^k$. $\boldsymbol{T}^k$ will be viewed as a group with this structure henceforward.

If a property is locally true for each $\alpha \in f$, $\alpha \in \boldsymbol{C}^k$, $f \in \boldsymbol{T}^k$ the property will be attributed to f. The identity element of any group under consideration will be called 1.

$\boldsymbol{P}^k = \{$set of all real power series of the form $\sum_{r=1}^{k} a_r x^r$, $a_1 \neq 0$, $1 \leqq k \leqq \infty\}$

$\boldsymbol{P}^\omega = \{\alpha \in \boldsymbol{P}^\infty$: α has a nonzero radius of convergence$\}$

$\boldsymbol{P}^0 = \{1, -1\}$.

Let ρ_k: $\boldsymbol{T}^k \to \boldsymbol{P}^k$, $k > 0$ be the mapping which assigns to each element of $\boldsymbol{T}^k$ its Taylor expansion to degree k about the origin. Put $\rho_0(f) = \pm 1$ according as f is the germ of a locally increasing or locally decreasing element of $\boldsymbol{C}^0$.

REMARK 1. If $0 \leqq k \leqq \infty$ or $k = \omega$, ρ_k is onto. This trivial for $k < \infty$ familiar for $k = \omega$ and true for $k = \infty$. For the last case see Borel [2].

The multiplication in $\boldsymbol{T}^k$ induces via the mapping ρ_k a group multiplication in $\boldsymbol{P}^k$. $\boldsymbol{P}^k$ will be viewed as a group with this multiplication. The elements of $\boldsymbol{P}^k$ are often referred to as jets in the literature. Note that composition of elements in $\boldsymbol{P}^k$ is not multiplication of polynomials but is substitution followed by truncation to degree k. Other homomorphisms which will be found useful are the mappings $\boldsymbol{T}^k \to \boldsymbol{T}^m$ and $\boldsymbol{P}^k \to \boldsymbol{P}^m$, $m \leqq k$ defined by the inclusion mapping of $\boldsymbol{C}^k \to \boldsymbol{C}^m$ followed by ϕ_m or $\rho_m \phi_m$.

For $F \in \boldsymbol{C}^k$, let F_+ (resp., F) be the restriction of F to $R_+ = \{x \in \boldsymbol{R}$: $x \geqq 0\}$ (resp., to $\boldsymbol{R}_- = \{x \in \boldsymbol{R}$: $x \leqq 0\}$). F_+(resp., F_-) is a local homeomorphism fixing 0 of $\boldsymbol{R}_+$ to $\boldsymbol{R}_+$ or $\boldsymbol{R}_+$ to $\boldsymbol{R}_-$(resp., $\boldsymbol{R}_-$ to $\boldsymbol{R}_-$ or $\boldsymbol{R}_-$ to $\boldsymbol{R}_+$). The notation f_+ and f_- will denote the corresponding notion for the germ f. f_+^{-1} (resp., f_-^{-1}) will denote $(f_+)^{-1}$(resp.,$(f_-)^{-1}$).

3. Periodic local homeomorphisms of $\boldsymbol{R}$.

Let

$\boldsymbol{T}_+^k =$ the set of elements of $\boldsymbol{T}_-^k$ which are increasing,

$\boldsymbol{P}_+^k = \rho_k(\boldsymbol{T}_+^k)$

$\boldsymbol{T}_-^k =$ the set of elements of $\boldsymbol{T}^k$ which are decreasing,

$\boldsymbol{P}_-^k = \rho_k(\boldsymbol{T}_-^k)$

REMARK 2.

(i) $\boldsymbol{T}_+^k \cap \boldsymbol{T}_-^k = \varnothing$, $\boldsymbol{T}_+^k \cup \boldsymbol{T}_-^k = \boldsymbol{T}^k$

(ii) $\boldsymbol{T}_+^k$ is a subgroup of $\boldsymbol{T}^k$ which is of index 2 and is therefore normal.

(iii) $\boldsymbol{T}_+^k \boldsymbol{T}_-^k = \boldsymbol{T}_-^k \boldsymbol{T}_+^k = \boldsymbol{T}_-^k$, $\boldsymbol{T}_-^k \boldsymbol{T}_-^k = \boldsymbol{T}_+^k$

The corresponding statements are true of $\boldsymbol{P}_+^k$ and $\boldsymbol{P}_-^k$.

Repeatedly used subsequently are following obvious facts:

If $f \in \boldsymbol{T}_+^k$, then $(f^{-1})_+ = f_+^{-1}$ and $(f^{-1})_- = f_-^{-1}$.

If $f \in \boldsymbol{T}_-^k$, then $(f^{-1})_+ = f_-^{-1}$ and $(f^{-1})_- = f_+^{-1}$.

PROPOSITION 1. *If* $g \in \boldsymbol{T}_+^k$, $g^m = 1$ *for some integer* $m \neq 0$, *then* $g = 1$.

COROLLARY. *If* $g \in \boldsymbol{T}^k$ *and* $g^m = 1$ *for some integer* $m \neq 0$, *then* $g^2 = 1$.

Proof of Proposition 1. Let $G \in \boldsymbol{C}^k$ such that $\phi_k(G) = g$ and $g^m = 1$. G is clearly monotone near 0. Thus for small $x > 0$, if $x < G(x)$, then $x < G(x) < \cdots < G^m(x) = x$. This is impossible. If $x > G(x)$ a similar argument applies. This constitutes a proof of this proposition.

It is to be noted that Proposition 1 follows trivially from a theorem in Bochner [1] in the case $k > 0$.

REMARK 2. $f \in \boldsymbol{T}_-^k$ and $f^2 = 1$, if and only if $f_- = f_+^{-1}$.

4. Factorization of Reversible Transformations of the Real line. In this section the possibility of factoring a reversible element of $\boldsymbol{T}^k$ is investigated.

Df: If $f \in \boldsymbol{T}_-^k$, then $\bar{f}$ is the element of $\boldsymbol{T}^0$ such that,

$$(\bar{f})_+ = f_+ \text{ and } (\bar{f})_- = f_+^{-1} .$$

LEMMA 1. *If* $f \in \boldsymbol{T}_-^k$ *and* $\rho_k(f)$ *is an involution in* $\boldsymbol{P}^k$, *then* $\bar{f}$ *is an involution in* $\boldsymbol{T}^k$.

Proof. $(f^{-1})_+ = f_-^{-1}$ and $(f^{-1})_- = f_+^{-1}$ since $f \in \boldsymbol{T}_-^k$. $\rho_k(f^2) = 1$ implies $\rho_k(f^{-1}) = \rho_k(f)$. Therefore the right derivatives of f_+ at zero are the same as the left derivatives of f_+^{-1} at zero. Therefore, $\bar{f} \in \boldsymbol{T}^k$ and $\bar{f}^2 = 1$ by Remark 2.

THEOREM 1. *If* $f \in \boldsymbol{T}_+^k (0 \leqq k \leqq \infty$ *or* $k = \omega)$, *then* f *is the product of two involutions in* $\boldsymbol{T}^m$, $(m \leqq k)$ *if and only if there exists* $g \in \boldsymbol{T}_-^m$ *such that* $gfg^{-1} = f^{-1}$ *and* $\rho_m(g)$ *is an involution in* $\boldsymbol{P}^m$.

Proof of Theorem 1. If $f = hk$, h, $k \in \boldsymbol{T}^m$ and $h^2 = k^2 = 1$, then by Remark 2 and Proposition 1 and its corollary one can conclude that either $h, k \in \boldsymbol{T}^m$ or $f = 1$. If $h, k \in \boldsymbol{T}_-^m$, set $g = h$. If $f = 1$ set g equal to any involution in $\boldsymbol{T}_-^m$. In either case $\rho_m(g)$ is also an involution.

Assume now that $f^{-1} = gfg^{-1}$ where $g \in \boldsymbol{T}_-^m$ and $\rho_m(g)$ is an involution. If one sets $(\bar{g})_+ = g_+$, $(\bar{g})_- = g_+^{-1}$ it is then easy to verify that $gfg^{-1} = f^{-1}$ implies $\bar{g}f\bar{g}^{-1} = f^{-1}$. This means that $f = \bar{g}(\bar{g}f)$ and both factors are involutions.

Some of the preceding material can be utilized to demonstrate the impossibility of factoring each reversible element of T^0 into the product of two involutions. This is embodied in the following proposition:

PROPOSITION 3. *There exists a reversible element of T^0 which* cannot *be factored into the product of two involutions.*

Proof of Proposition 3. Suppose there exists an element $f \in T^0_+$ such that f is reversible and $f_- = 1$ and $f_+ \neq 1$. It is clear that $gfg^{-1} = f^{-1}$ implies that $g \in T^0_+$. If such an element exists it cannot be factored as the product of two involutions by Theorem 1. An element of this type is constructed below:

Let $f = \phi_0(F)$ where $F_- = x$, $F \in C^0$

$$\text{and } F_+ = \begin{cases} F_0 & x \in (1/2, 1] \\ F_1 & x \in (1/4, 1/2] \\ \vdots & \\ F_k & x \in (2^{-k-1}, 2^{-k}] \\ \vdots & \vdots \end{cases}$$

and

$$F_{k+1} = \hat{F}_k \text{ and } \hat{F}_k(x) = (1/2\, F_k(2X))^{-1}$$

and

$$F_0(x) = \begin{cases} x + \dfrac{1}{M} \exp\left[(x-1)^{-1} - (x-1/2)^{-1}\right] & x \in (1/2, 1) \\ 1 & x = 1 \end{cases}$$

with M large enough to ensure that F_0 is one to one on its domain of definition. It can be easily verified that F is continuous on $(-\infty, 1]$. Moreover $2F(1/2\, x) = F^{-1}(x)$. Therefore $f = \phi_0(F)$ is reversible and not the product of two involutions. This completes the proof of the proposition.

REMARK 3. Composition in the group P^n is given as follows:
If

$$\alpha = \sum_{k=1}^{n} a_k x^k \quad \text{and} \quad \beta = \sum_{k=1}^{n} b_k x^k ,$$

then

$$\beta\alpha = \sum_{k=1}^{n} A_k x^k ,$$

where

$$A_k = \sum \frac{r!}{r_1!\ r_2! \cdots r_k!} b_r a_1^{r_1} a_2^{r_2} \cdots a_k^{r_k} ,$$
$$r_1 + 2r_2 + \cdots + kr_k = k$$

and

$$r = r_1 + r_2 + \cdots + r_k .$$

LEMMA 2. *Suppose* $\alpha \in \boldsymbol{P}^k$, $m \leqq k$

(a) *If* $\alpha = x + a_n x^m + o(x^m)$,

then

$$\alpha^{-1} = x - a_m x^m + o(x^m) .$$

(b) *If* $\alpha = -x + a_m x^m + o(x^m)$,

then

$$\alpha^{-1} = -x + (-1)^m a_m x^m + o(x^m) .$$

LEMMA 3. *Suppose* $\alpha \in \boldsymbol{P}^k$, α *is reversible,* $m \leqq k$ *and* $\alpha = x + a_m x^m + o(x^m)$, $a_m \neq 0$.

If $\beta\alpha\beta^{-1} = \alpha^{-1}$ and $\beta = \sum_{s=1}^{k} b_s x^s$, then $b_1 = -1$. Moreover, m must be even.

LEMMA 4. *If* $\alpha \in \boldsymbol{P}^2$ *and* $\alpha = -x + a_2 x^2$, *then* α *is an involution.*

The verification of Lemmas 2–4 is straightforward and will therefore be omitted.

Df: If $f \in \boldsymbol{T}^k$, $k > 0$, $m \leqq k$, let

$$f^{(m)}(0) = \left.\frac{d^m F(x)}{dx^m}\right|_{x=0} \qquad \text{for any } F \in \boldsymbol{C}^k$$

such that $\phi_k(F) = f$.

THEOREM 2. *If* $f \in \boldsymbol{T}_+^k$, $k \geqq 2$, f *reversible and* $\rho_k(f) \neq 1$ *then* f *can be factored as the product of two involutions in* $\boldsymbol{T}^2$.

Proof of Theorem. Let $f^{-1} = gfg^{-1}$; Lemma 3 and Lemma 4 show that $\rho_2(g)$ is an involution; now Theorem 1 applies to the image of f in $\boldsymbol{T}^2$.

COROLLARY. *If $f \in \boldsymbol{T}^\omega_+$ and f reversible in $\boldsymbol{T}^\omega$, then f can be factored as the product of two involutions in $\boldsymbol{T}^2$.*

This follows from Theorem 2 and the observation that if $f \in \boldsymbol{T}^\omega_+$ and $f^{(m)}(0) = 0$ for $m > 0$, then $f = 1$.

The question of factorization of reversible elements of $\boldsymbol{T}^\omega_-$ is settled positively in the following material.

LEMMA 5. *If $g, h \in \boldsymbol{T}^\omega_+$ and $g^2 = h^2$, then $g = h$.*

Proof of Lemma 5. *If* $g-h$ has zeros in every neighborhood of 0, then $g = h$ since g and h are analytic. Assume therefore that there exists a neighborhood $(0, \varepsilon]$, $\varepsilon > 0$, such that $g(x) > h(x)$ and $g(x)$, $h(x)$ are monotone in $(0, \varepsilon]$. Choose $x_0 > 0$ sufficiently small such that x_0, $g(x_0)$, $h(x_0)$, $g^2(x_0) \in (0, \varepsilon]$. Then $g(x_0) > h(x_0)$ and $g^2(x_0) > gh(x_0) > h^2(x_0)$. This contradicts $g^2 = h^2$.

THEOREM 3. *If $f \in \boldsymbol{T}^\omega_-$ and f is reversible in $\boldsymbol{T}^\omega$, then f is an involution in $\boldsymbol{T}^\omega$* (hence $1 \cdot f$ is the product of two involutions in $\boldsymbol{T}^\omega$).

Proof of Theorem 3. Suppose $f^{-1} = gfg^{-1}$. The proof of this theorem is divided into two cases according as $g \in \boldsymbol{T}^\omega_+$ or $g \in \boldsymbol{T}^\omega_-$.

Case I. $g \in \boldsymbol{T}^\omega_+$.
Moreover, $gfg^{-1} = f^{-1}$ iff $g_- f_+ g_+^{-1} = f_-^{-1}$ and $g_+ f_- g_-^{-1} = f_+^-$ iff $f_- = g_+ f_+^{-1} g_-^{-1}$ and $f_- = g_+^{-1} f_+^{-1} g_-$

$$\text{iff } \begin{cases} f_- = g_+ f_+^{-1} g_-^{-1} \\ g_+^2 = f_+^{-1} g_-^2 f_+ = (f_+^{-1} g_- f_+)^2 \end{cases} \qquad (\sharp) .$$

Since g_+ is locally increasing and analytic the foregoing implies that

$$g_+ = f_+^{-1} g_- f_+ \qquad (*) .$$

using Lemma 5.
Therefore

$$f_- = g_+ f_-^{-1} g_-^{-1} \qquad \text{by } (\sharp) .$$

$$f_- = (f_+^{-1} g_- f_+) f_+^{-1} g_-^{-1} \qquad \text{by } (*) .$$

Therefore $f_- = f_+^{-1}$ so f is an involution.

Case II. Suppose $g \in T_-^\omega$.

Therefore, $gfg^{-1} = f^{-1}$ is equivalent to $g_+f_-g_-^{-1} = f_-^{-1}$ and $g_-f_+g_+^{-1} = f_+^{-1}$. The foregoing is equivalent to $g_- = f_-g_+f_-$ and $g_- = f_+^{-1}g_+f_+^{-1}$. This means $g_-g_+ = (f_-g_+)^2$ and $g_-g_+ = (f_+^{-1}g_+)^2$. Therefore, $(f_-g_+)^2 = (f_+^{-1}g_+)^2$. Lemma 5 implies $f_-g_+ = f_+^{-1}g_+$. This means that $f_- = f_+^{-1}$. The last statement and Remark 2 implies that f is an involution. This concludes the proof of this theorem.

REMARK. It is known that Lemma 5 is false in T^0. It is not known if Lemma 5 is true in T^k $1 \leqq k \leqq \infty$. The truth of Lemma 5 for T^k would permit one to state the analogue of Theorem 6 for T^k. If Lemma 5 is false for T^k one could conclude that reversible elements of T^k could not be factored into the product of two involutions.

There is a square root lemma, weaker than Lemma 5, which will be proved which provides an additional criterion for the factorability of a reversible element of T^k into the product of two involutions.

LEMMA 6. *If $f, g \in T_+^k$, $f'(0) = a \neq 1$, $k \geqq 2$, $f^2 = g^2$, then $f = g$.*

Proof of Lemma 6. If $f^2 = g^2 = l$, then $l'(0) = a^2$. By Sternberg [5] Theorem 2 there exists $h \in T_+^{k-1}$ such that $hlh^{-1} = \phi_{k-1}(a^2x)$. It is now shown that the only differentiable square root of $\phi_{k-1}(a^2x)$ in T_+^{k-1} is $\phi_{k-1}(ax)$. Suppose there is another square root of $\phi_{k-1}(a^2x)$ given by $\phi_{k-1}(K(x))$. Consider some set $(0, \eta)$ on which $K^2(x) = a^2x$, $(\eta > 0)$. It is clear that $K(x) > ax$ or $K(x) < ax$ cannot hold in $(0, \eta)$. Therefore there is some point $x_0 \in (0, \eta)$ where $ax = K(x)$. Assume further that $a < 1$. If this is not the case one can apply the argument to l^{-1}.

In the argument following it will be convenient to utilize the notation $(F^m(y_0))'$ for $d/dx\, F^m(x)\,|_{x=y_0}$.

Let $x_k = K^k(x_0) = a^kx_0$. Since $a < 1$ $\lim_{k\to\infty} x_k = 0$ monotonically. $K'(x) \neq ax$ at some point in (x_1, x_0); otherwise $K(x) = ax$ on $(0, \eta)$ and there is nothing to prove. Assume, therefore, that there exists $y_0 \in (x_1, x_0)$ such that $K'(y_0) = \beta \neq a$. Let $Z_k = K^{2k-1}(y_0)$. Since $x_{2k} < Z_k < x_{2k-1}$ one observes that $\lim_{k\to\infty} Z_k = 0$. Since $K^2(x) = a^2x$, one observes that $K'(0) = a$ by the chain rule and the fact that $K'(0) > 0$. This implies $\lim_{k\to\infty} K'(Z_k) = a$.
But,

$$K'(Z_k) = K'\,K^{2k-1}(y_0) = \frac{(K^{2k}(y_0))'}{(K^{2k-1}(y_0))'}$$

$$= \frac{(K^{2k}(y_0))'}{(K^{2k-2}(y_0))'\,K(y_0))'}$$

Therefore,

$$K'(Z_k) = \frac{a^{2k}}{a^{2k-2}\beta} = \frac{a^2}{\beta} \neq a .$$

Therefore $\lim_{k\to\infty} K'(Z_k) \neq K'(0)$ which contradicts the continuity of $K'(x)$. This means that $\phi_{k-1}(ax)$ is the only square root of $\phi_{k-1}(a^2x)$ in $T_+^{k-1}(R, 0)$.

Since hfh^{-1} and hgh^{-1} are differentiable square roots of $\phi_{k-1}(a^2x)$, $hfh^{-1} = hgh^{-1}$ which implies $f = g$. This concludes the proof.

The foregoing leads to the following theorem.

THEOREM 4. *If $f \in T_-^k$, $k \geqq 2$, f reversible in T_-^k and $f^{-1} = gfg^{-1}$ $g \in T_+^k$ and $g'(0) \neq 1$, then f is an involution.*

Proof of Theorem 4. The proof utilizes the construction in the proof of Theorem 3 and will therefore be omitted.

5. Conclusions. In this paper Question (♯) of the introduction is shown to have answer "no" for continuous homeomorphisms and nearly "yes" for analytic homeomorphisms. For C^k homeomorphisms, $1 \leqq k \leqq \infty$, the reply to the Question is shown to depend on the existence of unique square roots in a particular group and the non-existence of an element $f \in T_+^k$ such that $\rho_k(f) = 1$ and f is reversible in T_+^k. It is not known whether $gfg^{-1} = f^{-1}$ in T^ω, $f \in T_+^\omega$, $f \neq 1$ implies that $\bar{g}$ is more than twice differentiable. One conjectures that this is not the case since it may be shown that $\alpha\beta\alpha^{-1} = \beta^{-1}$ in P^3, $\beta \neq 1$ and $\beta = x + b_3x^3$, does not ensure that $\alpha^2 = 1$.

One may generalize some of this material to higher dimensions. The techniques of this paper are applicable to some germs of transformations of R^n which fix a sufficiently nice $n - 1$ manifold containing the origin. This and related questions will be treated in a subsequent paper.

REFERENCES

1. S. Bochner, *Compact groups of differentiable transformations,* Annals of Math., **47** (1946), 639-653.

2. Emile Borel, *Sur quelques points de la theorie des fonctions,* Paris, Gauthier Villars et Fils, 1894.

3. A. B. Calica, *Reversible Homeomorphisms of the Real Line,* Thesis, Johns Hopkins University, Baltimore, 1970.

4. D. C. Lewis, Jr., *Reversible Transformations,* Pacific J. Math. Vol. **II**, No. 3, (1961), 1077-1087.

5. S. Sternberg, *Local C^n Transformations of the Real Line,* Duke Math. J. **24** (1957), 97-102.

Received May 22, 1970. This material is part of the author's dissertation written under the supervision of Professor D. C. Lewis, Jr. of Johns Hopkins University. This work was supported in part by a grant from the Naval Academy Research Council. The referee and Professor Samelson corrected a number of obscurities in the original exposition. In addition the proofs of Proposition 1 and Theorem 2 are elegant alternatives to the original, provided by Professor Samelson.

UNIVERSITY OF HAWAII

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

SEMI-GROUPS OF LOCAL LIPSCHITZIANS IN A BANACH SPACE

J. T. CHAMBERS AND S. OHARU

The purpose of this paper is to construct a nonlinear semi-group determined by a given (multi-valued) nonlinear operator A in a Banach space X, and to investigate the differentiability of this semi-group. The semi-group treated in this paper is the semigroup $\{T(t);\ t \geqq 0\}$ of nonlinear operators in X such that for each $\tau > 0$, $\{T(t);\ 0 \leqq t \leqq \tau\}$ is equi-Lipschitz continuous on bounded sets. In order that an operator A in X determine such a semi-group $\{T(t);\ t \geqq 0\}$ on $D(A)$ with $(d/dt)T(t)x \in AT(t)x$ for almost all $t \geqq 0$ and $x \in D(A)$, it is required that X have a uniformly convex dual, A be dissipative in a local sense, I-λA, λ positive and small, satisfy a range condition and an injectiveness condition, and finally the family of operators $(I-\lambda A)^{-n}$, $n = 1, 2, 3, \cdots$ be locally equi-bounded.

Let X be a Banach space and S a subset of X, and let $\{T(t);\ t \geqq 0\}$ be a family of nonlinear operators from S into itself satisfying the following conditions:

(i) $T(0) = I$ (the identity) and $T(t+s) = T(t)T(s)$ on S for $t, s \geqq 0$.

(ii) For $x \in S$, $T(t)x$ is strongly continuous in $t \geqq 0$.

Then the family $\{T(t);\ t \geqq 0\}$ is called a *semi-group on* S. The *infinitesimal generator* A_0 of the semi-group $\{T(t); t \geqq 0\}$ is defined by $A_0x = \lim_{h\to+0} h^{-1}\{T(h)x - x\}$ and the *weak infinitesimal generator* A' by $A'x = w\text{-}\lim_{h\to+0} h^{-1}\{T(h)x - x\}$, if the right sides exist, the notation "w-lim" means the weak limit in X.

An operator A in X is called a *D-operator* if for every bounded set B in X there exists a number $\omega_B \geqq 0$ such that

$$re < x' - y', f > \leqq \omega_B \| x - y \|^2 \quad \text{for } x, y \in B \cap D(A),\ x' \in Ax,\ y' \in Ay$$

and some $f \in F(x - y)$, where F denotes the duality mapping of X.

Our discussion requires that X have a *uniformly convex dual*. Then, if A is a D-operator satisfying some additional conditions, we obtain a semi-group $\{T(t);\ t \geqq 0\}$ on $\overline{D(A)}$ such that

$$\text{(A)} \qquad T(t)x = \lim_{\lambda\to+0} (I - \lambda A)^{-[t/\lambda]}x\ , \qquad x \in D(A)$$

and the convergence is uniform with respect to t in every finite interval;

(B) for every bounded set B in $D(A)$ and $\tau > 0$, there exists a number

$\omega_{B,\tau} \geqq 0$ such that $\| T(t)x - T(t)y \| \leqq e^{\omega_{B,\tau} t} \| x - y \|$ for $x, y \in B$ and $t \in [0, \tau]$.

The additional conditions on A are stated roughly as follows:

(1) The operator $(I - \lambda A)^{-1}$ must exist as a single-valued operator with domain $R(I - \lambda A)$, the range of $I - \lambda A$, for λ small; this is condition (I) of the paper.

(2) In order that the iterations of $(I - \lambda A)^{-1}$ be meaningful on $D(A)$, it is required that the range of $I - \lambda A$ contain $D(A)$; this is condition (R).

(3) The operators $(I - \lambda A)^{-k}$, $k = 1, 2, 3, \cdots$ must map bounded sets into bounded sets; this is the idea behind condition (E).

We note that if A is a dissipative operator, i.e., $\omega_B \equiv 0$ for every bounded set B in X, then (1) and (3) are satisfied.

Concerning the differentiability of the semi-group constructed we obtain, among other results, the following. If A is a D-operator satisfying (I), (R) and (E) and is maximal on $\overline{D(A)}$ in the sense explained in § 1, then there exists a uniquely determined semi-group $\{T(t); t \geqq 0\}$ on $\overline{D(A)}$ such that for each $x \in D(A)$ $(d/dt)T(t)x \in AT(t)x$ at almost all $t \geqq 0$.

Finally, we remark that for the Cauchy problem

$$(d/dt)u(t) \in Au(t), \ u(0) = x ,$$

where A is a D-operator in X satisfying (I), (R) and (E), we can construct the semi-group solution using the convergence (A). And conversely, in a reflexive Banach space, if A_0 is the infinitesimal generator of a semi-group $\{T(t); t \geqq 0\}$ satisfying (B), then A_0 is a D-operator in X and for $x \in D(A_0)$, $T(t)x$ is a solution of the Cauchy problem formulated for the operator A_0.

Section 1 deals with the notion of a D-operator and some of its properties. Section 2 concerns the abstract Cauchy problem. Section 3 contains the construction of the semi-group determined by the D-operator A. Finally, in Section 4, the question of the differentiability of the constructed semi-group is discussed.

The authors want to express their deep gratitude to Professor I. Miyadera for his many valuable suggestions.

0. Preliminaries. In this section we introduce some of the basic notions which are used in this paper.

Throughout this paper X denotes a Banach space. Let A be a multi-valued operator in X, that is, A assigns to each $x \in X$ a subset Ax of X. Ax may be empty for some $x \in X$. The *domain* of A, $D(A)$, is the set of all $x \in X$ such that $Ax \neq \emptyset$; the *range* of A, $R(A)$, is the set $\bigcup_{x \in X} Ax$. We write AS (or $A(S)$) for $\bigcup_{x \in S} Ax$, $S \subset X$.

Note that a single-valued operator is a special case of a multi-valued operator in which Ax, $x \in D(A)$, denotes the value of A at x or the singleton set consisting of this element, and Ax is the empty set if $x \notin D(A)$.

For subsets $S_1, S_2 \subset X$, $S_1 + S_2$ denotes the set $\{x+y;\ x \in S_1,\ y \in S_2\}$ where $S_1 + S_2 = \emptyset$ if $S_1 = \emptyset$ or $S_2 = \emptyset$. For a scalar λ and $S \subset X$, λS denotes the set $\{\lambda x;\ x \in S\}$, and we write $y + S$ for $\{y\} + S$.

Accordingly, for two operators A and B in X, we define the sum $A + B$ in X by $(A + B)x = Ax + Bx$, $D(A+B) = D(A) \cap D(B)$; the scalar multiplication λA in X by $(\lambda A)x = \lambda Ax$, $D(\lambda A) = D(A)$; and the product AB in X by $(AB)x = A(Bx)$, $D(AB) \subset D(B)$. We write $\gamma + \lambda A$ for the operator $\gamma I + \lambda A$, where I denotes the identity operator in X. For any positive integer k, we define the iteration A^k in X by $A^k x = A(A^{k-1}x)$, where $A^0 \equiv I$ and $D(A^k) \subset D(A)$.

Let A, $\tilde{A}$ be two operators in X. $\tilde{A}$ is an *extension* of A, and A is a *restriction* of $\tilde{A}$ (denoted $\tilde{A} \supset A$, $A \subset \tilde{A}$), if $Ax \subset \tilde{A}x$ for each $x \in X$, thus $D(A) \subset D(\tilde{A})$. If $S \subset X$, then by a restriction of A to S, $A|_S$, we mean the operator such that $D(A|_S) = D(A) \cap S$ and $A|_S x = Ax$ if $x \in S$.

If $S \subset X$, we denote the closure of S in X by $\bar{S}$. Let A be an operator in X, then B is called the *closure* of A, if $G(B) = \overline{G(A)}$, where $G(\cdot)$ denotes the graph of the operator. We write $B = \bar{A}$.

Let X^* be the dual space of X. We denote by $\langle x, f \rangle$ the pairing between $x \in X$ and $f \in X^*$. The *duality mapping* F *of* X is the mapping from X into X^* defined by

$$F(x) = \{f \in X^*;\ re\langle x, f\rangle = \|x\|^2 = \|f\|^2\}$$

for $x \in X$. If X^* is uniformly convex, then F is single-valued and uniformly continuous on bounded sets [4; Lemma 1.2].

We now state some standard definitions and collect some well-known results.

DEFINITION 1. An operator A in X is said to be *dissipative* if for each x, $y \in D(A)$ and $x' \in Ax$, $y' \in Ay$, there exists an $f \in F(x - y)$ such that $re\langle x' - y', f\rangle \leqq 0$. A is said to be an *m-dissipative operator in* X, if it is a dissipative operator in X and $R(I - \lambda_0 A) = X$ for some $\lambda_0 > 0$. Let $S \subset X$ and A be a dissipative operator in X, if every dissipative extension of A coincides on S with A, then A is said to be a *maximal dissipative operator on* S.

An m-dissipative operator A is maximal dissipative on $D(A)$. If X^* is strictly convex and A is a maximal dissipative opetator on S, then Ax is closed and convex for $x \in S$. If A is an m-dissipative operator, then $R(I - \lambda A) = X$ for all $\lambda > 0$ ([9; Lemma 4]).

DEFINITION 2. An operator A in X is said to be *demi-closed* if the following condition holds: if $\{x_n\} \subset D(A)$, $x_n \to x \in X$ (strong convergence) and if $y_n \in Ax_n$, such that $y_n \rightharpoonup y \in X$ (weak convergence) implies that $x \in D(A)$ and $y \in Ax$.

A demi-closed operator is closed. If X^* is uniformly convex and A is maximal dissipative on $\overline{D(A)}$, then A is demi-closed ([5; Lemma 3.7]).

DEFINITION 3. Let A be an operator in X. The operator A^0 defined by $A^0x = \{y \in Ax; \|y\| = \inf[\|u\|; u \in Ax]\}$ is called the *canonical restriction of A*.

If X^* is uniformly convex and A is an m-dissipative operator, then $D(A^0) = D(A)$ and A^0x is a non-empty closed convex set for $x \in D(A)$. If X and X^* are uniformly convex and Ax is closed and convex for $x \in X$, then A^0 is single-valued and $D(A^0) = D(A)$ ([5; Lemma 3.10]).

Finally, we list some notations which are used in this paper.

(1) Let $\{x_n\}$ be a sequence in X, then "$x_n \to x$", means that x_n converges to x in the norm topology, whereas, "$x_n \rightharpoonup x$", means that x_n converges to x in the weak topology.

(2) Let G be a single-valued operator in X and $B \subset X$, then by $\|G\|_{\mathrm{Lip}(B)}$, we mean the smallest Lipschitz constant for G on $B \cap D(G)$.

(3) We write J_λ for the resolvent $(I - \lambda A)^{-1}$ if it is well-defined and R_λ for the range $R(I - \lambda A) = \{x - \lambda y; x \in D(A), y \in Ax\}$.

(4) Let $K \subset X$. Then coK denotes the convex hull of K and $\overline{co}K$, the convex closure of K.

(5) For any nonempty set $S \subset X$, we write

$$|||S||| = \inf\{\|x\|; x \in S\} .$$

Thus for any operator A, $|||Ax|||$ is defined for $x \in D(A)$

1. D-operators. In this section we introduce the notion of a D-operator and establish some of its properties.

Let X be a Banach space and A an operator in X. If for every bounded set $B \subset X$ there exists a nonnegative number ω_B such that

$$re < x' - y', f > \leqq \omega_B \|x - y\|^2$$

for $x,\ y \in B \cap D(A)$, $x' \in Ax$, $y' \in Ay$ and for some $f \in F(x-y)$, then A is called a *D-operator*.

Put $B_n = \{x \in X;\ \|x\| < n\}$, $n = 1, 2, 3, \cdots$ If there exists a sequence $\{\omega_n\}$ of nonnegative numbers such that

$$re\langle x' - y', f\rangle \leqq \omega_n \|x - y\|^2$$

for $x,\ y \in B_n \cap D(A)$, $x' \in Ax$, $y' \in Ay$ and for some $f \in F(x-y)$, $n = 1, 2, 3, \cdots$, then A is a D-operator. If such a sequence is identically zero, then A is a dissipative operator. Note that if A is a D-operator, then $(A - \omega_n)|_{B_n}$ is a dissipative operator on B_n.

The next lemma by Kato [4; Lemma 1.1] gives a basic property of dissipative operators.

PROPOSITION 1.1. (Kato) *Let $x, y \in X$. Then there is a $\lambda_0 > 0$ such that $\|x\| \leqq \|x - \lambda y\|$ for $\lambda \in (0, \lambda_0)$ if and only if there is an $f \in F(x)$ such that $re\langle y, f\rangle \leqq 0$.*

Let A be a D-operator in X, then for every bounded set $B \subset D(A)$, we have that $(I - \lambda A)x \cap (I - \lambda A)y = \varnothing$ for $x, y \in B$, if $x \neq y$ and $\lambda \in (0, 1/\omega_B)$. In fact, for $x' \in Ax$, $y' \in Ay$, and some $f \in F(x-y)$ we have that

$$\begin{aligned}\|(x - \lambda x') - (y - \lambda y')\|\, \|x - y\| &\geqq re\langle (x - \lambda x') - (y - \lambda y'), f\rangle \\ &\geqq (1 - \lambda\omega_B)\|x - y\|^2 .\end{aligned}$$

Hence, we have $\|(x - \lambda x') - (y - \lambda y')\| \geqq (1 - \lambda\omega_B)\|x - y\|$, so, $(I - \lambda A)|_B$ has a Lipschitz continuous inverse and

$$\|(I - \lambda A|_B)^{-1}\|_{\mathrm{Lip}((I-\lambda A)B)} \leqq (1 - \lambda\omega_B)^{-1} \quad \text{for } \lambda \in (0, 1/\omega_B) .$$

However, in general, $(I - \lambda A)^{-1}$ is not a single-valued operator. For example, take X to be the real line and Ax, the function $x \sin x$. A is a D-operator, in fact, for the bounded set $[-M, M]$ we may take $\omega_{[-M,M]}$ to be $1 + M$. And

$$\|(I - \lambda A|_{[-M,M]})^{-1}\|_{\mathrm{Lip}((I-\lambda A)[-M,M])} \leqq (1 - \lambda(1 + M))^{-1} \text{ for } \lambda \in (0, 1/(1 + M)).$$

But $R(I - \lambda A) = X$ for $\lambda > 0$, and $(I - \lambda A)^{-1}$ can not defined as a single-valued operator on X, no matter how small we restrict $\lambda > 0$.

Hence, we make an additional assumption on the operator A:

(I) $(I - \lambda A)x \cap (I - \lambda A)y = \varnothing$ for $x,\ y \in X$ if $x \neq y$ and $\lambda \in (0, \lambda_0)$.

Condition (I) guarantees the existence of the resolvent $J_\lambda \equiv (I - \lambda A)^{-1}$ for $\lambda \in (0, \lambda_0)$ as a single-valued operator with $D(J_\lambda) = R(I - \lambda A)$. (I) corresponds to the assumption that $I - \lambda A$ is injective, if we are considering single-valued operators.

DEFINITION 1.1. Let G be an operator in X. G is said to be *locally bounded* if G maps bounded sets into bounded sets.

DEFINITION 1.2. Let $\{G_\gamma\}$, $\gamma \in \Gamma$, be a family of operators in X. $\{G_\gamma\}$, $\gamma \in \Gamma$, is said to be *locally equi-bounded*, if for every bounded set B, $\bigcup_{\gamma \in \Gamma} G_\gamma(B)$ is a bounded set.

PROPOSITION 1.2. *Let A be a D-operator in X satisfying* (I). *If $\{J_\lambda;\ \lambda \in (0, \lambda_0)\}$ is locally equibounded, then for every bounded set $B \subset X$ there exists a number $\tilde{\omega}_B \geqq 0$ such that*

$$\| J_\lambda \|_{\mathrm{Lip}(B)} \leqq (1 - \lambda \tilde{\omega}_B)^{-1} \qquad \text{for } \lambda \in \left(0, \min\left\{\lambda_0, \frac{1}{\tilde{\omega}_B}\right\}\right).$$

Proof. Let B be any bounded set in X, then $B_1 \equiv \bigcup_{\lambda \in (0, \lambda_0)} J_\lambda(B)$ is a bounded set and $B_1 \subset D(A)$. Hence, there exists a number $\tilde{\omega}_B \geqq 0$ such that $(1 - \lambda\tilde{\omega}_B)\| x - y \| \leqq \|(x - \lambda x') - (y - \lambda y')\|$ for $x, y \in B_1$, $x' \in Ax$, $y' \in Ay$ and $\lambda \in (0, \min\{\lambda_0, 1/\tilde{\omega}_B\})$. Thus, if $u, v \in B \cap R_\lambda$, then $J_\lambda u, J_\lambda v \in B_1$ and

$$(1 - \lambda\tilde{\omega}_B)\| J_\lambda u - J_\lambda v \| \leqq \| u - v \| \quad \text{for } \lambda \in \left(0, \min\left\{\lambda_0, \frac{1}{\tilde{\omega}_B}\right\}\right).$$

In the next proposition we impose two additional conditions on the operator A, which are essential to the construction of the semi-group in this paper.

PROPOSITION 1.3. *Let A be a D-operator in X satisfying* (I). *If*

(R) $$R(I - \lambda A) \supset D(A) \qquad \text{for } \lambda \in (0, \lambda_0),$$

and

(E) $$\{J_\lambda^n;\ \lambda \in (0, \lambda_0),\ n\lambda \in [0, \tau]\}$$

is locally equi-bounded for any $\tau > 0$, then for every bounded set $B \subset X$ and $\tau > 0$, there exists a number $\omega_{B,\tau} \geqq 0$ such that

$$\| J_\lambda^n \|_{\mathrm{Lip}(B)} \leqq (1 - \lambda\omega_{B,\tau})^{-n}$$

for $\lambda \in (0, \min\{\lambda_0, 1/\omega_{B,\tau}\})$ and $n\lambda \in [0, \tau]$.

Proof. Let B be any bounded set in X and $\tau > 0$. Set $B_1 = \bigcup_{\substack{\lambda \in (0,\lambda_0) \\ n\lambda \in [0,\tau]}} J_\lambda^n(B)$, then $B_1 \cup B$ is a bounded set and so there exists a number $\omega_{B,\tau} \geqq 0$ such that

$$\| J_\lambda \|_{\mathrm{Lip}(B_1 \cup B)} \leqq (1 - \lambda\omega_{B,\tau})^{-1} \quad \text{for } \lambda \in \left(0, \min\left\{\lambda_0, \frac{1}{\omega_{B,\tau}}\right\}\right).$$

Thus, if $u, v \in B \cap R_\lambda$, then $J_\lambda^{n-1} u, J_\lambda^{n-1} v \in B_1 \cup B$, provided $n\lambda \in [0, \tau]$. So that $\| J_\lambda^n u - J_\lambda^n v \| \leqq (1 - \lambda\omega_{B,\tau})^{-1} \| J_\lambda^{n-1} u - J_\lambda^{n-1} v \| \leqq (1 - \lambda\omega_{B,\tau})^{-n} \| u - v \|$ for $\lambda \in (0, \min\{\lambda_0, 1/\omega_{B,\tau}\})$.

The next proposition gives some sufficient conditions for (E).

PROPOSITION 1.4. *Let A be a D-operator satisfying* (I).

(a) *If there exist nonnegative numbers M and N such that*

$$\| J_\lambda x \| \leqq (1 + \lambda M) \| x \| + \lambda N \quad \textit{for } \lambda \in (0, \lambda_0) \textit{ and } x \in R_\lambda ,$$

then (E) holds.

(b) *If A is single-valued and* $\sup\{\|Ax\|; x \in D(A)\} < +\infty$, *then (E) holds.*

Proof. (a) Let B be any bounded set in X and $x \in B \cap R_\lambda$, $\lambda \in (0, \lambda_0)$. Then, it is easy to see that

$$\| J_\lambda^n x \| \leqq (1 + M\lambda)^n (\| x \| + n\lambda N) \leqq e^{Mn\lambda} (\sup \| x \| + n\lambda N) ,$$

which is bounded for $\lambda \in (0, \lambda_0)$ and $n\lambda \in [0, \tau]$.

(b) Take $x \in R_\lambda$, then $\| J_\lambda x \| \leqq \| x \| + \lambda \| A J_\lambda x \|$. Put

$$N = \sup_{x \in D(A)} \| Ax \| ,$$

then $\| J_\lambda x \| \leqq \| x \| + \lambda N$. Now apply (a), note that in this case $M = 0$.

We now wish to introduce a notion of maximal D-operator. Given a sequence of nondecreasing nonnegative numbers $\{\omega_n\}$, we consider the family of D-operators, $\mathscr{F}\{\omega_n\}$, consisting of all D- operators A in X such that there exist numbers $\omega_{B_n}(A) \leqq \omega_n$, $n = 1, 2, 3, \cdots$ with

$$re\langle x' - y', f \rangle \leqq \omega_{B_n}(A) \| x - y \|^2$$

for $x, y \in D(A) \cap B_n$, $x' \in Ax$, $y' \in Ay$ and some $f \in F(x - y)$, $n = 1, 2, 3, \cdots$. Recall that B_n denotes the open ball with radius n and center 0 in X. Note that if A is a D-operator, then there exists a sequence $\{\omega_n\}$ such that $A \in \mathscr{F}\{\omega_n\}$.

DEFINITION 1.3. If $A \in \mathscr{F}\{\omega_n\}$, then A is called a $(D, \{\omega_n\})$-*operator*. Let $S \subset X$ and A be a $(D, \{\omega_n\})$-operator in X. If every

$(D, \{\omega_n\})$-extension of A coincides on S with A, then A is said to be a maximal $(D, \{\omega_n\})$-operator on S.

PROPOSITION 1.5. *If A is a $(D, \{\omega_n\})$-operator in X and $S \subset X$, then there exists a maximal $(D, \{\omega_n\})$-operator $\tilde{A}$ on S such that $\tilde{A}\,|_S \supset A\,|_S$.*

Proof. Apply Zorn's Lemma.

We now show that if A is a maximal $(D, \{\omega_n\})$-operator on $\overline{D(A)}$ and furthermore if X^* is uniformly convex, then A is demi-closed and Ax is closed and convex. The uniform convexity of X^* gives the above properties which are essential in establishing the facts concerning the differentiability of the semi-groups constructed in this paper.

PROPOSITION 1.6. *Let X^* be uniformly convex. If A is a maximal $(D, (\omega_n)\}$-operator on $\overline{D(A)}$, then*

(a) *A is demi-closed,*

(b) *Ax is closed convex.*

Proof. (a) Let $\{x_k\}$ be a sequence such that $\{x_k\} \subset D(A)$, $x_k \to x_0 \in X$ and $Ax_k \ni y_k \rightharpoonup y$. We must show that $x_0 \in D(A)$ and $y \in Ax_0$. Define $\tilde{A}w = Aw$ if $w \neq x_0$ and $Aw \cup \{y\}$ if $w = x_0$. Then $\tilde{A} \supset A$ and $D(\tilde{A}) \subset \overline{D(A)}$. It is easy to see that $\tilde{A}$ is a $(D, \{\omega_n\})$-operator, Hence, by the maximality of A, $\tilde{A} = A$ and so $x_0 \in D(A)$ with $y \in Ax_0$.

(b) The same type of argument as in (a) easily establishes (b).

The next proposition states some basic properties of a demi-closed operator.

PROPOSITION 1.7. *Let X be a reflexive Banach space and A be a demi-closed operator in X. Let $\{x_n\} \subset D(A)$, $x_n \to x_0 \in X$,* and let *$\{y_n\}$ be a sequence in X such that $y_n \in Ax_n$ for each n. Then:*

(a) *if $\{y_n\}$ is bounded and V is the set of all weak cluster points of $\{y_n\}$, then $x_0 \in D(A)$, $V \neq \emptyset$, and $V \subset Ax_0$; if in particular, A is single-valued, then $y_n \rightharpoonup Ax_0$;*

(b) *if furthermore, X and X^* are uniformly convex, the canonical restriction A^0 is single-valued and if* $\limsup \|y_n\| \leqq ||| Ax_0 |||$, *then $x_0 \in D(A^0)$ and $y_n \to A^0 x_0$.*

Proof. (a) First, the reflexivity of X and the boundedness of $\{y_n\}$ imply that $V \neq \emptyset$. Let Y be the closed linear manifold determined by $\{y_n\}$. Then Y is a reflexive Banach space and Y^* is separable. Hence, for each $y \in V$ a subsequence $\{y_{n_i}\}$ can be found such that $y_{n_i} \rightharpoonup y$ in Y. Now, any $x^* \in X^*$ determines a $y^* \in Y^*$ such that $\langle x, x^* \rangle = \langle x, y^* \rangle$ for $x \in Y$; thus, since $y_n \in Y$, $\langle y_{n_i}, x^* \rangle \to \langle y, x^* \rangle$ for $x^* \in X^*$, and $y_{n_i} \rightharpoonup y$ in X. Since $x_{n_i} \to x_0$, $y_{n_i} \in Ax_{n_i}$, $y_{n_i} \rightharpoonup y$ and since A is demi-closed, we have that $x_0 \in D(A)$ and $y \in Ax_0$. This means that $V \subset Ax_0$. If in particular, A is single-valued, then $y = Ax_0$; hence all weak limits of subsequences of $\{y_n\}$ are same and equal to Ax_0. Therefore, it follows that $y_n \rightharpoonup Ax_0$.

(b) Since $\{y_n\}$ is bounded and A is demi-closed, there is a subsequence $\{y_{n_i}\}$ and a $y \in Ax_0$ such that $y_{n_i} \rightharpoonup y$. Thus, by assumption, we have

$$||| Ax_0 ||| \leqq \| y \| \leqq \liminf \| y_{n_i} \| \leqq \limsup \| y_{n_i} \| \leqq ||| Ax_0 ||| .$$

Since A^0 is single-valued, $y = A^0 x_0$ and $\lim \| y_{n_i} \| = \| A^0 x_0 \|$. But, X is uniformly convex; thus $y_{n_i} \to A^0 x_0$. Therefore, all strong limits of subsequences are same and equal to $A^0 x_0$, and it follows that $y_n \to A^0 x_0$.

PROPOSITION 1.8. *Let X and X^* be uniformly convex. Let A be a closed D-operator in X satisfying* (I), (3.9) (stated in Remark 3.1) and (E). *If $\tilde{A}$ is a maximal $((D, \{\omega_n\})$-extension of A on $\overline{D(A)}$ such that*

$$D(\tilde{A}) \subset \{z \in \overline{D(A)}; \| J_\lambda z - z \| = 0(\lambda) \text{ as } \lambda \downarrow 0\} ,$$

then $A^0 = \tilde{A}^0$.

Proof. First, note that $\tilde{A}$ is demi-closed, $\tilde{A}x$ is closed and convex, and so, $\tilde{A}^0$ is single-valued with $D(\tilde{A}^0) = D(\tilde{A})$. Take a sequence $\eta_k \downarrow 0$ and set $J_k = J_{\eta_k}$ and $A_k = \eta_k^{-1}[J_k - I]$. Let $x \in D(\tilde{A})$, then, since $D(A_k) \supset D(\tilde{A})$, we see that $\| A_k x \| = \eta_k^{-1} \| J_k x - J_k(x - \eta_k y) \| \leqq ||| \tilde{A}x ||| / (1 - \omega_B \eta_k)$ for $y \in \tilde{A}^0 x$, where ω_B is a constant associated with the closure B of $\{x - \eta_k y; y \in \tilde{A}x, \| y \| = ||| \tilde{A}x |||, k \text{ sufficiently large}\}$ through the D-operator A. Since $A_k x \in \tilde{A} J_k x$ for k, $A_k x \to \tilde{A}^0 x$ as $k \to +\infty$ by Proposition 1.7 (b), for each $x \in D(\tilde{A})$. Now take $z \in D(\tilde{A})$. Since $R(I - \lambda A) \supset \overline{D(A)}$ for $\lambda \in (0, \lambda_0)$ by assumption and since $\overline{D(A)} \supset D(\tilde{A})$, we see that $z \in R(I - \eta_k A)$ for k sufficiently large. Hence, there exist $x_k \in D(A)$ and $y_k \in Ax_k$ such that

$$z = x_k - \eta_k y_k .$$

But, $x_k = J_k z \to z$ as $k \to +\infty$; hence, by the closedness of A, $z \in D(A)$ and $\tilde{A}^0 z \in Az$. But $Az \subset \tilde{A}z$, so that $\tilde{A}^0 z \in A^0 z$. Also, $||| A^0 z ||| \leqq \| \tilde{A}^0 z \|$. Therefore, $v \in A^0 z \subset \tilde{A}z$ implies that $v = \tilde{A}^0 z$ because $\tilde{A}^0$ is single-valued. So, A^0 is also single-valued.

REMARK 1.1. Brezis and Pazy [1; Theorem 2.1] give the following result. Let X be a Hilbert space and A be a closed dissipative operator such that $R_\lambda \supset \overline{co}\, D(A)$ for all $\lambda > 0$, then A has a unique extension to a maximal dissipative operator $\tilde{A}$ satisfying $D(\tilde{A}) \subset \overline{D(A)}$, and, in fact, $D(\tilde{A}) = D(A)$ and $\tilde{A}^0 = A^0$.

2. Abstract Caucy problem. In this section we discuss the relationship between the abstract Cauchy problem formulated for a D-operator and the semi-group generated by such an operator.

The abstract Cauchy problem may be stated as follows:

Given an operator A in X and an element $x \in X$, find a X-valued function $u(t; x)$ on $[0, \infty)$ such that

(i) $u(t; x)$ is strongly absolutely continuous on every finite interval;

(ii) $u(0; x) = x$ and $(d/dt)u(t;x) \in Au(t; x)$ for almost all t.

We call this the abstract Cauchy problem, ACP, formulated to A.

PROPOSITION 2.1. *Let A be a D-operator in X. Then there is at most one solution of the ACP formulated to A with the initial value $x \in D(A)$.*

Proof. For $x \in D(A)$, suppose that $u(t; x)$ and $v(t; x)$ are solutions of the ACP formulated to A. By Kato's lemma [4; Lemma 1.3] we have that

$$\| u(t; x) - v(t; x)\|^2 = 2\int_0^t re\Big\langle \Big(\frac{d}{ds}\Big)u(s; x) - \Big(\frac{d}{ds}\Big)v(s; x), f(s)\Big\rangle ds$$

$$\leqq 2\omega_{x,\tau}\int_0^t \| u(s; x) - v(s; x)\|^2\, ds\ ,$$

where $\omega_{x,\tau}$ is a constant associated with the bounded set

$$B = \{u(t; x),\ v(t; x);\ t \in [0, \tau]\}$$

through the D-operator A and $f(s) \in F(u(s; x) - v(s; x))$, and also, note that $(d/ds)u(s; x) \in Au(s; x)$ and $(d/ds)v(s; x) \in Av(s; x)$ for almost all s. Hence, $u(t; x) \equiv v(t; x)$ for $t \in [0, \tau]$. Since τ is arbitrary, $u(t; x) \equiv v(t; x)$ for all $t \geqq 0$.

PROPOSITION 2.2. *If A is a D-operator in X such that for each $x \in D(A)$, there is a solution $u(t; x)$ to the ACP formulated to A satisfying the condition that for any sequentially compact set $K \subset D(A)$ and $\tau > 0$, $\{u(t; x);\ t \in [0, \tau],\ x \in K\}$ is bounded, then there is a semigroup $\{T(t); t \geqq 0\}$ defined on $\overline{D(A)}$ and such that $T(t)x \equiv u(t; x)$, $x \in D(A)$ and $t \in [0, \tau]$. Conversely, if X is reflexive and A_0 is the infinitesimal* *generator of a semi- group* $\{T(t); t \geqq 0\}$ *satisfying the condition that for every* $\tau > 0$ *and bounded set* B *there is a constant* $\omega_{B,\tau} \geqq 0$ *such that* $\| T(t) \|_{\mathrm{Lip}\,(B)} \leqq \exp(\omega_{B,\tau} t)$, $t \in [0, \tau]$, *then* A_0 *is a D-operator in* X *and for each* $x \in D(A_0)$, $T(t)x$ *is a solution of the ACP formulated to* A_0.

Proof. Take $x \in D(A)$ and $\tau > 0$, and put $T(t)x \equiv u(t; x)$, $t \in [0, \tau]$. Since $u(t; x) \in D(A)$ for almost all $t \in [0, \tau]$ and $u(t; x)$ is strongly continuous, $u(t; x) \in \overline{D(A)}$, i.e., $T(t)x \in \overline{D(A)}$ for all $t \in [0, \tau]$. Hence, $T(t)$ maps $D(A)$ into $\overline{D(A)}$. By Kato's lemma, for x, $y \in K$, a compact set, we have that

$$\| T(t)x - T(t)y \|^2 - \| x - y \|^2 = 2 \int_0^t re \left\langle \left(\frac{d}{ds}\right) T(s)x - \left(\frac{d}{ds}\right) T(s)y, f(s) \right\rangle ds$$
$$\leqq 2\omega_{K,\tau} \int_0^t \| T(s)x - T(s)y \|^2 \, ds ,$$

where $\omega_{K,\tau}$ is a number associated with the bounded set

$$\{T(t)x; \; t \in [0, \tau], \, x \in K\}$$

and $f(s) \in F(T(s)x - T(s)y)$, and also, note that $(d/ds)T(s)x \in AT(s)x$ and $(d/ds)T(s)y \in AT(s)y$ for almost all s. Therefore,

$$\| T(t)x - T(t)y \| \leqq \exp(\omega_{K,\tau} t) \, \| x - y \|, \quad x, \; y \in K, \; t \in [0, \tau] \, .$$

Now, take $z \in \overline{D(A)}$, then there exists a sequence $\{x_n\} \subset D(A)$ such that $x_n \to z$, and so, $\| T(t)x_n - T(t)x_m \| \leqq \exp(\omega_{K,\tau} t) \, \| x_n - x_m \|$ where $K = \{x_n\}$. Hence, define $T(t)z = \lim_{n \to \infty} T(t)x_n$, thus $T(t)$ maps $\overline{D(A)}$ into itself. The semi-group property follows from the uniqueness of the solution of the *ACP*. Conversely, take any bounded set B in $D(A_0)$, then for x, $y \in B$

$$re \langle h^{-1}(T(h)x - x) - h^{-1}(T(h)y - y), f \rangle \leqq h^{-1}(\exp(\omega_{B,\tau} h) - 1) \, \| x - y \|^2$$

where $h \in [0, \tau]$ and $f \in F(x - y)$. Letting $h \to +0$, we have that

$$re \langle A_0 x - A_0 y, f \rangle \leqq \omega_{B,\tau} \, \| x - y \|^2 ,$$

so A_0 is a D-operator. Let $x \in D(A_0)$, then

$$\sup \{h^{-1} \, \| T(h)x - x \|; \; 0 < h \leqq 1\} = M < +\infty$$

and $\| T(t+h)x - T(t)x \| \leqq M exp(\omega_{B,\tau} t) h$ for $t \in [0, \tau]$, $h \in (0, 1]$ and $B = \{T(h)x; \; h \in (0, 1]\}$. Thus, $T(t)x$ is strongly absolutely continuous on every finite interval. Since x is reflexive, $T(t)x$ is strongly differentiable for almost all $t \in [0, \tau]$ and $(d/dt)T(t) = A_0 T(t)x$ for almost all $t \in [0, \tau]$. Therefore, $T(t)x$ is a solution of the *ACP* formulated to A_0.

Combining the properties mentioned above we have the following:

PROPOSITION 2.3. *Let A be a D-operator in x. Then there is at most one semi-group $\{T(t); t \geqq 0\}$ on $\overline{D(A)}$ such that for each $x \in D(A)$, $T(t)x$ is a solution of the ACP formulated to A.*

3. Construction of the semi-groups. In this section, we construct the semigroup determined by a D-operator A which satisfies conditions (I), (R) and (E).

Throughout, it is assumed that X has a uniformly convex dual.

LEMMA 3.1. *Let A be a D-operator in X satisfying (I), (R) and (E). If $x \in D(A)$ and $\tau > 0$, then*

$$y(t; x) = \lim_{\lambda \to +0} (I - \lambda A)^{-[t/\lambda]} x \tag{3.1}$$

exists uniformly for $t \in [0, \tau]$.

Proof. Set $J_\lambda = (I - \lambda A)^{-1}$ and $A_\lambda = \lambda^{-1}(J_\lambda - I)$, $\lambda \in (0, \lambda_0)$. Let $x \in D(A)$ and $\tau > 0$. Set

$$B_{x,\tau} = \{J_h^m x;\ h \in (0, \lambda_0),\ mh \in [0, \tau]\} \cup \{x - hy;\ h \in (0, \lambda_0),\ y \in Ax,\ \|y\| \leqq |||Ax||| + 1\},$$

then $B_{x,\tau}$ is a bounded set by (E). Let $\omega_{B_{x,\tau}}$ be a number associated with this bounded set in the sense of Proposition 1.3. Then we have that

$$\begin{aligned}\|A_h J_h^{m-1} x\| &= h^{-1}\|J_h^m x - J_h^{m-1} x\| = h^{-1}\|J_h^m x - J_h^m(x - hy)\| \\ &\leqq (1 - h\omega_{B_{x,\tau}})^{-m}\|y\|\end{aligned}$$

for $y \in Ax$ with $\|y\| \leqq |||Ax||| + 1$. Hence, a positive number $C_{x,\tau}$ can be found such that

$$\|A_h J_h^{m-1} x\| \leqq (1 - h\omega_{B_{x,\tau}})^{-m} |||Ax||| \leqq C_{x,\tau}, \tag{3.2}$$

for h sufficiently small and $mh \in [0, \tau]$. Now, assume that $\lambda n \leqq h$ and $hm \leqq \tau$, where $h \in (0, \lambda_0)$ and m, n are integers. And let $k \leqq m$. Since

$$J_\lambda^{nk} x - J_\lambda^{n(k-1)} x = \lambda \sum_{p=0}^{n-1} A_\lambda J_\lambda^p J_\lambda^{n(k-1)} x,$$

we have

$$\begin{aligned}&(J_\lambda^{nk}x - J_\lambda^{n(k-1)}x) - (J_h^k x - J_h^{k-1}x)\\ &= \lambda \sum_{p=0}^{n-1} \{A_\lambda J_\lambda^{n(k-1)+p}x - A_h J_h^{k-1}x\} + (n\lambda - h)A_h J_h^{k-1}x \,.\end{aligned}$$

Thus, we can write

$$\begin{aligned}&\langle (J_\lambda^{nk}x - J_\lambda^{n(k-1)}x) - (J_h^k x - J_h^{k-1}x),\ F(J_\lambda^{nk}x - J_h^k x)\rangle\\ &= \lambda \sum_{p=0}^{n-1} \langle A_\lambda J_\lambda^{n(k-1)+p}x - A_h J_h^{k-1}x,\ F(J_\lambda^{n(k-1)+p+1}x - J_h^k x)\rangle\\ &\quad + \lambda \sum_{p=0}^{n-1} \langle A_\lambda J_\lambda^{n(k-1)+p}x - A_h J_h^{k-1}x,\ F(J_\lambda^{nk}x - J_h^k x)\\ &\quad - F(J_\lambda^{n(k-1)+p+1}x - J_h^k x)\rangle\\ &\quad + (n\lambda - h)\langle A_h J_h^{k-1}x,\ F(J_\lambda^{nk}x - J_h^k x)\rangle \equiv I_1 + I_2 + I_3 \,.\end{aligned}$$

We now estimate each term. Since A is a D-operator and $B_{x,\tau}$ is a bounded set, $I_1 \leqq \lambda \sum_{p=0}^{n-1} \omega_{B_{x,\tau}} \| J_\lambda^{n(k-1)+p+1}x - J_h^k x \|^2$. Since

$$\begin{aligned}&\| J_\lambda^{n(k-1)+p+1}x - J_\lambda^{nk}x \| \leqq \sum_{j=n(k-1)+p+1}^{nk-1} \| J_\lambda^j x - J_\lambda^{j+1}x \|\\ &\leqq \lambda \sum_{j=n(k-1)+1}^{nk-1} \| A_\lambda J_\lambda^j x \| \leqq \lambda(1 - \lambda\omega_{B_{x,\tau}})^{-nk+1} n \,|||\, Ax \,|||\\ &\leqq C_{x,\tau} n\lambda \leqq C_{x,\tau} h \,,\end{aligned}$$

we have

$$I_1 \leqq \omega_{B_{x,\tau}} h \| J_\lambda^{nk}x - J_h^k x \|^2 + \text{const}\,(x, \tau)h \,,$$

by using (3.2). Also, we have

$$I_2 \leqq 2C_{x,\tau}\lambda \sum_{p=0}^{n-1} \| F(J_\lambda^{nk}x - J_h^k x) - F(J_\lambda^{n(k-1)+p+1}x - J_h^k x) \| \,.$$

Employing the uniform continuity of F on bounded sets, we can find a function $\mathscr{E}(h) \equiv \mathscr{E}(h; x, \tau)$ such that $\mathscr{E}(h) \to 0$ as $h \to 0+$ and such that

$$\sup_{n\lambda \leqq h;\, hk \leqq \tau} \| F(J_\lambda^{nk}x - J_h^k x) - F(J_\lambda^{n(k-1)+p+1}x - J_h^k x) \| \leqq \mathscr{E}(h) \,.$$

Note that $\| J_\lambda^{n(k-1)+p+1}x - J_\lambda^{nk}x \| \leqq C_{x,\tau}h$. Also,

$$I_3 \leqq | n\lambda - h | \, \| A_h J_h^{k-1}x \| \, \| J_\lambda^{nk}x - J_h^k x \| \,.$$

Consequently,

$$\begin{aligned}&\| J_\lambda^{nm}x - J_h^m x \|^2 = \sum_{k=1}^{m} \{ \| J_\lambda^{nk}x - J_h^k x \|^2 - \| J_\lambda^{n(k-1)}x - J_h^{k-1}x \|^2 \}\\ &\leqq \sum_{k=1}^{m} 2\, re \langle (J_\lambda^{nk}x - J_h^k x) - (J_\lambda^{n(k-1)}x - J_h^{k-1}x),\ F(J_\lambda^{nk}x - J_h^k x)\rangle\\ &\leqq 2\omega_{B_{x,\tau}} h \sum_{k=1}^{m} \| J_\lambda^{nk}x - J_h^k x \|^2 + \psi(\lambda, h) \,,\end{aligned}$$

where $\psi(\lambda, h) \equiv \text{const}(x, \tau)(h + \mathscr{E}(h) + m|\lambda n - h|)$ and note that

$$2\, re\langle x - y, F(x)\rangle \geqq 2\,\|x\|^2 - 2\,\|x\|\,\|y\|$$
$$= \|x\|^2 - \|y\|^2 + (\|x\| - \|y\|)^2 \geqq \|x\|^2 - \|y\|^2 .$$

Hence, for each $t \in [0, \tau]$, we can write

$$\| J_\lambda^{n[t/h]}x - J_h^{[t/h]}x \|^2 \leqq 2\omega_{B_{x,\tau}} \int_0^t \| J_\lambda^{n[s/h]}x - J_h^{[s/h]}x \|^2 ds + \psi(\lambda, h) .$$

This is a Gronwall type inequality, and so, we have that

$$\| J_\lambda^{n[t/h]}x - J_h^{[t/h]}x \|^2 \leqq \psi(\lambda, h) \exp(2\omega_{B_{x,\tau}} t) .$$

Therefore, we have

$$\begin{aligned} &\| J_\lambda^{[t/\lambda]}x - J_h^{[t/h]}x \| \\ &\leqq \| J_\lambda^{[t/\lambda]}x - J_\lambda^{n[t/h]}x \| + \sqrt{\psi(\lambda, h)}\exp(\omega_{B_{x,\tau}}\tau) . \end{aligned} \tag{3.3}$$

First, take $\lambda = \varepsilon_\mu = 2^{-\mu}$, $h = \varepsilon_\nu = 2^{-\nu}$, $m = [t/\varepsilon_\nu]$ and $n = 2^{\mu-\nu}$. In this case $\psi(\varepsilon_\mu, \varepsilon_\nu) = \text{const}(x, \tau)(\varepsilon_\nu + \mathscr{E}(\varepsilon_\nu)) \to 0$ as $\nu \to \infty$, and $|[t/\varepsilon_\mu] - 2^{\mu-\nu}[t/\varepsilon_\nu]| \leqq 2^{\mu-\nu}$. So, we see that (by (3.2))

$$\| J_{\varepsilon_\mu}^{[t/\varepsilon_\mu]}x - J_{\varepsilon_\mu}^{2^{\mu-\nu}[t/\varepsilon_\nu]}x \| = 0(\varepsilon_\nu) ,$$

and hence

$$\| J_{\varepsilon_\mu}^{[t/\varepsilon_\mu]}x - J_{\varepsilon_\nu}^{[t/\varepsilon_\nu]}x \| \leqq 0(\varepsilon_\nu) + \sqrt{\psi(\varepsilon_\mu, \varepsilon_\nu)}\exp(\omega_{B_{x,\tau}}\tau) .$$

This means that $\{J_{\varepsilon_\nu}^{[t/\varepsilon_\nu]}x\}$ is a Cauchy sequence. We then set

$$y(t; x) = \lim_{\nu\to\infty} J_{\varepsilon_\nu}^{[t/\varepsilon_\nu]}x , \qquad t \in [0, \tau] , \tag{3.4}$$

Finally, we show that the existence of the limit is independent of the sequence chosen. Let $0 \leqq t < \tau$, and $0 < \lambda \leqq h < \min\{\lambda_0, \tau - t\}$. Taking, this time $m = [t/h] + 1$ and $n = [[t/\lambda]/[t/h] + 1]$ we observe that

$$\begin{cases} mh \leqq t + h,\ n\lambda \leqq h,\ |t - n\lambda m| \leqq 2\lambda + \tau\lambda/h, \\ |[t/\lambda] - nm|\lambda \leqq 3\lambda + \tau\lambda/h,\ m|n\lambda - h| \leqq 2h + 2\lambda + \tau\lambda/h . \end{cases} \tag{3.5}$$

Similarly, as above, taking $\lambda = \varepsilon_\nu$, then letting $\nu \to \infty$, we see using (3.4) and (3.5) that

$$\| y(t; x) - J_h^{[t/h]}x \| \leqq \text{const}(x, \tau)\sqrt{3h + \varepsilon(h)} .$$

LEMMA 3.2. *Let A be a D-operator in X satisfying (I), (R) and (E).*

(a) *For every bounded set B in $D(A)$ and $\tau > 0$, there exists a number $\omega_{B,\tau} \geqq 0$ such that*

$$\| y(t; x_1) - y(t; x_2) \| \leqq \exp(\omega_{B,\tau} t) \| x_1 - x_2 \|$$

for $t \in [0, \tau]$ *and* $x_1, x_2 \in B$.

(b) *For every* $x \in D(A)$ *and* $\tau_0 > 0$, *there exists a number* $\omega_{x,\tau} \geqq 0$ *such that*

$$\| y(t; x) - y(t'; x) \| \leqq | t - t' | \exp(\omega_{x,\tau}\tau) ||| Ax |||$$

for $t, t' \in [0, \tau]$.

Proof. (a) Let B be a bounded set in $D(A)$ and $\tau_0 > 0$. Take $x_1, x_2 \in B$, then by Proposition 1.3 we have that

$$\| J_h^{[t/h]} x_1 - J_h^{[t/h]} x_2 \| \leqq (1 - h\omega_{B,\tau})^{-[t/h]} \| x_1 - x_2 \|$$

for some $\omega_{B,\tau} \geqq 0$ and h sufficiently small. Now letting $h \to +0$, we obtain (a).

(b) Let $x \in D(A)$, $\tau > 0$ and set

$$\begin{aligned} B_{x,\tau} = & \{J_h^m x;\ h \in (0, \lambda_0),\ mh \in [0, \tau]\} \\ & \cup \{x - hy;\ h \in (0, \lambda_0),\ y \in Ax,\ \| y \| \leqq ||| Ax ||| + 1\}\ . \end{aligned}$$

Then, $B_{x,\tau}$ is a bounded set by (E). Now, let $\omega_{x,\tau}$ be a constant associated with this bounded set in the sense of Proposition 1.3 and let $0 \leqq t' < t \leqq \tau$. Then, by (3.2),

$$\begin{aligned} \| J_h^{[t/h]} x - J_h^{[t'/h]} x \| & \leqq \sum_{j=[t'/h]}^{[t/h]-1} \| J_h^{j+1} x - J_h^j x \| \leqq h \sum_{j=[t'/h]}^{[t/h]-1} \| A_h J_h^j x \| \\ & \leqq | [t/h] - [t'/h] | h (1 - h\omega_{x,\tau})^{-[t/h]} ||| Ax |||\ . \end{aligned}$$

Letting $h \to +0$, we have (b).

Consequently, we have the following main theorem:

THEOREM 3.1 *If* A *is a D-operator in* X *satisfying* (*I*), (*R*) *and* (*E*). *Then there exists a semi-group* $\{T(t)\}$ *on* $\overline{D(A)}$ *such that*

$$T(t)x = \lim_{\lambda \to +0} (I - \lambda A)^{-[t/\lambda]} x \ \text{for}\ t \geqq 0\ \text{and}\ x \in D(A)\ , \tag{3.7}$$

and the convergence is uniform with respect to t *in every finite interval.*

Proof. In view of Lemma 3.1, set $T(t)x = y(t; x)$ for $t \geqq 0$ and $x \in D(A)$. First, by using Lemma 3.2 (a), we can obtain a unique extension of $T(t)$ to $\overline{D(A)}$ by continuity, we denote this extension by the same symbol $T(t)$. Then each $T(t)$ maps $\overline{D(A)}$ into itself, and also for every bounded set B in $\overline{D(A)}$ and $\tau > 0$ there exists a number $\omega_{B,\tau} \geqq 0$ such that

$$\| T(t) \|_{\mathrm{Lip}(B)} \leqq \exp(\omega_{B,\tau} t) , \qquad t \in [0, \tau] . \tag{3.8}$$

To establish the semi-group property, first take $x \in D(A)$ and $t, s \geqq 0$ with $t + s \leqq \tau$. Let $B_{x,\tau}$ be the bounded set defined in the proof of Lemma 3.1 and $N(T(s)x)$ be a bounded neighborhood of $T(s)x$ (small), and then consider the bounded set $\overline{B_{x,\tau}} \cup N(T(s)x)$. Now using Proposition 1.3 and (3.8), it is seen that $\| T(t + s)x - T(t)T(s)x \|$ can be made arbitrarily small. (3.7) was established in Lemma 3.1.

REMARK 3.1. In Theorem 3.1, (3.7) holds for $x \in \overline{D(A)}$, if either of the following conditions is satisfied:

$$R(I - \lambda A) \supset \overline{D(A)} \quad \text{for } \lambda \in (0, \lambda_0), \quad \text{or} \tag{3.9}$$

$$A \text{ is closed.} \tag{3.10}$$

In fact, if (3.9) holds, then by Proposition 1.3 $\{J_\lambda^{[t/\lambda]}\}$ is equi-Lipschitz continuous on bounded sets in $\overline{D(A)}$. Hence, Lemma 3.1 implies the convergence (3.7) for all $x \in \overline{D(A)}$. Next, assume that A is closed. Let $x \in \overline{D(A)}$, $t \in [0, \tau]$, and then choose a sequence $\{x_n\} \subset D(A)$ with $x_n \to x$. Let $B = \overline{\{x_n\}}$, then by Proposition 1.2, we see that there is a number λ_B such that if $\lambda \in (0, \lambda_B)$, then $y_n(\lambda) = J_\lambda x_n \to v_\lambda \in X$. Hence, $Ay_n(\lambda) \ni \lambda^{-1}(y_n(\lambda) - x_n) \to \lambda^{-1}(v_\lambda - x)$. This means that

$$\lambda^{-1}(v_\lambda - x) \in Av_\lambda, \text{ i.e., } x \in (I - \lambda A)v_\lambda \subset R(I - \lambda A) .$$

Therefore, Proposition 1.3 implies that $\{J_\lambda^{[t/\lambda]}\}$ is equi-Lipschitz continuous on B, and so, Lemma 3.1 implies the convergence (3.7) for the x.

4. Differentiability of the Constructed Semi-Groups. The differentiability of the semi-group obtained by Theorem 3.1 is investigated in this section. The central part of the arguments is based on the results of Kato [4] and [5]. Throughout this section X is assumed to have a uniformly convex dual.

Let A be a D-operator in X satisfying (I), (R) and (E). Set $\varepsilon_n = 2^{-n}$ and $I_r = [0, r]$ for $r = 1, 2, 3, \cdots$ and define $J_n = (I - \varepsilon_n A)^{-1}$ and $A_n = \varepsilon_n^{-1}[J_n - I]$ for n with $\varepsilon_n \in (0, \lambda_0)$.

In view of (3.2), we note that for each r,

$$\| A_n J_n^{[t/\varepsilon_n]} x \| \leqq (1 - \varepsilon_n \omega_{B_{x,r}})^{-[t/\varepsilon_n]-1} ||| Ax ||| ,$$

n sufficiently large and $t \in [0, r]$, for the bounded set

$$B_{x,r} = \{J_n^{[t/\varepsilon_n]}x;\ t \in [0, r],\ n \text{ sufficiently large}\} \cup \{x - \varepsilon_n y;\ n \text{ large},\ y \in Ax,\ \| y \| \leqq ||| Ax ||| + 1\} .$$

Since

$$(1-\varepsilon_n\omega_{Bx,r})^{-1} = 1+\omega_{Bx,r}\varepsilon_n(1-\varepsilon_n\omega_{Bx,r})^{-1} \\ \leqq \exp(\omega_{Bx,r}\varepsilon_n(1-\varepsilon_n\omega_{Bx,r})^{-1})$$

for n sufficiently large and $t\in[0, r]$. Hence,

$$\| A_nJ_n^{[t/\varepsilon_n]}x\| \leqq \exp(\omega_{Bx,r}(r+\varepsilon_n)(1-\varepsilon_n\omega_{Bx,r})^{-1})\||Ax\||$$

for n sufficiently large and $t\in[0, r]$. Therefore, if we set $f_n(t; x) = A_nJ_n^{[t/\varepsilon_n]}x$ for $t\geqq 0$ and $x\in D(A)$, then $f_n(t; x)\in AJ_n^{[t/\varepsilon_n]+1}x$, and

(4.1) for every r, $\| f_n(t; x)\|$ is uniformly bounded with respect to n sufficiently large and $t\in[0, r]$.

Also, since

$$\int_0^{[t/\varepsilon_n]\varepsilon_n} A_nJ_n^{[s/\varepsilon_n]}x\,ds = \varepsilon_n\sum_{k=1}^{[t/\varepsilon_n]} A_nJ_n^{(k-1)}x \\ = [J_n^{[t/\varepsilon_n]}-I]\,J_nx + \varepsilon_n\{A_nx - A_nJ_n^{[t/\varepsilon_n]}x\}\,,$$

we have

$$\| [J_n^{[t/\varepsilon_n]}-I]J_nx - \int_0^t f_n(s; x)ds\| = O(\varepsilon_n)\,. \tag{4.2}$$

The main result of this section is the following:

THEOREM 4.1. *Let A be a demi-closed D-operator satisfying (I), (R) and (E), and $\{T(t)\}$ be the semi-group on $\overline{D(A)}$ obtained by Theorem 3.1. Then for $x\in D(A)$,*

(i) *$T(t)x\in D(A)$ for $t\geqq 0$,*

(ii) *there exists a function $f(\cdot\,; x)$ on $[0, \infty)$ such that*

$$f(t; x)\in \hat{A}T(t)x$$

for almost all $t\geqq 0$, where $\hat{A}x = \{y\in \overline{co}Ax;\ \| y\|\leqq \||Ax\||\}$, and

$$T(t)x - x = \int_0^t f(s; x)ds \qquad t\geqq 0\,. \tag{4.3}$$

Proof. Take $x\in D(A)$ and p with $1<p<+\infty$. Set $f_n(t; x) = A_nJ_n^{[t/\varepsilon_n]}x$, then by (4.1) $\{f_n(\cdot\,; x)|_{I_r};\ n$ sufficiently large$\}$ forms a bounded set of $L^p(I_r; X)$ for integer r. Thus by moving r and using the diagonal process, we find a subsequence $\{q\}\subset\{n\}$ and a function $f(\cdot\,; x)$ on $[0, \infty)$ such that $f_q(\cdot\,; x)|_{I_r}$ converges weakly to $f(\cdot\,; x)|_{I_r}$ in $L^p(I_r; X)$ for each integer r. Hence,

$$x^*\int_0^t f_q(s; x)ds \longrightarrow x^*\int_0^t f(s; x)ds$$

for all $x^*\in X^*$ and $t\geqq 0$. Thus (4.3) follows from (4.2). Write $V(t)$

for the set of all weak cluster points of $\{f_n(t; x); n\}$ for $t \geqq 0$, then Lemma 3.1 and Proposition 1.7 (a) imply that $T(t)x \in D(A)$, $V(t) \neq \emptyset$, and $V(t) \subset AT(t)x$ for $t \geqq 0$. Hence, by the same argument as in Kato [5; Lemma 8.2] we see that $f(t; x) \in \overline{co}\ AT(t)x$ for almost all $t \geqq 0$. And, in a similar way to Kato [5; Lemma 6.2], $\|f(t; x)\| \leqq |||AT(t)x|||$ for almost all $t \geqq 0$. Thus, it follows that $f(t; x) \in \hat{A}T(t)x$ for almost all $t \geqq 0$.

REMARK 4.1. Let A be a demi-closed D-operator in X satisfying (I), (R) and (E), and $\{T(t)\}$ be the semi-group obtained by Theorem 3.1, then $\{T(t)|_{D(A)}; t \geqq 0\}$ forms a semi-group on $D(A)$ by the above theorem. By (4.3), we see that the infinitesimal generator A_0 of $\{T(t)|_{D(A)}\}$ is densely defined in $D(A)$.

In view of these results and Proposition 1.6, we have the following.

THEOREM 4.2. *Let A be a maximal $(D, \{\omega_n\})$-operator on $\overline{D(A)}$ satisfying (I), (R) and (E). Then there is a uniquely determined semi-group $\{T(t)\}$ on $D(A)$ such that for each $x \in D(A)$,*

$$(d/dt)T(t)x \in A^0T(t)x \qquad \text{for almost all } t \geqq 0 .$$

THEOREM 4.3. *If A is a single-valued, demi-closed D-operator in X satisfying (I), (R) and (E). Then there is a uniquely determined semi-group $\{T(t)\}$ on $D(A)$ such that*

(a) *for $x \in D(A)$, $AT(t)x$ is weakly continuous in $t \geqq 0$ and*

$$T(t)x - x = \int_0^t AT(s)x\, ds \qquad \text{for } t \geqq 0 , \tag{4.4}$$

(b) *A is the weak infinitesimal generator and the infinitesimal generator A_0 is densely defined in $D(A)$.*

Proof. Using the notation in the proof of Theorem 4.1, we have that $V(t)$ is a singleton, since A is single-valued. And thus, by Proposition 1.7, w-$\lim f_n(t; x) = AT(t)x$ for $t \geqq 0$. The strong continuity of $T(t)x$ and the boundedness of $AT(t)x$ give that $AT(t)x$ is weakly continuous in $t \geqq 0$. Finally (4.4) follows directly from (4.2).

COROLLARY 4.1. *If A is a demi-closed D-operator in X satisfying (I), (R) and (E), and A^0 is single-valued, then there exists a unique semi-group $\{T(t)\}$ on $D(A)$ such that for $x \in D(A)$, $(d/dt)T(t)x = A^0T(t)x$ for almost all $t \geqq 0$.*

Proof. In this case, note that we have that $\hat{A} \equiv A^0$, where $\hat{A}$ is defined in Theorem 4.1 by $\hat{A}x = \{y \in \overline{co}Ax;\ \|y\| = |||Ax|||\}$.

COROLLARY 4.2. *If A is a demi-closed D-operator in X satisfying (I) and (R) and $\|J_\lambda x\| \leqq (1+M\lambda)\|x\| + N\lambda$ for $\lambda \in (0, \lambda_0)$, $x \in R_\lambda$, where M and N are nonnegative, then there is a semi-group $\{T(t)\}$ on $D(A)$ such that $(d/dt)T(t)x \in \overline{co}AT(t)x$ for almost all $t \geqq 0$ and $\|T(t)x\| \leqq e^{Mt}(\|x\| + Nt)$ for $t \geqq 0$.*

Proof. By Proposition 1.4, A satisfies condition (E) and also we have that $\|J_\lambda^{[t/\lambda]}x\| \leqq (1 + M\lambda)^{[t/\lambda]}(\|x\| + Nt)$, hence using Theorem 4.1 we have the assertion.

COROLLARY 4.3. *If A is a single-valued, demi-closed D-operator in X satisfying (I) and (R) and $\sup \|Ax\| = N < +\infty$, then A is the weak infinitesimal generator of a semi-group $\{T(t)\}$ on $D(A)$ such that $\|T(t)x\| \leqq \|x\| + Nt$ for $t \geqq 0$ and $x \in D(A)$ and*

$$\sup\{\|AT(t)x\|;\ t \geqq 0,\ x \in D(A)\} \leqq N\,.$$

Proof. Employ Proposition 1.4.

In the remainder of this section, we consider the case in which X is uniformly convex.

LEMMA 4.1. *Let A be a demi-closed D-operator in X satisfying (I), (R) and (E) such that A^0 is a single-valued operator with $D(A^0) = D(A)$. Then if $\{T(t)\}$ is the semi-group on $D(A)$ obtained by Theorem 4.1, we have for $x \in D(A)$,*

(a) *$|||AT(t)x|||$ is of bounded variation on every finite interval and has no positive jumps,*

(b) *the right derivative $D^+T(t)x$ exists and is strongly right-continuous in t, and $D^+T(t)x = A^0T(t)x$ for $t \geqq 0$,*

(c) *$A^0T(t)x$ is strongly continuous except possibly at a c countable number of points t.*

Proof. (a) Take $x \in D(A)$. Then by the same argument as in Kato [5; Lemma 6.6] we obtain that

$$e^{-\omega_{B,\tau}t}\,|||AT(t)x||| \leqq e^{-\omega_{B,\tau}t}\,|||AT(r)x|||$$

for all r and t with $0 \leqq r \leqq t \leqq \tau$. Thus, $|||AT(t)x|||$ is of bounded variation.

(b) Take $x \in D(A)$ and $t \geqq 0$. Choose a sequence $t_k \downarrow t$. Then by the proof of Kato [5; Theorem 7.5] we see that $\{A^0T(t_k)x\}$ contains a subsequence which converges strongly to $A^0T(t)x$. So, $A^0T(t)x$ is strongly right-continuous in t. But, since $T(t)x - x = \int_0^t A^0 T(s)x\,ds$ by Theorem 4.1, it follows that $D^+T(t)x = A^0T(t)x$ for each t.

(c) By (a) $\| A^0T(t)x \| = ||| AT(t)x |||$ is continuous except for a countable number of points t. In order to show that $A^0T(t)x$ is continuous except for those points, it suffices to repeat the same argument as in (b) with $t_k \uparrow t$. But the continuity at t of $\| A^0T(t)x \|$ assures that $\lim_k \| A^0T(t_k)x \| = \| A^0T(t)x \|$. Thus the uniform convexity implies that $A^0T(t)x$ is strongly continuous at the t.

Consequently, we have the following:

THEOREM 4.4. *Let X be uniformly convex. If A is a demi-closed D-operator in X satisfying (I), (R) and (E) such that A^0 is a single-valued operator with $D(A^0) = D(A)$, then A^0 is the infinitesimal generator of a unique semi-group $\{T(t)\}$ on $D(A)$ such that for $x \in D(A)$, $D^+T(t)x = A^0T(t)x$ for $t \geqq 0$, and $D^+T(t)$ is strongly right-continuous in $t \geqq 0$.*

The following results are the direct consequences of the above theorem.

COROLLARY 4.4. *Let X be uniformly convex. If A satisfies the assumptions of Theorem 4.2, then A^0 is the infinitesimal generator of a unique semi-group $\{T(t)\}$ on $D(A)$ such that for $x \in D(A)$, $T(t)x$ is strongly right-continuously differentiable in t and $D^+T(t)x = A^0T(t)x$ for $t \geqq 0$.*

COROLLARY 4.5. *Let X be uniformly convex. If A is a single-valued, demi-closed D-operator in X satisfying (I), (R) and (E), then A is the infinitesimal generator of a unique semi-group $\{T(t)\}$ on $D(A)$ such that for $x \in D(A)$, $T(t)x$ is strongly right-differentiable in t and $D^+T(t)x = AT(t)x$ for each $t \geqq 0$.*

REMARK 4.2. Let X be uniformly convex. If A is a closed dissipative operator in X satisfying (R), then A^0 is the infinitesimal generator of a unique semi-group $\{T(t)\}$ of contractions on $D(A)$ such that for $x \in D(A)$, $T(t)x$ is strongly right-continuously differentiable in t and $D^+T(t)x = A^0T(t)x$ for $t \geqq 0$. For details, see [10].

APPENDIX

A.1. After this paper was submitted for publication, Crandall and Liggett gave (in "Generation of semigroups of nonlinear transformations on general Banach spaces", to appear) a new method for constructing a semigroup of nonlinear contractions in a general Banach space. The main results in their paper can be extended straightforwardly to our case. As was stated in § 1, Propositions 1.2 and 1.3 are valid for general Banach spaces. Using these propositions in a similar way to their proof, we can obtain the assertion of Theorem 3.1, without assuming that X^* is uniformly convex. Also, we can obtain a similar result to theirs on the differentiability of semigroups of nonlinear contractions. For details, we shall publish elsewhere.

A.2. We did not give in the body of this paper any examples of D-operator satisfying conditions (I), (R) and (E). We state here a simple example of a D-operator which is not necessarily a dissipative operator.

Let Ω be a bounded domain with smooth boundary in R^N and let us consider the Cauchy problem

$$(\text{A.1})\qquad \begin{aligned}(\partial/\partial t)u_1 &= \Delta u_1 + \Phi u_2\ ,\\ (\partial/\partial t)u_2 &= \Delta u_2\end{aligned}$$

with the initial condition

$$u_1(0, s) = u_1(s)\ ,$$

$$u_2(0, s) = u_2(s)\ ,$$

over the Hilbert space $H = L_2(\Omega) \times L_2(\Omega)$ with the inner product

$$\langle u, v\rangle = \langle u_1, v_1\rangle + \langle u_2, v_2\rangle,\ u = \begin{pmatrix}u_1\\ u_2\end{pmatrix},\quad v = \begin{pmatrix}v_1\\ v_2\end{pmatrix}.$$

It is well-known that the operator Δ with domain $H^2(\Omega) \cap H_0^1(\Omega)$ is m-dissipative. We then assume that the operator Φ is locally bounded on X and Lipschitz continuous on bounded sets.

Now, let us define an operator A in H by the relation

$$Au = \begin{pmatrix}\Delta u_1 + \Phi u_2\\ \Delta u_2\end{pmatrix} \text{ for } u \in D(A) = \left\{u = \begin{pmatrix}u_1\\ u_2\end{pmatrix};\ u_1,\ u_2 \in H^2(\Omega) \cap H_0^1(\Omega)\right\}.$$

Then the problem (A.1) is understood as the ACP for A in the space H.

In the following, we demonstrate that A so defined is a demiclosed D-operator satisfying conditions (I), (R) and (E).

(a) Let B be any bounded set in H and $u,\ v \in B \cap D(A)$. Then

$$\begin{aligned}
&\langle Au - Av, u - v\rangle \\
&= \langle \varDelta(u_1 - v_1), u_1 - v_1\rangle + \langle \Phi u_2 - \Phi v_2, u_1 - v_1\rangle \\
&\quad + \langle \varDelta(u_2 - v_2), u_2 - v_2\rangle \\
&\leqq - \| u_1 - v_1 \|_1^2 - \| u_2 - v_2 \|_1^2 + \| \Phi u_2 - \Phi v_2 \| \, \| u_1 - v_1 \| \\
&\leqq \gamma_B \| u_2 - v_2 \| \, \| u_1 - v_1 \| \leqq \gamma_B/2 \cdot (\| u_1 - v_1 \|^2 + \| u_2 - v_2 \|^2) \\
&\leqq \omega_B \| u - v \|^2 ,
\end{aligned}$$

where γ_B is the smallest Lipschitz constant of Φ on the bounded set B. Hence, A is a D-operator.

(b) Let $v \in X$, $\lambda > 0$ and let us consider the equation

$$u - \lambda Au = v , \tag{A.2}$$

or equivalently,

$$\begin{cases} (I - \lambda\varDelta)u_1 - \lambda\Phi u_2 = v_1 , \\ \qquad\qquad (I - \lambda\varDelta)u_2 = v_2 , \end{cases} \tag{A.3}$$

Since $\varDelta$ is m-dissipative, we obtain a unique solution

$$u_2 = (I - \lambda\varDelta)^{-1} v_2 \tag{A.4}$$

of the second equation of (A.3). Substituting this into the first equation and using the m-dissipativity of $\varDelta$, we get

$$u_1 = (I - \lambda\varDelta)^{-1} [v_1 + \lambda\Phi(I - \lambda\varDelta)^{-1} v_2] . \tag{A.5}$$

Therefore, $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ is the unique solution of (A.2) and since $\lambda > 0$ and $v \in X$ were arbitrary, we see that $I - \lambda A$ is injective and $R(I - \lambda A) = H$ for all $\lambda > 0$. Hence, A satisfies (I) and (R).

(c) From (A.4) and $(A.5)$ it follows that

$$\begin{aligned}
&\| u \|^2 = \| u_1 \|^2 + \| u_2 \|^2 \\
&\leqq \| v_1 \|^2 + 2\lambda \| v_1 \| \, \| \Phi(I - \lambda\varDelta)^{-1} v_2 \| + \lambda^2 \| \Phi(I - \lambda\varDelta)^{-1} v_2 \|^2 + \| v_2 \|^2 \\
&\leqq \| v \|^2 + 2\lambda \| v \| \, \| \Phi(I - \lambda\varDelta)^{-1} v_2 \| + \lambda^2 \| \Phi(I - \lambda\varDelta)^{-1} v_2 \|^2 \\
&= (\| v \| + \lambda \| \Phi(I - \lambda\varDelta)^{-1} v_2 \|)^2 ,
\end{aligned}$$

or

$$\| J_\lambda v \| \leqq \| v \| + \lambda \| \Phi(I - \lambda\varDelta)^{-1} v_2 \| ,$$

where $u = J_\lambda v$, $\lambda > 0$ and $v \in X$. Now, let $v \in H$, $\tau > 0$, $\lambda > 0$, and let $n\lambda \in [0, \tau]$, then

$$\begin{aligned}
\| J_\lambda^n v \| &\leqq \| J_\lambda^{n-1} v \| + \lambda \| \Phi(I - \lambda\varDelta)^{-1} [J_\lambda^{n-1} v]_2 \| \\
&\leqq \| J_\lambda^{n-2} v \| + \lambda \{\| \Phi(I - \lambda\varDelta)^{-1} [J_\lambda^{n-2} v]_2 \| + \| \Phi(I - \lambda\varDelta)^{-1} [J_\lambda^{n-1} v]_2 \|\}
\end{aligned}$$

and inductively,

$$\leqq \| v \| + \lambda \sum_{i=0}^{n-1} \| \Phi(I - \lambda \Delta)^{-1} [J_\lambda^i v]_2 \| ,$$

where $[J_\lambda^i v]_2$ means the second component of $J_\lambda^i v$. But, since $[J_\lambda^i v]_2 = (I - \lambda \Delta)^{-i} v_2$,

$$\begin{aligned} \| J_\lambda^n v \| &\leqq \| v \| + \lambda \sum_{i=1}^{n} \| \Phi(I - \lambda \Delta)^{-i} v_2 \| \\ &\leqq \| v \| + \tau \sup_{0 \leqq i\lambda \leqq \tau} \| \Phi(I - \lambda \Delta)^{-i} v_2 \| . \end{aligned}$$

Let B be any bounded set in H. Since $(I - \lambda \Delta)^{-1}$ is a contraction on H, the set $\{(I - \lambda \Delta)^{-i} v_2 ; v \in B, 0 \leqq i\lambda \leqq \tau\}$ is bounded in $L_2(\Omega)$. On the other hand, Φ maps bounded sets into bounded sets by assumption, and hence

$$\sup \{\| \Phi((I - \lambda \Delta)^{-i} v_2 \| ; v \in B, 0 \leqq i\lambda \leqq \tau\} = M_{B,\tau} < +\infty .$$

Consequently,

$$\| J_\lambda^n v \| \leqq \sup \{\| v \| ; v \in B\} + \tau M_{B,\tau}$$

for $v \in B$, $\lambda > 0$ and $n\lambda \in [0, \tau]$, which means that A satisfies condition (E).

(d) Finally, we show that A is demi-closed. Assume that $u^{(n)} \in D(A)$, $u^{(n)} \to u$ and that $Au^{(n)} \rightharpoonup v$ in H. Then, $u_i^{(n)} \to u_i$, $i = 1, 2$, $\Phi u_2^{(n)} \to \Phi u_2$, and $\Delta u_2^{(n)} \rightharpoonup v_2$ in $L_2(\Omega)$. Since the closed linear operator Δ is demi-closed, we have that $v_2 = \Delta u_2$. Also, $\Delta u_1^{(n)} \rightharpoonup v_1 - \Phi u_2$; hence, $v_1 - \Phi u_2 = \Delta u_1$. Consequently, $v = Au$. This means that A is demi-closed.

From the above, it can be seen that other D-operators can be exhibited by replacing the operator Δ by any m-dissipative operator satisfying the assumption of Proposition 1.4 (a). Also, we can consider unbounded operators Φ by restricting the Hilbert space $H = H_1 \times H_2$ so that Φ is a locally bounded, locally Lipschitz continuous operator on a Hilbert space H_1 into another Hilbert space H_2.

References

1. H. Brezis and A. Pazy, *Semi-groups of nonlinear contractions on convex sets*, J. Functional Analysis, **6** (1970), 237–281.
2. M. G. Crandall and A. Pazy, *Semi.groups of nonlinear contractions and dissipative sets*, Jour. Func. Anal., **3** (1969), 376–418.
3. Y. Kōmura, *Differentiability of nonlinear semi-groups*, Jour. Math. Soc. Japan, **21** (1969), 375–402.
4. T. Kato, *Nonlinear semi-groups and evolution equations*, Jour. Math. Soc. Japan, **19** (1967), 503–520.
5. ———, *Accretive operators and nonlinear evolution equations in Banach spaces*, Proc. Symposium Nonlinear Func. Anal., Amer. Math. Soc., (1970), 138–161.

6. I. Miyadera, *On the convergence of nonlinear semi-groups*, Tôhoku Math. Jour., **21** (2), (1969), 221-236.
7. ———, *On the convergence of nonlinear semi-groups, II*, Jour. Math. Soc. Japan, **21** (1969), 403-412.
8. I. Miyadera and S. Oharu, *Approximation of semi-groups of nonlinear operators*, Tôhoku Math. Jour., **22** (1970), 24-47.
9. S. Oharu, *Note on the representation of semi-groups of nonlinear operators*, Proc. Japan Acad., **42** (1966), 1149-1154.
10. ———, *On the generation of semi-groups of nonlinear contractions*, J. Math. Soc. Japan, **22** (4), (1970), 526-550.

Received May 25, 1970 and in revised form April 19, 1971.

GEORGETOWN UNIVERSITY, WASHINGTON, D.C.,
WASEDA UNIVERSITY, TOKYO

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

FINITE DIMENSIONAL TORSION FREE RINGS

THOMAS J. CHEATHAM

In the class of rings with zero singular left ideal, several characterizations of rings with finite left Goldie dimension are given. They include: the direct limit of torsion free modules is torsion free; the direct limit of torsion free injective modules is injective; each absolutely pure torsion free module is injective; each module has a unique (up to isomorphism) torsion free covering module. The latter result gives a converse, in a special case, to a theorem of Mark Teply.

Throughout, R will denote an associative ring with identity and module, without further qualification, will mean unitary left R-module. For a module E, we use $S \subseteqq' E$ to denote that S is a large submodule of E[4, p. 60]; $Z(E)$ will denote the singular submodule of E, which consists of those elements in E whose annihilators are large left ideals in R.

DEFINITION 1. A module E is torsion free if $Z(E) = (0)$ and if $Z(R) = (0)$ we say R is a torsion free ring.

A submodule S of a module E is closed in E if $S \subseteqq' T \subseteqq E$ implies $T = S$. The following facts are easily verified.

LEMMA 1. *Let S be a submodule of a module E.*
(a) *If $Z(E/S) = (0)$, S is closed in E.*
(b) *If $Z(E) = (0)$, S is closed in E if and only if $Z(E/S) = (0)$.*

Proof. See Lemma 2.3 in [8].

DEFINITION 2. A module E has finite (Goldie) dimension if it contains no infinite direct sum of nonzero submodules. If the module R has finite dimension we call R a finite dimensional ring and write dim R is finite.

1. Torsion Free Rings.

Over an integral domain the direct limit of torsion free modules is torsion free. In this section we show that, in the class of torsion free rings, this property characterizes the finite dimensional rings. We also give two noetherian-like characterizations of such rings.

We record a theorem of F. Sandomierski [7] for easy reference.

THEOREM S. *Let $Z(R) = (0)$, and Q the maximal left quotient* *ring of* R[4, p. 106], *then the following are equivalent.*

(a) $QI = Q$ *for every* $I \subseteq' R$.

(b) *For* $I \subseteq' R$ *there are* $a_1, a_2, \cdots a_n \in I$ *such that* $\sum Ra_i \subseteq' R$.

(c) *For* $I \subseteq R$ *there are* $a_1, a_2, \cdots a_n \in I$ *such that* $\sum Ra_i \subseteq' I$.

(d) dim R *is finite.*

(e) Q *is a semisimple* (artinian) *ring.*

(f) Ker $(R \otimes_R E \to Q \otimes_R E) = Z(E)$ *for every module* E.

Sandomierski [7, Th. 2.5, p. 118] noted that if R has finite dimension the direct sum of torsion free injective modules is injective. If $Z(R) = (0)$ the converse is also known. In fact, under this assumption, it follows from Teply [10, Th. 2.1, p. 451] that $\dim R$ is finite if and only if any countable direct sum of torsion free injective modules is injective.

A set $\mathscr{S}$ of submodules of a module is directed if given $X, Y \in \mathscr{S}$ there is a $Z \in \mathscr{S}$ such that $X \cup Y \subseteq Z$. Clearly the union of a directed set of submodules is a submodule.

We will make use of the following lemma which is an unpublished remark of M. Teply.

LEMMA 2. *Let* $Z(R) = (0)$. *The union of a directed set of closed submodules of a torsion free module is a closed submodule if and only if dim* R *is finite.*

Proof. Assume that dim R is finite then, to show that the condition is necessary, we proceed as in [9, Prop. 2.1 (3)] using Theorem $S(b)$.

Conversely, let $E_1, E_2, \cdots$ be a countable family of torsion free injective modules. Then $E = \bigoplus_{n=1}^{\infty} E_n$ is a torsion free module so has a torsion free injective hull $I(E)$. But E can be written as the union of the chain $S_1 \subseteq S_2 \subseteq \cdots$ of injective (hence closed) submodules of $I(E)$, where $S_n = \bigoplus_{i=1}^{n} E_i$. Hence E is closed in $I(E)$. But $E \subseteq' I(E)$ so $E = I(E)$, i.e., E is injective. By a remark above dim R is finite.

The following useful result is well-known and trivial to prove.

LEMMA 3. *Let* $f, g \in Hom_R(E, F)$ *where* $Z(F) = (0)$.

(1) *If* f *and* g *agree on a large submodule of* E *then* $f = g$.

(2) *If* E *is injective,* $f(E)$ *is a direct summand of* F.

THEOREM 1. *Let* $Z(R) = (0)$. *Then the following statements are equivalent.*

(1) dim R *is finite.*

(2) *The direct limit of torsion free modules is torsion free.*

(3) *The direct limit of torsion free injective modules is injective.*

Proof. (1) implies (2). Let $\{E, f_a\}$ be the direct limit of the directed system of torsion free modules $\{E_a; f_a^b, A\}$. Then $E = \bigcup_{a \in A} \operatorname{Im} f_a$, so to show E is torsion free it suffices to show $\operatorname{Im} f_a$ is torsion free for each $a \in A$. But $\operatorname{Im} f_a \cong E_a/\operatorname{Ker} f_a$, and $\operatorname{Ker} f_a = \bigcup_{b \geqq a} \operatorname{Ker} f_a^b$. Furthermore, for each $a \in A$, $\{\operatorname{Ker} f_a^b \mid b \geqq a, b \in A\}$ is a directed set of submodules of the torsion free module E_a, and since $\operatorname{Im} f_a^b \subseteqq E_b$, $\operatorname{Ker} f_a^b$ is closed in E_a. By Lemma 2 $\operatorname{Ker} f_a$ is closed in E_a and hence by Lemma 1(b) $\operatorname{Im} f_a$ is torsion free. Hence (1) implies (2).

(2) implies (1). Let $\{C_a: a \in A\}$ be a directed set of closed submodules of a torsion free module E. For $a, b \in A$ such that $C_a \subseteqq C_b$ define a function $f_a^b: E/C_a \to E/C_b$ by $f_a^b(x + C_a) = x + C_b$. Clearly $f_a^b \in \operatorname{Hom}_R(E/C_a, E/C_b)$ and one easily checks that $\{E/C_a; f_a^b, A\}$ is a directed system of torsion free modules with direct limit $E/\bigcup_{a \in A} C_a$. So from (2) $E/\cup C_a$ is torsion free and hence (Lemma 1(a)) $\bigcup_{a \in A} C_a$ is closed in E. So (1) follows from Lemma 2.

(1) implies (3). Let $\{E, f_a\}$ be the direct limit of a directed system of torsion free injective modules $\{E_a; f_a^b, A\}$. Let I be any left ideal of R and $h \in \operatorname{Hom}_R(I, E)$. By Theorem $S(c)$ there is a finitely generated left ideal J of R such that $J \subseteqq' I$. From condition (2) we know that E is torsion free so by Lemma 3(1) we see that an extension of $f' = f|_J: J \to E$ to all of R will give the desired extension of f. To see that f' can be extended to R we proceed as follows.

Let $p: F \to J$ be an R-homomorphism of a finitely generated free module F onto J and identify J with $F/\operatorname{Ker} p$. It follows easily from Theorem S that F has finite dimension and that any submodule of F has a finitely generated large submodule. Thus let K be a finitely generated large submodule of $\operatorname{Ker} p$.

Now $f' \circ p(F)$ is a finitely generated submodule of $E = \bigcup_{a \in A} \operatorname{Im} f_a$ so there's an $a \in A$ such that $f' \circ p(F) \subseteqq \operatorname{Im} f_a$. Hence, by the projectivity of F, there's an $h \in \operatorname{Hom}_R(F, E_a)$ such that $f_a \circ h = f' \circ p$. Then $h(\operatorname{Ker} f' \circ p) \subseteqq \operatorname{Ker} f_a = \bigcup_{b \geqq a} \operatorname{Ker} f_a^b$ and since K is finitely generated $h(K) \subseteqq \operatorname{Ker} f_a^b$ for some $b \geqq a$, i.e., $f_a^b \circ h(K) = (0)$. Therefore $f_a^b \circ h$ induces an R-homomorphism g from $F/\operatorname{Ker} p = J$ into E_b. But E_b is injective so g can be extended to $g^* \in \operatorname{Hom}_R(R, E_b)$. Define f^* from R to E by $f^* = f_b \circ g^*$. Then f^* is an extension of f' so the desired extension of f.

(3) implies (1). Clearly (3) implies that any direct sum of torsion free injective modules is injective. So (1) follows by a previous remark. This completes the proof.

As a corollary we give an easy proof of (d) implies (e) in Theorem S.

COROLLARY. *A finite dimensional torsion free ring R has a semisimple* (artinian) *maximal left quotient ring Q.*

Proof. It suffices to show that every left ideal of Q is a direct summand of Q (as a Q-module). Since Q is von Neumann regular any finitely generated left ideal of Q is a direct summand. Hence any such ideal is left R-injective as Q is. But $Z({}_RQ) = (0)$ and it follows, via a direct limit argument and Theorem 1(3), that any left ideal of Q is R-injective, hence Q-injective. Thus any left ideal of Q is a direct summand.

B. Maddox [5] calls a module M absolutely pure if for every module E containing M as a submodule the sequence $0 \to G \otimes M \to G \otimes E$ is exact for every right R-module G. He showed that any direct sum of absolutely pure modules is absolutely pure. C. Megibben [6, Th. 3, p, 564] characterized left noetherian rings by the property "each absolutely pure module is injective." This result was also obtained indepently by Edgar Enochs.

We have a corresponding characterization of finite dimensionality in the class of torsion free rings.

THEOREM 2. *Let $Z(R) = (0)$. Then $\dim R$ is finite if and only if each absolutely pure torsion free module is injective.*

Proof. Assume that $\dim R$ is finite, and let E be an absolutely pure torsion free module and f an R-homomorphism of a left ideal I of R into E. Let J be a finitely generated left ideal of R such that $J \subseteq' I$. As remarked above it suffices to extend f', the restriction of f to J, to all of R. But this can be done by [6, Cor. 2, p. 562].

Conversely, it suffices to show that the direct sum of torsion free injective modules is injective. But any such sum is torsion free and absolutely pure, hence injective.

2. Torsion free covers.

The main result of this section (Theorem 3) gives a converse to [9, Th. 2.4, p. 459], in a special case. We begin with a definition.

DEFINITION 3. A torsion free cover of a module E is a homomorphism g from a torsion free module $T(E)$ onto E such that:

(1) Ker g contains no nonzero closed submodule of $T(E)$,

(2) given $f \in \mathrm{Hom}_R(F, E)$ where F is torsion free there is an $h \in \mathrm{Hom}_R(E, T(E))$ such that $g \circ h = f$.

This definition was given initially for modules over an integral domain by E. Enochs [3] who proved that, in this case, every module has a unique (up to isomorphism) torsion free covering module. B. Banaschewski [1, p. 59] gave the following construction for the cover of a module E over an integral domain R with quotient field K: $T(E) = \{f \in \mathrm{Hom}_R(K, I(E)) \mid f(1) \in E\}$; $g(f) = f(1)$.

M. Teply [8, p. 449] generalized the notion of a torsion free cover to a hereditary torsion theory $(\mathscr{T}, \mathscr{F})$ [2] of R-modules. He proved that each module has a $\mathscr{T}$-torsion free cover if $R \in \mathscr{F}$ and the direct sum of $\mathscr{T}$-torsion free injective modules is injective.

The Goldie torsion theory $(\mathscr{G}, \mathscr{F})$ is the torsion theory whose torsion class $\mathscr{G}$ is generated by all factor modules B/A where A is a large submodule of B. The Goldie torsion free class $\mathscr{F}$ is precisely the class of torsion free modules given by Definition 1. Teply's result shows that $\mathscr{G}$-torsion free covers exist if $Z(R) = (0)$ and dim R is finite. We prove the converse and show that Banaschewski's construction has an obvious analogy in this case.

THEOREM 3. *Let R be a ring with identity and maximal left quotient ring Q. If every left R-module has a torsion free cover then $Z(R) = (0)$ and dim R is finite. Moreover, the evaluation map from $T(E) = \{f \in \mathrm{Hom}_R(Q, I(E)) \mid f(1) \in E\}$ onto E is a torsion free cover of E.*

Proof. Let $(0) \to \mathrm{Ker}\, g \to T(R) \xrightarrow{g} R \to (0)$ be a torsion free cover of the module R. Since R is projective this sequence splits and hence $\mathrm{Ker}\, g$ is closed in $T(R)$. Then $\mathrm{Ker}\, g = (0)$ as g is a cover of R so $R \cong T(R) \in \mathscr{F}$, i.e., $Z(R) = (0)$.

To show that dim R is finite it suffices, by Theorem S, to show that Q is a semisimple ring. But, since $Z(R) = (0)$, Q is von Neumann regular so it suffices to show that Q is a finite dimensional ring. If not, there is an infinite set of nonzero elements $\{x_i \colon i \in I\}$ of Q such that the sum $B = \sum Qx_i$ is direct and a proper large left ideal of Q. Then $Q/B \neq (0)$ has a torsion free cover (as an R-module), say $g : F \to Q/B$.

Since $Z({}_RQ) = (0)$ the natural R-homomorphism $p : Q \to Q/B$ factors thru g, i.e., there exist $h \in \mathrm{Hom}_R(Q, F)$ such that $g \circ h = p$. Then $h(B) \subseteqq \mathrm{Ker}\ g$ and $p \neq 0$ implies $h \neq 0$ so by Lemma 3(1) $h(B) \neq (0)$ — we must note that B is a large R-submodule of Q. Then $h(Qx_j) \neq (0)$ for some index $j \in I$. By Lemma 3(2) $h(Qx_j)$ is a direct summand of F. Thus we have a nonzero closed submodule of F contained in

Ker g. This gives a contradiction so Q is a finite dimensional ring, hence semisimple.

Since Q is a semisimple ring it follows that any Q-module is torsion free when considered as an R-module. Hence any R-submodule of such a module is a torsion free R-module. Conversely, it follows from Theorem $S(f)$ that any torsion free R-module is an R-submodule of a Q-module (i.e., is Q-extendible in Banaschewski's terminology). Banaschewski [1, p. 63] established the existence of Q-extendible coverings.

Note that if $q_1, q_2, \cdots, q_n \in Q$, and $f_1, f_2, \cdots, f_n \in T(E) = \{f \in \mathrm{Hom}_R(Q, I(E)) \mid f(1) \in E\}$ are such that $\sum q_i f_i = 0 \in \mathrm{Hom}_R(Q, I(E))$ then the large left ideal $I = \{r \in R \mid r\ q_i \in R$ for all $i\}$ annihilates $\sum q_i \otimes f_i$ in $Q \otimes_R T(E)$. Thus $\sum q_i \otimes f_i \in Z(Q \otimes_R T(E)) = (0)$. This shows that the R-homomorphism $Q \otimes_R T(E) \rightarrow \mathrm{Hom}_R(Q, I(E))$ given by the Q-module structure of $\mathrm{Hom}_R(Q, I(E))$ is one-to-one. Therefore by [1, Prop. 3, p. 64] the evaluation map from $T(E)$ onto E is a torsion free cover of E. This completes the proof.

REMARK. This theorem shows that "torsion free" covers do not exist for the ring mentioned by Banaschewski [1, p. 66], that is, the ring of all functions on a set X with values in a field which are constant except on some finite set. Since $Z(R) = (0)$ the "torsion free" he is using agrees with our torsion free and covers do not exist.

This paper constitutes a portion of the author's doctoral dissertation at the University of Kentucky. The author is deeply indebted to Professor Edgar Enochs, his advisor, for his generous advice and encouragement.

REFERENCES

1. B. Banaschewski, *On covering of modules*, Math. Nachr., **31** (1966), 57-71.
2. S. E. Dickson, *A torsion theory for abelian categories*, Trans. Amer. Math. Soc., **121** (1966), 223-225.
3. E. Enochs, *Torsion free covering modules*, Proc. Amer. Math. Soc., **14** (1963), 884-889.
4. J. Lambek, *Lectures on rings and modules*, Blaisdell Publishing Co., Waltham, Mass., 1966.
5. B. Maddox. *Absolutely pure modules*, Proc. Amer. Math. Soc., **18** (1967), 155-158.
6. C. Megibben, *Absolutely pure modules*, Proc. Amer. Math. Soc., **26** (1970), 561-566.
7. F. Sandomierski, *Semisimple maximal quotient rings*, Trans. Amer. Math. Soc., **128** (1967), 112-120.
8. F. Sandomierski, *Nonsingular rings*, Proc. Amer. Math. Soc., **19** (1968), 225-230.
9. M. Teply, *Torsion free injective modules*, Pacific J, Math., **28** (1969), 441-453.
10. ———, *Some aspects of Goldie's torsion theory*, Pacific J. Math., **29** (1969), 447-459.

Received April 2, 1971 and in revised form July 15, 1971.

UNIVERSITY OF KENTUCKY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

A DUALITY BETWEEN TRANSPOTENCE ELEMENTS AND MASSEY PRODUCTS

BYRON DRACHMAN AND DAVID KRAINES

The purpose of this note is to show that if v is an element whose suspension is nonzero, and if u is dual to v, then the transpotence $\varphi_k(v)$ is defined and nonzero if and only if the k-Massey product $\langle u\rangle^k$ is defined and nonzero.

We wish to thank Dr. Samuel Gitler for a helpful conversation on this material.

1. Preliminaries.

1.1. *The Cobar Construction*: (Adams [1]). Let C be a simply connected DGA coalgebra over K with co-associative diagonal map where K is a commutative ring with unit. The Cobar Construction $\bar{F}(C)$ is the direct sum of the n-fold tensor products of the desuspension of $\bar{C} = \operatorname{Ker}(\varepsilon)$ where $\varepsilon\colon C \to K$ is the augmentation. Suppose C has a differential $\{d_n\colon C_n \to C_{n-1}\}$. A typical element is a linear combination of elements of the form

$$x = s^{-1}(c_1) \otimes \cdots \otimes s^{-1}(c_n) = [c_1 | \cdots | c_n]$$

where x has bidegree $(-n, m)$ and $m = \sum_{i=1}^{n}$ degree (c_i). The differential in $\bar{F}(C)$ is defined on elements of bidegree $(-1, *)$ by

$$d[c] = [-dc] + \sum_i (-1)^{\deg c_i'}[c_i' | c_i'']$$

where

$$\varDelta(c) = c \otimes 1 + 1 \otimes c + \sum_i c_i' \otimes c_i''$$

$\varDelta\colon C \to C \otimes C$ being the diagonal mapping of C. The differential is extended to all of $\bar{F}(C)$ by the requirement that $\bar{F}(C)$ be a DGA-algebra.

If C has a differential of degree $+1$ instead of -1, we no longer ask that C be a simply connected but only connected, and the element $[c_1 | \cdots | c_n]$ is assigned bidegree (n, m).

1.2. *The Bar Construction*. Let A be a connected associative DGA algebra over K. Let $\varepsilon\colon A \to K$ be the augmentation. Let $\bar{A} = \ker \varepsilon$. Then the Bar Construction $\bar{B}(A)$ is the direct sum of the n-fold tensor products of the suspension of $\bar{A}$. Let

$$\{d_n\colon A_n \to A_{n-1}\}$$

be the differential in A. $\bar{B}(A)$ is bigraded by assigning the element $[a_1 | \cdots | a_n]$ degree (n, m) where $m = \sum_{i=1}^{n} \deg a_i$. $\bar{B}(A)$ has a differential $d = d_E + d_I$ where

$$d_E([a_1 | \cdots | a_n]) = \sum_{i=1}^{n-1} (-1)^{u(i)} [a_1 | \cdots | a_i a_{i+1} | \cdots | a_n]$$

$$d_I([a_1 | \cdots | a_n]) = \sum_{i=1}^{n} (-1)^{u(i-1)} [a_1 | \cdots | \partial a_i | \cdots | a_n]$$

where

$$u(i) = i + \sum_{k-1}^{i} \deg a_k .$$

We also mention that $[a | \cdots (k) \cdots | a]$ is $\gamma_k[a]$, the kth divided power of $[a]$.

If instead of the above the differential of A has degree $+1$, we put the bidegree of $[a_1 | \cdots | a_n]$ to be $(-n, m)$. In this case we will always assume A is simply connected.

1.3. *The Suspension Map.* In the case of the Bar Construction the suspension map σ: $H_*(A) \to H_*(\bar{B}(A))$ is represented by $a \to [a]$. In the case of the Cobar Construction, σ: $H_*(PA) \to H_*(\bar{F}(A))$ is represented by $a \to [a]$ where PA is the subcomplex of primitive chains.

DEFINITION 1. The Massey Product $\langle u \rangle^k$. (Kraines [6]).

Let A be a DGA algebra over K. Suppose $a_1, \cdots, a_{k-1}$ are given in A such that a_1 is a cycle (or cocycle) and that

$$\partial a_n = \sum_{r=1}^{n-1} (-1)^{\deg a_r} a_r a_{n-r} \text{ for } n = 2, \cdots, k-1 .$$

Suppose u is represented by a_1. Then the Massey Product $\langle u \rangle^k$ is represented by the cycle

$$\sum_{r=1}^{k-1} (-1)^{\deg a_r} a_r \cdot a_{k-r}$$

THEOREM 1. (Kraines, [6]). *The operation $\langle u \rangle^k$ depends only on the class $\{a_1\} \in H(A)$.*

DEFINITION 2. (Gitler, [5]). Suppose that A is an associative DGA algebra. Suppose $x \in H(A)$ is such that $v^k = 0$. The transpotence $\varphi_k(v) \in H(\bar{B}(A))_{/Im\sigma}$ is defined as follows: If $b \in A$ represents v then there exists $M \in A$ such that $\partial M = -b^k$. $\varphi_k(v)$ is represented by

$$(-1)^w [b^{k-1} | b] + [M] \text{ where } w = (1)^{\deg b^{k-1}} + 1 .$$

2. Main Theorem.

THEOREM 2. *Let C be a co-associative DGA coalgebra over K and let A be the dual associative DGA algebra over K. Suppose $H(A; K)$ and $H(\bar{B}(A); K)$ are free and of finite type over K. Let v in $H(A)$ and v in $H(\bar{F}(C); K)$ be such that the Kronecker index $\langle \sigma(v), u\rangle$ is 1. Then $\varphi_k(v)$ is defined and is not zero in $H(\bar{B}(A); K)$ if and only if $\langle u\rangle^k$ is defined and not zero in $H(\bar{F}(C); K)$. In this case*

$$\langle \varphi_k(v), \langle u\rangle^k\rangle = 1 .$$

In order to prove this theorem we shall consider the Eilenberg-Moore Spectral Sequences with
$E^2 = \mathrm{Cotor}^{H(\bar{B}(A);K)}\ (K, K)$
$E^r \Rightarrow E^\circ H(\bar{F}(\bar{B}(A)); K) \approx H(A; K)$ as algebras, and dually,
$(E')^2 = \mathrm{Tor}^{H(\bar{F}(C);K)}\ (K, K)$
$(E')^r \Rightarrow E^\circ H(B(\bar{F}(\bar{C}); K) \approx H(C; K)$ as coalgebras.
We also note that the Kronecker Index $\langle\,,\rangle$: $C \otimes A \to K$ induces a pairing

$$\langle\ ,\ \rangle\colon\ \bar{F}(C) \otimes \bar{B}(A) \to K$$

LEMMA 1. *Let $b \in A$ represent $v \in H(A)$. Suppose $v^k = 0$. Then*

$$d_k[\varphi_k(v)] = [\sigma b]^k \ \mathit{in}\ E^k .$$

Proof. Let

$$V = \sum_{i=1}^{k-1} P(i)\,[[b^i \mid b]]\ \ ([[b]])^{k-i-1} \text{ where } P(i) = (-1)^{\deg b^i+1}$$

and the outside bars refer to the Cobar Construction and the inside bars refer to the Bar Construction.

Taking ∂V gives a telescoping series and so

$$\partial V = [\sigma b]^k + (-1)^w[\sigma(b^k)].\ \ \text{Here } (-1)^w = P(k-1) .$$

In E^1, V represents the class $(-1)^w[[b^{k-1} \mid b]] + [[M]] = [\varphi_k(v)]$.

The Lemma follows from the definition of a spectral sequence of a bi complex.

LEMMA 2. *Let $a \in \bar{F}(C)$ represent u. Then, by definition,*

$$\gamma_k[a] = [a \mid \cdots (k) \cdots \mid a] \in \bar{B}(\bar{F}(C)) .$$

If $\gamma_k[a]$ lives to E^{k-1} then $\langle u\rangle^k$ is defined and

$$d_k(\gamma_k[a]) = \langle u\rangle^k \ \mathit{in}\ (E')^k .$$

Proof. We first make an observation: Suppose $\langle u\rangle^t$ is defined. Let (a_i) be a defining system for $\langle u\rangle^t$. Let

$$W = \sum_{r=2}^{t} \sum_{i_1+\cdots+i_r=t} [a_{i_1} \mid \cdots \mid a_{i_r}] \in \bar{B}(\bar{F}(C)) \ .$$

Then

$$\partial W = \sum_{i=1}^{t-1} (-1)^{\deg a_i+1} [a_i a_{t-i}] \ .$$

Now to prove Lemma 2, we use induction on k. Suppose the lemma is true for $k-1$. Suppose $\gamma_k[a]$ lives to E_{k-1}. Since E is a spectral sequence of DGA coalgebras, and $d_{k-1}(\gamma_k[a]) = 0$, we have

$$\Delta\, d_{k-1}\gamma_k[a] = d_{k-1}^{\otimes}\ \Delta\,\gamma_k[a] = d_{k-1}^{\otimes} \sum_{i=0}^{k} \gamma_i[a] \otimes \gamma_{k-i}[a] = 0$$

where $d^{\otimes}$ is the differential in $E' \otimes E'$. That is, in particular when $i = k-1$ in the above, we see

$$d_{k-1}\gamma_{k-1}[a] \otimes [a] = 0 \text{ so } d_{k-1}\ \gamma_{k-1}[a] = 0 \ .$$

Now by inductive hypothesis, $\langle u\rangle^{k-1}$ is defined so there is a defining system $(a_1, \cdots, a_{k-1})$ for $\langle u\rangle^{k-1}$ and a cochain a_k such that

$$\delta a_k = \sum_{i=2}^{k-2} (-1)^{\deg a_{i-1}} a_{i-1}a_{k-i}$$

since $\langle u\rangle^{k-1} = d_{k-1}\,\gamma_{k-1}[a] = 0$.

The observation at the beginning of this lemma shows that

$$d_k\gamma_k[a] = \langle u\rangle^k \ .$$

We now give the proof of Theorem 2:

Assume $\varphi_k(v)$ is defined and nonzero. We are assuming $= 1 = \langle \sigma v, u\rangle$. Hence

$$\begin{aligned} 1 = \langle \sigma v, u\rangle = \langle \sigma b, a\rangle &= \langle [\sigma b]^k, \gamma_k[a]\rangle = \langle d_k\varphi_k(v), \gamma_k[a]\rangle \\ &= \langle \varphi_k(v), d_k\gamma_k[a]\rangle = \langle \varphi_k(v), \langle u\rangle^k\rangle \end{aligned}$$

by the duality of the two spectral sequences and Lemma 2.

It remains to be shown that if $\langle u\rangle^k$ is defined and nonzero, then so is $\varphi_k(v)$. Consider the map

$A \to \bar{F}(\bar{B}(A))$ defined by

$b \to [[b]]$.

This map is homotopy multiplicative (in fact is a *SHM* map) and is an equivalence. Hence $[[b^k]]$ differs from $[\sigma b]^k$ by a boundary. But $[\sigma b]^k = [\sigma b \mid \cdots (k) \cdots \mid \sigma b]$ is dual to $\gamma_k[a] = [a \mid \cdots (k) \cdots \mid a]$ in $\bar{B}\bar{F}(C)$, and so $d_k\gamma_k[a] = \langle u\rangle^k$ is not zero in E^k (Lemma 2) and so does not survive to E^∞, i.e., represents 0 in E^∞. The dual element $[\sigma b]^k$ represents 0 in E^∞, i.e., $[[b^k]] \sim [\sigma b]^k \sim 0$. Therefore $b^k \sim 0$ and so $\varphi_k(v)$ is defined.

We wish to mention two applications:

Al: Let $K = Z_p$ and let X be a $K(\pi, n)$ space (p an odd prime). Let $C = C^*(X; Z_p)$ and $A = C_*(X; Z_p)$ be cochain and chain complexes for X of finite type. In the notation of Cartain, $A = A_*(\pi, n; Z_p)$ ([2]). Cartan proved that $\langle \varphi_p(v), \beta P^m(u) \rangle = \langle \sigma v, u \rangle$. Now by Theorem 2, if

$$\langle \sigma v, u \rangle = 1$$

then $\langle \varphi_p(v), \langle u \rangle^p \rangle = 1$. Hence $\langle \varphi_p(v), \beta P^m u + \langle u \rangle^p \rangle = 0$. By Lemma 18 ([5]), $\langle u \rangle^p = c\beta P^m u$. This gives an easy proof of the fact that $c = -1$. (Compare Theorem 19 [5]).

A2: Now let $x = CP^{k-1}$. Then in $H^*(CP^{k-1}; Z) = P(v)_{/(v^k)}$ we have $v^k = 0$. Then $\varphi_k(v)$ is defined in $H^*(\Omega\ CP^{k-1}; Z)$ and by the Theorem 2, so is $\langle u \rangle^k$ in $H_{3k-2}(\Omega\ CP^{k-1}; Z)$ where $u \in H_2(\Omega\, CP^{k-1}, Z)$ and $\langle \varphi(v), \langle u \rangle^k \rangle = 1$. This gives another proof of the results of Stasheff ([7]).

REFERENCES

1. J. F. Adams, *On the Cobar Construction,* Colloquie de Topologie Algébraique, Louvain 1955.
2. H. Cartan et al, *Alqébres d'Eilenberq-MacLane et homotopie,* Séminaire Henri Cartan, 1954-55. Secrétariat mathématique, Paris, 1956.
3. A. Clark, *Homotopy commutativity and the Moore spectral Sequence,* Pacific J. Math., **15** (1965), 65-74.
4. B. Drachman, *A Diagonal Map for the Cobar Construction* Boletín de la Sociedad Matemática Mexicana, 1967.
5. S. Gitler, *Espacios Fibrados por H-Espacios.* Boletin de la Sociedad Matemática Mexicana, 1963.
6. D. Kraines, *Massey higher products.* Trans. Amer. Math. Soc., **124** no. 3, (1966). 431-449.
7. J. Stasheff, *Homotopy associativity of H-spaces II.* Trans. Amer. Math. Soc., **108** no. 2. (1963). 293-312.

Received September 28, 1970.

MICHIGAN STATE UNIVERSITY
AND
AARHUS UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

INTEGRAL REPRESENTATION OF EXCESSIVE FUNCTIONS OF A MARKOV PROCESS

RICHARD DUNCAN

Let X_t be a standard Markov process on a locally compact separable metric space E having a Radon reference measure. Let $\mathscr{S}$ denote the set of locally integrable excessive functions of X_t and $ex\mathscr{S}$ the set of elements lying on the extremal rays of $\mathscr{S}$. Then if $u \in ex\mathscr{S}$ is not harmonic, it is shown that there is an $x \in E$ such that $P_V u = u$ for all neighborhoods V of x where P_V is the hitting operator of V. A regularity condition is introduced which guarantees that two functions in $\mathscr{S}$ having the above property at x are proportional. A subset $\hat{E} \subset E$ and a metric topology on $\hat{E}$ are defined which allows one to represent each potential $p \in \mathscr{S}$ in the form $p(x) = \int u(x, y)\nu(dy)$ for some finite Borel measure $\nu \geqq 0$ on $\hat{E}$. Here the function $u: E \times \hat{E} \rightarrow [0, \infty]$ is measurable with respect to the product Borel field and has the property that for each $y \in \hat{E}$ the function $x \rightarrow u(x, y)$ is an extremal excessive function. In the course of this study a dual potential operator is introduced and some of its properties are investigated.

In § 2 we introduce the notation and assumptions which will be assumed to hold throughout the paper. Section 3 begins our study of $ex\mathscr{S}$ and using a result of Meyer [7] we show that to each function $u \in ex\mathscr{S}$ which is not harmonic we can associate a point $x \in E$ such that $P_V u = u$ for all open neighborhoods V of x. Here P_V is the hitting operator associated with V. We then say that u has support at x in analogy to the property introduced in axiomatic potential theory by Hervé [4]. We then discuss the axiom of proportionality, i.e., when is it true that if $u_1, u_2 \in ex\mathscr{S}$ have support at x, it follows that $u_1 = \alpha u_2$ for some $\alpha \geqq 0$. Some conditions are given which guarantee this property.

In § 4 we begin the discussion of representation of elements of $\mathscr{S}$. A uniform integrability condition on $\mathscr{S}$ is imposed and we define a suitable compact, convex set $\mathscr{K}$ in $\mathscr{S}$. Using the Choquet theorem and the characterization of $ex\mathscr{S}$ established in § 3, we define a subset $\hat{E} \subset E$ and a metric topology on $\hat{E}$ which allows us to represent each potential $p \in \mathscr{K}$ in the form $p(x) = \int u(x, y)\nu(dy)$ for some Borel measure $\nu \geqq 0$ on $\hat{E}$. Here $u: E \times \hat{E} \rightarrow [0, \infty]$ is a function measurable with respect to the product Borel field on $E \times \hat{E}$ and having the property that the function $x \rightarrow u(x, y)$ is an extremal excessive function for each $y \in \hat{E}$.

In § 5 the dual operator $\hat{U}$ is introduced, defined for a continious function on E with compact support by $\hat{U}f(y) = \int f(x)u(x, y)dx$. Some properties of $\hat{U}$ are investigated, and the integral representation is then extended to all potentials $p \in \mathscr{S}$.

2. Preliminaries and notation. The primary reference for the material in this paper will be Blumenthal and Getoor [2], and most of the notation will be taken from that book. Let therefore E be a locally compact separable metric space, and write $E_\Delta = E \cup \{\Delta\}$ where Δ is the point at infinity if E is not compact and an isolated point otherwise. We denote by $\mathscr{B}(E)$ and $\mathscr{B}(E_\Delta)$ the Borel sets of E and E_Δ respectively. Let $X = (\Omega, \mathscr{F}, \mathscr{F}_t, X_t, \theta_t, P^x)$ be a standard process with state space $(E, \mathscr{B}(E))$. Thus $X_t: \Omega \to E_\Delta$ for each t, $0 \leqq t \leqq \infty$, such that $X_s(\omega) = \Delta$ for all s $\geqq t$ if $X_t(\omega) = \Delta$. The path functions $t \to X_t(\omega)$, $\omega \in \Omega$, are right continuous on $[0, \infty)$ and have left-hand limits on $[0, \zeta)$ almost surely. Here $\zeta = \inf\{t: X_t = \Delta\}$ is the lifetime of X. The shift operators $\theta_t: \Omega \to \Omega$ are defined by $X_t \circ \theta_h = X_{t+h}$. For each $x \in E_\Delta$, P^x is a probability measure on the σ-algebra $\mathscr{F}$ such that $x \to P^x(\Lambda)$ is $\mathscr{B}(E_\Delta)$ measurable for each $\Lambda \in \mathscr{F}$ and $P^x(X_0 = x) = 1$. The reader is referred to [2] for the definitions of $\{\mathscr{F}_t\}$ and $\mathscr{F}$. Finally, X is assumed to be strong Markov and quasi-left continuous on $[0, \zeta)$.

If F is any topological space, we write $B(F)$ for the real-valued Borel measurable functions on F, and $bB(F)$ for the bounded elements of $B(F)$. If F is locally compact Hausdorff, $C_K(F)$ will denote the real-valued continuous functions on F with compact support. If L is any space of functions, L^+ will denote the nonnegative elements of L. If $f \in B(E)$ we extend f to E_Δ by setting $f(\Delta) = 0$.

We denote by P_t^α, $\alpha \geqq 0$, the α-transition operator so that $P_t^\alpha f(x) = e^{-\alpha t}E^x[f(X_t)]$ for $f \in bB(E)$. Set $P_t = P_t^0$. Our notation for the resolvent of the semi-group is $U^\alpha f(x) = \int_0^\infty P_t^\alpha f(x)dt = E^x\int_0^\infty e^{-\alpha t}f(X_t)dt$, and we put $U = U^0$, the potential operator. Recall that for $\alpha > 0$, $U^\alpha: bB(E) \to bB(E)$ is a bounded linear operator on the Banach space $bB(E)$ with the supremum norm, and $\|U^\alpha\| \leqq \alpha^{-1}$. If B is Borel, then $P_B^\alpha f(x) = E^x[e^{-\alpha T_B}f(X_{T_B}); T_B < \zeta]$ defines the α-hitting operators. Here $T_B = \inf\{t > 0; X_t \in B\}$ is the hitting time of B. Recall that if $B \in \mathscr{B}(E)$, then $B^r = \{x: P^x[T_B = 0] = 1\}$ is the set of regular points for B, and $B \cup B^r$ is the closure of B in the fine topology. Also if $D \in \mathscr{B}(E)$ and $D = D^r$ then for each $x \in E$ there is a decreasing sequence $\{G_n\}$ of open sets containing D such that $T_{G_n} \uparrow T_D$ a.s., P^x on $\{T_D < \infty\}$. A Borel set D for which $D = D^r$ is called finely perfect.

We let $\mathscr{S}^\alpha$ denote the α-excessive functions of X and set $\mathscr{S} =$

$\mathscr{S}^0$. Thus a nonnegative Borel function f is in $\mathscr{S}^\alpha$ if $P_t^\alpha f \leqq f$ for all $t \geqq 0$ and $P_t^\alpha f(x) \uparrow f(x)$ as $t \downarrow 0$ for all $x \in E$. Recall that the fine topology is the coarsest topology on E relative to which each $f \in \mathscr{S}^\alpha$ is continuous, $\alpha > 0$. Let $u \in \mathscr{S}$. Unless otherwise qualified, the statement $u = 0$ will mean that u is the zero function. Similarly, $u \neq 0$ will mean that u is not identically zero.

One basic assumption which will be assumed to hold throughout is the existence of a (Radon) reference measure. This is a Radon measure dx having the property that a set $B \in \mathscr{B}(E)$ is of potential zero, i.e., $U(x, B) = 0$ for all $x \in E$, if and only if $\int_B dx = 0$. This condition is satisfied if the elements in $\mathscr{S}^\alpha$ are lower semi-continuous for some $\alpha > 0$. If $f, g \in \mathscr{S}^\alpha$ and $f = g$ a.e., dx, then f and g are identical. Also, under this assumption each $f \in \mathscr{S}^\alpha$ is Borel measurable. An important situation where a reference measure exists is when there is a dual Markov process $\hat{X}_t$ as in Chapter VI of [2]. Here the resolvent kernel is of the form $U^\alpha f(x) = \int u^\alpha(x, y) f(y) \xi(dy)$ where $u^\alpha : E \times E \to \bar{R}^+$ is $\mathscr{B}(E) \times \mathscr{B}(E)$ measurable, $\xi(dy)$ is a Radon measure on E, and the function $x \to u^\alpha(x, y)$ is α-excessive for each $y \in E$, $\alpha \geqq 0$. Moreover, the resolvent of the dual process $\hat{X}_t$ is given by $\hat{U}^\alpha f(y) = \int u^\alpha(x, y) f(x) \xi(dx)$, and for each $x \in E$, the function $y \to u^\alpha(x, y)$ is α-excessive for $\hat{X}_t$. One can then define, analogous to X_t, a cofine topology for $\hat{X}_t$, and it turns out that the notion of semi-polar is equivalent in these two topologies. If D is Borel, then ${}^rD \backslash D^r$ is semi-polar, where ${}^rD$ denotes the set of points cofinely regular for D.

We make finally the following assumption on U: If f is a bounded Borel measurable function on E with compact support, then the function $x \to Uf(x)$ is finite. This condition is always satisfied by the operator U^α for $\alpha > 0$ and in fact the assumption is mainly a convenience that simplifies the notation. The reader can easily convince himself that all of the following results are true when stated in terms of α-potentials for $\alpha > 0$. Under this assumption each excessive function is the limit of an increasing sequence $\{Uf_n\}$ of finite potentials where each $f_n \geqq 0$ is in $B(E)$.

We fix once and for all a reference measure dx and, changing our notation slightly, we agree to denote by $\mathscr{S}$ the set of all excessive functions of X which are locally integrable with respect to dx. Now $\mathscr{S}$ is a convex, proper, pointed cone of functions on E and we denote by $ex\mathscr{S}$ the set of extreme rays of $\mathscr{S}$: $u \in ex\mathscr{S}$ if and only if for any representation of u in the form $u = u_1 + u_2$ with $u_1, u_2 \in \mathscr{S}$ it follows that $u_1 = \alpha u_2$ for some $\alpha \geqq 0$. We will draw heavily upon the following result found in Meyer [7, p. 59]:

THEOREM 2.1. *Let* $\{u_n\}$ *be a sequence of excessive functions. Then* *there is a subsequence* $\{u_{n'}\}$ *and an excessive function* u *such that* $u_{n'} \to u$ a.e., dx.

From now on all "almost everywhere (a.e.)" statements will be in reference to the measure dx.

3. Characterization of $ex\mathscr{S}$**.** We now want to give a characterization of the extremal rays of $\mathscr{S}$. For this we make the

DEFINITION 3.1. An excessive function $u \in \mathscr{S}$ is said to have support at $x \in E$ if for any open neighborhood V of x, $P_V u = u$. Also, u is said to be harmonic if $P_{K^c} u = u$ for all compact subsets $K \subset E$.

REMARK 3.2. If $u \in \mathscr{S}$ has a support at x, then u is harmonic in $E \backslash \{x\}$. In this connection, see Bauer [1, Chap. V].

We now prove

THEOREM 3.3. *Let* $u \in ex\mathscr{S}$*. Then if* u *is not harmonic,* u *has support at some* $x \in E$.

For the proof, we will need a series of lemmas.

LEMMA 3.4. *Let* $\{u_1^n\}$ *and* $\{u_2^n\}$ *be sequences of excessive functions in* $\mathscr{S}$ *such that* $u_1^n + u_2^n \to u$ *for some* $u \in \mathscr{S}$*. Then if* $u_1^n \to u_1$ *a.e., and* $u_2^n \to u_2$ *a.e., for* $u_1, u_2 \in \mathscr{S}$*, we have* $u_1^n \to u_1$ *and* $u_2^n \to u_2$ *on* $\{u < \infty\}$.

Proof. Of course $u = u_1 + u_2$ since they agree almost everywhere, hence everywhere. The important fact here is that if $v_n, v \in \mathscr{S}$ and $v_n \to v$ a.e., then $v \leqq \liminf v_n$ [*Proof*: We have by Fatou's lemma $\alpha U^\alpha(x, \liminf v_n) \leqq \liminf \alpha U^\alpha(x, v_n) \leqq \liminf v_n(x)$ for any $\alpha > 0$, so $\liminf v_n$ is super-median. If $\bar{v}$ is the excessive regularization of $\liminf v_n$, then $\bar{v} \leqq \liminf v_n$. But $\bar{v} = \liminf v_n$ a.e., and therefore $\bar{v} = v$ a.e., hence $\bar{v} = v$ everywhere so that $v \leqq \liminf v_n$]. Now if $u_1^n + u_2^n \to u = u_1 + u_2$, then on $\{u < \infty\}$, $A \equiv \{\limsup u_1^n > u_1\} \subset \{\liminf u_2^n < u_2\}$ since $x \in A$ and $u(x) < \infty$ implies there is a subsequence $\{n'\}$ such that $u_1^{n'}(x) \to \alpha > u_1(x)$ and hence $u_2^{n'}(x) \to \beta < u_2(x)$. Therefore $\liminf u_2^n(x) \leqq \liminf u_2^{n'}(x) < u_2(x)$. But $\{\liminf u_2^n < u_2\} = \phi$ by the above remark. Thus $A = \phi$ and for any $x \in E$ with $u(x) < \infty$ we have $\limsup u_1^n(x) \leqq u_1(x) \leqq \liminf u_1^n(x)$; therefore $u_1^n \to u_1$ and hence $u_2^n \to u_2$ on $\{u < \infty\}$.

LEMMA 3.5. *Suppose* $\{u_n\} \subset \mathscr{S}$ *and* $u_n \to \beta u$ *on* $\{u < \infty\}$ *with* $\beta > 0$ *and* $u_n \leqq u \in \mathscr{S}$ *for all* n*. Let* B *be Borel. Then if* $P_B u_n = u_n$ *for all* n*, we have* $P_B u = u$.

Proof. Since $u \in \mathscr{S}$ we always have $P_B u \leqq u$. To show $P_B u \geqq u$, consider a point $x \in E$ where $u(x) < \infty$. Then the measure $P_B(x, .)$ puts no mass on $\{u = \infty\}$. Since $u_n \leqq u$ for all n, the dominated convergence theorem implies $P_B(x, u_n) \to P_B(x, \beta u) = \beta P_B u(x)$. But $P_B(x, u_n) = u_n(x) \to \beta u(x)$ and since $\beta > 0$, $P_B u(x) = u(x)$. Hence $P_B u = u$ on $\{u < \infty\}$ and since $\{u = \infty\}$ has dx-measure zero, $P_B u = u$ everywhere.

LEMMA 3.6. *Suppose $u \in ex\mathscr{S}$ is not harmonic. Then there is a compact $K \subset E$ and a sequence $\{f_n\} \subset B^+(E)$ of Borel functions vanishing outside of K such that $Uf_n \leqq u$ for all n and $Uf_n \to u$ as $n \to \infty$ on $\{u < \infty\}$.*

Proof. Since $u \in ex\mathscr{S}$, there is a sequence $\{g_n\}$ of nonnegative Borel functions with $Ug_n \uparrow u$. Assume the conclusion is not true, and let $K \subset E$ be an arbitrary compact. Then $1 = I_K + I_{K^c}$ and hence $Ug_n = UI_K g_n + UI_{K^c} g_n \uparrow u$. Here I_B denotes the indicator function of B, for any $B \in \mathscr{B}(E)$. By Theorem (2.1) and Lemma (3.4) and the fact that $Ug_n \leqq u$ for all n, we can find a subsequence $\{n'\}$ and excessive functions $u_1, u_2 \in \mathscr{S}$ such that $UI_K g_{n'} \to u_1$ and $UI_{K^c} g_{n'} \to u_2$ on $\{u < \infty\}$ with $u = u_1 + u_2$. Since $u \in ex\mathscr{S}$, $u_2 = \beta u$ for some $\beta \geqq 0$. Now $\beta \neq 0$ since otherwise $UI_K g_{n'} \to u$ and $I_k g_{n'} = 0$ on K^c for all n'. Thus $UI_{K^c} g_{n'} \to \beta u$ on $\{u < \infty\}$ and $\beta > 0$. But for any $x \in E$,

$$\begin{aligned} P_{K^c} UI_{K^c} g_{n'}(x) &= E^x \int_{T_{K^c}}^{\infty} I_{K^c}(X_t) g_{n'}(X_t) dt = E^x \int_0^{\infty} I_{K^c}(X_t) g_{n'}(X_t) dt \\ &= UI_{K^c} g_{n'}(x) \ . \end{aligned}$$

Hence Lemma (3.5) implies that $P_{K^c} u = u$. But K was an arbitrary compact and u is therefore harmonic, giving a contradiction.

Proof of Theorem (3.3). Suppose $u \in ex\mathscr{S}$ is not harmonic. Then by Lemma (3.6) we can find a compact $K \subset E$ and a sequence $\{f_n\} \subset B^+(E)$ with each f_n vanishing outside of K and $Uf_n \to u$ on $\{u < \infty\}$, $Uf_n \leqq u$ for all n. We define recursively a decreasing sequence $\{B_j\}$ of nonempty Borel sets such that diameter $(B_j) \downarrow 0$ and such that for each $j > 0$ there is an $\alpha_j > 0$ and subsequence $\{n'\} \subset \{n\}$ with $UI_{B_j} f_{n'} \to \alpha_j u$ on $\{u < \infty\}$. Set $B_1 = K$ and assume B_j has been defined with a corresponding $\alpha_j > 0$ and subsequence $\{n'\} \subset \{n\}$. Since $\bar{B}_j \subset K$ is compact, we can find a finite Borel partition $\{C_i\}$ of B_j such that diameter $(C_i) < 1/j$ diameter (B_j) for each i. Then $I_{B_j} = \sum_i I_{C_i}$ and hence $UI_{B_j} f_{n'} = \sum_i UI_{C_i} f_{n'} \to \alpha_j u$. By Theorem (2.1) and Lemma (3.4), there is an i_0, a subsequence $\{n''\} \subset \{n'\}$, and excessive functions $u_1, u_2 \in \mathscr{S}$ with $u_1 \neq 0$ such that $UI_{C_{i_0}} f_{n''} \to u_1$ and $\sum_{i \neq i_0} UI_{C_i} f_{n''} \to u_2$ on $\{u < \infty\}$. Since $\alpha u = u_1 + u_2 \in exS$, $u_1 = \beta \alpha_j u$ for some $\beta > 0$. Let $B_{j+1} = C_{i_0}$ and $\alpha_{j+1} = \beta \alpha_j > 0$. Then diameter $(B_{j+1}) \leqq 1/j$ diameter (B_j)

and $UI_{B_{j+1}}f_{n''} \to \alpha_{j+1}u$ on $\{u < \infty\}$, thus completing the definition of the sequence $\{B_j\}$.

Consider now the decreasing sequence $\{\bar{B}_j\}$ of nonempty compact subsets of E, and let $x \in \bigcap_j \bar{B}_j$. Let V be any neighborhood of x. Since diameter $(\bar{B}_j) \downarrow 0$, there is some j_0 with $V \supset \bar{B}_{j_0} \supset B_{j_0}$, and hence $T_V \leqq T_{B_{j_0}}$ a.s. Now there is a subsequence $\{n'\} \subset \{n\}$ and an $\alpha_{j_0} > 0$ such that $UI_{B_{j_0}} f_{n'} \to \alpha_{j_0}u$ on $\{u < \infty\}$, $UI_{B_{j_0}}f_{n'} \leqq u$ for all n'. But for each $x \in E$

$$P_V UI_{B_{j_0}}f_{n'}(x) = E^x\int_{T_V}^{\infty} I_{B_{j_0}}(X_t)f_{n'}(X_t)dt = UI_{B_{j_0}}f_{n'}(x)$$

since $T_V \leqq T_{B_{j_0}}$ a.s., Lemma (3.5) implies that $P_V u = u$ and the proof is complete.

We list here a property of $ex\mathscr{S}$.

PROPOSITION 3.7. (i) *If $u \in S$ has support at x, there is a sequence $\{x_n\}$ with $x_n \to x$ and $u\{x_n\} \uparrow \|u\| = \{\sup u(y) : y \in E\}$.*

(ii) *If u is harmonic and E is not compact, there is a sequence $\{x_n\}$ such that $x_n \to \Delta$ and $u(x_n) \uparrow \|u\|$.*

Proof. (i) Suppose not. Then there is a neighborhood V of x and a constant $M < \|u\|$ such that $u(x) \leqq M$ for all $x \in V$. Let G be a neighborhood of x with $\bar{G} \subset V$. Then $u(X_{T_G}) \leqq M$ a.s., on $\{T_G < \infty\}$ since $X_{T_G} \in G \subset G^r \subset \bar{G} \subset V$ a.s., on $\{T_G < \infty\}$. But $u(y) = P_G u(y) = E^y[u(X_{T_G}); T_G < \infty]$ and hence $u(y) \leqq M$ for all $y \in E$, a contradiction.

(ii) Same proof as in (i) using neighborhoods of infinity.

Recall that a point $x \in E$ is polar if $P^y[T_x < \infty] = 0$ for all $y \in E$ where T_x is the hitting time of $\{x\}$. It follows from (3.5) of [2, Chap. II] that if $u \in \mathscr{S}$, then $\{u = \infty\}$ is polar. As a converse to this result, we prove

PROPOSITION 3.8. *Assume $U^\alpha : C_K(E) \to C(E)$ for some $\alpha \geqq 0$. Then if x is polar and $0 \neq u \in ex\mathscr{S}$ has support at x, $\|u\| = \infty$.*

Proof. Suppose x is polar and let $0 \neq u \in ex\mathscr{S}$ have support at x with $\|u\| = M < \infty$. Let $\{G_n\}$ be a decreasing sequence of open sets containing x with $\bigcap_n \bar{G}_n = \{x\}$. Let $y \in E$ be distinct from x. Then $T_{G_n} \uparrow \infty$ a.s., P^y and $u(y) = P_{G_n}u(y) = E^y[u(X_{T_{G_n}})] \leqq MP^y[T_{G_n} < \infty]$. By (4.24) of [2, Chap. II], $X_{T_{G_n}} \to \Delta$ a.s., P^y as $n \to \infty$. Since $X_{T_{G_n}} \in \bar{G}_n$ on $\{T_{G_n} < \infty\}$ a.s., it follows that $T_{G_n} = \infty$ a.s. P^y for large n. Hence $P^y[T_{G_n} < \infty] \downarrow 0$ as $n \to \infty$ and therefore $u(y) = 0$. Since $y \neq x$ was arbitrary, $u(y) = 0$ for all $y \neq x$ and hence $u = 0$ as dx does not charge the polar set $\{x\}$. This contradicts the fact that $u \neq 0$, thus completing the proof.

We now investigate the following uniqueness problem: When is it true that if $u_1, u_2 \in ex\mathscr{S}$ have support at x, then $u_1 = \alpha u_2$ for some $\alpha \geqq 0$? For this we make the following

DEFINITION 3.9. (i) If u has support at $x \in E$, then u is said to be regular at x if $P_D u = u$ for all finely perfect sets $D = D^r$ containing x of the form $D = G^r$ where G is finely open.

(ii) A family $\mathscr{U} \subset ex\mathscr{S}$ of excessive functions is said to be regular if any $u \in \mathscr{U}$ which has support at x is regular at x.

PROPOSITION 3.10. *Suppose $u \in \mathscr{S}$ has support at $x \in E$ and has the following property: For every decreasing sequence $\{G_n\}$ of open sets containing x with $\lim_n T_{G_n} = T$ a.s., we have $P_{G_n} u \rightarrow P_T u$. Then u is regular at x.*

Proof. Let D be a finely perfect set containing x, and let $y \in E$ be arbitrary. Then there is a decreasing sequence $\{G_n\}$ of open sets containing D such that $T_{G_n} \uparrow T_D$ a.s. P^y on $\{T_D < \infty\}$; hence $P_{G_n} u(y) \rightarrow P_D u(y)$. But each G_n is a neighborhood of x, therefore $P_{G_n} u(y) = u(y)$ for all n, and it follows that $u(y) = P_D u(y)$. Since y was arbitrary, $P_D u = u$.

REMARK 3.11. If $u \in \mathscr{S}$ has support at x and is regular at x, then $P_V u = u$ for all finely open V containing x.

We now prove the main result concerning regularity.

THEOREM 3.12. *Suppose $\mathscr{U} \subset ex\mathscr{S}$ is regular, and let $x \in E$. Then up to a nonnegative multiplicative constant, there is at most one $u \in \mathscr{U}$ having support at x. Moreover, if $u \in \mathscr{S}$ has support at x and is regular at x, then $u \in ex\mathscr{S}$.*

Proof. We first show that if $u_1, u_2 \in \mathscr{S}$ have support at x and are regular at x, then $u_1 \leqq u_2$ or $u_2 \leqq u_1$. Indeed, set $D_1 = \{u_1 < u_2\}^r$ and $D_2 = D_1^{cr} \subset \{u_2 \leqq u_1\}$. Now D_1 and D_2 are finely perfect and since $E = D_1 \cup D_2$, x must be regular for one of these sets. Assume that $x \in \{u_1 < u_2\}^r = D_1$ (the other case is treated similarly). Since u_1 and u_2 are finely continuous, $u_1 = P_{D_1} u_1 \leqq P_{D_1} u_2 = u_2$, i.e., $u_1 \leqq u_2$. Let now $\beta = \sup\{\alpha \geqq 0 \colon \alpha u_1 \leqq u_2\} \geqq 1$. We claim that if $\beta = \infty$ then $u_1 = 0$. For in this case $u_2 = \infty$ on $\{u_1 > 0\}$. But $u_2 \in \mathscr{S}$ and hence $\int_{\{u_1 > 0\}} dx = 0$, for otherwise there would exist a compact $K \subset \{u_1 > 0\}$ such that $\int_K dx > 0$ which would imply that $\int_K u_2 dx = \infty$. Thus $u_1 = 0$ a.e., hence $u_1 = 0$ everywhere. Assume therefore that $\beta < \infty$. Then $\beta u_1 \leqq u_2$.

On the other hand, if $\varepsilon > 0$, there is an $x \in E$ such that $u_2(x) < (\beta + \varepsilon)u_1(x)$. But $(\beta + \varepsilon)u_1$ and u_2 also have support at x and are regular at x, implying that $u_2 \leqq (\beta + \varepsilon)u_1$. Since $\varepsilon > 0$ was arbitrary, $u_2 \leqq \beta u_1$ and therefore $\beta u_1 = u_2$, proving the first part of the theorem.

To prove the second part, assume that $u \in \mathscr{S}$ has support at x and is regular at x. Then if $u = u_1 + u_2$ with $u_1, u_2 \in \mathscr{S}$, we have $u = P_D u = P_D u_1 + P_D u_2 = u_1 + u_2$ for all finely perfect D containing x. But $P_D u_i \leqq u_i$ $(i = 1, 2)$ and hence $P_D u_i = u_i$. Thus u_1 and u_2 have support at x and are regular at x. The preceding proof implies that $u_1 = \alpha u_2$ for some $\alpha \geqq 0$ and therefore $u \in ex\mathscr{S}$.

Suppose $ex\mathscr{S}$ has the following property: If $u \in ex\mathscr{S}$ has support at x, then u is locally bounded and continuous on E. Using Proposition (3.10), it is easy to see that $ex\mathscr{S}$ is regular. We show that in certain cases a form of continuity is actually necessary for regularity to hold.

PROPOSITION 3.13. *Assume X is a Hunt process. Let x_0 be regular for $\{x_0\}$ and suppose $u \in ex\mathscr{S}$ has support at x_0 and $u(x_0) \neq 0$. Then u is the unique* (up to a nonnegative multiplicative constant) *element in $ex\mathscr{S}$ having support at x_0 if and only if $u(x) \leqq u(x_0) < \infty$ for all $x \in E$.*

Proof. Since x_0 is not polar and $u(x_0) \neq 0$, it follows that the excessive function $P_{x_0}u(x) = E^x[u(x_{T_{x_0}})] \leqq u(x)$ is not identically zero, has support at x_0 and is regular there, and is therefore in $ex\mathscr{S}$ from Theorem (3.12). If $u(x_0) = \infty$, then $E^x[u(x_{T_{x_0}})]$ could only take the values 0 and ∞ since $X_{T_{x_0}} = x_0$ a.s., on $\{T_{x_0} < \infty\}$. But then $P_{x_0}u = 0$ a.e. since $P_{x_0}u \in \mathscr{S}$, and hence $P_{x_0}u = 0$, a contradiction. Now the uniqueness assumption on u implies that $u = \alpha P_{x_0}u$ for some $\alpha \geqq 0$ and since $0 < P_{x_0}u(x_0) = u(x_0) < \infty$ it follows that $\alpha = 1$ and therefore $u(x) = P_{x_0}u(x) = E^x[u(X_{T_{x_0}})] \leqq u(x_0) < \infty$ for all $x \in E$.

Conversely, assume $u(x) \leqq u(x_0) < \infty$ for all $x \in E$. Let $\{G_n\}$ be a decreasing sequence of open sets containing x_0 such that $\bigcap_n \bar{G}_n = \{x_0\}$. Then $T_{G_n} \uparrow T_{x_0}$ a.s. Since X is a Hunt process, $X_{T_{G_n}} \to X_{T_{x_0}} = x_0$ and $\lim_n u(X_{T_{G_n}}) \geqq u(X_{T_{x_0}}) = u(x_0)$ on $\{T_{x_0} < \infty\}$. But $u(x) \leqq u(x_0)$ for all $x \in E$ and hence $\lim_n u(X_{T_{G_n}}) = u(x_0)$ on $\{T_{x_0} < \infty\}$. The bounded convergence theorem now implies that $u(x) = P_{G_n}u(x) = E^x[u(X_{T_{G_n}})] \to E^x[u(X_{T_{x_0}})] = P_{T_{x_0}}u(x)$ for each $x \in E$ and the proof is complete.

The property of regularity is not shared by all standard processes (consider translation to the right on the line), and we now seek other conditions which guarantee the uniqueness property announced in Theorem (3.12). First let us state this property explicitly.

(A) Let $x \in E$ be arbitrary. If $u_1, u_2 \in ex\mathscr{S}$ have support at x, then $u_1 = \alpha u_2$ for some $\alpha \geqq 0$.

This property was first studied by Hervé [4] in axiomatic potential and is known as the axiom of proportionality. We introduce now a property that will guarantee (A) in a large number of cases.

(B) Suppose $u \in ex\mathscr{S}$ has support at x, and let D be finely perfect set containing x. Then $P_D u$ has support at x.

Note that the property includes the case $P_D u = 0$. We will state explicitly when (B) is assumed to hold.

Let $T = \inf\{t: X_t \neq X_0\}$. A point $x \in E$ is called an instantaneous point if $P^x[T = 0] = 1$. It is easy to see that if dx does not charge singletons, then the points of E are instantaneous.

THEOREM 3.14. *Assume* (B) *and that* dx *does not charge singletons. Let* $u \in ex\mathscr{S}$ *have support at* x_0 *and suppose that either* x_0 *is polar or* $u(x_0) = 0$. *Then if* $D = D^r$ *contains* x_0, *we have* $P_D u = u$ *or* $P_D u = 0$.

Proof. Let $v = u - P_D u \geqq 0$. Then (B) implies $P_V v = v$ for all open neighborhoods V of x_0. It follows that if $B \subset E$ is any Borel set such that x_0 is in the interior of B^c, then $P_{B^c} v = v$. Let $E' = E \backslash \{x_0\}$ and consider the standard process $\tilde{X}_t$ defined by $\tilde{X}_t = X_t$ if $t < T_{x_0}$ and $\tilde{X}_t = \Delta$ if $t \geqq T_{x_0}$. Then $\tilde{X}_t$ has state space E' and transition function $\tilde{P}_t f(x) = E^x[f(X_t); t < T_{x_0}]$. Let d be a metric on E compatible with the topology and suppose $x \in E'$. Then there is a closed ball $B \subset E'$ with center x such that x_0 is in the interior of B^c. Thus if $y \in E'$, $\tilde{P}_{B^c} v(y) = E^y[v(X_{T_{B^c}}); T_{B^c} < T_{x_0}] \leqq E^y[v(X_{T_{B^c}})] = v(y)$. Since v is nonnegative and finely continuous, it follows from [2, Chap. II, (5.9)] that v is excessive for $\tilde{X}_t$. Therefore if we denote by $\{\tilde{U}^\alpha\}$ the resolvent operators for $\tilde{X}_t$, we have

$$\alpha \tilde{U}^\alpha v(x) = \alpha E^x \int_0^{T_{x_0}} e^{-\alpha t} v(X_t) dt \leqq v(x)$$

for all $x \in E'$. Now if $x \neq x_0$,

$$\begin{aligned} \alpha U^\alpha v(x) &= \alpha \tilde{U}^\alpha v(x) + \alpha E^x \int_{T_{x_0}}^\infty e^{-\alpha t} f(X_t) dt \\ &\leqq v(x) + E^x[e^{-\alpha T_{x_0}} \alpha U^\alpha v(X_{T_{x_0}})]. \end{aligned}$$

If x_0 is polar, the third term in the inequality is zero. If $u(x_0) = 0$, then $P_D u(x_0) = 0$ and $\alpha U^\alpha u(x_0) \leqq u(x_0) = 0$; similarly $\alpha U^\alpha P_D u(x_0) = 0$. It follows that $\alpha U^\alpha v(x_0) = 0$ and therefore $E^x[e^{-\alpha T_{x_0}} \alpha U^\alpha v(X_{T_{x_0}})] = 0$ since $X_{T_{x_0}} = x_0$ a.s., on $\{T_{x_0} < \infty\}$. Thus in both cases $\alpha U^\alpha v(x) \leqq v(x)$ for all $x \neq x_0$. We now define a function $\tilde{v}$ by $\tilde{v}(x) = v(x)$ if $x \neq x_0$, $\tilde{v}(x_0) = \infty$. Sin e x_0 has dx-measure zero, $\{x_0\}$ has zero measure with respect to the measures $\alpha U^\alpha(x, .)$, $x \in E$. It follows that $\alpha U^\alpha v(x) = \alpha U^\alpha \tilde{v}(x) \leqq \tilde{v}(x)$ for all $x \in E$ and therefore $\lim_{\alpha \to \infty} \alpha U^\alpha \tilde{v}(x) = v(x)$ is in $\mathscr{S}$. Thus we have a decomposition of u in the form $u = v + P_D u$

where v and $P_D u$ are in $\mathscr{S}$. Since $u \in ex\mathscr{S}$, $P_D u = u$ for some $\alpha \geqq 0$. If $\alpha = 0$ or $P_D u = 0$, then $u = 0$. We claim that if $P_D u \neq 0$, then $D \cap \{0 < u < \infty\} \neq \phi$. For if otherwise, $D = D \cap \{u = 0\} \cup D \cap \{u = \infty\}$, a disjoint union. But $\{u = \infty\}$ is polar, hence $T_D = T_{D \cap \{u=0\}}$ a.s., and therefore $X_{T_D} \in \{u = 0\}$ a.s., on $T_D < \infty$. Hence $P_D u(x) = E^x[u(X_{T_D})] = 0$ for all $x \in E$, a contradiction. Thus if $\alpha > 0$ and $P_D u \neq 0$, there is a point $x \in D$ with $0 < u(x) < \infty$ and hence $\alpha P_D u(x) = \alpha u(x) = u(x)$ implying that $\alpha = 1$, or $P_D u = u$.

COROLLARY 3.15. *Assume* (B) *and suppose points are polar and that* $\mathscr{S}$ *has the following property: if* $u \in ex\mathscr{S}$ *has support at* x *and* $u \neq 0$, *then* $u(x) \neq 0$. *Then* $ex\mathscr{S}$ *is regular.*

Proof. If points are polar, then dx certainly does not charge singletons. If $0 \neq u \in ex\mathscr{S}$ has support at x and $D = D^r$ contains x, then $P_D u = u$ or $P_D u = 0$ by Theorem (3.14). But $P_D u(x) = u(x) \neq 0$ and therefore $P_D u = u$, proving that $ex\mathscr{S}$ is a regular.

According to Theorem (3.3), to each $u \in ex\mathscr{S}$ which is not harmonic we can associate a point $x \in E$ such that u has support at x. We want to consider the case where to each $u \in ex\mathscr{S}$ which is not harmonic, there is a unique point x at which u has its support. In axiomatic potential theory this property holds by virtue of the sheaf properties of the harmonic functions in that theory. Here, however, we do not have the property that if G_1 and G_2 are open and u is harmonic in G_1 and G_2, then u is harmonic in $G_1 \cup G_2$. For a Hunt process this property holds if u is locally bounded (cf. Meyer [7]).

For the moment we content ourselves with the following results.

PROPOSITION 3.18. *Assume* $\mathscr{U} \subset ex\mathscr{S}$ *is regular. If* $u \in \mathscr{U}$ *has support at* x_1 *and* x_2, *then* $u(x_1) = u(x_2)$.

Proof. Suppose $u(x_1) < \delta < u(x_2)$. Then $V = \{u < \delta\}$ is finely open and contains x_1. Now $u(X_{T_V}) \leqq \delta$ a.s., on $\{T_V < \infty\}$ since u is finely continuous; hence $u(x_2) = P_V u(x_2) = E^{x_2}[u(X_{T_V}); T_V < \infty] \leqq \delta$, a contradiction.

DEFINITION 3.19. $\mathscr{U} \subseteqq ex\mathscr{S}$ is separating if to each $u \in \mathscr{U}$ there is a unique $x \in E$ such that u has support at x.

From Proposition (3.7), it follows that if $\mathscr{U} \subseteqq ex\mathscr{S}$ contains no harmonic functions and each $u \in \mathscr{U}$ has the property that its supremum is approached in any neighborhood of one and only one point in E, then $\mathscr{U}$ is separating. The following proposition justifies the terminology.

PROPOSITION 3.20. *Assume* $\mathscr{U} \subseteqq ex\mathscr{S}$ *is regular and contains no harmonic functions.*

(i) *Suppose* $\mathscr{U}$ *has the property that if* $u \in \mathscr{U}$ *has support at* x, *then* $0 < u(x) < \infty$. *Then* $\mathscr{U}$ *is separating if* $\mathscr{S}$ *separates points.*

(ii) *Suppose* $\mathscr{U}$ *has the following property: If* $u \in \mathscr{U}$ *has support at* x *and if* $y \neq x$, *there is a function* $v \in \mathscr{S}$ *and a Borel set* $D = D^r$ *containing* x *such that* $v \geqq u$ *on* D *and* $v(y) < v(x)$. *Then* $\mathscr{U}$ *is separating.*

Proof. (i) It suffices to consider the case where $u \in \mathscr{U}$ has support at two distinct points x and y. By Proposition (2.16), $u(x) = u(y) = \beta > 0$. Let $v \in \mathscr{S}$ satisfy $v(x) > v(y)$. Then there is an $\alpha > 0$ such that $\alpha v(x) > \beta > \alpha v(y)$. Now $V = \{\alpha v > u\}$ is finely open and contains x. Therefore $\alpha v > P_V \alpha v \geqq P_V u = u$, i.e., $u \leqq \alpha v$. But $\alpha v(y) < u(y)$, a contradiction.

(ii) Suppose u has support at x and y, $x \neq y$ and let v and D be as in the hypothesis. We have from Hunt's theorem [2, p. 141], $u = P_D u = \inf\{s \in \mathscr{S}: s \geqq u \text{ on } D\}$. Thus, from $v \geqq u$ on D it follows that $u \leqq v$ and hence $u(x) \leqq v(x) < v(y) = u(y)$. But $u(x) = u(y)$ by Proposition (2.16), a contradiction.

4. Representation of excessive functions. In this section we prove a representation theorem, in integral form, for a certain class of potentials of the standard process X. In the next section we extend this representation to all potentials in $\mathscr{S}$. Recall that $\mathscr{S}$ denotes the set of all excessive functions that are locally integrable with respect to the reference measure dx. We now topologize $\mathscr{S}$ as a subset of $M^+(E)$, the nonnegative Radon measures on E: to each $u \in \mathscr{S}$ we associate the measure $u(x)dx$. This topology on $\mathscr{S}$ is locally convex and it is given by the family of semi-norms $\{p_f: f \in C_K(E)\}$ defined by $p_f(u) = \int fu\,dx$. Thus a sequence $\{u_n\} \subset \mathscr{S}$ converges to $u \in \mathscr{S}$ if and only if $\int fu_n dx \to \int fu\,dx$ for all $f \in C_K(E)$. Moreover, because of the hypotheses on the state space E, $\mathscr{S}$ is metrizable (Cf. Choquet [3]).

A cap of $\mathscr{S}$ is a compact subset of $\mathscr{S}$ of the form $\{h \leqq 1\}$ where h is a map of $\mathscr{S}$ into $[0, \infty]$, linear in the sense that $h(0) = 0$, $h(u + v) = h(u) + h(v)$ for $u, v \in \mathscr{S}$, and $h(\alpha u) = \alpha h(u)$ for $u \in \mathscr{S}$, $\alpha \in R^+ = [0, \infty)$. In order to guarantee the existence of a sufficient number of caps of $\mathscr{S}$, we will make a special assumption. Recall that a sequence $\{\nu_n\}$ of nonnegative Radon measures on E is bounded if the sequence $\{\nu_n(f)\}$ is bounded for each $f \in C_K^+(E)$. Our special assumption, which holds in the situation discussed in [7, Chap. II], is as follows:

(4.1) Suppose $\{u_n\} \subset \mathscr{S}$ is a bounded sequence in $M^+(E)$ and $u_n \to u$ a.e., for some $u \in \mathscr{S}$. Then there is a subsequence $\{u_{n'}\} \subset \{u_n\}$ such that $u_{n'} \to u$ in $\mathscr{S}$.

It follows that $\mathscr{S}$ is a closed subset of $M^+(E)$, for if $\{u_n\}$ is a sequence of excessive functions in $\mathscr{S}$ and $u_n \to \nu$ in $M^+(E)$ for some $\nu \in M^+(E)$, then by Theorem (2.1) we can find a subsequence $\{u_{n'}\} \subset \{u_n\}$ and an excessive function u such that $u_{n'} \to u$ a.e. But for each $f \in C_K(E)$ we have by Fatou's lemma $\int fudx = \int f \liminf u_{n'}dx \leqq \liminf \int fu_{n'}dx = \int fd\nu(x)$ so that $u \in \mathscr{S}$. By (4.1) there is a subsequence $\{u_{n''}\} \subset \{u_{n'}\}$ such that $u_{n''} \to u$ in $\mathscr{S}$ and therefore $\int fudx = \int fd\nu(x)$ for all $f \in C_K(E)$, implying that $d\nu(x) = u(x)dx$. Note that (4.1) is satisfied if $\mathscr{S}$ has the following property: If $\{u_n\} \subset \mathscr{S}$ and $u_n \to u$ a.e. for some $u \in \mathscr{S}$, then for each compact $K \subset E$, there is a subsequence $\{u_{n'}\} \subset \{u_n\}$ which is uniformly integrable over K.

Now (4.1) implies that $\mathscr{S}$ is well-capped, i.e., $\mathscr{S}$ is the union of its caps (Meyer [6, Chap. XI]). Thus Choquet's representation theorem applies (cf. [3]). Let $\mathscr{S}'$ denote the continuous linear forms on $\mathscr{S}$. Then if $v \in \mathscr{S}$, there is a nonnegative Radon measure ν carried by $ex\mathscr{S}$ such that for $l \in \mathscr{S}'$, $l(v) = \int_{ex\mathscr{S}} l(u)\nu(du)$.

Let now $\{K_n\}$ be an increasing sequence of compact subsets of E with $K_n \subseteq K_{n+1}$ and $E = \bigcup_n K_n$. Let $\{f_n\}$ be a sequence of nonnegative continuous functions with compact support such that for each $n, f_n(x) = 1$ for all $x \in K_n$. Choose numbers $\alpha_n > 0$ such that $\sum_n \alpha_n \int f_n dx = 1$, and denote by $h\colon \mathscr{S} \to [0, \infty]$ the functional defined by $h(u) = \sum_n \alpha_n \int f_n u dx$. It is clear that $h(0) = 0$, $h(u + v) = h(u) + h(v)$ for $u, v \in \mathscr{S}$, and $h(\beta u) = \beta h(u)$ for $\beta \geqq 0$. If we let $\mathscr{K} = \{u\colon h(u) \leqq 1\} = \{u\colon \sum_n \alpha_n \int f_n u dx \leqq 1\}$, then (4.1) implies that $\mathscr{K}$ is a compact, convex set in $\mathscr{S}$. Therefore, if $\hat{\mathscr{S}}$ is the convex, proper cone generated by $\mathscr{K}$, $\hat{\mathscr{S}}$ will have compact base $\mathscr{K}$ and will be σ-compact. Note that $\hat{\mathscr{S}} = \{u \in \mathscr{S}\colon h(u) < \infty\}$ and that if $v \in \mathscr{S}$ is bounded, then $v \in \hat{\mathscr{S}}$. final-ly, we denote by $\mathscr{B}(\mathscr{K})$ the Borel sets of $\mathscr{K}$.

LEMMA 4.2. *Suppose $\{u_j\}$ is a sequence of excessive functions in $\mathscr{K}$ such that $u_j \to u$ in $\mathscr{K}$ for some $u \in \mathscr{K}$. Then for each integer $n > 0$ and $\alpha > 0$ we have $U^\alpha(x, u_j \wedge n) \xrightarrow{j} U^\alpha(x, u \wedge n)$ for all $x \in E$.*

Proof. Consider an integer $n > 0$ and $\alpha > 0$. We show first that $\int_B u_j \wedge ndx \to \int_B u \wedge ndx$ for all Borel sets $B \subset E$ having compact closure. Assume this is not the case so that there is an $\varepsilon > 0$ and a subsequence $\{j'\} \subset \{j\}$ with $\left|\int_B u_{j'} \wedge n dx - \int_B u \wedge n dx\right| \geqq \varepsilon$ for some Borel set B with compact closure and for all j'. By Theorem (2.1) and (4.1) we can find a subsequence $\{j''\} \subset \{j'\}$ and an excessive function u such that $u_{j''} \to u$ a.e. as $j'' \to \infty$ and that $\int f u_{j''} dx \to \int f \tilde{u} dx$ for all $f \in C_K(E)$. It follows that $\int f \tilde{u} dx = \int f u dx$ for all $f \in C_K(E)$ and therefore $u = \tilde{u}$ a.e., hence everywhere. Thus $u_{j''} \to u$ a.e., and $\int_B u_{j''} \wedge n dx \to \int_B u \wedge n dx$, giving the desired contradiction.

Fix $x \in E$. Then the Borel measure $B \to U^\alpha(x, B)$ is absolutely continuous with respect to dx, and $U^\alpha(x, E) = U^\alpha 1(x) \leqq 1/\alpha < \infty$. Since $\int_B u_j \wedge n dx \xrightarrow{j} \int_B u \wedge n dx$ for all $B \in \mathscr{B}(E)$ with compact closure, it follows that $U^\alpha(x, u_j \wedge n) \to U^\alpha(x, u \wedge n)$ as $j \to \infty$ and the proof is complete.

THEOREM 4.3. *The map $\Phi: E \times \mathscr{K} \to \bar{R}^+ = [0, \infty]$ defined by $\Phi(x, u) = u(x)$ is $\mathscr{B}(E) \times \mathscr{B}(\mathscr{K})$ measurable.*

Proof. It is sufficient to show that for each $\alpha > 0$, the map Φ^α: $E \times \mathscr{K} \to \bar{R}^+$ defined by $\Phi^\alpha(x, u) = U^\alpha(x, u) = U^\alpha u(x)$ is $\mathscr{B}(E) \times \mathscr{B}(\mathscr{K})$ measurable since for each $x \in E$ and $u \in \mathscr{K}$, $\alpha\Phi^\alpha(x, u) = \alpha U^\alpha u(x) \uparrow u(x) = \Phi(x, u)$ as $\alpha \to \infty$. Let $\alpha > 0$, and for each integer $n > 0$ define the map $\Phi_n^\alpha: E \times \mathscr{K} \to \bar{R}^+$ by $\Phi_n^\alpha(x, u) = U^\alpha(x, u \wedge n)$. For fixed $u \in \hat{\mathscr{S}}$ the map $x \to \Phi_n^\alpha(x, u)$ is $\mathscr{B}(E)$ measurable, and Lemma (4.2) implies that for fixed $x \in E$ the map $u \to \Phi_n^\alpha(x, u)$ is continuous on $\mathscr{K}$. Since $\mathscr{K}$ is a compact metric space, if follows that Φ_n^α is $\mathscr{B}(E) \times \mathscr{B}(\mathscr{K})$ measurable. But $\Phi_n^\alpha(x, u) = U^\alpha(x, u \wedge n) \uparrow U^\alpha(x, u)$ as $n \to \infty$ and therefore Φ^α is $\mathscr{B}(E) \times \mathscr{B}(\mathscr{K})$ measurable, completing the proof of Theorem (4.3).

COROLLARY 4.4. *Let $B \in \mathscr{B}(E)$. Then the map $P_B: E \times \mathscr{K} \to \bar{R}^+$ defined by $P_B(x, u) = \int P_B(x, dy)u(y) = P_B u(x)$ is $\mathscr{B}(E) \times \mathscr{B}(\mathscr{K})$ measurable.*

Proof. Let $H = \{\varphi \in B(E \times \mathscr{K})$: $(x, u) \to \int P_B(x, dy)\varphi(y, u)$ is $\mathscr{B}(E) \times \mathscr{B}(\mathscr{K})$ measurable$\}$. Then H contains all functions of the form $\varphi_1(x)\varphi_2(u)$ where $\varphi_1 \in B(E)$ and $\varphi_2 \in B(\mathscr{K})$. Moreover, if $\{\varphi_n\}$ is an increasing sequence of functions in H with $\varphi = \lim \varphi_n$, then the monotone convergence theorem implies that φ is in H. Hence, by the monotone class theorem, $B(E \times \mathscr{K}) \subset$ H. Since the function $(x, u) \to u(x)$ is in $B(E \times \mathscr{K})$, the result follows.

COROLLARY 4.5. (i) *Suppose $\nu \geqq 0$ is a finite Borel measure* *on $\hat{\mathscr{S}}$ carried by $\mathscr{K}$, and $v(x) = \int u(x)\nu(du)$. Then $v \in \hat{\mathscr{S}}$.*

(ii) *Suppose $\nu \geqq 0$ is a finite Borel measure on $\hat{\mathscr{S}}$ carried by $\mathscr{K}$ and v is an excessive function such that $l(v) = \int l(u)\nu(du)$ for all $l \in \mathscr{S}'$. Then $v \in \hat{\mathscr{S}}$ and $v(x) = \int u(x)\nu(du)$ for all $x \in E$.*

Proof. (i) Note first that the integral makes sense by the joint measurability of the map $(x, u) \to u(x)$. We have by Fubini's theorem $\alpha U^\alpha(x, v) = \int \alpha U^\alpha(x, u)v(du) \leqq \int u(x)\nu(du) = v(x)$ since $\alpha U^\alpha(x, u) \leqq u(x)$ for all $u \in \mathscr{K}$, $\alpha \geqq 0$. Also, since $\alpha U^\alpha(x, u) \uparrow u(x)$ as $\alpha \to \infty$ for all $u \in \mathscr{K}$, the monotone convergence theorem implies that $\alpha U^\alpha(x, v) \uparrow v(x)$ as $\alpha \to \infty$ so that v is excessive. To see that $v \in \hat{\mathscr{S}}$, use Fubini's theorem to write

$$\begin{aligned} h(v) &= \sum_n \alpha_n \int f_n(x) \int u(x)\nu(du)dx = \sum_n \alpha_n \int \left(\int f_n(x)u(x)dx \right) \nu(du) \\ &= \int \left(\sum_n \alpha_n \int f_n(x)u(x)dx \right) \nu(du) = \int h(u)\nu(du) < \infty \end{aligned}$$

since $h(u) \leqq 1$ for all $u \in \mathscr{K}$.

(ii) Since $p_f \in \mathscr{S}'$ for each $f \in C_K(E)$, we have

$$\int f(x)v(x)dx = \int \left(\int f(x)u(x)dx \right) \nu(du)$$

for all $f \in C_K(E)$. On the other hand, the function $\tilde{v}(x) = \int u(x)\nu(du)$ is in $\hat{\mathscr{S}}$ by (i), and for each $f \in C_K(E)$,

$$\begin{aligned} \int f(x)\tilde{v}(x)dx &= \int f(x) \int u(x)\nu(du)dx = \int \left(\int f(x)u(x)dx \right) \nu(du) \\ &= \int f(x)v(x)dx \end{aligned}$$

and therefore $\tilde{v} = v$ a.e., and hence everywhere since $\tilde{v}$ and v are excessive.

Consider again our increasing sequence $\{K_j\}$ of compact subsets of E with $\mathring{K}_j \subset K_{j1}$ and $E = \bigcup_j K_j$. For each j, define $\Psi_j : E \times \mathscr{K} \to \bar{R}^+$ by $\Psi_j(x, u) = P_{K_j^c}(x, u)$, a $\mathscr{B}(E) \times \mathscr{B}(\mathscr{K})$ *measurable function, and set* $\Psi(x, u) = \lim_j \downarrow P_{K_j^c}(x, u)$. From Fubini's theorem, the map $u \to h(\Psi_j(., u)) = \sum_n \alpha_n \int f_n \Psi_j(x, u)dx$ is $\mathscr{B}(\mathscr{K})$ measurable, and therefore $h(\Psi(., u)) = \lim_j \downarrow h(\Psi_j(., u)$ is $\mathscr{B}(\mathscr{K})$ measurable. Therefore the set $\mathscr{P} = \{u \in \mathscr{K} : h(\Psi(., u)) = 0\}$ is a Borel subset of $\mathscr{K}$. It is clear that $u \in \mathscr{P}$ if and only if $u \in \mathscr{K}$ and $P_{K_j^c}u \downarrow 0$ a.e., as $j \to \infty$ for all increasing sequences $\{K_j\}$ of compacts such that $\mathring{K}_j \subset K_{j+1}$ and $E = \bigcup_j K_j$. Finally, we put $\hat{\mathscr{P}} = ex\mathscr{K} \cap \mathscr{P} \backslash \{0\}$ where $ex\mathscr{K}$ is the set of extreme points of the compact, convex set $\mathscr{K}$. Then $\hat{\mathscr{P}} \subset \{u \in$

$\mathscr{P}: h(u) = 1\}$. See Meyer [6, Chap. XI]. We make the following assumption on $\hat{\mathscr{P}}$, which is valid if $\hat{\mathscr{P}}$ is regular and separating:

(4.6) $\hat{\mathscr{P}}$ is separating and the proportionality axiom holds.

Note that $\hat{\mathscr{P}}$ contains no harmonic elements for if $u \in \hat{\mathscr{P}}$ is harmonic, then $u = P_{K_j^c} u \downarrow 0$ a.e., for a sequence $\{K_j\}$ of compacts with $\mathring{K}_j \subset K_{j+1}$ and $E = \bigcup_j K_j$. Thus $u = 0$ a.e., hence everywhere and $0 \notin \hat{\mathscr{P}}$. Therefore, according to Theorem (3.3) and the assumption (4.6), to each $u \in \hat{\mathscr{P}}$ we can associate a unique $y \in E$, the point at which u has its support. We indicate this relation by setting $u = u_y$. Consider now the map $\Gamma: \hat{\mathscr{P}} \to E$ defined by $\Gamma(u_y) = y$. Define $\hat{E} = \Gamma(\hat{\mathscr{P}}) \subset E$. Then Γ is one-one onto $\hat{E}$. Moreover, we can give $\hat{E}$ the topology which makes Γ a homeomorphism between $\hat{\mathscr{P}}$ and $\hat{E}$. It is easily seen that this topology is given by the metric $d: \hat{E} \times \hat{E} \to R^+$ defined by $d(x, y) = \rho(u_x, u_y)$ where ρ is the metric on $\mathscr{K}$. In other words, the topology on $\hat{E}$ is defined by the family of semi-norms $\{p_f: f \in C_K(E)\}$ given for $y \in \hat{E}$ by $p_f(y) = \int f u_y dx$.

Consider now the function $u: E \times \hat{E} \to \bar{R}^+$ defined by $u(x, y) = u_y(x)$. This function is $\mathscr{B}(E) \times \mathscr{B}(\hat{E})$ measurable since it is the restriction of the $\mathscr{B}(E) \times \mathscr{B}(\mathscr{K})$-measurable map $(x, u) \to u(x)$ to the set $E \times \hat{\mathscr{P}}$ and $\hat{\mathscr{P}}$ is Borel in $\mathscr{K}$. We come now to the main result of this development. Recall that an excessive function $p \in \mathscr{S}$ is called a potential if $P_{K_n^c} p \downarrow 0$ a.e., for all increasing sequences $\{K_n\}$ of compacts such that $\mathring{K}_n \subseteqq K_{n+1}$ and $E = \bigcup_n K_n$.

THEOREM 4.7. *There is a subset $\hat{E} \subseteqq E$ with a metric topology and a function $u: E \times \hat{E} \to \bar{R}^+$ which is $\mathscr{B}(E) \times \mathscr{B}(\hat{E})$ measurable and having the property that the function $x \to u(x, y)$ is an extremal excessive function for each $y \in \hat{E}$. Each potential $p \in \hat{\mathscr{S}}$ has a representation of the form*

$$p(x) = \int u(x, y) \nu(dy)$$

for some finite Borel measure $\nu \geqq 0$ on $\hat{E}$.

Proof. The only statement to prove is the last sentence of the theorem. If $p \in \hat{\mathscr{S}}$, then by Choquet's theorem there is a nonnegative Radon measure μ carried by $ex\mathscr{K}$ such that $l(p) = \int_{ex\mathscr{K}} l(u)\mu(du)$ for $l \in \mathscr{S}'$; therefore $p(x) = \int_{ex\mathscr{K}} u(x)\mu(du)$ by Corollary (3.4). Since $\hat{\mathscr{P}} \subset \mathscr{K}$ is Borel, $p(x) = \int_{\hat{\mathscr{P}}} u(x)\mu(du) + \int_{\mathscr{F}} u(x)\mu(du)$ where $\mathscr{F} = ex\mathscr{K} \backslash \hat{\mathscr{P}}$. Now $\int_{\mathscr{P}} u(x)\mu(du) = \int_{\Gamma(\hat{\mathscr{P}})} \Gamma^{-1}(y)(x)\mu \circ \Gamma^{-1}(dy) = \int_{\hat{E}} u(x, y)\nu(dy)$ where $\nu = \mu \circ \Gamma^{-1}$ is a Borel measure on $\hat{E}$.

It remains to show that $\int_{\mathscr{F}} u(x)\mu(du) = 0$. Let $\{K_n\}$ be an increasing sequence of compacts such that $\mathring{K}_n \subset K_{n+1}$ and $E = \bigcup_n K_n$. Then Fubini's theorem yields

$$P_{K_n^c} p(x) = \int_{\hat{\mathscr{P}}} P_{K_n^c} u(x)\mu(du) + \int_{\mathscr{F}} P_{K_n^c} u(x)\mu(du) .$$

Now $P_{K_n^c} p \downarrow 0$ a.e., and hence $\int_{\mathscr{F}} \lim \downarrow P_{K_n^c} u(x)\mu(du) = 0$ a.e., or $\int_{\mathscr{F}} \Psi(x, u)\mu(du) = 0$ a.e.

Using Fubini's theorem again, we can write

$$0 = h\Big(\int_{\mathscr{F}} \Psi(., u)\mu(du)\Big) = \int_{\mathscr{F}} h(\Psi(., u))\mu(du).$$

Thus $\mu\{u \in \mathscr{F} \colon h(\Psi(., u)) > 0;\} = \mu\{ex\mathscr{K} \backslash \hat{\mathscr{P}}\} = 0$ and therefore μ is carried by $\hat{\mathscr{P}}$, completing the proof of Theorem (4.7).

We are going to improve Theorem (4.7), but before this we consider a related notion which is of independent interest.

5. Dual operator and the representation theorem.

We introduce now a dual operator associated with the potential operator U.

DEFINITION 5.1. The linear operator $\hat{U}\colon C_K(E) \to C(\hat{E})$ is defined for $f \in C_K(E)$ by $\hat{U}f(y) = \int f(x)u(x, y)dx$ and is called the dual operator of U.

The fact that $\hat{U}f(y)$ is a continuous function on $\hat{E}$ follows from the observation that $\hat{U}f(y) = \int f(x)u(x, y)dx = \int f(x)u_y(x)\,dx = p_f(y)$ where p_f is a semi-norm defining the topology on $\hat{E}$.

We want to investigate some of the properties of $\hat{U}$. The results obtained here are analogus to the case where a dual process exists as in [2, Chap. VI] or [7, Chap. II]. Now Meyer [5] has shown that $\mathscr{S}$, and therefore $\hat{\mathscr{S}}$, is a lattice in its own order, i.e., the order defined for $u, v \in \mathscr{S}$ by $u < v$ if and only if there is an $s \in \mathscr{S}$ such that $v = u + s$. The Choquet-Meyer Uniqueness Theorem [3] then implies that each $u \in \hat{\mathscr{S}}$ is represented by a unique nonnegative Radon measure carried by $ex\mathscr{K}$.

If ν is a signed Borel measure on $\hat{E}$ having finite total variation, we denote by $U\nu(x)$ the function $x \to \int u(x, y)\nu(dy)$. If $\nu \geqq 0$ is finite, then $U\nu \in \hat{\mathscr{S}}$ from Corollary (4.5).

PROPOSITION 5.2. (i) *If ν is a signed Borel measure on $\hat{E}$ of* *finite total variation, and if* $U\nu = 0$ *a.e., then* $\nu = 0$.

(ii) *If* $K \subset \hat{E}$ *is compact, then the restrictions of the functions in image* (U) *to* K *is dense in* $C(K)$.

Proof. (i) If ν is a such a measure, write $\nu = \nu_1 - \nu_2$ where ν_1 and ν_2 are finite and nonnegative. Then $\int u(x, y)\nu_1(dy) = \int u(x, y)\nu_2(dy)$ a.e., or $U\nu_1 = U\nu_2$ a. e.. But each of these functions is in $\hat{\mathscr{S}}$, hence $U\nu_1 = U\nu_2$. The Choquet-Meyer uniqueness theorem then implies $\nu_1 = \nu_2$ and therefore $\nu = \nu_1 - \nu_2 = 0$.

(ii) Let $K \subset \hat{E}$ be compact. Let ν be a Radon measure on K and suppose that $\int_K \hat{U}f(y)\nu(dy) = 0$ for all continuous functions f with compact support. Then $0 = \int_K \hat{U}f(y)\nu(dy) = \int\left(\int u(x, y)f(x)dx\right)\nu(dy) = \int f(x)dx\int u(x, y)\nu(dy) = \int f(x)\,U\nu(x)dx$ for all $f \in C_K(E)$. But then $U\nu = 0$ a.e., and hence by (ii), $\nu = 0$. The result now follows from the Hahn-Banach Theorem.

We now make the following observations: The set $\hat{E} = \hat{\mathscr{P}} \subset \mathscr{K}$ is a subset of the compact set $\mathscr{K}$, and therefore $F = \hat{E}^a$, the closure of $\hat{E}$ in $\mathscr{K}$, is a compact subset of $\mathscr{K}$. Note that $0 \notin F$. We claim that if $f \in C_K(E)$, then the function $\hat{U}f$ extends uniquely to a continuous function on F which we continue to denote by $\hat{U}f$. This follows from the previously mentioned fact that $\hat{U}f(y) = p_f(u_y)$ and p_f is one of the semi-norms defining the topology on F. Note that if $u \in F\backslash\hat{E}$, then $\hat{U}f(u) = \int f(x)u(x)dx$. In the terminology of [7], F is a "Martin Compactification" of the space $\hat{E}$. Finally, recall that $M^+(F)$ denotes the nonnegative Radon measures on F, and that any finite nonnegative Borel measure ν on $\hat{E}$ can be regarded as an element $\tilde{\nu} \in M^+(F)$ by the formula $\tilde{\nu}(B) = \tilde{\nu}(B \cap \hat{E})$ for $B \in \mathscr{B}(F)$. We now generalize Theorem (4.7).

THEOREM 5.3. *There is a subset* $\hat{E} \subset E$, *a metric topology on* $\hat{E}$ *making* $\hat{E}$ *a dense subset of a compact metric space* F, *and a function* $u: E \times \hat{E} \to [0, \infty]$ *having the following properties: The function* u *is* $\mathscr{B}(E) \times \mathscr{B}(\hat{E})$ *measurable and for each* $y \in \hat{E}$, *the function* $x \to u(x, y)$ *is an extremal excessive function. Each potential* $p \in \mathscr{S}$ *has a representation of the form*

$$p(x) = \int_{\hat{E}} u(x, y)\nu(dy)$$

for some uniquely determined finite Borel measure $\nu \geqq 0$ *on* $\hat{E}$. *For any* $f \in C_K(E)$, $\hat{U}f$ *has a unique continuous extension to* F.

Proof. According to Theorem (4.7) and the preceding remarks, the only part of the theorem to prove is the representation for potentials $p \in \mathscr{S}$. We show that if $p \in \mathscr{S}$ is potential, then $p \in \hat{\mathscr{S}}$ and hence the representation holds from Theorem (4.7). But if $p \in \mathscr{S}$, then $p_n(x) = (p \wedge n)(x)$ is an element of $\hat{\mathscr{S}}$ and therefore $p_n(x) = \int u(x, y)\nu_n(dy)$ for some finite Borel measure $\nu_n \geqq 0$ on $\hat{E}$. Let $f \in C_K^+(E)$. Then

$$\begin{aligned}\int f(x)p_n(x)dx &= \int f(x)\Big(\int u(x, y)\nu_n(dy)\Big)dx \\ &= \int\Big(\int f(x)u(x, y)dx\Big)\nu_n(dy) = \int \hat{U}f(y)\nu_n(dy) \\ &= \int_F \hat{U}f(y)\tilde{\nu}_n(dy) \uparrow \int f(x)p(x)dx < \infty .\end{aligned}$$

Since F is compact and $0 \notin F$, we can find a finite number $\{f_i\}$ of function in $C_K^+(E)$ such that $\sum_i p_{f_i}(u) > 0$ for all $u \in F$. But $p_{f_i}(u) = \int f_i u dx = \hat{U}f_i(u)$ on F and therefore $\sum_i \hat{U}f_i > 0$ on F. But then

$$\int_F \sum_i \hat{U}f_i(y)\tilde{\nu}_n(dy) = \int \sum_1 f_i(x)p_n(x)dx \uparrow \int \sum_i f_i(x)p(x)dy < \infty$$

as $n \to \infty$. Hence $\tilde{\nu}_n(F) \leqq M < \infty$ for some finite $M > 0$, and $\{\tilde{\nu}_n\}$ is bounded set in $M^+(F)$ and hence is pre-compact in the vague topology. There exists therefore a finite Radon measure $\nu \in M^+(F)$ and a subsequence $\{n'\}$ such that $\tilde{\nu}_{n'}(g) \to \nu(g)$ for all $g \in C(F)$. Since $\hat{U}f \in C(F)$ for $f \in C_K^+(E)$, we have

$$\int \hat{U}f(y)\tilde{\nu}_{n'}(dy) = \int f(x)p_{n'}(x)dx \uparrow \int_F \hat{U}f(u)\nu(du) = \int f(x)p(x)dx .$$

Now $\hat{U}f(u) = \int f(x)u(x)dx$ for $u \in F$ and therefore

$$\int f(x)p(x)dx = \int\Big(\int f(x)u(x)dx\Big)\nu(du) = \int f(x)\Big(\int u(x)\nu(du)\Big)dx .$$

Here we use the joint measurability of the function $(x, u) \to u(x)$ and Fubini's theorem. Since this equation holds for all $f \in C_K^+(E)$, it follows that $p(x) = \int_F u(x)\nu(du)$ a.e., and hence everywhere since each function is excessive by Corollary (4.5). Since $\nu(F) < \infty$ and $F \subset \mathscr{K}$, the same Corollary implies that $p \in \hat{\mathscr{S}}$, thus completing the proof of Theorem (5.3).

Recall that an excessive function $v \in \mathscr{S}$ is said to be harmonic if $P_B v = v$ whenever B is the complement of a compact subset of E. Now according to [2, p. 272], each $u \in \mathscr{S}$ has a unique representation of the form $u = p + v$ where p is a potential and v is an harmonic excessive function: The reader can easily convince himself that the proof given in the cited reference is equally valid under our assumptions. If we let $\mathscr{R} = \{u \in ex\mathscr{S}: u \text{ is harmonic}\}$ and $P = \{u \in ex\mathscr{S}: u \text{ is a potential}\}$, then the following corollary is an immediate consequence of the above fact

COROLLARY 5.4. (i) *Each $u \in \mathscr{S}$ has a unique representation of the form $u(x) = \int u(x, y)\nu(dy) + v(x)$ where $\nu \geqq 0$ is a finite Borel measure on $\hat{E}$ and $v \in \mathscr{S}$ is harmonic.*

(ii) *$ex\mathscr{S} = P \cup \mathscr{R}$. Of course, $P \cap \mathscr{R} = \{0\}$.*

REMARK 5.5. In § 3 we introduced the assumption (4.6) and we now show how to obtain a representation as in Theorem (5.3) under the single assumption that to each $x \in E$ there is at most one $u \in ex\mathscr{S}$ having support at x. Define $\hat{E} = \{x \in E$: there is a $u \in \hat{\mathscr{P}}$ having support at $x\}$ and write $x \sim y$ if and only if there is $u \in \hat{\mathscr{P}}$ having support at x and y. It is easy to see that $\sim$ is an equivalence relation on $\hat{E}$ and we put $\tilde{E} = \hat{E}/\sim$, the set of equivalence classes of $\hat{E}$. We denote by $\tilde{x}$ the equivalence class containing x. If we define $\tilde{\Gamma}: \tilde{E} \to \hat{\mathscr{P}}$ by $\tilde{\Gamma}(\tilde{x})$ = the unique $u \in \hat{\mathscr{P}}$ having support at x, then $\tilde{\Gamma}$ is one-one onto $\hat{\mathscr{P}}$, and the metric d on E defined by $d(\tilde{x}, \tilde{y}) = \rho(\tilde{\Gamma}(\tilde{x}), \tilde{\Gamma}(\tilde{y}))$, where ρ is the metric on $\hat{\mathscr{P}}$, endows $\tilde{E}$ with a topology that makes $\tilde{\Gamma}$ a homeomorphism between $\mathscr{P}$ and $\tilde{E}$. Imitating the proof of Theorem (4.7) we obtain an analogous representation with the space $\hat{E}$ replaced by $\tilde{E}$. Of course $\tilde{E}$ is no longer a subset of E, but rather a set of equivalence classes of points of E. Note that $\hat{\mathscr{P}}$ is separating if and only if $x \sim y$ implies that $x = y$.

REMARK 5.6. Denote by $\hat{E}'$ the subset $\hat{E} \subset E$ equipped with the subspace topology, i.e., the topology induced by E. A natural question to ask is if there is any relation between $\hat{E}'$ and $\hat{E} = \hat{\mathscr{P}}$ as topological spaces. We show that is a dual process exists as in Chapter VI of [2], then the map $\Gamma': \hat{\mathscr{P}} \to E'$ defined by $\Gamma'(u_x) = x$ is a homeomorphism so that $\hat{E} = \hat{E}'$ as topological spaces. Now the dual process $\tilde{X}_t$ has a potential operator $\tilde{U}$ of the form $\tilde{U}f(y) = \int g(x, y)f(x)dx$, and it follows from [7, Chap. III, T7 and T10] that $g(x, y) = u(x, y)$ for $y \in \hat{E} = \hat{\mathscr{P}}$. In other words, $\hat{E} = \{y \in E: x \to g(x, y)$ is an extremal potential$\}$ and therefore $\hat{U}f(y) = \tilde{U}f(y)$ for all $y \in \hat{E}$ and $f \in C_K(E)$. If $u_{y_n} \to u_{y_0}$ in $\hat{\mathscr{P}}$, then $\hat{U}f(y_n) = \tilde{U}f(y_n) \to \tilde{U}f(y_0) = \hat{U}f(y_0)$ for each $f \in C_K(E)$. Now it is easy to see that the operator $\tilde{U}: C_K(E) \to C(E)$ has an image which separates points of E so that $y_n \to y_0$ in E, hence E'. Thus Γ' is continuous. On the other hand, if $y_n \to y_0$ in E' then $\hat{U}f(y_n) = \tilde{U}f(y_n) \to \tilde{U}f(y_0) = \hat{U}f(y_0)$ for all $f \in C_K(E)$ by the continuity of $\tilde{U}f$. Thus $u(x, y_n) \to u(x, y_0)$ in $\hat{\mathscr{P}}$ and Γ'^{-1} is continuous, proving that Γ' is a homeomorphism.

REFERENCES

1 H. Bauer, *Harmonishe Raume and ihre Potentialtheorie,* Lecture Notes in Mathematics, 22. Springer-Verlag, Berlin, 1966.

2. R. M. Blumenthal and R. K. Getoor, *Markov Processes and Potential Theory,* Academic Press, New York, 1968.

3. G. Choquet, *Lectures on Analysis* (3 volumes), W. A. Benjamin, Inc., New York, 1969.

4. R. M. Hervé, *Rechrches axiomatiques sur la théorie des fonctions surharmoniques et du potentiel,* Ann. Inst. Fourier **12**, (1962), 415-571.

5. P. A. Meyer, *Processus de Markov,* Lecture Notes in Mathematics, 26, Springer-Verlag, Berlin, 1967.

6. ———, *Probabilitiés et Potential,* Hermann, Paris, 1966.

7. ———, *Processus de Markov: la frontière de Martin,* Lecture Notes in Mathematics, 77. Spring-Verlag, Berlin, 1968.

Received September 1, 1970. This paper was part of a thesis submitted in partial fulfillment of the requirements for the Ph. D. degree at the University of California, San Diego. The author wishes to express his appreciation to Professor R. K. Getoor for his advice and encouragement in the preparation of this paper.

UNIVERSITY OF MONTREAL

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

AN EXTENSION OF SOME RESULTS OF TAKESAKI IN THE REDUCTION THEORY OF VON NEUMANN ALGEBRAS

GEORGE A. ELLIOTT

Briefly, the results in this paper are that both for measurable fields of von Neumann algebras and for families of measurable fields of operators, pointwise isomorphism implies isomorphism.

In the special case when half the measurable fields considered are constant, these results were established by Takesaki. If the Borel space on which the fields are defined is standard, the results can be established by classical means; in the case considered by Takesaki they are due to von Neumann.

For the results of the present paper, two new tools seem to be needed. The first is a measurable choice theorem of Aumann which generalizes the classical one. This has already been applied to reduction theory by Flensted-Jensen. The second is a criterion for a von Neumann algebra containing the diagonal operators to be decomposable: it should consist of decomposable operators. This answers a question of Dixmier.

We shall use the terminology of reduction theory developed in [2], Chapitre II.

2. LEMMA (Aumann). *Let T be a Borel space and let X be a standard Borel space. Let G be a Borel subset of $T \times Y$ such that the projection of G onto T is all of T. Let there be given a finite measure on T. Then there exists a measurable map $g\colon T \rightarrow X$ such that $(t, g(t)) \in G$ for almost all $t \in T$.*

Proof. See [1]. The proof is by reduction to the case that T is standard.

3. THEOREM. *Let T be a Borel space, and suppose given a finite measure on T and a measurable field of Hilbert spaces on T with direct integral H. Let A and B be decomposable von Neumann algebras in H. If for each $t \in T$ there is a spatial isomorphism of $A(t)$ onto $B(t)$ then there exists a decomposable spatial isomorphism of A onto B. This statement also holds with the word "spatial" removed.*

Proof. The proof of the first assertion is the same as the proof of Lemma 2 on page 179 of [2], with the exception that 2 above is used instead of the more well known measurable choice theorem for standard measures.

The second assertion is reduced to the first by tensoring with the scalars on a separable infinite dimensional Hilbert space, just as in Theorem 3 of [4].

4. LEMMA. *Let T be a Borel space, and suppose given a finite measure on T and a measurable field of Hilbert spaces on T with direct integral H. Then a von Neumann algebra in H containing the diagonal operators is decomposable if it consists of decomposable operators.*

This answers affirmatively the question on page 174 of [2].

Proof. We may suppose that the field of Hilbert spaces is constant. By [2], page 178, Corollaire, it then follows that the algebra of all decomposable operators is spatially isomorphic to $Z \otimes B$ with Z a commutative algebra and B the algebra of all operators on a separable Hilbert space. Since the algebra of diagonal operators is countably decomposable (the measure of T is finite), so is Z; therefore both Z and B and hence also $Z \otimes B$ have a countable separating set of vectors.

Let $\xi_1, \xi_2, \cdots$ be a countable separating set of vectors for the algebra of all decomposable operators, such that $\Sigma \|\xi_i\|^2 < \infty$. Then for any operator x we have $(\Sigma \|x\xi_i\|^2)^{1/2} < \infty$. On the algebra of decomposable operators this expression defines a norm, which on bounded sets determines the strong topology. We shall denote this norm by N.

Let A be a von Neumann algebra containing the algebra of diagonal operators, and consisting of decomposable operators. To show that A is decomposable, we must show that A is countably generated over the algebra of diagonal operators ([2], page 174, Théorème 2). Writing as before the algebra of decomposable operators as $Z \otimes B$ with Z commutative and B a type I_n factor, n countable, let $x_1, x_2, \cdots$ be a sequence strongly dense in the unit ball of $1 \otimes B$. For each $k = 1, 2, \cdots$ let y_k be an element of A which is at minimal distance from x_k with respect to the norm N of the preceding paragraph (such y_k exists because bounded weakly closed sets of A are weakly compact, and N is weakly lower semicontinuous). Then $y_1, y_2, \cdots$ generate A over the algebra of diagonal operators. For if e is a diagonal projection then for each $k = 1, 2, \cdots$ the distance from ey_k to ex_k with respect to N is minimal. Hence, if $e_1, \cdots, e_p$ are diagonal projections with sum 1, and if $k_1, \cdots, k_p = 1, 2, \cdots$ then the distance of $e_1 y_{k_1} + \cdots + e_p y_{k_p}$ to $e_1 x_{k_1} + \cdots + e_p x_{k_p}$ with respect to N is minimal. The assertion follows, because the operators $e_1 x_{k_1} + \cdots + e_p x_{k_p}$ as above are strongly dense in the unit ball of decomposable operators, and the strong topology on this unit ball is metrized by N.

5. THEOREM. *Let T be a Borel space, and suppose given a finite measure on T and a measurable field of Hilbert spaces on T with direct integral H. Let (x_i) and (y_i) be families of decomposable operators in H. If for each $t \in T$ the families $(x_i(t))$ and $(y_i(t))$ are simultaneously unitarily equivalent, then (x_i) and (y_i) are simultaneously unitarily equivalent, with the equivalence implemented by a decomposable unitary operator.*

Proof. Suppose first that the families (x_i) and (y_i) are countable. Then the conclusion may be deduced as in A 82, page 348 of [3], using again 2 instead of the classical measurable choice theorem.

If the families (x_i) and (y_i) are not countable, by 4 it is still true that the von Neumann algebra A generated by the x_i and the diagonal operators is decomposable. It follows that there exists a countable family $w_1, w_2, \cdots$ in A generating A over the diagonal operators. We may suppose that the x_i form a sub involutive algebra, containing the diagonal operators. Then the x_i are strongly dense in A, and the x_i of norm $\leqq 1$ are strongly dense in the unit ball of A. As shown in the proof of 4, the strong topology on the unit ball of A is metrizable. We may suppose that $w_1, w_2, \cdots$ lie in the unit ball of A. Then there exists a countable subfamily of (x_i) which generates A over the diagonal operators (namely, the union of sequences converging strongly to each w_k, $k = 1, 2, \cdots$). By the first paragraph of the proof this countable subfamily of (x_i) is simultaneously unitarily equivalent to the corresponding subfamily of (y_i), by a decomposable unitary operator, say v.

We claim that $vx_iv^* = y_i$ for every i. The subfamily of (x_i) such that $vx_iv^* = y_i$ contains the diagonal operators and also a set (the above countable subfamily) which generates A over the diagonal operators. It therefore contains a sub involutive algebra dense in A. By metrizability of the strong topology on bounded sets, this subfamily is closed under strong limits (use Proposition 4, page 160 of [2]). It follows that $vx_iv^* = y_i$ for all i.

6. REMARKS. Once 5 has been reduced by use of 4 to the case that the families are countable, the proof can also be finished by a variant of the method of Takesaki, in [4] (in which not just a measur-

able choice but a Borel choice is made).

On the other hand, although Takesaki was able to prove his special case of 3 by a Borel choice argument, the author does not see how to extend this approach and was forced to be content in the proof of 3 with making a measurable choice.

REFERENCES

1. R. J. Aumann, *Measurable utility and measurable choice theorem* (Research program in Game Theory and Mathematical Economics, No. 30), Hebrew University, Jerusalem, 1967.

2. J. Dixmier, *Les algèbres d'opérateurs dans l'espace hilbertien*, 2e édition, Gauthier-Villars, Paris, 1969.

3. ———, *Les C*-algèbres et leurs représentations*, 2e édition, Gauthier-Villars, Paris, 1969.

4. M. Takesaki, *Remarks on the reduction theory of von Neumann algebras*, Proc. Amer. Math. Soc., **20** (1969), 434-438.

Received December 22, 1970.

QUEEN'S UNIVERSITY AT KINGSTON, CANADA
AND
THE INSTITUTE FOR ADVANCED STUDY, PRINCETON, NEW JERSEY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

DIRECTED GRAPHS AS UNIONS OF PARTIAL ORDERS

PETER C. FISHBURN AND JOEL H. SPENCER

The index of an irreflexive binary relation R is the smallest cardinal number $\sigma(R)$ such that R equals the union of $\sigma(R)$ partial orders. With $s(n)$ the largest index for an R defined on n points, it is shown that $s(n)/\log_2 n \to 1$ as $n \to \infty$. The index function is examined for symmetric R's and almost transitive R's, and a characterization for $\sigma(R) \leqq 2$ is presented. It is shown also that

$$\inf \{n: s(n) > 3\} \leqq 13 ,$$

but the exact value of $\inf \{n: s(n) > 3\}$ is presently unknown.

1. Introduction. A binary relation on a set X is a subset of ordered pairs xy in $X \times X$. A directed graph (hereafter *digraph*[1]) $G = (X, R)$ is a nonempty set X and an irreflexive ($xx \notin R$) binary relation R on X. If $\phi \subset Y \subseteqq X$ then $G \mid Y$ is the digraph obtained from $G = (X, R)$ by deleting all points in X-Y.

A *partial order* P on X is an irreflexive and transitive ($xy \in P$ & $yz \in P \Rightarrow xz \in P$) binary relation on X. A digraph $G = (X, R)$ is *resolved* by a set of partial orders on X if and only if R equals the *union* of the partial orders in the set. Since $\{xy\}$ is a partial order when $xy \in R$, every G is resolved by some set of partial orders.

The *index*[2] of a digraph $G = (X, R)$ is the smallest cardinal number $\sigma(R)$ such that R is resolved by $\sigma(R)$ partial orders on X. Clearly $\sigma(R) = 1$ if and only if R is a partial order. $\sigma(\{ab, ba\}) = 2$, and $\sigma(R) = 3$ for the cyclic triangle $R = \{ab, bc, ca\}$. The smallest X that we know of that admits an R with $\sigma(R) = 4$ has 13 points. (See Figure 1.) In connection with a later characterization of $\sigma \leqq 2$ we present an R with $\sigma(R) = 2$ where R cannot be the union of two *disjoint* partial orders.

Our definition of $\sigma(R)$ is motivated by Dushnik and Miller's definition [2] of the *dimension of a partial order* P on X as the smallest cardinal number $D(P)$ such that P equals the *intersection* of $D(P)$ linear orders on X. *A linear order* L on X is a complete ($x \neq y \Rightarrow xy \in L$ or $yx \in L$) partial order, and a *chain* in X is a linear

[1] We shall sometimes refer to a binary relation as a digraph, omitting explicit mention of the set on which the relation is defined.

[2] It is tempting to use "dimension" instead of "index," but since the former term is used for a number of other concepts in the theory of binary relations we favor the latter here. It would be proper to write $\sigma(G)$ instead of $\sigma(R)$, but since $\sigma(R) = \sigma(R')$ if R is isomorphic to R' the specific omission of X will cause no problems.

order on a subset of X. A number of facts about $D(P)$ are summarized in [1], which gives other references.

This paper examines the index function σ for digraphs. The next section focuses on large values for $\sigma(R)$. Our first theorem, based on a theorem in Folkman [4], shows that $\sigma(R)$ can be arbitrarily large for both *symmetric* ($xy \in R \Rightarrow yx \in R$) and asymmetric ($xy \in R \Rightarrow yx \notin R$) digraphs. The second theorem examines the behavior of σ in the following way. Let

$$s(n) = \sup\{\sigma(R)\colon R \text{ is an irreflexive binary relation on } n \text{ points}\} ,$$

the largest σ for a digraph with n points. When u is a real-valued function on $\{1, 2, \cdots\}$ and $u(n)$ remains bounded as n gets large, we write $u = 0(1)$ according to popular convention. Theorem 2 states that

$$\log_2 n - \frac{1}{2}\log_2\log_2 n + 0(1) \geqq s(n) \geqq \log_2 n - \frac{3}{2}\log_2\log_2 n - 0(1) .$$

This gives another proof that σ can be arbitrarily large, and shows that $s(n)/\log_2(n)$ approaches 1 as n gets large.

The rest of the paper is mostly concerned with small values of σ. Section 3 presents an (X, R) with $|X| = 13$ and $\sigma(R) = 4$. We do not presently know the smallest X that admits an R with $\sigma(R) = 4$.

Symmetric digraphs (X, S) are examined in § 4, where we give a necessary and sufficient condition for $\sigma(S) \leqq 2$. Suppose that P is a partial order on X and

$$S = \{xy\colon xy \in X \times X \;\&\; x \neq y \;\&\; xy \notin P \;\&\; yx \notin P\} .$$

Then S is a symmetric digraph. We note that when S is defined in this way, then $D(P) \leqq 2$ if and only if $\sigma(S) \leqq 2$, and

$$D(P) \leqq n \Rightarrow \sigma(S) \leqq 2(n-1) .$$

The question of whether $\sigma(S) \leqq n \Rightarrow D(P) \leqq f(n)$ for some function f is presently open.

A binary relation R is almost transitive[3] if and only if $(ab \in R \;\&\; bc \in R \;\&\; a \neq c) \Rightarrow ac \in R$. Section 5 proves that $\sigma(R) \leqq 2$ when R is an almost transitive digraph.

Section 6 then gives a general characterization of $\sigma(R) \leqq 2$ that is stated in terms of a partition of the subset of R whose elements

[3] Harary, Norman and Cartwright [7, p. 7] call this transitivity, but we use the modifier to distinguish it from the more common use of "transitivity" in which a, b and c do not have to be distinct.

are involved in nontransitive adjacent pairs such as xy, $yz \in R$ & $xz \notin R$.

2. Digraphs with large indices.

THEOREM 1. *If n is a positive integer then there are asymmetric and symmetric digraphs whose indices exceed n.*

Our proof is based on a specialization of Theorem 2 in Folkman [4]. A *graph* (X, E) is a nonempty set X and a set E of unordered pairs $\{x, y\}$ with $x, y \in X$ and $x \neq y$. A *triangle of* (X, E) is a set $\{\{a, b\}, \{b, c\}, \{a, c\}\} \subseteqq E$. A *partition* of X is a set of mutually disjoint subsets of X whose union equals X.

LEMMA 1 (Folkman). *Let m be a positive integer. Then there is a graph (X, E) that includes no triangles, and every partition $\{C_1, \cdots, C_k\}$ of X with $k \leqq m$ contains a C_i such that $a, b \in C_i$ for some $\{a, b\} \in E$.*

Proof of Theorem 1. Let (X, E) be such a graph for $m = 2^n$. Let (X, R) be any digraph for which $xy \in R$ or $yx \in R$ if and only if $\{x, y\} \in E$. Suppose that R is the union of partial orders $P_1, \cdots, P_n$ on X. Since E has no triangles, any subset of a P_i is a partial order and hence we can assume $P_i \cap P_j = \varnothing$ when $i \neq j$. Letting $A(x) = \{i$: for some $y \in X$, $xy \in P_i\}$, partition X so that x and y are in the same element of the partition if and only if $A(x) = A(y)$. The number of elements in the partition does not exceed 2^n. Thus, by Lemma 1, the partition contains an element Y with $x, y \in Y$ and $\{x, y\} \in E$. Then $A(x) = A(y)$. Since $xy \in R$ or $yx \in R$, take $xy \in P_j$ for definiteness with $j \in A(x)$. Since $j \in A(y)$ also, there is a $z \in X$ such that $yz \in P_j$. Transitivity then implies that $xz \in P_j$ and hence that E includes a triangle, which contradicts our initial hypothesis. Therefore $\sigma(R) > n$. By the definition of R it can be taken to be either asymmetric or symmetric (or neither).

Henceforth in this section all logarithms are to base 2 unless indicated otherwise. $[r]$ = (largest integer $\leqq r$) and $\{r\}$ = (smallest integer $\geqq r$).

THEOREM 2. $\log n - 1/2 \log\log n + 0(1) \geqq s(n) \geqq \log n - 3/2 \log\log n - 0(1)$.

We show first the upper bound, using two preparatory lemmas.

LEMMA 2. *In any digraph $G = (H, R)$ with $|H| = m$ there exists $D \subseteqq H$ such that $|D| \geqq \{\log_4 m\} = \{1/2 \log m)$ and $\sigma(G|D) \leqq 2$.*

Proof. We use induction on m, the lemma being obvious for small values of m. Fix $x \in H$. Split $H^* = H - \{x\}$ into four parts:

$$T_1 = \{y \in H^*: xy \notin R \ \& \ yx \notin R\} \qquad S_1 = \varnothing$$

$$T_2 = \{y \in H^*: xy \in R \ \& \ yx \notin R\} \qquad S_2 = \{x\} \times D_2$$

$$T_3 = \{y \in H^*: xy \notin R \ \& \ yx \in R\} \qquad S_3 = D_3 \times \{x\}$$

$$T_4 = \{y \in H^*: xy \in R \ \& \ yx \in R\} \qquad S'_4 = \{x\} \times D_4 , \quad S''_4 = D_4 \times \{x\} .$$

Some $|T_i| \geqq \{(m-1)/4\}$. By induction find $D_i \subseteqq T_i$ with

$$|D_i| \geqq \{\log_4 |T_i|\} \geqq \{\log_4 \{(m-1)/4\}\} = \{\log_4 m\} - 1$$

and $G|D_i = P_1 \cup P_2$. Then set $D = D_i \cup \{x\}$. $G|D = (P_1 \cup S_i) \cup (P_2 \cup S_i)$ except for $i = 4$ when $G|D = (P_1 \cup S'_4) \cup (P_2 \cup S''_4)$.

LEMMA 3. *In any digraph $G = (X, R)$ with $|X| = n$ there is a partition $\{D_1, \cdots, D_t\}$ of X such that $t < 3n/\log n$ and $\sigma(G|D_i) \leqq 2$ for each i.*

Proof. Given G, by Lemma 2 find D_1 such that

$$|D_1| = x_1 \geqq \{\log_4 n\} .$$

By induction find D_i such that

$$|D_i| = x_i \geqq \left\{\log_4\left(n - \sum_{j=1}^{i-1} x_j\right)\right\} .$$

From elementary calculus we can show $\sum_{i=1}^{t} x_i \geqq n$ for

$$t \leqq (2+\varepsilon)n/\log n .$$

We now show the upper bound for Theorem 2. Let $G = (X, R)$ with $|X| = n$. Take $D_1, \cdots, D_t$ as in Lemma 3. Let $\{A_i^*, B_i^*\}$ be a partition of $\{1, \cdots, t\}$ for $i = 1, \cdots, s$ such that for all $1 \leqq j \neq k \leqq t$ there exists $i, 1 \leqq i \leqq s$, such that $j \in A_i^*$ & $k \in B_i^*$. By Spencer [12] we may take

$$s = \log t + 1/2 \log\log t + 0(1) \leqq \log n - 1/2 \log\log n + 0(1) .$$

$\{A_i^*, B_i^*\}$ induces a partition $\{A_i, B_i\}$ of X with

$$A_i = \bigcup_{j \in A_i^*} D_j , \qquad B_i = \bigcup_{j \in B_i^*} D_j .$$

Then set

$$P_i = \{xy\colon x \in A_i \ \&\ y \in B_i \ \&\ xy \in R\} \quad \text{for } i = 1, \cdots, s\,.$$

Since $\sigma(G \mid D_i) \leqq 2$, $G \mid D_i = P'_i \cup P''_i$. Set

$$P' = \bigcup_{i=1}^{s} P'_i,\ P'' = \bigcup_{i=1}^{s} P''_i\,.$$

Then $R = P' \cup P'' \cup P_1 \cup \cdots \cup P_s$, giving the upper bound of Theorem 2.

We turn to the lower bound of the theorem, again using two preliminary lemmas. A complete asymmetric digraph is a *tournament*.[4] We shall show that a "random" tournament $T = (X, R)$ with $|X| = n$ has $\sigma(T) \geqq \log n - 3/2 \log\log n - 0(1)$. Intuitively speaking, we show that all $P \subseteqq T$ are essentially bipartite.

Let T^n be the set of tournaments with $X = \{1, 2, \cdots, n\}$. We say that $T = (X, R) \in T^n$ has *property* α if and only if there are A, $B \subseteqq X$ with $|A| = |B| \geqq 3 \log n$ and $A \times B \subseteqq R$. T has *property* β if and only if there is an $A \subseteqq X$ and a linear order L on A such that $|A| \geqq (\log n)^2$ and

$$(*) \qquad |R \cap L| \leqq \frac{1}{3}\binom{|A|}{2}\,.$$

LEMMA 4. *For n sufficiently large there exists $T \in T^n$ satisfying neither property α nor property β.*

Proof. If $T \in T^n$ has property α, there are $A, B \subseteqq X$ with $|A| = |B| = [3 \log n]$ and $A \times B \subseteqq R$. Set $t = [3 \log n]$. For fixed A and B, 2^{-t^2} is the proportion of $T \in T^n$ that satisfy this condition. There are less than n^{2t} choices of A and B, so less than $n^{2t}2^{-t^2}$ of the $T \in T^n$ satisfy α. $n^{2t}2^{-t^2} \to 0$ as $n \to \infty$.

If $T \in T^n$ has property β, there exists $A \subseteqq X$ and L on A such that $|A| = [(\log n)^2]$ and (*) holds. There are less than $n^{(\log n)^2}$ choices of A and then $[(\log n)^2]!$ choices of L. Given A and L, the proportion of $T \in T^n$ satisfying (*) is the probability of at most $\binom{t}{2}/3$ heads in $\binom{t}{2}$ flips of a fair coin where $t = |A| \sim (\log n)^2$. This probability is approximately $p^{-\binom{t}{2}}$ where $p = 3^{1/3}\ (3/2)^{2/3} > 1$. Thus the proportion of $T \in T^n$ satisfying β is less than

$$n^{(\log n)^2}\,[(\log n)^2]!\ p^{-\binom{t}{2}},\ \text{which} \longrightarrow 0 \text{ as } n \longrightarrow \infty\,.$$

Thus for n sufficiently large some $T \in T^n$ can satisfy neither α nor β.

---

[4] See Moon [**9**] for extensive discussion of tournaments. See also [**3**, **10**, **11**] for resulted to the present paper.

LEMMA 5. *If* $T_1, \cdots, T_n \subseteqq \{1, \cdots, s\}$ *then there are* $n/\binom{s}{s/2}$ T_i *which are mutually comparable.*[6]

Proof. We use a technique due to Lubell [8]. There are $s!$ maximal chains of subsets of $\{1, \cdots, s\}$ under the ordering of $\subset$. If $|T_i| = a$ then T_i is in $a!\ (s-a)! \geqq (s/2)!^2 = s!/\binom{s}{s/2}$ maximal chains. Thus some maximal chain must contain $n\,[s!/\binom{s}{s/2}]/s!$ T_i.

In the following proof of the lower bound of Theorem 2 we use the fact that $1/\binom{s}{s/2} \sim \sqrt{\pi/2}\ \sqrt{s}\ 2^{-s}$.

Let $G = (X, R)$ be a tournament that satisfies neither α nor β (Lemma 4). Suppose that $R = P_1 \cup \cdots \cup P_s$. Define

$$\begin{aligned} W_i &= \{x \in X\colon |\{y \in X\colon xy \in P_i\}| > 3\ \log n\} \\ L_i &= \{x \in X\colon |\{y \in X\colon yx \in P_i\}| > 3\ \log n\} \\ R_i &= X - W_i - L_i \end{aligned}$$

for $1 \leqq i \leqq s$. (We split X into winners, losers, and the rest.) By Lemma 4, $W_i \cap L_i = \varnothing$. For $x \in X$ set

$$T_x = \{i\colon x \in W_i \cup R_i\} \subseteqq \{1, \cdots, s\}\ .$$

By Lemma 5 find $V \subseteqq X$ such that $|V| \geqq n\sqrt{\pi/2}\sqrt{s}\ 2^{-s}$ and $T_x \subseteqq T_y$ or $T_y \subseteqq T_x$ whenever $x, y \in V$. Induce a linear order L on V by setting $xy \in L$ if $T_x \subset T_y$: when $T_x = T_y$, L is defined in any fixed manner.

Now assume $s < \log n - 3/2 \log\log n - 7$. Then $|V| \geqq 2^7\sqrt{\pi/2}$ $(\log n)^2$. Set

$$Z_i = L \cap P_i \qquad\qquad 1 \leqq i \leqq s\ .$$

Given $xy \in Z_i$, $T_x \subseteqq T_y$ so that we cannot have $x \in W_i$ & $y \in L_i$. And since $W_i \cap L_i = \varnothing$ we cannot have $x \in L_i$ & $y \in W_i$. Therefore

$$Z_i = \{xy \in Z_i\colon x \text{ or } y \in R_i\} \cup \{xy \in Z_i\colon x, y \in W_i\} \cup \{xy \in Z_i\colon x, y \in L_i\}\ .$$

There are at most $6 \log n\ |V|$, $3\ \log n\ |V|$ and $3\ \log n\ |V|$ ordered pairs in the first, second and third parts respectively of this decomposition of Z_i. Thus $|Z_i| \leqq 12\ \log n\ |V|$. Since G does not have property β it follows that

$$\frac{1}{3}\binom{|V|}{2} \leqq |R \cap L| \leqq \sum_{i=1}^{n} |Z_i| \leqq 12\,(\log n)^2\,|V|$$

and hence that $|V| \leqq 72\,(\log n)^2 + 1$. Since this contradicts $|V| \geqq 2^7$ $\sqrt{\pi/2}\,(\log n)^2$ it must be true that $s \geqq \log n - 3/2 \log\log n - 0(1)$.

---

[5] T_i and T_j are mutually comparable if and only if $T_i \subseteq T_j$ or $T_j \subseteq T_i$.

This completes the proof of Theorem 2.

If a sufficiently good bound could be placed on

$$\{xy \in P_i\colon\ x \text{ or } y \in R_i \text{ or } x, y \in W_i \text{ or } x, y \in L_i\}$$

then one could prove $s(n) = \log n - 1/2 \log\log n + o(\log\log n)$. One might even show that $s(n) = \log n - 1/2 \log\log n + 0(1)$.

3. A digraph with $\sigma = 4$ and $|X| = 13$. Although the theorems of the preceding section show that there are digraphs with large indices, they are of little use in attempting to discover the smallest X that admits an R for which $\sigma(R) = n$. Figure 1 shows the smallest X that we know of for which $\sigma(R) = 4$.

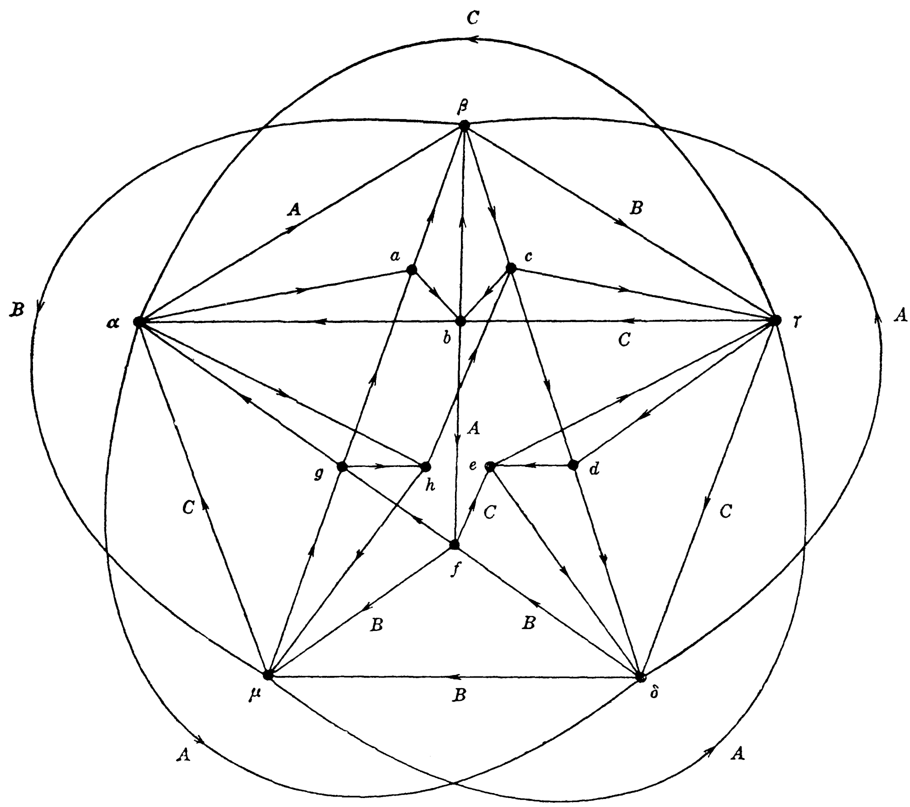

FIGURE 1

Assume that $\sigma(R) = 3$ for Figure 1, with A, B and C three partial orders whose union equals R. Then one of A, B and C must contain exactly one of $\alpha\beta$, $\beta\gamma$, $\gamma\delta$, $\delta\mu$ and $\mu\alpha$ and the other two must each contain exactly two of these ordered pairs in alternating fashion.

Suppose for example that $\alpha\beta \in A$, $\beta\gamma \in B$, $\gamma\delta \in C$, $\delta\mu \in B$, $\mu\alpha \in C$. Then $\gamma\alpha$, $\delta\beta$, $\mu\gamma$, $\alpha\delta$, and $\beta\mu$ must be respectively in C, A, A, A, and B. Then $\gamma b \in C$ and δf, $f\mu \in B$. Since $\gamma b \in C$ and $f\mu \in B$, $bf \in A$. Since $bf \in A$ and $\delta f \in B$, $fe \in C$. By the cyclic triangle $\{fe, e\delta, \delta f\}$, $e\delta$ must be in A. But since $\delta\beta \in A$ this implies $e\beta \in A$, which is false. A similar contradiction to $\sigma = 3$ is obtained when any alternative assignment is made for $\alpha\beta, \beta\gamma, \cdots, \mu\alpha$.

4. Indices of symmetric digraphs. In this section we consider symmetric ($xy \in S \Rightarrow yx \in S$) digraphs (X, S). For any binary relation R, $R^* = \{xy\colon yx \in R\}$, the *converse* or *dual of* R.

A graph (X, E) is a *comparability graph* if and only if there is a partial order P on X such that $\{x, y\} \in E$ if and only if $xy \in P \cup P^*$. Ghouila-Houri [5] and Gilmore and Hoffman [6] provide characterizations of comparability graphs. When (X, S) is a symmetric digraph, $(X, E(S))$ will denote the graph in which $\{x, y\} \in E(S)$ if and only if $xy \in S$.

THEOREM 3. *Suppose that (X, S) is a symmetric digraph. Then $\sigma(S) \leqq 2$ if and only if $(X, E(S))$ is a comparability graph.*

Proof. If $(X, E(S))$ is a comparability graph then $S = P \cup P^*$ for a partial order P, and thus $\sigma(S) \leqq 2$. Conversely, if $S = P_1 \cup P_2$ with P_1 and P_2 partial orders, then $P_2 = P_1^*$.

In [1] it is shown that if (X, P) is a transitive digraph (so that P is a partial order) and if $S = \{xy\colon x \neq y \ \&\ xy \notin P \cup P^*\}$ then $D(P) \leqq 2$ if and only if $(X, E(S))$ is a comparability graph. Hence, as a corollary to Theorem 3 we have $D(P) \leqq 2$ if and only if $\sigma(S) \leqq 2$. Our next theorem extends this in one direction.

THEOREM 4. *Suppose that P on X is a partial order and let $S = \{xy\colon x \neq y \ \&\ xy \notin P \cup P^*\}$. Then $D(P) \leqq n \Rightarrow \sigma(S) \leqq 2(n-1)$ for $n > 1$.*

Proof. The theorem is true for $n = 2$. Using induction, assume it's true for all $n < m$ and suppose $D(P) = m$ with $P = \bigcap_1^m L_i$ where each L_i is a linear order. Let $P' = \bigcap_2^m L_i$ and

$$S' = \{xy\colon x \neq y \ \&\ xy \notin P' \cup (P')^*\} .$$

Since $D(P') \leqq m - 1$, the induction hypothesis gives $\sigma(S') \leqq 2(m-2)$. Clearly $S' \subseteqq S$ and $S - S' = (P' \cap L_1^*) \cup ((P')^* \cap L_1)$. Since $P' \cap L_1^*$ is a partial order (the intersection of two partial orders) and $(P')^* \cap L_1$ is a partial order, $\sigma(S) \leqq \sigma(S') + 2 \leqq 2(m-2) + 2 = 2(m-1)$.

5. **Almost transitive digraphs.** The proof of the next theorem has several similarities to Szpilrajn's proof [13] of the theorem that any partial order P on X can be extended to a linear order L with $P \subseteqq L$. We recall that R is almost transitive if and only if $(ab \in R$ & $bc \in R$ & $a \neq c) \Rightarrow ac \in R$.

THEOREM 5. *$\sigma(R) \leqq 2$ if (X, R) is an almost transitive digraph.*

Proof. Assume that (X, R) is an almost transitive digraph. Let $A = \{ab: ab \in R \text{ \& } ba \notin R\}$, the asymmetric part of R. Let $A^+ = \{ab: ab \in A \text{ or } \{aa_1, a_1a_2, \cdots, a_nb\} \subseteqq A$ for distinct $a_1, \cdots, a_n$ in X that are different from a and $b\}$, the almost transitive closure of A. Clearly $A^+ \subseteqq R$ and A^+ is almost transitive.

To show that A^+ is a partial order it suffices to show that it is asymmetric. To the contrary suppose that $xy \in A^+$ and $yx \in A^+$. Then from the definition of A^+ and almost transitivity for R it follows easily that there is a $c \in X$ for which $cx \in A$ and $xc \in R$, which contradicts the definition of A. Hence A^+ is a partial order.

Let $\mathscr{P} = \{P: P$ is a partial order on X & $A^+ \subseteqq P \subseteqq R\}$. It follows easily from Zorn's lemma that there is a $P^* \in \mathscr{P}$ such that $P^* \subset P$ for no $P \in \mathscr{P}$. Letting P^* be maximal in this sense we now prove that

$$ab, ba \in R \Rightarrow ab \in P^* \text{ or } ba \in P^* .$$

To the contrary suppose that each of ab and ba is in R and neither is in P^*. Then let

$$W = \{xy: x \neq y \text{ \& } (xa \in P^* \text{ or } x = a) \text{ \& } (by \in P^* \text{ or } y = b)\} ,$$

and let $V = P^* \cup W$, so that $P^* \subset V$. We show that V is a partial order (clearly $A^+ \subseteqq V \subseteqq R$), thus contradicting the maximality of P^*. V is irreflexive since P^* and W are irreflexive. For transitivity take $xy, yz \in V$. If both xy and yz are in P^* then $xz \in P^*$ by the transitivity of P^*.

Suppose next that $xy \in P^*$ and $yz \in W$. The latter gives $(ya \in P^*$ or $y = a)$, from which $xa \in P^*$ follows, and it gives also $(bz \in P^*$ or $z = b)$, from which $xz \in V$ follows unless $x = z$. But if $x = z$ we have $xa \in P^*$ and $(bx \in P^*$ or $x = b)$, which give $ba \in P^*$, contradicting the hypothesis that $ba \notin P^*$. Hence $xy \in P^*$ & $yz \in W \Rightarrow xz \in V$. Similarly, $xy \in W$ & $yz \in P^* \Rightarrow xz \in V$.

The final case for transitivity is $xy, yz \in W$. Then $(xa \in P^*$ or $x = a)$ and $(bz \in P^*$ or $z = b)$ so that $xz \in W$ unless $x = z$. But if $x = z$ then $[(xa \in P^* \text{ or } x = a) \text{ \& } (bx \in P^* \text{ } or \text{ } x = b)] \Rightarrow (ba \in P^* \text{ or } b = a)$, which is false. Hence V is a partial order, a contradiction to the maximality of P^*, and therefore

$$ab,\ ba \in R \Longrightarrow ab \in P^* \text{ or } ba \in P^* .$$

Finally, let $Q = R - P^*$ so that $R = P^* \cup Q$. Q is irreflexive since R is irreflexive. Suppose that $xy, yz \in Q$. Then, since both xy and yz are in R but not A, yx and zy are in R and must be in P^* by the preceding analysis. Therefore $zx \in P^*$ and $z \neq x$. Then, by almost transitivity of R, $xz \in R$ and thus $xz \in Q$ since P^* is asymmetric.

Thus $R = P^* \cup Q$, the union of two partial orders.

6. A partition characterization for $\sigma \leqq 2$. Given a digraph (X, R) let K be the set of all ordered pairs of pairs in R that deny transitivity, so that

$$xyKyz \text{ if and only if } xy \in R \ \&\ yz \in R \ \&\ xz \notin R ,$$

and let V be the subset of R involved in these intransitivities so that

$$V = \{xy\colon xyKyz \text{ or } zxKxy \text{ for some } z \in X\} .$$

Suppose that $\sigma(R) \leqq 2$. If $xyKyz$ then xy and yz must be in different resolving partial orders, so that the digraph (V, K) must be bipartite or 2-colorable. Moreover, if xy and yz are in V and in the same resolving partial order and if $xz \in V$ also, then transitivity requires that xz be in this partial order. These two necessary conditions for $\sigma(R) \leqq 2$ are reflected in A1 and A2 of Theorem 6. Their insufficiency for $\sigma(R) \leqq 2$ is noted later. (Note that $\sigma(R) = 1$ if and only if $V = \varnothing$.)

THEOREM 6. *Suppose that (X, R) is a digraph and $V \neq \varnothing$. Then $\sigma(R) = 2$ if and only if V can be partitioned into V_1 and V_2 so that*

A1. $xyKyz \Rightarrow xy$ *and* yz *are in different* V_i ,

A2. $xy, yz \in V_i \ \&\ xz \in V \Rightarrow xz \in V_i$,

A3. $xy \in R - V \Rightarrow$ (1) *and* (2) *do not hold simultaneously*:

(1) $(yz \in V_2 \ \&\ xz \in V_1)$ *or* $(zx \in V_2 \ \&\ zy \in V_1)$, *for some* $z \in X$,

(2) $(yw \in V_1 \ \&\ xw \in V_2)$ *or* $(wx \in V_1 \ \&\ wy \in V_2)$, *for some* $w \in X$.

If $R = P_1 \cup P_2$ then $V_i = P_i \cap V$ for $i = 1, 2$ are easily seen to satisfy A1 through A3, and $V_1 \cap V_2 = \varnothing$.

Before proving sufficiency we show that A1 and A2 are not sufficient for $\sigma = 2$. All directed edges in the 13-point asymmetric

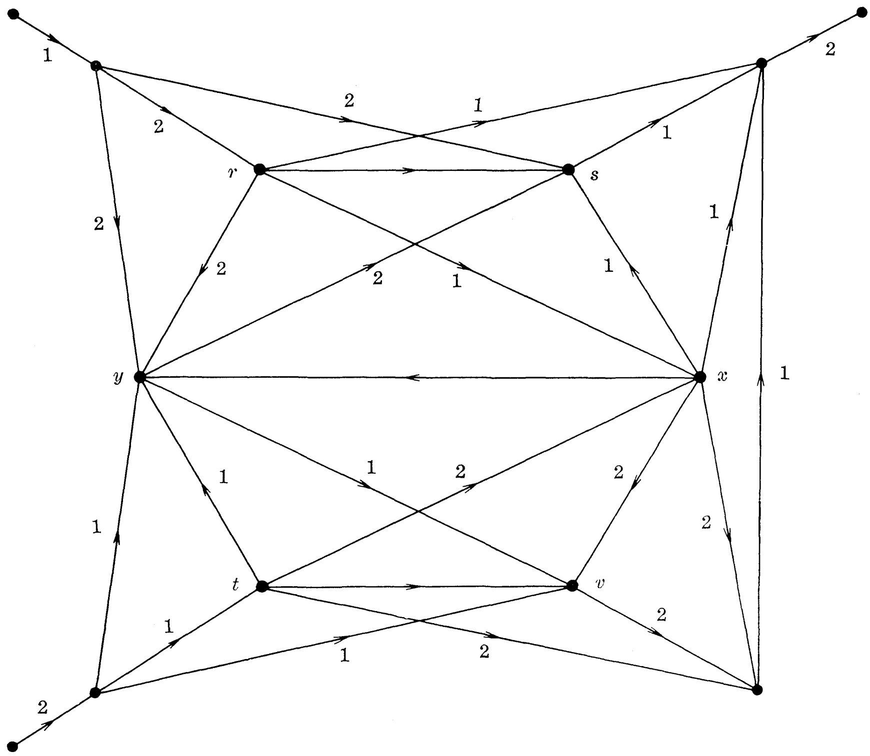

FIGURE 2

digraph of Figure 2 are in V except for xy, rs and tv, and A1 and A2 hold. Labels 1 and 2 for P_1 and P_2 are assigned to the edges in V in the only way consistent with A1 and A2, beginning with P_1 in the upper left corner. For $\sigma(R) = 2$ we require rs and tv in both P_1 and P_2, but xy violates A3 and cannot be assigned either

$$P_1\,[rx \in P_1 \;\&\; ry \notin P_1] \text{ or } P_2\,[tx \in P_2 \;\&\; ty \notin P_2]\;.$$

By deleting the edge xy from Figure 2 we obtain an R with $\sigma(R) = 2$ where R is not the union of two *disjoint* partial orders.

Sufficiency Proof for Theorem 6. With $V \neq \varnothing$ let A1, A2 and A3 hold. For $i = 1, 2$ let

$$S_i = \{xy\colon xy \in R - V \;\&\; \text{(i) holds}\}\;.$$

Let $R^0 = R - V - S_1 - S_2$ and for $i = 1, 2$ define P_i by

$$P_i = V_i \cup S_i \cup R^0\;.$$

Since $P_i \subseteqq R$, it is irreflexive. We now prove that P_1 is transitive. The proof for P_2 is similar.

Assume that $xy, yz \in P_1$. Then $xz \in R$, for if both xy and yz are in V_1 then $xz \in R$ by A1, and if one of xy and yz is in $S_1 \cup R^0$ then $xz \in R$ by the definitions. Thus $xz \in P_1$ unless $xz \in V_2 \cup S_2$. $xz \in V_2$ is contradicted in all cases:

1. $xy, yz \in V_1 \Rightarrow xz \notin V_2$, by A2;
2. $xy \in V_1$ & $yz \in S_1 \Rightarrow xz \notin V_2$, by A3;
3. $xy \in V_1$ & $yz \in R^0 \Rightarrow xz \notin V_2$, by A3;
4. xy, $yz \in S_1 \cup R^0$. Then $ax \in R \Rightarrow ay \in R \Rightarrow az \in R$ and $za \in R \Rightarrow ya \in R \Rightarrow xa \in R$. Hence neither $axKxz$ nor $xzKza$ can hold.

It remains to show that $xz \notin S_2$. Assume $xz \in S_2$ to the contrary and for definiteness take $zw \in V_1$ and $xw \in V_2$ (Figure 3). We note first

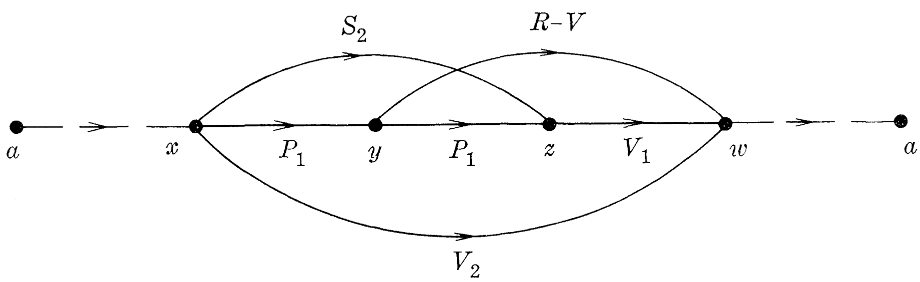

FIGURE 3

that $yw \notin V_2$, for $yw \in V_2 \Rightarrow yz \in S_2$. Moreover, $yw \notin V_1$, for $yw \in V_1$ & $xy \in V_1$ contradict A2, and $yw \in V_1$ & $xy \in S_1 \cup R^0$ contradict the definition of S_2 along with A3. Hence $yw \in R - V$. Now if $ax \in V_1$ then $ay \in R$ and hence (since $yw \in R - V$) $aw \in R$; and if $wa \in V_1$ then $za \in R$ and hence (since $xz \in R - V$) $xa \in R$. Since $xw \in V_2$ requires either $axKxw$ with $ax \in V_1$ or $xwKwa$ with $wa \in V_1$, and since $ax \in V_1$ contradicts $axKxw$ (since $aw \in R$) and $wa \in V_1$ contradicts $xwKwa$ (since $xa \in R$), the proof is complete.

REFERENCES

1. K. A. Baker, P. C. Fishburn, and F. S. Roberts, *Partial Orders of Dimension* 2, *Interval Orders, and Interval Graphs*, RAND Paper P-4376, The RAND Corporation, 1970.
2. B. Dushnik and E. W. Miller, *Partially Ordered Sets*, Amer. J. Math., **63** (1941), 600-610.
3. P. Erdös and L. Moser, *On the Representation of Directed Graphs as Unions of Orderings*, Math. Inst. Hung. Acad. Sci., **9** (1964), 125-132.
4. J. Folkman, *Graphs with Monochromatic Complete Subgraphs in Every Edge Coloring*, SIAM J. Appl. Math., **18** (1970), 19-24.
5. A. Ghouila-Houri, *Caractérisation des graphes nonorientés dont on peut orienter les arêtes de manière à obtenir le graphe d'une relation d'ordre*, C. R. Acad. Sci. Paris **254** (1962), 1370-1371.

6. P. C. Gilmore and A. J. Hoffman, *A Characterization of Comparability Graphs and of Interval Graphs*, Canad. J. Math., **16** (1964), 539-548.
7. F. Harary, R. Z. Norman, and D. Cartwright, *Structural Models*: *An Introduction to the Theory of Directed Graphs*, Wiley, New York, 1965.
8. D. Lubell, *A Short Proof of Sperner's Lemma*, J. Combinatorial Theory, **1** (1966), 299.
9. J. W. Moon, *Topics on Tournaments*, Holt, New York, 1969.
10. K. B. Reid and E. T. Parker, *Disproof of a conjecture of Erdös and Moser on tournaments*, J. Combinatorial Theory, **9** (1970), 225-238..
11. J. Spencer, *Notes on Combinatorial Mathematics*: *Transitive Subtournaments*, RAND Memorandum RM-6156-PR, The RAND Corporation, 1969.
12. ———, *Minimal Completely Separating Systems*, J. Combinatorial Theory, **8** (1970), 446-447.
13. E. Szpilrajn, *Sur extension de l'ordre partiel*. Fundamenta Math., **16** (1930), 386-389.

Received July 22, 1970 and in revised form October 12, 1970.

THE INSTITUTE FOR ADVANCED STUDY AND THE RESEARCH ANALYSIS CORPORATION
AND
THE RAND CORPORATION, SANTA MONICA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

ZERO DIVISORS IN DIFFERENTIAL RINGS

HOWARD E. GORMAN

Let R be a commutative ordinary differential ring with 1. Let A be a commutative differential R-algebra satisfying the ascending chain condition on radical differential ideals. Let M be a differentially finitely generated R-module. We obtain the following results on the zero divisors of A and M in R. (i) If R satisfies the ascending chain condition on radical differential ideals and if A has zero nilradical, then the assassinator of A in R is finite and consists of differential ideals; it is contained in the support of A in R, and the minimal members of each set comprise exactly the minimal prime ideals which contain the annihilator of A in R; (ii) If $R \subseteqq A$ and I is a radical differential ideal of A, then we obtain the assassinator of A/I in R from the assassinator of A/I in A by intersecting with R; (iii) If R is noetherian, then the set of zero divisors of M in R is a unique union of prime differential ideals of R, each of which is maximal among annihilators in R of nonzero elements of M; (iv) If I is the annihilator or power annihilator of M in R, then any prime ideal of R minimal over I is the annihilator of a nonzero element of M. In the above, (iii) and (iv) require an additional hypothesis to be made explicit later.

These results (except (ii)) are well known for finite modules over noetherian rings.

2. Preliminaries. In what follows, all rings are commutative and all modules are unitary. R will always be a differential ring with 1, with fixed derivation denoted by "$\,'\,$". By a *differential module* M over R, one means an R-module M together with an additive map from M to M, again denoted by "$\,'\,$", which satisfies $(rm)' = r'm + rm'$ for each $r \in R$ and $m \in M$. If $x \in M$, the successive derivatives of x will be denoted by $x', x'', \cdots, x^{(n)}, \cdots$. By a *differential algebra* A over R, one means a differential module A which is a ring and for which the module derivation is a ring derivation. By an ideal of A, we always mean an algebra ideal.

Let M be any R-module and $T \subseteqq M$ a subset. We denote the zero divisors of T in R by $\mathscr{Z}_R(T)$ and the annihilator of T in R by $\mathscr{A}_R(T)$. The *assassinator* of M in R, written $\operatorname{Ass}_R M$, is the set of prime ideals of R which are the annihilators of nonzero elements of M. The *support* of M in R, written $\operatorname{Supp}_R M$, is the set of prime ideals P of R such that $M_P \neq 0$.

Now let R be a differential ring and M a differential R-module.

Denote by $[T]/R$ the smallest differential submodule of M containing T. We call M *d-finitely generated* if there exists $n \geqq 0$ and $x_1, \cdots, x_n$ in M such that $M = [x_1, \cdots, x_n]/R$.

Let $S \subseteqq R$ be a multiplicatively closed set with $0 \notin S$. Then the derivations on R and M extend by the usual quotient formula to make M_s into a differential R_s-module. (See [2; Lemma 1].)

Assume, in addition, that M is a differential R-algebra. Denote by $\{T\}/M$ the smallest radical differential ideal containing T. The following fact is a trivial consequence of [5; Lemma 1.3]. Let Rad $M = 0$ (i.e., M has zero nilradical), and let T be a subset of either R or M. Then $\mathscr{A}_R(T)$ and $\mathscr{A}_M(T)$ are radical differential ideals.

3. The assassinator.

We begin by stating the first main theorem.

THEOREM 1. *Let R be a differential ring and A a differential R-algebra. Let R and A satisfy the ascending chain condition on radical differential ideals, and let* Rad $A = 0$. *Then* $\operatorname{Ass}_R A$ *is finite, consists of differential prime ideals, and is contained in* $\operatorname{Supp}_R A$. *The minimal members of each of these sets are the same and coincide with the prime ideals of R minimal over $\mathscr{A}_R(A)$.*

Before proving Theorem 1, we need a series of lemmas.

LEMMA 1. *Let R be a differential ring satisfying the ascending chain condition on radical differential ideals. Let A be a nonzero differential R-algebra with* Rad $A = 0$. *Then* $\operatorname{Ass}_R A \neq \varnothing$.

Proof. For any nonzero $a \in A$, $\mathscr{A}_R(a)$ is a proper, radical differential ideal of R. By hypothesis, there are ideals of R maximal among annihilators of nonzero elements of A. That these ideals are prime is well known [4; Theorem 6].

LEMMA 2. *Let R be a differential ring, and let M be a differential R-module. Let T be a subset of M, and suppose that $\mathscr{A}_R(T)$ is a differential ideal. Then:*

(i) $\mathscr{A}_R(T) = \mathscr{A}_R([T]/R)$;

(ii) *if M is a differential R-algebra, then* $\mathscr{A}_R(T) = \mathscr{A}_R([T]/M)$; *if, in addition,* Rad $M = 0$, *then* $\mathscr{A}_R(T) = \mathscr{A}_R(\{T\}/M)$.

Proof. Let $y \in \mathscr{A}_R(T)$. Then, since $xy' + x'y = 0$ for any $x \in T$, and $y' \in \mathscr{A}_R(T)$, we see that $x'y = 0$. Hence, $x''y + x'y' = 0$. The above argument applied to y' instead of to y would have resulted in $x'y' = 0$. Hence $x''y = 0$. Continuing in this way, we see that

$x^{(k)}y = 0$ for each nonnegative integer k. Now since an arbitrary element of $[T]/R$ has the form $\sum_{i,j} a_{i,j}x_j{}^{(i)}$, and an arbitrary element of $[x]/M$ (when M is an R-algebra) has the form $\sum_{i,j} b_{ij}x_i^{(i)} + \sum_{i,j} c_{ij}x_j{}^{(i)}$, for $b_{ij} \in M$ and a_{ij}, $c_{ij} \in R$, for every i and j, we see that $\mathscr{A}_R(T) \subseteqq \mathscr{A}_R([T]/R)$ and $\mathscr{A}_R(T) \subseteqq \mathscr{A}_R([T]/M)$. Since the opposite inclusions are clear, we have equality.

Now assume that M is an R-algebra and that Rad $M = 0$. By the above, we will be through once we show that $\mathscr{A}_R([T]/M) \subseteqq \mathscr{A}_R(\{T\}/M)$. Now $\mathscr{A}_M(\mathscr{A}_R([T]/M))$ is a radical differential ideal of M. Since it contains T, it contains $\{T\}/M$; i.e., $\{T\}/M$ annihilates $\mathscr{A}_R([T]/M)$; therefore, $\mathscr{A}_R([T]/M)$ annihilates $\{T\}/M$. This completes the proof.

LEMMA 3. *Let R be a differential ring, and let A be a differential R-algebra satisfying the ascending chain condition on radical differential ideals, and such that* Rad $A = 0$. *Let P be a prime ideal of R containing $\mathscr{A}_R(A)$. Then $P \in \mathrm{Supp}_R$* A.

Proof. Since A satisfies the ascending chain condition on radical differential ideals, there must be $a_1, \cdots, a_r$ in A such that $A = \{a_1, \cdots, a_r\}/A$. Suppose that $A_P = 0$. Then there are $s_i \in R - P$ such that $s_i a_i = 0$ for each i. Let $s = \prod_{i=1}^r s_i$. Then $sa_i = 0$ for each i. Since Rad $A = 0$, $\mathscr{Z}_A(s)$ is a radical differential ideal of A containing each a_i, and so must equal A. But then $sA = 0$; i.e., $s \in \mathscr{A}_R(A)$, which contradicts $s \notin P$. This completes the proof.

LEMMA 4. *Let R be a differential ring satisfying the ascending chain condition on radical differential ideals, and let A be a differential R-algebra with* Rad $A = 0$. *Then* $\mathrm{Ass}_R A \subseteqq \mathrm{Supp}_R A$, *and each member of* $\mathrm{Supp}_R A$ *contains a member of* $\mathrm{Ass}_R A$. *In particular, both sets have the same minimal elements.*

Proof. That $\mathrm{Ass}_R A \subseteqq \mathrm{Supp}_R A$ is just [1; § 1, °3, Prop. 7(i)]. Now let $Q \in \mathrm{Supp}_R A$. Then $A_Q \neq 0$ as an R_Q-algebra. By Lemma 1, $\mathrm{Ass}_{R_Q}(A_Q) \neq \varnothing$. Let $P_1 \in \mathrm{Ass}_{R_Q}(A_Q)$ with $P_1 = \mathscr{Z}_{R_Q}(a/1)$. Since Rad $(A_Q) = 0$, P_1 is a differential ideal. Let $P = \{r \in 1 \mid r/1 \in P_1\}$. Then P is a prime differential ideal of R and $P \subseteqq Q$. We claim that $P \in \mathrm{Ass}_R A$. By hypothesis, $P = \{p_1, \cdots, p_n\}/R$ for some $p_1 \cdots, p_n \in R$. Since $p_i a/1 = 0$, there are $s_i \in R - P$ such that $p_i s_i a = 0$ for each i. Hence, if $s = \prod_{i=1}^n s_i$, $p_i \in \mathscr{Z}_R(sa)$ for each i. Since $\mathscr{Z}_R(sa)$ is a radical differential ideal of R, $P \subseteqq \mathscr{Z}_R(sa)$. On the other hand, if $x \in \mathscr{Z}_R(sa)$, then $xa/1 = 0$; i.e., $x/1 \in P_1$; i.e., $x \in P$. Hence $P = \mathscr{Z}_R(sa) \in \mathrm{Ass}_R A$, and we are done.

LEMMA 5. *Let R be a differential ring, and let A be a differential R-algebra satisfying the ascending chain condition on radical differential ideals. Assume that* Rad $A = 0$. *Then A has a normal series*

$$A = A_0 \supseteqq A_1 \supseteqq \cdots \supseteqq A_n = 0$$

where

(i) *A_i is a radical differential ideal of A for each i;*

(ii) *$\mathscr{A}_R(A_{i-1}/A_i) = \mathrm{Ass}_R\,(A_{i-1}/A_i)$ for each i, and both consist of a single prime differential ideal P_i of R.*

Proof. Let $B \neq A$ be a radical differential ideal of A. Then A/B is a differential R-algebra satisfying the ascending chain condition on radical differential ideals, and $\mathrm{Rad}\,(A/B) = 0$. Since $A/B \neq 0$, we are guaranteed by Lemma 1 that there exists in $\mathrm{Ass}_R\,(A/B)$ a differential prime ideal $P = \mathscr{Z}_R(x)(x \in A/B$ and nonzero) which is maximal among the annihilator of nonzero elements of A/B. Let $B_1 = \varphi^{-1}(\{x\}/(A/B))$ where φ is the canonical homomorphism of A onto A/B. Then B_1 is a radical differential ideal of A, $B \subsetneqq B_1$ and $B_1/B \cong \{x\}/(A/B)$ so that $\mathscr{A}_R(B_1/B) = P$ by Lemma 2. Now suppose that $Q \in \mathrm{Ass}_R(B_1/B)$. Then $Q = \mathscr{Z}_R(b_1)$ for some $b_1 \in B_1/B$. Since $Pb_1 = 0$, $P \subseteqq Q$; hence, by the maximality of P, $P = Q$ and $\mathrm{Ass}_R\,(B_1/B)$ consists of the single prime P.

Starting with $B = 0$ and using the above method, we construct an increasing chain of radical differential ideals of A satisfying the conclusions of the lemma. By hypothesis, this chain must stop; i.e., at some stage, $B_1 =$ A, and we are done.

Proof of Theorem 1. We follow the notation of Lemma 5. By [1; § 1, °1, Prop. 3],

$$\mathrm{Ass}_R\, A \subseteqq \cup_{i=1}^{n} \mathrm{Ass}_R\,(A_{i-1}/A_i) = \{P_1, \cdots, P_n\}$$

so that $\mathrm{Ass}_R\, A$ is finite and consists of differential ideals. By Lemma 4, $\mathrm{Ass}_R\, A \subseteqq \mathrm{Supp}_R\, A$, and each has the same minimal elements. (In fact, since $P_i \in \mathrm{Supp}_R(A_{i-1}/A_i)$ by Lemma 4 and since $0 \neq (A_{i-1}/A_i)_{P_i} = (A_{i-1})_{P_i}/(A_i)_{P_i}$ each $P_i \in \mathrm{Supp}_R\, A_{i-1} \subseteqq \mathrm{Supp}_R\, A$.) That these minimal elements coincide with the prime ideals of R minimal over $\mathscr{A}_R(A)$ follows from the following two facts: The minimal elements of $\mathrm{Ass}_R\, A$, and so of $\mathrm{Supp}_R\, A$, contain $\mathscr{A}_R(A)$; the primes minimal over $\mathscr{A}_R(A)$ are members of $\mathrm{Supp}_R\, A$ by Lemma 3. This completes the proof.

COROLLARY. *Let the hypotheses be as in Theorem* 1. *Then* $\mathrm{Supp}_R\, A$ *consists of exactly the prime ideals of R which contain $\mathscr{A}_R(A)$.*

We remark that if R contains the rational numbers and satisfies the ascending chain condition on radical differential ideals, then any quotient by a differential ideal of the differential polynomial ring over R in a finite number of differential indeterminates also satisfies the ascending chain condition on radical differential ideals.

If we assume that $R \subseteqq A$, we get the following result with no chain condition assumptions on R.

THEOREM 2. *Let R be a differential ring contained in the differential R-algebra A. Assume that A satisfies the ascending chain condition on radical differential ideals. Let I be a radical differential ideal of A. Then:* (i) *I can be written uniquely as $I = \bigcap_{i=1}^{n} P_i$ where the P_i are prime differential ideals of A;* (ii) *if $Q_i = P_i \cap R$, then*

$$\operatorname{Ass}_A(A/I) = \{P_1, \cdots, P_n\} \text{ and } \operatorname{Ass}_R(A/I) = \{Q_1, \cdots, Q_n\}.$$

Proof. We note that (i) is well known and proved more directly in [5; Theorem 7.5]. Now A/I, viewed as an A-algebra, satisfies the hypotheses of Lemma 1 and Theorem 1. Let $P_1, \cdots, P_n$ be the unique elements of $\operatorname{Ass}_A(A/I)$ minimal over $\mathscr{A}_A(A/I)$. Since $1 \in A$, $\mathscr{A}_A(A/I) = I$, and since I is a radical ideal, $I = \bigcap_{i=1}^{n} P_i$. This proves (i).

Since the P_i are minimal over $\mathscr{A}_A(A/I)$, they are minimal members of $\operatorname{Ass}_A(A/I)$ by Theorem 1. On the other hand, let $P = \mathscr{Z}_A(a_1) \in \operatorname{Ass}_A(A/I)$, with $a_1 \in A/I$. Let $a \in A$ be mapped to a_1. Then $a \notin P_j$ for some $j = 1, \cdots, n$. But $Pa \subseteqq I \subseteqq P_j$, so that $P \subseteqq P_j$; i.e., $P = P_j$. Hence $\operatorname{Ass}_A(A/I) = \{P_1, \cdots, P_n\}$.

Now let $P_i = \mathscr{Z}_A(a_i)$, $a_i \in A/I$ for each i. Then $Q_i = P_i \cap R$ must be $\mathscr{Z}_R(a_i)$ for each i; i.e., $Q_i \in \operatorname{Ass}_R(A/I)$.

To complete the proof, we must show that any $Q \in \operatorname{Ass}_R(A/I)$ is one of the Q_i. Localize A and R at Q. Then A_Q is an R_Q-algebra satisfying the hypotheses of the theorem and I_Q is a radical differential ideal of A_Q. Further, I_Q is a proper ideal of A_Q for, since $I \cap R \subseteqq Q$, we see that $(I \cap R)_Q = I_Q \cap R_Q \subseteqq Q_Q$; i.e., $R_Q \not\subseteq I_Q$. Since each P_i is prime, $I_Q = (\bigcap_{i=1}^{n} P_i)_Q = \bigcap_{i=1}^{r} (P_i)_Q$ where we have assumed that $P_1, \cdots, P_r$ are exactly those among $P_1, \cdots, P_n$ such that $(P_i)_Q \neq A_Q$. Note that $r > 0$ by Lemma 1 since $A_Q/I_Q \neq 0$. By the initial argument in this part of the theorem, $\operatorname{Ass}_{R_Q}(A_Q/I_Q) = \{(P_1)_Q, \cdots, (P_r)_Q\}$. Since $Q_Q \subseteqq \mathscr{Z}_{R_Q}(A_Q/I_Q)$, $Q_Q \subseteqq (P_i)_Q \cap R_Q = (Q_i)_Q$ for some i. Since Q_Q is maximal, $Q_Q = (Q_i)_Q$; i.e., $Q = Q_i$, and the proof is complete.

4. **The case for modules.** The situation for modules is less complete. However, under the restriction given below, we can gain some information about $\operatorname{Ass}_R(M)$ when M is a d-finitely generated R-module.

We say that the differential R-module M satisfies the property (♯) *if ideals of R maximal among the annihilators of nonzero elements of M are differential ideals. We say that M satisfies the property* (♯♯) *if* M/N *satisfies the property* (♯) *for every differential submodule N of M.*

THEOREM 3. *Let R be a noetherian differential ring and M a nonzero, d-finitely generated R-module which satisfies the property* (♯♯). *Then* $\mathrm{Ass}_R M$ *is finite.*

Proof. The assassinator of nonzero modules over noetherian rings is never empty. Using the condition (♯♯) and Lemma 2(i), we modify the proof of Lemma 5 to prove an analogue of Lemma 5 in which the A_i are replaced by differential R-modules. The result now follows as in the first part of Theorem 1.

Further progress in this direction is limited by the fact that prime ideals of R containing $\mathscr{A}_R(M)$ need not be in $\mathrm{Supp}_R M$. The correct modification is given in Lemma 7. (For example, let $R = Z$, the integers, with the trivial derivation. Let M be generated over $Z/2Z$ by 1 and the set $\{x/2^n\}$ for $n = 0, 1, 2, \cdots$, and have derivation defined by $(x/2^n)' = x/2^{n+1}$. Then $M = [1, x]/Z$. Now $\mathscr{A}_Z(M) = 0$; but if $P = 3Z$, $M_P = 0$.)

The following discussion indicates what is still true if we assume only the condition (♯). We shall need the result [2; Th. 1]:[1]

THEOREM A. *Let R be a noetherian differential ring, and let M be a d-finitely generated R-module. Then M satisfies the ascending chain condition on differential submodules.*

We can now prove

THEOREM 4. *Let R be a noetherian differential ring and M a d-finitely generated R-module which satisfies the property* (♯). *Then* $\mathscr{Z}_R(M)$ *is expressible uniquely as the union of a finite number of differential prime ideals, each of which is maximal among the annihilators of nonzero elements of M.*

Proof. Each nonzero $x \in M$ has an annihilator ideal, and $\mathscr{Z}_R(M)$ is clearly their union. Each such annihilator is contained in a maximal one which is prime, and differential by assumption. Let $\{P_\lambda\}_{\lambda \in \Lambda}$ be the set of these maximal annihilators, and let $P_\lambda = \mathscr{Z}_R(x_\lambda)$, $x_\lambda \in M$

[1] This theorem, in different language, is originally due to J. Johnson, *Differential dimension polynomials and a fundamental theorem on differential modules*, Amer. J. Math., **91** (1969), 239.

for each $\lambda \in \Lambda$. The differential submodule N of M generated by the x_λ's is d-finitely generated by Theorem A. Let $N = [x_1, \cdots, x_n]/R$, with the $x_1, \cdots, x_n$ chosen from among the x_λ's. Then, for any λ, $x_\lambda = \sum_{i,j} r_{ij} x_i^{(j)}$, with $r_{ij} \in R$ for each i and j, and only a finite number of values for j appearing. Since, by Lemma 2, $P_i = \mathscr{A}_R([x_i])$ for each $i = 1, 2, \cdots, n$, this implies that $P_\lambda \supseteqq \cap_{i=1}^n P_i$. This implies, by maximality, that P_λ is one of the P_i's. Hence, $\mathscr{Z}_R(M) = \cup_{i=1}^n P_i$.

To show uniqueness, we remark that if Q were a member of another such union, then $Q \subseteqq \cup_{i=1}^n P_i$ implies that Q equals one of the P_i's [4; Th. 8]. This proves the theorem.

For any R-module M, define $\mathscr{P}\mathscr{A}_R(M)$, the *power annihilator* of M in R, to be the set of r in R such that for every $m \in M$, there is a positive integer n with $r^n m = 0$. Then $\mathscr{P}\mathscr{A}_R(M)$ is an ideal which contains both $\mathscr{A}_R(M)$ and its radical. (If M is finitely generated, it equals this radical.)

LEMMA 6. *Let M be a differential R-module. Let $a \in M$ and $r \in R$, and suppose that $ra = 0$. Then, for every nonnegative integer n, we have $r^{n+1} a^{(n)} = 0$.*

Proof. We proceed by induction, the case $n = 0$ being satisfied by hypothesis.

If $r^n a^{(n-1)} = 0$, then $r^n a^{(n)} + n r^{n-1} r' a^{(n-1)} = 0$.

On multiplying through by r, we have the result.

LEMMA 7. *Let R be a differential ring M a d-finitely generated R-module. Let $P \subseteqq R$ be a prime ideal containing $\mathscr{P}\mathscr{A}_R(M)$. Then $M_P \neq 0$.*

Proof. Let $M = [m_1, \cdots, m_r]/R$, and assume that $M_P = 0$. Then there is an $s \in R - P$ such that $sm_i = 0$ for each i. By Lemma 6, $s^k m_i^{(k-1)} = 0$ for each i and k, and so, for every $m \in M$, there is a positive integer t with $s^t m = 0$; i.e., $s \in \mathscr{P}\mathscr{A}_R(M)$. This contradicts $s \notin P$.

LEMMA 8. *Let M be any R-module. Let $I = \mathscr{A}_R(M)$ (resp., $I = \mathscr{P}\mathscr{A}_R(M)$), and let P be a prime ideal of R containing I. Assume that $M_P \neq 0$. Then $I_P \subseteqq \mathscr{A}_{R_P}(M_P) \subseteqq P_P$ (resp., $I_P \subseteqq \mathscr{P}\mathscr{A}_{R_P}(M_P) \subseteqq P_P$).*

Proof. The first inclusion is clear in both cases. We prove the second inclusion, $\mathscr{P}\mathscr{A}_{R_P}(M_P) \subseteqq P_P$. Let $x/t \in \mathscr{P}\mathscr{A}_{R_P}(M_P)$ with $x \in R$ and $t \in R - P$, and let $m \in M$ be such $m/1 \neq 0$. If $(x/t)^r m/1 = 0$, then there is an $s \in R - P$ with $sx^r m = 0$. If $x \notin P$, then $sx^r \in R - P$, so that $sx^n m = 0$ implies that $m/1 = 0$, a contradiction. Hence, $x \in P$, and we are done.

If M is any R-module, it is well known that any prime ideal of R minimal over $\mathscr{A}_R(M)$ is contained in $\mathscr{Z}_R(M)$ (See [4; Th. 84]). A minor variant of the proof in the reference proves.

LEMMA 9. *Let M be any R-module. Let P be a prime ideal minimal over $\mathscr{P}\mathscr{A}_R(M)$. Then $P \subseteqq \mathscr{Z}_R(M)$.*

THEOREM 5. *Let R be a noetherian differential ring and M a d-finitely generated R-module. Let $I = \mathscr{A}_R(M)$ (resp., $\mathscr{P}\mathscr{A}_R(M)$), and let P be a minimal prime ideal over I. Assume that $M_P \neq 0$ (note Lemma 7 in this regard) and that M_P satisfies the property (♯). Then P is a differential ideal and $P \in \operatorname{Ass}_R M$.*

Proof. M_P is a nonzero, d-finitely generated module over R_P. Since P is minimal over I, Lemma 8 implies that P_P is minimal over $\mathscr{A}_{R_P}(M_P)$ (resp., $\mathscr{P}\mathscr{A}_{R_P}(M_P)$). By Lemma 9 and the remark preceding it, $P_P \subseteqq \mathscr{Z}_{R_P}(M_P)$. It follows from Theorem 4 and the maximality of P_P that there is an $x \in M$ such that $P_P = \mathscr{Z}_{R_P}(x/1)$. Further, P_P is a differential R_P-ideal. Since $P = \{r \in R \mid r/1 \in P_P\}$, P is a differential ideal also. Since P_P is finitely generated, this implies the existence of an $s \in R - P$ such that $sPx = 0$. But then $P = \mathscr{Z}_R(sx)$. For if $ysx = 0$, for some $y \in R$, then $(y/1)(x/1) = 0$; i.e., $y/1 \in P_P$. It follows that $y \in P$, and we are done.

EXAMPLE. Let S be a noetherian ring containing the rational numbers and equipped with the trivial derivation. Let R be the ring of formal power series over S in the indeterminate z, equipped with the derivation defined by $z' = z$. Since every prime ideal of R is of the form PR or $PR + zR$, where P is a prime ideal of S, R satisfies the condition (♯♯) for any R-module. Let x be an indeterminate, and let $M_1 = R[x^{-1}]$, viewed as a differential R-module by the derivation $(x)' = r$ for some unit $r \in S$. Since $x^{-(n+1)} = (x^{-1})^{(n)}$ times a unit of S, M_1 is d-finitely generated over R by 1 and x^{-1}. Let M be any quotient module of M_1 by a differential submodule. Then M and R satisfy the hypotheses of Theorems 3, 4, and 5. Notice that if M_1 is considered as a ring, Rad M_1 need not be zero.

REFERENCES

1. N. Bourbaki, *Algebre Commutative*, Chap. 4, Hermann, Paris, 1961.
2. H. E. Gorman, *Differential rings and modules*, to appear in Scripta Math.
3. ———, *Radical regularity in differential rings*, Canadian J. Math., **XXIII** No. 2 (1971), 197-201.

4. I. Kaplansky, *Commutative Rings*, Allyn and Bacon, 1970.
5. ———, *An Introduction to Differential Algebra*, Hermann, Paris, 1957.
6. A. Seidenberg, *Differential ideals in rings of finitely generated type*, Amer. J. Math., **89** (1967), 22-42.

Received March 17, 1970 and in revised form April 29, 1971. The work on this paper was partially supported by NSF Grant 14023.

STANFORD UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

A NOTE ON THE LÖWNER DIFFERENTIAL EQUATIONS

MAURICE HEINS

Charles Löewner in Memoriam

The object of the present note is to indicate a derivation of the Löwner differential equations [1] based on the derivation of an associated differential equation for Green's function of the variable region relative to the defining parameter. Decisive in our treatment is the use of a certain normalized minimal positive harmonic function on the variable region. In fact, our starting point was the feeling that the Poisson kernel asserted its presence so strongly in the Löwner differential equations that the concomitant presence of a normalized minimal positive harmonic function on the variable region should appear naturally in the study of the question. We shall see that this is the case. A technical advantage of the present approach is that the "tip" lemmas of the classical proof are dispensed with.

It would be of interest to see whether the indicated method, which is available for other families of harmonic functions monotone justifying in a parameter, has useful applications to the theory of harmonic functions.

2. Let γ be a Jordan arc with parametric domain $[0, T]$ such that $0 < |\gamma(t)| < 1$ for $0 \leqq t < T$ and $|\gamma(T)| = 1$. Let A_t denote the complement of the set $\gamma(\{t \leqq s < T\})$ with respect to the open unit disk, $0 \leqq t \leqq T$. Let g_t denote Green's function for A_t with pole at 0. The continuous dependence of g_t on the parameter t is an elementary matter (minimal property of Green's function, the Phragmén-Lindelöf boundary maximum principle). We let $\alpha(t)$ denote $\lim_{z\to 0} [g_t(z) + \log|z|]$. We note that $\alpha: t \to \alpha(t)$ is an increasing continuous function which satisfies $\alpha(T) = 0$. We reparametrize γ, as in the original Löwner argument, by composing γ with

$$t \longrightarrow \operatorname{inv}\ \alpha[t + \alpha(0)],\ 0 \leqq t \leqq -\alpha(0),$$

so that for the new γ we have $T = -\alpha(0)$ and $\alpha(t) = \alpha(0) + t$. [The notation "inv" is used to denote the inverse of a univalent function.]

We let G be defined by

$$G(z, t) = g_t(z),\ z \in A_t,\ 0 \leqq t \leqq T.$$

Given a function F having as domain a subset of $C \times R$, we denote by $D_2F(a, b)$ the derivative of $t \to F(a, t)$ at b and by $D_1F(a, b)$ and $\bar{D}_1F(a, b)$ the complex differential coefficients of $z \to F(z, b)$ at a, the obvious conventions holding. Our first step is to establish the existence of D_2G and to obtain information about it. We remark that the logarithmic singularity of g_t at 0 is harmless. The difference quotient

$$\frac{G(z, t) - G(z, s)}{t - s}, \quad 0 \leqq s < t \leqq T, \tag{1}$$

defines a positive harmonic function on A_s which takes the value 1 at 0 and vanishes continuously at each point of the frontier of A_t.

To control the limiting behavior of (1) as $(s, t) \to (\sigma, \sigma)$, $0 \leqq \sigma \leqq T$, we make use of the boundary behavior of the Riemann mapping function for a simply-connected Jordan region and the following standard lemma of Harnack type.

LEMMA. *Let* $m(z) = (1 - |z|)\ (1 + |z|)^{-3}$ *and* $M(z) = (1 + |z|) \times (1 - |z|)^{-3}$. *Let a and b be points of the semi-circular disk* $\{\mathrm{Im}\, z > 0, |z| < 1\}$. *Let u be nonnegative and harmonic on this set and vanish continuously on the diameter. Then*

$$[M(b)]^{-1} \frac{u(b)}{\mathrm{Im}\, b} \leqq [m(a)]^{-1} \frac{u(a)}{\mathrm{Im}\, a} \cdot$$

[A proof of this lemma is readily given with the aid of Schwarzian reflexion and the Poisson integral for a circular disk.]

Suppose that $(s_n, t_n) \to (\sigma, \sigma)$, where $0 \leqq s_n < t_n \leqq 1$. Then some subsequence of the sequence of difference quotients (1), given by $s = s_n$ and $t = t_n$, converges, uniformly on compact subsets of A_σ, to a positive harmonic function on A_σ which takes the value 1 at 0. Using the boundary behavior of the Riemann mapping function when a Jordan boundary lies at hand and the stated lemma, we see that the limit function in question vanishes continuously at each point of the frontier of A_σ, the "tip" $\gamma(\sigma)$ excepted.

We introduce the normalized Riemann mapping function f_t, mapping the open unit disk onto A_t and satisfying $f_t(0) = 0$, $f_t'(0) > 0$. From the continuity of $t \to g_t$, we infer the continuity of $t \to \mathrm{inv} f_t$ and thence the continuity of $t \to f_t$. Of course, the term "continuity" is to be construed in the sense of uniform limits on compact subsets. We let $\kappa(t)$ denote the unique preimage of $\gamma(t)$ with respect to the continuous extension of f_t to the closed unit disk. If h is a positive harmonic function on A_t taking the value 1 at 0 and vanishing continuously at each point of the frontier of A_t, the "tip" $\gamma(t)$ excepted, then

$$h[f_t(z)] = \operatorname{Re}\left[\frac{\kappa(t)+z}{\kappa(t)-z}\right], \quad |z|<1. \tag{2}$$

It follows that there is at most one h having the stated property. Using (2) as a defining condition for h, we see that such h exist. We denote the unique h in question (which is a normalized minimal positive harmonic function on A_t) by h_t.

Combining the results of the preceding two paragraphs we conclude that the difference quotient (1) tends to h_σ as $(s, t) \to (\sigma, \sigma)$ and that, in fact, the uniformity of the limit process holds on compact subsets of A_σ. We let H be defined by

$$H(z, t) = h_t(z), \quad z \in A_t, \quad 0 \leqq t \leqq T.$$

We see that the following differential equation, which will serve as a basis for the derivation of the Löwner differential equations, holds:

$$D_2 G = H. \tag{3}$$

Continuity of $t \to h_t$ and $\kappa\colon t \to \kappa(t)$. A second application of the boundary behavior of the Riemann mapping function for Jordan regions and the lemma yields the continuity of $t \to h_t$, $0 \leqq t \leqq T$. It suffices to establish the fact that if $t_n \to \sigma$, then some subsequence of (h_{t_n}) tends to h_σ. Using the continuity of $t \to h_t$, the continuity of $t \to f_t$, and (2) we shall now conclude the continuity of κ. Indeed, if $t_n \to \sigma$ and $\kappa(t_n) \to \alpha$, we obtain, using (2), the equality

$$h_\sigma[f_\sigma(z)] = \operatorname{Re}\left(\frac{\alpha+z}{\alpha-z}\right), \quad |z|<1,$$

and hence $\alpha = \kappa(\sigma)$. The continuity of κ follows.

3. The Löwner differential equations. The equations bear on the functions F, Ψ, and Θ, which will now be introduced.

F. We define F by $F(z, t) = f_t(z)$, $|z| < 1$, $0 \leqq t \leqq T$. It is convenient to have available Φ defined by $\Phi(z, t) = \operatorname{inv} f_t(z)$, $z \in A_t$, $0 \leqq t \leqq T$. Its role is auxiliary. The function Φ is useful as a link between F and $G = -\log |\Phi|$.

Ψ. The function Ψ is defined by $\Psi(z, t) =$ $\operatorname{inv} f_t[f_0(z)]$, $|z| < 1$, $0 \leqq t \leqq T$. This is the first function studied by Löwner in his classical paper. There is an identity involving Φ and Ψ:

$$\Psi(z, t) = \Phi[f_0(z), t]. \tag{4}$$

Θ. The function Θ is specified by the requirement that, $z \to \Theta(z, t)$ is the inverse of $z \to \Psi(z, t)$, $|z| < 1, 0 \leqq t \leqq T$. From

$$z = \Psi[\Theta(z, t), t] = \text{inv } f_t \ \{f_0[\Theta(z, t)]\}$$

for (z, t) in the domain of Θ, we obtain for such (z, t) the identity

$$F(z, t) = f_t(z) = f_0[\Theta(z, t)]. \tag{5}$$

The equation (3) yields a corresponding equation for Φ. Indeed, let $\tilde{h}_t$ denote the analytic function with domain A_t satisfying $\tilde{h}_t(0) = 1$, $\operatorname{Re} \tilde{h}_t = h_t$, and let $\tilde{H}$ be defined by $\tilde{H}(z, t) = \tilde{h}_t(z)$, $z \in A_t$, $0 \leqq t \leqq T$. Clearly, the function $\tilde{H}$ is continuous on its domain. To derive an equation bearing on Φ, we introduce G_1 having the same domain as G which satisfies $G_1(z, t) = G(z, t) + \log |z|$, $z \neq 0$, $G_1(0, t) = \alpha(0) + t$, and thereupon $\tilde{G}_1$ with the same domain and satisfying the condition that $z \to \tilde{G}_1(z, t)$ is the analytic function with real part $z \to G_1(z, t)$ satisfying $\tilde{G}_1(0, t) = G_1(0, t)$. It is readily verified that

$$D_2 G_1 = H,$$

the limit process being uniform in the sense indicated above. It follows, in view of the normalization made on $\tilde{G}_1$, that

$$D_2 \tilde{G}_1 = \tilde{H}.$$

Using the relation

$$\Phi(z, t) = z \exp\, [-\tilde{G}_1(z, t)],$$

we are led to the equation

$$D_2 \Phi = -\, \tilde{H}\Phi. \tag{6}$$

From (4) and (6) we obtain

$$\begin{aligned} D_2\Psi(z, t) &= -\tilde{H}[f_0(z), t]\Psi(z, t) \\ &= -\ \tilde{h}_t[f_0(z)]\Psi(z, t) \\ &= -\ [(\tilde{h}_t \circ f_t) \circ (\text{inv } f_t \circ f_0)(z)]\Psi(z, t), \end{aligned}$$

and, consequently, the Löwner equation

$$D_2\Psi(z, t) = -\frac{\kappa(t) + \Psi(z, t)}{\kappa(t) - \Psi(z, t)}\, \Psi(z, t), \tag{7}$$

(z, t) in the domain of Ψ.

The equation for F. From (6) and the continuity of $\tilde{H}, \Phi, D_1\Phi$, $\tilde{D}_1\Phi$ (trivially, since it vanishes), we conclude that Φ has the C' property and so is differentiable. Since F is continuous on its domain, $D_1\Phi$ is nowhere zero, and the identity,

$$\Phi[F(z, t), t] = z,$$

$|z| < 1$, $0 \leqq t \leqq T$, prevails, it is a standard matter of the differential calculus to conclude the existence of D_2F and the identity,

$$D_1\Phi[F(z, t), t]D_2F(z, t) + D_2\Phi[F(z, t), t] = 0, \quad |z| < 1, 0 \leqq t \leqq T. \tag{8}$$

It is elementary that

$$D_1\Phi[F(z, t), t]D_1F(z, t) = 1, \tag{9}$$

for the same (z, t). From (8) and (6) we obtain

$$D_1\Phi[F(z, t), t]D_2F(z, t) = z\,\frac{\kappa(t) + z}{\kappa(t) - z},$$

and thereupon using (9) the equation

$$D_2F(z, t) = D_1F(z, t)z\,\frac{\kappa(t) + z}{\kappa(t) - z}, \tag{10}$$

$|z| < 1$, $0 \leqq t \leqq T$.

That Θ satisfies the equation (10), Θ replacing F, on its domain, is immediate from (10), and the identities obtained from (5) by differentiation. The C' property of F follows from the continuity of $D_1\Phi$, $D_2\Phi$ and F on their respective domains and the identities (8) and (9) as well as the non-vanishing of $D_1\Phi$. The C' property Θ is now concluded with the aid of (5).

Reference

1. K. Löwner, *Untersuchungen über schlichte konforme Abbildungen des Einheitskreises I*, Math. Ann., **89** (1923), 103-121.

Received December 7, 1970.

University of Illinois

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

SEMI-ORTHOGONALITY IN RICKART RINGS

LOUIS HERMAN

This note initiates a study of the semi-orthogonality relation on the lattice of principal left ideals generated by idempotents of a Rickart ring. It will be seen that two left ideals in a von Neumann algebra are semi-orthogonal if and only if their unique generating projections are non-asymptotic. Connections between semi-orthogonality, dual modularity, von Neumann regularity, and algebraic equivalence will be established; those Rickart rings with a superabundance of semi-orthogonal left ideals will be characterized.

A *regular ring* is a ring A with identity in which each element $a \in A$ is *regular* in the sense that $aba = a$ for some element $b \in A$. A *Rickart ring* is a ring A with identity in which the left (and right) annihilator of each element is a principal left (right) ideal generated by an idempotent. Regular rings and Baer rings, as defined by Kaplansky [4], are special cases of Rickart rings: in particular, then, a von Neumann algebra is a Rickart ring. Rickart rings are called Baer rings in [2]. Throughout this note, A will denote a Rickart ring. $L(M)$ and $R(M)$ will denote respectively the left and right annihilators of a subset M of A. The letters e, f, g, h and k will denote idempotents and the letters E, F, G, H and K will denote the left ideals they generate.

Ordered by set inclusion, the set $L(A)$ of principal left ideals generated by idempotents forms a lattice. If E and F form a modular pair in $L(A)$, we shall write $(E, F)M$; if E and F form a dual modular pair in $L(A)$, we shall write $(E, F)M^*$. Following S. Maeda [6], we shall say that two left ideals E and F in $L(A)$ are *semi-orthogonal*, $E \sharp F$, if they are generated by orthogonal idempotents. Maeda shows that the semi-orthogonality relation $\sharp$ on $L(A)$ has these properties: (1) If $E \sharp E$, then $E = (0)$; (2) If $E \sharp F$, then $F \sharp E$; (3) If $E_1 \leqq E$ and $E \sharp F$, then $E_1 \sharp F$; (4) If $E \sharp F$ and $E \vee F \sharp G$, then $E \sharp F \vee G$; (5) If $E \leqq F$, then there is a left ideal G in $L(A)$ such that $E \vee G = F$ and $E \sharp G$.

The results herein form a portion of the author's dissertation, submitted to the Graduate School of the University of Massachusetts and directed by Professor D. J. Foulis.

2. Semi-orthogonal left ideals.

In this section, we give geometric meaning to Maeda's canonical semi-orthogonality relation in $L(A)$.

THEOREM 1. *Let $E = Ae$ and $F = Af$. Then the following conditions are equivalent*:

(1) $E \sharp F$.

(2) $E \cap F = (0)$ *and* $e(1 - f)$ *is regular in* A.

(3) $E \oplus F = E \vee F$ *in* $L(A)$.

Proof. The proofs of (1) implies (2) and of (3) implies (1) are routine. To see that (2) implies (3), we suppose that $e(1 - f)xe(1 - f) = e(1 - f)$ for some $x \in A$. Put $g = (1 - f)xe(1 - f)$. Then $fg = 0 = gf$ and $eg = e(1 - f)xe(1 - f) = e(1 - f) = e - ef$. Then $g^2 = (1 - f)xe(1 - f)g = (1 - f)xeg = (1 - f)xe(1 - f) = g$ and $(f + g)^2 = f + fg + gf + g = f + g$.

We claim that $E \oplus F = A(f + g)$. But $f = (f + g) - g(f + g) \in A(f + g)$ and $e = ef + eg = e(f + g) \in A(f + g)$. Thus $E \oplus F \leqq A(f + g)$. Conversely, $f + g = f + (1 - f)xe(1 - f) = (1 - f)xe + (1 - xe + fxe)f \in E \oplus F$. Hence $E \oplus F = A(f + g) \in L(A)$.

We can find perspicacious geometric and topological interpretations for each of these equivalent conditions in the ring of bounded operators on a Hilbert space or, more generally, in any von Neumann algebra. In such a ring, any left annihilator is a principal left ideal generated by a unique projection (= self-adjoint idempotent). Let e and f denote the unique generating projections of E and F respectively: we shall identify these projections with their ranges.

If $e \wedge f = 0$, e and f are said to be *asymptotic* if $\sup |\langle \alpha, \beta \rangle| = 1$, where $\|\alpha\| = 1 = \|\beta\|$, $\alpha \in e$, $\beta \in f$; otherwise e and f are said to be *non-asymptotic*. It is known [5, p. 166 and pp. 172–174] that these conditions are equivalent: (1) e and f form a non-asymptotic pair; (2) The projection map of the subspace $e \oplus f$ onto e is continuous; (3) The vector sum of e and f is a closed subspace; (4) $(e, f)M^*$ in the projection lattice of the ring of all bounded operators on the underlying Hilbert space. The relation of semi-orthogonality to non-asymptoticity is provocative; for, by modifying results of Jacob Feldman [1, pp. 12–14], it is easy to verify that $E \sharp F$ if and only if e and f form a non-asymptotic pair.

Our next result, though appearing an immediate consequence of Theorem 1 (2), seems to require a measure of prestidigitatorial skill with idempotents.

COROLLARY 1. *ef is regular if and only if $(1 - f)(1 - e)$ is regular.*

Proof. We prefer to demonstrate the obviously equivalent statement: If $e(1 - f)$ is regular, then so is $f(1 - e)$. To this end, choose an idempotent h with $Ah = Ae \cap Af$. Put $e_1 = e + h - eh$ and $f_1 = f + h - fh$. Then e_1 and f_1 are idempotent generators for Ae and Af respectively and $h = he_1 = e_1h = hf_1 = f_1h$. By direct computation, we have $e_1(1 - f_1) = e(1 - f)(1 - h)$ and $f_1(1 - e_1) = f(1 - e)(1 - h)$. Since $e(1 - f)$ is regular, $e(1 - f)xe(1 - f) = e(1 - f)$ for some $x \in A$. Then, an easy computation shows $e_1(1 - f_1)[(1 - f)x]e_1(1 - f) = e_1(1 - f_1)$; thus $e_1(1 - f_1)$ is regular.

Put $e_0 = e_1(1 - h)$ and $f_0 = f_1(1 - h)$. Then $e_0(1 - f_0) = e_1(1 - f_1)$ is regular. Moreover, if $z \in Ae_0 \cap Af_0 \leqq Ae_1 \cap Af_1 = Ah$, then $z = zh$ $(ze_0)h = ze_1(1-h)h = 0$; so $Ae_0 \cap Af_0 = (0)$. Then by Theorem 1 (2), we have $Ae_0 \sharp Af_0$.

Consequently, $f(1 - e)(1 - h) = f_1(1 - e_1) = f_0(1 - e_0)$ is regular. Then $f(1 - e)(1 - h)yf(1 - e)(1 - h) = f(1 - e)(1 - h)$ for some element $y \in A$. But this means that $f(1 - e)(1 - h)yf(1 - e) - f(1 - e) = f(1 - e)(1 - h)yf(1 - e)h - f(1 - e)h$ is an element of $A(1 - e) \cap Ah = A(1 - e) \cap Ae \cap Af = (0)$. Thus $f(1 - e)[(1 - h)y]f(1 - e) = f(1 - e)(1 - h)yf(1 - e) = f(1 - e)$, showing that $f(1 - e)$ is regular in A.

COROLLARY 2. *If $E \sharp F$, then $(E, F)M$ and $(E, F)M^*$ in $L(A)$.*

Proof. A proof that E and F form a modular pair is given by Maeda [6, Lm. 1]. Now suppose that $Ae \sharp Af$ with $Af \leqq Ag \leqq Ae \oplus Af$. Then $g = xe + yf$ for some elements x and y in A. Then $xe = g - yf \in Ae \cap Ag$ and we have $g = xe + yf \in (Ae \cap Ag) \oplus Af$. Thus $Ag \leqq (Ae \cap Ag) \oplus Af$. Since the opposite inclusion is evident, $Ag = (Ae \cap Ag) \oplus Af$. Hence $(Ae, Af)M^*$.

3. **Equivalence of left ideals.** Two left ideals E and F in $L(A)$ are *semi-orthogonally perspective* via G, $G: E \sim F$, if $E \oplus G = E \vee F = G \oplus F$ with $E \sharp G$ and $G \sharp F$. The importance of this relation is exemplified in the following result:

THEOREM 1. *If $G: E \sim F$, then the mapping $E_0 \rightarrow \varphi(E_0) = (E_0 \oplus G) \cap F$ is a lattice isomorphism of the principal lattice ideal generated by E in $L(A)$ onto the principal lattice ideal generated by F in $L(A)$. Under this mapping, moreover, semi-orthogonal left ideals contained in E correspond with semi-orthogonal left ideals contained in F.*

Proof. The proof is entirely lattice theoretic. Define a mapping ψ by $F_0 \rightarrow (G \oplus F_0) \cap E$ for each $F_0 \leqq F$; clearly both φ and ψ are isotone maps. By Corollary 2.2, we have $(F, G)M^*$ and $(G, E)M$. With these modularity relations, it is easy to compute $(\psi \circ \varphi)(E_0) = E_0$ for all $E_0 \leqq E$. Similarly $(\varphi \circ \varphi)(F_0) = F_0$ for all $F_0 \leqq F$. Thus φ is a lattice isomorphism with ψ its inverse mapping.

Now suppose $E_1, E_2 \leqq E$ with $E_1 \# E_2$. Since $E \# G$, $E_1 \oplus E_2 \# G$ also. Then $E_1 \oplus G \# E_2$ and we may compute $\varphi(E_1) \oplus G = [(E_1 \oplus G) \cap F] \oplus G = (E_1 \oplus G) \cap (F \oplus G) = (E_1 \oplus G) \cap (E \oplus G) = E_1 \oplus G \# E_2$, since $(F, G)M^*$. Thus $\varphi(E_1) \# E_2 \oplus G$, so that $\varphi(E_1) \# \varphi(E_2)$. Conversely, if $F_1, F_2 \leqq F$ with $F_1 \# F_2$, a similar argument shows $\psi(F_1) \# \psi(F_2)$.

LEMMA 1. [7, Th. 2]. *Let $eA = aA$ and $Af = Aa$. Then there exists a unique element $a^+ \in A$ such that*

(1) $aa^+ = e$.

(2) $fa^+ = a^+$.

Moreover,

(3) $a^+a = f$.

(4) $Ae = Aa^+$.

(5) $fA = a^+A$.

(6) $a = aa^+a$.

(7) $a^+ = a^+aa^+$.

Two idempotents e and f are *algebraically equivalent* via a and $b(a, b\colon e \sim f)$ if $e = ab$, $f = ba$, $a \in eAf$ and $b \in fAe$. This is easily seen to be an equivalence relation. The idempotents e and f are algebraically equivalent if and only if Ae and Af are isomorphic A-modules; moreover, in that case, the mapping $x \to bxa$ is a ring isomorphism of eAe onto fAf [4, pp. 21-23].

Notice that by Lemma 1, if $eA = aA$ and $Af = Aa$, then e and f are algebraically equivalent via a, a^+. This observation enables us to relate algebraic equivalence in A to semi-orthogonal perspectivity in $L(A)$.

THEOREM 2. *If $Ae \sim Af$, then $e \sim f$.*

Proof. Suppose $Ag\colon Ae \sim Af$. Put $a = e(1 - g)$ and $b = f(1 - g)$; then a and b are regular by Theorem 2.1 (2). An easy computation shows $eA = RL(e) = RL(e(1 - g)) = RL(a) = aA$ and similarly $fA = bA$. Moreover, $Ae \oplus Ag = Ag \oplus Af$ implies $R(a) = R(b)$; thus $Aa = LR(a) = LR(b) = Ab$. Choose an idempotent h with $Ah = Aa = Ab$. Then by our observation above, $e \sim h$ and $h \sim f$. Hence $e \sim f$.

For semi-orthogonal left ideals, the converse of Theorem 2 is also valid. We prove this as a first consequence of Lemma 2. With $Ae \# Af$, this fundamental lemma establishes a bijection of eAf onto, what might be termed, the set of relative semi-orthocomplements of Af in $Ae \oplus Af$.

LEMMA 2. *Let $E = Ae$ and $F = Af$ with $E \# F$.*

(1) *If $G \oplus F = E \oplus F$ with $G \in L(A)$, then $G = A(e - a)$ for some*

unique $a \in eAf$.

(2) *If* $a \in eAf$, *then there exists a left ideal* $G \in L(A)$ *such that*

(i) $G = A(e - a)$.

(ii) $G \oplus F = E \oplus F$.

(iii) $E \vee G = E \oplus LR(a)$.

(iv) $E \cap G = E \cap L(a)$.

Proof. To prove (1), let g be an idempotent generator for G. Choose w and x in A such that $e = wg + xf$. Then $e = ewg + exf$. Put $a = exf$. Then $e - a = ewg \in G$; so $A(e - a) \leqq G$. Conversely, $g = ye + zf = y(e - a) + ya + zf = yewg + ya + zf$ for some $y, z \in A$. But $g - yewg = ya + zf \in G \cap F = (0)$, so that $g = yewg = y(e - a)$. Hence $G = Ag \leqq A(e - a)$.

If also $b \in F = Af$ with $e - b \in G$, then $a - b = (e - b) - (e - a) \in G \cap F = (0)$; so $a = b$. This establishes the uniqueness of a.

To prove (2), let e_0 and f_0 denote orthogonal idempotent generators for E and F respectively. Put $g = e_0 - e_0 a$ and $G = Ag$. Since $ae_0 = afe_0 = aff_0e_0 = 0$, we find that $g = g^2$. Thus $G \in L(A)$. Now $g = e_0(e - a)$ and $e - a = e(e_0 - e_0 a) = eg$ implies $G = Ag = A(e - a)$, proving (i). The remaining parts of (2) are straightforward computations.

THEOREM 2. *Let* $Ae \# Af$. *Then* $Ae \sim Af$ *if and only if* $e \sim f$.

Proof. Suppose $a, b \colon e \sim f$. Put $G = A(e - a)$ and $H = A(f - b)$. Then by Lemma 2 (2), $G \oplus Af = Ae \oplus Af = Ae \oplus H$. But $e - a = ab - a = a(b - f) = -a(f - b)$ and $f - b = ba - b = b(a - e) = -b(e - a)$, showing that $G = A(e - a) = A(f - b) = H$. Thus $Ae \oplus G = Ae \oplus Af = G \oplus Af$.

4. Regularity. In this section, we characterize those Rickart rings A in which $E \cap F = (0)$ implies $E \# F$ for all E and F in $L(A)$. It will be convenient in the two lemmas and in Theorem 1 to adopt some notation. Let a and b denote regular elements with $Ae = Aa$ and $fA = bA$. Choose a^+ and b^+ by Lemma 3.1 so that $a^+a = e$ and $bb^+ = f$; choose idempotent generators g and h of $LR(ab)$ and $RL(ab)$ respectively. In the context of Rickart $*$-semigroups, Theorem 1 is due to D. J. Foulis [2].

LEMMA 1. *If* eb *or* af *is regular, then so is* ab.

Proof. Suppose eb is regular. Choose an idempotent generator k for Aeb and choose $(eb)^+$ so that $(eb)^+eb = k$. Put $x = (eb)^+a^+h$. Then $xab = (eb)^+a^+hab = (eb)^+a^+ab = (eb)^+eb = k$. Then $abxab = abk = (ae)bk = a(eb)k = a(eb) = (ae)b = ab$, showing that ab is regular. The argument for af is similar.

LEMMA 2. *If ab is regular, so are eb and af.*

Proof. Choose $(ab)^+$ so that $ab(ab)^+ = h$. Let k denote an idempotent generator of $LR(ef)$ and put $x = kb(ab)^+$. Then $afx = afkb(ab)^+ = (ae)fkb(ab)^+ = a(ef)kb(ab)^+ = a(ef)b(ab)^+ = (ae)fb(ab)^+ = afb(ab)^+ = ab(ab)^+ = h$. Hence $afxaf = haf = habb^+ = abb^+ = af$, showing that af is regular. Similarly eb is regular.

THEOREM 1. *ab is regular if and only if ef is regular.*

Proof. If ab is regular, then so is eb by Lemma 2. Since eb is regular, so is ef by Lemma 2 again, applied with $a = e$.

Conversely, if ef is regular, then so is eb by Lemma 1, applied with $a = e$. Then since eb is regular, so is ab by Lemma 1 again.

THEOREM 2. *These conditions are equivalent:*

(1) *ef is regular for every idempotent e and f.*

(2) *If a and b are regular, then so is ab.*

(3) *If $E \cap F = (0)$, then $E \sharp F$.*

Moreover, if A is a matrix ring, we may add

(4) *A is a regular ring.*

Proof. The equivalence of (1) and (2) is a consequence of Theorem 1. That (1) implies (3) is a consequence of Theorem 2.1 (2). Using the notation of the proof of Corollary 2.1, we may show that (3) implies (1); with $E = Ae$ and $F = Af$, we have $Ae_0 \cap Af_0 = (0)$ as before. Then by (3), $Ae_0 \sharp Af_0$. Consequently, $e_1(1 - f_1) = e_0(1 - f_0)$ is regular by Theorem 2.1, and hence $e(1 - f)$ is regular. Thus (3) implies $e(1 - f)$ is regular for every idempotent e and f, and this is evidently equivalent to (1).

Let us now suppose that A is a Rickart matrix ring of order $\geqq 2$. If A is a regular ring, then $E \cap F = (0)$ implies $E \sharp F$ for all E and F in $\boldsymbol{L}(A)$ by Theorem 2.1. Conversely, if this condition holds for all E and F in $\boldsymbol{L}(A)$, we show that A is a regular ring. To this end, let e_{ij}, $1 \leqq i, j \leqq n$, be a family of matrix units for A. We shall show that $e_{11}Ae_{11}$ and hence A, which is isomorphic to the $n \times n$ matrix ring over $e_{11}Ae_{11}$, is a regular ring.

Let $e_{11}xe_{11}$ denote an arbitrary element in $e_{11}Ae_{11}$; put $a = e_{11}xe_{12}$ and choose idempotent generators e and f for $RL(a)$ and $LR(a)$ respectively. Since $R(f) = R(a)$, $ae_{ii} = 0$ for $i \neq 2$ implies $fe_{ii} = 0$ for $i \neq 2$; since $L(e) = L(a)$, $e_{22}a = 0$ implies $e_{22}e = 0$. Thus $fe = f(\Sigma e_{ii})e = (\Sigma fe_{ii})e = (fe_{22})e = f(e_{22}e) = 0$, showing that $Ae \cap Af = (0)$. Moreover $f(1 - e) = f$ is regular. Hence $Ae \sharp Af$.

Now let e_0 and f_0 denote orthogonal idempotents generating Ae and Af respectively. Put $g = e_0 - e_0a$. Then, as in the proof of Lemma 3.2, $a = e(1 - g)$ and $Ag = A(e - a)$. Thus $Ae \cap Ag = Ae \cap L(a) = Ae \cap L(e) = (0)$. Then by hypothesis, $Ae \# Ag$. But this means that $a = e(1 - g)$ is regular in A. Choose an element b in A with $aba = a$.

Then

$$(e_{11}xe_{12})b(e_{11}xe_{12}) = aba = a = e_{11}xe_{12}$$

or equivalently

$$(e_{11}xe_{12})b(e_{11}xe_{11}) = e_{11}xe_{11}.$$

Thus

$$(e_{11}xe_{11})(e_{12}be_{11})(e_{11}xe_{11}) = e_{11}xe_{11},$$

showing that $e_{11}xe_{11}$ is a regular element of $e_{11}Ae_{11}$.

Hence $e_{11}Ae_{11}$ is a regular ring.

Recall that two left ideals in a von Neumann algebra A are semi-orthogonal if and only if their unique generating projections are non-asymptotic. Therefore, a von Neumann matrix algebra with no asymptotic pairs of projections must be regular and hence finite dimensional [8, pp. 85–87]. The definitive result in the general case is due to D. M. Topping [9]. Topping shows that in a von Neumann algebra these conditions are equivalent: (1) A has no asymptotic pairs of projections; (2) A contains no infinite orthogonal sequence of non-abelian projections; (3) A is the direct sum of an abelian subalgebra and a finite dimensional subalgebra. As a consequence of this result, a type II_1 von Neumann algebra may contain asymptotic pairs of projections, although its projection lattice is necessarily modular. Thus semi-orthogonality and dual modularity are in general distinct concepts. Using Foulis' characterization of dual modularity in terms of range-closedness, this same example shows that the product of two projections in a von Neumann algebra may have a closed range without being $*$-regular.

A simple proof, in the spirit of this paper, of (1) implies (2) in Baer $*$-rings would be worthwhile; for this would show that a complete $*$-regular ring can contain no infinite orthogonal sequence of non-abelian projections and hence no infinite orthogonal sequence of equivalent projections. A complete $*$-regular ring must, therefore, be of finite type. This is a difficult step in Irving Kaplansky's proof [3] that an orthocomplemented complete modular lattice is a continuous geometry.

References

1. J. Feldman, *Isomorphisms of rings of operators*, Thesis, University of Chicago, 1954.

2. D. J. Foulis, *Relative inverses in Baer *-semigroups*, Mich. Math. J., **10** (1963), 65-84.
3. I. Kaplansky, *Any orthocomplemented complete modular lattice is a continuous geometry*, Ann. of Math., **61** (1955), 524-541.
4. ———, *Rings of Operators*, New York: W. A. Benjamin, Inc., 1968.
5. G. W. Mackey, *On infinite dimensional linear spaces*, Trans. Amer. Math. Soc., **57** (1945), 155-207.
6. S. Maeda, *On relatively semi-orthocomplemented lattices*, J. Sci. Hiroshima Univ., **24A** (1960), 155-161.
7. D. D. Miller and A. H. Clifford, *Regular* D-*classes in semigroups*, Trans. Amer. Math. Soc., **82** (1956), 270-280.
8. J. von Neumann, *Continuous Geometry*, Princeton: Princeton Univ. Press, 1960.
9. D. M. Topping, *Asymptoticity and semimodularity in projection lattices*, Pacific. J. Math., **20** (1967), 317-325.

Received October 23, 1970.

UNIVERSITY OF MASSACHUSETTS
AND
KANSAS STATE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

ON THE SOLUTION OF LINEAR G.C.D. EQUATIONS

DAVID JACOBSON AND KENNETH S. WILLIAMS

Let Z denote the domain of ordinary integers and let $m(\geqq 1)$, $n(\geqq 1)$, $l_i(i=1, \cdots, m)$, $l_{ij}(i=1, \cdots, m; j=1, \cdots, n) \in Z$. We consider the solutions $x \in Z^n$ of

$$\text{(1)} \qquad \textbf{G.C.D.}\,(l_{11}x_1 + \cdots + l_{1n}x_n + l_1, \cdots, l_{m1}x_1 + \cdots + l_{mn}x_n + l_m, c) = d\,,$$

where $c(\neq 0)$, $d(\geqq 1) \in Z$ and G.C.D. denotes "greatest common divisor". Necessary and sufficient conditions for solvability are proved. An integer t is called a *solution modulus* if whenever x is a solution of (1), $x + ty$ is also a solution of (1) for all $y \in Z^n$. The positive generator of the ideal in Z of all such solution moduli is called the minimum modulus of (1). This minimum modulus is calculated and the number of solutions modulo it is derived.

1. Introduction. Let Z denote the domain of ordinary integers and let $m(\geqq 1)$, $n(\geqq 1)$, $l_i(i = 1, \cdots, m)$, $l_{ij}(i = 1, \cdots, m; j = 1, \cdots, n) \in Z$. We write $\boldsymbol{l} = (l_1, \cdots, l_m)$ and for each $i = 1, \cdots, m$ we write $\boldsymbol{l}_i = (l_{i1}, \cdots, l_{in})$ and $\boldsymbol{l}'_i = (l_{i1}, \cdots, l_{in}, l_i)$ so that $\boldsymbol{l} \in Z^m$, each $\boldsymbol{l}_i \in Z^n$, and each $\boldsymbol{l}'_i \in Z^{n+1}$. If $\boldsymbol{x} = (x_1, \cdots, x_n) \in Z^n$ we write in the usual way $\boldsymbol{l}_i \cdot \boldsymbol{x}$ for the linear expression $l_{i1}x_1 + \cdots + l_{in}x_n$. We let L denote the $m \times n$ matrix whose ith row is $\boldsymbol{l}_i$ and L' denote the $m \times (n + 1)$ matrix whose ith row is $\boldsymbol{l}'_i$.

Henceforth in this paper we will write the abbreviation G.C.D. for "greatest common divisor" of a finite sequence of integers, not all zero, and consider the solutions $\boldsymbol{x} \in Z^n$ of

$$\text{(1.1)} \qquad \text{G.C.D.}\,(\boldsymbol{l}_1 \cdot \boldsymbol{x} + l_1, \cdots, \boldsymbol{l}_m \cdot \boldsymbol{x} + l_m, c) = d\,,$$

where $c(\neq 0)$, $d(\geqq 1) \in Z$. A number of authors have either used or proved results concerning special cases of this equation (see for example [1], [5]) so that it is of interest to give a general treatment. This equation is clearly connected with the system

$$\text{(1.2)} \qquad \boldsymbol{l}_i \cdot \boldsymbol{x} + l_i \equiv 0 \pmod{d} \ (i = 1, \cdots, m)\,.$$

If we denote the number of incongruent solutions modulo d of (1.2) by $N(d, L')$, then $N(d, L') > 0$ is a necessary condition for the solvability of (1.1). A complete treatment of the system (1.2) has been given by Smith [4]. Let D_i = greatest common divisor of the determinants of all the $i \times i$ submatrices in L $(i = 1, \cdots, \min(m, n))$, D'_i = greatest common divisor of the determinants of all the $i \times i$ submatrices in L' ($i = 1, \cdots, \min(m, n+1)$), γ_i = greatest common divisor of d and $\frac{D_i}{D_{i-1}}$, $i = 1, \cdots, \min(m, n)$, where $D_0 = 1$, and γ'_i = greatest common divisor of d and $\frac{D'_i}{D'_{i-1}}$, $i = 1, \cdots, \min(m, n)$, where $D'_0 = 1$. Smith has shown that (1.2) is solvable if and only if

$$\prod_{i=1}^{\min(m,n)} \gamma_i = \prod_{i=1}^{\min(m,n)} \gamma'_i$$

and

$$\frac{D'_{n+1}}{D'_n} \equiv 0 \pmod{d}, \text{ if } m > n .$$

When solvable he shows that

$$N(d, L') = \gamma d^{\max(n-m, 0)} ,$$

where

$$\gamma = \prod_{i=1}^{\min(m,n)} \gamma_i .$$

We show in Theorem 1 that the conditions

$$(1.3) \qquad d \mid c, N(d, L') > 0, \text{G.C.D. } (\boldsymbol{l}_1, \cdots, \boldsymbol{l}_m, d) = \text{G.C.D. } (\boldsymbol{l}'_1, \cdots, \boldsymbol{l}'_m, c)$$

are both necessary and sufficient for solvability of (1.1). When (1.1) is solvable, (1.3) shows that the quantity g = G.C.D. $(\boldsymbol{l}_1, \cdots, \boldsymbol{l}_m, d)$ is a factor of $\boldsymbol{l}_i$, l_i ($i = 1, \cdots, m$), c and d. Cancelling this factor throughout we obtain the equation

$$\text{G.C.D. } (\boldsymbol{l}_1/g \cdot \boldsymbol{x} + l_1/g, \cdots, \boldsymbol{l}_m/g \cdot \boldsymbol{x} + l_m/g, c/g) = d/g .$$

This equation is equivalent to (1.1) in the sense that every solution of this equation is a solution of (1.1) and vice-versa. Thus we can suppose without loss of generality that

$$\text{G.C.D. } (\boldsymbol{l}_1, \cdots, \boldsymbol{l}_m, d) = 1 .$$

The solution set of (1.1) is denoted by $\mathscr{S}_d^c \equiv \mathscr{S}_d^c(L')$ that is,

$$(1.4) \quad \mathscr{S}_d^c \equiv \mathscr{S}_d^c(L') = \{\boldsymbol{x} \in Z^n \mid \text{G.C.D. } (\boldsymbol{l}_1 \cdot \boldsymbol{x} + l_1, \cdots, \boldsymbol{l}_m \cdot \boldsymbol{x} + l_m, c) = d\}.$$

Moreover when $\mathscr{S}_d^c \neq \varnothing$, we have

$$d \mid c, N(d, L') > 0, \text{G.C.D. } (\boldsymbol{l}'_1, \cdots, \boldsymbol{l}'_m, c) = 1 ,$$

and we write e for the integer c/d.

If $t \in Z$, $\boldsymbol{a} = (a_1, \cdots, a_n) \in Z^n$ and $\boldsymbol{b} = (b_1, \cdots, b_n) \in Z^n$, we say that

$\boldsymbol{a}$ and $\boldsymbol{b}$ are congruent modulo t (writing $\boldsymbol{a} \equiv \boldsymbol{b} \pmod{t}$) if and only if $a_i \equiv b_i \pmod{t}$ for each $i = 1, \cdots, n$. This congruence $\equiv$ is an equivalence relationship on Z^n. *If* $\mathscr{S}_d^c \neq \varnothing$, *any integer* t *for which this* equivalence relationship is preserved on $\mathscr{S}_d^c (\subseteq Z^n)$ is called a *solution modulus* of (1.1). Thus a solution modulus t has the property that if $\boldsymbol{x} \in \mathscr{S}_d^c$ then $\boldsymbol{x} + t\boldsymbol{y} \in \mathscr{S}_d^c$ for all $\boldsymbol{y} \in Z^n$. Clearly 0 and $\pm c$ are solution moduli. In Theorem 2 it is shown that the set of all solution moduli with respect to $\mathscr{S}_d^c$ viz.,

$$\mathfrak{M}_d^c \equiv \mathfrak{M}_d^c(L') = \{t \in Z \mid \boldsymbol{x} + t\boldsymbol{y} \in \mathscr{S}_d^c \text{ for all } \boldsymbol{x} \in \mathscr{S}_d^c \text{ and all } \boldsymbol{y} \in Z^n\} ,$$

is a principal ideal of Z. The positive generator of this ideal is denoted by $M_d^c(L')$ and called the *minimum modulus* of the equation (1.1). We show

$$M_d^c \equiv M_d^c(L') = d \prod_{p \mid e, N(pd, L') > 0} p \, . \tag{1.5}$$

(Here and throughout this paper the empty product is to be taken as 1). The product in (1.5) is taken over precisely those primes $p \mid e$ for which the system of congruences $\boldsymbol{l}_i \cdot \boldsymbol{x} + l_i \equiv 0 \pmod{pd}$ $(i = 1, \cdots, m)$ is solvable.

In § 5 we consider the problem of evaluating $\mathfrak{N}_d^c \equiv \mathfrak{N}_d^c(L')$, the number of incongruent solutions $\boldsymbol{x}$ of (1.1) modulo the minimum modulus M_d^c, from which the number of solutions modulo a given modulus can be determined. In Theorem 4 we derive a technical formula which allows the evaluation of $\mathfrak{N}_d^c$ in some important cases (see § 6). In particular we prove that if G.C.D. $(d, e) = 1$ then

$$\mathfrak{N}_d^c = N(d, L') \prod_{p \mid e, N(pd, L') > 0} p^n \left(1 - \frac{1}{p^{r(p,L)}}\right) , \tag{1.6}$$

where $r(p, L)$ is the rank of the matrix $L^{(p)}$ obtained from L by replacing each entry l_{ij} by its residue class modulo p in the finite field Z_p.

Finally in § 7 an alternative approach is given which enables us to generalize a recent result of Stevens [6].

2. **A necessary and sufficient condition for $\mathscr{S}_d^c \neq \varnothing$.** We begin by dealing with the case $d = 1$. We prove

LEMMA 1. $\mathscr{S}_1^c \neq \varnothing$ *if and only if*

$$\text{G.C.D.}\, (\boldsymbol{l}_1', \cdots, \boldsymbol{l}_m', c) = 1 \, . \tag{2.1}$$

Proof. The necessity of (2.1) is obvious. Thus to complete the proof it suffices to show that if (2.1) holds then $\mathscr{S}_1^c \neq \varnothing$. In view of (2.1) for each prime $p \mid c$ there must be some l_i or $l_{ij} \not\equiv 0 \pmod{p}$.

If some $l_i \not\equiv 0 \pmod p$ we let $\boldsymbol{x}^\dagger(p) = \boldsymbol{0}$, otherwise we have some $l_{ij} \not\equiv 0 \pmod p$ and we let $\boldsymbol{x}^\dagger(p) = (0, \cdots, 0, x_j, 0, \cdots, 0)$, where the j^{th} entry x_j is any solution of $l_{ij}x_j \equiv 1 \pmod p$, so that in both cases we have

$$\text{G.C.D.}\ (\boldsymbol{l}_1\cdot\boldsymbol{x}^\dagger(p) + l_1, \cdots, \boldsymbol{l}_m\cdot\boldsymbol{x}^\dagger(p) + l_m, p) = 1\ .$$

We now determine $\boldsymbol{x}$ by the Chinese remainder theorem so that $\boldsymbol{x} \equiv \boldsymbol{x}^\dagger(p) \pmod p$, for all $p\,|\,c$. Hence we have

$$\begin{aligned}\text{G.C.D.}\ &(\boldsymbol{l}_1\cdot\boldsymbol{x} + l_1, \cdots, \boldsymbol{l}_m\cdot\boldsymbol{x} + l_m, \prod_{p|c} p)\\ &= \prod_{p|c} \text{G.C.D.}\ (\boldsymbol{l}_1\cdot\boldsymbol{x} + l_1, \cdots, \boldsymbol{l}_m\cdot\boldsymbol{x} + l_m, p)\\ &= \prod_{p|c} \text{G.C.D.}\ (\boldsymbol{l}_1\cdot\boldsymbol{x}^\dagger(p) + l_1, \cdots, \boldsymbol{l}_m\cdot\boldsymbol{x}^\dagger(p) + l_m, p)\\ &= 1\ ,\end{aligned}$$

proving that $\boldsymbol{x} \in \mathscr{S}_1^c$.

Now we use Lemma 1 to handle the general case $d \geqq 1$. We prove

THEOREM 1. *$\mathscr{S}_d^c \neq \varnothing$ if and only if*

$$d\,|\,c,\ N(d, L') > 0,\ \text{G.C.D.}\ (\boldsymbol{l}_1, \cdots, \boldsymbol{l}_m, d) = \text{G.C.D.}\ (\boldsymbol{l}'_1, \cdots, \boldsymbol{l}'_m, c). \tag{2.2}$$

Proof. The necessity is obvious. Thus to complete the proof we must show that if (2.2) holds then $\mathscr{S}_d^c \neq \varnothing$. As $N(d, L') > 0$ there exists $\boldsymbol{k} \in Z^n$ and $\boldsymbol{h} = (h_1, \cdots, h_m) \in Z^m$ such that

$$\boldsymbol{l}_i\cdot\boldsymbol{k} + l_i = dh_i,\ i = 1, \cdots, m\ . \tag{2.3}$$

We write $d_1 = d/g$, $\boldsymbol{g}_i = \boldsymbol{l}_i/g \in Z^n$, $\boldsymbol{g}'_i = \boldsymbol{l}'_i/g \in Z^{n+1}$, $g_i = l_i/g \in Z$ $(i = 1, \cdots, m)$ where $g = \text{G.C.D.}\ (\boldsymbol{l}_1, \cdots, \boldsymbol{l}_m, d)$ and suppose that

$$\text{G.C.D.}\ (\boldsymbol{g}_1, \cdots, \boldsymbol{g}_m, \boldsymbol{h}, e) > 1\ , \tag{2.4}$$

where $e = c/d$. Then there exists a prime p such that

$$\boldsymbol{g}_i \equiv \boldsymbol{0}\ (i = 1, \cdots, m),\ \boldsymbol{h} \equiv \boldsymbol{0},\ e \equiv 0 \pmod p)\ . \tag{2.5}$$

Now from (2.3) we have

$$\boldsymbol{g}_i\cdot\boldsymbol{k} + g_i = d_1h_i,\ i = 1, \cdots, m\ ,$$

and so appealing to (2.5) we deduce $g_i \equiv 0 \pmod p$ $(i = 1, \cdots, m)$, giving $\boldsymbol{g}'_i \equiv \boldsymbol{0} \pmod p$ $(i = 1, \cdots, m)$. Thus we have G.C.D. $(\boldsymbol{g}'_1, \cdots, \boldsymbol{g}'_m, d_1e) \equiv 0 \pmod p$, which contradicts G.C.D. $(\boldsymbol{g}'_1, \cdots, \boldsymbol{g}'_m, d_1e) = 1$. Hence our assumption (2.4) is incorrect and we have G.C.D. $(\boldsymbol{g}_1, \cdots, \boldsymbol{g}_m, \boldsymbol{h}, e) = 1$. Thus by Lemma 1 there exists $\boldsymbol{\lambda} \in Z_n$ such that

$$\text{G.C.D.}\ (\boldsymbol{g}_1\cdot\boldsymbol{\lambda} + h_1, \cdots, \boldsymbol{g}_m\cdot\boldsymbol{\lambda} + h_m, e) = 1$$

and so $\boldsymbol{x} = d_1\boldsymbol{\lambda} + \boldsymbol{k} \in \mathscr{S}_d^c$.

3. Throughout the rest of this paper we suppose that $\mathscr{S}_d^c \neq \varnothing$ and G.C.D. $(\boldsymbol{l}_1, \cdots, \boldsymbol{l}_m, d) = 1$. Thus by Theorem 1 we have $d|c$, $N(d, L') > 0$ and G.C.D. $(\boldsymbol{l}'_1, \cdots, \boldsymbol{l}'_m, c) = 1$. Also throughout this paper corresponding to any $\boldsymbol{x} \in \mathscr{S}_d^c$ we define $\boldsymbol{u} \in Z^m$ by $\boldsymbol{u} = (u_1, \cdots, u_m)$, where $\boldsymbol{l}_i \cdot \boldsymbol{x} + l_i = du_i (i = 1, \cdots, m)$, so that G.C.D. $(\boldsymbol{u}, e) = 1$. The following lemmas will be needed later.

LEMMA 2. (i) *If $\boldsymbol{x} \in \mathscr{S}_d^c$ and p is a prime dividing e for which the system of simultaneous congruences*

$$\boldsymbol{l}_i \cdot \boldsymbol{z} + u_i \equiv 0 \pmod{p}, \ i = 1, \cdots, m \ , \tag{3.1}$$

is solvable then $N(pd, L') > 0$.

(ii) *Conversely if p is a prime dividing e for which $N(pd, L') > 0$ then there exists $\boldsymbol{x} \in \mathscr{S}_d^c$ such that* (3.1) *is solvable.*

Proof. (i) For $\boldsymbol{x} \in \mathscr{S}_d^c$ and z a solution of (3.1) we let $\boldsymbol{w} = \boldsymbol{x} + d\boldsymbol{z}$. Then for $i = 1, \cdots, m$ we have

$$\begin{aligned} \boldsymbol{l}_i \cdot \boldsymbol{w} + l_i &= (\boldsymbol{l}_i \cdot \boldsymbol{x} + l_i) + d\boldsymbol{l}_i \cdot \boldsymbol{z} \\ &= d(u_i + \boldsymbol{l}_i \cdot \boldsymbol{z}) \\ &\equiv 0 \pmod{pd} \ , \end{aligned}$$

showing that $N(pd, L') > 0$.

(ii) We define v_i by $\boldsymbol{l}_i \cdot \boldsymbol{w} + l_i = pdv_i$ $(i = 1, \cdots, m)$ and claim that

$$\text{G.C.D. } (\boldsymbol{l}_1, \cdots, \boldsymbol{l}_m, pv_1, \cdots, pv_m, e) = 1 \ . \tag{3.2}$$

For if not there is a prime $p'|e$ such that

$$\boldsymbol{l}_i \equiv \boldsymbol{0}, \ pv_i \equiv 0 \pmod{p'} \ (i = 1, \cdots, m) \ .$$

Thus from $\boldsymbol{l}_i \cdot \boldsymbol{w} + l_i = d\,pv_i$ we have $l_i \equiv 0 \pmod{p'}$ $(i = 1, \cdots, m)$, giving $\boldsymbol{l}'_i \equiv \boldsymbol{0} \pmod{p'}$ $(i = 1, \cdots, m)$, which contradicts G.C.D. $(\boldsymbol{l}'_1, \cdots, \boldsymbol{l}'_m, \text{de}) = 1$. Hence (3.2) is valid and so by Lemma 1 we can find $\boldsymbol{t} \in Z^n$ such that

$$\text{G.C.D. } (\boldsymbol{l}_1 \cdot \boldsymbol{t} + pv_1, \cdots, \boldsymbol{l}_m \cdot \boldsymbol{t} + pv_m, e) = 1 \ .$$

We set $\boldsymbol{x} = \boldsymbol{w} + d\,\boldsymbol{t}$ so that for $i = 1, \cdots, m$ we have

$$\boldsymbol{l}_i \cdot \boldsymbol{x} + l_i = d(\boldsymbol{l}_i \cdot \boldsymbol{t} + pv_i) \ ,$$

giving

$$\begin{aligned} &\text{G.C.D. } (\boldsymbol{l}_1 \cdot \boldsymbol{x} + l_1, \cdots, \boldsymbol{l}_m \cdot \boldsymbol{x} + l_m, c) \\ &\quad = d \text{ G.C.D. } (\boldsymbol{l}_1 \cdot \boldsymbol{t} + pv_1, \cdots, \boldsymbol{l}_m \cdot \boldsymbol{t} + pv_m, e) \\ &\quad = d \ , \end{aligned}$$

so that $\boldsymbol{x} \in \mathscr{S}_d^c$. Finally taking $z = -\boldsymbol{t}$ we see that the system

$$\boldsymbol{l}_i \cdot z + u_i \equiv 0 \pmod{p} \quad (i = 1, \cdots, m)$$

is solvable, as $u_i = \boldsymbol{l}_i \cdot \boldsymbol{t} + pv_i$.

LEMMA 3. *Let t be a positive integer, A a subset of Z^n which consists of $A(t)$ distinct congruence classes modulo t. Now if t' is a positive integer such that $t \mid t'$ then A consists of $(t'/t)^n A(t)$ congruence classes modulo t'.*

Proof. It suffices to prove that a congruence class C modulo t of A consists of $(t'/t)^n$ classes modulo t'. This is clear for if $\boldsymbol{x} \in C$ then so does $\boldsymbol{x} + t\boldsymbol{y}_i$, $(i = 1, \cdots, (t'/t)^n)$, where the $\boldsymbol{y}_i$ are incongruent modulo t'/t, moreover the $\boldsymbol{x} + t\boldsymbol{y}_i$ are incongruent modulo t' and every member of C is congruent modulo t' to one of them.

4. The minumum modulus. In this section we determine the minimum modulus M_d^c. We prove

THEOREM 2. *If $\mathscr{S}_d^c \neq \varnothing$ and G.C.D. $(\boldsymbol{l}_1, \cdots, \boldsymbol{l}_m, d) = 1$ the minimum modulus M_d^c with respect to $\mathscr{S}_d^c$ is given by*

$$M_d^c = d \prod_{p \mid e, N(pd, L') > 0} p \,. \tag{4.1}$$

Proof. As $\mathscr{S}_d^c \neq \varnothing$, $\mathfrak{M}_d^c$—the set of all solution moduli with respect to $\mathscr{S}_d^c$—is well-defined and moreover $\mathfrak{M}_d^c$ is non-empty as 0 and $\pm c$ belong to $\mathfrak{M}_d^c$. The proof will be accomplished by showing that $\mathfrak{M}_d^c$ is a principal ideal of Z generated by $d \prod_{p \mid e, N(pd, L') > 0} p$.

(i) We begin by showing that $\mathfrak{M}_d^c$ is an ideal of Z. It suffices to prove that if $t_1 \in \mathfrak{M}_d^c$ and $t_2 \in \mathfrak{M}_d^c$ then $t_1 - t_2 \in \mathfrak{M}_d^c$. For any $\boldsymbol{x} \in \mathscr{S}_d^c$ and any $\boldsymbol{y} \in Z^n$ we have $\boldsymbol{x} + t_1\boldsymbol{y} \in \mathscr{S}_d^c$, as $t_1 \in \mathfrak{M}_d^c$. Hence as $t_2 \in \mathfrak{M}_d^c$ we have

$$(\boldsymbol{x} + t_1\boldsymbol{y}) + t_2(-\boldsymbol{y}) \in \mathscr{S}_d^c \,,$$

that is

$$\boldsymbol{x} + (t_1 - t_2)\, \boldsymbol{y} \in \mathscr{S}_d^c \,,$$

so that

$$t_1 - t_1 \in \mathfrak{M}_d^c \,.$$

(ii) Next we show that $k = d \prod_{p \mid e, N(pd, L') > 0} p \in \mathfrak{M}_d^c$.
For $\boldsymbol{x} \in \mathscr{S}_d^c$ and any $\boldsymbol{y} \in Z^n$ we have

$$\begin{aligned} \text{G.C.D.}\ (\boldsymbol{l}_1\cdot(\boldsymbol{x}+k\boldsymbol{y})+l_1, \cdots, \boldsymbol{l}_m\cdot(\boldsymbol{x}+k\boldsymbol{y})+l_m, c) &= \text{G.C.D.}\ (\boldsymbol{l}_1\cdot\boldsymbol{x}+l_1+k(\boldsymbol{l}_1\cdot\boldsymbol{y}), \cdots, \boldsymbol{l}_m\cdot\boldsymbol{x}+l_m+k(\boldsymbol{l}_m\cdot\boldsymbol{y}), de) \\ &= d\ \text{G.C.D.}\ (u_1+k_1\,(\boldsymbol{l}_1\cdot\boldsymbol{y}), \cdots, u_m+k_1\,(\boldsymbol{l}_m\cdot\boldsymbol{y}), e)\ , \end{aligned}$$

where $k_1 = k/d$. To complete the proof we must show that for all $\boldsymbol{y} \in Z^n$ we have

$$\text{G.C.D.}\ (u_1+k_1\,(\boldsymbol{l}_1\cdot\boldsymbol{y}), \cdots, u_m+k_1\,(\boldsymbol{l}_m\cdot\boldsymbol{y}), e) = 1\ .$$

Suppose that this is not the case. Then there exists $\boldsymbol{y}_0 \in Z^n$ and a prime $p|e$ such that $u_i + k_1\,(\boldsymbol{l}_i\cdot\boldsymbol{y}_0) \equiv 0 \pmod p$ for $i = 1, \cdots, m$. Let $\boldsymbol{z} = \boldsymbol{x} + k\boldsymbol{y}_0$ so that for $i = 1, \cdots, m$ we have

$$\begin{aligned} \boldsymbol{l}_i\cdot\boldsymbol{z}+\boldsymbol{l}_i &= \boldsymbol{l}_i\cdot\boldsymbol{x}+l_i+k\,(\boldsymbol{l}_i\cdot\boldsymbol{y}_0) \\ &= d\,(u_i+k_1\,(\boldsymbol{l}_i\cdot\boldsymbol{y}_0))\ , \end{aligned}$$

that is,

$$\boldsymbol{l}_i\cdot\boldsymbol{z}+l_i \equiv 0 \pmod{pd}\ ,$$

so that $N(pd, L') > 0$. Hence as $p|e$ we have $p|k_1$ and so $p|u_i$ for $i = 1, \cdots, m$. This is the required contradiction as G.C.D. $(u_1, \cdots, u_m, e) = 1$, since $\boldsymbol{x} \in \mathscr{S}_d{}^c$.

(iii) In (i) we showed that $\mathfrak{M}_d^c$ is an ideal of Z and since Z is a principal ideal domain, $\mathfrak{M}_d^c$ is principal. Thus by the definition of the minimum modulus M_d^c we have $\mathfrak{M}_d^c = (M_d^c)$. In (ii) we showed that $k \in \mathfrak{M}_d^c$ so that $M_d^c|k$. Hence to show that $M_d^c = k$ we have only to show that $k|M_d^c$.

Now for all $\boldsymbol{x} \in \mathscr{S}_d{}^c$ and all $\boldsymbol{y} \in Z^n$ we have

$$\text{G.C.D.}\ (\boldsymbol{l}_1\cdot(\boldsymbol{x}+M_d^c\,\boldsymbol{y})+l_1, \cdots, \boldsymbol{l}_m\cdot(\boldsymbol{x}+M_d^c\,\boldsymbol{y})+l_m, c) = d\ .$$

Hence

$$\text{G.C.D.}\ (du_1+M_d^c\,\boldsymbol{l}_1\cdot\boldsymbol{y}, \cdots, du_m+M_d^c\,\boldsymbol{l}_m\cdot\boldsymbol{y}, d\,e) = d\ ,$$

and so we must have

$$M_d^c\,\boldsymbol{l}_i\cdot\boldsymbol{y} \equiv 0 \pmod d\ ,$$

for all $\boldsymbol{y} \in Z^n$ and all i $(1 \leqq i \leqq m)$. Taking in particular $\boldsymbol{y} = (0, \cdots, 0, 1, 0, \cdots, 0)$, where the 1 appears in the j^{th} place we must have for $i = 1, \cdots, m$ and $j = 1, \cdots, n$

$$M_d^c\,l_{ij} \equiv 0 \pmod d\ ,$$

that is

$$\text{G.C.D.}\ (M_d^c\,l_{11}, \cdots, M_d^c\,l_{mn}) \equiv 0 \pmod d$$

or

$$M_d^c \text{ G.C.D. } (\boldsymbol{l}_1, \cdots, \boldsymbol{l}_m) \equiv 0 \pmod d .$$

But G.C.D. $(\boldsymbol{l}_1 \cdots, \boldsymbol{l}_m, d) = 1$ so we must have $M_d^c \equiv 0 \pmod d$. Thus it suffices to prove that

$$k_1 | \pi_d^c, \textit{ where } k_1 = k/d = \prod_{p|e, N(pd, L')>0} p \textit{ and } \pi_d^c = M_d^c/d .$$

We suppose that $k_1 \nmid \pi_d^c$ so that there exists a prime $p|e$ for which the system $\boldsymbol{l}_i \cdot \boldsymbol{w} + l_i \equiv 0 \pmod{pd}$ $(i = 1, \cdots, m)$ is solvable yet $p \nmid \pi_d^c$. By Lemma 2 (ii) there exists $z \in Z^n$ such that for some $\boldsymbol{x} \in \mathscr{S}_d^c$ we have

$$\boldsymbol{l}_i \cdot \boldsymbol{z} + u_i \equiv 0 \pmod p, \; i = 1, \cdots, m .$$

As $p \nmid \pi_d^c$ we can define λ by $\pi_d^c \lambda \equiv 1 \pmod p$ and let $\boldsymbol{y} = \lambda z$ so that for $i = 1, \cdots, m$ we have

$$u_i + \pi_d^c \boldsymbol{l}_i \cdot \boldsymbol{y} \equiv 0 \pmod p . \tag{4.2}$$

But as M_d^c is the minimum modulus and $\boldsymbol{x} \in \mathscr{S}_d^c$ we must have

$$\text{G.C.D. } (\boldsymbol{l}_1 \cdot (\boldsymbol{x} + M_d^c \boldsymbol{y}) + l_1, \cdots, \boldsymbol{l}_m \cdot (\boldsymbol{x} + M_d^c \boldsymbol{y}) + l_m, c) = d ,$$

that is

$$\text{G.C.D. } (u_1 + \pi_d^c \boldsymbol{l}_1 \cdot \boldsymbol{y}, \cdots, u_m + \pi_d^c \boldsymbol{l}_m \cdot \boldsymbol{y}, e) = 1 ,$$

which is contradicted by (4.2). Hence $\pi_d^c = \prod_{p|e, N(pd, L')>0} p$ and this completes the proof.

We note the following important corollary of Theorem 2.

COROLLARY 1. *$\boldsymbol{x} \in Z^n$ is a solution of*

$$\text{G.C.D. } (\boldsymbol{l}_1 \cdot \boldsymbol{x} + l_1, \cdots, \boldsymbol{l}_m \cdot \boldsymbol{x} + l_m, c) = d \tag{4.3}$$

if and only if

$$\text{G.C.D. } (\boldsymbol{l}_1 \cdot \boldsymbol{x} + l_1, \cdots, \boldsymbol{l}_m \cdot \boldsymbol{x} + l_m, M_d^c) = d . \tag{4.4}$$

Proof. (i) Suppose $\boldsymbol{x}$ is a solution of (4.3). Then we can define u_i $(i = 1, \cdots, m)$ by $\boldsymbol{l}_i \cdot \boldsymbol{x} + l_i = du_i$ and we have

$$\text{G.C.D. } (u_1, \cdots, u_m, e) = 1 .$$

Hence we deduce

$$\text{G.C.D. } (u_1, \cdots, u_m, \prod_{p|e, N(pd, L')>0} p) = 1$$

and so

$$\text{G.C.D.}\ (\boldsymbol{l}_1\cdot\boldsymbol{x} + l_1, \cdots, \boldsymbol{l}_m\cdot\boldsymbol{x} + l_m, d \prod_{p\mid e, N(pd,L')>0} p) = d\ ,$$

which by Theorem 2 is

$$\text{G.C.D.}\ (\boldsymbol{l}_1\cdot\boldsymbol{x} + l_1, \cdots, \boldsymbol{l}_m\cdot\boldsymbol{x} + l_m, M_d^c) = d\ .$$

(ii) Conversely suppose $\boldsymbol{x}$ is a solution of (4.4). Then there exist u_i $(i = 1, \cdots, m)$ such that $l_i\cdot x + l_i = du_i$ and

$$\text{G.C.D.}\ (u_1, \cdots, u_m, \prod_{p\mid e, N(pd,L')>0} \mathrm{p}) = 1\ .$$

Suppose however that

$$\text{G.C.D.}\ (u_1, \cdots, u_m, e) \neq 1\ .$$

Then there exists a prime p such that

$$u_i \equiv 0\ (i = 1, \cdots, m), e \equiv 0\ (\mathrm{mod}\ p), N(pd, L') = 0.$$

But for $i = 1, \cdots, m$ we have

$$\boldsymbol{l}_i\cdot\boldsymbol{x} + l_i = du_i \equiv 0\ (\mathrm{mod}\ pd)\ ,$$

that is $N(pd, L') > 0$, which is the required contradiction. Hence we have

$$\text{G.C.D.}\ (u_1, \cdots, u_m, e) = 1$$

and so

$$\text{G.C.D.}\ (\boldsymbol{l}_1\cdot\boldsymbol{x} + l_1, \cdots, \boldsymbol{l}_m\cdot\boldsymbol{x} + l_m, c) = d\ .$$

5. Number of solutions with respect to the minimum modulus. We begin by evaluating $\mathfrak{N}_1^c$, that is, the number of solutions of (1.1), when $d = 1$, which are incongruent modulo M_1^c. We prove

THEOREM 3. $\mathfrak{N}_1^c = \prod_{p\mid c, N(p,L')>0} p^n\left(1 - \frac{1}{p^{r(p,L)}}\right)$, *where* $r(p, L)$ *is the rank of the matrix* $L^{(p)}$ *obtained from* L *by replacing each entry* l_{ij} *by its residue class modulo* p *in the finite field* Z_p.

Proof. By Corollary 1 the required number of solutions $\mathfrak{N}_1^c$ is just the number of solutions taken modulo M_1^c of

$$\text{G.C.D.}\ (\boldsymbol{l}_1\cdot\boldsymbol{x} + l_1, \cdots, \boldsymbol{l}_m\cdot\boldsymbol{x} + l_m, M_1^c) = 1\ .$$

Thus as $M_1^c = \prod_{p\mid c, N(p,L')>0} p$ is a product of distinct primes, a standard argument involving use of the Chinese remainder theorem shows that this number $\mathfrak{N}_1^c$ is just $\prod_{p \mid M_1^c} \mathfrak{N}(p)$, where $\mathfrak{N}(p)$ is the number of solutions $\boldsymbol{x}$ taken modulo p of

$$\text{G.C.D.}\ (\boldsymbol{l}_1 \cdot \boldsymbol{x} + l_1, \cdots, \boldsymbol{l}_m \cdot \boldsymbol{x} + l_m, p) = 1 \,. \tag{5.1}$$

Now $\boldsymbol{x}$ is a solution of (5.1) if and only if $\boldsymbol{x}^{(p)}$ is not a solution of the system (T denotes transpose)

$$L^{(p)}\boldsymbol{x}^{(p)T} + \boldsymbol{l}^{(p)T} = \boldsymbol{0}^T \,.$$

Since $N(p, L') > 0$, this system is consistent over the field Z_p and has $p^{n-r(p,L)}$ solutions. Thus the number of solutions (modulo p) of (5.1) is $p^n - p^{n-r(p,L)} = p^n\left(1 - \dfrac{1}{p^{r(p,L)}}\right)$, giving

$$\mathfrak{N}_1^c = \prod_{p \mid c, N(p,L') > 0} p^n \left(1 - \frac{1}{p^{r(p,L)}}\right)$$

as required.

In the proof of Theorem 2 we have seen that any solution modulus M of (1.1) is a multiple of M_d^c. As $\mathscr{S}_d^c$ consists of $\mathfrak{N}_d^c$ congruence classes modulo M_d^c, Lemma 3 shows that $\mathscr{S}_d^c$ consists of $(M/M_d^c)^n\mathfrak{N}_d^c$ congruence classes modulo M. Hence by Theorem 3 we have

COROLLARY 2. *The number of solutions $\boldsymbol{x}$ of* (1.1), *with* $d = 1$, *determined modulo M—a multiple of M_d^c—is*

$$M^n \prod_{p \mid c, N(p,L') > 0} \left(1 - \frac{1}{p^{r(p,L)}}\right) .$$

As a consequence of Corollary 2 we have the linear case of a result recently established by Stevens [6]. A generalization of this result is proved in § 7.

COROLLARY 3. (Stevens) *The number of solutions of*

$$\text{G.C.D.}\ (a_1x_1 + b_1, \cdots, a_nx_n + b_n, c) = 1 \,,$$

taken modulo c, is

$$c^n \prod_{p \mid c} \left(1 - \frac{\nu_1(p) \cdots \nu_n(p)}{p^n}\right) ,$$

where $\nu_i(p)(i = 1, \cdots, n)$ is the number of incongruent solutions modulo p of $a_ix_i + b_i \equiv 0 \pmod{p}$.

Proof. The system

$$a_i x_i + b_i \equiv 0 \pmod{p} \ (i = 1, \cdots, n) ,$$

is solvable if and only if

$$\text{G.C.D. } (a_i, p) \mid b_i \ (i = 1, \cdots, n) ,$$

that is, if and only if

$$p \nmid a_i \text{ or } p \mid \text{G.C.D. } (a_i, b_i) \ (i = 1, \cdots, n) .$$

Hence by Corollary 2 the required number of solutions is

$$c^n \prod_{p \mid c}{}' \left(1 - \frac{1}{p^{r(p)}}\right) , \tag{5.2}$$

where the dash (′) denotes that the product is taken over all p such that $p \nmid a_i$ or $p \mid$ G.C.D. (a_i, b_i) $(1 \leqq i \leqq n)$ and $r(p)$ is the number of a_i $(i = 1, \cdots, n)$ not divisible by p. As

$$\nu_i(p) = \begin{cases} 1, \ p \nmid a_i , \\ 0, \ p \mid a_i, \ p \nmid b_i , \\ p, \ p \mid a_i, \ p \mid b_i , \end{cases}$$

for $i = 1, \cdots, n$, (5.2) is just

$$c^n \prod_{p \mid c} \left(1 - \frac{\nu_1(p) \cdots \nu_n(p)}{p^n}\right) ,$$

which is the required result.

We now turn to the general case $d \geqq 1$. Let p be a prime and let E denote an equivalence class of $\mathscr{S}_d^c$ consisting of elements of $\mathscr{S}_d^c$ which are congruent modulo d. We assert that if $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in E$ then the system $\boldsymbol{l}_i \cdot \boldsymbol{z}^{(1)} + u_i^{(1)} \equiv 0 \pmod{p}$ $(i = 1, \cdots, n)$ is solvable if and only if the system $\boldsymbol{l}_i \cdot \boldsymbol{z}^{(2)} + u_i^{(2)} \equiv 0 \pmod{p}$ $(i = 1, \cdots, n)$ is solvable. As $\boldsymbol{x}^{(1)} \equiv \boldsymbol{x}^{(2)} \pmod{p}$ there exists $\boldsymbol{t} \in Z^n$ such that $\boldsymbol{x}^{(2)} = \boldsymbol{x}^{(1)} + d\boldsymbol{t}$. Hence for $i = 1, \cdots, n$ we have

$$\begin{aligned} du_i^{(2)} &= \boldsymbol{l}_i \cdot \boldsymbol{x}^{(2)} + l_i \\ &= \boldsymbol{l}_i \cdot \boldsymbol{x}^{(1)} + l_i + d\boldsymbol{l}_i \cdot \boldsymbol{t} \\ &= du_i^{(1)} + d\boldsymbol{l}_i \cdot \boldsymbol{t} \end{aligned}$$

giving

$$u_i^{(2)} = u_i^{(1)} + \boldsymbol{l}_i \cdot \boldsymbol{t} .$$

If there exists $\boldsymbol{z}^{(1)} \in Z^n$ such that $\boldsymbol{l}_i \cdot \boldsymbol{z}^{(1)} + u_i^{(1)} \equiv 0 \pmod{p}$ $(i = 1, \cdots, n)$ letting $\boldsymbol{z}^{(2)} = \boldsymbol{z}^{(1)} - \boldsymbol{t}$ we have $\boldsymbol{l}_i \cdot \boldsymbol{z}^{(2)} + u_i^{(2)} = \boldsymbol{l}_i \cdot \boldsymbol{z}^{(1)} - \boldsymbol{l}_i \cdot \boldsymbol{t} + u_i^{(1)} + \boldsymbol{l}_i \cdot \boldsymbol{t} \equiv 0 \pmod{p}$, which completes the proof of the assertion. Hence the solvability of the system

$$\boldsymbol{l}_i \cdot \boldsymbol{z} + u_i \equiv 0 \pmod{p} \quad (i = 1, \cdots, n)$$

depends only on the equivalence class E to which $\boldsymbol{x}$ (recall $\boldsymbol{l}_i \cdot \boldsymbol{x} + l_i = du_i$) belongs. Thus we can define a symbol $\delta_p(E)$ as follows:

$$\delta_p(E) = \begin{cases} 1, & \text{if for some } \boldsymbol{x} \in E \text{ (and thus for all } \boldsymbol{x} \in E) \text{ the system} \\ & \boldsymbol{l}_i \cdot \boldsymbol{z} + u_i \equiv 0 \pmod{p} \ (i = 1, \cdots, m) \text{ is solvable,} \\ 0, & \text{otherwise.} \end{cases}$$

We now prove the following result.

THEOREM 4. $\mathfrak{N}_d^c = \sum_{j=1}^{N(d,L')} \left\{ \prod_{p \mid e, N(pd, L') > 0} p^n \left(1 - \frac{1}{p^{r(p,L)}}\right)^{\delta_p(E^{(j)})} \right\}$, *where the* $E^{(j)}$ *denote the* $N(d, L')$ *congruence classes modulo* d *in* $\mathscr{S}_d^c$.

Proof. We let

$$\mathscr{S} = \{\boldsymbol{x} \in Z^n \mid \boldsymbol{l}_i \cdot \boldsymbol{x} + l_i \equiv 0 \pmod{d}, i = 1, \cdots, m\}$$

so that we have $\mathscr{S}_d^c \subseteqq \mathscr{S}$. Now $\mathscr{S}$ consists of $N(d, L')$ congruence classes modulo d and if we restrict this equivalence relation modulo d to $\mathscr{S}_d^c$, we show that $\mathscr{S}_d^c$ also contains the same number of classes. We write $E(\boldsymbol{x})$ (resp. $E'(\boldsymbol{x})$) for the equivalence class to which $\boldsymbol{x} \in \mathscr{S}_d^c$ (resp. $\boldsymbol{x} \in \mathscr{S}$) belongs. From the proof of Theorem 1 we see that for each $\boldsymbol{x} \in \mathscr{S}$ there exists $\boldsymbol{\lambda} \in Z^n$ such that $\boldsymbol{x} + d\boldsymbol{\lambda} \in \mathscr{S}_d^c$. We define a mapping f from the set of equivalence classes of $\mathscr{S}$ into the set of equivalence classes of $\mathscr{S}_d^c$ as follows: For $\boldsymbol{x} \in \mathscr{S}$

$$f(E'(\boldsymbol{x})) = E(\boldsymbol{x} + d\boldsymbol{\lambda}) .$$

This mapping is well-defined for if $\boldsymbol{x}' \in \mathscr{S}$ is such that $E'(\boldsymbol{x}') = E'(\boldsymbol{x})$ then $E(\boldsymbol{x}' + d\boldsymbol{\lambda}') = E(\boldsymbol{x} + d\boldsymbol{\lambda})$. f is onto for if $\boldsymbol{x} \in \mathscr{S}_d^c$ then $f(E'(\boldsymbol{x})) = E(\boldsymbol{x})$ and is also one-to-one, for if $f(E'(\boldsymbol{x})) = f(E'(\boldsymbol{y}))$, then $E(\boldsymbol{x} + d\boldsymbol{\lambda}) = E(\boldsymbol{y} + d\boldsymbol{\lambda}')$, that is $\boldsymbol{x} \equiv \boldsymbol{y} \pmod{d}$, giving $E'(\boldsymbol{x}) = E'(\boldsymbol{y})$. Thus the number of equivalence classes of $\mathscr{S}_d^c$ is the same as the number of equivalence classes of $\mathscr{S}$, that is $N(d, L')$.

Since $d \mid M_d^c$, each equivalence class E of $\mathscr{S}_d^c$, consists of a certain number of distinct classes in $\mathscr{S}_d^c$ modulo M_d^c. We now determine this number. If $\boldsymbol{x} \in E$, $\boldsymbol{x} + d\boldsymbol{t}$ also belongs in E if and only if it belongs in $\mathscr{S}_d^c$, that is, if and only if,

$$\text{G.C.D.}\ (\boldsymbol{l}_1 \cdot (\boldsymbol{x} + d\boldsymbol{t}) + l_1, \cdots, \boldsymbol{l}_m \cdot (\boldsymbol{x} + d\boldsymbol{t}) + l_m, c) = d ,$$

that is, if and only if,

$$\text{G.C.D.}\ (u_1 + \boldsymbol{l}_1\cdot\boldsymbol{t}, \cdots, u_m + \boldsymbol{l}_m\cdot\boldsymbol{t}, e) = 1\ . \tag{5.3}$$

Thus the number of distinct classes modulo M_d^c contained in E is just the number of distinct classes modulo $\pi_d^c = M_d^c/d$ which satisfy (5.3). But the minimum modulus of (5.3) is $\prod_{p|e} p^{\delta_p(E)}$. By lemma 2 (i) $\delta_p(E) = 1$ implies $N(pd, L') > 0$, so that $\prod_{p|e} p^{\delta_p(E)}$ divides $\prod_{p|e, N(pd,L')>0} p = \pi_d^c$. Writing $\prod_{p|e}^{+}$ for $\prod_{p|e,N(pd,L')>0}$ and $\prod_{p|e}^{0}$ for $\prod_{p|e,N(pd,L')=0}$, the required number of classes is by Corollary 2

$$\begin{aligned}
&= \prod_{p|e}^{+} p^n \cdot \prod_{p|e}\left(1 - \frac{1}{p^{r(p,L)}}\right)^{\delta_p(E)} \\
&= \prod_{p|e}^{+} p^n\left(1 - \frac{1}{p^{r(p,L)}}\right)^{\delta_p(E)} \cdot \quad \prod_{p|e}^{0}\left(1 - \frac{1}{p^{r(p,L)}}\right)^{\delta_p(E)} \\
&= \prod_{p|e}^{+} p^n\left(1 - \frac{1}{p^{r(p,L)}}\right)^{\delta_p(E)} ,
\end{aligned}$$

as $N(pd, L') = 0$ implies $\delta_p(E) = 0$.

Finally letting $E^{(1)}, \cdots, E^{(h)}$ denote the $h = N(d, L')$ distinct equivalence classes in $\mathscr{S}_d^c$ we deduce that the total number of incongruent solutions modulo M_d^c of (1.1) is

$$\sum_{j=1}^{N(d,L')} \left\{ \prod_{p|e,N(pd,L')>0} p^n \left(1 - \frac{1}{p^{r(p,L)}}\right)^{\delta_p(E^{(j)})} \right\} .$$

We remark that $r(p, L) \neq 0$, for $p|e$ and $\delta_p(E) = 1$. Otherwise, if $r(p, L) = 0$, $\boldsymbol{l}_i \equiv \boldsymbol{0} \pmod p$ $(i = 1, \cdots, m)$. But as $\delta_p(E) = 1$ then for $\boldsymbol{x} \in E$ the system $l_i\cdot z + u_i \equiv 0 \pmod p$ $(i = 1, \cdots, m)$ is solvable contradicting G.C.D. $(u_1, \cdots, u_m, e) = 1$.

6. Some special cases. We note a number of interesting cases of our results.

COROLLARY 4. *If* G.C.D. $(d, e) = 1$ *then the number* $\mathfrak{N}_d^c$ *of solutions of* (1.1) *modulo* M_d^c *is*

$$\mathfrak{N}_d^c = N(d, L') \prod_{p|e,N(pd,L')>0} p^n \left(1 - \frac{1}{p^{r(p,L)}}\right) .$$

Proof. By Theorem 4 it suffices to show that if G.C.D. $(d, e) = 1$, $p|e$, $N(pd, L') > 0$ then for all $\boldsymbol{x} \in \mathscr{S}_d^c$ we have $\delta_p(E) = 1$, that is the system $\boldsymbol{l}_i\cdot\boldsymbol{z} + u_i \equiv 0 \pmod p$ is solvable. Let $\boldsymbol{w}$ be a solution of $\boldsymbol{l}_i\cdot\boldsymbol{w} + l_i \equiv 0 \pmod{pd}$, say $\boldsymbol{l}_i\cdot\boldsymbol{w} + l_i = pdv_i$ $(i = 1, \cdots, m)$. As $p \nmid d$ we can define $z = d^{-1}(\boldsymbol{w} - \boldsymbol{x})$, where $dd^{-1} \equiv 1 \pmod p$ so that for $i = 1, \cdots, m$ we have

$$\begin{aligned} \boldsymbol{l}_i \cdot \boldsymbol{z} + u_i &= d^{-1}(\boldsymbol{l}_i \cdot \boldsymbol{w} - \boldsymbol{l}_i \cdot \boldsymbol{x}) + u_i \\ &= d^{-1}(pdv_i - l_i - du_i + l_i) + u_i \\ &= dd^{-1}(pv_i - u_i) + u_i \\ &\equiv 0 \pmod{p} , \end{aligned}$$

as required.

COROLLARY 5. *If* $N(d, L') = 1$ *then the number* $\mathfrak{N}_d^c$ *of solutions of* (1.1) *modulo* M_d^c *is*

$$\mathfrak{N}_d^c = \prod_{p \mid e, N(pd, L') > 0} p^n \left(1 - \frac{1}{p^{r(p,L)}}\right) . \tag{6.1}$$

In particular $N(d, L') = 1$ when L is invertible (mod d), and so $\mathfrak{N}_d^c$ is given by (6.1). Moreover if L is invertible modulo $d \prod_{p|e} p$ or c, then (1.1) is solvable and $\mathfrak{N}_d^c = \prod_{p|e}(p^n - 1)$.

Proof. This is immediate from Theorem 4 since by Lemma 2(ii), $\delta_p(E) = 1$ for all $p|e$, $N(pd, L') > 0$. Also (1.1) is solvable when L is invertible modulo $d \prod_{p|e} p$ as

$$\text{G.C.D. } (\boldsymbol{l}_1, \cdots, \boldsymbol{l}_m, d) = \text{G.C.D. } (\boldsymbol{l}'_1, \cdots, \boldsymbol{l}'_m, c) = 1 .$$

COROLLARY 6. *If* L *is invertible modulo* $\prod_{p|e, N(pd, L')>0} p$ *then the number of solutious of* (1.1) *modulo* M_d^c *is*

$$\mathfrak{N}_d^c = N(d, L') \prod_{p|e, N(pd, L')>0} (p^n - 1) .$$

Proof. Let p be any prime such that $p|e$ and $N(pd, \mathrm{L}') > 0$. Then L is invertible modulo p and so for any $\boldsymbol{x} \in \mathscr{S}_d^c$ the system

$$\boldsymbol{l}_i \cdot \boldsymbol{z} + u_i \equiv 0 \pmod{p} \; (1 = 1, \cdots, n)$$

is solvable and so $\delta_p(E^{(j)}) = 1$, $\mathrm{j} = 1, \cdots, N(d, L')$. Moreover as L is invertible modulo p we have $r(p, L) = n$ and the result follows from Theorem 4.

COROLLARY 7. *If*

$$\text{G.C.D. } (a_1, \cdots, a_n, d) = 1 \tag{6.2}$$

the equation

$$\text{G.C.D. } (a_1x_1 + \cdots + a_nx_n + b, c) = d \tag{6.3}$$

is solvable if and only if

$$d \mid c, \; \text{G.C.D. } (a_1, \cdots, a_n, b, c) = 1 . \tag{6.4}$$

The minimum modulus of (6.3) *is*

$$d \prod_{p|c/d}{}' p$$

and the number of solutions $\boldsymbol{x}$ *modulo this minimum modulus is*

$$d^{n-1} \prod_{p|c/d}{}' (p^n - p^{n-1}) ,$$

where the dash (′) *means that the product is taken over those primes* $p|c/d$ *such that* G.C.D. $(a_1, \cdots, a_n, p) = 1$.

Proof. According to Smith [4] or Lehmer [3] the number of solutions $\boldsymbol{x}$ taken modulo d of

$$a_1x_1 + \cdots + a_nx_n + b \equiv 0 \pmod{d}$$

is d^{n-1} G.C.D. $(a_1, \cdots, a_n, d)$ if G.C.D. $(a_1, \cdots, a_n, d)$ divides b and 0 otherwise. Thus as G.C.D. $(a_1, \cdots, a_n, d) = 1$, we have $N(d, L') = d^{n-1}$ and so by Theorem 1 (6.3) is solvable if and only if

$$d|c, \text{G.C.D. } (a_1, \cdots, a_n, b, c) = 1 .$$

Now if (6.3) is solvable and $p|c/d$ then

$$\text{G.C.D. } (a_1, \cdots, a_n, pd)|b$$

if and only if

$$\text{G.C.D. } (a_1, \cdots, a_n, p) = 1 ,$$

in view of (6.2) and (6.4). Thus by Theorem 2 the minimum modulus is

$$d \prod_{p|c/d}{}' p .$$

Finally for $p|c/d$, G.C.D. $(a_1, \cdots, a_n, p) = 1$ we have $r(p, L) = 1$ and moreover the congruence $a_1x_1 + \cdots + a_nx_n + u \equiv 0 \pmod{p}$ is always solvable so that $\delta_p(E^{(j)}) = 1, j = 1, \cdots, d^{n-1}$. Hence by Theorem 4 the number of solutions is

$$d^{n-1} \prod_{p|c/d}{}' p^n \left(1 - \frac{1}{p}\right) .$$

We remark that in particular ([5])

$$\text{G.C.D. } (ax + b, c) = 1$$

is solvable if and only if G.C.D. $(a, b, c) = 1$, has minimum modulus $\prod_{p|c, p\nmid a} p$, and has $\prod_{p|c,p\nmid a} (p - 1)$ solutions x modulo the minimum modulus.

COROLLARY 8. *There is a unique solution of* (1.1) *modulo* M_d^c *if and only if*

(i) $N(d, L') = 1$ *and there is no prime* p *such that*

$$p \mid e, N(pd, L') > 0 \ ,$$

or

(ii) $N(d, L') = 1$ *and the only prime* p *such that* $p \mid e, N(pd, L') > 0$, *is* $p = 2$, *and* $r(2, L) = 1$, $n = 1$.

Proof. If (1.1) possesses a unique solution modulo M_d^c, Theorem 4 shows that S can consist only of a single congruence class modulo d. Hence $N(d, L') = 1$. Also by Theorem 4 if there is no prime p such that $p \mid e$ and $N(pd, L') > 0$ then $\mathfrak{N}_d^c = 1$. Suppose however that there is such a prime p. Then by Corollary 5 we have

$$1 = \prod_{p \mid e, N(pd, L') > 0} (p^n - p^{n-r(p,L)}) \ .$$

This occurs if and only if

$$p^n - p^{n-r(p,L)} = 1 \ , \tag{6.5}$$

for all $p \mid e$ with $N(pd, L') > 0$. But the left-hand side of (6.5) is divisible by p unless $r(p, L) = n$. Then $p^n = 2$ and we have $p = 2$, $n = 1$, $r(p, L) = r(2, L) = 1$, which proves the theorem.

7. Another method. Although the formula of Theorem 4 applies to some important cases in § 6, this formula seems difficult to evaluate even for example in the diagonal case

$$\text{G.C.D.}\ (a_1x_1 + b_1, \cdots, a_nx_n + b_n, c) = d \ .$$

The inherent difficulty is in determining for a given prime p which solutions of this equation have the property that the system $a_iz_i + u_i \equiv 0 \pmod p$ $(i = 1, \cdots, n)$ is solvable. We now present another method which in conjunction with previous results yields the diagonal case.

We consider the set $\mathfrak{U}$ of $\boldsymbol{u} \in Z^m$ with G.C.D. $(\boldsymbol{u}, e) = 1$ for which the system

$$\boldsymbol{l}_i \cdot \boldsymbol{x} + l_i \equiv du_i \pmod c \ (i = 1, \cdots, n) \text{ is solvable} \ . \tag{7.1}$$

It is clear that if $\boldsymbol{u} \in \mathfrak{U}$ and $\boldsymbol{u} \equiv \boldsymbol{u}' \pmod e$ then $\boldsymbol{u}' \in \mathfrak{U}$. We denote by K_d^c the number of distinct classes modulo e contained in $\mathfrak{U}$. Let $\mathfrak{N}$ denote the number of solutions $\boldsymbol{x}$ of (1.1) modulo c. We prove

THEOREM 5. $\mathfrak{N} = K_d^c N_c(L^*)$ *where* L^* *is the* $m \times (n+1)$ *matrix*

$[L\colon 0]$.

Proof. If $\boldsymbol{x} \in \mathscr{S}_d^c$ then there exists $\boldsymbol{u} \in Z^n$ such that $\boldsymbol{l}_i \cdot \boldsymbol{x} + l_i = du_i$ $(i = 1, \cdots, m)$ and G.C.D. $(\boldsymbol{u}, e) = 1$. If $\boldsymbol{x}, \boldsymbol{x}' \in \mathscr{S}_d^c$ are such that $\boldsymbol{x} \equiv \boldsymbol{x}' \pmod{e}$ then $du_i \equiv du_i' \pmod{c}$, that is $u_i \equiv u_i' \pmod{e}$.

Conversely if G.C.D. $(\boldsymbol{u}, e) = 1$ and $\boldsymbol{x}$ satisfies $\boldsymbol{l}_i \cdot \boldsymbol{x} + l_i \equiv du_i$ (mod c) $(i = 1, \cdots, m)$ then $\boldsymbol{l}_i \cdot \boldsymbol{x} + l_i = d(u_i + \lambda_i e)$ and $\boldsymbol{x} \in \mathscr{S}_d^c$ as G.C.D. $(\boldsymbol{u} + \boldsymbol{\lambda} e, e) =$ G.C.D. $(\boldsymbol{u}, e) = 1$.

Thus $\boldsymbol{x} \in \mathscr{S}_d^c$ if and only if $\boldsymbol{x}$ is a solution of $\boldsymbol{l}_i \cdot \boldsymbol{x} + l_i \equiv du_i$ (mod c), where G.C.D. $(\boldsymbol{u}, e) = 1$. Now there are K_d^c incongruent classes of $\boldsymbol{u}$ modulo e, with G.C.D. $(\boldsymbol{u}, e) = 1$, for which (7.1) is solvable. For each one of these, (7.1) has $N_c(L\colon 0)$ incongruent solutions modulo c. Hence we have

$$\mathfrak{N} = K_d^c N_c(L^*)$$

as required.

We now obtain the following interesting result.

COROLLARY 9. *If $\boldsymbol{h} \in Z^n$ and $e_1, \cdots, e_n$ are divisors of e then the system*

$$u_i \equiv h_i \pmod{e_i} \quad (i = 1, \cdots, n) \tag{7.2}$$

has a solution $\boldsymbol{u} = (u_1, \cdots, u_n)$ such that G.C.D. $(\boldsymbol{u}, e) = 1$ *if and only if* G.C.D. $(e_1, \cdots, e_n, h_1, \cdots, h_n, e) = 1$. *When this holds* (7.2) *has*

$$\prod_{i=1}^{n} (e/e_i) \prod_{p \mid e}{}' \left(1 - \frac{1}{p^{r(p)}}\right)$$

distinct solutions $\boldsymbol{u}$ modulo e, for which G.C.D. $(\boldsymbol{u}, e) = 1$, *where $r(p) =$ number of e_i $(i = 1, \cdots, n)$ not divisible by p, and the dash (′) means that the product is taken over those primes $p \mid e$ such that $p \nmid e_i$ or $p \mid$* G.C.D. (e_i, h_i) $(i = 1, \cdots, n)$.

Proof. The system (7.2) has a solution $\boldsymbol{u}$ such that G.C.D. $(\boldsymbol{u}, e) = 1$ if and only if

$$\text{G.C.D. } (e_1 x_1 + h_1, \cdots, e_n x_n + h_n, e) = 1 \tag{7.3}$$

is solvable, which by Lemma 1 is the case if and only if G.C.D. $(e_1, \cdots, e_n, h_1, \cdots, h_n, e) = 1$. Applying Theorem 5 to (7.3) we have $\mathfrak{N} = K_1^e N_e(L^*)$ and we note that K_1^e is the number of distinct solutions $\boldsymbol{u}$ modulo e of (7.2) for which G.C.D. $(\boldsymbol{u}, e) = 1$. However $N_e(L*)$ is the number of solutions $\boldsymbol{x}$ modulo e such that $e_i x_i \equiv 0 \pmod{e}$ $(i = 1, \cdots, n)$. Clearly $N_e(L^*) = \prod_{i=1}^{n} e_i$. By Corollary 2

$$\mathfrak{N} = e^n \prod_{p \mid e, N(p,L')>0} \left(1 - \frac{1}{p^{r(p,L)}}\right),$$

where

$$L' = \begin{pmatrix} e_1 & & & h_1 \\ & \ddots & & \vdots \\ & & e_n & h_n \end{pmatrix}.$$

Now $N(p, L') > 0$ if and only if the system $e_i w_i + h_i \equiv 0 \pmod p$ $(i = 1, \cdots, n)$ is solvable, that is, if and only if G.C.D. $(p, e_i) \mid h_i$ or if and only if $p \nmid e_i$ or $p \mid$ G.C.D (e_i, h_i) $(i = 1, \cdots, n)$. Also $r(p, L)$ is just the number of the e_i $(i = 1, \cdots, n)$ not divisible by p. This completes the proof.

We now obtain a generalization of Steven's result [6] (see Corollary 3).

COROLLARY 10. *The equation*

$$\text{G.C.D. } (a_1x_1 + b_1, \cdots, a_nx_n + b_n, c) = d\ ,$$

where

$$\text{G.C.D. } (a_1, \cdots, a_n, d) = 1\ ,$$

is solvable if and only if

$$d \mid c,\ \text{G.C.D. } (a_i, d) \mid b_i\ (i = 1, \cdots, n)\ ,$$

$$\text{G.C.D. } (a_1, \cdots, a_n, b_1, \cdots, b_n, c) = 1.$$

The number of solution modulo c is given by

$$\prod_{i=1}^{n} \text{G.C.D. } (a_i, d) \cdot (c/d)^n \cdot \prod_{p \mid c/d} \left(1 - \frac{\nu_1(p) \cdots \nu_n(p)}{p^n}\right),$$

where $\nu_i(p)$ $(i = 1, \cdots, n)$ is the number of incongruent solutions modulo p of $\dfrac{a_i}{\text{G.C.D. } (a_i, d)} x + \dfrac{b_i}{\text{G.C.D. } (a_i, d)} \equiv 0 \pmod p$.

Proof. The necessary and sufficient conditions for solvability are immediate from Theorem 1. When solvable we calculate the number $\mathfrak{N}$ of solutions modulo c using Theorem 5. Thus we require the number of distinct $\boldsymbol{u}$ modulo e with G.C.D. $(\boldsymbol{u}, e) = 1$ such that

$$a_ix_i + b_i \equiv du_i \pmod{de}\ (i = 1, \cdots, n)$$

is solvable, that is,

$$(a_i/d_i)x_i + (b_i/d_i) \equiv (d/d_i)u_i \pmod{d/d_i \cdot e}$$

where $d_i = \text{G.C.D}\ (a_i, d)\ (i = 1, \cdots, n)$.
This is solvable if and only if

$$\text{G.C.D.}\ ((a_i/d_i), (d/d_i)e) \mid (d/d_i)u_i - (b_i/d_i) (i = 1, \cdots, n) ,$$

that is, if and only if,

$$(d/d_i)u_i \equiv (b_i/d_i)\ (\text{mod G.C.D.}\ ((a_i/d_i), e)\ (i = 1, \cdots, n) .$$

This system is equivalent to

$$u_i \equiv h_i\ (\text{mod G.C.D.}\ (a_i/d_i, e))\ (i = 1, \cdots, n) ,$$

where $h_i = (d/d_i)^{-1}b_i/d_i$ and $(d/d_i)^{-1}$ is an inverse of d/d_i modulo G.C.D. $(a_i/d_i, e)$ since G.C.D. $(d/d_i, a_i/d_i, e) = 1$. Thus by Corollary 9 the number of such $\boldsymbol{u}$ is

$$\prod_{i=1}^{n} \frac{e}{\text{G.C.D.}\ ((a_i/d_i), e)} \prod_{p|e}{}' \left(1 - \frac{1}{p^{r(p)}}\right) ,$$

where the dash (′) means that the product is taken over those $p|e$ such that $p|a_i/d_i$ or $p|\text{G.C.D.}\ (a_i/d_i, b_i/d_i)$, $i = 1, \cdots, n$, as $p|\text{G.C.D.}$ $(a_i/d_i, e, h_i)$ if and only if $p|\text{G.C.D.}\ (a_i/d_i, e, b_i/d_i)$ because $(d/d_i)h_i \equiv b_i/d_i$ (mod G.C.D. $(a_i/d_i, e)$ and G.C.D. $(d/d_i, a_i/d_i) = 1$ $(i = 1, \cdots, n)$. Also $r(p)$ is the number of a_i/d_i $(i = 1, \cdots, n)$ not divisible by p.

Next we need the number of incongruent $\boldsymbol{x}$ modulo de such that

$$a_i x_i \equiv 0 \pmod{de}\ (i = 1, \cdots, n) .$$

This is just

$$\begin{aligned} &\prod_{i=1}^{n} \text{G.C.D.}\ (a_i, de) \\ &= \prod_{i=1}^{n} d_i\ \text{G.C.D.}\ (a_i/d_i, (d/d_i)e) \\ &= \prod_{i=1}^{n} d_i\ \text{G.C.D.}\ (a_i/d_i, e) . \end{aligned}$$

Hence by Theorem 5 the required number of solutions is

$$\prod_{i=1}^{n} (d_i\ e) \cdot \prod_{p|e}{}' \left(1 - \frac{1}{p^{r(p)}}\right) ,$$

where the dash (′) means that the product is taken over those $p|e$ such that $p|a_i/d_i$ or $p|\text{G.C.D.}\ (a_i/d_i, b_i/d_i)$, $i = 1, \cdots, n$. This number is

$$\prod_{i=1}^{n} d_i \cdot e^n \cdot \prod_{p|e} \left(1 - \frac{\nu_1(p) \cdots \nu_n(p)}{p^n}\right),$$

as

$$\nu_i(p) = \begin{cases} 1, & p \nmid a_i/d_i, \\ 0, & p \mid a_i/d_i,\ p \nmid b_i/d_i\ , \\ p, & p \mid a_i/d_i,\ p \mid b_i/d_i\ . \end{cases}$$

Finally we state that all formulas are easily modified if we do not assume $g = \text{G.C.D.}\ (\boldsymbol{l}_1, \cdots, \boldsymbol{l}_m, d) = 1$ (See introduction, Theorem 1). For example we list

THEOREM 2′. *If $\mathscr{S}_d^c \neq \varnothing$ the minimum modulus M_d^c with respect to* (1.1) *is given by*

$$M_d^c = d_1 \prod_{p \mid e, N(pd_1, L'/g) > 0} p\ .$$

COROLLARY 4′. *If* G.C.D. $(d, e) = 1$ *then the number $\mathfrak{N}_d^c$ of solutions of* (1.1) *modulo M_d^c is*

$$\mathfrak{N}_d^c = N(d, L'/g) \prod_{p \mid e, N(pd_1, L'/g) > 0} p^n \left(1 - \frac{1}{p^{r(p, L/g)}}\right)\ .$$

REFERENCES

1. T. M. Apostol, *Euler's $\varnothing$ − function and separable Gauss sums*, Proc. Amer. Math. Soc., **24** (1970), 482-485.
2. L. E. Dickson, *History of the Theory of Numbers*, Chelsea N.Y., (1952), 88-93.
3. D. N. Lehmer, *Certain theorems in the theory of quadratic residues*, Amer Math. Monthly, **20** (1913), 155-156.
4. H. J. S. Smith, *On systems of linear indeterminate equations and congruences*, Phil. Trans. Lond., **151** (1861), 293-326. (Collected Mathematical Papers Vol. 1, Chelsea N. Y. (1965), 367-409.)
5. R. Spira, *Elementary problem no. E1730*, Amer. Math. Monthly, **72** (1965), 907.
6. H. Stevens, *Generalizations of the Euler $\varnothing$ − function*, Duke Math. J., **38** (1971), 181-186.

Received November 30, 1970, and in revised form April, 1971. This research was supported by a National Research Council of Canada Grant (No. A-7233).

CARLETON UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

ON RANK 3 PROJECTIVE PLANES

MICHAEL J. KALLAHER

One of the unsolved problems in the theory of projective planes is the following: Is every finite projective plane with a transitive collineation group desarguesian? This problem is investigated under the additional hypothesis that the group has rank 3. It is proven that if a projective plane $\mathscr{P}$ of order $n > 2$ has a rank 3 collineation group then $\mathscr{P}$ is non-desarguesian and either (i) n is odd and $n = m^4$, or (ii) n is even and $n = m^2$ with $m = 0 \pmod 4$.

One of the older problems in the theory of projective planes is the following: If a finite projective plane $\mathscr{P}$ has a collineation group G transitive on its points, is the plane desarguesian? So far only under one or more additional assumptions has the answer been shown to be yes. The basis for these results is the theorem, due to Wagner [10], that if the transitive group G contains a central collineation, then $\mathscr{P}$ is desarguesian.

Ostrom and Wagner [8] showed that if the group G is doubly transitive then $\mathscr{P}$ is desarguesian, and G contains all elations of $\mathscr{P}$. Higman and McLaughlin [4] investigated the problem in the case when the group G is transitive on the flags of $\mathscr{P}$. They proved that under certain restrictions on the order of $\mathscr{P}$ the plane is desarguesian. Keiser [6] and Wagner [10] have showned that under restrictions on the order of $\mathscr{P}$ and the order of G the plane is desarguesian.

The rank of a permutation group G transitive on a set Ω is the number of orbits of G_P, P a point of Ω, in Ω. Hence a transitive group G has rank 2 on a set Ω if and only if G is doubly transitive on Ω. G has rank 3 if and only if for every point $P \in \Omega$ G_P has two orbits besides G_P. Ostrom and Wagner have thus answered the question when the group G has rank 2. It is then natural to ask: If a finite projective plane has a transitive collineation group of rank 3, is the plane desarguesian?

Investigating this question we have found that a more appropriate question is: Which finite projective planes have rank 3 collineation groups? For we will prove in this article the following

MAIN THEOREM. *Let $\mathscr{P}$ be a finite projective plane of order n with rank 3 collineation group G. Then n satisfies one of the statements*:

(i) $n = 2$

(ii) *n is odd and $n = m^4$*

(iii) *n is even and* $n = m^2$ *with* $m \equiv 0 \pmod 4$
Furthermore only in case (i) *is* $\mathscr{P}$ *desarguesian.*

The proof consists in showing that G must be a nonsolvable flag-transitive group of even order. It is not known whether $\mathscr{P}$ can actually exist in cases (ii) or (iii). Hence the theorem leads to the following conjecture:

CONJECTURE. The only finite projective plane having a rank 3 collineation group is the desarguesian plane of order two.

The desarguesian plane of order two (Fig. 1) has the rank 3 collineation group G generated by the collineations $\sigma = (P_1P_7P_6P_4P_3P_2P_5)$ and $\tau = (P_7P_6P_3)(P_4P_5P_2)$. Note that G is solvable, sharply flag-transitive, and has order 21.

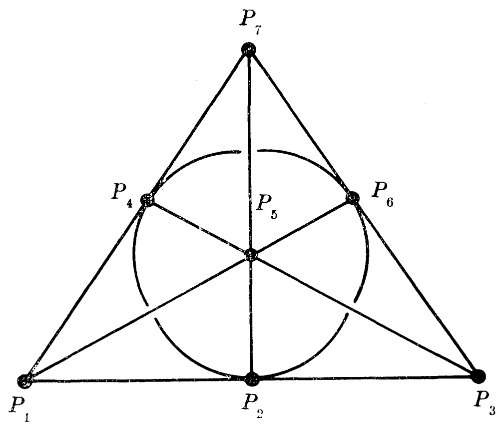

FIGURE 1

We wish to thank Professor Ostrom for many helpful suggestions and for reading a preliminary draft.

2. Definitions and results required later. We assume the reader is familiar with the theory and terminology of projective planes as appears, for example, in Chapters 3–5 of Dembowski [3]. Also a familiarity with the simpler aspects of permutation group theory (as in Chapter 1 of Wielandt [12]) will be assumed.

A *transitive collineation group* on a projective plane $\mathscr{P}$ is one which is transitive on the points (and hence on the lines by Result II below). A *rank* 3 *collineation group* of $\mathscr{P}$ is a transitive collineation group which has rank 3 as a permutation group on the points of $\mathscr{P}$. A flag of $\mathscr{P}$ is an incident point-line pair; i.e., a pair P, l with P a point, l a line, of $\mathscr{P}$ and $P \in l$. A flag-transitive collineation group is one which is transitive on the flags of $\mathscr{P}$.

Use will be made of a number of results concerning permutations and collineations. These are listed below for convenience:

Result I (Ostrom [27]; Dembowski [3], p. 214): If $\mathscr{P}$ is a finite projective plane having a collineation group G which is transitive on nonincident point-line pairs, then G is doubly transitive on the points of $\mathscr{P}$.

Result II (Dembowski [2], Hughes [5], Parker [9]): A collineation group of a finite projective plane has equally many point and line orbits.

Result III (Higman and McLaughlin [4]): Let $\mathscr{P}$ be a finite projective plane of order n with a flag-transitive collineation group G. If n is odd and not a fourth power, then $\mathscr{P}$ is desarguesian and G contains all elations of $\mathscr{P}$ (See also Dembowski [3], p. 212).

Result IV (Higman and McLaughlin [4]): Let G be a flag-transitive collineation group of a desarguesian projective plane $\mathscr{P}$ of order n. G contains all elations of $\mathscr{P}$ with precisely two exceptions:
(i) $n = 2$ and G has order 21.
(ii) $n = 8$ and G has order 657.

Result V (Wagner [10]): If $\mathscr{P}$ is a finite projective plane having a transitive collineation group G which contains a nontrivial central collineation, then $\mathscr{P}$ is desarguesian and G contains all elations of $\mathscr{P}$.

Result VI (Keiser [6]): Let $\mathscr{P}$ be a finite projective plane of order n with a transitive collineation group G. If G is nonsolvable and if $n = m^2$ with $m \equiv 2 \pmod 4$ or $m \equiv 3 \pmod 4$, then $\mathscr{P}$ is desarguesian and G contains all the elations of $\mathscr{P}$.

Results VII (Dembowski [3], p. 212): Let $\mathscr{P}$ be a finite projective plane of order n. If G is a collineation group which is solvable and primitive on the points of $\mathscr{P}$, then $n^2 + n + 1$ is a prime.

For the next result we note that a permutation group on a set Ω is *regular* (*sharply transitive*) if it is transitive and G_P consists only of the identity for each point $P \in \Omega$. G is a *Frobenius group* on Ω if (i) it is transitive, (ii) G_P is nontrivial for each point $P \in \Omega$, and (iii) for every two distinct points $P, Q \in \Omega$ $G_{P,Q}$ consists only of the identity.

Result VIII (Wielandt [11], 11.6): A transitive permutation group of prime degree is solvable if and only if it is either regular or a Frobenius group (Due to E. Galois).

3. The investigation. In this section we prove that if a finite projective plane $\mathscr{P}$ has a rank 3 group G of collineations then G is flag-transitive on $\mathscr{P}$, G is nonsolvable (if $n > 2$), and $\mathscr{P}$ is not desarguesian (if $n > 2$). We start with

LEMMA 1. *Let $\mathscr{P}$ be a finite projective plane and G a rank 3 group of collineation of $\mathscr{P}$. Then for every point P of $\mathscr{P}$, G_P permutes the lines through P in one or two orbits.*

Proof. G_P permutes the points different than P in two orbits $\mathscr{O}_1$ and $\mathscr{O}_2$. Define the sets $\mathscr{M}_i$, $i = 1, 2$, by: $\mathscr{M}_i$ consists of the lines through P such that l contains at least one point in $\mathscr{O}_i$. G_P is transitive on $\mathscr{M}_1$ and $\mathscr{M}_2$, and every line through P is either in $\mathscr{M}_1$ or $\mathscr{M}_2$. If $\mathscr{M}_1 \cap \mathscr{M}_2$ is empty, then G_P permutes the lines through P in two orbits. If $\mathscr{M}_1 \cap \mathscr{M}_2$ is nonempty, then G_P is transitive on the lines through P.

THEOREM 1. *Let $\mathscr{P}$ be a finite projective plane and G a rank 3 group of collineations of $\mathscr{P}$. Then G_P is flag-transitive.*

Proof. Assume G is not flag-transitive. For every point P of $\mathscr{P}$ G_P has 3 point orbits in $\mathscr{P}$. Hence by Result II G_P has three line orbits in $\mathscr{P}$. If all the lines through P are in a single orbit under G_P, then G is flag-transitive contrary to our assumption. Thus by Lemma 1, two of these line orbits consist of the lines through P. Thus the third line orbit of G_P must consist of all the lines of $\mathscr{P}$ which do not go through P. Hence for every point P, G_P is transitive on all lines not through P.

Let (P, l) be a non-incident point-line pair of $\mathscr{P}$ and (Q, m) another non-incident point-line pair. There exists a collineation $\sigma \in G$ such that $P\sigma = Q$. Let $\bar{l}$ be the image of l under σ. $Q \notin \bar{l}$ since $P \notin l$. Then there exists a collineation $\tau \in G_Q$ such that $\bar{l}\tau = m$. Then $P\sigma\tau = Q\tau = Q$ and $l\sigma\tau = \bar{l}\tau = m$. This proves G is transitive on non-incident point-line pairs.

Result I implies that G is doubly transitive on the points of $\mathscr{P}$. This is a contradiction since G has rank 3 on the points of $\mathscr{P}$. Thus G is flag-transitive.

LEMMA 2. *Let $\mathscr{P}$ be a finite projective plane of order n and G a rank 3 group of collineations of $\mathscr{P}$, and let P a point of $\mathscr{P}$.*

(i) *G_P permutes the points not equal to P in two orbits $\mathscr{O}_1$ and $\mathscr{O}_2$ of lengths $k_1(n + 1)$ and $k_2(n + 1)$ respectively, where $k_1 + k_2 = n$.*

(ii) *G_P permutes the lines not through P in two orbits $\mathscr{L}_1$* and *$\mathscr{L}_2$ of lengths s_1 and s_2 respectively, with $s_1 + s_2 = n^2$.*

Proof. G_P is transitive on the lines through P. Thus every line through P intersects the non-trivial point orbits $\mathscr{O}_1$ in the same number of points. Thus $|\mathscr{O}_1| = k_1(n+1)$ where k_1 is the number of points of $\mathscr{O}_1$ on a line l through P. Similarly $|\mathscr{O}_2| = k_2(n+1)$, k_2 the number of points of $\mathscr{O}_2$ on a line l through P. This proves (i) since for every line l through P, a point $\neq P$ on l is either in $\mathscr{O}_1$ or $\mathscr{O}_2$.

G_P has three point orbits. Hence it has three line orbits (Result II). One of these consists of the lines through P. The other two line orbits are made up of the lines not through P and there are n^2 such lines. This gives (ii).

LEMMA 3. *Let $\mathscr{P}$ be a finite projective plane of order n and G a rank* 3 *group of collineations of $\mathscr{P}$. If G is solvable, then*

(i) *$n^2 + n + 1$ is a prime;*

(ii) *G acts as a Frobenius group on $\mathscr{P}$;*

(iii) *n is even, $|G| = 1/2(n^2+n+1)n(n+1)$, and $|G_P| = 1/2n(n+1)$.*

If $|G|$ is odd, then we also have

(iv) *$n = 2m$, m odd.*

Proof. Since G is flag-transitive, G is primitive on the points of $\mathscr{P}$ (Dembowski [3], p. 212). By Result VII $n^2 + n + 1$ is a prime. Result VIII implies that G is either regular on $\mathscr{P}$ or it is a Frobenius group on $\mathscr{P}$. Since G is clearly not regular it must be a Frobenius group on $\mathscr{P}$. This proves (i) and (ii).

By Lemma 2 we have for every point $P \in \mathscr{P}$

$$|G_P| = k_1(n+1)\,|G_{P,Q}| = k_2(n+1)\,|G_{P,R}|\,,$$

where $Q \in \mathscr{O}_1$ and $R \in \mathscr{O}_2$, and $k_1 + k_2 = n$. But G a Frobenius group on $\mathscr{P}$ implies $|G_{P,Q}| = 1 = |G_{P,R}|$. Hence $|G_P| = k_1(n+1) = k_2(n+1)$ and thus $k_1 = k_2 = n/2$. This implies n is even since k_1 is an integer, and we have $|G| = 1/2(n^2+n+1)(n(n+1))$. This proves (iii).

If $|G|$ is odd, then $n/2$ is odd and this proves (iv).

LEMMA 4. *Let $\mathscr{P}$ be a finite projective plane of order n with a rank* 3 *group of collineations. If $n > 2$, then $|G|$ is even.*

Proof. Assume $|G|$ is odd. Then G is solvable (by the Feit-Thompson theorem) and the previous lemma implies $n^2 + n + 1$ is a prime, $n = 2m$ with m odd, and $|G| = (n^2+n+1)m(n+1)$. Also for each point P, G_P has two line orbits $\mathscr{L}_1$ and $\mathscr{L}_2$ of lengths s_1 and s_2 respectively with

$$(1) \qquad s_1 + s_2 = n^2 = 4m^2$$

(Lemma 2). Since $|G_P| = m(n+1) = m(2m+1)$, we have

$$(2) \qquad s_i \mid m(2m+1)\,, \qquad i = 1, 2.$$

Let $s = \text{g.c.d.}\,(s_1, s_2)$ and define integers t_1 and t_2 by

$$(3) \qquad s_i = st_i\ , \qquad i = 1, 2\ .$$

Then $\text{g.c.d.}\,(t_1, t_2) = 1$ and $s(t_1 + t_2) = 4m^2$ (by (1)). Since s is odd (for otherwise $|G_P|$ is even), $s \mid m^2$. Since $s \mid s_1$ (2) implies $s \mid m$ since $\text{g.c.d.}\,(m, 2m + 1)$. Hence

$$(4) \qquad m = su$$

for some integer u, and

$$(5) \qquad t_1 + t_2 = 4su^2\ .$$

If $v = \text{g.c.d.}\,(t_1, \text{s}) > 1$, then $v \mid t_2 - a$ contradiction to the fact that $\text{g.c.d.}\,(t_1, t_2) = 1$. If $w = \text{g.c.d.}\,(t_1, u) > 1$, then $w \mid t_2$- again a contradiction. Hence $1 = \text{g.c.d.}\,(t_1, s) = \text{g.c.d.}\,(t_1, u)$, which implies $\text{g.c.d.}\,(t_1, m) = 1$. Similarly $\text{g.c.d.}\,(t_2, m) = 1$ and thus

$$(6) \qquad t_i \mid 2m + 1\ , \qquad i = 1, 2\ .$$

Then we have, using (5)

$$2(2m + 1) \geqq t_1 + t_2 = 4su^2\ .$$

Applying (4) we get

$$1 \geqq 2su(u - 1)\ .$$

Therefore $u = 1$ and

$$(7) \qquad m = s\ .$$

But then

$$(8) \qquad t_1 + t_2 = 4m\ .$$

If $t_1 = t_2$, then $t_1 = 2m$ and this contradicts (6). Thus without loss of generality we may assume $t_1 < t_2$. Then from (8) we get $t_1 < 2m$, $t_2 > 2m$ and this implies $t_2 = 2m + 1$ (by (6)) and $t_1 = 2m - 1$ (by (8)). Hence $2m - 1 \mid 2m + 1$, which implies $m = 1, n = 2$. This proves the lemma.

REMARK. The example at the end of §1 shows that if $n = 2$ $|G|$ can be odd.

By combining Lemma 3 and Lemma 4 we can show that the rank 3 group G is nonsolvable if $n > 2$:

THEOREM 2. *Let $\mathscr{P}$ be a finite projective plane of order n with a rank 3 group G of collineations. If $n > 2$, then G is nonsolvable.*

Proof. Assume G is solvable and $n > 2$. By Lemma 3(ii) G acts as a Frobenius group on $\mathscr{P}$. Lemma 4 implies $|G|$ is even. Hence G has an element σ of order 2. Either σ is a central collineation or it fixes a subplane of $\mathscr{P}$ pointwise (Baer [1]). In both cases σ fixes more than two points. But $G_{P,Q}$ consists only of the identity for every two distinct points P, Q of $\mathscr{P}$. This gives a contradiction. Therefore G is nonsolvable if $n > 2$.

Our last result in this section shows that $\mathscr{P}$ is desarguesian only when $n = 2$.

THEOREM 3. *Let $\mathscr{P}$ be a finite projective plane with a rank 3 group G of collineations. $\mathscr{P}$ is desarguesian if and only if $n = 2$.*

Proof. Assume $\mathscr{P}$ is desarguesian. By Theorem 1 G is flag-transitive. G cannot contain all the elations of $\mathscr{P}$. For the group H generated by the elations of $\mathscr{P}$ is doubly transitive on the points of $\mathscr{P}$. H a subgroup of G implies G is doubly transitive on the points of $\mathscr{P}$-again contradicting the fact that G has rank 3 on the points of $\mathscr{P}$. By Result IV either $n = 2$ and G has order 21, or $n = 8$ and G has order 657. But the second case cannot occur since $n > 2$ implies $|G|$ is even (Lemma 4). Thus $n = 2$ and G has order 21.

Conversely if $n = 2$, then $\mathscr{P}$ is desarguesian, and the example at the end of §1 shows that in this case a rank 3 group does occur.

4. Proof of the main theorem.

We now prove the main theorem stated in §1. Assume n is odd. If n is not a fourth power, then Theorem 1 and Result III implies that $\mathscr{P}$ is desarguesian and G contains all elations of $\mathscr{P}$. But by Theorem 3 this is impossible.

Assume n is even and $n > 2$. Lemma 4 $|G|$ is even. If n is not a square, then an element in G of order 2 must be an elation (Baer [1]) and $\mathscr{P}$ is desarguesian by Result V. This contradicts Theorem 3. Hence n is a square. If $n = m^2$ with $m \equiv 2 \pmod 4$, then $\mathscr{P}$ is desarguesian by Result VI since G is nonsolvable (Corollary 2.1). This contradicts Theorem 3 again. The proof of the main theorem is complete.

REFERENCES

1. R. Baer, *Projectivities with fixed points on every line of the plane*, Bull. Amer. Math. Soc., **52** (1946), 273-286.

2. P. Dembowski, *Verallgemeinerungen von Transitivitätsklassen endlicher projektiver Ebenen*, Math. Z., **69** (1958), 59-89.

3. ———, *Finite Geometries*, Springer-Verlag, Berlin-Heidelberg, 1968.

4. D. G. Higman and J. E. McLaughlin, *Geometric ABA-groups*, Ill. J. Math., **5** (1961), 382-397.

5. D. R. Hughes, *Collineations and generalized incidence matrices*, Trans. Amer. Math. Soc., **86** (1957), 284-296.

6. V. H. Keiser, *Finite projective planes with non-solvable transitive collineation groups*, Amer. Math. Monthly, **74** (1967), 556-559.

7. T. G. Ostrom, *Dual transitivity in finite projective planes*, Proc. Amer. Math. Soc., **9** (1958), 55-56.

8. T. G. Ostrom and A. Wagner, *On projective and affine planes with transitive collineation groups*, Math. Z., **71** (1959), 186-199.

9. E. T. Parker, *On collineations of symmetric designs*, Proc. Amer. Math., Soc., **8** (1957), 350-351.

10. A. Wagner, *On Perspectivities of finite projective planes*, Math. Z., **71** (1959), 113-123.

11. H. Wielandt, *Finite permutation groups*, Acedemic Press, New York and London, 1964.

Received January 12, 1971. This work was partially supported by NSF Contract No. GP. 17461.

WASHINGTON STATE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

ON SOLVABLE O*-GROUPS

D. P. MINASSIAN

The purpose of this paper is to prove the existence of O*-groups of arbitrary solvable length, as well as of nonsolvable O*-groups.

By a *partial order* for a group G we mean a reflexive, antisymmetric and transitive relation, $\leqq$, on G such that if g and h are elements of G and $g \leqq h$, then $xgy \leqq xhy$ for all x and y in G. If also any two elements g and h of G are *comparable* (i.e., either $g \leqq h$ or $h \leqq g$), then the partial order for G is called a *total* (or *full*, or *linear*) *order*. The group G is an *O*-group* if any partial order for G is included in some total order for G.

A group G is *solvable of length* n, where n is a positive integer, if the derived chain of G reaches the unit subgroup, E, in n steps:

$$G = G^1 \supsetneqq G^2 \supsetneqq \cdots \supsetneqq G^n \supsetneqq G^{n+1} = E,$$

where G^{i+1} is the derived group of G^i (denoted below by $G^{i+1} = [G^i, G^i]$).

It has been shown that non-abelian free groups are not O*-groups ([1], [2], [3], [4], [6]). Further, Kargapolov [5] and Kargapolov, Kokorin and Kopytov [6] have produced solvable groups which are not O*-groups even though they admit a full order: these are the free r-step solvable groups on k generators for $r \geqq 3$ and $k \geqq 2$. In view of these results one may ask if there exist solvable O*-groups of arbitrary length, and nonsolvable O*-groups. The answers are affirmative.

THEOREM. *For every positive integer m there exists an O^*-group G that is solvable of length m.*

Proof. Let F be the free group on k generators for some fixed $k \geqq 2$. Let F_i be the ith term in the lower central series for F, where $F_1 = F$, and let F^i be the ith derived group for F, where $F^1 = F$. Consider F/F_i, the free nilpotent group of class i with k generators. By varying i we shall obtain the desired groups G of the theorem.

We first claim that F/F_i is torsion-free for every positive integer i. If not, then for some i there exists an element $a \in F$ and a positive integer p such that $a \notin F_i$, but $a^p \in F_i$. Now $a \in F_h - F_{h+1}$ for some positive integer $h \leqq i-1$. Thus $a^p \in F_i \subseteqq F_{h+1}$, and so F_h/F_{h+1} is not torsion-free. On the other hand, Witt's theorem (see, e.g., [8, p. 41]) states that F_h/F_{h+1} is a free abelian group (and hence torsionfree), a contradiction. Thus F/F_i is torsion-free, as claimed.

Malcev [9] has shown that a torsion-free nilpotent group is an O^*-group. Hence F/F_i is an O^*-group for every positive integer i. Now for every such i the solvable length of F/F_i is finite, since F/F_i is nilpotent. Thus we shall complete our proof by establishing the following lemma.

LEMMA. *For every positive integer m, there exists an integer n such that solvable length of F/F_n is m.*

Proof. We first note that for every positive integer i, there exists an integer j such that $F_j \not\supseteq F^i$. This follows from the fact that $F^i \neq E$ for each i (hence F is not solvable), together with the theorem of Magnus (cf. [8, p. 38]) which asserts that $\bigcap_{i=1}^{\infty} F_i = E$. We next show that for each i and j,

$$(F/F_j)^i = F^i F_j/F_j. \tag{1}$$

Indeed, it is readily seen that if A and B are subgroups of a group G and if B is invariant under conjugation by elements of A, then $(AB/B)^2 = A^2B/B$. From this, an induction on i shows that for a normal subgroup N of a group G it is true that $(G/N)^i = G^iN/N$ for all i, which implies the desired result.

Note that for each i there exists J such that for $j \geqq J$, the solvable length of F/F_j exceeds i. This follows from (1) and the fact that, by our first assertion, we can choose J such that $F_J \not\supseteq F^i$. In particular, then, the solvable length of F/F_j is unbounded with increasing j. Note also that the solvable length of F/F_{j+1} exceeds the solvable length of F/F_j by at most 1. For if F/F_j is solvable of length $r-1$, then $(F/F_j)^r = E$. Thus, by (1) we have $F^rF_j/F_j = E$, which implies $F^r \subseteqq F_j$. On the other hand, F/F_{j+1} has solvable length $\leqq r$ since (again using (1))

$$\begin{aligned}(F/F_{j+1})^{r+1} &= [(F/F_{j+1})^r, (F/F_{j+1})^r] \\ &= [F^rF_{j+1}/F_{j+1}, F^rF_{j+1}/F_{j+1}] \subseteqq [F_j/F_{j+1}, F_j/F_{j+1}] = E,\end{aligned}$$

where $\subseteqq$ holds since both F^r and F_{j+1} are subsets of F_j, and the final equality derives from the fact that F_j/F_{j+1} is abelian by Witt's theorem (above). The lemma follows at once from these results and the fact that $F/F_2 = F_1/F_2$ has solvable length 1 by Witt's theorem.

The proof of the theorem is now complete.

COROLLARY. *There exist nonsolvable O^*-groups.*

Proof. Kargapolov [5] and Kokorin [7] have shown that the restricted direct product of O*-groups is an O*-group. Thus the restricted direct product, $G = \prod_{i=1}^{\infty} F/F_i$, of the groups F/F_i is an O*-group. If G were solvable of length m, then each F/F_i would have solvable length $\leqq m$; for if a subgroup $H \subseteqq G$, then $H^k \subseteqq G^k$ for every k. Since this contradicts the fact noted above that the solvable length of F/F_j is unbounded with increasing j, G is a non-solvable O*-group.

Note. The mapping φ of F into the *unrestricted* (or *complete*) direct product, $\prod_{i=1}^{\infty} F/F_i$, of the groups F/F_i given by

$$\varphi(a) = (aF_1, \cdots, aF_n, \cdots) \text{ for every } a \in F$$

is a monomorphism by Magnus' theorem, above. Since F is not an O*-group (see [1], [4], or [6]), we have an immediate example of a subdirect product of O*-groups which is not itself an O*-group. (In [5], Kargapolov uses some of the groups F/F_i to show that the class of O*-groups is not closed under formation of unrestricted direct products.)

Acknowledgement. I acknowledge the generously donated help of Professor E. F. Krause (University of Michigan), my adviser for the doctoral dissertation on which this paper is based.

REFERENCES

1. L. Fuchs, *On orderable groups*, Proc. Internat. Conf. Theory of Groups (Canberra, 1965), 89-98, Gordon and Breach, New York, 1967.
2. L. Fuchs and E. Sasiada, *Note on orderable groups*, Annales Universitatis Scientiarum Budapestinensis de Rolando Eotvos Nominatae, Sectio Mathematica (1964), 13-17.
3. W. C. Holland, *Extensions of Ordered Algebraic Structures*, Doctoral Thesis, Tulane Univerity, 1961.
4. H. Hollister, *Groups in which every maximal partial order is isolated*, Proc. Amer. Math. Soc., **19** (1968), 467-469.
5. M. I. Kargapolov, *Fully orderable groups* (Russian), Algebra i Logika Sem., (6) **2** (1963), 5-14.
6. M. I. Kargapolov, A. I. Kokorin, and V. M. Kopytov, *On the theory of orderable groups* (Russian), Algebra i Logika Sem., (6) **4** (1965), 21-27.
7. A. I. Kokorin, *Ordering a direct product of ordered groups* (Russian), Ural. Gos. Univ. Mat. Zap., (3) **4** (1963), 95-96.
8. A. G. Kurosh. *The Theory of Groups*, vol. ii, Chelsea, New York, 1956.
9. A. I. Malcev, *On the completion of group order* (Russian), Trudy Mat. Inst. Steklov, **38** (1951), 173-175.

Received August 25, 1970 and in revised form April 12, 1971.

BUTLER UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

GENERATORS OF THE MAXIMAL IDEALS OF $A(\bar{D})$

NILS ØVRELID

Let $A = A(\bar{D})$ be the sup norm algebra of functions continuous in $\bar{D}$ and holomorphic in D, where D is a bounded, strictly pseudoconvex domain in C^n. This paper gives necessary and sufficient local conditions that a subfamily of A generates the maximal ideal $\mathscr{M}_w(\bar{D})$ of functions in A vanishing at $w \in \bar{D}$. In particular, it shows that $\mathscr{M}_w(\bar{D})$ is generated by $z_1 - w_1, \cdots, z_n - w_n$ when $W \in D$.

In [3], Gleason shows that if m is an (algebraically) finitely generated maximal ideal of a commutative Banach algebra A, the maximal ideal space $\mathscr{M}_A$ can be given an analytic structure near m, in terms of which the Gelfand transforms of the elements of A are holomorphic functions.

In a sense, the results of this paper go in the opposite direction. We consider a bounded domain D in C^n, with C^2 strictly pseudoconvex boundary, and study the algebra $A = A(\bar{D})$ of functions continuous on $\bar{D}$ and holomorphic in D. By a recent result, Henkin [4], Kerzman [7], Lieb [9], A equals the closure in $C(\bar{D})$ of the algebra $O(\bar{D})$ of functions holomorphic in some neighbourhood of $\bar{D}$, from which it follows that $\mathscr{M}_A \approx \bar{D}$.

We first fix the notation. If $w \in \bar{D}$, $\mathscr{M}_w$ denotes the maximal ideal of the ring O_w of germs of holomorphic functions at w, while $\mathscr{M}_w(\bar{D})$ is the maximal ideal in A of functions vanishing at w. If f is a function on some neighbourhood of w, f_w denotes the germ of f at w.

THEOREM 1. *Let $w \in D$, and $f_1, \cdots, f_N \in A$. Then $f_1, \cdots, f_N$ generate $\mathscr{M}_w(\bar{D})$ if and only if*

(1) *${f_1}_w, \cdots, {f_N}_w$ generate $\mathscr{M}_w$, and*

(2) *w is the only common zero of $f_1, \cdots, f_N$ in $\bar{D}$.*

COROLLARY. *If $w \in D$, $z_1 - w_1, \cdots, z_n - w_n$ generate $\mathscr{M}_w(\bar{D})$.*

Below we give the more general theorem 2, which also gives a similar characterization of generators of $\mathscr{M}_w(\bar{D})$ when $w \in \partial D$. When $n = 2$, Kerzman and Nagel [8] have shown that $z_1 - w_1$ and $z_2 - w_2$ generate $\mathscr{M}_w(\bar{D})$ when $w \in D$, as well as similar results for algebras with Hölder norms. I want to thank Dr. Kerzman for sending me a copy of his thesis [7], where these results are stated.

The main tool in the proof is the following result, which is proved in [11]:

LEMMA 1. *Suppose* $u \in C^\infty_{(0,q)}(D)$ *is bounded, with* $\bar\partial u = 0$, $q \geqq 1$. *Then there exists a* $v \in C^\infty_{(0,q-1)}(D)$ *with* $\bar\partial v = u$, *such that* v *has a continuous extension to* $\bar D$.

A closely related result is given in Lieb [10], while a stronger result for (0, 1)-forms, involving Hölder estimates, is given in Kerzman [7].

It is convenient to prove first a more general result. If U is open in $\bar D$, let $H(U)$ denote functions in $C(U)$ that are holomorphic in $D \cap U$. When $w \in \bar D$, we define $H_w = \varinjlim_{U \ni w} H(U)$, so H_w is the space of germs at w of continuous functions on $\bar D$ that are holomorphic in D. It is easy to see that H is the sheaf of A-holomorphic functions in the sense of [2].

PROPOSITION 1. *Let* D *be as above,* $w \in \bar D$, *and suppose* $f_1, \cdots, f_N$ *have* w *as their only common zero. We let* I *denote the ideal in* A *generated by* $f_1, \cdots, f_N$, *and* I_w *the ideal in* H_w *generated by* $f_{1w}, \cdots, f_{Nw}$. *If* $f \in A$ *and* $f_w \in I_w$, *then* $f \in I$.

Proof. By assumption, we may write $f = \sum_{i=1}^N g_i \cdot f_i$ on a neighbourhood U of w in $\bar D$, with $g_1, \cdots, g_N \in H(U)$. We want to write $f = \sum_{i=1}^N h_i \cdot f_i$, with $h_1, \cdots, h_N \in A$, and shall first solve the problem differentiably. As the sets $N_i = \{z \in \bar D \backslash \{w\} : f_i(z) = 0\}$, $i = 1, \cdots, N$, are closed in $C^n \backslash \{w\}$, it is well known how to construct $\tilde\varphi_1, \cdots, \tilde\varphi_N$ with $\tilde\varphi_i = 0$ on a neighbourhood of N_i, $i = 1, \cdots, N$, that form a C^∞ partition of unity on $C^n \backslash \{w\}$. Choose $\varphi_0 \in C_0^\infty(U')$, where $U' \cap \bar D = U$, with $\varphi_0 = 1$ on a neighbourhood U_1 of w, and define $\varphi_i = (1 - \varphi_0) \cdot \tilde\varphi_i$, $i = 1, \cdots, N$.

If we define

$$g'_i = \varphi_0 \cdot g_i + \frac{\varphi_i \cdot f}{f_i}, \text{ clearly } \sum_{i=1}^N g'_i \cdot f_i = f \text{ on } \bar D.$$

The $g'_i s \in C^\infty(D) \cap C(\bar D)$, and are holomorphic in $U_1 \cap D$.

We want to use Lemma 1 to modify the $g'_i s$ to get h_i's in A. To handle the combinatorial difficulties, we apply the homological argument of [6].

NOTATION. $L_r = \{u \in C^\infty_{(0,r)}(D)$, u and $\bar\partial u$ have bounded coefficients$\}$, while $L_r^s = L_r \otimes_C \bigwedge^s C^N$, $0 \leqq r, s$.

If we choose a basis $e_1, \cdots, e_N$ in C^N, the elements in L_r^s may be written uniquely as $\sum_{|I|=s} u_I \otimes e^I$, where $u_I \in L_r$, $e^I = e_{i_1} \wedge \cdots \wedge e_{i_s}$, and we sum over strictly increasing sequences $I = (i_1, \cdots, i_s)$. We define $\bar\partial$ on L_r^s by $\bar\partial(u \otimes \omega) = (\bar\partial u) \otimes \omega$ and linearity. Clearly

$\bar{\partial} L_r^s \subset L_{r+1}^s$, and lemma 1 gives:

LEMMA 1′. *If* $k \in L_r^s$ *and* $\bar{\partial} k = 0$, $r \geqq 1$, *there exists a* $k' \in L_{r-1}^s$, such that $\bar{\partial} k' = k$, *and* k' *has a continuous extension to* $\bar{D}$.

The product determined by $(u \otimes \omega) \cdot (u' \otimes \omega') = (u \wedge u') \otimes (\omega \wedge \omega')$ is clearly a bilinear map $L_r^s \times L_{r'}^{s'} \to L_{r+r'}^{s+s'}$.

Let $e_1^*, \cdots, e_N^*$ be the reciprocal basis to $e_1, \cdots, e_N$, so $\langle e_i^*, e_j \rangle = \delta_{ij}$. We define $P_f \colon L_r^s \to L_r^{s-1}$ by

$$P_f(d \otimes \omega) = \sum_{i=1}^{N} (f_i \cdot u) \otimes (e_i^* \lrcorner\, \omega), \text{ and linearity.}$$

(For the definition of $\lrcorner$, se [12] Ch. 1.)
$P_f \colon L_r^1 \to L_r^0$ maps $\sum_{i=1}^N u_i \otimes e_i$ to $\sum_{i=1}^N f_i \cdot u_i$; in particular, $P_f g' = f$, when $g' = \sum_{i=1}^N g_i' \otimes u_i$.

A simple computation gives $P_f^2 = 0$, while the derivation property of $\lrcorner$ gives

$$\text{(i)} \qquad P_f(k \cdot k') = (P_f k) \cdot k' + (-1)^s k \cdot P_f k'$$

when $k \in L_r^s$.

Let $M_r^s = \{k \in L_r^s \colon k|_{U_1} = 0\}$.

LEMMA 2. *The complex* $0 \leftarrow M_r^0 \xrightarrow{Pf} M_r^1 \xrightarrow{Pf} \cdots \xrightarrow{Pf} M_r^N \leftarrow 0$ *is exact.*

Proof. Let $\varphi \in C^\infty(C^N)$ be zero near w and one outside U_1. We put $k_0 = \sum_{i=1}^N (\varphi \cdot \tilde{\varphi}_i)/f_i \otimes e_i$. Clearly $k_0 \in L_0^1$, and $P_f k_0 \in L_0^0$ is identically one in $D \backslash U_1$. If $k \in M_r^s$ and $P_f k = 0$, $k_0 \cdot k \in M_r^{s+1}$, and by (*i*), $P_f(k_0 \cdot k) = (P_f k_0) \cdot k = k$.

As $f_1, \cdots, f_N$ are holomorphic in D, P_f and $\bar{\partial}$ commute.

LEMMA 3. *If* $k \in M_r^s$ *and* $P_f k = \bar{\partial} k = 0$, *there exists a* $k' \in L_r^{s+1}$, *with* $P_f k' = k$ *and* $\bar{\partial} k' = 0$.

This is trivially true when $r > n$, and the proof goes by downward induction on r. Suppose the lemma is valid for $r + 1$. By Lemma 2, there exists a $k_1 \in M_r^{s+1}$ with $P_f k_1 = k$. Clearly $\bar{\partial} M_r^{s+1} \subset M_{r+1}^{s+1}$, while $P_f \bar{\partial} k_1 = \bar{\partial} P_f k_1 = 0$. Using the induction hypothesis, we can find $k_2 \in L_{r+1}^{s+2}$ with $P_f k_2 = \bar{\partial} k_1$ and $\bar{\partial} k_2 = 0$. By Lemma 1′, $k_2 = \bar{\partial} k_3$, with $k_3 \in L_r^{s+2}$. If we put $k' = k_1 - P_f k_3$, we get $k' \in L_r^{s+1}$, with $\bar{\partial} k' = \bar{\partial} k_1 - P_f \bar{\partial} k_3 = 0$, and $P_f k' = P_f k_1 - P_f^2 k_3 = k$. This completes the induction step.

Proof of Proposition 1. As the g_i's are holomorphic in $U_1 \cap D$, $\bar\partial g' \in M_1^1$. Applying Lemma 1′ and Lemma 3, we find a $k \in L_0^2$, with $\bar\partial P_f k = P_f \bar\partial k = \bar\partial g'$, such that k is continuous on $\bar D$. If $h = g' - P_f k$, $\bar\partial h = 0$. Writing $h = \sum_{i=1}^N h_i \otimes e_i$, this means that $h_1, \cdots, h_N \in A$, and $\sum_{i=1}^N h_i \cdot f_i = f$.

THEOREM 2. *Let $w \in \bar D$, and let M_w denote the unique maximal ideal of H_w. The family $(f_i)_{i \in I}$ in A generates $\mathscr{M}_w(\bar D)$ if and only if*

(1) *$(f_{i_w})_{i \in I}$ generates M_w, and*

(2) *w is the only common zero of functions f_i in $\bar D$*

Proof. I. The sufficiency of (1) and (2): If $f \in \mathscr{M}_w(\bar D)$, we have $f_w \in M_w$, and by (1) f_w belongs to some ideal $[f_{i_1,w}, \cdots, f_{i_{M,w}}]$. As $(z_1 - w_1)_w, \cdots, f(z_n - w_n)_w$ belong to M_w, the functions $z_i - w_i$; $i = 1, \cdots, n$, may be expressed as linear combinations of functions $f_{i_{M+1}}, \cdots, f_{i_P}$ in the family on some open neighbourhood V of w in $\bar D$. Then $f_{i_{M+1}}, \cdots, f_{i_P}$ have w as their only common zero in V. By condition (2) and the compactness of $\bar D \backslash V$, there exist $f_{i_{P+1}}, \cdots, f_{i_N}$ in the family with no common zeroes outside V. Now proposition 1 implies that $f \in [f_{i_1}, \cdots, f_{i_N}]$.

II. The necessity of (1) and (2): If $(f_i)_{i \in I}$ generate $\mathscr{M}_w(\bar D)$, condition (2) follows from the fact that A separates points in $\bar D$. Condition (1) follows from

PROPOSITION 2. *The germs at w of elements in $\mathscr{M}_w(\bar D)$ generate M_w.*

The following proof of Proposition 2 was kindly communicated to me by Dr. R. M. Range, and replaces a more complicated argument of my own:

When $w \in D$, $z_1 - w_1, \cdots, z_n - w_n$ generate $\mathscr{M}_w = M_w$. Thus we may assume $w \in \partial D$, and consider an $f \in H(U \cap \bar D)$ with $f(w) = 0$, where U is some neighbourhood of w in C^n. We choose $\varphi \in C_0^\infty(U)$ such that $\varphi \equiv 1$ on a smaller neighbourhood V of w. As D is strictly pseudoconvex, we may extend it inside V to a strictly pseudoconvex domain D' containing w. As $\bar\partial(\varphi \cdot f)$ vanishes on $V \cap D$, it may be extended by zero to a smooth, bounded, $\bar\partial$-closed $(0, 1)$-form ω on D'. By Lemma 1, the equation $\bar\partial g = \omega$ has a solution in $C^\infty(D') \cap C(\bar D')$, and we may assume $g(w) = 0$. As g is holomorphic in $D' \cap V$, we may write it near w as $g = \sum_{i=1}^n g_i(z_i - w_i)$, with $g_1, \cdots, g_n$ holomorphic. Thus $f_w = (\varphi \cdot f - g)_w + \sum_{i=1}^n g_{i_w}(z_i - w_i)_w$, and $\varphi \cdot f - g|_{\bar D} \in \mathscr{M}_w(\bar D)$.

When $w \in D$ and I is finite, Theorem 2 reduces to theorem 1. If $w \in \partial D$, it follows from Gleason's result that $\mathscr{M}_w(\bar{D})$ is not finitely generated. If M_w were finitely generated, it would by Proposition 2 be generated by finitely many elements of A, which implies by the argument of I that $\mathscr{M}_w(\bar{D})$ must be finitely generated. Thus M_w is not finitely generated when $w \in \partial D$. (This may also be proved in a more direct fashion).

Note. The Corollary to Theorem 1 has also been proved by G. M. Henkin in Bull. Acad. Polon. Sci., **24** (1971) 37–42, and by I. Lieb in Math. Ann., **190** (1970–71) 6–44, which contains a detailed version of [10].

References

1. J. Dieudonné, *Foundations of Modern Analysis*, Academic Press 1960.

2. T. W. Gamelin, *Uniform Algebras*, Prentice Hall 1969.

3. A. Gleason, *Finitely generated ideals in Banach algebras*, J. Math. Mech., **13** (1964), 125–32.

4. G. M. Henkin, *Integral representations of functions holomorphic in strictly pseudoconvex domains and some applications*, Mat. Sbornik, **78** (1969), 611–32. (Math. USSR Sbornik Vol. **7** (1969).)

5. L. Hörmander, *Introduction to Complex Analysis in Several Variables*, van Nostrand 1966.

6. ———, *Generators of some rings of analytic functions*, Bull. Amer. Math. Soc., **73** (1967), 943–49.

7. N. Kerzman, *Hölder and L^p-estimates for solutions of $\bar{\partial}u = f$ in strongly pseudoconvex domains*, Comm. Pure. Appl. Math. **24** (1971) 301–80.

8. N. Kerzman and A. Nagel, *Finitely generated ideals in certain function algebras* Journal of Functional Analysis. **7** (1971) 212–15.

9. I. Lieb, *Ein approximationssatz auf streng pseudoconvexen gebieten*, Math. Ann., **184** (1969), 56–60.

10. ———, *Beschränktkeitsaussagen für den d''-operator*, Nachr. Ak. Wiss. Gött. II. Math. Phys. Kl. 1970, 1, 7 p.

11. N. Øvrelid, *Integral representation formulas and L^p-estimates for the $\bar{\partial}$-equation*, (to appear in Math. Scand.).

12. S. Sternberg, *Lectures on Differential Geometry*, Prentice Hall 1964.

Received August 6, 1970. Dr. Range's proof received June 28, 1971.

UNIVERSITETET I OSLO

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

A SEMILATTICE DECOMPOSITION INTO SEMIGROUPS HAVING AT MOST ONE IDEMPOTENT

MOHAN S. PUTCHA AND JULIAN WEISSGLASS

A semigroup S is said to be *viable* if $ab = ba$ whenever ab and ba are idempotents. The main theorem of this article proves in part that S is a viable semigroup if and only if S is a semi-lattice of $\mathscr{S}$-indecomposable semigroups having at most one idempotent.

Furthermore, each semigroup appearing in the decomposition has a group ideal whenever it has an idempotent. Also included as part of the main theorem is the more elementary result that S is viable if and only if every $\mathscr{J}$-class contains at most one idempotent.

Throughout S will denote a semigroup and $E = E(S)$ the set of idemotents of S.

DEFINITION. Let $a, b \in S$. We say $a | b$ if there exist $x, y \in S$ such that $ax = ya = b$. The set-valued function $\mathfrak{M}$ on S is defined by $\mathfrak{M}(a) = \{e | e \in E, a | e\}$. The relation δ on S is defined by $a\ \delta\ b$ if $\mathfrak{M}(a) = \mathfrak{M}(b)$.

Our first goal is to show that if S is viable then δ is a congruence on S and S/δ is the semilattice described above.

LEMMA 1. *Let S be viable. If $ab = e \in E$, then $bea = e$.*

Proof. $(bea)^2 = beabea = bea$. Hence $bea \in E$. But cleary $abe = e \in E$. Hence $bea = abe = e$.

LEMMA 2. *Let S be viable. Suppose $a \in S$ and $e \in E$. Then $a | e$ if and only if $e \in S^1aS^1$.*

Proof. If $a | e$, then $e \in S^1aS^1$ by definition. Conversely assume $e = sat$ with $s, t \in S^1$. By (1), $ates = e$ and $tesa = e$. Therefore $a | e$.

THEOREM 3. *Let S be viable. Then*

(i) *δ is a congruence relation on S containing Green's relation $\mathscr{H}$.*

(ii) *S/δ* is a semilattice and

(iii) *each δ-class contains at most one idempotent and a group ideal whenever it contains an idempotent.*

Proof. (i) Clearly δ is an equivalence relation. We will show that δ is right compatible. Assume $a\ \delta\ b$. If $ac | e \in E$, then

$acx = e$ for some $x \in S$. By (1), $cxea = e$. Hence $a|e$. Thus $b|e$, so $yb = e$ for some $y \in S$. Therefore $ybcxea = e$, so $bc|e$ by (2). Hence $\mathfrak{M}(ac) \subseteqq \mathfrak{M}(bc)$. Similary $\mathfrak{M}(bc) \subseteqq \mathfrak{M}(ac)$ and hence $ac\ \delta\ bc$. That δ is left compatible follows analogously. Consequently, δ is a congruence. It is immediate that $\mathscr{H} \subseteqq \delta$.

(ii) To show S/δ is a band, let $a \in S$. If $a^2|e \in E$ then by (2), $a|e$. Hence $\mathfrak{M}(a^2) \subseteqq \mathfrak{M}(a)$. Suppose $a|e \in E$, say $ax = ya = e$, $x, y \in S$. Then $ya^2x = e$. Again using (2), $a^2|e$. Thus, $\mathfrak{M}(a^2) = \mathfrak{M}(a)$ and $a\ \delta\ a^2$. So S/δ is a band. Now let $a, b \in S$. If $e \in \mathfrak{M}(ab)$, then there exist $x, y \in S$ such that $abx = yab = e$. Hence $ya(ba)bx = e$, and by (2), $e \in \mathfrak{M}(ba)$. Therefore $\mathfrak{M}(ab) \subseteqq \mathfrak{M}(ba)$. By symmetry, $\mathfrak{M}(ba) \subseteqq \mathfrak{M}(ab)$. Hence $ab\ \delta\ ba$ and S/δ is a semilattice.

(iii) Suppose, $e_1\ \delta\ e_2$ with $e_1, e_2 \in E$. Then $e_1 \in \mathfrak{M}(e_1) = \mathfrak{M}(e_2)$, so $e_2|e_1$. Similarly $e_1|e_2$. Hence $e_1\ \mathscr{H}\ e_2$ and by [2], Lemma 2.15, $e_1 = e_2$. Thus each δ-class contains at most one idempotent. Now suppose A is a δ-class containing an idempotent e. Let $a \in A$. Since $e \in \mathfrak{M}(e) = \mathfrak{M}(a) = \mathfrak{M}(a^2)$, there exists $x \in S$ such that $a^2x = e$. Now $a\ \delta\ a^2$ implies $ax\ \delta\ a^2x$, so $ax\ \delta\ e\ \delta\ a$. Hence $ax \in A$ and $a(ax) = e$ implies e is a right zeroid of A. Similarly e is a left zeroid and by [2], §2.5, Exercise 6, A has a group ideal.

A semigroup is said to be *$\mathscr{S}$-indecomposable* if it has no proper semilattice decomposition.

COROLLARY 4. *If the viable semigroup S is $\mathscr{S}$-indecomposable then $S/\delta = 1$ and is either idempotent-free or has a group ideal and exactly one idempotent.*

LEMMA 5. *Assume I is an idempotent-free ideal of S. Then S is viable if and only if the Rees factor semigroup S/I is viable.*

Proof. Assume S is viable and that $ab, ba \in E(S/I)$. If $ab \in I$, then $ba = b(ab)a \in I$, so $ab = ba$ in S/I. So we may assume ab and ba are not in I. But then $ab, ba \in E(S)$. Hence $ab = ba$ in S and so in S/I. Therefore S/I is viable. Conversely, let $ab, ba \in E(S)$. Since S/I is viable $ab = ba$ in S/I. But $ab, ba \notin \mathrm{I}$ since I is idempotent-free. Hence $ab = ba$ in S and S is viable.

A semigroup S is said to be E-inversive if for every $a \in S$ there exists $x \in S$ such that $ax \in E$.

THEOREM 6. *The following are equivalent.*

(i) *Every $\mathscr{J}$-class of S contains at most one idempotent*

(ii) *S is viable.*

(iii) *S is a smilattice of $\mathscr{S}$-indecomposable semigroups each of* *which contains at most one idempotent and a group ideal whenever it contains an idempotent.*

(iv) *S is a semilattice of semigroups having at most one idempotent.*

(v) *S is viable and E-inversive or an ideal extension of an idempotent-free semigroup by a viable E-inversive semigroup.*

Proof. (i) $\Rightarrow$ (ii) If ab and ba are idempotents then $ab = a(ba)b \in S^1baS^1$. Similarly $ba \in S^1abS^1$. Hence $ab \mathscr{J} ba$, so $ab = ba$.

(ii) $\Rightarrow$ (iii) By Tamura [3], S is a semilattice of $\mathscr{S}$-indecomposable semigroups. Since subsemigroups of viable semigroups are viable, each component is viable. The result follows from (4).

(iii) $\Rightarrow$ (iv) a fortiori

(iv) $\Rightarrow$ (i) Suppose $e, f \in E$ with $e \mathscr{J} f$. Then e and f are in the same component of the given semilattice decomposition. Hence $e = f$.

(ii) $\Rightarrow$ (v) Let $I = \{a \in S \mid \mathfrak{M}(a) = \varnothing\}$. If I is empty then S is E-inversive. Otherwise, I is obviously an idempotent-free δ-class of S. Moreover if $ax \mid e$ or $xa \mid e$, $e \in E$, then by (2), $a \mid e$. Hence, $a \in I$ implies $ax, xa \in I$ so that I is an ideal of S. By (5), S/I is viable. Since S/I has a zero, it is E-inversive. In fact, every nonzero element of S/I divides a nonzero idempotent of S/I.

(v) $\Rightarrow$ (ii) Follows from (5).

REMARK. Observe that the semilattice decomposition of (iii) in general will not be isomorphic to S/δ since in fact S may be idempotent free. Also, $\mathscr{J}$ may be replaced $\mathscr{D}$ in the theorem.

LEMMA 7. *S is an ideal extension of a group by a nil semigroup if and only if S is a subdirect product of a group and a nil semigroup.*

Proof. Suppose S is an ideal extension of a group G by a nil semigroup N. Let e be the identity of G. It is easy to see that e is central in S. It is well known that S is a subdirect product of subdirectly irreducible semigroups S_α $(\alpha \in \Omega)$. Let $\sigma_\alpha \colon S \to S_\alpha$ be the natural map. Let $e_\alpha = e\sigma_\alpha$. Then e_α is a central idempotent in S_α and hence is zero or 1 (cf. [1]). If $e_\alpha = 0$, then $\sigma_\alpha(G) = 0$ and hence $S_\alpha = \sigma_\alpha(S)$ is a nil semigroup. If $e_\alpha = 1$, then all of S_α is contained in $\sigma_\alpha(G)$ and hence S_α is a group. Consequently each S_α is a nil semigroup or a group. Let $\Omega_1 = \{\alpha \mid \alpha \in \Omega,\ S_\alpha \text{ is nil}\}$ and let $\Omega_2 = \{\alpha \mid \alpha \in \Omega,\ S_\alpha \text{ is a group}\}$. Let $\psi_i = \prod_{\alpha \in \Omega_i} \sigma_\alpha \colon S \to \prod_{\alpha \in \Omega_i} S_\alpha$ be defined for $i = 1, 2$. One can check that S is a subdirect product of $S\psi_1$ and $S\psi_2$ with $S\psi_1$ a nil semigroup and $S\psi_2$ a group.

Conversely, suppose S is a subdirect of a group G and a nil semigroup N. Consider S embedded in $G \times N$. Let e be the identity of G. There exists $a \in N$ such that $(e, a) \in S$. There exists a positive integer k such that $a^k = 0$. Hence $(e, 0) = (e, a^k) = (e, a)^k \in S$. If $g \in G$, there exists $b \in N$ such that $(g, b) \in S$. Thus $(g, 0) = (e, 0)(g, b) \in S$. Hence $G \times \{0\} \subseteqq S$ and $G \times \{0\}$ is an ideal of S. Let $(g, a) \in S$. Since $a \in N$, there exists a positive integer k such that $a^k = 0$. Hence $(g, a)^k = (g^k, a^k) = (g^k, 0) \in G \times \{0\}$. Therefore S is an ideal extension of the group $G \times \{0\}$ by a nil semigroup.

COROLLARY 8. *The following are equivalent.*

(i) *S is viable and a power of each element lies in a subgroup.*

(ii) *S is a semilattice of semigroups which are ideal extensions of groups by nil semigroups.*

(iii) S is a semilattice of semigroups each of which is a subdirect product of a nil semigroup.

Moreover the decompositions in (ii) *and* (iii) *are the same and coincide with the δ-decomposition as specified in Theorem* 3.

A semigroup S is *separative* if $x^2 = xy = y^2$ $(x, p \in S)$ implies $x = y$.

COROLLARY 9. *The following are equivalent.*

(i) *S is viable, separative and a power of each element of S lies in a subgroup.*

(ii) *S is a semilattice of groups.*

Proof. (i) $\Rightarrow$ (ii) By (8), it suffices to show that if T is separative and an ideal extension of a group G by a nil semigroup, then $T = G$. Let e be the identity of G. Then e is central in T. If $T \neq G$, then there exists $a \in T$, $a \notin G$ with $a^2 \in G$. Then $a^2 = (ae)^2 = a(ae)$. Thus $a = ae \in G$, a contradiction. Hence $T = G$.

(ii) $\Rightarrow$ (i) Obvious.

REFERENCES

1. M. Chacron and G. Thierrin, *σ-reflexive semigroups and rings*, Canad. Math. Bull., (to appear).

2. A. H. Clifford and G. B. Preston, *The algebraic theory of semigroups*, vol. **1**, Amer. Math. Soc., Providence, Rhode Island, 1961.

3. T. Tamura, *Another proof of a theorem concerning the greatest semilattice decomposition of a semigroup*, Proc. Japan Academy, **40** (10), 1964, 777-780.

Received October 31, 1970.

UNIVERSITY OF CALIFORNIA, SANTA BARBARA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

RINGS OF QUOTIENTS AND π-REGULARITY

R. M. RAPHAEL

Throughout this paper rings are understood to be commutative with 1, and subrings are understood to have the same identity as their over-rings. Familiarity with the Utumi-Lambek concept of complete ring of quotients $Q(R)$, of a commutative ring R, is assumed. $Q(R)$ is commutative and it contains a copy of the classical ring of quotients of R (denoted $Q_{Cl}(R)$), obtained by localizing R at its set of nonzero-divisors. Any ring lying between R and $Q(R)$ is called a ring of quotients of R. R is π-regular if for $r \in R$ there exists $r' \in R$ and a positive integer n such that $r^n = (r^n)^2 r'$. This paper investigates the question: if $Q(R)$ is π-regular, under what conditions are all rings of quotients of R π-regular?

The characterization obtained is applied to the case of semiprime rings. Examples are given, followed by some results directed at the problem of characterizing internally those rings R for which $Q(R)$ is π-regular. The author is indebted to the referee for posing the latter question, and for his criticisms. The terminology and notation are consistent with Lambek's Lectures on Rings and Modules.

PROPOSITION 1. (Bourbaki-Storrer, [(6, 5.6), (1, p. 173, 16(d))]. *If R is a commutative ring then the following are equivalent:*

(1) *R is π-regular,*

(2) *$R/\mathrm{rad}\ R$ is regular, where rad R is the prime radical of R,*

(3) *all prime ideals of R are maximal ideals.*

COROLLARY 2. *A semiprime π-regular ring is regular.*

Let R be a ring and let S be an over-ring of R. An element s of S is called integrally dependent on R if there exist elements $r_0, r_1, \cdots, r_{n-1}$ in R such that $s^n + r_{n-1}s^{n-1} + \cdots + r_1 s + r_0 = 0$. The set of all elements of S which are integrally dependent on R is a ring called the integral closure of R in S, and if this is all of S then S is called an integral extension of R.

PROPOSITION 3. [7, p. 259]. *Let R, S be rings, S an integral extension of R. If P is a prime ideal of S, P is maximal in S if and only if $P \cap R$ is a maximal ideal in R.*

DEFINITION 4. A ring is *classical* if it coincides with its classical ring of quotients. Equivalently, each of its elements is a unit or a zero-divisor.

LEMMA 5. *A π-regular ring is a classical ring.*

Proof. Let r be a nonzero-divisor in R, a π-regular ring. Then there exists $r' \in R$ and an integer n such that $r^n(1 - r^n r') = 0$. Since r does not divide zero, neither does r^n so $1 - r^n r' = 0$ which shows that r is a unit.

The main result.

PROPOSITION 6. *Let R be a commutative ring with complete ring of quotients $Q(R)$ which is π-regular. The following are equivalent*:

(1) *$Q(R)$ is integral over R,*
(2) *every ring of quotients of R is π-regular,*
(3) *every ring of quotients of R is classical,*
(4) *$R[q]$ is π-regular for all $q \in Q(R)$,*
(5) *$R[q]$ is classical for all $q \in Q(R)$,*
(6) *the units of $Q(R)$ are integral over R.*

Proof. Clearly $(2) \Rightarrow (4) \Rightarrow (5)$ and $(2) \Rightarrow (3) \Rightarrow (5)$. $(1) \Rightarrow (2)$. If S is a ring of quotients of R, then S is integral over R. Any prime ideal of S contracts to a prime ideal of R which is maximal in R by Proposition 1. Thus by Proposition 3 all prime ideals in S are maximal and by Proposition 1, S is π-regular.

$(5) \Rightarrow (6)$. Let q be a unit in $Q(R)$ with inverse q'. Since $R[q]$ is classical q is either a zero-divisor or a unit in $R[q]$. If it were a zero-divisor in $R[q]$ then it would be both a unit and a zero-divisor in $Q(R)$, an impossibility. Thus q' lies in $R[q]$, and $q' = r_n q^n + \cdots + r_1 q + r_0$ for some $r_i \in R$, $i = 0, 1, \cdots, n$. Now $1 = qq' = r_n q^{n+1} + \cdots + r_1 q^2 + r_0 q$. If one multiplies both sides of the equation by $(q')^{n+1}$ and transposes one obtains the equation $(q')^{n+1} - r_0(q')^n - r_1(q')^{n-1} - \cdots - r_{n-1}(q') - r_n = 0$ which shows that q' is integrally dependent on R. Since every unit is the inverse of a unit (6) is established.

$(6) \Rightarrow (1)$. Let $q \in Q(R)$ Since $Q(R)$ is π-regular there is a $q' \in Q(R)$ such that $q^n = (q^n)^2 q'$. Let $e = q^n q'$, $u = q^n + 1 - q^n q'$. One verifies immediately that $e = e^2$, that u is a unit with inverse $u^{-1} = q^n(q')^2 + 1 - q^n q'$ and that $q^n = ue$. Now e is integral over R, and by (6) u is, so q^n is integral over R, which implies in turn that q is integral over R.

PROPOSITION 7. [4, p. 42]. *Let R be a semiprime ring. Then $Q(R)$ is regular.*

PROPOSITION 8. *Let R be semiprime and let $Q(R)$ be its complete ring of quotients. Then the following are equivalent:*

(1) *$Q(R)$ is integral over R,*

(2) *all rings of quotients of R are regular,*

(3) *all rings of quotients of R are classical.*

Proof. $Q(R)$ is regular so by Proposition 6, $(2) \Rightarrow (3) \Rightarrow (1)$. $(1) \Rightarrow (2)$. Let S be a ring of quotients of R. $Q(R)$ is semiprime so S is as well. By Proposition 6, S is π-regular. Therefore by Corollary 2, S is a regular ring.

EXAMPLE 9. Boolean rings. A ring is Boolean if each element is idempotent. Thus a Boolean ring is regular. Rings of quotients of Boolean rings are discussed in [3, 2.4] where it is shown that a Boolean ring coincides with its complete ring of quotients if and only if it is complete when viewed as a partially ordered set. Furthermore the complete ring of quotients of a Boolean ring is Boolean. Thus if R is a non-complete Boolean algebra, $Q(R)$ is a proper extension of R, which clearly satisfies condition (1) of Proposition 8.

EXAMPLE 10. In Fine-Gillman-Lambek [2, 4.3] the rings $Q_L(X)$ and $Q_F(X)$ are introduced and it is shown that the former is the complete ring of quotients of the latter. To realize $Q_L(X)$ one considers the set of all locally constant continuous real-valued functions whose domains of definition are dense open subsets of a completely regular Hausdorff space X, and divides out by the equivalence relation which identifies two functions which agree on the intersection of their domains. $Q_F(X)$ is the subring determined by the functions with finite range. $Q_F(X)$ is regular. It is not difficult to see that the two rings differ if X is the real field in its usual topology.

Let $g \in Q_L(X)$ and suppose that $g^n + g^{n-1}f_{n-1} + \cdots + f_0 = 0$ for some $f_i \in Q_F(X)$, $i = 0, 1, \cdots, n-1$. We may assume that all the functions are defined on the domain D given by the intersection of their individual domains. Each f_i is defined on a finite clopen partition Π_i of D, on the elements of which it is fixed. Let Π be the common refinement of the Π_i. Then Π is finite and each f_i is fixed on the elements of Π. Since g must satisfy the above polynomial it can assume only a finite number of different values on a given element of Π. Thus g restricted to D has finite range and therefore lies in $Q_F(X)$. Thus the elements of $Q_L(X) - Q_F(X)$ are not integral over

$Q_F(X)$. Thus we have examples of regular rings for which the conditions of Proposition 8 fail.

Proposition 6 demands the π-regularity of $Q(R)$ thus raising the question: for which rings is the complete ring of quotients π-regular? In the Noetherian case the classical ring of quotients is Noetherian and it coincides with the complete ring of quotients. Thus [6, 5.5 and 5.7] the complete ring of quotients is π-regular if and only if it is Artinian. Furthermore Small [5] has shown that a Noetherian ring R has Artinian classical ring of quotients if and only if R satisfies the following 'regularity' condition: if $\bar{r}$ is not a zero-divisor in $R/\text{rad}\, R$, then r is not a zero-divisor in R. We examine the question of $Q(R)$'s π-regularity in the light of this condition. By $\bar{R}$ and $\overline{Q(R)}$ we denote $R/\text{rad}\, R$ and $Q(R)/\text{rad}\, Q(R)$ respectively. The following diagram (with the obvious maps) is commutative

$$\begin{array}{ccc} R & \longrightarrow & Q(R) \\ \downarrow & & \downarrow \\ \bar{R} & \longrightarrow & \overline{Q(R)} \end{array}$$

and $\bar{R} \rightarrow \overline{Q(R)}$ is a monomorphism since $\text{rad}\,(Q(R)) \cap R = \text{rad}\, R$.

LEMMA 11. *If $Q(R)$ is π-regular and* rad R *is nilpotent then R satisfies the regularity condition.*

Proof. Let $\bar{r}$ be a nonzero-divisor in $\bar{R}$. If $\bar{r}$ is a zero-divisor in $\overline{Q(R)}$, then there exists $s \in Q(R) \backslash \text{rad}\, Q(R)$ such that $rs \in \text{rad}\, Q(R)$. There is a dense ideal D in R such that $sD \subset R$. Suppose that $sD \subset R$. Since rad R is nilpotent, $(\text{rad}\, R)^k = (0)$ for some integer k. Thus $s^k D^k = (0)$. But D^k is dense so $s^k = 0$, contradicting the fact that $s \notin \text{rad}\, Q(R)$. Thus there exists $d \in D$ such that $sd \in R \backslash (\text{rad} R)$. Now $r(sd) \in \text{rad}\, R$ contradicting the fact that $\bar{r}$ is not a zero-divisor in $\bar{R}$. Thus $\bar{r}$ is a nonzero-divisor in $\overline{Q(R)}$. But $\overline{Q(R)}$ is regular by Proposition 1, so $\bar{r}$ is invertible in $\overline{Q(R)}$. Thus there is a $q \in Q(R)$ such that $rq - 1 \in \text{rad}\, Q(R)$, from which it is easy to see that r is a unit in $Q(R)$, and therefore not a zero-divisor in R.

LEMMA 12. *If* rad R *is nilpotent then $\overline{Q(R)}$ is a ring of quotients of $\bar{R}$. Furthermore if R satisfies the regularity condition then $\overline{Q(R)}$ contains $Q_{Cl}(\bar{R})$.*

Proof. Let $\bar{q}$ be a nonzero element of $\overline{Q(R)}$. $qD \subset R$, for some dense ideal D of R. Suppose that $qD \subset \text{rad}\, R$. There exists an integer k such that $(\text{rad}\, R)^k = (0)$ so $q^k D^k = (0)$ yielding $q^k = 0$, a contradition. Thus there is a $d \in D$ such that $qd \in R\backslash \mathrm{rad}\, R$ yielding $\bar{q}\bar{d} \neq \bar{0}$ in $\bar{R}$, and $\overline{Q(R)}$ is a ring of quotients of $\bar{R}$. [4, p. 46 no. 5].

If the regularity condition holds and $\bar{r}$ is a nonzero-divisor in $\bar{R}$, then r is a nonzero-divisor in R and $rq = 1$ for some $q \in Q(R)$. But then $\bar{r}\bar{q} = \bar{1}$ showing the nonzero-divisors in $\bar{R}$ have inverses in $\overline{Q(R)}$. Thus $\overline{Q(R)} \supset Q_{Cl}(\bar{R})$.

PROPOSITION 13. *If* $\mathrm{rad}\, R$ *is nilpotent and* $Q_{Cl}(\bar{R}) = Q(\bar{R})$ *then* $Q(R)$ *is π-regular if and only if* R *satisfies the regularity condition.*

Proof. Lemma 11 gives one implication. If $Q_{Cl}(\bar{R}) = \overline{Q(R)}$ then by Lemma 12 $Q(\bar{R}) = \overline{Q(R)}$. But $Q(\bar{R})$ is regular by Proposition 7. Thus by Proposition 1, $Q(R)$ is π-regular.

The above proposition applies to the Noetherian case. More generally if R is commutative with maximum condition on annihilatar ideals then:

(a) $\mathrm{rad}\, R$ is nilpotent [3]

(b) $\bar{R}$ satisfies the maximum condition annihilator ideals [5, 1.16], and

(c) $Q(R) = Q_{Cl}(R)$, [4, p. 114, 5(g)].

By condition (b), condition (c) also holds for the ring $\bar{R}$. This together with condition (a) makes Proposition 13 meaningful for these rings as well.

REFERENCES

1. N. Bourbaki, *Algèbra Commutative*, Chapitres 1, 2. Paris: Herman 1961.

2. Fine, Gillman, and Lambek, *Rings of Quotients of Rings of Continuous Functions*, Montreal, McGill University Press, 1965.

3. Herstein and Small, *Nil rings satisfying certain chain conditions*, Canad. J. Math., **16** (1964), 771-776.

4. J. Lambek, *Lectures on Rings and Modules*, Waltham, Toronto, London: Blaisdell, 1966.

5. Small, *Orders in Artinian Rings*, J. Algebra, **4** (1966), 13-41.

6. H. H. Storrer, *Epimorphism en von kommutativen Ringen*, Comm. Math. Helv., **43** (1968), 378-401.

7. Zariski and Samuel; Commutative Algebra, 1. Van Nostrand, 1958.

Received May 12, 1970. Based partially on a thesis supervised by J. Lambek at McGill University with N.R.C. support.

DALHOUSIE UNIVERSITY AND SIR GEORGE WILLIAMS UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

INFINITE MATRICES SUMMING EVERY ALMOST PERIODIC SEQUENCE

JAMIL A. SIDDIQI

Necessary and sufficient conditions are given for infinite matrices to sum every almost periodic sequence and their basic properties as summability matrices are studied. It is then shown that these matrices enter naturally in the problem of the determination of the jump or total quadratic jump of normalized functions of bounded variation on the circle in terms of the limits of matrix transforms of certain functions of their Fourier-Stieltjes coefficients. The results obtained generalize the classical theorems of Fejér and Wiener as also the extensions of theorems of Wiener given by Lozinskiĭ, Keogh, Petersen and Matveev. Applications are made to the study of coefficient properties of holomorphic functions in the unit disk with positive real part.

1. R. H. C. Newton [11] proved that a regular matrix $A = (a_{n,k})$ sums every periodic sequence if and only if $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} \exp(2\pi ikt)$ exists for each rational t. Vermes [15] generalized this result by proving that an arbitrary matrix $A = (a_{n,k})$ sums every periodic sequence if and only if for every rational t, (1) $\sum_{k=0}^{\infty} a_{n,k} \exp(2\pi ikt)$ converges and (2) $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} \exp(2\pi ikt)$ exists.

The set P of all periodic sequences of complex numbers is a linear subspace of l_∞ that is not closed in the usual norm topology of the Banach space l_∞ since P is meager in l_∞. Berg and Wilansky [3] proved that the closure Q of P in l_∞ is the set of all semi-periodic sequences. (A sequence $x = \{x_k\}$ is called semi-periodic if for any $\varepsilon > 0$, there exists an integer r such that $|x_k - x_{k+rn}| < \varepsilon$ for every n and k). Berg [2], gave a characterization of infinite matrices summing every semi-periodic sequence which is rather involved. We first show that these matrices can be characterized simply as follows:

THEOREM 1. *An infinite matrix $A = (a_{n,k})$ sums every semi-periodic sequence if and only if* (1) $\|A\| = \sup_{n\geq 0} \sum_{k=0}^{\infty} |a_{n,k}| < \infty$ *and* (2) $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} \exp(2\pi ikt)$ *exists for all rational* t.

Proof. If $x \in Q$, then for any $\varepsilon > 0$, there exists a $y \in P$ such that $\|x - y\|_\infty < \varepsilon$. If y is of period r, there exist constants $\lambda_1, \cdots, \lambda_r$ such that

$$\sum_{\nu=1}^{r} \exp(2\pi ik\nu/r) \cdot \lambda_\nu = y_k, \quad k = 0, 1, \cdots, r-1$$

so that

$$\left| \sum_{k=0}^{\infty} a_{m,k}\, x_k - \sum_{k=0}^{\infty} a_{n,k}\, x_k \right| \leqq \left| \sum_{k=0}^{\infty} (a_{m,k} - a_{n,k})\,(x_k - y_k) \right|$$
$$+ \left| \sum_{k=0}^{\infty} (a_{m,k} - a_{n,k}) \Big(\sum_{\nu=1}^{r} \exp\,(2\pi i k\nu/r)\, \lambda_\nu \Big) \right|$$
$$\leqq 2\,||\,A\,||\,\varepsilon + \varepsilon$$

for n and m sufficiently large. Hence $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k}\, x_k$ exists. Thus (1) and (2) are sufficient. The necessity of (1) can be established as in Berg and Wilansky [2] (see also the proof of Theorem 2 below where similar arguments have been given to prove the necessity of Theorem 2 (1) independently of the use of Theorem 1), and that of (2) is immediate since $\{\exp\,(2\pi ikt)\}$ is periodic when t is rational.

2. A sequence $x = \{x_k\}$ of complex numbers is called almost periodic if to any $\varepsilon > 0$, there corresponds an integer $N = N(\varepsilon) > 0$ such that among any N consecutive integers there exists an integer r with the property $|\,x_k - x_{k+r}\,| < \varepsilon$ for all k. If we denote by AP the set of all almost periodic sequences of complex numbers, then clearly AP is a linear subspace of l_∞ and $P \subset \bar{P} = Q \subset AP \subset l_\infty$. Also AP is a closed subspace of l_∞. For if $\{x^{(n)}\}$ is a Cauchy sequence in AP, there exists an $x = \{x_k\} \in l_\infty$ such that $\lim_{n\to\infty} ||\,x^{(n)} - x\,||_\infty = 0$. Given any $\varepsilon > 0$, we can choose an n such that $|\,x_k^{(n)} - x_k\,| < \varepsilon/3$ for every k. Since $x^{(n)} \in AP$, there exists an integer $N = N(\varepsilon)$ such that among N consecutive integers there is an integer r such that $|\,x_k^{(n)} - x_{k+r}^{(n)}\,| < \varepsilon/3$ for every k so that

$$|\,x_k - x_{k+r}\,| \leqq |\,x_k - x_k^{(n)}\,| + |\,x_k^{(n)} - x_{k+r}^{(n)}\,| + |\,x_{k+r}^{(n)} - x_{k+r}\,|$$
$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

for every k. Thus AP is a Banach space. We note that $Q \subsetneqq AP$ since if t is irrational, then $\{\exp\,(2\pi ikt)\}$ is almost periodic but not semi-periodic.

Infinite matrices summing every almost periodic sequence in AP can be characterized as follows:

THEOREM 2. *An infinite matrix $A = (a_{n,k})$ sums every almost periodic sequence if and only if* (1) $||\,A\,|| = \sup_{n\geqq 0} \sum_{k=0}^{\infty} |\,a_{n,k}\,| < \infty$ *and* (2) $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} \exp\,(2\pi ikt)$ *exists for all* t.

Proof. Suppose that A sums every almost periodic sequence. Since for each t, $\{\exp\,(2\pi ikt)\} \in AP$, (2) holds. To prove the necessity of (1), we first observe that if $y \in l_1$, its norm $||\,y\,||_{(AP)^*}$ is identical with its l_1-norm. For if $y = \{y_k\} \in l_1$, we define a sequence $\tilde{x}$ of period n by the rule: $\tilde{x}_k = \operatorname{sgn} y_k$ for $k \leqq n$ so that

$$\| y \|_{(AP)^*} = \sup_{\substack{x \in AP \\ \| x \| = 1}} | y(x) | \geqq | y(\tilde{x}) | \geqq \sum_1^n | y_k| + \sum_{n+1}^\infty \tilde{x}_k y_k ,$$

where

$$\left| \sum_{n+1}^\infty \tilde{x}_k y_k \right| \leqq \sum_{n+1}^\infty | y_k | \longrightarrow 0$$

as $n \to \infty$. Thus $\| y \|_{(AP)^*} \geqq \| y \|_{l_1}$. Clearly $\| y \|_{(AP)^*} \leqq \| y \|_{l_1}$ so that $\| y \|_{(AP)^*} = \| y \|_{l_1}$.

For each fixed n, put

$$y_N(x) = \sum_{k=0}^N a_{n,k} x_k, \text{ where } x \in AP .$$

$y_N \in (AP)^*$ and $\lim_{N\to\infty} y_N(x)$ exists for each $x \in AP$. By the uniform boundedness principle,

$$\| y_N \|_{(AP)^*} = \| y_N \|_{l_1} = \sum_1^N | a_{n,k} | \leqq M_n < \infty$$

for each N so that $\sum_{k=0}^\infty | a_{n,k} | < \infty$ for each n. If we put

$$z_n(x) = \sum_{k=0}^\infty a_{n,k} x_k, \ x \in AP ,$$

then $z_n \in (AP)^*$ and $\lim_{n\to\infty} z_n(x)$ exists for each $x \in AP$. Applying once more the uniform boundedness principle, we get

$$\| A \| = \sup_{n \geqq 0} \sum_{k=0}^\infty | a_{n,k} | < \infty .$$

Thus (1) holds.

To prove the sufficiency of conditions (1) and (2), we note that if $x = \{x_k\} \in AP$, there exists a sequence $\{\sum_0^N b_j \exp(2\pi i \lambda_j k)\} \in AP$ such that for all k,

$$\left| x_k - \sum_0^N b_j \exp(2\pi i \lambda_j k) \right| < \varepsilon .$$

Now

$$\begin{aligned} \left| \sum_{k=0}^\infty a_{m,k} x_k - \sum_{k=0}^\infty a_{n,k} x_k \right| &\leqq \left| \sum_{k=0}^\infty (a_{m,k} - a_{n,k}) \Big(x_k - \sum_0^N b_j \exp(2\pi i \lambda_j k) \Big) \right| \\ &\quad + \left| \sum_{k=0}^\infty (a_{m,k} - a_{n,k}) \sum_0^N b_j \exp(2\pi i \lambda_j k) \right| \\ &\leqq 2 \| A \| \varepsilon + \varepsilon \end{aligned}$$

for m and n sufficiently large. Thus $\lim_{n\to\infty} \sum_{k=0}^\infty a_{n,k} x_k$ exists.

We call a matrix $A = (a_{n,k})$ satisfying the conditions (1) and (2) of Theorem 2, an almost periodic matrix. We now establish a few properties of these matrices. We recall that the set of all sequences summable by a given matrix $A = (a_{n,k})$ is called its convergence field and is denoted by (A). If (A) contains all convergent sequences then A is called conservative. It is known that A is conservative if and only if (1) $\|A\| < \infty$, (2) $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} = \alpha$ exists and (3) $\lim_{n\to\infty} a_{n,k} = \alpha_k$ exists for each fixed k. We have:

PROPOSITION 1. *An almost periodic matrix is always conservative.*

Proof. It is sufficient to show that $\lim_{n\to\infty} a_{n,k} = \alpha_k$ exists. If we put $K_n(t) = \sum_{k=0}^{\infty} a_{n,k} \exp(2\pi ikt)$, then $\{K_n\}$ is a sequence of continuous functions on $[0, 1]$ such that $\lim_{n\to\infty} K_n(t) = K(t)$ exists for each t and $|K_n(t)| \leqq \|A\| < \infty$ for all n and all t. By bounded convergence theorem,

$$\lim_{n\to\infty} a_{n,k} = \lim_{n\to\infty} \int_0^1 K_n(t) e^{-2\pi ikt}\, dt = \int_0^1 K(t)\, e^{-2\pi ikt}\, dt$$

exists for each k.

The converse is easily seen to be false.

A conservative matrix $A = (a_{n,k})$ is called multiplicative if there exists an $m > 0$ such that $\lim_{n\to\infty} x_n = x$ implies $\lim_{n\to\infty} A_n(x) = \lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} x_k = mx$ and then A is called m-multiplicative. Since

$$\lim_{n\to\infty} A_n(x) = \alpha x + \sum_{k=0}^{\infty} \alpha_k (x_k - x) ,$$

it follows that a conservative matrix $A = (a_{n,k})$ is multiplicative if and only if $\lim_{n\to\infty} a_{n,k} = 0$ for each k. An examination of the proof of Proposition 1 shows that an almost periodic matrix $A = (a_{n,k})$ is multiplicative if and only if

$$\int_0^1 K(t)\, e^{-2\pi ikt}\, dt = 0 \qquad \textit{for all } k = 0, \pm 1, \pm 2, \cdots$$

so that, by the uniqueness of Fourier expansion, if and only if $K(t) = 0$ a.e. Thus we have:

PROPOSITION 2. *An almost periodic matrix $A = (a_{n,k})$ is multiplicative if and only if $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} \exp(2\pi ikt) = 0$ a.e. in $(0, 1)$.*

It may be remarked that there exist multiplicative almost periodic matrices for which the above limit is not zero for all $t \in (0, 1)$. The positive matrix $A = (a_{n,k})$ where $a_{n\,2k} = 0$, $a_{n\,2k+1} = n^k/(n+1)^{k+1}$ for $k = 0, 1, 2, \cdots$ is one such matrix. We also have:

PROPOSITION 3. *An almost periodic matrix* $A = (a_{n,k})$ *is regular if and only if* (1) $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} = 1$ *and* (2) $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} \exp(2\pi ikt) = 0$ *a. e. in* $(0, 1)$.

We call an almost periodic matrix normal if (1) $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} = 1$ and (2) $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} \exp(2\pi ikt) = 0$ for all $t \in (0, 1)$. Clearly a normal almost periodic matrix is regular.

3. A sequence $x = \{x_k\}$ is said to be $\mathscr{F}_A$ summable where $A = (a_{n,k})$ if $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} x_{k+p}$ exists uniformly in $p = 0, 1, 2, \cdots$. An obvious modification of the reasoning used in the proof of Theorem 2 yields the following:

THEOREM 3. *Let* A *be a given matrix. Then every almost periodic sequence is summable* $\mathscr{F}_A$ *if and only if* A *is an almost periodic matrix.*

In particular, a sequence $x = \{x_k\}$ of complex numbers is called almost convergent if $\lim_{n\to\infty}(n+1)^{-1} \sum_{k=0}^{n} x_{k+p}$ exists uniformly in $p = 0, 1, \cdots$ i.e., if it is summable $\mathscr{F}_A$ where A is the matrix of the arithmetic mean. Every almost periodic sequence is almost convergent (cf. Theorem 3) but not conversely. Lorentz [8] has proved that a matrix $A = (a_{n,k})$ sums every almost convergent sequence to its almost convergence limit if and only if (1) A is regular and (2) $\lim_{n\to\infty} \sum_{k=0}^{\infty} |\Delta a_{n,k}| = 0$ where $\Delta a_{n,k} = a_{n,k} - a_{n,k+1}$ for $k = 0, 1, \cdots$. He calls matrices $A = (a_{n,k})$ satisfying (1) and (2) strongly regular. A simple modification of his proof of the above characterization yields the following:

THEOREM 4. *A matrix* $A = (a_{n,k})$ *sums every almost convergent sequence if and only if* (1) A *is conservative and* (2) $\lim_{n\to\infty} \sum_{k=0}^{\infty} |\Delta(a_{n,k} - \alpha_k)| = 0$, *where* $\alpha_k = \lim_{n\to\infty} a_{n,k}$.

A natural problem in this connection is to determine whether there exist matrices that sum every almost periodic sequence without necessarily summing every almost convergent sequence. The fact that there exist almost convergent sequences that are not almost periodic does not resolve the problem since, a priori, it is not clear that the convergence field of an almost periodic matrix does not contain all almost convergent sequences. This is settled by the following:

THEOREM 5. *There exists a normal almost periodic matrix* $A = (a_{n,k})$ *such that* $|A| = (|a_{n,k}|)$ *is also almost periodic but* A *is not strongly regular.*

Proof. Let $A = (a_{n,k})$ be defined as follows:

$$a_{n,0} = 0 \,,$$
$$a_{n,k} = \frac{1}{n} \quad \text{for } 1 \leqq k \leqq n \,,$$
$$a_{n,k} = \exp\{i\pi(k-n)\log(k-n)\}/n \quad \text{for } n < k \leqq 2n \,,$$
$$a_{n,k} = 0 \quad \text{for } k > 2n.$$

Clearly $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} = 1$, since $\sum_{k=0}^{\infty} a_{n,k} = 1 + (1/n) \sum_{n=1}^{2n} \exp \{i\pi(k-n)\log(k-n)\}$ and the partial sums $s_n(x)$ of the series $\Sigma \exp\{i\pi k \log k + ikx\}$ are $O((n)^{1/2})$ uniformly in x (cf. Zygmund [17] p. 199). Also $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} \exp(2\pi ikt) = 0$ for all $t \in (0, 1)$ since, in view of the above cited result,

$$\sum_{k=0}^{\infty} a_{n,k} \exp(2\pi ikt) = \frac{1}{n} \frac{e^{i2\pi t}(1 - e^{i2\pi nt})}{1 - e^{i2\pi t}} + O\left(\frac{\sqrt{n}}{n}\right) (n \to \infty) \,.$$

Also since $\| A \| = 2$, it follows that A is normal almost periodic. For $t \in (0, 1)$

$$\sum_{k=0}^{\infty} | a_{n,k} | \exp(2\pi ikt) = \frac{1}{n} e^{2\pi it} \frac{(1 - e^{4\pi int})}{1 - e^{2\pi it}} = o(1) \ (n \to \infty)$$

and

$$\sum_{k=0}^{\infty} | a_{n,k} | \to 2 \ (n \to \infty) \,,$$

so that $| A |$ is also almost periodic. However A is not strongly regular. In fact,

$$\sum_{k=0}^{\infty} | a_{n,k} - a_{n,k+1} | = \frac{2}{n} \sum_{1}^{n-1} \left| \sin \frac{\pi}{2} \left\{ \log\left(1 + \frac{1}{k}\right)^k + \log(k+1) \right\} \right| + o(1) \,.$$

Since

$$k \log\left(1 + \frac{1}{k}\right) \to 1$$

as $k \to \infty$, we have

$$\sin \frac{\pi}{2} \left\{ k \log\left(1 + \frac{1}{k}\right) + \log(k+1) \right\} = \cos \frac{\pi}{2} \log(k+1) + o(1)$$

so that

$$\frac{2}{n} \sum_{1}^{n-1} \left| \sin \frac{\pi}{2} \left\{ k \log\left(1 + \frac{1}{k}\right) + \log(k+1) \right\} \right|$$
$$= \frac{2}{n} \sum_{1}^{n-1} \left| \cos \frac{\pi}{2} \log(k+1) \right| + o(1) \ (n \to \infty) \,.$$

We assert that

$$\frac{1}{n}\sum_{1}^{n-1}\left|\cos\frac{\pi}{2}\log(k+1)\right|$$

does not tend to zero. In fact, if we put

$$u_k = \left|\cos\frac{\pi}{2}\log(k+1)\right| - \left|\cos\frac{\pi}{2}\log k\right|,$$

then we have

$$|u_k| \leqq 2\left|\sin\frac{\pi}{4}\log(k^2+k)\right|\left|\sin\frac{\pi}{4}\log\left(1+\frac{1}{k}\right)\right| = O\left(\frac{1}{k}\right)(k\to\infty).$$

It is known (cf. Zygmund [17], p. 78) that if a series Σu_k is summable $(C, 1)$ and $u_k = 0(1/k)$, then Σu_k is convergent. Hence, if in our case the series Σu_n were summable $(C, 1)$ to zero i.e., if

$$\frac{1}{n}\sum_{1}^{n-1}\left|\cos\frac{\pi}{2}\log(k+1)\right|$$

were to tend to zero as $n\to\infty$, the series Σu_k would be convergent which is not the case since

$$\sum_{1}^{n} u_k = \left|\cos\frac{\pi}{2}\log(n+1)\right|$$

does not tend to a limit as $n\to\infty$.

As a corollary of Theorem 5, we get that there exist sequences that are almost convergent without being almost periodic.

4. Let $V[0, 2\pi]$ denote the class of all normalized functions F of bounded variation in $[0, 2\pi]$ such that $F(x+2\pi) - F(x) = F(2\pi) - F(0)$ for all x and let $\{C_n\}$ be the sequence of Fourier-Stieltjes coefficients of F. We now show that almost periodic matrices enter naturally in the solution of the problem of the determination of the jump or the total quadratic jump of a function $F\in V[0, 2\pi]$ by means of the limits of the matrix transforms of $\{C_k e^{ikx}\}$ or $\{|C_k|^2\}$ respectively.

THEOREM 6. *Let $A = (a_{n,k})$ be such that $\|A\| < \infty$. Then for every $F\in V[0, 2\pi]$ and for every $x\in[0, 2\pi]$, the sequence $\{C_k e^{ikx}\}$ is summable A or $\mathscr{F}_A$ to $(2\pi)^{-1}D(x)$ where, $D(x) = F(x+0) - F(x-0)$, if and only if A is normal almost periodic.*

Proof. We prove the assertion for summability A, the proof for summability $\mathscr{F}_A$ being similar. The condition is necessary, for if we choose F: $F(t) = 2\pi$ for $0 < t \leqq 2\pi$ and $F(0) = 0$, then $C_k = 1$ for all k, $D(0) = 2\pi$ and $D(x) = 0$ for $0 < x < 2\pi$ so that $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} \exp(2\pi ikx) = 0$ for all $x \in (0, 1)$ and $\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} = 1$.

Suppose that A is a normal almost periodic matrix. Then

$$\sum_{k=0}^{\infty} a_{n,k} C_k e^{ikx} = \sum_{k=0}^{\infty} a_{n,k} \frac{1}{2\pi} \sum_{j=0}^{\infty} D(x_j) e^{ik(x-x_j)} + (2\pi)^{-1} \int_0^{2\pi} K_n\left(\frac{x-t}{2\pi}\right) dF_c(t) ,$$

where $K_n(t) = \sum_{k=0}^{\infty} a_{n,k} \exp(2\pi ikt)$, $\{x_j\}$ are the points of jump of F in $[0, 2\pi)$ and F_c is the continuous part of F. Clearly the first term on the right tends to $D(x)/2\pi$ as $n \to \infty$. The second term on the right tends to 0 as $n \to \infty$, for, given an $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\int_{x-\delta}^{x+\delta} \left| dF_c(t) \right| < \frac{\varepsilon}{2} \| A \|^{-1}$$

so that

$$\left| \int_{x-\delta}^{x+\delta} K_n\left(\frac{x-t}{2\pi}\right) dF_c(t) \right| < \frac{\varepsilon}{2}$$

and, by bounded convergence theorem,

$$\left| \left(\int_0^{x-\delta} + \int_{x+\delta}^{2\pi} \right) K_n\left(\frac{x-t}{2\pi}\right) dF_c(t) \right| < \frac{\varepsilon}{2}$$

for large n. Thus $\{C_k e^{ikx}\}$ is summable A to $D(x)/2\pi$.

Theorem 6 generalizes a theorem of Fejér [4] (cf. also Zygmund [17] p. 107) and, in particular, it shows that in Fejér's theorem the summability $(C, 1)$ can be replaced by almost convergence.

THEOREM 7. *Let $A = (a_{n,k})$ be such that $\| A \| < \infty$. Then for every $F \in V[0, 2\pi]$, the sequence $\{|C_k|^2\}$ is summable A and $\mathscr{F}_A$ to $(4\pi^2)^{-1} \sum_{j=0}^{\infty} | D(x_j)|^2$, where $\{x_j\}$ are the points of jump of F in $[0, 2\pi)$ if and only if A is a normal almost periodic matrix.*

Proof. If we put $F^*(x) = (2\pi)^{-1} \int_0^{2\pi} F(x + t)\, d\bar{F}(t)$, then $F^* \in V[0, 2\pi]$, $F^*(+0) - F^*(-0) = (2\pi)^{-1} \sum_{j=0}^{\infty} | D(x_j) |^2$ and the Fourier-Stielt,jes coefficients of F^* are $\{|C_k|^2\}$. Applying Theorem 6 to F^* at $x = 0$ we get the proof of the sufficiency part of the above theorem.

To prove the necessity part, we observe if $\{C_k\}$ and $\{C'_k\}$ are the Fourier-Stieltjes coefficients of F and F' in $V[0, 2\pi]$, then

$$\{C_k \bar{C}'_k + \bar{C}_k C'_k\} = \left\{ \frac{1}{2} (|C_k + C'_k|^2 - | C_k - C'_k|^2) \right\}$$

is summable A to $(4\pi^2)^{-1}\sum_{l=0}^{\infty}\{D(y_l)\,\overline{D'(y_l)}+\overline{D(y_l)}\,D'(y_l)\}$, where $\{y_l\}$ denotes the set of all points of jump of F and F'. On replacing F' by iF', we get that $\{C_k\bar{C}_k'-\bar{C}_kC_k'\}$ is summable A to $(4\pi^{-2})$ $\sum_{l=0}^{\infty}\{D(y_l)\ \overline{D'(y_l)}-\overline{D(y_l)}\ D'(y_l)\}$ so that $\{C_k\bar{C}_k'\}$ is summable A to $(4\pi^2)^{-1}\sum_{l=0}^{\infty}\{D(y_l)\overline{D'(y_l)}\}$. If we choose $F'\in V[0,2\pi]$ such that $F'(t)=0$ for $0\leqq t<x$, $F'(t)=2\pi$ for $x<t\leqq 2\pi$, then $C_k'=e^{-ikx}$ so that $\{C_k\,e^{ikx}\}$ is summable A to $D(x)/2\pi$ for each $x\in[0,2\pi]$ and Theorem 6 applies. Thus we conclude that A is normal almost periodic.

Theorem 7 generalizes a theorem of Wiener [16] (cf. also Zygmund [17] p. 108) and, in particular, it shows that in Wiener's theorem the summability $(C,1)$ can be replaced by almost convergence.

As an immediate consequence we have the following:

THEOREM 8. *For functions $F\in V[0,2\pi]$, the following conditions are equivalent:*

(1) *F is continuous,*

(2) *$\{|C_k|^2\}$ is summable A or $\mathscr{F}_A$ to 0 by a normal almost periodic matrix A,*

(3) *$\{|C_k|\}$ is summable A or $\mathscr{F}_A$ to 0 by a normal almost periodic matrix $A=(a_{n,k})$ for which $\lim_{n\to\infty}\sum_{k=0}^{\infty}|a_{n,k}|\exp(2\pi ikt)=0$ for all $t\in(0,1)$.*

Proof. The equivalence of (1) and (2) is a direct consequence of Theorem 7. Suppose that F is continuous. Then the convolution F^* as defined in the proof of Theorem 7 is continuous and belongs to $V[0,2\pi]$. If we go through the steps of the proof of Theorem 6 for F^* with $x=0$ and $D^*(0)=(2\pi)^{-1}\sum_{j=0}^{\infty}|D(x_j)|^2$ and note that the Fourier-Stieltjes coefficients of F^* are $\{|C_k|^2\}$, we conclude that the hypothesis $\lim_{n\to\infty}\sum_{k=0}^{\infty}|a_{n,k}|\exp(2\pi ikt)=0$ for all $t\in(0,1)$ without the requirement that $\lim_{n\to\infty}\sum_{k=0}^{\infty}|a_{n,k}|$ exists, assures that $\lim_{n\to\infty}\sum_{k=0}^{\infty}|a_{n,k}|\,|C_k|^2=0$. Applying Schwarz inequality, we get that $\lim_{n\to\infty}\sum_{k=0}^{\infty}|a_{n,k}|\,|C_k|=0$ and consequently that $\{|C_k|\}$ is summable A to 0. Similarly we show that $\{|C_k|\}$ is summable $\mathscr{F}_A$ to 0. Thus (1) implies (3). Suppose that $\{|C_k|\}$ is summable A to 0. If we write $C_k=C_k'+C_k''$, where C_k' and C_k'' are respectively the Fourier-Stieltjes coefficients of the saltus part and the continuous part of F, we have

$$\left|\sum_{k=0}^{\infty}a_{n,k}(|C_{k+p}|-|C_{k+p}'|)\right|\leqq\sum_{k=0}^{\infty}|a_{n,k}|\,|C_{k+p}''|\leqq B\sum_{k=0}^{\infty}|a_{n,k}|\,|C_{k+p}''|^2\,.$$

Since the last term tends to zero in view of the equivalence of (1) and (2) already proved and since the almost periodic sequence $\{|C_k'|\}$ is summable $\mathscr{F}_A$, we have

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k} \,|\, C_{k+p} \,| = 0$$

uniformly in p. Similarly we prove that

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} |\, a_{n,k} \,|\, |C_{k+p} \,| = L \text{ (say)} ,$$

exists uniformly in p. If we set

$$\sigma_{n,p} = \sum_{k=0}^{\infty} a_{n,k} \,|\, C_{k+p} \,| ,$$

we see that

$$\begin{aligned} \sum_{p=0}^{\infty} |\, a_{n,p} \,|\, \sigma_{n,p} &= \sum_{p=0}^{\infty} |\, a_{n,p} \,| \sum_{k=0}^{\infty} a_{n,k} \,|\, C_{k+p} \,| \\ &= \sum_{k=0}^{\infty} a_{n,k} \sum_{p=0}^{\infty} |\, a_{n,p} \,|\, |\, C_{k+p} \,| . \end{aligned}$$

If for an $\varepsilon > 0$, we choose an $N = N(\varepsilon)$ such that for all $n \geqq N$

$$|\, \sigma_{n,p} \,| < \varepsilon, \quad \left| \left(\sum_{p=0}^{\infty} |\, a_{n,p} \,|\, |\, C_{k+p} \,| \right) - L \right| < \varepsilon$$

uniformly in p and k respectively, it follows that for $n \geqq N$

$$\begin{aligned} \left| \sum_{k=0}^{\infty} a_{n,k} \right| L &\leqq \left| \sum_{k=0}^{\infty} a_{n,k} \left[L - \sum_{p=0}^{\infty} |\, a_{n,p} \,|\, |\, C_{k+p} \,| \right] \right| \\ &\quad + \left| \sum_{k=0}^{\infty} a_{n,k} \sum_{p=0}^{\infty} |\, a_{n,p} \,|\, |\, C_{k+p} \,| \right| \\ &\leqq \|\, A \,\|\, \varepsilon + \sum_{p=0}^{\infty} |\, a_{n,p} \,|\, \sigma_{n,p} \\ &\leqq \|\, A \,\|\, \varepsilon + \|\, A \,\|\, \varepsilon = 2 \,\|\, A \,\|\, \varepsilon . \end{aligned}$$

Making $n \to \infty$, we get $L \leqq 2 \,\|\, A \,\|\, \varepsilon$ so that $L = 0$. Thus $\{|\, C_k \,|\}$ is summable $\mathscr{F}_{|A|}$ to 0. Hence $\{|\, C_k \,|^2\}$ is summable $\mathscr{F}_{|A|}$ to 0 and therefore summable $\mathscr{F}_A$ to 0. Since (1) and (2) have already been shown to be equivalent, we conclude that F is continuous. Thus (3) implies (1).

Theorem 8 generalizes a theorem of Wiener [16] (cf. Zygmund [17] p. 108) and contains as special cases various generalizations of that theorem including those given by Lozinskiĭ [9] and Matveev (cf. Bari [1] p. 256).

Theorem 8 can be reformulated in the following strengthened forms which we state separately.

THEOREM 9. *For* $F \in V[0, 2\pi]$ *to be continuous, it is necessary* *that* $\{|C_k|^2\}$ *should be summable* $\mathscr{F}_A$ *to* 0 *by each normal almost periodic matrix* A *and sufficient that* $\{|C_k|^2\}$ *should be summable* A *to* 0 *by some normal almost periodic matrix* A.

THEOREM 10. *For* $F \in V[0, 2\pi]$ *to be continuous, it is necessary that* $\{|C_k|^2\}$ *and* $\{|C_k|\}$ *should be summable* $\mathscr{F}_{|A|}$ *to* 0 *by each normal almost periodic matrix* $A = (a_{n,k})$ *for which* (1) $\lim_{n\to\infty} \sum_{k=0}^{\infty} |a_{n,k}| \exp(2\pi ikt) = 0$ *for all* $t \in (0, 1)$ *and sufficient that either* $\{|C_k|^2\}$ *or* $\{|C_k|\}$ *should be summable* A *by some normal almost periodic matrix satisfying* (1).

It may be pointed out that the assertion regarding summability $\mathscr{F}_{|A|}$ in Theorem 10 has been established in the course of the proof of Theorem 8. Theorem 10 generalizes the following strengthened form of Wiener's theorem given by Keogh and Petersen [7].

THEOREM A. *For* $F \in V[0, 2\pi]$ *to be continuous, it is necessary that* $\{|C_k|^2\}$ *and* $\{|C_k|\}$ *should be almost convergent to zero and sufficient that either* $\{|C_k|^2\}$ *or* $\{|C_k|\}$ *should be summable to zero by some summation method which contains almost convergence.*

Since every strongly regular matrix $A = (a_{n,k})$ is an almost periodic matrix satisfying (1) and the $(C, 1)$ matrix is strongly regular, the direct proposition of Theorem A is a particular case of the corresponding assertion in Theorem 10. We have already remarked earlier (§ 3) that Lorentz [8] has shown that a matrix sums all almost convergent sequences to their almost convergence limits if and only if it is strongly regular. The sufficiency part of Theorem A is therefore also a special case of the corresponding assertion in Theorem 10.

Lorentz [8] has proved that (a) if A is regular, then summability $\mathscr{F}_A$ implies almost convergence and that (b) if A is strongly regular, then summability $\mathscr{F}_A$ and almost convergence are equivalent. Although not explicitly stated by Lorentz, it follows that summability $\mathscr{F}_A$ and almost convergence are equivalent if and only if A is strongly regular. For, if A is not strongly regular, there exists an almost convergent sequence that is not summable A and hence a fortiori not summable $\mathscr{F}_A$. Hence if A is not strongly regular, summability $\mathscr{F}_A$ is strictly weaker then almost convergence. Since there exist non-strongly regular normal almost periodic matrices satisfying (1) (cf. Theorem 5), Theorem 10 is sharper than Theorem A in both directions.

A particularly interesting corollary of Theorem 10 is the following:

COROLLARY. *For a continuous $F \in V[0, 2\pi]$ with Fourier coefficients $\{c_k\}$, we have $\sum_p^{n+p} |c_k| = o(\log n)$ $(n \to \infty)$ uniformly in p.*

This result is significant since there exist continuous functions of bounded variation for which $c_k \neq o(1)$ $(k \to \infty)$.

5. In Theorems 6 and 7 of the preceding section, we started with matrices A satisfying the condition that $\|A\| < \infty$ and then found the necessary and sufficient condition in order that these be effective in the problem of the determination of jump or total quadratic jump of functions belonging to $V[0, 2\pi]$. However, as we shall see below, this restriction is not necessary. In fact, if we call a matrix $A = (a_{n,k})$ for which

$$K_n:\ K_n(t) = \sum_{n=0}^{\infty} a_{n,k} \exp(2\pi ikt)$$

is continuous in $[0, 1]$ for each n, a matrix with continuous kernel, we have the following:

THEOREM 11. *Let $A = (a_{n,k})$ be a matrix with continuous kernel $\{K_n\}$. Then for every $F \in V[0, 2\pi]$ and for every $x \in [0, 2\pi]$, the sequence $\{C_k e^{ikx}\}$ is summable A or $\mathscr{F}_A$ to $(2\pi)^{-1}D(x)$, where*

$$D(x) = F(x+0) - F(x-0)$$

if and only if

(i) $\sup_{N \geqq 0} \max_t |K_n^N(t)| = M_n < \infty$ *for every* n,

(ii) $\sup_{n \geqq 0} \max_t |K_n(t)| = M < \infty$,

(iii) $\lim_{n\to\infty} K_n(t) = 0$ for $t \in (0, 1)$ *and* $=1$ *otherwise,*

where $K_n^N(t) = \sum_{k=0}^{N} a_{n,k} \exp(2\pi ikt)$, $N = 0, 1, \cdots$.

Proof. If A sums every sequence $\{C_k e^{ikx}\}$ for each x in $[0, 2\pi]$, and for each $F \in V[0, 2\pi]$, it follows that for each fixed n the sequence of continuous functions $\{K_n^N\}$ converges weakly in $C[0, 1]$ so that, by the uniform boundedness principle, we get (i). Since

$$\sum_{k=0}^{\infty} a_{n,k} C_k = \lim_{N\to\infty} \frac{1}{2\pi}\int_0^{2\pi} K_n^N\left(\frac{-t}{2\pi}\right) dF(t) = \frac{1}{2\pi}\int_0^{2\pi} K_n\left(\frac{-t}{2\pi}\right) dF(t)$$

and

$$\lim_{n\to\infty}\int_0^{2\pi} K_n\left(\frac{-t}{2\pi}\right) dF(t)$$

exists for all $F \in V[0, 2\pi]$, again, by the uniform boundedness principle, we get (ii). If for each $t \in [0, 1]$, we choose F: $F(x) = 0$ in $[0, t]$, $F(x) = 2\pi$ in $(t, 2\pi]$, we get $C_k = e^{-ikt}$ so that (iii) holds. Thus conditions (i), (ii) and (iii) are necessary. The proof of the sufficiency of these conditions is the same as in case of Theorem 6, if we observe that the continuity of K_n and (i) assure that

$$\sum_{k=0}^{\infty} a_{n,k} C_k e^{ikx} = \frac{1}{2\pi}\int_0^{2\pi} K_n\left(\frac{x-t}{2\pi}\right) dF(t) .$$

The assertion for summability F_A can be similarly proved.

We call a matrix $A = (a_{n,k})$ normal Fejér effective if it satisfies conditions (i), (ii) and (iii) of Theorem 11.

We can similarly prove the following analogues of Theorems 7 and 8 respectively.

THEOREM 12. *Let $A = (a_{n,k})$ be a matrix with continuous kernel. Then for every $F \in V[0, 2\pi]$, the sequence $\{|C_k|^2\}$ is summable A and $\mathscr{F}_A$ to $(4\pi^2)^{-1}\sum_{j=0}^{\infty}|D(x_j)|^2$, where $\{x_j\}$ are the points of jump of F in $[0, 2\pi)$, if and only if A is a normal Fejér effective matrix.*

THEOREM 13. *For functions $F \in V[0, 2\pi]$, the following conditions are equivalent:*

(1) *F is continuous,*

(2) *$\{|C_k|^2\}$ is summable A or $\mathscr{F}_A$ to 0 by a normal Fejér effective matrix A with continuous kernel,*

(3) *$\{|C_k|\}$ is summable A or $\mathscr{F}_A$ to 0 by a normal Fejér effective matrix $A = (a_{n,k})$ with continuous kernel, for which $\lim_{n\to\infty}\sum_{k=0}^{\infty}|a_{n,k}|\exp(2\pi ikt) = 0$ for all $t \in (0, 1)$.*

Theorems analogous to Theorems 9 and 10 can also be established.

A normal almost periodic matrix is clearly a normal Fejér effective matrix since the hypothesis $\|A\| < \infty$ implies that the conditions (i) and (ii) of Theorem 11 are satisfied. But the converse is not true. Consider the matrix $A = (a_{n,k})$, where

$$a_{n,0} = 0 ,$$

$$a_{n,k} = \frac{1}{n} \qquad \text{for } 1 \leqq k \leqq n ,$$

$$a_{n,k} = \frac{\exp\{i\pi(k-n)\log(k-n)\}}{n^{3/4}(k-n)^{1/2+\alpha}}\left(0 < \alpha < \frac{1}{2}\right) \qquad \text{for } k > n .$$

It can be verified that A is a normal Fejér effective matrix with continuous kernel that does not satisfy the condition $\|A\| < \infty$, since even $\sum_{k=0}^{\infty}|a_{n,k}| = \infty$ so that applying Theorem 2 one concludes that the matrix A is not an almost periodic matrix. It follows that for the validity of the theorems of this section we need normal matrices that may not be conservative.

6. Hayman [6] and Petersen [12] have applied Wiener's theorem and its generalization Theorem A respectively to the study of coefficient properties of holomorphic functions with positive real part. We can apply Theorem 10 to obtain the following:

THEOREM 14. *Let* $\psi(z) = 1 + \sum_{n=1}^{\infty} b_n z^n = u + iv$ *be holomorphic in* $|z| < 1$ *satisfying the condition* $u > 0$ *there and let*

$$\psi(z) = \int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z}\, dg(\theta)$$

be the Herglotz representation of ψ *where* g *is non-decreasing on* $[0, 2\pi]$. *Let* g_1 *denote the saltus part of* g.

a. *If* $A = (a_{n,k})$ *is a normal almost periodic matrix for which* $\lim_{n\to\infty} \sum_{k=0}^{\infty} |a_{n,k}| \exp(2\pi ikt) = 0$ *for all* $t \in (0, 1)$, *then there exists a complex Borel measure* μ *uniquely determined by* g_1 *and* A, *defined on the disk* $\Delta = \{w: |w| \leqq 2\}$ *such that*

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k}\, \phi(b_{n+p}) = \int_{\Delta} \phi(w) d\mu \qquad \text{for all } \phi \in C(\Delta)$$

uniformly in p, *where* $C(\Delta)$ *denotes the space of all complex continuous functions on* Δ.

b. *If, moreover,* $\lim_{n\to\infty} \sum_{k=0}^{\infty} |a_{n,k}|$ *exists, then there exists a positive Borel measure* ν *uniquely determined by* g_1 *and* A, *defined on* Δ *such that*

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} |a_{n,k}|\, \phi(b_{n+p}) = \int_{\Delta} \phi(w) d\nu \qquad \text{for all } \phi \in C(\Delta)$$

uniformly in p.

c. *If we define* χ_E: $\chi_E(b_k) = 1$ *if* $b_k \in E$ *and* $= 0$ *if* $b_k \notin E$ *where* E *is a Borel set, then under the same assumption on* $A = (a_{n,k})$ *as in* **b**, *we have*

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} |a_{n,k}|\, \chi_E(b_{k+p}) = \nu(E)$$

uniformly in p.

Proof. If $\{\theta_\nu\}$ denote the points of discontinuity and $\{\alpha_\nu\}$ the jumps of g in $[0, 2\pi)$, and $g_2 = g - g_1$, we get

$$\begin{aligned}\psi(z) &= \int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z}\, dg(\theta) = \int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z}\, dg_1(\theta) + \int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z}\, dg_2(\theta) \\ &= \sum_{\nu=0}^{\infty} \alpha_\nu \frac{1+ze^{-i\theta_\nu}}{1-ze^{-i\theta_\nu}} + \int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z}\, dg_2(\theta) \\ &= \sum_{n=0}^{\infty} c_n z^n + \sum_{n=0}^{\infty} d_n z^n ,\end{aligned}$$

where $c_n = \sum_{\nu=0}^{\infty} 2\alpha_\nu e^{-in\theta_\nu} \, (n \geqq 1)$. It is known (cf. Hayman [5] p. 12) that $|b_n| \leqq 2$. Since

$$1 = \psi(0) = \int_0^{2\pi} dg(\theta) = \sum \alpha_\nu + g_2(2\pi) - g_2(0) \geqq \sum \alpha_\nu ,$$

$$|c_n| \leqq 2 \sum \alpha_\nu = 2c_0 \leqq 2 \, .$$

If we set $\phi(Re^{i\theta}) = (3-R)\,\phi\,(2e^{i\theta})\,(2 < R < 3)$ and $\phi(Re^{i\theta}) = 0 (R \geqq 3)$, then ϕ is continuous in the whole plane, $\phi = 0$ for $|w| \geqq 3$ and hence ϕ is uniformly continuous. This extension of ϕ outside the disk $|w| \leqq 2$ does not alter $\phi\,(b_n)$ and $\phi\,(c_n)$ since $|b_n| \leqq 2$ and $|c_n| \leqq 2$.

If we put

$$L\,(\phi) = \lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k}\,\phi\,(c_{k+p}) \, ,$$

then this limit exists uniformly in p, since $\{\phi(c_k)\}$ is almost periodic. We now show that

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k}\,[\phi(c_{k+p}) - \phi(b_{k+p})] = 0$$

uniformly in p. Since ϕ is uniformly continuous, given any $\varepsilon > 0$, we can choose $\delta > 0$ so that *if* $|w - w'| < \delta$, then $|\phi(w) - \phi(w')| < \varepsilon$. Now

$$\begin{aligned}\sum_{k=0}^{\infty} |a_{n,k}|\,|\phi\,(b_{k+p}) - \phi\,(c_{k+p})| &\leqq \Big(\sum_{|d_{k+p}|<\delta} + \sum_{|d_{k+p}|\geqq\delta} \Big) \\ &\qquad \times |a_{n,k}|\,|\phi\,(b_{k+p}) - \phi(c_{k+p})| \\ &\leqq \varepsilon\,\|A\| + 2M \sum_{|d_{k+p}|\geqq\delta} |a_{n,k}| \, .\end{aligned}$$

Since A is normal almost periodic and is such that $\lim_{n\to\infty} \sum_{k=0}^{\infty} |a_{n,k}| \exp(2\pi ikt) = 0$ for all $t \in (0, 1)$ and $\{d_n\}$ are the Fourier-Stieltjes coefficients of continuous part of g, it follows from Theorem 10 and the inequalities

$$\delta \sum_{|d_{k+p}|\geqq\delta} |a_{n,k}| \leqq \sum_{|d_{k+p}|\geqq\delta} |a_{n,k}|\,|d_{k+p}| \leqq \sum_{k=0}^{\infty} |a_{n,k}|\,|d_{k+p}|$$

that $\lim_{n\to\infty} \sum_{|d_{k+p}|\geqq\delta} |a_{n,k}| = 0$ uniformly in p. It follows that

$$L(\phi) = \lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k}\,\phi(c_{k+p}) = \lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k}\,\phi(b_{k+p})$$

exists uniformly in p and depends only on g_1. Since L is a bounded linear functional on the space of all continuous functions in the plane with compact support, there exists a complex Borel measure μ in the plane such that

$$L(\phi) = \lim_{n\to\infty} \sum_{k=0}^{\infty} a_{n,k}\,\phi\,(b_{k+p}) = \int_\Delta \phi\,(w)\,d\mu \, .$$

This establishes (a), and (b) follows from it.

Suppose E is a set, whose frontier has ν-measure zero. Then we can find a compact set $K \subset \operatorname{Int} E$ such that $\nu(K) > \nu(E) - \varepsilon$. Further, we can construct ϕ continuous in the plane such that $0 \leqq \phi \leqq 1$, $\phi(K) = 1$ and $\phi(cE) = 0$. Then

$$\sum_{k=0}^{\infty} | a_{n,k} | \phi(b_{k+p}) \leqq \sum_{k=0}^{\infty} | a_{n,k} | \chi_E(b_{k+p})$$

so that

$$\varliminf_{n\to\infty} \sum_{k=0}^{\infty} | a_{n,k} | \chi_E(b_{k+p}) \geqq \int_{\Delta} \phi(w) d\nu \geqq \nu(K) > \nu(E) - \varepsilon .$$

Similarly there exists a bounded open set $U \supset E$ such that $\nu(U) < \nu(E) + \varepsilon$. Choose a continuous function ψ in the plane such that $0 \leqq \psi \leqq 1$, $\psi(E) = 1$ and $\psi(cU) = 0$. Then

$$\sum_{k=0}^{\infty} | a_{n,k} | \psi(b_{k+p}) \geqq \sum_{k=0}^{\infty} | a_{n,k} | \chi_E(b_{k+p})$$

so that

$$\varlimsup_{n\to\infty} \sum_{k=0}^{\infty} | a_{n,k} | \chi_E(b_{k+p}) \leqq \int \psi(w) d\nu \leqq \nu(U) < \nu(E) + \varepsilon .$$

Since ε is arbitrary, we get

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} | a_{n,k} | \chi_E(b_{k+p}) = \nu(E) ,$$

uniformly in p.

We remark that in the above theorem we can replace the normal almost periodic matrix by a Fejér effective matrix with continuous kernel.

Finally, I would like to express my thanks to Professor B. Kuttner for kindly reading the manuscript of this paper and making valuable suggestions.

References

1. N. Bari, *A treatise on trigonometric series*, vol. **1**. Oxford: Pergamon Press 1964.
2. I. D. Berg *The algebra of semiperiodic sequences*, Michigan Math. J., **10**, (1963), 237-239.
3. I. D. Berg, and A. Wilansky, *Periodic, almost-periodic and semi-periodic sequences*, Michigan Math. J. **9** (1962), 363-368.
4. L. Fejér, *Uber die Bestimmung des Sprunges einer Funktionen aus ihrer Fourierreihe*, J. Reine Angew. Math. **142** (1913), 165-168.
5. W. K. Hayman, *Multivalent Functions*, Cambridge, 1958.
6. ———, *On functions with positive real part*, J. London Math. Soc., **36** (1961), 35–48.

7. F. R. Keogh and G. M. Petersen, *A strengthened form of a theorem of Wiener*, Math. Zeitschrift, **71** (1959), 31–35.
8. G. G. Lorentz, *A contribution to the theory of divergent sequences*, Acta Mathematica, **80** (1948), 169–190.
9. S. M. Lozinskiĭ, *On a theorem of N. Wiener*, C. R. (Doklady) Acad. Sci. URSS (N.S.), **49** (1945), 542–545.
10. ———, *On a theorem of N. Wiener II* C. R. (Doklady) Acad. Sci. URSS (N.S), **53** (1946), 687–690.
11. R. H. C. Newton, *On the summability of periodic sequences* I, II, Proc. Kon. Nederl. Akad. Wetensch. Ser A. **57** (1954), 533–544, 545–549.
12. G. M. Petersen, *On functions with positive real part*, J. London Math. Soc. **36** (1961), 49–51.
13. J. A. Siddiqi, *Coefficient properties of certain Fourier series*, Math. Ann. **181** (1969), 242–254.
14. ———, *On mean values for almost periodic functions*, Arch. der Math. **20** (1969), 648–655.
15. P. Vermes, *Infinite matrices summming every periodic sequence*, Proc. Kon. Nederl. Akad. Wetensch. Ser A. **58** (1955), 627–633.
16. N. Wiener, *The quadratic variation of a function and its Fourier coefficients*, Massachusetts J. Math. **3** (1924), 72–94.
17. A. Zygmund, *Trigonometric Series* vol. 1, Cambridge 1959.

Received December 7, 1970. The research was supported in part by the National Research Council of Canada Grant A-4057 and in part by the grant of a fellowship of the Summer Research Institute of the Canadian Mathematical Congress.

UNIVERSITE DE SHERBROOKE, SHERBROOKE, CANADA
SUMMER RESEARCH INSTITUTE OF CANADIAN MATHEMATICAL CONGRESS
UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER, CANADA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

UNIFORM CONVERGENCE FOR MULTIFUNCTIONS

R. E. SMITHSON

Let $\mathscr{F}$ be a collection of multivalued functions on a topological space into uniform space. The topology of uniform convergence is defined on $\mathscr{F}$, and it is shown that for point compact functions this topology is larger than the pointwise topology. Some results are given on uniform convergence of nets in $\mathscr{F}$. It is also shown that if $\mathscr{F}$ consists of point compact continuous functions on a compact space, then the compact open topology and topology of uniform convergence are the same. Finally the following Ascoli theorem for multifunctions is obtained. Theorem: Let $\mathscr{C}$ be the set of point compact, continuous multifunctions on a compact regular space into a T_2-uniform space. Then $\mathscr{F} \subset \mathscr{C}$ is compact if and only if (i) $\mathscr{F}$ is closed in $\mathscr{C}$, (ii) $\mathscr{F}[x]$ has compact closure for each x and (iii) $\mathscr{F}$ is equicontinuous.

1. Introduction. In [2] the topology of pointwise convergence and the compact open topologies were defined for sets of multivalued functions. Basic properties of these topologies (such as separation axioms, etc.) were studied and some characterizations of compact sets were obtained. The purpose of the present paper is to continue the development of the basic topologies to topologies generated by uniformities, and hence, to a discussion of uniform convergence.

Thus let $(X, \mathscr{T})$ be a topological space, and let $(Y, \mathscr{V})$ be a uniform space. Let $\mathscr{F}$ be a family of multivalued functions on X into Y. We construct a uniformity for $\mathscr{F}$ as follows. If $V \in \mathscr{V}$, define $W(V)$ by: $W(V) = \{(F, G) \in \mathscr{F} \times \mathscr{F} \mid \text{for all } x \in X,\ (y, G(x)) \cap V \neq \phi$ for all $y \in F(x)$, and $(F(x), y') \cap V \neq \phi$ for all $y' \in G(x)\}$.

The cumbersomeness of this definition is a result of the fact that $F(x)$ and $G(x)$ are subsets of Y rather than points. Note, that if F and G are singlevalued functions, then $W(V) = \{(F, G) \mid (F(x), G(x)) \in V\}$. We could have used the set $\{(F, G) | F(x) \times G(x) \subset V\}$, but then we might not get many pairs in members of the uniformity. Another possibility is sets of the form $\{(F, G) \mid F(x) \times G(x) \cap V \neq \phi\}$. This could give a reasonable definition, but this condition is somewhat weaker than the one chosen, and in some cases allows too many pairs (F, G) in the set. Now we let $\mathfrak{W}$ be the uniformity generated by the collection of all such sets $W(V)$. The topology generated by $\mathfrak{W}$ is called the topology of uniform convergence (we sometimes abbreviate this to the u.c. topology or simply u.c.) and we obtain the following relationship between this topology and the topology of pointwise convergence defined in [2].

DEFINITION. A multifunction F: $X \to Y$ on a topological space X into a topological space Y is called *point compact* if and only if $F(x)$ is compact for each $x \in X$.

If $(Y, \mathscr{V})$ is a uniform space, we shall assume that Y has the topology generated by the uniformity $\mathscr{V}$. (See Kelley [1] for notation and definitions.) Further, if $A \subset X$, then A^* and A° denote the closure and interior of A respectively. Finally, in this paper the terms function and multifunction will be synonymous.

LEMMA 1. *If each member of $\mathscr{F}$ is point compact, then the topology of uniform convergence is larger than the topology of pointwise convergence.*

Proof. From [2] we have that the pointwise topology is generated by sets of the form $\{F \in \mathscr{F} \mid F(x) \subset U\}$ or $\{F \in \mathscr{F} \mid F(x) \cap U \neq \phi\}$ where $x \in X$ and U is an open subset of Y. Thus let $\mathscr{O} = \{F \in \mathscr{F} \mid F(x) \subset U\}$, and let $H \in \mathscr{O}$. We shall show that there is a $V \in \mathscr{V}$ such that $W(V)[H] \subset \mathscr{O}$. Since $H(x)$ is compact and U is open, there exists a member $V \in \mathscr{V}$ such that $V[H(x)] \subset U$ [1; pg. 199, Th. 33]. Now suppose that $G \in W(V)[H]$; then $(H, G) \in W(V)$ and hence, if $y' \in G(x)$, then $(H(x), y') \cap V \neq \phi$. Therefore there is a $y \in H(x)$ such that $(y, y') \in V$. Thus $y' \in V[y] \subset V[H(x)] \subset U$. That is, $G(x) \subset U$, and so $W(V)[H] \subset U$. Now suppose that $\mathscr{O} = \{F \mid F(x) \cap U \neq \phi\}$, and let $H \in \mathscr{O}$. Then there exists $y \in H(x) \cap \mathscr{O}$ and there exists a $V \in \mathscr{V}$ such that $V[y] \subset \mathscr{O}$. If $G \in W(V)[H]$, then $(y, G(x)) \cap U \neq \phi$. Thus there is a $y' \in G(x)$ such that $(y, y') \in V$ and hence, $y' \in V[y] \subset \mathscr{O}$. That is, $G(x) \cap U \neq \phi$ and $W(V)[H] \subset \mathscr{O}$. These two results show that the topology of uniform convergence is larger than the topology of pointwise convergence.

We can now use this result to get the following.

THEOREM 2. *Let $\mathscr{F}$ be the set of point compact multifunctions on X into Y. Then a net $\{F_\alpha, \alpha \in D\}$ converges uniformly to $F \in \mathscr{F}$ if and only if $\{F_\alpha, \alpha \in D\}$ is a Cauchy net, relative to $\mathscr{W}$, and converges pointwise to F.*

Proof. That uniform convergence implies pointwise convergence follows from Lemma 1, and if a net converges uniformly, it is a Cauchy net with respect to $\mathscr{W}$.

Now suppose that $\{F_\alpha, \alpha \in D\}$ is a Cauchy net with respect to $\mathscr{W}$ and suppose $F_\alpha \to F$ pointwise.

Now let $W(V) \in \mathscr{W}$, $V \in \mathscr{V}$. We need to show that there is a $\beta \in D$ such that $F_\alpha \in W(V)[F]$ for all $\alpha \geq \beta$. Let V' and V_1 be closed symmetric members of $\mathscr{V}$ such that $V' \circ V' \subset V_1 \subset V$.

Since $\{F_\alpha, \alpha \in D\}$ is Cauchy, there exists a β such that $(F_\alpha, F_r) \in W(V')$ for all $\alpha, \gamma \geq \beta$.

Now let $x \in X$, and let $y' \in F(x)$. If $V_0 \in \mathscr{V}$, then there exists β' such that $\gamma \geq \beta'$ implies that $F_\gamma(x) \cap V_0[y'] \neq \phi$. Thus there is a net $\{y_\gamma, \gamma \in D'\}$, $y_\gamma \in F_\gamma(x)$ such that $y_\gamma \to y'$. Now if $\alpha > \beta$ and $\gamma > \max(\beta', \beta)$ then, $(F_\alpha(x), y_\gamma) \cap V' \neq \phi$. Thus there is a net $(y_{\alpha\gamma}, y_\gamma) \in V'$, $\gamma \in D'$ in $F_2(x)$. Now if $\{y_{\alpha\gamma} \mid \gamma \in D'\} = A$ is finite there is a γ_0 and a D_0', a cofinal subset of D' such that $(y_{\alpha\gamma_0}, y_\gamma) \in V'$, $\gamma \in D_0'$. Otherwise $A \subset F_\alpha(x)$ is infinite and since $F_\alpha(x)$ is compact A has a limit point $y_0 \in F_\alpha(x)$. Then in either case there is a $y_0 \in F_\alpha(x)$, such that $(y_0, y') \in V'$. We have shown that $(F_\alpha(x), y') \cap V' \neq \phi$ for all $x \in X$, where $y' \in F(x)$, and for all $\alpha \geq \beta$.

The next step is to show that $(y, F(x)) \cap V' \neq \phi$ for all x, and all $\alpha \geq \beta$ where $y \in F_\alpha(x)$. For this let $x \in X$ and $V_0 \in \mathscr{V}$ with V_0 symmetric. Then there exists a β' such that $F_\gamma(x) \subset V_0[F(x)]$ for $\gamma > \beta'$. Thus there is a net (y, y'_α) with $y'_\alpha \in F_\gamma(x)$ and $y_\gamma \in F(x)$, $\gamma \in D'$.

Now for $y \in F_\alpha(x)$, $(y, F_\gamma(x)) \cap V' \neq \phi$ for $\alpha, \gamma \geq \beta$ and all x. Hence, for all $x \in X$, $(y, y_\gamma) \in V' \circ V'$ where $y_\gamma \in F(x)$. If y_0 is a limit point of $\{y_\gamma, \gamma \in D'\}$ in $F(x)$, then since V_1 is closed $(y, y_0) \in V_1 \subset V$ and so $(y, F(x)) \cap V \neq \phi$ for all $x \in X$, $y \in F_\alpha(x)$, $\alpha > \beta$.

These two together imply that $(F_\alpha, F) \in W(V)$ for $\alpha > \beta$ and so $\{F_\alpha, \alpha \in D\}$ converges uniformly to F.

DEFINITIONS. Let $F: X \to Y$ be a multifunction on a topological space X into a topological space Y.

(1) The function F is *upper semi-continuous* (u.s.c.) if and only if whenever $F(x) \subset V$, an open subset of Y, there is an open set $U \subset X$ such that $x \in U$ and $F(U) \subset V$.

(2) The function F is lower *semi-continuous* (l.s.c.) if and only if whenever $F(x) \cap V \neq \phi$, V open, there is an open set U, such that $x \in U$ and $F(x') \cap V \neq \phi$ for all $x' \in U$.

(3) The function F is *continuous* if and only if it is both u.s.c. and l.s.c.

THEOREM 3. *Suppose the net $\{F_\alpha, \alpha \in D\}$ converges uniformly to F. If each F_α is* u.s.c. (l.s.c., continuous) *and if F is point compact, then F is* u.s.c. (l.s.c., continuous).

Proof. Let $x \in X$ and $F(x) \subset S$, an open subset of Y. Let $V \in \mathscr{V}$ such that $V[F(x)] \subset S$, and $V' \in \mathscr{V} \ni V' \circ V' \subset V$ and let $\beta \in D$ be such that $F_\alpha \in W(V')[F]$ for all $\alpha > \beta$. Now let $\alpha > \beta$ and then $F_\alpha(x) \subset V'[F(x)]$ (we may assume the $F_\alpha(x)$ is contained in the interior of $V'[F(x)]$) and so there exists an open set $U \subset X$ such that $F_\alpha(U) \subset V'[F(x)]$. Also since $(F, F_\alpha) \in W(V')$ we have $(y', F_\alpha(x')) \cap V' \neq \phi$ for all x' and $y' \in F(x')$ (in particular this holds for all $x' \in U$). But then there exists a $y \in F(x)$ such that $(y', y) \in V' \circ V' \subset V$ and hence $F(x') \subset V[F(x)] \subset S$ for all $x' \in U$ and thus F is u.s.c.

Now suppose that $F(x) \cap S \neq \phi$, and let $y \in F(x) \cap S$. Further, let $V \in \mathscr{V}$ be such that $V[y] \subset S$ and let $V' \in \mathscr{V}$ and $V' \circ V' \subset V$ (we assume that V and V' are symmetric). Let $\beta \in D$ such that $F_\alpha \in W(V')[F]$ for all $\alpha > \beta$. Let α be fixed, $\alpha > \beta$, and then $F_\alpha(x) \cap V'[y] \neq \phi$. So let $x \in U \subset X$ be an open set such that $F_\alpha(x') \cap V'[y'] \neq \phi$ for all $x' \in U$, where $y' \in F_\alpha(x) \subset V'[y]$. Now there is a $y'' \in F(x')$ such that $(y'', y') \in V'$ and hence, $(y'', y) \in V' \circ V' \subset V$. Hence, $F(x') \cap S \neq \phi$ for all $x' \in U$, and so F is l.s.c.

The above two parts show that if each F_α is both u.s.c. and l.s.c., then so is F. Consequently if F_α is continuous for each α, then F is continuous.

COROLLARY. *Let $\{F_\alpha, \alpha \in D\}$ be a net of* u.s.c. (l.s.c.) *functions into a T_2-space such that for each $V \in \mathscr{V}$ there is a $\beta \in D$ such that, $\alpha > \beta$, and $y_1, y_2 \in F_\alpha(x)$ implies that $(y_1, y_2) \in V$. If $\{F_\alpha, \alpha \in D\}$ converges uniformly to F, then F is a continuous single-valued function.*

LEMMA 4. *Suppose that $F: X \to Y$ is a continuous point compact function on the space X into the regular space Y. Let $K \subset X$ be compact, and let U be an open subset of Y such that $F(x) \cap U \neq \phi$ for all $x \in K$. Then there exists a compact set $C \subset U \cap F(K)$ such that $F(x) \cap C \neq \phi$ for all $x \in K$.*

Proof. Let $x \in K$ and let $y \in F(x) \cap U$. Then there is an open set $V_x \subset Y$ such that $y \in V_x \subset V_x^* \subset U$, and an open set $W_x \subset X$ such that $x \in W_x$ and if $x' \in W$, then $F(x') \cap V_x \neq \phi$. Pick such a V_x and W_x for each $x \in K$. Thus the family $\mathscr{W} = \{W_x : x \in K\}$ is an open cover of K and so there is a finite subcover, $W_{x_1}, \cdots, W_{x_k}$. Let V_{x_1} be the set corresponding to W_{x_1} as above. Then the set $C' = \bigcup_{i=1}^k V_{x_i}^* \subset U$ is closed and $F(x) \cap C' \neq \phi$ for all $x \in K$. Finally, since F is u.s.c. and point compact, $F(K)$ is compact and so $C = C' \cap F(K)$ is the desired set.

REMARK. If we merely require that F be l.s.c. in Lemma 4, the proof given shows that there is a closed subset C' of U such that $F(x) \cap C' \neq \phi$ for all $x \in K$.

LEMMA 5. *Let $\mathscr{F}$ be the family of continuous, point compact functions on a compact space $(X, \mathscr{T})$ into the uniform space $(Y, \mathscr{V})$. Then the topology of uniform convergence is the same as the compact open topology.*

Proof. Let $K \subset X$ be compact and let $\mathcal{O}_1, \mathcal{O}_2$ be open subsets of Y. From [2] the compact open topology is generated by sets of the form $T = \{F \in \mathcal{F}: F(K) \subset \mathcal{O}_1$ and $F(x) \cap \mathcal{O}_2 \neq \phi$ for all $x \in K\}$. First we shall show that T is open in the topology of uniform convergence. For this let $F \in T$. Since K is compact, F is continuous and $F(x)$ compact, $F(K)$ is compact and so there is a member $U_1 \in \mathcal{V}$ such that $V_1[F(K)] \subset \mathcal{O}_1$ Also, by Lemma 4 there is a compact set $C \subset F(K) \cap \mathcal{O}_2$ such that $F(x) \cap C \neq \phi$ for all $x \in K$. Let V_2 be a member of $\mathcal{V}$ such that $V_2[C] \subset \mathcal{O}_2$, and let V be a symmetric member of $\mathcal{V}$ such that $V \subset V_1 \cap V_2$. Then if $G \in W(V)[F]$, we get $G(K) \cap \mathcal{O}_1$ and $G(x) \cap \mathcal{O}_2 \neq \phi$ for all $x \in K$. Hence $W(V)[F] \subset T$, and so T is open with respect to the topology of uniform convergence.

Now let $V \in \mathcal{V}$ and consider the set $W(V)[F]$. Let V' be a closed symmetric member of the uniformity such that $V' \circ V' \subset V$. If $x \in X$, then, since F is point compact, there exists a finite set $\{y_1, \cdots, y_k\} \subset F(x)$ such that $F(x) \subset \bigcup_1^n V'[y_i]^0$. Further, by the continuity of F, there exists a closed, hence compact, neighborhood K of x such that $F(K) \subset \bigcup_1^k V'[y_i]^0$, and $F(z) \cap V'[y_i]^0 \neq \phi$ for all $i = 1, \cdots, k$ and $z \in K$. Since X is compact we obtained a finite cover $K_1, \cdots, K_m$ of X of such sets together with corresponding sets $\bigcup_1^k V'[y_{i_1}]^0, \cdots, \bigcup_1^{k_m} V'[y_{i_m}]^0$. Set $S_j = \bigcup_1^{k_j} V'[y_{i_j}]^0$ and define a set U_j as follows:

$$U_j = \{G \in \mathcal{F}: G(K_j) \subset S_j \text{ and } G(x) \cap V'[y_{i_j}]^0 \neq \phi \\ \text{for all } i_j \text{ and } x \in K_j\} .$$

Note that $F \in U_j$ for each $j = 1, \cdots, m$. Further, let $G \in \bigcap_1^m U_j$ and let $y \in G(x)$ where $x \in K_j$. Since $G \in U_j$, $G(x) \subset S_j$ and hence, $y \in V'[y_{i_j}]^0$ for some y_{i_j}. Also $F(x) \cap V'[y_{i_j}] \neq \phi$ follows from the construction of the K_j's. Hence, $(F(x), y) \cap V \neq \phi$. Finally, if $y \in F(x)$ for $x \in K_j$, then, since $G(x) \cap V'[y_{i_j}]^0 \neq \phi$, there exists a $y' \in G(x)$ such that $(y, y') \in V' \circ V'$. Thus $(y, G(x)) \cap V \neq \phi$, and the u.c. topology is contained in the compact open topology. Hence, the lemma follows.

DEFINITION. Let $\mathcal{F}$ be a family of functions on the space X into the uniform space $(Y, \mathcal{V})$ the family $\mathcal{F}$ is *equicontinuous at* $x \in X$ if and only if for each $V \in \mathcal{V}$ there is a neighborhood U of x such that for all $F \in \mathcal{F}$ (1) $F(U) \subset V[F(x)]$, and (2) for each $z \in U$, $G(z) \cap V[y] \neq \phi$ for all $y \in F(x)$.

LEMMA 6. *Let $\mathcal{F}$ be a collection of point compact functions on the space X into the uniform space $(Y, \mathcal{V})$ which is equicontinuous at x. Then the pointwise closure of $\mathcal{F}$ in the family of point compact functions on X into Y is also equicontinuous at x.*

Proof. Let F be a point compact function which is in the pointwise closure of $\mathscr{F}$. Let V be a closed member of $\mathscr{V}$ and let $V' \in \mathscr{V}$ be closed and symmetric and suppose $V' \circ V' \subset V$. Let U be a neighborhood of x such that for each $G \in \mathscr{F}$, $G(U) \subset V'[G(x)]$ and $G(z) \cap V'[y] \neq \phi$ for $z \in U$ and $y \in G(x)$. Now if $F(U) \not\subset V[F(x)]$, there is a $z \in U$ and a $y \in F(z)$ such that $y \notin V[F(x)]$. Since $F(x)$ is compact, there is a finite set $\{y_1, \cdots, y_k\} \subset F(x)$ such that $F(x) \subset \bigcup_i^k V'[y_i]^\circ = S$. Further there is an open set T such that $y \in T$ and $T \cap V[y_i] = \phi$ for each i. Now set $W(S) = \{H: H(x) \subset S\}$ and $W(T) = \{H: H(z) \cap T \neq \phi\}$. Then $W = W(S) \cap W(T)$ is a pointwise open set containing F such that $W \cap \mathscr{F} = \phi$. This is a contraction and so we conclude that $F(U) \subset V[F(x)]$. A similar argument will show that $F(z) \cap V[y] \neq \phi$ for all $y \in F(x)$ and $z \in U$. Hence the lemma follows.

We say that the family $\mathscr{F}$ is *equicontinuous* in case it is equicontinuous at each point. Then if $\mathscr{F}$ is equicontinuous, each member of $\mathscr{F}$ is l.s.c., and if each member of $\mathscr{F}$ is point compact, thene ach member of $\mathscr{F}$ is u.s.c. and hence, each member of $\mathscr{F}$ is continuous.

Let $\mathscr{F}$ be a family of functions from the topological space X into the topological space Y. A topology $\mathscr{T}$ on $\mathscr{F}$ is said to be *jointly continuous* (j.c.) if and only if the function $P: \mathscr{F} \times X \rightarrow Y$ defined by $P(F, x) = F(x)$ is continuous.

LEMMA 7. *If $\mathscr{F}$ is an equicontinuous collection of point compact functions, then the pointwise topology for $\mathscr{F}$ is jointly continuous.*

Proof. First we shall show that P is u.s.c. For this suppose that $F(x) \subset W$ where W is an open subset of Y and $(F, x) \in \mathscr{F} \times X$. Since $F(x)$ is compact, there is a symmetric V in the uniformity such that $V \circ V[F(x)] \subset W$, and such that $F(x) \subset \bigcup V[y]^\circ = S$ for $y \in F(x)$. Since $\mathscr{F}$ is equicontinuous, there is a neighborhood U of x such that $G(U) \subset V[G(x)]$ for all $G \in \mathscr{F}$. Further, let $T = \{G \in \mathscr{F}: G(x) \subset S\}$. Then $(F, x) \in T \times U$ and $P(T \times U) \subset W$. A similar argument shows P is l.s.c. and so the pointwise topology is jointly continuous.

COROLLARY. *If F is an equicontinuous family of point compact functions on the compact space X into the uniform space $(Y, \mathscr{V})$, then the* u.c. *topology, the pointwise topology and the compact open topology are all the same.*

Proof. This follows from Lemmas 5 and 7 together with Propositions 6 and 7 of [2].

We need one more lemma before stating one of the main theorems of this paper.

LEMMA 8. *If $\mathscr{F}$ is compact relative to a* j.c. *topology $\mathscr{T}$, and if each member of $\mathscr{F}$ is point compact, then $\mathscr{F}$ is equicontinuous.*

Proof. Let $V \in \mathscr{V}$, the uniformity for Y; let V' be a symmetric member of $\mathscr{V}$ such that $V' \circ V' \subset V$, and let $x \in X$. Since P is continuous, we can find a neighborhood $W(F)$ for each $F \in \mathscr{F}$ and a neighborhood $U(F)$ of x such that there is a finite subset $\{y_1, \cdots, y_k\}$ of $F(x)$ and such that for $G \in W(F)$, $G(z) \subset \bigcup_1^k V'[y_i]$, and $G(z) \cap V'[y_i] \neq \phi$ for each y_i. Then there is a finite subcover $W(F_1), \cdots, W(F_n)$ of $\mathscr{F}$ with corresponding sets $U(F_1), \cdots, U(F_n)$. Let $U = \bigcap_{i=1}^n U(F_j)$. Then if $G \in \mathscr{F}$, $G \in W(F_j)$ for some j, and if $z \in U$, then $G(z) \subset \bigcup_i^k V'[y_i]$ where $y_1 \in F_j(x)$. Thus, since $G(x) \cap V'[y_i] \neq \phi$, for each $y \in G(z)$ there is a $y' \in G(x)$ such that $(y, y') \subset V' \circ V' \subset V$ and so $G(z) \subset V[G(x)]$ for all $z \in U$. On the other hand let $y \in G(x)$. Since $G(x) \subset \bigcup_1^k V'[y_i]$ and $G(z) \cap V'[y_i] \neq \phi$ for all y, we get $G(z) \cap V[y] \neq \phi$. Hence, $\mathscr{F}$ is equicontinuous at x, and the lemma follows.

Now by combining the above results with Theorem 3 in [2] we get the following Ascoli Theorem.

THEOREM 9. *Let $\mathscr{C}$ be the set of all continuous, point compact functions on a compact regular space into a T_2, uniform space. Let $\mathscr{C}$ have the topology of uniform convergence. Then a subset $F \subset \mathscr{C}$ is compact if and only if*

(i) *$\mathscr{F}$ is closed in $\mathscr{C}$,*

(ii) *$F[x] = \cup \{F(x)\colon F \in \mathscr{F}\}$ has compact closure for each $x \in X$, and*

(iii) *$\mathscr{F}$ is equicontinuous.*

We can extend many of the above results in the following way. Let $\mathscr{A}$ be a family of subsets of X. Then in the definition of $W(V)$ replace $x \in X$ by $x \in A$ for some $A \in \mathscr{A}$. Then generate a uniformity by these sets. This gives us the topology of uniform convergence on members of $\mathscr{A}$. (If $\mathscr{A} = \{X\}$, there is no difference, and if $\mathscr{A}$ is all singletons we get pointwise convergence). In particular if $\mathscr{A}$ is the set of compact sets then we obtain the topology of uniform convergence on compacta. Then if we use the topology of uniform convergence on compacta in place of the u.c. topology, we can obtain results analogous to Theorem 9 for functions on locally compact spaces.

REFERENCES

1. J. L. Kelley, *General Topology*, Van Nostrand, New York, N. Y., 1955.

2. R. E. Smithson, *Topologies on Sets of Relations*, Journal Nat. Sci. and Math. Government College (Lahore) to appear.

Received April 15, 1969.

UNIVERSITY OF WYOMING

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

MULTIPLICITY TYPE AND CONGRUENCE RELATIONS IN UNIVERSAL ALGEBRAS

T. P. WHALEY

For given multiplicity types μ and μ' we consider the possibility of always replacing a universal algebra $\langle A; F\rangle$ of multiplicity type μ with an algebra $\langle A; G\rangle$ of multiplicity type μ' which has exactly the same congruence relations.

In [1] Gould considered the corresponding problem for subalgebra structures. There he completely determined those types μ' which could replace a given type μ. His results were very positive; e.g., any countable type with finitely many nonzero entries can always be replaced by a type representing a single operation. We do not completely determine which types can replace a given type in the congruence ralation sense, but give necessary conditions which show that simplifications as in the subalgebra case are impossible. We also show that no two finite types are interchangeable with respect to congruence relations.

In this paper we shall be concerned only with the congruence relations of the algebras considered so we may disregard nullary operations. Thus we alter the notion of multiplicity type as follows:

DEFINITION 1.1. By the *multiplicity type* of an algebra $\mathscr{A}$ we mean the sequence $\mu = \langle \mu_1, \mu_2, \cdots, \mu_n, \rightarrow\rangle_{n \in \omega}$ where $\mathscr{A}$ has exactly μ_i operations of rank i for $i = 1, 2, \cdots$.

DEFINITION 1.2. We denote the set of all congruence relations of the algebra $\mathscr{A}$ by $\Theta(\mathscr{A})$. If a, b are elements of the algebra $\mathscr{A}$, we denote by $\Theta(a, b)$ the smallest congruence relation of $\mathscr{A}$ which contains (a, b).

DEFINITION 1.3. If $\mu = \langle \mu_1, \mu_2, \cdots, \mu_n, \rightarrow\rangle$ and $\mu' = \langle \mu'_1, \mu'_2, \cdots, \mu'_n, \rightarrow\rangle$ are sequences of cardinal numbers, we write $\mu \leqq \mu'$ provided for any algebra $\mathscr{A} = \langle A; F\rangle$ of multiplicity type μ there is an algeba $\mathscr{A}' = \langle A; F'\rangle$ of multiplicity type μ' such that $\Theta(\mathscr{A}) = \Theta(\mathscr{A}')$.

2. A necessary condition for $\mu \leqq \mu'$.

The purpose of this section is to prove the following theorem which shows that in considering congruence relations, as contrasted with subalgebras, the number of operations present is very crucial.

THEOREM 2.1. *If $\mu \leqq \mu'$, then $\Sigma \mu_i \leqq \Sigma \mu'_i$.*

To prove the theorem, we construct, for each cardinal $m \geqq 2$, an algebra $\mathscr{A}_m = \langle A_m, f_\xi \rangle_{1 \leq \xi \leq m}$ with each f_ξ a unary operation and such that $\Theta(\mathscr{A}') \neq \Theta(\mathscr{A}_m)$ if $\mathscr{A}' = \langle A_m, G \rangle$ is of type μ' where $\Sigma \mu'_i < m$.

To construct $\mathscr{A}_m$, let $A_m = \{a_\xi : 0 \leqq \xi \leqq m\} \cup \{b_\xi : 0 \leqq \xi \leqq m\}$ where the a'_ξs and b'_ξs are all distinct. Let $B_m = A_m \sim \{a_0, b_0\}$. Now if $1 \leqq \xi \leqq m$ let f_ξ be defined by $f_\xi(a_0) = a_\xi, f_\xi(b_0) = b_\xi$, and $f_\xi(x) = x$ if $x \in B_m$. Then $\mathscr{A}_m = \langle A_m; f_\xi \rangle_{1 \leq \xi \leq m}$.

LEMMA 2.2. *Letting id_x denote the identity relation on X we have*

(i) *if $x, y \in B_m$, then* $\theta(x, y) = id_{A_m} \cup \{(x, y), (y, x)\}$,

(ii) $\theta(a_0, b_0) = id_{A_m} \cup \{(a_\xi, b_\xi) : 0 \leqq \xi \leqq m\} \cup \{(b_\xi, a_\xi) : 0 \leqq \xi \leqq m\}$

(iii) *if $x \in B_m$, then*

$$\theta(a_0, x) = id_{A_m} \cup [\{x\} \cup \{a_\xi : 0 \leqq \xi \leqq m\}]^2$$

and

$$\theta(b_0, x) = id_{A_m} \cup [\{x\} \cup \{b_\xi : 0 \leqq \xi \leqq m\}]^2 .$$

Proof. Both (i) and (ii) are clear. For (iii) we consider $\theta(a_0, x)$. Now since $(a_0, x) \in \theta(a_0, x)$ we have $(f_\xi(a_0), f_\xi(x)) = (a_\xi, x) \in \theta(a_0, x)$ for each ξ. By transitivity we get $(a_\xi, a_\eta) \in \theta(a_0, x)$ for each ξ, η with $1 \leqq \xi \leqq m$ and $1 \leqq \eta \leqq m$. Thus $[\{x\} \cup \{a_\xi : 0 \leqq \xi \leqq m\}]^2 \subseteqq \theta(a_0, x)$. Since $id_{A_m} \cup [\{x\} \cup \{a_\xi : 0 < \xi \leqq m\}]^2$ is a congruence relation, the proof is completed. The claim for $\theta(b_0, x)$ follows by symmetry.

LEMMA 2.3. *Let f be a unary operation on A_m which preserves the congruence relations of $\mathscr{A}_m$* (i.e. adding f as an operation would not affect the congruence relations). *Then $f \mid B_m = id_{B_m}$* or *$f \mid B_m$ is constant.*

Proof. Let us assume that $f \mid B_m \neq id_{B_m}$. Then there is some $x \in B_m$ such that $f(x) = y \neq x$. Suppose $y \in B_m$. Let $z \in B_m, z \notin \{x, y\}$. Then $(f(z), (f(x)) = (f(z), y) \in \theta(x, z)$. By Lemma 2.2. part (i) we get $f(z) = y$. Also we have $(f(x), f(z)) = (f(z), f(y)) = (y, f(y)) \in \theta(x, y) \cap \theta(z, y)$. Thus $f(y) = y$ and $f \mid B_m$ is constant.

Now if $y \notin B_m$, we have for any $z \in B_m$ that $(f(x), f(z)) = (y, f(z)) \in \theta(x, z)$ so $f(z) = y$ and $f \mid B_m$ is constant.

LEMMA 2.4. *Let f be a unary operation on A_m which preserves the congruence relations of $\mathscr{A}_m$. Then if $f \mid B_m$* is *constant, f is constant.*

Proof. Suppose first that $f \mid B_m \equiv b$ for some $b \in B_m$. Without loss of generality we may assume that $f \mid B_m \equiv a_2$. Then $(f(a_1), f(b_0)) = (a_2, f(b_0)) \in \theta(a_1, b_0)$ so by Lemma 2.3. (iii) we have $f(b_0) = a_2$. Also,

$(f(a_0), f(a_2)) = (f(a_0), a_2) \in \theta(a_0, a_2)$ so $f(a_0) \in \{a_\xi \mid 0 \leqq \xi \leqq m\}$. Furthermore $(f(a_0), f(b_0)) = (f(a_0), a_2) \in \theta(a_0, b_0)$ giving $f(a_0) = a_2$.

Now assume that $f \mid B_m \equiv a_0$ (the case $f/B_m \equiv b_0$ follows by symmetry). Then $(f(a_1), f(a_0)) = (a_0, f(a_0)) \in \theta(a_1, a_0)$ so $f(a_0) \in \{a_\xi \mid 0 \leqq \xi \leqq m\}$. Also we have $(f(b_1), f(b_0)) = (a_0, f(b_0)) \in \theta(b_1, b_0)$ so $f(b_0) = a_0$. Finally $(f(b_0), f(a_0)) = (a_0, f(a_0)) \in \theta(b_0, a_0)$ so $f(a_0) = a_0$.

LEMMA 2.5. *If f is a unary operation on A_m which preserves the congruence relations of $\mathscr{A}_m$, then $f = id_{A_m}$, f is constant, or $f = f_\xi$ for some $\xi, 1 \leqq \xi \leqq m$.*

Proof. Assume that f is not constant. By Lemmas 2.3 and 2.4 we know that $f \mid B_m = id_{B_m}$. Thus $(f(b_1), f(b_0)) = (b_1, f(b_0)) \in \theta(b_1, b_0)$ so $f(b_0) \in \{b_\xi \mid 0 \leqq \xi \leqq m\}$. Similarly $f(a_0) \in \{a_\xi \mid 0 \leqq \xi \leqq m\}$. Since $(f(a_0), f(b_0)) \in \theta(a_0, b_0)$, we know that for some $\xi, 0 \leqq \xi \leqq m$, we have $(f(a_0), f(b_0)) = (a_\xi, b_\xi)$.

NOTATION. Suppose g is an n-ary operation on the set X. If $1 \leqq k \leqq n$, if $c_1, \cdots, c_{n-k} \in X$, if $\{1, \cdots, n\} = \{i_1, \cdots, i_{n-k}\} \cup \{j_1, \cdots, j_k\}$, then we denote by $g[i_1, \cdots, i_{n-k}; c_1, \cdots, c_{n-k}]$ the k-ary operation on X defined by

$$g[i_1, \cdots, i_{n-k}; c_1, \cdots, c_{n-k}](x_1, \cdots, x_k) = g(y_1, \cdots, y_n)$$

where

$$y_j = \begin{cases} c_s & \text{if} \quad j = i_s \\ x_j & \text{if} \quad j = j_s \,. \end{cases}$$

More informally $g[i_1, \cdots, i_{n-k}; c_1, \cdots, c_{n-k}]$ is obtained by holding each c_j fixed in the i_j coordinate of g.

REMARK. An operation preserves the congruence relations of an algebra if and only if each of its unary translations preserves the congruence relations of the algebra. Thus if g is an operation on A_m which preserves the congruence relations of $\mathscr{A}_m$, then a given unary translation of g must be the identity map, a constant map or else one of the f_ξ. It is the purpose of the next lemma to show that only one f_ξ can be so obtained from a given operation g.

LEMMA 2.6. *Let g be an n-ary operation of A_m which preserves the congruence relations of $\mathscr{A}_m$. If $g[i_1, \cdots, i_{n-1}; c_1, \cdots, c_{n-1}] = f_\xi$ for some $i_1, \cdots, i_{n-1}; c_1, \cdots, c_{n-1}, \xi$ where $1 \leqq i_j \leqq n$, $c_j \in A_m$, and $1 \leqq \xi \leqq m$, then $g[i_1, \cdots, i_{n-1}; d_1, \cdots, d_{n-1}] = f_\xi$ for each $d_1, \cdots, d_{n-1} \in A_m$.*

Proof. The proof is by induction on n. The first case to consider is $n = 2$. Without loss of generality we assume that $g[1, c] = f_1$ and show that $g[1, d] = f_1$ for each $d \in A_m$. We consider the cases $c = a_1$, $c = a_0$, and $c = a_2$. The case $c = a_j$ with $j > 2$ would be handled just as $c = a_2$, and $c = b_i$ would be symmetric to $c = a_i$.

Case 1. $c = a_1$: We are assuming now that $g(a_1, x) = f_1(x)$ for all $x \in A_m$. Since $(g(b_1, a_0), g(a_1, a_0)) = (g(b_1, a_0), a_1) \in \theta(b_1, a_1)$, we have $g(b_1, a_0) = a_1$ or $g(b_1, a_0) = b_1$. If $g(b_1, a_0) = b_1$, then $g[1, b_1] \equiv b_1$. But this would give $(g(b_1, a_2), g(a_1, a_2) = (b_1, a_2)) \in \theta(b_1, a_1)$, a contradiction. Thus $g(b_1, a_0) = a_1$. From this we see that $g[1, b_1] = f_1$ or $g[1, b_1] \equiv a_1$. However, if $g[1, b_1] \equiv a_1$, then $(g(b_1, a_2), g(a_1, a_2)) = (a_1, a_2) \in \theta(b_1, a_1)$, a contradiction. Therefore, $g[1, b_1] = f_1$.

Now since $(g(a_0, b_0), g(a_1, b_0)) = (g(a_0, b_0), b_1) \in \theta(a_0, a_1)$ gives $g(a_0, b_0) = b_1$, we have $g[1, a_0] \equiv b_1$ or $g[1, a_0] = f_1$. Noting that $(g(a_0, a_0), g(a_1, a_0)) = (g(a_0, a_0), a_1) \in \theta(a_0, a_1)$ we see that $g(a_0, a_0) \in \{a_\xi \mid 0 \leqq \xi \leqq m\}$. Thus $g[1, a_0] \neq b_1$ so $g[1, a_0] = f_1$. By symmetry we get $g[1, b_0] = f_1$.

For $2 \leqq \xi \leqq m$ we have $(g(a_\xi, b_0), g(a_1, b_0)) = (g(a_\xi, b_0), b_1) \in \theta(a_\xi, a_1)$ so $g(a_\xi, b_0) = b_1$. Thus $g[1, a_\xi] \equiv b_1$ or $g[1, a_\xi] = f_1$. Now $(g(a_\xi, a_1), g(a_1, a_1)) = (g(a_\xi, a_1), a_1) \in \theta(a_\xi, a_1)$ so $g(a_\xi, a_1) \neq b_1$. Thus $g[1, a_\xi] = f_1$, and by symmetry $g[1, b_\xi] = f_1$.

Case 2. $c = a_0$: We are assuming that $g(a_0, x) = f_1(x)$ for each $x \in A_m$. Thus we have $(g(a_1, b_0), g(a_0, b_0)) = (g(a_1, b_0), b_1) \in \theta(a_1, a_0)$. Hence $g(a_1, b_0) = b_1$ so $g[1, a_1] \equiv b_1$ or $g[1, a_1] = f_1$. If $g[1, a_1] \equiv b_1$, then $(g(a_1, a_2), g(a_0, a_2)) = (b_1, a_2) \in \theta(a_1, a_0)$, a contradiction. Thus $g[1, a_1] = f_1$, and we have Case 1.

Case 3. $c = a_2$: Now we are assuming that $g(a_2, x) = f_1(x)$ for each $x \in A_m$. Here we have $(g(a_0, b_0), g(a_2, b_0)) = (g(a_0, b_0), b_1) \in \theta(a_0, a_2)$ so $g(a_0, b_0) = b_1$. Thus $g[1, a_0] \equiv b_1$ or $g[1, a_0] = f_1$. If $g[1, a_0] \equiv b_1$, then $(g(a_0, a_2), g(a_2, a_2)) = (b_1, a_2) \in \theta(a_0, a_2)$, a contradiction. Therefore, $g[1, a_0] = f_1$, and we have Case 2. This completes the step $n = 2$ in the induction argument.

Let us assume that the lemma holds for $n = k$ and that g is $(k+1)$ary with $g[i_1, i_2, \cdots, i_k; c_1, \cdots, c_k] = f_\xi$. Without loss of generality we take $g[1, 2, \cdots, k; c_1, \cdots, c_k] = f_\xi$. Thus $g(c_1, \cdots, c_k, x) = f_\xi(x)$ for all $x \in A$. Applying the induction hypothesis to the k-ary operation $g[1, c_1]$ we get $g(c_1, d_2, d_3, \cdots, d_k, x) = f_\xi(x)$ for all $d_j \in A_m$ and all $x \in A_m$. Now we apply the case $n = 2$ to the binary operation $g[2, 3, \cdots, k; d_2, d_3, \cdots, d_k]$ to get $g(d_1, d_2, \cdots, d_k, x) = f_\xi(x)$ for arbitrary elements $d_1, \cdots, d_k$.

COROLLARY 2.7. *Let g be an n-ary operation on A_m which preserves the congruence relations of $\mathscr{A}_m$. Then all nonconstant, nonidentity unary translations of g are equal.*

Proof. Suppose $g[i_1, \cdots, i_{n-1}; c_1, \cdots, c_{n-1}] = f_\xi$. By Lemma 2.6 any unary translation of g obtained by fixing these same coordinates is equal to f_ξ. Let $\{1, \cdots, n\} = \{i_1, \cdots, i_{n-1}, j\}$. Then for any $x_1, \cdots, x_n \in A_m$ we have

$$\begin{aligned} g(x_1, \cdots, x_n) &= g[i_1, \cdots, i_{n-1}; y_1, \cdots, y_{n-1}](x_j) \\ &= f_\xi(x_j) \end{aligned}$$

where $y_k = x_{i_k}$ for $k = 1, \cdots, n-1$.

Now consider a unary translation of g obtained by fixing another set of $n-1$ coordinates; say $g[j_1, \cdots, j_{n-1}; d_1, \cdots, d_{n-1}]$ where $j = j_k$ for some $k, 1 \leqq k \leqq n-1$. Then for any $x \in A_m$

$$\begin{aligned} g[j_1, \cdots, j_{n-1}; d_1, \cdots, d_{n-1}](x) &= g(y_1, \cdots, y_n) \\ &= f_\xi(y_j) \\ &= f_\xi(y_{j_k}) \\ &= f_\xi(d_k) \end{aligned}$$

where

$$y_i = \begin{cases} d_s & \text{if} \quad i = j_s \\ x & \text{if} \quad i \notin \{j_1, \cdots, j_{n-1}\} \,. \end{cases}$$

Thus $g[j_1, \cdots, j_{n-1}; d_1, \cdots, d_{n-1}]$ is constant.

Proof of Theorem 2.1. Suppose μ and μ' are sequences with $\Sigma\mu'_i < \Sigma\mu_i = m$. Clearly $\mu \geqq \varepsilon_m = \langle m, 0, 0, \cdots\rangle$. Since $\geqq$ is transitive, it is enough to show that $\mu' \ngeqq \varepsilon_m$. Now $\mathscr{A}_m$ is an algebra of type ε_m. Suppose $\mathscr{A} = \langle A_m; G\rangle$ is an algebra of type μ' such that $\Theta(\mathscr{A}) = \Theta(\mathscr{A}_m)$. Then $\Theta(\mathscr{A}_m) = \Theta(\bar{\mathscr{A}})$ where $\bar{\mathscr{A}} = \langle A_m; \bar{G}\rangle$ with $\bar{G}$ consisting of the unary translations of G. Now by Lemma 2.5 each element of $\bar{G}$ is the identity, a constant, or one of the f'_ξs. By Lemma 2.6 at most $\Sigma\mu'_i$ of the f'_ξs can be so obtained. Let η be such that $f_\eta \notin \bar{G}$. Then $(a_\eta, b_\eta) \notin \theta(a_0, b_0)$ in $\bar{\mathscr{A}}$. This contradiction shows that $\mu' \ngeqq \varepsilon_m$ and completes the proof of Theorem 2.1.

3. Finite Types.

DEFINITION 3.1. A sequence μ is said to be finite if $\Sigma\mu_i$ is finite.

NOTATION. For a finite sequence μ, we let $l(\mu) = n$ if n is the largest interger such that $\mu_n \neq 0$.

LEMMA 3.2. *If $\mathscr{D} = \langle D; F\rangle$ is an algebra of finite multiplicity type μ and if $C = \{c_1, \cdots, c_n\}$ is a finite subset of D, then the number* *of translations of the operations of F which can be obtained by fixing only elements from C is at most*

$$P(\mu, n) = \sum_{k=0}^{l(\mu)-1}\left(\sum_{i=k+1}^{l(\mu)}\binom{i}{k}\mu_i\right)^{nk}$$

where

$$\binom{i}{k} = \frac{i!}{k!(i-k)!}$$

Proof. Suppose $1 \leqq j < k \leqq l(\mu)$ and $f \in F$ is k-ary. By fixing j elements from C in j of the coordinates of f we obtain a $(k-j)$-ary translation of f. There are $\binom{k}{j}$ ways to choose the j coordinates and n^j ways to choose the constants. Thus for fixed j and k we get $\mu_k \cdot \binom{k}{j} \cdot n^j$ translations in this way. For a fixed j we then obtain $\sum_{k=j+1}^{l(\mu)}\binom{k}{j}\mu_k^{nk}$ translations from F by fixing exactly j elements from C. Now summing on j we obtain the desired number of translations.

For a fixed sequence μ and a fixed nonnegative integer k we now construct an algebra $\mathscr{D} = \langle D; F\rangle$ of multiplicity type μ whose congruence relations can not be realized by operations of type μ' if $P(\mu', k) < P(\mu, k)$. Let $m = P(\mu, k)$, and recall from § 2 that $\mathscr{A}_m = \langle A_m; f_\xi\rangle_{1<\xi<m}$. Let $C = \{c_1, \cdots, c_k\}$ be such that c_i's are distinct and $C \cap A_m = \phi$. Now we take $D = C \cup A_m$. If g is an operation on D, by a C translation of g we mean a translation of g obtained by fixing some of the coordinates of g with elements from C. Now any application of g to elements of D may be regarded in a unique way as either an application of g to elements of C or else as an application of a C translation of g to elements of A_m (including as a C-translation g itself). We thus define F by telling what the elements of F do to elements of C and telling what the C-translations of elements of F do to elements of A_m. If $f \in F$ and F has rank l, then we shall have $f(c_{i_1}, \cdots, c_{i_l}) = c_{i_l}$ if each $c_{i_j} \in C$. Now by Lemma 3.2 there will be at most m C-translations of elements of F. Let us denote these C-translations by $\{g_\xi \mid 1 \leqq \xi \leqq m\}$. Now if g_ξ is j-ary and $x_s \in A_m$ for $1 \leqq s \leqq j$, then we take $g_\xi(x_1, \cdots, x_j) = f_\xi(x_j)$.

The following lemma is clear from the construction of $\mathscr{D}$.

LEMMA 3.3. *If $\theta \in \Theta(\mathscr{A}_m)$, then $\theta \cup id_C \in \Theta(\mathscr{D})$. Conversely, if $\theta \subseteqq (A_m)^2$ and if $\theta \cup id_C \in \Theta(\mathscr{D})$, then $\theta \in \Theta(\mathscr{A}_m)$.*

LEMMA 3.4. *If g is an n-ary operation on D which preserves the congruence relations of $\mathscr{D}$ and if $g(x_1, \cdots, x_n) = c \in C$ for some*

$x_1, \cdots, x_n \in A_m$, *then* $g(y_1, \cdots, y_n) = c$ *for all* $y_1, \cdots, y_n \in A_m$.

Proof. The pair $(g(x_1, \cdots, x_n), g(y_1, \cdots, y_n))$ is in the congruence relation of $\mathscr{D}$ generated by $\{(x_1, y_1), \cdots, (x_n, y_n)\}$. But by Lemma 3.3 this congruence relation is contained in $(A_m)^2 \cup id_C$.

LEMMA 3.5. *Let g be an n-ary operation on D which preserves the congruence relations of $\mathscr{D}$. If h is a unary translation of g, then $h \mid A_m = id_{A_m}$, $h \mid A_m$ is constant, or $h \mid A_m = f_\xi$ for some ξ, $1 \leqq \xi \leqq m$.*

Proof. By Lemma 3.4 if $h \mid A_m$ is not constant, then $(h \mid A_m)$: $A_m \rightarrow A_m$ and thus $h \mid A_m$ preserves the congruence relations of $\mathscr{A}_m$. The conclusion now follows from Lemma 2.5.

LEMMA 3.6. *Let g be an n-ary operation on D which preserves the congruence relations of $\mathscr{D}$. Then there is at most one ξ such that f_ξ is a unary translation of $g \mid (A_m)^n$.*

Proof. This follows from Lemma 3.3 and Lemma 2.6.

THEOREM 3.7. *If μ and μ' are finite sequences such that $\mu < \mu'$, then $P(\mu, k) \leqq P(\mu', k)$ for each nonnegative integer k.*

Proof. Suppose $P(\mu', k) < P(\mu, k)$. The algebra $\mathscr{D}$ is of multiplicity type μ. Suppose $\mathscr{D}' = \langle D; G\rangle$ is of multiplicity type μ' and that $\Theta(\mathscr{D}) = \Theta(\mathscr{D}')$. Let $\bar{G}$ be the set of all C-translations of elements of G (again including the elements of G). Then $\Theta(\mathscr{D}) = \Theta(\mathscr{D}'')$ where $\mathscr{D}'' = \langle D; \bar{G}\rangle$. Now in $\mathscr{D}$ we have

$$\theta(a_0, b_0) = id_D \cup \{a_i, b_i) \mid 1 < i \leqq m\}^2 .$$

Thus for each i, $1 \leqq i \leqq m$, we must have $g \in \bar{G}$ such that some unary translation of g is f_i. However, by Lemma 3.6, there is at most one i for a given $g \in \bar{G}$. Furthermore, by Lemma 3.2, the number of elements in $\bar{G}$ is $P(\mu, k) < P(\mu, k) = m$. This contradiction shows that such a G does not exist and thus concludes the proof.

REMARK. For a fixed finite sequence μ, $P(\mu, n)$ is a polynomial in n of degree $l(\mu) - 1$ having positive coefficients. The coefficient of n^k in $P(\mu, n)$ is $\sum_{i=k+1}^{l(\mu)} \binom{i}{k} \mu_i$. Hence the following corollaries follow easily from Theorem 3.7.

COROLLARY 3.8. *If μ and μ' are finite sequences such that $\mu \leqq \mu'$ then $l(\mu) \leqq l(\mu')$.*

COROLLARY 3.9. *If μ and μ' are finite sequences such that $\mu < \mu'$ and if n is the largest integer for which $\mu_n \neq \mu'_n$, then $\mu_n < \mu'_n$.*

COROLLARY 3.10. *If μ and μ' are finite sequences, then $\mu \nleqq \mu'$, $\mu' \nleqq \mu$,* or *$\mu = \mu'$. Thus among finite types, $\leqq$ is a partial ordering.*

REMARK. While a complete characterization of the relation $\leqq$ such as that given by Gould for the case of subalgebras would be of interest, it seems that the results given here indicate that such a result would not be as easily applied as is the subalgebra result. For example, in general we can not reduce the number of operations even by increasing rank.

REFERENCE

1. M. Gould, *Multiplicity type and subalgebra structure in universal algebras*, Pacific J. Math. **26** (1968), 469-485.

Received July 15, 1970. This work was partially supported by an NSF COSIP grant.

SOUTHWESTERN AT MEMPHIS

PACIFIC JOURNAL OF MATHEMATICS
Vol. 39, No. 1, 1971

GLOBALIZATION THEOREMS FOR LOCALLY FINITELY GENERATED MODULES

ROGER WIEGAND

Each commutative ring has a coreflection $\hat{R}$ in the category of commutative regular rings. We use the basic properties of $\hat{R}$ to obtain globalization theorems for finite generation and for projectivity of R-modules.

1. Preliminaries. A detailed description of the ring $\hat{R}$ may be found in [8]. Here we list without proofs the facts that will be needed. We assume that everything is unitary, but not necessarily commutative. However, R will always denote an arbitrary *commutative* ring. All unspecified tensor products are taken over R. For each $a \in R$ and each $P \in \operatorname{Spec}(R)$, let $a(P)$ be the image of a under the obvious map $R \to R_P/PR_P$. Then $\hat{R}$ is the subring $\prod_P R_P/PR_P$ consisting of finite sums of elements $[a, b]$, where $[a, b]$ is the element whose P^{th} coordinate is 0 if $b \in P$ and $a(P)/b(P)$ if $b \notin P$. There is a natural homomorphism $\varphi\colon R \to \hat{R}$ taking a to $[a, 1]$. The ring $\hat{R}$ is regular (in the sense of von Neumann). The statement that $\hat{R}$ is a coreflection means simply that each homomorphism from R into a commutative regular ring factors uniquely through φ.

The map $\operatorname{Spec}(\varphi)\colon \operatorname{Spec}(\hat{R}) \to \operatorname{Spec}(R)$ is one-to-one and onto; for each $P \in \operatorname{Spec}(R)$ we let $\hat{P}$ be the corresponding prime ($=$ maximal) ideal of $\hat{R}$.

If A is an R-module and $P \in \operatorname{Spec}(R)$, then A_P/PA_P and $(A \otimes \hat{R})_{\hat{P}}$ are vector spaces over R_P/PR_P and $\hat{R}_{\hat{P}}$ respectively. The map $\varphi\colon R \to \hat{R}$ induces an isomorphism $R_P/PR_P \cong \hat{R}_{\hat{P}}$, and, under the identification, A_P/PA_P and $(A \otimes \hat{R})_{\hat{P}}$ are isomorphic vector spaces.

2. Globalization theorems.

LEMMA. *If $A \otimes \hat{R} = 0$ and A_R is locally finitely generated then $A = 0$.*

Proof. For each prime P, $A_P/PA_P = 0$, by the last paragraph of §1. Since A_P is finitely generated over R_P, Nakayama's lemma implies that $A_P = 0$ for each $P \in \operatorname{Spec}(R)$. Therefore $A = 0$.

THEOREM 1. *Assume $(A \otimes \hat{R})$ is finitely generated over $\hat{R}$, and that A_R is either locally free or locally finitely generated. Then A_R is finitely generated.*

Proof. Assume A_R is locally free. Then, for each prime P, A_P is a direct sum of, say, κ copies of R_P. Then A_P/PA_P is a direct sum of κ copies of R_P/PR_P. But since $(A \otimes \hat{R})$ is finitely generated over $\hat{R}$, A_P/PA_P is finite dimensional over R_P/PR_P. Thus κ is finite, and we conclude that A_R is locally finitely generated.

Now, if A_R is not finitely generated, we can express A as a well-ordered union of submodules A_α, each of which requires fewer generators than A. We will get a contradiction by showing that some $A_\alpha = A$. Let $B_\alpha = \mathrm{Im}(A_\alpha \otimes \hat{R} \to A \otimes \hat{R})$. Since

$$A \otimes \hat{R} = \varinjlim_{\alpha} (A_\alpha \otimes \hat{R}) , \qquad A \otimes \hat{R} = \bigcup_{\alpha} B_\alpha .$$

Since the B_α are nested and $(A \otimes \hat{R})$ is finitely generated over $\hat{R}$, some $B_{\alpha_0} = A \otimes \hat{R}$, that is, $A_{\alpha_0} \otimes \hat{R} \twoheadrightarrow A \otimes \hat{R}$. Let $C = A/A_{\alpha_0}$. Then $C \otimes \hat{R} = \mathrm{Coker}\,(A_{\alpha_0} \otimes \hat{R} \to A \otimes \hat{R}) = 0$, and C_R is certainly locally finitely generated. By the lemma, $C = 0$, and $A_{\alpha_0} = A$.

THEOREM 2. *Let A_R be finitely generated and flat, and assume $(A \otimes \hat{R})$ is $\hat{R}$-projective. Then A_R is projective.*

Proof. By Chase's theorem [3, Theorem 4.1] it is sufficient to show that A_R is finitely related. Let $0 \to K \to F \to A \to 0$ be an exact sequence, with F_R free of finite rank. This sequence splits locally, so K is locally finitely generated. Since A_R is flat, the long exact sequence of Tor shows that $0 \to K \otimes \hat{R} \to F \otimes \hat{R} \to A \otimes \hat{R} \to 0$ is exact. This sequence splits, so $(K \otimes \hat{R})$ is finitely generated over $\hat{R}$. By Theorem 1, K_R is finitely generated.

3. Applications. The following result generalizes the well-known fact that over a noetherian ring every finitely generated flat module is projective.

PROPOSITION 1. *If R has* a.c.c. *on intersections of prime ideals then every finitely generated flat R-module is projective.*

Proof. In [8] these rings are characterized as those for which $(A \otimes \hat{R})$ is $\hat{R}$-projective for *every* finitely generated A_R. The conclusion follows from Theorem 2.

Suppose A_R is locally finitely generated. For each prime ideal P let $r_A(P)$ denote the number of generators required for A_P over R_P. By Nakayama's lemma, $r_A(P) = d_A(\hat{P})$, the dimension of $(A \otimes \hat{R})_{\hat{P}}$ as a vector space over $\hat{R}_{\hat{P}}$. Since the map $\hat{P} \to P$ is continuous, it follows that if r_A is continuous on $\mathrm{Spec}\,(R)$ then d_A is continuous on $\mathrm{Spec}\,(\hat{R})$. Using these observations we can give easy proofs of the following two theorems:

THEOREM 3 (Bourbaki [1, Th. 1]): *Assume A_R is finitely generated and flat, and that r_A is continuous. Then A_R is projective.*

THEOREM 4 (Vasconcelos [7, Prop. 1.4]): *Assume A_R is projective and locally finitely generated, and that r_A is continuous. Then A_R is finitely generated.*

Proof of Theorem 3. By Theorem 3 we may assume R is regular. A proof of Theorem 3 in this case may be found in [5], but we include one here for completeness. For each $k \geqq 0$ let

$$U_k = \{P \in \operatorname{Spec}(R) \mid r_A(P) = k\} .$$

By hypothesis the sets U_k are clopen, and we let e_k be the idempotent with support U_k. Then $A = A\,e_0 \oplus \cdots \oplus A\,e_n$, and $r_{A e_k}$ is constant on $\operatorname{Spec}(Re_k)$. Therefore we may assume r_A is constant on $\operatorname{Spec}(R)$, say $r_A(P) = n$ for all P. Given a prime P, choose $a_1, \cdots, a_n \in R$ such that $a_1(P), \cdots, a_n(P)$ span A_P. Then $a_1(Q), \cdots, a_n(Q)$ span R_Q for all Q in some neighborhood of P. (Here we need A_R finitely generated.) In this way we get a partition of $\operatorname{Spec}(R)$ into disjoint clopen sets $V_1, \cdots, V_m$ together with elements $a_{ij} \in R$ such that $a_{ij}(P), \cdots, a_{nj}(P)$ span A_P for each $P \in V_j$. Let e_j be the idempotent with support V_j, and set $b_i = \Sigma_j e_j a_{ij}$. Then, if P_R is free on $u_1, \cdots, u_n$, the map $P \to A$ taking u_i to b_i is an isomorphism locally, and therefore globally.

Proof of Theorem 4. By Theorem 1 and the proof of Theorem 3 we can assume R is regular and $r_A(P) = n$ for all P. Write $A = \bigoplus \sum_{i \in I} Re_i$, $e_i^2 = e_i \neq 0$, by [4]. Given $P \in \operatorname{Spec}(R)$, since $(Re_i)_P$ is 0 if $e_i \in P$ and R_P if $e_i \notin P$, we see that there are precisely n indices i for which $e_i \notin P$. For each n-element subset $J \subseteqq I$ let

$$U(J) = \{P \in \operatorname{Spec}(R) \mid e_j \notin P \text{ for each } j \in J\} .$$

These open sets cover $\operatorname{Spec}(R)$, so $\operatorname{Spec}(R) = U(J_1) \cup \cdots \cup U(J_m)$. If $j \notin J_1 \cup \cdots \cup J_m$ then e_j is in every prime ideal, contradicting $e_j \neq 0$. Therefore $|I| \leqq mn$, and A_R is finitely generated.

As a final application we give the following:

PROPOSITION 2. *Let $0 \to A \to B \to C \to 0$ be an exact sequence of flat R-modules Assume A_R is finitely generated and $(B \otimes \hat{R})_{\hat{R}}$ is projective. Then A_R is projective.*

Proof. Since C_R is flat, $0 \to A \otimes \hat{R} \to B \otimes \hat{R} \to C \otimes \hat{R} \to 0$ is exact. Since $\hat{R}$ is semihereditary $(A \otimes \hat{R})$ is R-projective. By Theorem 2, A_R is projective.

If B_R is projective this proposition contains no new information. (In fact, a trivial extension of Chase's Theorem shows that the sequence splits.) On the other hand, if we let M_R be projective, take $f \in R$, and let $B = M_f = \{[m/f^n]\}$, then B_R is not in general projective; but by the second corollary to Theorem 5 (next section), $B \otimes \hat{R}$ is $\hat{R}$-projective.

4. Epimorphisms. Suppose M is a multiplicative subset of R, and let $S = M^{-1}R$. Since $S \otimes \hat{R}_{\hat{P}} = S_P/PS_P$ for each prime P, we see that $S \otimes \hat{R}_{\hat{P}}$ is $\hat{R}_{\hat{P}}$ if $P \cap M = \varnothing$, and 0 if $P \cap M \neq \varnothing$. If we could show that $(S \otimes \hat{R})_{\hat{R}}$ is finitely generated, it would follow easily that $S \otimes \hat{R} = \hat{R}/K$, where K is the intersection of those primes $\hat{P}$ for which $P \cap M = \varnothing$. We give an indirect proof of this fact in a more general setting.

Suppose R and S are commutative rings and that $\alpha\colon R \to S$ is an epimorphism in the category of rings. By a theorem of Silver [6] this is equivalent to the natural map $S \otimes S \to S$ being an isomorphism. It is known [8] that $R \to \hat{R}$ is an epimorphism, and it follows readily that the natural maps $f\colon S \to S \otimes \hat{R}$ and $g\colon R \to S \otimes \hat{R}$ are epimorphisms.

THEOREM 5. *Let R and S be commutative rings and let $\alpha\colon R \to S$ be an epimorphism in the category of rings. Then there is a unique ring homomorphism $\beta\colon \hat{S} \to S \otimes \hat{R}$ making the following diagram commute*:

$$\begin{tikzcd} & \hat{S} \arrow[dd, "\beta"] \\ \hat{R} \arrow[ur, "\hat{\alpha}"] \arrow[dr, "g"'] & \\ & S \otimes \hat{R} \end{tikzcd}$$

Moreover, β is an isomorphism, and $\hat{\alpha}$ and g are surjections with kernel $K = \cap\{\hat{P} \mid S_P \neq PS_P\}$.

Proof. We first show that $S \otimes \hat{R}$ is regular. Suppose A and B are $(S \otimes \hat{R})$-modules. Then by Silver's Theorem $B = S \otimes_R B$, and by [2, p. 165] we have

$$A \otimes_{S\otimes\hat{R}} B = A \otimes_{S\otimes\hat{R}} (S \otimes_R B) = (A \otimes_S S) \otimes_{\hat{R}\otimes R} B = A \otimes_{\hat{R}} B .$$

It follows that tensor products over $S \otimes \hat{R}$ are exact, and therefore

$S \otimes \hat{R}$ is regular. Hence there is a unique map $\beta: \hat{S} \to S \otimes \hat{R}$ such that $\beta\varphi_S = f$, where $\varphi_S: S \to \hat{S}$ is the natural map. Consider the diagram:

$$\begin{tikzcd}
R \arrow[rr, "\alpha"] \arrow[dd, "\varphi_R"'] & & S \arrow[dl, "f"'] \arrow[dd, "\varphi_S"] \\
& S \otimes \hat{R} \arrow[dr, shift left, "\gamma"] & \\
\hat{R} \arrow[ur, "g"] \arrow[rr, "\hat{\alpha}"'] & & \hat{S} \arrow[ul, shift left, "\beta"]
\end{tikzcd}$$

Here γ is defined by the equations $\gamma f = \varphi_S$, $\gamma g = \hat{\alpha}$. Now $\gamma\beta\varphi_S = \gamma f = \varphi_S$ and $\beta\gamma f = \beta\varphi_S = f$. Since φ_S and f are both epimorphisms, we see that $\gamma = \beta^{-1}$. Also, $B\hat{\alpha} = B\gamma g = g$, as required. Uniqueness of β follows from the fact that $\hat{\alpha}$ is an epimorphism (since both α and φ_S are).

Next, we show $\hat{\alpha}$ is onto. To simplify notation, we assume R is regular and $\alpha: R \to S$ is an epimorphism. Then $S \otimes S \xrightarrow{\mu} S$ is an isomorphism. But then $S_P \otimes_{R_P} S_P \to S_P$ is an isomorphism for each $P \in \operatorname{Spec}(R)$. If $s \in S_P$ then $1 \otimes s - s \otimes 1 \in \ker\ \mu_P = 0$. It follows that the dimension of S_P as a vector space over R_P is either 0 or 1. Therefore α_P is surjective for each P, $(\alpha(1) = 1)$, and we conclude that α is surjective.

Finally, we compute $\ker g = K$. If $P \in \operatorname{Spec}(\hat{R})$, then

$$K \subseteqq \hat{P} \Longleftrightarrow K_{\hat{P}} = 0 \Longleftrightarrow \hat{S}_{\hat{P}} \neq 0 \Longleftrightarrow S \otimes \hat{R}_{\hat{P}} \neq 0 \Longleftrightarrow S_P/PS_P \neq 0 \ .$$

COROLLARY 1. *Let M be a multiplicative subset of R and let $S = M^{-1}R$. Then $S \otimes \hat{R}$ is a cyclic $\hat{R}$-module, and $S \otimes \hat{R}$ is $\hat{R}$-projective if and only if $\{\hat{P} \mid M \cap P \neq \varnothing\}$ is closed in* Spec $(\hat{R})$.

Proof. Let K be as in Theorem 5. Then $S \otimes \hat{R} = \hat{R}/K$ is $\hat{R}$-projective if and only if K is a principal ideal, that is, if and only if the set of primes containing K is open in Spec $(\hat{R})$. But

$$\hat{P} \supseteqq K \Longleftrightarrow PS_P \neq S_P \Longleftrightarrow M \cap P = \varnothing \ .$$

The next corollary shows that Theorem 2 is false if A_R is not assumed to be finitely generated.

COROLLARY 2. *For each $f \in R$, $R_f \otimes \hat{R}$ is $\hat{R}$-projective.*

Proof. Set $M = \{f^n : n \geqq 0\}$. Then $P \cap M \neq \emptyset$ if and only if $\varphi(f) \in \hat{P}$. Thus K is the principal ideal of $\hat{R}$ generated by $\varphi(f)$, and $\hat{R}/K$ is $\hat{R}$-projective.

REFERENCES

1. N. Bourbaki, *Eléments de mathématique. Algèbre commutative,* Chapter II, Hermann, Paris, 1961.
2. H. Cartan and S. Eilenberg, *Homological algebra,* Princeton Univ. Press, Princeton, 1956.
3. S. U. Chase, "Direct products of modules", Trans. Amer. Math. Soc., **97** (1960), 457-473.
4. I. Kaplansky, Projective modules, Ann. of Math., **68** (1958), 372-377.
5. R. S. Pierce, *Modules over commutative regular rings,* Memoirs Amer. Math. Soc., **70** (1967).
6. L. Silver, *Non-commutative localizations and applications,* J. Algebra, **7** (1967), 44-76.
7. W. Vasconcelos, *On projective modules of finite rank,* Proc. Amer. Math. Soc., **22**, No. 2 (1969), 430-433.
8. R. Wiegand, *Modules over universal regular rings,* (to appear in the Pacific J. Math.).

Received March 19, 1971. This research was partially supported by NSF Grant GP-19102.

UNIVERSITY OF WISCONSIN

Pacific Journal of Mathematics
Vol. 39, No. 1 May, 1971