Pacific Journal of Mathematics

IN THIS ISSUE—

Vol. 49, No. 1 **May, 1973**

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

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

FREE LATTICE-ORDERED MODULES

A. BIGARD

The aim of this paper is to show that the theory of free lattice-ordered groups developed by E. C. Weinberg in the abelian case can be generalized to modules over a totally ordered Ore domain A. The main result is that for every torsion-free ordered A-module M, there exists a free f-module over M. The generalization given will be seen to be, in a certain sense, the best possible.

All rings and modules considered will be assumed to be unital. Let A be a partially ordered ring and A_+ its order. If M is a left A-module, we say that $P \subseteqq M$ is an order on M if:

$P + P \subseteqq P$, $A_+P \subseteqq P$, and $P \cap -P = \{0\}$. If such a P is given, we say that M is a partially ordered (or ordered) module. If P is a total order on M, that is, if $M = P \cup -P$, we say that M is a totally ordered module. Let M and N be partially ordered A-modules and let f be a mapping from M to N. Then f is an o-homomorphism if f is a monotonic homomorphism of A-modules. The o-homomorphism f is an o-isomorphism if f is one-to-one and if f^{-1} is an o-homomorphism.

1. Some properties of f-modules.

In this section, A will denote a directed p.o. ring. An A-module M which is lattice-ordered by the order P is called a lattice-ordered module or l-module. Products of lattice-ordered modules are defined in a natural way. If M and N are l-modules, an homomorphism f from M to N is called a l-homomorphism if, for $x, y \in M$:

$$f(x \vee y) = f(x) \vee f(y) \quad \text{and} \quad f(x \wedge y) = f(x) \wedge f(y) .$$

An f-module M is a lattice-ordered module which is a subdirect product of totally ordered modules. This definition was first introduced in [1] and [3].

Recall that a convex l-subgroup S in a commutative l.o. group G is called prime if G/S is totally ordered. The following theorem gives useful characterizations of f-modules.

THEOREM 1. *Let M be a lattice-ordered module over a unital directed ring A. The following are equivalent*:

(1) *M is an f-module.*

(2) *For $x, y \in M$ and $0 \leqq \lambda \in A$, $\lambda(x \vee y) = \lambda x \vee \lambda y$ and $\lambda(x \wedge y) = \lambda x \wedge \lambda y$.*

(3) $x \wedge y = 0$ *implies* $\lambda x \wedge y = 0$ *for all* $0 \leqq \lambda \in A$.
(4) *Every minimal prime subgroup of* M *is a submodule.*

Proof. (1) implies (2): This is clear since (2) is satisfied in a totally ordered module.

(2) implies (3): If $x \wedge y = 0$, then we have:

$$0 \leqq \lambda x \wedge y \leqq (\lambda \vee I)x \wedge (\lambda \vee I)y = (\lambda \vee I)(x \wedge y) = 0 .$$

(3) implies (4): Let S be a minimal prime subgroup. Then, $x \in S$ if and only if there exists $y \notin S$ with $x \wedge y = 0$. [2]. Thus, if $x \in S$ and $0 \leqq \lambda \in A$, we have $\lambda x \in S$. Since A is directed, S is a submodule.

(4) implies (1): Let $(S_i)_{i \in I}$ be the family of all minimal prime subgroups of M. Then each quotient M/S_i is a totally ordered module and M is a subdirect product of these modules.

If A is not unital, then (1), (3), and (4) are equivalent but condition (2) is weaker (see [3]).

In the sequal, we shall be concerned mainly with torsion-free modules, that is modules in which $\lambda x = 0$ implies $\lambda = 0$ or $x = 0$. The following property is useful:

PROPOSITION 1. *If A is totally ordered, every torsion-free f-module F is a subdirect product of torsion-free totally ordered modules.*

Let S be a minimal prime subgroup of F. Suppose that $\lambda \neq 0$ and $\lambda x \in S$. We may assume $\lambda > 0$, as A is totally ordered. As in the proof of Theorem 1, there exists $y \notin S$ with $\lambda x \wedge y = 0$. This implies $\lambda(x \wedge y) = \lambda x \wedge \lambda y = 0$, and hence $x \wedge y = 0$. As $y \notin S$, we obtain $x \in S$. This proves that M/S is torsion-free and the theorem follows.

As in the theory of ordered groups, P is an isolated order on M if $\lambda > 0$ and $\lambda x \in P$ implies $x \in P$.

PROPOSITION 2. *Every torsion-free f-module is isolated.*

Proof. If $\lambda > 0$ and $\lambda x \geqq 0$, we have $\lambda(-x \vee 0) = -\lambda x \vee 0 = 0$, hence $-x \vee 0 = 0$ and $x \geqq 0$.

Conversely, it is clear that when A is totally ordered, every isolated module is torsion-free.

2. Embedding an order in a total order. In this section, we consider only torsion-free modules over a totally ordered unital ring A. This is not as restrictive as it seems, since the existence of a nontrivial torsion-free module implies that A has no zero divisors, and an f-ring with no zero divisors is totally ordered.

LEMMA 1. *Let M be a torsion-free A-module. For every $x \in M$, A_+x is an order.*

Proof. Suppose that $y \in A_+x \cap -A_+x$, so that $y = \lambda x = -\mu x$. The relation $(\lambda + \mu)x = 0$ implies $\lambda + \mu = 0$ or $x = 0$. In the first case, $\lambda = -\mu \in A_+ \cap -A_+$ so in each case $y = 0$.

LEMMA 2. *Let P and Q be two orders on M. Then $P - Q$ is an order if and only if $P \cap Q = 0$.*

Proof. The condition is necessary, since $P \cap Q \subseteqq (P-Q) \cap (Q-P)$. For the converse, suppose $P \cap Q = 0$ and let $y \in (P - Q) \cap (Q - P)$. Then $y = p - q = q' - p'$, and $p + p' = q + q' \in P \cap Q = 0$. Hence, $p = -p' \in P \cap -P = 0$, $q = -q' \in Q \cap -Q = 0$, and it follows that $y = 0$.

The ring A is said to be a left Ore domain if A admits a left quotient field. Equivalently, A has no zero divisors and satisfies the following condition:

(I) If $\rho \neq 0$ and $\sigma \neq 0$, $A\rho \cap A\sigma \neq 0$.

Clearly, when A is totally ordered, this condition can be replaced by the following:

(II) If $0 < \rho$ and $0 < \sigma$, $A_+\rho \cap A_+\sigma \neq 0$.

THEOREM 2. *Let A be a totally ordered ring with no divisors of zero. The following are equivalent:*

(1) *A is a left Ore domain.*

(2) *In a torsion-free A-module, every order is contained in a total order.*

(3) *In a torsion-free A-module, every order is contained in an isolated order.*

Proof. (1) implies (2): By Zorn's lemma, every order is contained in a maximal order. It remains to show that each maximal order P is total. If not, suppose $b \notin P \cup -P$. As $P \subset P + A_+b$ (strictly), $P + A_+b$ fails to be an order. By Lemma 2, $P \cap -A_+b \neq 0$ and there exists $\rho > 0$ with $\rho b \in -P$. Similarly, $P - A_+b$ is not an order, $P \cap A_+b \neq 0$, and there exists $\sigma > 0$ with $\sigma b \in P$. By condition (II), there exists $\lambda > 0$ and $\mu > 0$ with $\lambda\rho = \mu\sigma > 0$. Hence $\lambda\rho b = \mu\sigma b \in P \cap -P = 0$. This implies $b = 0$, which is a contradiction. Hence P is a total order.

(2) implies (3): This is clear from Proposition 2.

(3) implies (1): Consider A as a left-module on itself. Take $0 < \rho$ and $0 < \sigma$. If $A_+\rho \cap A_+\sigma = 0$, $A_+\rho - A_+\sigma$ is an order by Lemma 2. Hence it is contained in an isolated order P, and thus $\rho 1 \in P$ and $\sigma(-1) \in P$. Then $1 \in P$ and $-1 \in P$, which is a contradiction.

COROLLARY 1. *Let A be a totally ordered left Ore domain. Let f be an o-homomorphism of the torsion-free module M ordered by P into a torsion-free totally ordered module T. There exists a total order P_0 which contains P such that $f(x) > 0$ implies $x \in P_0$.*

To see that $S = \{x \mid f(x) > 0\} \cup \{0\}$ is an order on M, note that $S + S \subseteqq S$ and $S \cap -S = \{0\}$. Also for $\lambda > 0$ and $0 \neq x \in S$, $f(x) > 0$ and hence $f(\lambda x) = \lambda f(x) > 0$ since T is torsion-free. As $P \cap -S = 0$, $P + S$ is an order by Lemma 2. The corollary then follows from Theorem 1.

COROLLARY 2. *Let A be a totally ordered left Ore domain and let M be a torsion-free A-module ordered by P. The intersection of all total orders containing P is the set $\bar{P}$ of elements $x \in M$ for which there exists $\lambda > 0$ with $\lambda x \in P$.*

Each total order containing P is isolated and hence contains $\bar{P}$. Suppose $x \notin \bar{P}$, so that $P \cap A_+x = 0$. By Lemma 2, $P - A_+x$ is an order. By Theorem 2, $P - A_+x$ is contained in a total order Q. Since $-x \in Q$ and $x \neq 0$, $x \notin Q$.

THEOREM 3. *Let A be a totally ordered left Ore domain. If M is an A-module ordered by P, these are equivalent:*

(1) *P is isolated.*

(2) *M is torsion-free and P is an intersection of total orders.*

(3) *M can be embedded in a direct product of totally ordered torsion-free modules.*

(4) *M can be embedded in a torsion-free f-module.*

Proof. (1) implies (2): This follows directly from Corollary 2, as $P = \bar{P}$.

(2) implies (3): Let $(P_i)_{i \in I}$ be the set of all total orders containing P. If we denote by M_i the module M ordered by P_i, there is a canonical embedding of M into the direct product of the modules M_i.

(3) implies (4): Clear.

(4) implies (1): This follows from Proposition 1.

3. Free f-modules. Let A be a totally ordered left Ore domain, and let M be a torsion-free A-module ordered by P. A torsion-free f-module L will be called *free over* M if:

(1) There exists an injective o-homomorphism φ from M to L.

(2) For every torsion-free f-module F and every o-homomorphism f from M to F, there exists a unique l-homomorphism $\bar{f}$ from L to F such that $\bar{f} \circ \varphi = f$.

It is not difficult to show that L is determined up to an l-isomorphism. To show that such an L exists, we use the two following lemmas:

LEMMA 3. *If $x_{\alpha\beta} (\alpha \in R, \beta \in S)$ and $x_{\gamma\delta} (\gamma \in U, \delta \in V)$ are two finite families of elements in a lattice-ordered module,*

$$\bigvee_{\alpha \in R} \bigwedge_{\beta \in S} x_{\alpha\beta} - \bigvee_{\gamma \in U} \bigwedge_{\delta \in V} x_{\gamma\delta} = \bigvee_{(\alpha,\sigma) \in R \times (V^{S \times U})} \bigwedge_{(\beta,\gamma) \in S \times U} (x_{\alpha\beta} - x_{(\gamma)(\sigma(\beta,\gamma))}) .$$

Proof.

$$\bigvee_R \bigwedge_S x_{\alpha\beta} - \bigvee_U \bigwedge_V x_{\gamma\delta} = \bigvee_R \bigwedge_S \bigwedge_U \bigvee_V (x_{\alpha\beta} - x_{\gamma\delta}) = \bigvee_R \bigwedge_{S \times U} \bigvee_V (x_{\alpha\beta} - x_{\gamma\delta})$$
$$= \bigvee_R \bigvee_{\sigma \in (V^{S \times U})} \bigwedge_{S \times U} (x_{\alpha\beta} - x_{(\gamma)(\sigma(\beta,\gamma))})$$
$$= \bigvee_{(\alpha,\sigma) \in R \times (V^{S \times U})} \bigwedge_{(\beta,\gamma) \in S \times U} (x_{\alpha\beta} - x_{(\gamma)(\sigma(\beta,\gamma))}) .$$

LEMMA 4. *Let N be a f-module and K a submodule of N. The f-submodule generated by K is the set K' of all elements $\bigvee_{\alpha \in R} \bigwedge_{\beta \in S} x_{\alpha\beta}$ with $x_{\alpha\beta} \in K$.*

Proof. By Lemma 3, K' is an l-subgroup of N. If $\lambda \geqq 0$, it follows from Theorem 1 that: $\lambda \bigvee_R \bigwedge_S x_{\alpha\beta} = \bigvee_R \bigwedge_S \lambda x_{\alpha\beta}$. Since the ring is assumed to be directed, K' is a submodule.

THEOREM 4. *Let $(P_i)_{i \in I}$ be the set of all total orders on M containing P, and denote by M_i the module M ordered by P_i. Let φ be the canonical map of M into $\prod_{i \in I} M_i$. Then the f-submodule L of $\prod_{i \in I} M_i$ generated by $\varphi(M)$ is free over M.*

Proof. Suppose f is an o-homomorphism from M into a torsion-free f-module F. If $x \in L$, then by Lemma 4, $x = \bigvee_R \bigwedge_S \varphi(x_{\alpha\beta})$ where $x_{\alpha\beta} \in M$.

Let $\bar{f}(x) = \bigvee_R \bigwedge_S f(x_{\alpha\beta})$. To show that $\bar{f}$ is a mapping, it is sufficient to show, by Lemma 3, that $\bigvee_R \bigwedge_S \varphi(x_{\alpha\beta}) = 0$ implies $\bigvee_R \bigwedge_S f(x_{\alpha\beta}) = 0$.

By Proposition 1, we may assume that F is totally ordered. By

Corollary 1 of Theorem 2, there exists a total order P_0 containing P such that $f(x) > 0$ implies $x \in P_0$.

If $\bigvee_R \bigwedge_S f(x_{\alpha\beta}) > 0$, there exists $\alpha \in R$ such that for each $\beta \in S$, $f(x_{\alpha\beta}) > 0$, which implies $x_{\alpha\beta} \in P_0$. It follows that $\bigvee_R \bigwedge_S x_{\alpha\beta} > 0$ (modulo P_0) and $\bigvee_R \bigwedge_S \varphi(x_{\alpha\beta}) \neq 0$. Alternatively, if $\bigvee_R \bigwedge_S f(x_{\alpha\beta}) < 0$, there exists for each $\alpha \in R$, a $\beta \in S$ such that $f(x_{\alpha\beta}) < 0$. Thus $x_{\alpha\beta} \in -P_0$ and it follows that $\bigvee_R \bigwedge_S x_{\alpha\beta} < 0$ (with respect to P_0). Hence $\bigvee_R \bigwedge_S \varphi(x_{\alpha\beta}) \neq 0$. Now, it is clear that $\bar{f}$ is a mapping. By Lemma 3, $\bar{f}$ is a group homomorphism. The theorem follows easily.

COROLLARY. *Let A be a totally ordered ring with no divisors of zero. The following are equivalent*:

(1) *A is a left Ore domain.*

(2) *For every torsion-free ordered module M, there exists a free f-module over M.*

Proof. By Theorem 4, (1) implies (2). Conversely, if φ is the o-homomorphism of M into the free f-module L over M, the positive cone of M is a subset of $Q = \{x \mid \varphi(x) \geqq 0\}$, which is an isolated order. Thus, (2) implies (1) by Theorem 2.

Note that φ is an o-isomorphism of M into L if and only if M is isolated.

It is now easy to construct the free f-module over an arbitrary set E. Let M be the free module generated by E, and trivially order M by $P = \{0\}$. The free f-module L generated by M is a free f-module over E, with obvious definitions.

REFERENCES

1. A. Bigard, *Contribution à la théorie des groupes réticulés,* Thèse Sc. Math., Paris, 1969.

2. P. Conrad, *Introduction à la théorie des groupes réticulés,* Secrétariat mathématique, Paris, 1967.

2. S. A. Steinberg, *Lattice-ordered rings and modules*, Ph. D. Thesis, Urbana, Illinois, 1970.

4. E. C. Weinberg, *Free lattice-ordered abelian groups,* Math. Annalen t, **151** (1963), 187-199 and t. **159** (1965), 217-222.

Received July 19, 1972 and in revised form May 4, 1973.

CENTRE UNIVERSITAIRE DU MANS

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

SUBNORMAL OPERATORS IN STRICTLY CYCLIC OPERATOR ALGEBRAS

RICHARD BOLSTEIN AND WARREN WOGEN

It is shown that a subnormal operator cannot belong to a strictly cyclic and separated operator algebra unless it is normal and has finite spectrum. Further, a subnormal operator not of this type cannot have a strictly cyclic commutant.

1. Let $\mathscr{H}$ be a complex Hilbert space, and let $\mathscr{A}$ be a subset of the algebra $\mathscr{B}(\mathscr{H})$ of all bounded linear operators on $\mathscr{H}$. A vector $x \in \mathscr{H}$ with the property that $\mathscr{A}x = \{Ax: A \in \mathscr{A}\}$ is the full Hilbert space is said to be a *strictly cyclic vector* for $\mathscr{A}$, and $\mathscr{A}$ is said to be *strictly cyclic* if such a vector exists. A vector x is called a *separating vector* for $\mathscr{A}$ if no two distinct operators in $\mathscr{A}$ agree at x. The set $\mathscr{A}$ is said to be *strictly cyclic and separated* if there is a vector x which is both strictly cyclic and separating for $\mathscr{A}$.

Strictly cyclic operator algebras have recently been investigated by Mary Embry [2] and Alan Lambert [3]. Let $\mathscr{A}'$ denote the *commutant* of the set $\mathscr{A}$, that is, $\mathscr{A}'$ is the set of all bounded linear operators which commute with every operator in $\mathscr{A}$. Note that if x is a *cyclic vector* for $\mathscr{A}$ (meaning $\mathscr{A}x$ is dense in $\mathscr{H}$), then x is separating for $\mathscr{A}'$.

LEMMA 1. *Let $\mathscr{A}$ be a strictly cyclic subset of $\mathscr{B}(\mathscr{H})$. If $\mathscr{A}$ is abelian, then it is maximal abelian, $\mathscr{A} = \mathscr{A}'$. Thus, a strictly cyclic abelian subset is automatically a weakly closed algebra.*

This lemma, which indicates the severity of the condition of strict cyclicity, is a sharper form of a result of Lambert [3].

Proof. Let x be strictly cyclic for $\mathscr{A}$, and let $B \in \mathscr{A}'$. Then there exists $A \in \mathscr{A}$ such that $Ax = Bx$. But $\mathscr{A} \subset \mathscr{A}'$ by hypothesis, so $A \in \mathscr{A}'$. Since x is separating for $\mathscr{A}'$, we have $B = A \in \mathscr{A}$, and the proof is complete.

If $\mathscr{A}$ is strictly cyclic and abelian, then it is strictly cyclic and separated by Lemma 1. Mary Embry [2] showed that the converse holds if $\mathscr{A}$ is the commutant of a single operator. Thus, if A is normal and $\{A\}'$ is strictly cyclic and separated, then $\{A\}'$ consists of normal operators by Fuglede's theorem. In a private communication to the authors, Mary Embry asked if "normal" could be replaced by "subnormal" in this statement. An operator is called *subnormal* if it is the restriction of a normal operator to an invariant subspace. To this end, we show that if A is subnormal then strict cyclicity of $\{A\}'$ already forces A to be normal, and, moreover, its spectrum is a finite set. Thus, the commutant of a subnormal operator cannot be strictly cyclic and separated unless the underlying Hilbert space is finite-dimensional (since the commutant is then abelian and hence the operator, which is normal, must have simple spectrum). More generally, it is shown that a uniformly closed subalgebra $\mathscr{A}$ of $\mathscr{B}(\mathscr{H})$ which has a separating vector x with the property that $\mathscr{A}x$ is a closed subspace of $\mathscr{H}$ (this is the case if x is also strictly cyclic) contains no subnormal operators except possibly for normal operators with finite spectrum.

2. Let μ be a finite positive Borel measure in the plane with compact support X, let $H^2(\mu)$ be the closure of the polynomials in $L^2(\mu)$, and put $H^\infty(\mu) = H^2(\mu) \cap L^\infty(\mu)$. The next theorem, which is used to derive the main result, may be of independent interest.

THEOREM 1. *$H^\infty(\mu) = H^2(\mu)$ if, and only if, X is finite.*

Proof. The sufficiency is trivial. Assume now that X is infinite. Note that the inclusion map of $H^\infty(\mu)$ into $H^2(\mu)$ is continuous. We will show that the inverse map is not continuous, and hence, by the Open Mapping Theorem, that $H^\infty(\mu) \neq H^2(\mu)$.

Since X is compact and infinite, its set X' of accumulation points is compact and nonempty. Choose $\lambda_0 \in X'$ such that $|\lambda_0| = \max\{|\lambda|: \lambda \in X'\}$, and let $D_1 = \{\lambda: |\lambda| \leqq |\lambda_0|\}$. By the choice of λ, $X \backslash D_1$ is a countable set. Therefore, we can choose a closed disk D_2 which contains D_1 and is tangent to D_1 at λ_0, in such a way that the boundary of D_2 intersects X only at λ_0. Now note that we may as well assume that D_2 is the closed unit disc Δ, and that $\lambda_0 = 1$.

Now $X \backslash \Delta$ is a countable set $\{y_1, y_2, \cdots\}$, and if this set infinite, we must have $\lim y_n = 1$. Let $K = \Delta \cup (X \backslash \Delta)$. Then K is a compact set which does not separate the plane. Define a sequence of functions $\{f_n\}$ on K by

$$f_n(z) = \begin{cases} z^n: & z \in \Delta \\ 0: & z = y_i, 1 \leqq i \leqq n \\ 1: & z = y_i, i > n \,. \end{cases}$$

Then, for each n, f_n is continuous on K and analytic in its interior. By Mergelyan's theorem, each f_n is the uniform limit on K of a sequence of polynomials. Hence each $f_n \in H^\infty(\mu)$.

Let χ denote the function which has the value 1 at the point 1 and the value zero elsewhere. Clearly, $f_n \to \chi$ pointwise, and hence in the metric of $L^2(\mu)$ by dominated convergence. In particular, $\chi \in H^\infty(\mu)$. However, the point 1 is an accumulation point of the support of μ, and hence $\|f_n - \chi\|_\infty = 1$ for every n. Thus, $\{f_n\}$ converges to χ in $H^2(\mu)$ but not in $H^\infty(\mu)$.

THEOREM 2. *Let S be a subnormal operator on the Hilbert space $\mathscr{H}$, let $\mathscr{A}$ be the uniformly closed algebra generated by S. If $\mathscr{A}$ has a separating vector x such that $\mathscr{A}x$ is a closed subspace of $\mathscr{H}$, then the spectrum of S is a finite set, and hence S is normal.*

Proof. Let $\mathscr{B}$ be the uniformly closed algebra generated by S and the identity operator I. Since $\mathscr{B}x$ is the sum of $\mathscr{A}x$ and the one-dimensional space spanned by x, and since we assume that $\mathscr{A}x$ is closed, we also have that $\mathscr{B}x$ is a closed subspace of $\mathscr{H}$.

Now $\mathscr{B}x$ is invariant under S and the restriction operator $S_0 = S \mid \mathscr{B}x$ is subnormal. Since the uniformly closed algebra $\mathscr{B}_0$ generated by S_0 and I contains $\mathscr{B} \mid \mathscr{B}x$, it follows that x is a strictly cyclic vector for $\mathscr{B}_0$, that is, $\mathscr{B}_0 x = \mathscr{B}x$. By the representation theorem for subnormal operators with a cyclic vector, Bram [1], S_0 is unitarily equivalent to the operator of multiplication by the identity function on some $H^2(\mu)$ space. Furthermore, the unitary equivalence can be constructed so that x corresponds to the constant function 1.

Now $\mathscr{B}_0$ corresponds via the unitary equivalence to the algebra of multiplication operators $M_\phi : f \to \phi f$ on $H^2(\mu)$, where ϕ belongs to the $L^\infty(\mu)$-closure of the polynomials. Since any such function ϕ belongs to $H^\infty(\mu)$, it follows that the constant function 1 is a strictly cyclic vector for $\{M_\phi : \phi \in H^\infty(\mu)\}$, and hence that $H^\infty(\mu) = H^2(\mu)$. By Theorem 1, $H^2(\mu)$ is finite-dimensional.

It follows that $\mathscr{B}x$ is finite-dimensional, and, since $\mathscr{A} \subset \mathscr{B}$, so is $\mathscr{A}x$. Since x separates $\mathscr{A}$, it follows that $\mathscr{A}$ is finite-dimensional. So there is a polynomial p such that $p(S) = 0$. Since $p(\sigma(S)) = \sigma(p(S)) = \{0\}$, $\sigma(S)$ in finite and hence S is normal.

COROLLARY 1. *Let $\mathscr{A}$ be a uniformly closed subalgebra of $\mathscr{B}(\mathscr{H})$ which has a separating vector x such that $\mathscr{A}x$ is a closed subspace of $\mathscr{H}$. (This is the case if $\mathscr{A}$ is strictly cyclic and separated.) Then $\mathscr{A}$ contains no subnormal operator with infinite spectrum.*

Proof. Suppose $S \in \mathscr{A}$ is subnormal, and let $\mathscr{A}(S)$ be the uniformly closed algebra generated by S. Since $\mathscr{A}(S) \subset \mathscr{A}$, x separates $\mathscr{A}(S)$. Since the linear transformation $A \to Ax$ of $\mathscr{A}$ onto $\mathscr{A}x$ is continuous and one-to-one, and since $\mathscr{A}x$ is closed by hypothesis, the transformation has a continuous inverse by the Open Mapping Theorem.

Therefore, $\mathscr{A}(S)x$ is closed, and the result follows from Theorem 2.

COROLLARY 2. *The commutant of a subnormal operator S is strictly cyclic if, and only if, S is normal and has finite spectrum.*

Proof. Suppose $\{S\}'$ has a strictly cyclic vector x. Then x separates $\{S\}''$, and it follows from [2, Lemma 2.1 (i)] that $\{S\}''x$ is a closed subspace. Thus, by Corollary 1, S has finite spectrum and hence is normal.

Conversely, if $\sigma(S) = \{\lambda_1, \cdots, \lambda_n\}$, then each λ_j is an eigenvalue and $\mathscr{H}$ is the direct sum of the corresponding eigensubspaces $\mathscr{H}_j$. It follows that $\{S\}' = \mathscr{B}(\mathscr{H}_1) \oplus \cdots \oplus \mathscr{B}(\mathscr{H}_n)$. Hence any vector $x = x_1 + \cdots + x_n$ where $0 \neq x_j \in \mathscr{H}_j$, $j = 1, \cdots, n$, is strictly cyclic for $\{S\}'$.

COROLLARY 3. *Let S be a subnormal operator on a Hilbert space $\mathscr{H}$. If $\{S\}'$ is strictly cyclic and separated, then $\mathscr{H}$ is finite-dimensional.*

Proof. By Corollary 2, S is normal, its spectrum is finite, and $\{S\}' = \mathscr{B}(\mathscr{H}_1) \oplus \cdots \oplus \mathscr{B}(\mathscr{H}_n)$ with notation as in the proof of that corollary. If x is strictly cyclic for $\{S\}'$, then $x = x_1 + \cdots + x_n$ where $0 \neq x_j \in \mathscr{H}_j$, all j. If some $\mathscr{H}_j$ has dimension greater than 1, then there is a nonzero operator B_j on $\mathscr{H}_j$ which annihilates x_j, and hence there is a nonzero $B \in \{S\}'$ such that $Bx = 0$. Therefore, if $\{S\}'$ is strictly cyclic and separated, each $\mathscr{H}_j$ is one-dimensional and hence $\mathscr{H} = \mathscr{H}_1 \oplus \cdots \oplus \mathscr{H}_n$ is finite-dimensional.

COROLLARY 4. *Let S be a subnormal operator on a Hilbert space $\mathscr{H}$. If $\{S\}''$ is strictly cyclic, then $\mathscr{H}$ is finite-dimensional.*

Proof. If x is strictly cyclic for $\{S\}'' \subset \{S\}'$, then it is strictly cyclic and separating for $\{S\}'$ and the result follows from Corollary 3.

An operator A is said to be *strictly cyclic* if the weakly closed algebra generated by A and I has this property. Since this algebra is contained in the second commutant of A, it follows that the second commutant of a strictly cyclic operator is strictly cyclic. In view of Corollary 4, we have

COROLLARY 5. *There exist no strictly cyclic subnormal operators on an infinite-dimensional Hilbert space.*

REFERENCES

1. J. Bram, *Subnormal operators*, Duke Math. J., **22** (1955), 75-94.
2. Mary R. Embry, *Strictly cyclic operator algebras on a Banach space*, Pacific J. Math., **45** (1973),
3. Alan Lambert, *Strictly cyclic operator algebras*, Pacific J. Math., **39** (1971), 717-727.

Received July 7, 1972.

GEORGE MASON UNIVERSITY
AND
UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

IRREDUCIBLE SUMS OF SIMPLE MULTIVECTORS

HERBERT BUSEMANN AND D. EDWARD GLASSCO II

Denoting by $V^n(F)$ the n-dimensional vector space over the field F of characteristic 0, let $V_r^n(F)$ be the linear space of all r-vectors $\tilde{R}$ over $V^n(F)$ and $G_r^n(F)$ the Grassmann cone of the simple r-vectors R in $V_r^n(F)$. The sum $\tilde{R} = \sum_{i=1}^k R_i (R_i \in G_r^n(F))$ is irreducible if $\tilde{R}$ is not the sum of fewer than k elements of $G_r^n(F)$. (Duality reduces the interesting cases to $2 \leqq r \leqq n/2$.) Such sums are trivial only for $r = 2$, because $\bigwedge_{i=1}^k R_i \neq 0$ while always sufficient for irreducibility is then also necessary. Extension of F does not influence irreducibility if $r = 2$ but it can for $r > 2$.

The sets $W_r^n(F, k)$ of those $\tilde{R}$ in $V_r^n(F)$ which are irreducible sums of k terms behave as expected when $r = 2$, but have the most surprising properties for larger r. Although $V_3^6(F) = \bigcup_{k=1}^3 W_3^6(F, k)$ and $W_3^6(F, 3) \neq \phi$, the sets $W_3^6(\boldsymbol{R}$ or $\boldsymbol{C}, 2)$ have interior points as sets in $V_3^6(\boldsymbol{R}$ resp. $\boldsymbol{C})$ and so does $W_3^6(\boldsymbol{R}, 3)$ but $W_3^6(\boldsymbol{C}, 3)$ does not.

The paper is based on the thesis [1] with the same title by the second author.

The smallest number k for which $V_r^n(F, k) = \bigcup_{i=1}^k W_r^n(F, i)$ coincides with $V_r^n(F)$ is denoted by $N(F, n, r)$ which by duality equals $N(F, n, n - r)$. Obviously $N(F, n, r) \leqq \binom{n}{r}$. But in spite of various inequalities relating these numbers which show that $\binom{n}{r}$ is much too large, the precise value of $N(F, n, r)$ is known only in the two cases implied by the above statements: namely $N(F, n, 2) = [n/2]$ and $N(F, 6, 3) = 3$.

The values $N(\boldsymbol{C}, 7, 3) = 5$, $N(\boldsymbol{C}, 8, 3) = 7$, and $N(\boldsymbol{C}, 9, 3) = 10$ have been claimed but questioned, see Schouten [3, p. 27] and [1].

The purpose of our investigation is to elucidate why the case $r = 2$ is so much simpler than $2 < r < n - 2$. In addition to the already mentioned facts we show that $V_2^n(F, k)$ is an algebraic variety, because, if $\tilde{R}^{(i)}$ is the ith exterior power of $\tilde{R}$, then $\tilde{R}^{(k+1)} = 0$ is necessary and sufficient for $\tilde{R} \in V_r^n(F, k)$ when $r = 2$, but merely necessary when $r > 2$. This implies $\dim V_2^n(\boldsymbol{R}$ resp. $\boldsymbol{C}, k) < \dim V_2^n(\boldsymbol{R}$ resp. $\boldsymbol{C}, k + 1)$ for $1 \leqq k < [n/2]$ in contrast to the case $n = 6, r = 3$. In fact we show that $V_r^n(\boldsymbol{R}$ or $\boldsymbol{C}, k)$ is for $r > 2$, $k > 1$, and $n \geqq (k - 1)r + 3$ not even a closed set.

An irreducible representation $\tilde{R} = \sum_{i=1}^k R_i$, $k > 1$, is for $r = 2$ never unique, but for $r > 2$ it is (up to a permutation) if $\bigwedge_{i=1}^k R_i \neq 0$ and $k \leqq r$. The condition $k \leqq r$ is probably superfluous but enters —like $n \geqq (k - 1)r + 3$ (instead of $n \geqq r + 3$) above—because we use the Plücker relations for simple vectors which get out of hand for large k. A coordinate-free approach would therefore be preferable, but in many cases we were not able to devise one.

We will continue using capitals (R, S, T) with a tilde and with or without subscripts for general multivectors and omit the tilde only when the vectors are known or assumed to be simple.

2. **Results for general F, n, r, k.** The following agreement will prove convenient. $e_1, e_2, \cdots$ are used for elements of a base. If two spaces $V^m \subset V^n$ occur, then the base $e_1, \cdots, e_n$ of V^n is chosen so that $e_1, \cdots, e_m$ is a base of V^m. We begin with some simple remarks.

(2.1) *If $R \in G_r^n$ then $R = R' + S \wedge e_n$ with $R' \in G_r^{n-1}$ and $S \in G_{r-1}^{n-1}$.*

For, with suitable $v_i \in V^{n-1}$ and β_i

$$R = \bigwedge_{i=1}^{r} (v_i + \beta_i e_n)$$

$$= \bigwedge_{i=1}^{r} v_i + [\sum_{i=1}^{r} (-1)^{n-i}\beta_i v_1 \wedge \cdots \wedge v_{i-1} \wedge v_{i+1} \wedge \cdots \wedge v_r] \wedge e_n .$$

If the v_i are dependent, the bracket reduces to one term; if not, the bracket is an $(r-1)$-vector in the r-space spanned by $v_1, \cdots, v_r$ and hence is simple.

We apply (2.1) to prove

(2.2) *$\tilde{R} \in W_r^n(F, k)$ if and only if $\tilde{R} \wedge e_{n+1} \wedge \cdots \wedge e_{n+m} \in W_{r+m}^{n+m}(F, k)$.*

It suffices to prove this for $m = 1$. We show if $\tilde{R} \in W_r^n(k)$ and $\tilde{R} \wedge e_{n+1} \in W_{r+1}^{n+1}(l)$, then $l = k$. Trivially $\tilde{R} \wedge e_{n+1} \in V_{r+1}^{n+1}(k)$, whence $l \leqq k$. By (2.1) and the hypothesis $\tilde{R} \wedge e_{n+1} = \sum_{i=1}^{l} R_i = \sum_{i=1}^{l} (R_i' + S_i \wedge e_{n+1})$ with $R_i \in G_{r+1}^{n+1}$, $R_i' \in G_{r+1}^n$, and $S_i \in G_r^n$. Therefore, $\Sigma R_i' = 0$ and $\tilde{R} \wedge e_{n+1} = (\Sigma S_i) \wedge e_{n+1}$, which implies $\tilde{R} = \sum_{i=1}^{l} S_i$ and $k \leqq l$.

Corollary 2.3. *$N(F, n+1, r+1) \geqq N(F, n, r)$.*

Anticipating $N(F, n, 2) = [n/2]$ we see that both equality and inequality occur. $N(2m, 2m-2) = N(2m, 2) > N(2m-1, 2) = N(2m-1, 2m-3)$. Similarly $N(2m+1, 2m-1) = N(2m, 2m-2)$. Also $N(n, r) \geqq [(n-r+2)/2]$, but this lower bound is for $r > 2$ too small to be useful.

A consequence of (2.1) is the generalization

(2.4) *If $\tilde{R} \in V_r^n(k)$, then $\tilde{R} = \tilde{R}' + \tilde{S} \wedge e_n$ with $\tilde{R}' \in V_r^{n-1}(k)$ and*

$\tilde{S} \in V_{r-1}^{n-1}(k)$.

By hypothesis $\tilde{R} = \sum_{i=1}^{l} R_i (l \leqq k, R_i \in G_r^n)$. Applying (2.1) to each R_i yields $\tilde{R} = \sum_{i=1}^{l} (R_i' + S_i \wedge e_n) = \sum_{i=1}^{l} R_i' + (\sum_{i=1}^{l} S_i) \wedge e_n$ with $R_i' \in G_r^{n-1}$ and $S_i \in G_{r-1}^{n-1}$, which is the assertion.

With $k = N(F, n, r)$ we deduce from (2.4):

(2.5) $N(F, n, r) \leqq N(F, n-1, r) + N(F, n-1, r-1)$.

For $r = 2$ equality holds when n is even and inequality holds when n is odd.

A linear map $f\colon U^m \to V^n$ induces a homomorphism $f^*\colon U_r^m \to V_r^n$ given by $f^*(u_1 \wedge \cdots \wedge u_r) = f(u_1) \wedge \cdots \wedge f(u_r)$. The map f^* is surjective when f is. We note

(2.6) *If* $f^*(R_1) + \cdots + f^*(R_k)$ *is irreducible in* V_r^n, *then so is* $R_1 + \cdots + R_k$ *in* U_r^m.

We apply this first to the projection $f\colon V^{n+i} \to V^n$ defined by

$$f\colon \sum_{i=1}^{n+k} \alpha^i e_i \longrightarrow \sum_{i=1}^{n} \alpha^i e_i$$

and find:

(2.7) *If* $R_i \in G_r^n(F)$ *and* $\sum_{i=1}^{k} R_i$ *is irreducible in* $V_r^n(F)$, *then it is irreducible in* $V_r^{n+k}(F)$.

Hence

(2.8) $N(F, n+1, r) \geqq N(F, n, r)$.

The case $r = 2$ shows again that both inequality and equality can occur in (2.8). Next we apply (2.6) to the map $f\colon V^{n+k} \to V^{n+1}$ given by

$$f\colon \sum_{i=1}^{n+k} \alpha^i e_i \longrightarrow \sum_{i=1}^{n} \alpha^i e_i + \left(\sum_{i=n+1}^{n+k} \alpha^i\right) e_{n+1}$$

and find using (2.2):

(2.9) *If* $\sum_{i=1}^{k} R_i$ *is irreducible in* $V_r^n(F)$, *then* $\sum_{i=1}^{k} R_i \wedge e_{n+i}$ *is irreducible in* $V_{r+1}^{n+k}(F)$.

Two important facts will now be proved together:

THEOREM 2.10. *If* $\bigwedge_{i=1}^{k} R_i \neq 0$, *then* $\sum_{i=1}^{k} R_i$ *is irreducible. The* *converse holds only for* $r = 2$.

THEOREM 2.11. *If* $\tilde{R} \in V_r^n(F, k)$ *then* $\tilde{R}^{(k+1)} = 0$. *The converse holds only for* $r = 2$.

If r is odd then $\tilde{R}^{(i)} = 0$ for any $i > 1$ so that $\tilde{R}^{(k+1)} = 0$ imposes no condition. If r is even the relation $(\sum_{i=1}^k R_i)^{(k+1)} = 0$ is obvious, so that the first part of (2.11) holds. Since

(2.12) $(\sum_{i=1}^k R_i)^{(k)} = k! \bigwedge_{i=1}^k R_i$ *for even* r

it follows that $\sum_{i=1}^k R_i \in W_r^n(k)$ when $\bigwedge_{i=1}^k R_i \neq 0$. Applying (2.9) we see that this also holds for odd r.

If $\bigwedge_{i=1}^k R_i = 0$, $r = 2$, and $R_i = v_i \wedge w_i$ then one of the v_i or w_i depends on the rest, say $v_k = \sum_{i=1}^{k-1} \lambda_i v_i + \sum_{i=1}^k \mu_i w_i$ so that

$$\sum_{i=1}^{k} R_i = \sum_{i=1}^{k-1} [v_i \wedge w_i + (\lambda_i v_i + \mu_i w_i) \wedge w_k] .$$

Each bracket represents a simple vector because it is a 2-vector in the space spanned by v_i, w_i, and w_k.

That $\bigwedge_{i=1}^k R_i \neq 0$ is necessary for irreducibility only when $r = 2$ follows from (2.2). This establishes (2.10).

It remains only to prove the second part of (2.11). Let $r = 2$ and $\tilde{R}^{(k+1)} = 0$. Then $\tilde{R} \in W_r^n(k + i)$ with $i \geqq 1$ is impossible because (2.10) and (2.12) would imply $\tilde{R}^{(k+i)} \neq 0$. That $\tilde{R}^{(k+1)} = 0$ is not sufficient for $\tilde{R} \in V_r^n(F, k)$ is obvious for odd r and follows from (2.2) for even $r > 2$.

Corollaries of (2.10) resp. (2.11) are:

(2.13) $N(F, n, 2) = [n/2]$.

(2.14) *If* $\tilde{R} \in W_2^n(F, k)$ *then also* $\tilde{R} \in W_2^n(F_0, k)$ *for any extension field* F_0 *of* F. *This is not true for* $r > 2$.

The latter means that for each $n - 2 > r > 2$ there are $\tilde{R}$, $k' < k$, $F \subset F_0$ with $\tilde{R} \in W_r^n(F, k)$ and $\tilde{R} \in W_r^n(F_0, k')$, and follows from (5.9) and (2.2). Note: The first part of (2.14) does not mean, for example, that $\tilde{R} \in V_2^n(F)$, $\tilde{R} \in W_2^n(F_0, 2)$, hence $\tilde{R} = R_1 + R_2$ with $R_i \in G_2^n(F_0)$, imply $R_i \in G_2^n(F)$, but only that $R_i' \in G_2^n(F)$ with $\tilde{R} = R_1' + R_2'$ exist, compare (4.3).

Whereas in (2.2) and (2.9) the number of summands is the same in hypothesis and assertion, it is different in the next theorem which is therefore harder to prove.

THEOREM 2.15. *Let* $\tilde{R} \in W_r^n(F, k)$, $E_i = \bigwedge_{l=1}^r e_{n+(i-1)r+l} (i = 1, \cdots, j)$, *then* $\tilde{R} + \sum_{i=1}^j E_i \in W_r^{n+rj}(F, k + j)$.

Evidently it suffices to prove this for $j = 1$, or with $E = E_1$ that $\tilde{R} + E \in W_r^{n+r}(k + 1)$. Let $\tilde{R} + E = \sum_{i=1}^m S_i$, $S_i \in G_r^{n+r}$, and denote by S'_i the projection of S_i on V_r^n. Then S'_i is simple and $\tilde{R} = \sum_{i=1}^m S'_i$. Therefore, $\tilde{R} \in W_r^n(k)$ implies $m \geqq k$ and that for $m = k$ all $S'_i \neq 0$. We show that $m = k$ is impossible.

There are at least two S_i which do not lie in G_r^n. For, $S_i \in G_r^n$, if $i > 1$, would entail $S_1 = S'_1 + E$ with $S'_1 \wedge E \neq 0$, but $S'_1 + E$ is not simple by (2.10). Assume that S_1 and S_2 do not lie in G_r^n. For $w = \sum_{i=1}^{n+r} \alpha^i e_i$, put $w' = \sum_{i=1}^n \alpha^i e_i$ and $w'' = \sum_{i=n+1}^{n+r} \alpha^i e_i$.
Then

$$S_i = \bigwedge_{j=1}^r w_{ij} = \bigwedge_{j=1}^r (w'_i + w''_j)$$

and we may assume further that $w''_{11} \neq 0$ and $w''_{21} \neq 0$.

There are subscripts i, j, k, l with $i \neq k$ such that $w''_{ij} \wedge w''_{kl} \neq 0$. Otherwise $w''_{11} \wedge w''_{kl} = 0$ for $k \neq 0$ so that $w''_{kl} = \lambda_{kl} w''_{11}$ for $k \neq 1$. Similarly $w''_{kl} = \mu_{kl} w''_{21}$ for $k \neq 2$, so that $w''_{1l} = \mu_{1l} \lambda_{21} w''_{11}$.

This, with $\lambda_{11} = 1$ and $\lambda_{1l} = \mu_{1l} \lambda_{21}$, gives

$$S_i = \bigwedge_{j=1}^r (w'_j + \lambda_{ij} w''_{11}) \ .$$

But then ΣS_i cannot produce E. Thus we may assume (with a possible change of notation) that $w''_{11} \wedge w''_{21} \neq 0$. Then $e_1 \wedge \cdots \wedge e_n \wedge w_{11} \wedge w_{21} = e_1 \wedge \cdots \wedge e_n \wedge w''_{11} \wedge w''_{21} \neq 0$ and there is a base $\{e'_i\}$ of V^{n+r} with $e'_i = e_i$ for $i \leqq n$, $e'_{n+1} = w_{11}$, and $e'_{n+2} = w_{21}$. Then with the original $\tilde{R}$, $E, S_1, \cdots, S_k$,

$$(\tilde{R} + E) \wedge e'_{n+1} \wedge e'_{n+2} = (S_3 + \cdots + S_k) \wedge e'_{n+1} \wedge e'_{n+2} \ ,$$

i.e.,

$$\tilde{R} \wedge e'_{n+1} \wedge e'_{n+2} \in W_{r+2}^{n+r}(k - 1)$$

contradicting (2.2) and (2.7),

3. The sets $V_r^n(F_t, k)$. Let F_t be a topological field. Obviously $G_r^n(F_t) = V_r^n(F_t, 1) = W_r^n(F_t, 1)$ is a closed set in $V_r^n(F_t)$. It is clear that for $k < N(F_t, n, r)$ the set $V_r^n(F_t, k)$ cannot be open, but one might expect it to be closed. This is true for $r = 2$, see below, but in general not for $r > 2$. To show the latter it is not necessary to study general n and $r > 2$ because of the following:

THEOREM 3.1. *If for a topological field* F_t *the set* $V_r^n(F_t, k)$ *is* *not closed in* $V_r^n(F_t)$ *then for* $m \geqq n$, $s \geqq r$, $m - s \geqq n - r$ *and* $j \geqq 0$ *the set* $V_s^{m+js}(F_t, k + j)$ *is not closed in* $V_s^{m+js}(F_t)$.

First let $j = 0$, $m \geqq n$ and $\tilde{R} \in V_r^n(F_t, k)$. By (2.7) $\tilde{R} \in V_r^m(F_t, k)$ so that the latter is not closed. For any m we conclude from $\tilde{R} \in V_r^m(F_t, k)$ and (2.2) that

$$\tilde{R} \wedge e_{m+1} \wedge \cdots \wedge e_{m+h} \in V_{r+k}^{m+k}(F_t, k)\ .$$

Since $V_r^m(F_t, k)$ is not closed there are $\tilde{R}_\nu$ in $V_r^m(F_t, k)(\nu = 1, 2, \cdots)$ such that $\tilde{R}_\nu \rightarrow \tilde{R} \in W_r^m(F_t, k')$ with $k' > k$.
Then by (2.2)

$$\tilde{R}_\nu \wedge e_{m+1} \wedge \cdots \wedge e_{m+h} \longrightarrow \tilde{R} \wedge e_{m+1} \wedge \cdots \wedge e_{m+h} \in W_{r+h}^{m+h}(F_t, k')$$

so that $V_{r+h}^{m+h}(F_t, k)$ is not closed. This settles the case $j = 0$ or that $V_s^m(F_t, k)$ is not closed.

With the notation of (2.15) we see with the same argument

$$V_s^{m+js}(F_t, k + j) \ni \tilde{R}_\iota + \sum_{i=1}^{j} E_i \longrightarrow \tilde{R} + \sum_{i=1}^{j} E_i \in W_s^{m+js}(F_t, k' + j)$$

which proves (3.1).

In §5 it will be shown that $N(F, 6, 3) = 3$ and $V_3^6(\boldsymbol{R}$ resp. $\boldsymbol{C}, 2)$ is not closed in $V_3^6(\boldsymbol{R}$ resp. $\boldsymbol{C})$. Probably no $V_r^n(\boldsymbol{R}$ resp. $\boldsymbol{C}, k)$ with $3 \leqq r \leqq n - 3$ and $1 < k < N(\boldsymbol{R}$ resp. $\boldsymbol{C}, n, r)$ is closed, but from (3.1) we obtain (with $2 + j = k$) this best possible result only for $k = 2$.

THEOREM 3.2. *The sets* $V_r^n(\boldsymbol{R}, k)$ *and* $V_r^n(\boldsymbol{C}, k)$ *are not closed in* $V_r^n(\boldsymbol{R})$ *resp.* $V_r^n(\boldsymbol{C})$ *when* $k \geqq 2$, $r \geqq 3$, *and* $n \geqq (k - 1)r + 3$.

The mentioned best result would require a direct treatment of the case $k > 2$ instead of reduction to $k = 2$. The fact that we use Plücker relations in §5, which become very involved for large n, r, k, is responsible for our incomplete result in the case $k > 2$.

We now discuss the case $r = 2$. The by (2.11) necessary and sufficient condition $\tilde{R}^{(k+1)} = 0$ for $\tilde{R} \in V_2^n(F, k)$ amounts to polynomial conditions on the components α^{ik} of $\tilde{R} = \sum_{1 \leqq i < k \leqq n} \alpha^{ik} e_i \wedge e_k$. *The set* $V_2^n(F, k)$ *is therefore an algebraic cone in* $V_2^n(F)$ and hence closed when F carries a topology.

It is also clear that for $1 \leqq k < k' \leqq [n/2]$ the set $V_2^n(F, k)$ is a proper subset of $V_2^n(F, k')$ and plausible but, since we do not know whether $V_2^n(F, k)$ is an irreducible manifold, not *a priori* certain, that the *dimension in the sense of algebraic geometry* (denoted by *a-dim*) and consequently in the case of $\boldsymbol{R}$ resp. $\boldsymbol{C}$ also the *topological dimension* (=dim), of $V_2^n(F, k)$ is less than that of $V_2^n(F, k')$. That a proof is necessary may be seen from the case $r = 3$ (see §§5 and 6). In spite of $N(F, 6, 3) = 3$ the sets $W_3^6(\boldsymbol{R}$ resp. $\boldsymbol{C}, 2)$ and $W_3^6(\boldsymbol{R}, 3)$ have nonempty interiors in $V_3^6(\boldsymbol{R}$ resp. $\boldsymbol{C})$ so that

$$\dim V_3^6(\boldsymbol{R} \text{ resp. } \boldsymbol{C},\ 2) = \dim V_3^6(\boldsymbol{R} \text{ resp. } \boldsymbol{C})$$
$$\dim W_3^6(\boldsymbol{R}, 2) = \dim W_3^6(\boldsymbol{R}, 3) = \dim V_3^6(\boldsymbol{R}) = 20\ .$$

But $W_3^6(\boldsymbol{C}, 3)$ has no interior points and hence by a theorem in dimension theory (see [2, p. 46])

$$\dim W_3^6(\boldsymbol{C}, 3) < \dim W_3^6(\boldsymbol{C}, 2) = \dim V_3^6(\boldsymbol{C}) = 40\ .$$

Although we need only the expression for $\tilde{R}^{(k)}$ in the case $r = 2$, we give, owing to its potential usefulness, the *expression of* $\bigwedge_{i=1}^k \tilde{R}_i$ *of k different r-vectors in terms of the components of the* $\tilde{R}_i$. The rather long proof can be found in [1, p. 51].

Put $J = \{j_1, \cdots, j_r\}$ where $1 \leqq j_1 < \cdots < j_r \leqq n,\ n \geqq kr$.

Let $\alpha_i^J = \alpha_i^{j_1 \cdots j_r} (i = 1, \cdots, k)$ be indeterminates and define for a permutation π of $\{1, \cdots, r\}$

$$\alpha_i^{\pi(J)} = \alpha_i^{j_{\pi(1)} \cdots j_{\pi(r)}} = \operatorname{sgn} \pi \alpha_i^J\ .$$

If $\{H = h_1, \cdots, h_{kr}\}$ with $1 \leqq h_1 < \cdots < h_{kr} \leqq n$ and $J_1 \cup \cdots \cup J_k = H$ (disregarding order) then $J_\nu \cap J_\mu = \phi$ for $\nu \neq \mu$ and $J_1, \cdots, J_k$ in this order is a permutation of H whose sign is denoted by

$$\begin{bmatrix} J_1 \cdots J_k \\ H \end{bmatrix} .$$

We then define

$$F^H(\alpha_1, \cdots, \alpha_k) = \sum_{J_1 \cup \cdots \cup J_k = H} \begin{bmatrix} J_1 \cdots J_k \\ H \end{bmatrix} \alpha_1^{J_1} \cdots \alpha_k^{J_k}$$

where α_ν stands for $\{\alpha_\nu^J : J \subset H\}$. If π is a permutation of $\{1, \cdots, k\}$, then

$$F^H(\alpha_{\pi(1)}, \ldots, \alpha_{\pi(k)}) = (\operatorname{sgn} \pi)^r F^H(\alpha_1, \cdots, \alpha_k)\ .$$

THEOREM 3.3. *If* $\tilde{R}_i = \sum_{J \subset N} \alpha_i^J e_J$ *with* $N = \{1, \cdots, n\}$, *then* $\bigwedge_{i=1}^k \tilde{R}_i = \sum_{H \subset N} F^H(\alpha_1, \cdots, \alpha_k) e_H$.

Consequently, if $Q^H(\alpha)$ originates from $F^H(\alpha_1, \cdots, \alpha_k)$ through replacing each α_i by the same $\alpha = \{\alpha^J\}$, then we obtain

COROLLARY 3.4. *If* $\tilde{R} = \sum_{J \subset N} \alpha^J e_J$, *then* $\tilde{R}^{(k)} = k! \sum_{H \subset N} Q^H(\alpha) e_H$.

From (3.4) one deduces with the conventions $\binom{0}{2} = \binom{1}{2} = 0$,

$$a\text{-dim } V_2^n(F, k) = \binom{n}{2} - \binom{n-2k}{2}, \; 1 \leqq k \leqq \left[\frac{n}{2}\right]$$

(see [1, p. 65]). Hence

$$\dim V_2^n(\boldsymbol{R}, k) = \binom{n}{2} - \binom{n-2k}{2}, \; \dim V_2^n(\boldsymbol{C}, k) = 2\left[\binom{n}{2} - \binom{n-2k}{2}\right],$$

and so

$$\dim V_2^n(\boldsymbol{R} \text{ resp. } \boldsymbol{C}, k) < \dim V_2^n(\boldsymbol{R} \text{ resp. } \boldsymbol{C}, k+1) \text{ for } 1 \leqq k < \left[\frac{n}{2}\right]$$

in contrast to the case $r = 3$.

4. **Uniqueness.** Let $R_i \in G_r^n(F)$, $(i = 1, \cdots, k)$. The sum $\sum_{i=1}^k R_i$ is called *unique* in $V_r^n(F)$ if $S_i \in G_r^n(F)(i = 1, \cdots, k)$ and $\Sigma R_i = \Sigma S_i$ imply that $S_{\pi(i)} = R_i (i = 1, \cdots, k)$ for a suitable permutation π of $\{1, \cdots, k\}$.

Obviously:

(4.1) *If $\sum_{i=1}^k R_i$ is unique then it is irreducible.*

(4.2) *If $\sum_{i=1}^k R_i$ is irreducible resp. unique then so is $\sum_{i=1}^j R_i$ for $j < k$.*

The converse of (4.1) does not hold; in particular:

(4.3) *If $r = 2$, $k > 1$, $\bigwedge_{i=1}^k R_i \neq 0$ then $\sum_{i=1}^k R_i$ is not unique, i.e., no irreducible sum of 2-vectors is unique.*

Because of (4.2) it suffices to observe that

$$e_1 \wedge e_2 + e_3 \wedge e_4 = (e_1 + e_3) \wedge e_2 + e_3 \wedge (-e_2 + e_4) .$$

However, if $r > 2$ and $\bigwedge_{i=1}^k R_i \neq 0$, then $\sum_{i=1}^k R_i$ probably is unique. Because the Plücker relations are hard to handle for large k we were able to prove only:

THEOREM 4.4. *If $r > 2$, $k \leqq r$, and $\bigwedge_{i=1}^k R_i \neq 0$, then $\sum_{i=1}^k R_i$ is unique.*

Here both the field F and the dimension n of the space (except that $n \geqq rk$ is, of course, implied) are deliberately omitted because they are immaterial.

First we convince ourselves that n is unimportant and at the end of the proof we indicate why F is.

Let $R_i = v_{(i-1)r+1} \wedge \cdots \wedge v_{ir} (i = 1, \cdots, k)$, $\bigwedge v_i \neq 0$. It suffices to prove (for a given F) that ΣR_i is unique in the space V spanned by $v_1, \cdots, v_{kr}$. For, let also

$$\tilde{R} = \sum_{i=1}^{k} R_i = \Sigma R_i^*, \ R_i^* \in V_r^n, \ V^n \supset V .$$

Under projection of V^n on V let $R_i^* \to R_i'$. Then $\Sigma R_i^* \to \Sigma R_i'$, $\tilde{R} \to \tilde{R}$, $\bigwedge R_i^* \to \bigwedge R_i'$ so that $\tilde{R} = \Sigma R_i'$ and if (4.4) holds in V then $\{R_i'\}$ is a permutation of $\{R_i\}$. Therefore, $\bigwedge R_i' \neq 0$ and hence $\bigwedge R_i^* \neq 0$. If $R_i^* = v^*_{(i-1)r+1} \wedge \cdots \wedge v^*_{ir}$ then because (4.4) holds in the space spanned by $v_1^*, \cdots, v_{rk}^*$ we have $R^*_{\pi(i)} = R_i$ for a suitable permutation π of $\{i, \cdots, k\}$.

In the proof of (4.4) we therefore assume that $n = rk$ and $\tilde{R} = \sum_{i=1}^k R_i$ with

$$R_i = e_{(i-1)r+1} \wedge \cdots \wedge e_{ir} = e_{L(i)} ,$$

where

$$L(\nu) = \{(\nu - 1)r + 1, \cdots, \nu r\} \quad (\nu = 1, \cdots, k) .$$

We further put

$$I = \{i_1, \cdots, i_k\} \quad \text{with} \quad 1 \leqq i_1 < \cdots < i_k \leqq rk ,$$

also

$$I(\nu) = I/\{i_\nu\}, \ I(\nu, \mu) = I/\{i_\nu, i_\mu\}, \text{ etc} .$$

It will prove convenient and causes no ambiguities to use $I(\nu)$ for $\{i_1, \cdots, i_{\nu-1}, i_{\nu+1}, \cdots, i_k\}$ even if i_ν is not defined.

Ω is the set of all I with $i_\nu \in L(\nu) (\nu = 1, \cdots, k)$, and $I(\nu) \in \Omega$ means $i_\mu \in L(\mu)$ for $\mu \neq \nu$.

We also use

$$e_{L(\nu)} = \textstyle\bigwedge_{i \in L(v)} e_i, \quad e_I = \bigwedge_{i \in I} e_i, \text{ etc.}$$

The sign depends on the order but will prove irrelevant. Finally $E(\nu)$ and $F(\nu)$ are the spaces spanned by the e_i with $i \in L(\nu)$ or $i \notin L(\nu)$ respectively.

From now on we will often use the *Plücker relations* (see [3, p. 23] and [4, p. 27]) which in our type of notation may be stated as follows:

Let $P = \{p_1, \cdots, p_r\}$, $1 \leqq p_i \leqq n$, $P(i) = P/\{p_i\}$,

$$\alpha^P = \alpha^{p_1 \cdots p_r} = \operatorname{sgn} \pi \alpha^{p_{\pi(1)} \cdots p_{\pi(r)}}$$

for a permutation π of $\{1, \cdots, r\}$ and similarly for Q. The vector

$$\frac{1}{r!}\sum_{P}\alpha^{P}e_{P}=\sum_{1\leq p_1<\cdots<p_r\leq n}\alpha^{p_1\cdots p_r}e_{p_1}\wedge\cdots\wedge e_{p_r}\in V_r^n$$

is simple if and only if for any P, Q

Plücker: $\alpha^{P}\alpha^{Q}+\sum_{i=1}^{r}(-1)^{i}\alpha^{P(r)q_i}\alpha^{p_rQ(i)}=0$.

We prove several lemmas beginning with

(4.5) *Let $\tilde{T}=\sum\gamma^{I}e_{I}$ and suppose $\tilde{T}\wedge\tilde{R}=0$. If $k<r$, or $k=r$ but $\tilde{T}$ is simple, then $\gamma^{I}\neq 0$ only if $I\in\Omega$. Thus simple $\tilde{T}\neq 0$ implies $\gamma^{I}\neq 0$ for at least one $I\in\Omega$.*

If $k<r$ the assertion follows from

$$\tilde{R}\wedge\tilde{T}=\sum_{\nu=1}^{k}\left[e_{L(\nu)}\wedge\left(\sum_{I\cap L(\nu)=\phi}\gamma^{I}e_{I}\right)\right],$$

$$e_{L(\nu)}\wedge e_{I}\neq e_{L(\mu)}\wedge e_{I}\quad\text{for}\quad\nu\neq\mu\quad\text{and}\quad k<r,$$

and the observation that $I\cap L(\nu)=\phi$ for some ν is equivalent to $I\notin\Omega$.

If $k=r$ then the terms in $\tilde{R}\wedge\tilde{T}$ with $e_{L(1)}$ as a factor are

$$e_{L(1)}\wedge\left[\sum_{i_1>r}\gamma^{I}e_{I}+(-1)^{r}\gamma^{L(1)}(e_{L(2)}+\cdots+e_{L(k)})\right].$$

Therefore, $\gamma^{I}=0$ if $i_1>r$ (hence $I\notin\Omega$) and $\gamma^{L(\nu)}+(-1)^{r}\gamma^{L(1)}=0$ for $\nu>1$. Generally $\gamma^{I}=0$ if, $I\notin\Omega$ and I is no $L(\nu)$; moreover,

$$\gamma^{L(\mu)}+(-1)^{r}\gamma^{L(\nu)}=0\quad\text{if}\quad\mu\neq\nu.$$

If r is even then $\gamma^{L(\nu)}=0$ for all ν so that $\gamma^{I}\neq 0$ only for $I\in\Omega$.

If r is odd then $\gamma^{L(\nu)}=\lambda$ for all ν and thus

$$\tilde{T}=\lambda\tilde{R}+\sum_{I\in\Omega}\gamma^{I}e_{I}.$$

We show $\tilde{T}$ is simple only if $\lambda=0$ which completes the assertion.

Let $I\in\Omega$ and assume $i_1\neq r$. Then with

$$L(1,r)i_s=\{1,\cdots,r-1,i_s\}$$

one of the Plücker relations for the simplicity of $\tilde{T}$ is

$$0=\gamma^{L(1)}\gamma^{I}+\sum_{s=1}^{r}(-1)^{s}\gamma^{L(1,r)i_s}\gamma^{rI(s)}=\lambda\gamma^{I},$$

for $\gamma^{L(1,r)i_s}=0$ because $L(1,r)\in\bigcup_{\nu=1}^{r}L(\nu)\cup\Omega$ for $s>1$, and $L(1,r)i_1$ contains a repeated index (since $i_1\neq r$). If $i_1=r$ just permute $L(1)$ so that r is not the last element. Thus $\gamma^{I}=0$ for all $I\in\Omega$, or

$\lambda = 0$. Since $\tilde{T} = \lambda\tilde{R}$ would not be simple we must have $\lambda = 0$.

Let $H = \{h_1, \cdots, h_{k-1}\}$ with $1 \leqq h_1 < \cdots < h_{k-1} \leqq rk$.

(4.6) *If* $\tilde{S} = \sum \beta^H e_H$ *and* $\tilde{R} \wedge \tilde{S}$ *is simple then, for*

$$I \in \Omega,\ \beta^{I(s)}\beta^{I(t)} = 0$$

if $s \neq t$.

The terms in the expansion

$$\tilde{R} \wedge \tilde{S} = \Sigma \alpha^{i_1 \cdots i_{r+k-1}} e_{i \cdots i_{r+k-1}}$$

which contain e_I as a factor are given by

$$(4.7)\quad e_I \wedge [\pm \beta^{I(1)} e_{L(1,i_1)} \pm \cdots \pm \beta^{I(k)} e_{L(k,i_k)}]$$

where $L(\nu, i_\nu) = L(\nu)/\{i_\nu\}$. Consider the Plücker relation for $\tilde{R} \wedge \tilde{S}$ beginning with

$$\alpha^{I,L(s,i_s)}\alpha^{I,L(t,i_t)} = \beta^{I(s)}\beta^{I(t)}\ .$$

The terms not written down all vanish. The first $k-1$ that follow vanish because the first factor has a repeated superscript. From the $(k+1)$st term on, the last element of $L(i, i_s)$ is the first superscript of the second factor α which then vanishes because it does not appear in (4.7). (This requires $r \geqq 3$. The first α also vanishes and for a similar reason.)

The following is the decisive step in our long argument:

(4.8) *If both* $S = \Sigma \beta^H e_H$ *and* $\tilde{R} \wedge S$ *are simple and some* $\beta^{I(1)} \neq 0$ $(I(1) \in \Omega)$ *then* $\beta^{i_1 \nu_1 \cdots \nu_{k-2}} = 0$ *for* $i_1 \in L(1)$ *and any* $\nu_s (s = 1, \cdots, k-2)$. *Briefly* $S \in F(1)_{k-1}$.

Take any $i_1 \in L(1)$ and join it to $I(1)$. This produces an $I \in \Omega$. We prove inductively.

$$\beta^{\nu_1 \cdots \nu_\lambda I(k-\lambda, \cdots, k)} = 0 \quad \text{for all} \quad \nu_s \quad \text{and} \quad \lambda \leqq k-2\ .$$

If $k = 2$ we have $\beta^{i_1} = \beta^{I(2)} = 0$ by (4.6) and $\beta^{I(1)} \neq 0$.

If $k \geqq 3$ we make

Step 1. Consider the Plücker relation

$$\begin{aligned} 0 = \beta^{I(1)}\beta^{\nu I(k-1,k)} &- \beta^{I(1,k)\nu}\beta^{i_k I(k-1,k)} + \beta^{I(1,k)i_1}\beta^{i_k \nu I(1,k-1,k)} \\ &- \beta^{I(1,k)i_2}\beta^{i_k \nu I(2,k-1,k)} \pm \cdots \pm \beta^{I(1,k)i_{k-2}}\beta^{i_k \nu I(k-2,k-1,k)}\ . \end{aligned}$$

Except for order

$$i_k I(k-1, k) = I(k-1), \quad \text{and} \quad I(1, k)i_1 = I(k)$$

so that the second and third terms vanish by (4.6). The remaining terms vanish because the sets $I(1, k)i_2, \cdots, I(1, k)i_{k-2}$ contain repeated elements. If $k = 3$ we are finished. If $k > 3$ we make

Step 2. Take the Plücker relation

$$\begin{aligned} 0 = {} & \beta^{I(1)}\beta^{\nu\mu I(k-2,k-1,k)} - \beta^{I(1,k)\nu}\beta^{i_k\mu I(k-2,k-1,k)} \\ & + \beta^{I(1,k)\mu}\beta^{i_k\nu I(k-2,k-1,k)} - \beta^{I(1,k)i_1}\beta^{i_k\nu\mu I(1,k-2,k-1,k)} \\ & + \beta^{I(1,k)i_2}\beta^{i_k\nu\mu I(2,k-2,k-1,k)} - \cdots \pm \beta^{I(1,k)i_{k-3}}\beta^{i_k\nu\mu I(k-3,k-2,k-1,k)} . \end{aligned}$$

Except for order

$$i_k\mu I(k-2, k-1, k) = \mu I(k-2, k-1)$$

and

$$i_k\nu I(k-2, k-1, k) = \nu I(k-2, k-1)$$

so that the second and third terms vanish by Step 1 (that $k-1, k$ are replaced by $k-2, k-1$ is immaterial since the argument of Step 1 is the same for any permutation of $\{2, \cdots, k\}$). The fourth term vanishes because of (4.6) and $I(1, k)i_1 = I(k)$. In all following terms the sets of superscripts in the first factor β contain repeated elements and these terms vanish also. This completes the argument in case $k = 4$. If $k > 4$, the process clearly continues.

(4.9) *If both $S = \Sigma\beta^H e_H$ and $\tilde{R} \wedge S$ are simple and $\beta^{I(t)} \neq 0$ for some $I(t) \in \Omega$ then $S \in F(t)_{k-1}$. If $S = w_1 \wedge \cdots \wedge w_{k-1}$ then each $w_i \in F(t)$.*

The first part is a consequence of (4.8). The second statement follows from the general lemma.

(4.10) *If $m < n$, $V^m \subset V^n$, $v_j \in V^n$, $\bigwedge_{j=1}^{s} v_j \neq 0$ and $\bigwedge_{j=1}^{s} v_j \in V_s^m$ then $v_j \in V^m (j = 1, \cdots s)$.*

Setting $v_i = v_i' + \beta_i e_n$ $(i = 1, \cdots, s)$ in the proof of (2.1) yields the case $m = n - 1$ from which the general case follows.

THEOREM 4.11. *If $k \leqq r$, $S = \bigwedge_{i=1}^{k-1} w_i \neq 0$ and $\tilde{R} \wedge S$ is simple then $w_i \in F(t)$ for a suitable t.*

If $S = \Sigma\beta^H e_H$ then it suffices by (4.9) to show that $\beta^{I(t)} \neq 0$ for a suitable $I(t) \in \Omega$. Because $\tilde{R} \wedge S$ is a simple $(r + k - 1)$-vector there is a vector $v = \sum_{i=1}^{rk} \delta^i e_i$ such that $\tilde{R} \wedge S \wedge v = 0$ and $S \wedge v \neq 0$. If

$T = S \wedge v = \Sigma \gamma^J e^J$ then

$$\gamma^J = \Sigma \pm \beta^{J(t)} \delta^{j_t} .$$

By (4.5) there is at least one $I \in \Omega$ with $\gamma^I \neq 0$, hence $\beta^{I(t)} \neq 0$ for some t.

After these preparations we are ready to prove (4.4). First observe

(4.12) *If* $\tilde{R} = \sum_{\nu=1}^{k} v_{L(\nu)} (v_i \in V)$ *then* $\bigwedge_{i=1}^{rk} v_i \neq 0$.

Each $v_i \neq 0$ because $\sum_{i=1}^{k} e_{L(\nu)} = \tilde{R}$ is irreducible. Assme $\bigwedge v_i = 0$ and let $\{w_j\}$, $1 \leqq j \leqq \lambda < kr$, be a maximal set of independent v_i. Since the w_i span a proper subspace V' of V, an e_μ with $w_1 \wedge \cdots \wedge w_\lambda \wedge e_\mu \neq 0$ exists, and the w_i together with e_μ span a space V'' with $V' \subset V'' \subset V$. Now $\Sigma v_{L(\nu)}$ is irreducible in V and V', and therefore (see (2.2) and (2.7)) $\Sigma_{L(\nu)} \wedge e_\mu$ is irreducible in V'' and in V. But if $\mu \in L(s)$ then $\tilde{R} \wedge e_\mu = \sum_{\nu \neq s} e_{L(\nu)} \wedge e_\mu$.

(4.13) *If* $\tilde{R} = \sum_{\nu=1}^{k} v_{L(\nu)} (v_i \in V)$, $k \leqq r$, $I = \{i_1, \cdots, i_k\} \in \Omega$ *then* $v_{i_t} \in E(\pi(t))$ *for a suitable permutation* π *of* $\{1, \cdots, k\}$.

First $v_{I(s)} \neq 0$ by (4.12). Next

$$\tilde{R} \wedge v_{I(s)} = v_{L(s)} \wedge v_{I(s)} \neq 0$$

is simple. Therefore (4.11) yields $v_{i_\nu} \in F(\pi(s))$, $(\nu \neq s)$ for a suitable number $\pi(s)$, $(1 \leqq \pi(s) \leqq k)$. We must show that $\pi(s)$ defines a permutation of $\{1, \cdots, k\}$ or that $\pi(s) \neq \pi(t)$ for $s \neq t$. Assume $\pi(s) = \pi(t)$ for some $s \neq t$. Then $v_{i_\nu} \in F(\pi(s))$ for $\nu = 1, \cdots, k$ because $I(s) \cup I(t) = I$ whence $v_I \in F(\pi(s))_k$ and $\tilde{R} \wedge v_I = e_{L(\pi(s))} \wedge v_I \neq 0$ contradicting $\tilde{R} \wedge v_I = 0$. Thus $v_{i_t} \in \bigcap_{i_t \in I(s)} F(\pi(s)) = \bigcap_{s \neq t} F(\pi(s)) = E(\pi(t))$.

This establishes the uniqueness of $\Sigma e_{L(\nu)}$. For, consider $I \in \Omega$ and put $I' = \{j_1, i_2, \cdots, i_k\}$ with $j_1 \in L(1)$. Then $I' \in \Omega$. Since $v_{i_\nu} \in E(\pi(\nu))$ for $\nu > 1$ it follows from (4.13) that $v_{j_1} \in E(\pi(1))$. Thus $v_{j_1} \in E(\pi(1))$ for all $j_1 \in L(1)$ and $v_{L(1)} = \alpha_1 e_{L(\pi(1))}$.

Generally, $v_{L(\nu)} = \alpha_\nu e_{L(\pi(\nu))}$ whence $\alpha_\nu = 1 (\nu = 1, \cdots, k)$ and uniqueness follows.

The condition $v_i \in V$ which entered the proof of (4.12) because we applied (2.2) can now be eliminated; $\Sigma e_{L'\nu)}$ retains its form after extension of the underlying field and therefore remains unique after the extension. This justifies the formulation of (4.4) which does not mention a field.

5. The case $n = 6$, $r = 3$. The remainder of the paper deals with the case $n = 6$, $r = 3$ whose importance was noted in connection

with (3.2). We first show $N(F, 6, 3) = 3$ which may be new for $F \neq C$. Our inequalities (2.3) and (2.5) give only

$$2 = N(F, 5, 2) \leqq N(F, 6, 3) \leqq N(F, 5, 3) + N(F, 5, 2) = 4 .$$

With $e_{ijk} = e_i \wedge e_j \wedge e_k$ we prove:

(5.1) $\tilde{S} = e_{145} + e_{246} + e_{356}$ *is irreducible; whence* $N(F, 6, 3) \geqq 3$.

This proof rests on the observation:

(5.2) *If* $R = \tilde{S} \wedge \sum_{i=1}^{6} \beta^i e_i$ *is simple then* $\beta^1 = \beta^2 = \beta^3 = 0$.

(The converse is trivial but not needed.) If

$$R = \sum_{1 \geqq i<j<k<l<\geqq 6} \alpha^{ijkl} e_{ijkl}$$

then

$$\sum_{1<j<k<6} \alpha^{1jk6} e_{1jk6} = e_{16} \wedge [\beta^6 e_{45} - \beta^1(e_{24} + e_{35})] .$$

Therefore one of the Plücker relations for R is

$$0 = \alpha^{1624}\alpha^{1635} - \alpha^{1623}\alpha^{4165} + \alpha^{1625}\alpha^{4613} = (\beta^1)^2 .$$

Similarly, $\beta^2 = \beta^3 = 0$.

Assume $\tilde{S}$ were reducible, $\tilde{S} = v_{123} + v_{456}$ (where again $v_{ijk} = v_i \wedge v_j \wedge v_k$) with $v_i = \sum_{k=1}^{6} \beta_i^k e_k$. Then $\tilde{S} \wedge v_i$ is simple, so that by (5.2) $\beta_i^k = 0$ for $k \leqq 3$, whence

$$\tilde{S} = [\det(\beta_k^{3+i}) + \det(\beta_{3+k}^{3+i})] e_{456} ,$$

which is false because $\tilde{S} \wedge e_6 = e_{1456} \neq 0$.

To show $N(F, 6, 3) \leqq 3$ we need the lemma:

(5.3) *Given* $\tilde{R}_i \in V_2^4(F)(i = 1, \cdots, m)$ *there are* $\lambda_i \in F$ *and* $R_i \in G_2^4(F)(i = 0, \cdots m)$ *such that* $\tilde{R}_i = R_i + \lambda_i R_0 (i = 1, \cdots, m)$.

If $\tilde{R}_i$ is simple then $R_i = \tilde{R}_i$, $\lambda_i = 0$ will do, so we assume that no $\tilde{R}_i$ is simple. G_2^4 is a quadratic cone and a hypersurface in $V_2^4(F)$. Therefore, $R_0 \in G_2^4$ exists such that the tangent hyperplane of G_2^4 at R_0 does not contain any $\tilde{R}_i$, and no line through $\tilde{R}_i$ and R_0 intersects G_2^4 (as a locus in V_2^4 completed to a projective space) at infinity. Then the line through $\tilde{R}_i$ and R_0 intersects G_2^4 in a second point R_i' so that

$$\tilde{R}_i = (1 - \lambda_i)R_i' + \lambda_i R_0 = R_i + \lambda_i R_0 .$$

This argument does not require extending F because it amounts to solving a quadratic equation of which one root is F.

Now let $\tilde{R} = \sum_{1\leqq i<j<k\leqq 6} \alpha^{ijk} e_{ijk} \in V_3^6(F)$ be given. A simple calculation shows that either $\tilde{R} \in V_3^6(F, 2)$ or a base $\{\bar{e}_i\}$ exists in terms of which

$$\tilde{R} = \sum_{1\leqq i<j\leqq 4} \beta^{ij5}\bar{e}_{ij5} + \sum_{1\leqq i<j\leqq 4} \beta^{ij6}\bar{e}_{ij6} = \tilde{S}_1 \wedge \bar{e}_5 + \tilde{S}_2 \wedge \bar{e}_6$$

with $\tilde{S}_i \in V_2^4$. By (5.3) there are $S_i \in G_2^4$ and $\lambda_i \in F$ such that

$$\begin{aligned}\tilde{R} &= (S_1 + \lambda_1 S_0) \wedge \bar{e}_5 + (S_2 + \lambda_2 S_0) \wedge \bar{e}_6 \\ &= S_1 \wedge \bar{e}_5 + S_2 \wedge \bar{e}_6 + S_0 \wedge (\lambda_1 \bar{e}_5 + \lambda_2 \bar{e}_6) \in V_3^6(F, 3) .\end{aligned}$$

Thus:

$$N(F, 6, 3) = 3 . \tag{5.4}$$

By a similar argument we prove

$$3 \leqq N(F, 7, 3) \leqq 5 . \tag{5.5}$$

The left inequality follows from $3 = N(6, 3) \leqq N(7, 3)$, see (2.8). For the right inequality one shows (see [1, p. 90]) that either $\tilde{R} \in V_3^7(2)$ or with a suitable base $\{\bar{e}_i\}$

$$\tilde{R} = \sum_{1\leqq i<j\leqq 4} \beta^{ij5}\bar{e}_{ij5} + \sum_{1\leqq i<j\leqq 4} \beta^{ij6}\bar{e}_{ij6} + \sum_{1\leqq i<j\leqq 4} \beta^{ij7}\bar{e}_{ij7} + \sum_{i=1}^{5} \beta^{i67}\bar{e}_{i67} .$$

The last sum is simple and applying (5.3) to the first three terms on the right yields $N(F, 7, 3) \leqq 5$. This method does not extend to $N(n, 3)$ with $n > 7$.

We now study a *special type* of $\tilde{R} \in V_3^6(C)$ which will confirm some of the important assertions made previously.

Let Y be the set of triples

$$Y = \{123, 126, 135, 156, 234, 246, 345, 456\}$$

and suppose that the α^I, $I \in Y$, satisfy the inequalities

$$\begin{aligned} &\alpha^{123}\alpha^{156} + \alpha^{126}\alpha^{135} \neq 0 , \qquad \alpha^{123}\alpha^{246} + \alpha^{126}\alpha^{234} \neq 0 , \\ &\alpha^{123}\alpha^{345} + \alpha^{135}\alpha^{234} \neq 0 , \qquad \alpha^{234}\alpha^{456} + \alpha^{246}\alpha^{345} \neq 0 , \\ &\alpha^{135}\alpha^{456} + \alpha^{156}\alpha^{345} \neq 0 , \qquad \alpha^{126}\alpha^{456} + \alpha^{156}\alpha^{246} \neq 0 , \end{aligned} \tag{5.6}$$

and that the roots λ, μ of

$$(\alpha^{123} x - \alpha^{234})(\alpha^{156} x - \alpha^{456}) + (\alpha^{126} x + \alpha^{246})(\alpha^{135} x + \alpha^{345}) = 0 \tag{5.7}$$

are distinct. They are different from zero.

THEOREM 5.8. *If $\tilde{R} = \sum_{I \in Y} \alpha^I e_I \in V_3^6(C)$ and the α^I satisfy (5.6), then $\tilde{R} = R_\beta + R_\gamma$, $R_\beta = \sum_{I\in Y} \beta^I e_I$, $R_\gamma = \sum_{I \in Y} \gamma^I e_I$, where R_β and R_γ are simple with $R_\beta \wedge R_\gamma \neq 0$. Hence the representation $R_\beta + R_\gamma$ is* *unique* (*by* (4.4)).

If λ, μ *are the solutions of* (5.7) *then the* β^I *and* $\gamma^I (I \in Y)$ *are given by*

$$\beta^{1ij} = \frac{\mu\alpha^{1ij} - \alpha^{4ij}}{\mu - \lambda}, \quad \gamma^{1ij} = \frac{\alpha^{4ij} - \lambda\alpha^{1ij}}{\mu - \lambda},$$
$$\beta^{4ij} = \lambda\beta^{1ij}, \qquad \gamma^{4ij} = \mu\gamma^{1ij}.$$

No β^I *or* $\gamma^I (I \in Y)$ *vanishes.*

This representation was found by using Plücker relations (see [1, pp. 98–106]), but after it is explicitly given one readily verifies that R_β and R_γ are simple and that $R_\beta \wedge R_\gamma \neq 0$. In fact, it is easy to factor R_β and R_γ, see [4, p. 21]: Since $\beta^I \neq 0$ if $I \in Y$, letting $\nu = (\beta^{123})^{-2/3}$ we find

$$R_\beta = u \wedge v \wedge w = \Sigma u^i e_i \wedge \Sigma v^i e_i \wedge \Sigma w^i e_i$$

with

$$u^i = \nu\beta^{23i},\ v^i = -\nu\beta^{13i},\ w^i = \nu\beta^{12i}$$

(see also [1, p. 102]), and similarly for R_γ.

First we confirm the statement in the introduction that *irreducibility may depend on the field.*

(5.9) *If the* α^I *in* (5.8) *are real and* λ, μ *are not, then* $\tilde{R} \in W_3^6(\boldsymbol{C}, 2)$ *but* $\tilde{R} \in W_3^6(\boldsymbol{R}, 3)$.

For because $R_\beta + R_\gamma$ is unique, $\tilde{R} \in V_3^6(\boldsymbol{R}, 2)$ is impossible, and this with $N(\boldsymbol{R}, 6, 3) = 3$ entails the assertion.

Next we observe that the vector

$$\tilde{R}(\eta) = e_1 \wedge (e_2 + e_5) \wedge e_6 + e_1 \wedge e_3 \wedge e_5 + \eta e_2 \wedge e_4 \wedge e_6$$
$$+ (e_2 + e_5) \wedge e_3 \wedge e_4 \qquad (\eta \neq 0)$$

is a special case of (5.8) and that λ, μ are real when $\eta < 0$. Letting $\eta \to 0^-$ we find

$$\tilde{R}(0^-) = e_1 \wedge (e_2 + e_5) \wedge e_6 + e_1 \wedge e_3 \wedge e_5 + (e_2 + e_5) \wedge e_3 \wedge e_4$$

which by (5.1) lies in $W_3^6(\boldsymbol{R} \text{ or } \boldsymbol{C}, 3)$. Therefore:

(5.10) *The sets* $V_3^6(\boldsymbol{R}, 2)$ *resp.* $V_3^6(\boldsymbol{C}, 2)$ *are not closed in* $V_3^6(\boldsymbol{R})$ *resp.* $V_3^6(\boldsymbol{C})$.

Theorem (3.2) whose proof used (5.10) is therefore completely established.

We now prove a surprising fact for $\boldsymbol{C}$ which has no analogue for $\boldsymbol{R}$ (see (6.3)):

(5.11) *The interior of $W_3^6(\boldsymbol{C}, 3)$ as a set in $V_3^6(\boldsymbol{C})$ is empty.*

We show that if $\tilde{R} = R_1 + R_2 + R_3$, $R_i \in G_3^6(C)$, is irreducible then it is the limit of elements in $V_3^6(\boldsymbol{C}, 2)$.

R_i and $R_j (i \neq j)$ have no nonvanishing 2-vector as a common factor, because $R_i + R_j$ would then be simple. Thus two cases are to be considered:

(1) $R_i \wedge R_j \neq 0$ for some i, j, say $R_1 \wedge R_2 \neq 0$,

(2) R_i and R_j have for $i \neq j$ a vector $v_k \neq 0$ (but no 2-vector $\neq$ 0) as a common factor where (i, j, k) is a permutation of $(1, 2, 3)$.

In the latter case the v_i are either parallel or no two v_i are parallel. If they were parallel we could choose e_6 parallel to the v_i so that $\tilde{R} = \tilde{S} \wedge e_6$ with $\tilde{S} \in V_2^5(C)$, and ΣR_i would be reducible since $N(F, 5, 2) = 2$. If no two v_i are parallel then with suitable u_i

$$R_1 = u_1 \wedge v_2 \wedge v_3,\ R_2 = u_2 \wedge v_1 \wedge v_3,\ R_3 = u_3 \wedge v_1 \wedge v_2\ .$$

The vectors u_i, v_j are independent, for otherwise ΣR_i would be a 3-vector in a space of dimension less than 6 and by $N(F, 5, 3) = 2$ reducible. The proof of (5.10) shows $\tilde{R}$ can be approximated by elements of $V_3^6(\boldsymbol{C}, 2)$.

In case (1) there are vectors $w_1, \cdots, w_6, v_1, v_2, v_3$ such that w_1, w_2, w_3 are parallel to R_1, w_4, w_5, w_6 are parallel to R_2, $R_3 = v_1 \wedge v_2 \wedge v_3$ and

$$v_1 = a_1 w_1 + a_4 w_4,\ v_2 = a_2 w_2 + a_5 w_5,\ v_3 = a_3 w_3 + a_6 w_6\ .$$

If $\bigwedge_{i=1}^6 w_i \neq 0$ then $\tilde{R} = \sum_{I \in Y} \alpha^I w_I$ and $\prod_{i=1}^6 a_i \neq 0$ is equivalent to (5.6), so we have a special case of (5.8) (see [1, p. 83]) and hence $\tilde{R} \in V_3^6(\boldsymbol{C}, 2)$ contrary to the hypothesis. If $\bigwedge_{i=1}^6 w_i = 0$ and/or $\prod_{i=1}^6 a_i = 0$ we can choose w_i' and a_i' arbitrarily close to w_i resp. a_i such that $\bigwedge_{i=1}^6 w_i' \neq 0$ $\prod_{i=1}^6 a_i' \neq 0$ and $\lambda \neq \mu$, so that $\tilde{R}$ is the limit of elements in $V_3^6(\boldsymbol{C}, 2)$.

For $kr \leqq n$ let $Z_r^n(F, k)$ *be the set of* $\tilde{R} = \sum_{i=1}^k R_i$ *with* $\bigwedge_{i=1}^k R_i \neq 0$. Then $Z_r^n(F, k) \subset W_r^n(F, k)$ by (2.10).

(5.12) *$Z_r^n(\boldsymbol{R}$ resp. $\boldsymbol{C}, k)$ is dense in $V_r^n(\boldsymbol{R}$ resp. $\boldsymbol{C}, k)$.*

This is nearly obvious: If $\tilde{R} = \sum_{i=1}^j R_i \in Z_r^n(j)$, $j < k$, then $R_{j+1}, \cdots, R_k$ exist with $\bigwedge_{i=1}^k R_i \neq 0$ and

$$\tilde{R} = \lim (\tilde{R} + \delta \sum_{i=j+1}^k R_i) \text{ as } \delta \to 0\ .$$

If $\tilde{R} = \sum_{i=1}^j R_i \in W_r^n(j)$, $j \leqq k$, $\bigwedge_{i=1}^j R_i = 0$ and $R_i = \bigwedge_{h=1}^r v_{(i-1)r+h}$ then

$w_i \to v_i$ with $\bigwedge_{i=1}^{jr} w_i \neq 0$ exist and $\sum_{i=1}^{j} \bigwedge_{h=1}^{r} w_{(i-1)r+h} \to \tilde{R}$.

Because

$$V_3^6(C) = Z_3^6(C, 2) \cup V_3^6(C, 2)/Z_3^6(C, 2) \cup W_3^6(C, 3)\,,$$

(5.11, 12) show that $Z_3^6(C, 2)$ is dense in $V_3^6(C)$, so that $V_3^6(C)/Z_3^6(C, 2)$ has no interior points, and hence has dimension less than 40 ($=$ dim $V_3^6(C)$), see [2, p. 46]. In the next section we will see that $Z_3^6(C, 2)$ is open. Thus

(5.13) *The set* $Z_3^6(C, 2)$ *is open and dense in* $V_3^6(C)$, *hence*

$$V_3^6(C)/Z_3^6(C, 2)$$

is closed and $\dim Z_3^6(C, 2) = 40$, $\dim V_3^6(C)/Z_3^6(C, 2) < 40$.

Note that $W_3^6(C, 3) \subset V_3^6(C)/Z_3^6(C, 2)$ and that therefore *the closure of* $W_3^6(C, 3)$ *has dimension less than* 40.

6. The sets $Z_3^6(\boldsymbol{R}$ resp. $\boldsymbol{C}, 2)$ and $W_3^6(\boldsymbol{R}, 3)$. We now prove that $Z_3^6(\boldsymbol{R}$ resp. $\boldsymbol{C}, 2)$ is open. Actually our next theorem provides much more information which will allow us to show that $W_3^6(\boldsymbol{R}, 3)$ has a nonempty interior.

THEOREM 6.1. *Let* $F = \boldsymbol{R}$ or $\boldsymbol{C}$. *If* $R_1, R_2 \in G_3^6(F)$ *and* $R_1 \wedge R_2 \neq 0$ *then there is a neighborhood* $\tilde{U}(\tilde{R}_0)$ *of* $\tilde{R}_0 = R_1 + R_2$ *in* $V_3^6(F)$ *such that for* $\tilde{R} \in \tilde{U}(\tilde{R}_0)$ *there are simple* R_1', R_2' *with* $\tilde{R} = R_1' + R_2'$. *Furthermore, given neighborhoods* $U_i(R_i)$ *of* R_i *in* $G_3^6(F)$ *there is a neighborhood* $\tilde{U}'(\tilde{R}_0) \subset \tilde{U}(\tilde{R}_0)$ *such that* $\tilde{R} \in \tilde{U}'(\tilde{R}_0)$ *implies* $R_i' \in U_i(R_i)$ *and* $R_1' \wedge R_2' \neq 0$. *Consequently* $\tilde{R} \in Z_3^6(F, 2)$ *and by* (4.4) $R_1' + R_2'$ *is unique.*

Necessary for (6.1) to work is that $G_3^6 \times G_3^6$ and V_3^6 have the same dimension, which is the case because a-dim $G_r^n = r(n-r) + 1$, a-dim $V_r^n = \binom{n}{r}$ and $2[3(6-3)+1] = 20 = \binom{6}{3}$. But this argument is far from sufficient as similar situations for other dimensions show; the structure of G_3^6 enters.

Since $R_1 \wedge R_2 \neq 0$ we can choose a base so that $R_1 = e_{123}$, $R_2 = e_{456}$. A neighborhood of R_1 on G_3^6 consists of the simple $R_1' = \Sigma \beta^I e_I = \sum_{1 \leq i<j<k \leq 6} \beta^{ijk} e_{ijk}$ with β^{123} close to 1 and the remaining β^I close to 0, so $\beta^{123} \neq 0$ may be assumed. Similarly for $R_2' = \Sigma \gamma^I e_I$ and $\gamma^{456} \neq 0$. The components of $\tilde{R}_0$ are $1, 0, \cdots, 0, 1$.

The special properties of G_3^6 arise from the Plücker relations which (see [1, p. 69]) with $\lambda = (\beta^{123})^{-1}$ are equivalent to

$$\beta^{ijk} = \pm \lambda \begin{vmatrix} \beta^{i\rho j} & \beta^{i\rho k} \\ \beta^{i\sigma j} & \beta^{i\sigma k} \end{vmatrix} 1 \leqq i \leqq 3,\ 4 \leqq j < k \leqq 6,\ (i, \rho, \sigma) = \pi(1, 2, 3)\ ,$$

$$\beta^{456} = \lambda^2 \begin{vmatrix} \beta^{124} & \beta^{125} & \beta^{126} \\ \beta^{134} & \beta^{135} & \beta^{136} \\ \beta^{234} & \beta^{235} & \beta^{236} \end{vmatrix} = B\ .$$

Similarly with $\mu = (\gamma^{456})^{-1}$

$$\gamma^{ijk} = \pm \mu \begin{vmatrix} \gamma^{i\rho k} & \gamma^{j\rho k} \\ \gamma^{i\sigma k} & \gamma^{j\sigma k} \end{vmatrix} 1 \leqq i < j \leqq 3,\ 4 \leqq k \leqq 6,\ (\rho, \sigma, k) = \pi(4, 5, 6)\ ,$$

$$\gamma^{123} = \mu^2 \begin{vmatrix} \gamma^{145} & \gamma^{245} & \gamma^{345} \\ \gamma^{146} & \gamma^{246} & \gamma^{346} \\ \gamma^{156} & \gamma^{256} & \gamma^{356} \end{vmatrix} = C\ .$$

If now $\Sigma\alpha^I e_I = \tilde{R} = R_1' + R_2' = \Sigma\beta^I e_I + \Sigma\gamma^I e_I$ then $\alpha^I = \beta^I + \gamma^I$, and substitution gives

$$\alpha^{123} = \beta^{123} + C\ ,$$

$$\alpha^{ijk} = \beta^{ijk} \pm \mu \begin{vmatrix} \gamma^{i\rho k} & \gamma^{j\rho k} \\ \gamma^{i\sigma k} & \gamma^{j\sigma k} \end{vmatrix} 1 \leqq i < j \leqq 3,\ 4 \leqq k \leqq 6,\ (\rho, \sigma, k) = \pi(4, 5, 6)\ ,$$

$$\alpha^{ijk} = \pm \lambda \begin{vmatrix} \beta^{i\rho j} & \beta^{j\rho k} \\ \beta^{i\sigma j} & \beta^{i\sigma k} \end{vmatrix} + \gamma^{ijk},\ 1 \leqq i \leqq 3,\ 4 \leqq j < k \leqq 6,\ (i, \rho, \sigma) = \pi(1, 2, 3),$$

$$\alpha^{456} = B + \gamma^{456}\ .$$

Thus the 20 components of $\tilde{R}$ are expressed in terms of β^{123}, the nine β^I with $1 \leqq i \leqq 3, 4 \leqq j < k \leqq 6$, the nine γ^I with $1 \leqq i < j \leqq 3, 4 \leqq k \leqq 6$, and γ^{456}. Evaluation of the functional determinant at $(1, 0, \cdots, 0, 1)$ gives the value 1. Therefore, the implicit function theorem is applicable and yields the assertion. The details of the calculation may be found in the thesis, [1, pp. 93–97].

As a corollary we have

(6.2) *The set* $Z_3^6(\boldsymbol{R}$ *resp.* $\boldsymbol{C}, 2)$ *is open in* $V_3^6(\boldsymbol{R}$ *resp.* $\boldsymbol{C})$.

But in contrast to (5.11):

(6.3) *The interior of* $W_3^6(\boldsymbol{R}, 3)$ *is not empty.*

For take any $\tilde{R}_0 = \sum_{I \in Y} \alpha^I e_I$ of (5.8) for which the α^I are real but λ, μ are not. Then

$$\tilde{R}_0 = R_\beta + R_\gamma \quad \text{with} \quad R_\beta \wedge R_\gamma \neq 0\ .$$

By (6.1) for $\tilde{R} \in \tilde{U}'(\tilde{R}_0) \subset V_3^6(\boldsymbol{C})$

$\tilde{R} = R_1' + R_2'$ with $R_1' \wedge R_2' \neq 0$, R_1' close to R_β and R_2' close to R_γ so that R_1' and R_2' cannot be real either. Also $R_1' + R_2'$ is unique by (4.4) and this implies as in the proof of (5.9) that

$$\tilde{R} \in W_3^6(\boldsymbol{R}, 3) \quad \text{for} \quad \tilde{R} \in \tilde{U}'(\tilde{R}_0) \cap V_3^6(\boldsymbol{R}) ,$$

where we consider $V_3^6(\boldsymbol{R})$ as a subset of $V_6^3(\boldsymbol{C})$.

REFERENCES

1. D. E. Glassco, *Irreducible Sums of Simple Multivectors*, Dissertation, University of Southern California, 1971.

2. W. Hurewicz and H. Wallmann, *Dimension Theory*, 1st edition, Princeton, 1941.

3. J. Schouten, *Tensor Analysis for Physicists*, 2nd edition, Oxford, 1954.

4. B. L. Van der Waerden, *Einführung in die Algebraische Geometrie*, Berlin, 1938, New York, 1945.

Received July 17, 1972.

UNIVERSITY OF SOUTHERN CALIFORNIA
AND
CALIFORNIA POLYTECHNIC STATE UNIVERSITY AT SAN LUIS OBISPO

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

COUNTABLY COMPACT GROUPS AND FINEST TOTALLY BOUNDED TOPOLOGIES

W. W. COMFORT AND VICTOR SAKS

The first main result formalizes the general principle that each totally bounded group G is dense in some group H, not much larger than G, in which every subset of small cardinality has a complete accumulation point. For example: If G is totally bounded and $|G| = \mathfrak{n} \geqq \aleph_0$, then G is dense in a countably compact group H such that $|H| \leqq \mathfrak{n}^{\aleph_0}$. A corollary: If K is an infinite compact group with weight not exceeding $2^{\mathfrak{n}}$, then K contains a dense, countably compact subgroup H with $|H| \leqq \mathfrak{n}^{\aleph_0}$.

The following results are given in § 2: If t is the finest totally bounded topological group topology on an infinite Abelian group G, then every subgroup of G is t-closed and (G, t) is not pseudocompact (both conclusions can fail for G non-Abelian); a closed subgroup of a pseudocompact group need not be pseudocompact; if $\{(G_i, t_i): i \in I\}$ are nontrivial Abelian groups with their finest totally bounded topologies and $(G, \mathscr{T})$ is their product, then $\mathscr{T} = t$ if and only if $|I| < \aleph_0$.

1. Countably compact groups. Throughout this section the word group refers to a topological group which satisfies the Hausdorff separation axiom. Such spaces are known to be completely regular topological spaces. A group is said to be totally bounded if for each non-empty open subset U of G there is a finite subset $\{x_k: k < n\}$ of G for which $G = \bigcup_{k<n} x_k \cdot U$. Each subgroup of a compact group is totally bounded, and Weil [26] has shown the converse: Each totally bounded group G is (homeomorphic with) a dense subgroup of a compact group and this compactification is unique to within a topological isomorphism leaving G fixed pointwise. We refer to this compactification of G as the Weil completion of G, and we denote it by the symbol $\bar{G}$.

A completely regular Hausdorff space X is countably compact if each of its infinite subsets has an accumulation point, and pseudocompact if each continuous, real-valued function on X is bounded (equivalently: each locally finite family of open subsets of X is finite). It is easy to see that each countably compact space is pseudocompact, and (as in [6], for example) that each pseudocompact group is totally bounded. Examples abound of pseudocompact groups which are not countably compact; see for example Kister [18] or H. Wilcox [28].

A number of theorems in the works of Itzkowitz [16], [17], and H. Wilcox [28] are devoted to showing that (in various settings and under various hypotheses) between a totally bounded group G and its Weil completion $\bar{G}$ there is a pseudocompact group which is in a certain precise sense not much larger than G. The principal result of this section is a theorem of compactness type which has a number of corollaries improving these results. We show, specifically, that the groups of Itzkowitz and H. Wilcox may be chosen countably compact.

Notation. If $\mathfrak{m}$ is a cardinal and S is a set, then

$$\mathscr{P}_{\mathfrak{m}}(S) = \{A \subset S : |A| < \mathfrak{m}\} .$$

If $\mathfrak{m}$ and $\mathfrak{n}$ are cardinals then

$$\mathfrak{n}^{\underline{\mathfrak{m}}} = \sum \{\mathfrak{n}^{\mathfrak{k}} : \mathfrak{k} \text{ is a cardinal and } \mathfrak{k} < \mathfrak{m}\} .$$

It is well-known and easy to prove that if $\mathfrak{m}$ and $\mathfrak{n}$ are infinite cardinals and $\mathfrak{m} \leqq \mathfrak{n}^+$, then

$$|\mathscr{P}_{\mathfrak{m}}(\mathfrak{n})| = \mathfrak{n}^{\underline{\mathfrak{m}}} \geqq \mathfrak{m} .$$

(Here as usual, the symbol $\mathfrak{n}^+$ denotes the smallest cardinal greater than $\mathfrak{n}$.)

Recall that if X is a space and $Y \subset X$, then a point p of X is a complete accumulation point of Y provided that

$$|U \cap Y| = |Y|$$

for each neighborhood U of p in X. It is obvious that if X is compact, then each infinite subset of X has a complete accumulation point in X.

The weight and density character of a space X are denoted wX and dX, respectively.

DEFINITION. Let $\mathfrak{m}$ and $\mathfrak{n}$ be cardinal numbers with $\aleph_0 \leqq \mathfrak{m} < \mathfrak{n}$. The space X is $[\mathfrak{m}, \mathfrak{n}]$-compact in the sense of complete accumulation points provided: If $Y \subset X$ and $\mathfrak{m} \leqq |Y| < \mathfrak{n}$, then Y has a complete accumulation point in X.

The term we have just used is often defined as above except that it is required that $|Y|$ be a regular cardinal. Even this weaker property is strong enough to yield a compactness condition of covering type; in the interest of completeness we give a proof below. For positive results in the converse direction, see Alexandroff and Urysohn [2] and Aleksandrov [1], and for negative results in the converse direction see Mishchenko [21].

We note that in our terminology the spaces which are $[\aleph_0, \aleph_1]$-compact in the sense of complete accumulation points are the countably compact spaces: each countably infinite subset has a (complete) accumulation point.

PROPOSITION. *Let $\mathfrak{m}$ and $\mathfrak{n}$ be cardinal numbers with $\aleph_0 \leqq \mathfrak{m} < \mathfrak{n}$ and let X have the property that if $Y \subset X$ and $\mathfrak{m} \leqq |Y| < \mathfrak{n}$ and $|Y|$ is regular then Y has a complete accumulation point in X. Then for each open cover $\mathscr{U}$ of X such that $\mathfrak{m} \leqq |\mathscr{U}| < \mathfrak{n}$ and $|\mathscr{U}|$ is regular, there is a cover $\mathscr{V} \subset \mathscr{U}$ and $|\mathscr{V}| < |\mathscr{U}|$.*

Proof. Suppose that $\mathscr{U} = \{U_\xi: \xi < |\mathscr{U}|\}$ is a counterexample, and for $\xi < |\mathscr{U}|$ choose recursively $p_\xi \in X$ and $f(\xi) < |\mathscr{U}|$ as follows:

(i) $p_0 \in X \backslash U_0$ and $p_0 \in U_{f(0)}$;

(ii) if p_η and $f(\eta)$ have been defined for all $\eta < \xi$, then

$$p_\xi \in X \backslash \bigcup_{\eta<\xi} (U_\eta \cup U_{f(\eta)}) \quad \text{and} \quad p_\xi \in U_{f(\xi)} .$$

Then with $Y = \{p_\xi: \xi < |\mathscr{U}|\}$ we have $|Y| = |\mathscr{U}|$ so there is a complete accumulation point p of Y in X. If η is chosen so that $p \in U_\eta$, then there exists $\xi > \eta$ such that $p_\xi \in U_\eta$. This contradiction completes the proof.

THEOREM 1.1. *Let G be a totally bounded group such that $|G| = \mathfrak{n} \geqq \aleph_0$ and let $\mathfrak{m}$ be a regular cardinal for which $\mathfrak{m} \leqq \mathfrak{n}^+$. Then there is a group H, with $G \subset H \subset \bar{G}$, such that H is $[\aleph_0, \mathfrak{m}]$-compact in the sense of complete accumulation points and $|H| \leqq \mathfrak{n}^{\underline{\mathfrak{m}}}$.*

Proof. For $A \in \mathscr{P}_{\mathfrak{m}}(\bar{G})$ with $|A| \geqq \aleph_0$ let p_A be a complete accumulation of A in $\bar{G}$, and for $S \subset \bar{G}$ let

$$F(S) = S \cup \{p_A: A \in \mathscr{P}_{\mathfrak{m}}(S), |A| \geqq \aleph_0\} ;$$

and for $S \subset \bar{G}$ let $\langle S \rangle$ denote the subgroup of $\bar{G}$ generated by S.

Now let $H_0 = G$ and $H_1 = \langle F(H_0) \rangle$ and recursively, if $\xi < \mathfrak{m}$ and H_η has been defined for $\eta < \xi$, let

$$H_\xi = \langle F(\bigcup_{\eta<\xi} H_\eta) \rangle .$$

We show by induction $|H_\xi| \leqq \mathfrak{n}^{\underline{\mathfrak{m}}}$ for $\xi < \mathfrak{m}$. This is true for $\xi = 0$ because

$$|H_0| = \mathfrak{n} \leqq \mathfrak{n}^{\underline{\mathfrak{m}}}$$

and for $\xi = 1$ because

$$|F(H_0)| \leqq |\mathscr{P}_{\mathfrak{m}}(H_0)| = \mathfrak{n}^{\underline{\mathfrak{m}}} .$$

If $|H_\eta| \leqq \mathfrak{n}^{\underline{\mathfrak{m}}}$ for $\eta < \xi$ then $|\bigcup_{\eta<\xi} H_\eta| \leqq |\xi| \cdot \mathfrak{n}^{\underline{\mathfrak{m}}} = \mathfrak{n}^{\underline{\mathfrak{m}}}$, so

$$|F(\bigcup_{\eta<\xi} H_\eta)| \leqq (\mathfrak{n}^{\underline{\mathfrak{m}}})^{\underline{\mathfrak{m}}} = \mathfrak{n}^{\underline{\mathfrak{m}}} ;$$

this last equality holds because $\mathfrak{m}$ is regular (see Bachmann [3], pp.

152–153). Thus $|H_\xi| \leqq \mathfrak{n}^{\underline{\mathfrak{m}}}$.

Now let

$$H = \bigcup_{\xi < \mathfrak{m}} H_\xi .$$

Then H is a group and $G \subset H \subset \bar{G}$ and

$$|H| \leqq \sum_{\xi < \mathfrak{m}} |H_\xi| \leqq \mathfrak{m} \cdot \mathfrak{n}^{\underline{\mathfrak{m}}} = \mathfrak{n}^{\underline{\mathfrak{m}}} .$$

And if $A \in \mathscr{P}_{\mathfrak{m}}(H)$ with $|A| \geqq \aleph_0$ then because $\mathfrak{m}$ is regular there is $\xi < \mathfrak{m}$ such that $A \subset H_\xi$, and we have

$$p_A \in F(H_\xi) \subset H_{\xi+1} \subset H ;$$

thus H is $[\aleph_0, \mathfrak{m}]$-compact in the sense of complete accumulation points.

Our first corollary, but with "pseudocompact" in place of "countably compact", was given in the general case by H. Wilcox [28] and earlier, for Abelian groups G, by Itzkowitz [16], [17].

COROLLARY 1.2. *Let G be a totally bounded group such that $|G| = \mathfrak{n} \geqq \aleph_0$. Then there is a countably compact group H such that $G \subset H \subset \bar{G}$ and $|H| \leqq \mathfrak{n}^{\aleph_0}$.*

Proof. This follows from Theorem 1.1, upon taking $\mathfrak{m} = \aleph_1$.

COROLLARY 1.3. *For each infinite cardinal $\mathfrak{n}$ there is a totally bounded group H which is $[\aleph_0, \mathfrak{n}^+]$-compact in the sense of complete accumulation points but not compact, and for which $dH \leqq \mathfrak{n}$ and $|H| \leqq 2^{\mathfrak{n}}$.*

Proof. Let K be the compact group $2^{2^{\mathfrak{n}}}$. According to a well-known result of Hewitt [11] and Pondiczery [22] there is a dense subset S of K with $|S| = \mathfrak{n}$. Let G be the subgroup of K generated by S, so that $|G| = \mathfrak{n}$ and $K = \bar{G}$ by Weil's theorem. The result now follows from Theorem 1.1, upon taking $\mathfrak{m} = \mathfrak{n}^+$ and noting that

$$\mathfrak{n}^{\underline{\mathfrak{m}}} = \mathfrak{n}^{\mathfrak{n}} = 2^{\mathfrak{n}} .$$

COROLLARY 1.4. *There is a separable, countably compact group which is not compact.*

Proof. This is the case $\mathfrak{n} = \aleph_0$ of Corollary 1.3.

COROLLARY 1.5. *Let $\mathfrak{n}$ be a cardinal and let K be a compact group such that $\aleph_0 \leqq wK \leqq 2^{\mathfrak{n}}$. If $\mathfrak{m}$ is a regular cardinal for which $\mathfrak{m} \leqq \mathfrak{n}^+$, then K contains a dense subgroup H which is $[\aleph_0, \mathfrak{m}]$-compact* *in the sense of complete accumulation points such that* $|H| \leqq \mathfrak{n}^{\mathfrak{m}}$.

Proof. According to Kuzminov [20] there is a continuous function f from the compact group 2^{wK} onto K. (For an English-language proof that K is the continuous image of $2^{\mathfrak{n}}$ for some cardinal $\mathfrak{n}$ when K is compact and Abelian, see Hewitt and Ross [12], pp. 423–424. That $\mathfrak{n}$ may be chosen to be wK follows from general topological considerations as in Engelking [9] p. 162.) Again by the theorem of Hewitt [11] and Pondiczery [22] there is a dense subset D of 2^{wK} with $|D| \leqq \mathfrak{n}$. Then $f[D]$ is dense in K, so there is a dense subgroup G of K such that $|G| = \mathfrak{n}$. Then $K = \bar{G}$ by Weil's theorem, so the result follows from Theorem 1.1.

The following two corollaries, with "countably compact" in place of "pseudocompact", are given by H. Wilcox [28]. The first of these is given by Itzkowitz [16], [17] for the case in which K is Abelian.

COROLLARY 1.6. *Let* $\mathfrak{n}$ *be a cardinal and let* K *be a compact group such that* $\aleph_0 \leqq wK \leqq 2^{\mathfrak{n}}$. *Then* K *contains a dense, countably compact subgroup* H *such that* $|H| \leqq \mathfrak{n}^{\aleph_0}$.

Proof. This follows from Corollary 1.5, upon taking $\mathfrak{m} = \aleph_1$.

We note that Corollary 1.6 may be proved by appealing to Wilcox's theorem in place of the result of Kuzminov. If K is given as in Corollary 1.6 and H is a dense, pseudocompact subgroup of K with $|H| \leqq \mathfrak{n}^{\aleph_0}$ (as afforded by H. Wilcox [28]) then according to Corollary 1.2 above applied to the pair (H, K) there is a countably compact group H' for which

$$H \subset H' \subset \bar{H} = K$$

and $|H'| \leqq (\mathfrak{n}^{\aleph_0})^{\aleph_0} = \mathfrak{n}^{\aleph_0}$.

COROLLARY 1.7. *Assume the generalized continuum hypothesis. If* $\mathfrak{n}$ *is a cardinal and* K *is an infinite compact group such that* $|K| = 2^{2^{\mathfrak{n}}}$, *then there is a dense, countably compact subgroup* H *of* K *such that* $|H| \leqq \mathfrak{n}^{\aleph_0}$.

Proof. It is known that $|K| = 2^{wK}$. (A direct proof is given by H. Wilcox [27]. Earlier Hulanicki [14] [15], using essentially an argument of Čech and Pospísil [4], showed that $|K| = 2^{\theta K}$ where θK denotes the smallest cardinal which is the cardinality of a family $\mathscr{U}$ of open subsets of G such that $|\bigcap \mathscr{U}| = 1$. Since $\theta K = wK$—see Hewitt and Ross [13], pp. 99–100—we have again $|K| = 2^{wK}$.) From the generalized continuum hypothesis it follows that $wK = 2^{\mathfrak{n}}$, so Corollary 1.6 applies.

We have shown in this section that several of the pseudocompact groups considered in [28] may in fact be taken to be countably compact. We close with an example showing that not all of the conclusions of [28] may be strengthened in this manner.

We continue the notational convention used earlier: If G is a group and $S \subset G$, then $\langle S\rangle$ denotes the subgroup of G generated by S. For $x \in G$ we write $\langle x\rangle$ in place of $\langle\{x\}\rangle$.

Here and later the symbol $\boldsymbol{T}$ denotes the circle group

$$\boldsymbol{T} = \{z\colon z \text{ is a complex number and } |z| = 1\} ,$$

and $\boldsymbol{Q}$ is the "rational subgroup" of $\boldsymbol{T}$—i.e.,

$$\boldsymbol{Q} = \{z \in \boldsymbol{T}\colon \arg z \text{ is rational}\} .$$

DEFINITION. Let G be a group and $x \in G$. Then x is a metric element of G if $\mathrm{cl}_G\langle x\rangle$ is metrizable.

THEOREM 1.8. *Let M be the set of metric elements of the group $\boldsymbol{T}^{\aleph_1}$. Then M is not a countably compact group.*

Proof. We have $\boldsymbol{Q} \subset \boldsymbol{T}$, and hence $\boldsymbol{Q}^{\aleph_1} \subset \boldsymbol{T}^{\aleph_1}$. It is easy to see that every element of $\boldsymbol{T}^{\aleph_1}$ is the limit of a sequence of elements of $\boldsymbol{Q}^{\aleph_1}$. (In detail: Let $p \in \boldsymbol{T}^{\aleph_1}$ and for $\xi < \aleph_1$ and each integer $n > 0$ let $q_\xi^{(n)}$ be chosen in $\boldsymbol{Q}$ so that

$$|q_\xi^{(n)} - p_\xi| < 1/n .$$

Then $q^{(n)} \in \boldsymbol{Q}^{\aleph_1}$, and $q^{(n)} \longrightarrow p$.) Thus it suffices to show

(a) $\boldsymbol{Q}^{\aleph_1} \subset M$; and

(b) $M \subsetneqq \boldsymbol{T}^{\aleph_1}$.

For (a) let $x \in \boldsymbol{Q}^{\aleph_1}$ and let $S \subset \aleph_1$ have the property that $|S| \leqq \aleph_0$ and for each $\xi < \aleph_1$ there is $\eta \in S$ such that $x_\xi = x_\eta$. We claim that the natural projection $\pi\colon \boldsymbol{T}^{\aleph_1} \longrightarrow \boldsymbol{T}^S$ is one-to-one on $\langle x\rangle$. If $x^m \neq x^n$ there is $\xi < \aleph_1$ for which $x_\xi^m \neq x_\xi^n$, and then choosing $\eta \in S$ such that $x_\xi = x_\eta$ have

$$(\pi(x^m))_\eta = (x^m)_\eta = (x^m)_\xi \neq (x^n)_\xi = (x^n)_\eta = (\pi(x^n))_\eta ;$$

thus $\pi(x^m) \neq \pi(x^n)$ and the claim is established.

We claim next that π is one-to-one on the closure in $\boldsymbol{T}^{\aleph_1}$ of $\langle x\rangle$. Indeed if $p, q \in \mathrm{cl}\,\langle x\rangle$ with $p_\xi \neq q_\xi$ for some $\xi < \aleph_1$, then upon choosing $\eta \in S$ such that $x_\xi = x_\eta$ we note that the projections π_ξ and π_η from $\boldsymbol{T}^{\aleph_1}$ onto $\boldsymbol{T}_\xi$ and $\boldsymbol{T}_\eta$ respectively agree on x, hence on $\langle x\rangle$, hence on cl $\langle x \rangle$, hence at p and q. Thus

$$(\pi(p))_\eta = \pi_\eta(p) = \pi_\xi(p) = p_\xi \neq q_\xi = \pi_\xi(q) = \pi_\eta(q) = (\pi(q))_\eta ,$$

so $\pi(p) \neq \pi(q)$.

Thus π is a one-to-one, continuous function from the compact group cl $\langle x \rangle$ into the metrizable group $\boldsymbol{T}^S$.

Thus the function π, when restricted to the compact group cl $\langle x \rangle$, is a one-to-one continuous function into the metrizable group $\boldsymbol{T}^S$. This restricted function is then a homeomorphism, cl $\langle x \rangle$ is metrizable, and $x \in M$. Assertion (a) is proved.

For (b) it suffices to cite from [12] pp. 407–408 the familiar fact that there exists $x \in \boldsymbol{T}^{\aleph_1}$ such that cl $\langle x \rangle = \boldsymbol{T}^{\aleph_1}$. Since $\boldsymbol{T}^{\aleph_1}$ is not metrizable, we have $x \in \boldsymbol{T}^{\aleph_1} \backslash M$.

2. **Finest totally bounded topologies.** Throughout this section the word group refers simply to a non-empty set together with a multiplication and inversion satisfying the usual group axioms; no topology is assured. Topological groups are denoted by the symbols $(G, \mathscr{T})$, (G, t) and the like. It is assumed that these satisfy the Hausdorff separation axiom.

It is known (see for example Dixmier [8], p. 296 ff.; Kurosh [19], p. 157; von Neumann [25]; T. Wilcox [29]; and Hewitt and Ross [12], pp. 348–351) that there are groups G with the property that for no topology $\mathscr{T}$ on G is $(G, \mathscr{T})$ a totally bounded topological group. But if G is an Abelian group then, because there are sufficiently many homomorphisms from G to the circle group $\boldsymbol{T}$ to separate points of G, the group G may be embedded algebraically into a product of copies of $\boldsymbol{T}$ and therefore there is a totally bounded topology $\mathscr{T}$ on G relative to which $(G, \mathscr{T})$ is a topological group. According to Comfort and Ross [5], the totally bounded group topologies on the Abelian groups G are precisely the topologies induced on G by point-separating group of homomorphisms into $\boldsymbol{T}$; the finest such topology is the one induced by the group of all such homomorphisms.

It is well-known [8] that if a (not necessarily Abelian) group G admits a totally bounded group topology $\mathscr{T}$ then it admits a (necessarily unique) finest such topology. We denote this latter topology on G, when it exists, by the symbol t. It is not difficult to see that (G, t) has the property that each homomorphism from (G, t) to a totally bounded group is continuous. Indeed t may be defined as follows: Let $\{(H_i, f_i): i \in I\}$ be a listing of all pairs (H, f) with H a totally bounded topological group and f a homomorphism from G onto a dense subset of H, and let

$$e: G \longrightarrow P = \prod_{i \in I} H_i$$

be defined by the rule

$$(ex)_i = f_i(x) ;$$

then e is a one-to-one map because (G, id) is one of the pairs (H_i, f_i), and t is the topology induced on G (more precisely: on $e[G]$) by P.

It is clear from the foregoing remarks that the finest totally bounded group topology t on G is characterized by the property that each homomorphism from G to a totally bounded group is t-continuous.

In this section we prove that for each infinite Abelian group G the topological group (G, t) is not pseudocompact. This improves an observation made in 1.8 of [5]. We show also that a product of infinitely many nontrivial totally bounded Abelian topological groups does not have its finest totally bounded topology.

LEMMA 2.1. *Let G be an Abelian group and H a subgroup of G. Then H is t-closed in G.*

Proof. If $x \in G \backslash H$ then H and xH are different elements of G/H so there is a homomorphism χ from G/H into $\boldsymbol{T}$ such that $\chi(xH) \neq 1$. If φ denotes the natural mapping from G onto G/H then $\chi \circ \varphi$ is a homomorphism from G to $\boldsymbol{T}$ and

$$x \notin (\chi \circ \varphi)^{-1}(\{1\}) \supset H .$$

The result now follows from the fact that $\chi \circ \varphi$ is t-continuous, so that $(\chi \circ \varphi)^{-1}(\{1\})$ is a closed subset of (G, t).

THEOREM 2.2. *Let G be an infinite Abelian group. Then (G, t) is not a pseudocompact topological group.*

Proof. It is well-known and easy to prove from standard structure theorems (see for example [12], page 227) that there is a subgroup H of G such that $|G/H| = \aleph_0$. If (G, t) were pseudocompact then G/H in the usual quotient topology would be pseudocompact (being the continuous image of G), a Hausdorff space (because H is closed by Lemma 2.1), and countable. Since G/H is a pseudocompact, Lindelöf space it is countably compact ([10], Exercise 3D); indeed, it is compact ([10], Theorem 8.2 and Exercise 5H). But this is impossible, since an infinite countably compact group has cardinality at least $2^{\aleph_0}$ ([12], page 31).

REMARKS 2.3. (a) An early version of this paper showed only that (G, t) as in Theorem 2.2 could not be countably compact, and left unsettled the question whether (G, t) might be pseudocompact.

We are grateful to Lew Robertson for formulating the argument given above, which shows in effect that a pseudocompact group never contains a closed, normal subgroup of countably infinite index.

(b) If the word Abelian is omitted from Lemma 2.1 or from Theorem 2.2, the resulting statements are false. It has been pointed out to us by Lew Robertson that according to a result of van der Waerden [24] the real, special orthogonal group $\mathrm{SO}(3, \boldsymbol{R})$, which is an infinite, compact, connected, Lie group, admits no discontinuous homomorphism into any compact group. (It follows from (22.13), (22.14), and (22.22.h) of [12] that the complex special linear group $\mathrm{SL}(2, \boldsymbol{C})$ admits no algebraic isomorphism, continuous or discontinuous, into any compact group. Such a group is said to be minimal almost periodic; see [25].) According to the discussion preceding 2.1, then, this compact, metric topology on the (non-Abelian) group $\mathrm{SO}(3, \boldsymbol{R})$ is the finest totally bounded topology t on $\mathrm{SO}(3, \boldsymbol{R})$. Since $\mathrm{SO}(3, \boldsymbol{R})$ contains copies of $\boldsymbol{T}$—and hence also non-closed copies of $\boldsymbol{Q}$—not every subgroup of $\mathrm{SO}(3, \boldsymbol{R})$ is t-closed.

There is another property relating to finest totally bounded topologies which, though it fails for the non-Abelian group $\mathrm{SO}(3, \boldsymbol{R})$, holds for each Abelian group: According to Theorem 2.2 those closed copies of $\boldsymbol{T}$ inside $\mathrm{SO}(3, \boldsymbol{R})$ do not inherit their own finest totally bounded topology. But if H is any (necessarily t-closed) subgroup of a topological group (G, t) with G Abelian then the topology induced on H is its finest totally bounded topology. To prove this it suffices, according to Theorem 1.7 of [5], to show that each homomorphism $f: H \to \boldsymbol{T}$ is continuous in this induced topology. Because $\boldsymbol{T}$ is divisible such a homomorphism f extends to a homomorphism $\bar{f}: G \to \boldsymbol{T}$; since $\bar{f}$ is t-continuous on G its restriction to H is also continuous.

Our next result answers a question suggested by 2.1 and 2.2. The construction follows an argument given in Theorem 2.3 of [28], and is clearly susceptible to substantial generalization.

THEOREM 2.4. *There is an Abelian pseudocompact group with a closed subgroup which is not pseudocompact.*

Proof. Let $K = \boldsymbol{T}^{\aleph_1}$ and let

$$H = \{x \in K : |\{\xi < \aleph_1 : x_\xi \neq 1\}| \leqq \aleph_0\} .$$

(H is an example of what Corson [7] calls a Σ-space.) That H is countably compact is seen as in [7] or [18]: If $A \subset H$ and $|A| \leqq \aleph_0$ then for some countable subset S of $\aleph_1$ we have

$$A \subset (\textstyle\prod_{\xi \in S} \boldsymbol{T}_\xi) \times \prod_{\xi \in \aleph_1 \setminus S} \{1\}_\xi \subset H ,$$

so that each countable subset of H is contained in a compact subspace of H.

Let $z \in \boldsymbol{T}$ have the property that $z^n \neq 1$ for each integer n and let p be that element of K for which $p_\xi = z$ for all $\xi < \aleph_1$; and let J denote the subgroup of K generated by p.

Now let G be the subgroup of K generated by H and J. Clearly

$$(*) \qquad G = \{x \in K \colon |\{\xi < \aleph_1 \colon x_\xi \neq z^n\}| \leqq \aleph_0 \text{ for some integer } n\} .$$

We complete the proof by showing

(a) the group G is pseudocompact;

(b) J is a closed subgroup of G; and

(c) J is not pseudocompact.

For (a) we note that H is dense in G (because $H \subset G \subset K$ and H is dense in K) and that H is countably compact and hence pseudocompact. Thus G is pseudocompact.

For (b) we note that since the subgroup of $\boldsymbol{T}$ generated by z is dense in $\boldsymbol{T}$, we have

$$(**) \qquad \mathrm{cl}_K J = \{x \in K \colon x_\xi = x_\eta \quad \text{for all} \quad \xi, \eta < \aleph_1\} .$$

From (*) and (**) it follows that

$$\mathrm{cl}_K J \cap G = J ,$$

so that J is closed in G.

For (c) we note that J is (homeomorphic with) the group $\{z^n \in T \colon n$ is an integer$\}$. This countable, infinite group is obviously not pseudocompact.

The proof is complete.

THEOREM 2.5. *Let $\{(G_i, t_i) \colon i \in I\}$ be a family of groups G_i each with its finest totally bounded topology t_i. If $|I| < \aleph_0$ and $(G, \mathscr{T})$ is the product of the spaces (G_i, t_i), then the totally bounded topology $\mathscr{T}$ is the finest totally bounded group topology for G.*

Proof. It suffices to treat the case $I = \{1, 2\}$. Let f be a homomorphism from $G = G_1 \times G_2$ to a totally bounded group H, and let U be a neighborhood in H of the identity e of H. Let V be a neighborhood of e such that $V^2 \subset U$, and let f_1 and f_2 be defined from G_1 and G_2 respectively to H by the rules

$$f_1(g_1) = f(g_1, e_2),\ f_2(g_2) = f(e_1, g_2) .$$

Because t_1 and t_2 are the finest totally bounded topologies on G_1 and G_2 respectively, the homomorphisms f_1 and f_2 are continuous. Thus there are neighborhoods W_1 and W_2 of the identity elements e_1 and e_2

such that

$$f_1[W_1] \subset V \quad \text{and} \quad f_2[W_2] \subset V .$$

It is now clear that

$$f[W_1 \times W_2] \subset V^2 \subset U .$$

We conclude that f is continuous on $(G, \mathscr{T})$. Thus $\mathscr{T}$ is the finest totally bounded topology for G.

Our final result is in contrast with Theorem 2.5.

THEOREM 2.6. *Let $\{(G_i, \mathscr{T}_i): i \in I\}$ be a family of totally bounded Abelian groups with $|G_i| \geqq 2$ for $i \in I$. If $|I| \geqq \aleph_0$ and $(G, \mathscr{T})$ is the product of the spaces $(G_i, \mathscr{T}_i)$, then the totally bounded topology $\mathscr{T}$ is not the finest totally bounded group topology for G.*

Proof. Let e_i be the identity element of G_i, and let

$$H = \{x \in G: |\{i \in I: x_i \neq e_i\}| < \aleph_0\} .$$

Then H is a dense, proper subgroup of $(G, \mathscr{T})$. The result now follows from Lemma 2.1.

REFERENCES

1. P. S. Aleksandrov, *Some results in the theory of topological spaces, obtained within the last twenty-five years*, Uspehi Mat. Nauk, **15**, 2 (1960), 25–95 (= Russian Math. Surveys 15 No. 2 (1960), 23–83).

2. P. Alexandroff and P. S. Urysohn, *Mémoire sur les espaces topologiques compacts*, Verh. Akad. Wetensch. Amsterdam, **14** (1929), 1–96.

3. Heinz Bachmann, *Transfinite Zahlen*, Springer-Verlag, Berlin, 1955.

4. Eduard Čech and Bedřich Pospisil, I. *Sur les espaces compacts, II, Sur les caractères des points dans les espaces $\mathscr{L}$*, Publ. Fac. Sci. Univ. Masaryk 1938, No. 258, 14 pp. Also in Topological papers of Eduard Čech, 495–505, Publ. House Czech. Acad. Sci. Prague, 1968.

5. W. W. Comfort and K. A. Ross, *Topologies induced by groups of characters*, Fundamenta Math., **55** (1964), 283–291.

6. ———, *Pseudocompactness and uniform continuity in topological groups*, Pacific J. Math., **16** (1966), 483–496.

7. H. H. Corson, *Normality in subsets of product spaces*, Amer. J. Math., **81** (1959), 785–796.

8. J. Dixmier, *Les C*-algèbres et leurs representations*, Gauthier-Villars, Paris, 1969.

9. R. Engelking, *Outline of General Topology*, North-Holland Publishing Company, Amsterdam, 1968.

10. Leonard Gillman and Meyer Jerison, *Rings of Continuous Functions*, D. Van Nostrand Company, Inc., Princeton, N. J., 1960.

11. E. Hewitt, *A remark on density character*, Bull. Amer. Math. Soc., **52** (1946), 641–643.

12. Edwin Hewitt and Kenneth A. Ross, *Abstract Harmonic Analysis I*, Springer-Verlag, Berlin, 1963.

13. Edwin Hewitt and Kenneth A. Ross, *Abstract Harmonic Analysis II*, Springer-Verlag, Berlin, 1970.
14. A. Hulanicki, *On cardinal numbers related with locally compact groups*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astr. Phys., **6** (1558), 67–70.
15. ———, *On the power of compact spaces*, Colloq. Math., **7** (1960), 199–200.
16. G. L. Itzkowitz, *Extensions of Haar measure for compact connected Abelian groups*, Bull. Amer. Math. Soc., **71** (1965), 152–156.
17. ———, *Extensions of Haar measure for compact connected Abelian groups*, Indag. Math., **27** (1965), 190–207.
18. J. M. Kister, *Uniform continuity and compactness in topological groups*, Proc. Amer. Math. Soc., **13** (1962), 37–40.
19. A. G. Kurosh, *The Theory of Groups*, Vol. 2, Chelsea Publishing, New York, 1956.
20. V. Kuzminov, *On a hypothesis of P. S. Aleksandrov in the theory of topological groups*, Doklady Akad. Nauk SSSR, **125** (1959), 727–729.
21. A. S. Mishchenko, *On finally compact spaces*, Doklady Akad. Nauk SSSR, **145** (1962), 1224–1227 (= Soviet Math. Doklady 3 (1962), 1199–1202).
22. E. S. Pondiczery, *Power problems in abstract spaces*, Duke Math. J., **11** (1944), 835–837.
23. Victor Saks, *Countably compact groups*, Doctoral dissertation, Wesleyan University, 1972.
24. B. L. van der Waerden, *Stetigkeitssätze für halbeinfache Liesche Gruppen*, Math. Zeitschrift, **36** (1933), 780–786.
25. J. von Neumann, *Almost periodic functions in a group, I*, Trans. Amer. Math. Soc., **36** (1934), 445–492.
26. André Weil, *Sur les espaces à structure uniforme et sur la topologie générale*, Publ. Math. Univ. Strasbourg, Hermann, Paris, 1937.
27. Howard J. Wilcox, *Pseudocompact groups*, Doctoral dissertation, University of Rochester, Rochester, New York, 1965.
28. ———, *Pseudocompact groups*, Pacific J. Math., **19** (1966), 365–379.
29. Theodore W. Wilcox, *On the structure of maximally almost periodic groups*, Bull. Amer. Math. Soc., **73** (1967), 732–734.

Received July 7, 1972. Parts of this paper appeared in [23], written by the second-listed author under the supervision of the first-listed author. We are grateful to the referee for several suggestions improving our exposition. The first author gratefully acknowledges support received from the National Science Foundation under grant NSF-GP-39263.

WESLEYAN UNIVERSITY
AND
UNIVERSITY OF COSTA RICA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

MAXIMAL INVARIANT SUBSPACES OF STRICTLY CYCLIC OPERATOR ALGEBRAS

MARY R. EMBRY

A *strictly cyclic* operator algebra $\mathscr{A}$ on a complex Banach space $X(\dim X \geqq 2)$ is a uniformly closed subalgebra of $\mathscr{L}(X)$ such that $\mathscr{A}x = X$ for some x in X. In this paper it is shown that (i) if $\mathscr{A}$ is strictly cyclic and intransitive, then $\mathscr{A}$ has a maximal (proper, closed) invariant subspace and (ii) if $A \in \mathscr{L}(X)$, $A \neq zI$ and $\{A\}'$ (the commutant of A) is strictly cyclic, then A has a maximal hyperinvariant subspace.

1. Notation and terminology. Throughout the paper X is a complex Banach space of dimension greater than one and $\mathscr{L}(X)$ is the algebra of continuous linear operators on X. $\mathscr{A}$ will denote a uniformly closed subalgebra of $\mathscr{L}(X)$ which is *strictly cyclic* and x_0 will be a *strictly cyclic vector* for $\mathscr{A}$: that is, $\mathscr{A}x_0 = X$. We do not insist that the identity element I of $\mathscr{L}(X)$ be an element of $\mathscr{A}$.

If $\mathscr{B} \subset \mathscr{L}(X)$, then the *commutant* of $\mathscr{B}$ is $\mathscr{B}' = \{E: E \in \mathscr{L}(X)$ and $EB = BE$ for all B in $\mathscr{B}\}$. We shall use the terminology of "invariant" and "transitive" as follows: if $M \subset X$ and $\mathscr{B} \subset \mathscr{L}(X)$, then (i) M is *invariant* under $\mathscr{B}$ if $\mathscr{B}M = \{Bm: B \in \mathscr{B}$ and $m \in M\} \subset M$, (ii) M is an *invariant subspace* for $\mathscr{B}$ if M is invariant under $\mathscr{B}$ and M is a closed, nontrivial ($\neq \{0\}$, X) linear subspace of X, (iii) $\mathscr{B}$ is *transitive* if $\mathscr{B}$ has no invariant subspace and *intransitive* if $\mathscr{B}$ has an invariant subspace. Further, if $A \in \mathscr{L}(X)$ and $\{A\}'$ is intransitive, then each invariant subspace of $\{A\}'$ is called a *hyperinvariant subspace* of A. Finally an invariant subspace of $\mathscr{B}$ is *maximal* if it is not properly contained in another invariant subspace of $\mathscr{B}$.

2. Introduction. Strictly cyclic operator algebras have been studied by A. Lambert, D. A. Herrero, and the auther of this paper. (See for example [2]-[6].) One of the major results in [2, Theorem 3.8], [3, Theorem 2], and [6, Theorem 4.5] is that a transitive subalgebra of $\mathscr{L}(X)$ containing a strictly cyclic algebra is necessarily strongly dense in $\mathscr{L}(X)$. In each of three developments the following is a key lemma: The only dense linear manifold invariant under a strictly cyclic subalgebra of $\mathscr{L}(X)$ is X. In Lemma 1 we shall present a generalization of this lemma which will be useful in the study of maximal invariant subspaces and noncyclic vectors of a strictly cyclic algebra $\mathscr{A}$.

LEMMA 1. *If M is invariant under $\mathscr{A}$ and $x_0 \in \bar{M}$, then $M = X$.*

(It should be noted that we do not require M to be linear nor do we require, as was done in Lemma 3.4 of [2], that $I \in \mathscr{A}$. The proof given here is a slight modification of that given in [2].)

Proof. We shall show that $\mathscr{A}x_0 \subset M$ and thus $X = \mathscr{A}x_0 \subset M$. Let $\{x_n\}$ be a sequence in M such that $\lim_{n\to\infty} x_n = x_0$. By [2, Lemma 3.1 (ii)] there exists a sequence $\{A_n\}$ in $\mathscr{A}$ such that $A_n x_0 = x_0 - x_n$ and $\lim_{n\to\infty} ||A_n|| = 0$. Thus for n sufficiently large, $||A_n|| < 1$ and $(I - A_n)^{-1} = \sum_{k=0}^{\infty} (A_n)^k$. Consequently, $\mathscr{A}(I - A_n)^{-1} \subset \mathscr{A}$ and since $x_0 = (I - A_n)^{-1}x_n$, we have $\mathscr{A}x_0 = \mathscr{A}(I - A_n)^{-1}x_n \subset \mathscr{A}x_n \subset M$, as desired.

For the sake of future reference we restate and reprove the transitivity theorem.

THEOREM 1. *If $\mathscr{A}$ is a strictly cyclic transitive subalgebra of $\mathscr{L}(X)$, then $\mathscr{A}$ is strongly dense in $\mathscr{L}(X)$.*

Proof. Using Lemma 1 we can show (as in [2, Lemma 3.5]) that each densely defined linear transformation commuting with $\mathscr{A}$ is everywhere defined and continuous. Further, again using Lemma 1, we can show that if $E \in \mathscr{A}$ and $z \in \sigma(E)$, then either $zI - E$ is not one-to-one or does not have dense range. Thus if $\mathscr{A}$ is transitive, necessarily $E = zI$. Consequently, it follows from [1, p. 636 and Cor. 2.5, p. 641] that $\mathscr{A}$ is strongly dense in $\mathscr{L}(X.)$

3. **Maximal invariant subspaces.** In [2, Theorem 3.1] it is shown that every strictly cyclic, separated operator algebra $\mathscr{A}$ has a maximal invariant subspace. ($\mathscr{A}$ is separated by x_0 if $A = 0$ whenever $A \in \mathscr{A}$ and $Ax_0 = 0$.) Theorem 2 allows us to obtain the same result without the hypothesis that $\mathscr{A}$ be separated, provided $\mathscr{A}$ is intransitive.

THEOREM 2. *An intransitive, strictly cyclic subalgebra $\mathscr{A}$ of $\mathscr{L}(X)$ has a maximal invariant subspace.*

Proof. Let $\mathscr{M} = \{M: M$ is an invariant subspace of $\mathscr{A}\}$. By hypothesis $\mathscr{M} \neq \emptyset$. We shall order $\mathscr{M}$ by set inclusion and show that each linearly ordered subset of $\mathscr{M}$ has an upper bound in $\mathscr{M}$. To this end we let $\{M_\alpha\}$ be a linearly ordered subset of $\mathscr{M}$. Then $\bigcup_\alpha M_\alpha$ is invariant under $\mathscr{A}$. By Lemma 1, if $\overline{\bigcup_\alpha M_\alpha} = X$, then $\bigcup_\alpha M_\alpha = X$ and consequently $x_0 \in M_\alpha$ for some value of α. Since this last implies that $X = \mathscr{A}x_0 \subset \mathscr{A}M_\alpha \subset M_\alpha$ and contradicts the fact that M_α is a proper closed linear subspace of X, we see that $\bigcup_\alpha M_\alpha$ is not dense in X. Thus $\overline{\bigcup_\alpha M_\alpha}$ is an element of $\mathscr{M}$ and is an upper bound for $\{M_\alpha\}$. By the Maximality Principle $\mathscr{M}$ has a maximal element.

Lemma 1 and the Maximality Principle can be combined to arrive at other similar results. For example, (i) if $\mathscr{A}$ is intransitive and strictly cyclic, then $\mathscr{A}$ has a proper maximal invariant subset (this will be discussed further in §4) and (ii) if X is a Hilbert space and $\mathscr{A}$ has a reducing subspace (that is, an invariant subspace of $\mathscr{A}$ which is also invariant under $\mathscr{A}^* = \{A^*: A \in \mathscr{A}\}$), then $\mathscr{A}$ has a maximal reducing subspace.

In [2, Theorem 3.7] it is shown that if A is not a scalar multiple of I and $\{A\}'$ is strictly cyclic, then A has a hyperinvariant subspace. This result combined with Theorem 2 yields the following:

COROLLARY 1. *If A is not a scalar multiple of I and $\{A\}'$ is strictly cyclic, then A has a maximal hyperinvariant subspace.*

We shall now turn our attention to intransitive, strictly cyclic operator algebras on a Hilbert space X. If M is a closed linear subspace of X, P_M will denote the orthogonal projection of X onto M and $M^\perp$ the orthogonal complement of M: $M^\perp = \{y: \langle y, m\rangle = 0$ for all m in $M\}$. Furthermore, $\mathscr{A}^* = \{A^*: A \in \mathscr{A}\}$.

In the Hilbert space situation we are able to conclude that $\mathscr{A}^*/M$ is strongly dense in $\mathscr{L}(M^\perp)$ when M is a maximal invariant subspace for $\mathscr{A}$. This remains an open question if X is an arbitrary Banach space and is a particularly interesting one if X is reflexive. For in that case if M is a maximal invariant subspace of $\mathscr{A}$, then $M^\perp = \{x^*: x^*(M) = 0\}$ is a minimal invariant subspace of $\mathscr{A}^*$.

THEOREM 3. *Let $\mathscr{A}$ be a strictly cyclic operator algebra on a Hilbert space X. If M is a maximal invariant subspace of $\mathscr{A}$, then* (i) $(I - P_M)\mathscr{A}(I - P_M)x_0 = M^\perp$ *and* (ii) $\mathscr{A}^*(I - P_M)$ *is strongly dense in* $\mathscr{L}(M^\perp)$.

Proof. Note first that $(I - P_M)\mathscr{A}(I - P_M) = (I - P_M)\mathscr{A}$, so that (i) is immediate. Since M is a maximal invariant subspace for $\mathscr{A}$, $M^\perp$ is a minimal invariant subspace for $\mathscr{A}^*$. Thus each of $\mathscr{A}^*(I - P_M)$ and $(I - P_M)\mathscr{A}(I - P_M)$ is transitive on $M^\perp$. Thus the uniform closure of $(I - P_M)\mathscr{A}(I - P_M)$ in $\mathscr{L}(M^\perp)$ is transitive and by (i) is strictly cyclic; hence by Theorem 1 $(I - P_M)\mathscr{A}(I - P_M)$ is strongly dense in $\mathscr{L}(M)$, which concludes our proof of (ii).

THEOREM 4. *Let X be a Hilbert space, $A \in \mathscr{L}(X)$ and $\{A\}'$ strictly cyclic. If M is a maximal invariant subspace for $\{A\}'$, then there exists a multiplicative linear functional f on $\{A\}''$ such* *that for each E in $\{A\}''$, $(E - f(E)I)(X) \subset M$.*

Proof. As we noted in the proof of Theorem 3,

$$\mathscr{B} = (I - P_M)\{A\}'(I - P_M)$$

is strongly dense in $\mathscr{L}(M^{\perp})$ and thus its commutant consists of the scalar multiples of the identity operator on $M^{\perp}$. Since $\{A\}'' \subset \{A\}'$ and M is invariant under $\{A\}'$, we know that $(I - P_M)\{A\}''(I - P_M)$ is contained in the commutant of $\mathscr{B}$ on $M^{\perp}$ and hence $(I - P_M)\{A\}''(I - P_M) \subset \{z(I - P_M)\}$. Thus for E in $\{A\}''$, there exists a complex number z such that $(I - P_M)E(I - P_M) = z(I - P_M)$. Therefore, $(I - P_M)(E - zI) = 0$ since M is invariant under $\{A\}''$; or equivalently $(E - zI)(X) \subset M$. Since M is a proper subset of X, it is now obvious that the number z for which $(E - zI)(X) \subset M$ is unique. Define $f(E) = z$.

That f is linear follows immediately from the fact that $f(E)$ is the unique number for which $(E - f(E)I)(X) \subset M$. Furthermore, since M is invariant under $\{A\}''$, $(FE - f(E)F)(X) \subset M$ for all $E, F \in \{A\}''$. Consequently (by uniqueness again), $0 = f(FE - f(E)F) = f(FE) - f(E)f(F)$ and thus we see that f is multiplicative.

COROLLARY 2. *Let $A \in \mathscr{L}(X)$ where X is a Hilbert space. If the range of $A - zI$ is dense in X for each complex z, then $\{A\}'$ is not strictly cyclic.*

Proof. Except for one minor technicality, Corollary 2 follows immediately from Theorem 4. For, if $\{A\}'$ is strictly cyclic and intransitive, by Theorem 4 there exists a complex number $f(A)$ such that the range of $A - f(A)I$ is contained in a proper subspace of X. By Corollary 1 the only other way in which $\{A\}'$ can be strictly cyclic is when $A = zI$ for some complex z, in which case the range of $A - zI$ is certainly not dense in X.

In [2, Lemma 3.6] and [3, Proposition 2], it is shown that if $E \in \mathscr{A}'$, where $\mathscr{A}$ is strictly cyclic and $z \in \sigma(E)$, then either $zI - E$ is not one-to-one or $zI - E$ does not have dense range. Corollary 2 now adds to our knowledge of $\sigma(A)$ where $\{A\}'$ is strictly cyclic: in this case we know that for at least one value of z, the range of $A - zI$ is nondense. Indeed we have the stronger result:

COROLLARY 3. *Let $A \in \mathscr{L}(X)$ where X is a Hilbert space. If $\{A\}'$ is strictly cyclic, then there exists a common eigenvector for $\{A^*\}''$.*

Proof. The case in which $\{A\}' = \mathscr{L}(X)$ is trivial. Thus we assume $A \neq zI$. By Theorem 4 if $E \in \{A\}''$, there exists a complex number $f(E)$ such that $(E - f(E)I)(X) \subset M$ where M is a maximal invariant subspace of $\{A\}'$. Therefore, $E^*(I - P_M)x_0 = f(E)^*(I - P_M)x_0$ and $(I - P_M)x_0 \neq 0$ since x_0 is cyclic for $\{A\}'$ and M is a proper invariant subspace for $\{A\}'$.

4. **Noncyclic vectors of $\mathscr{A}$.** In this last section of this paper we shall discuss briefly several properties of the set of noncyclic vectors of a strictly cyclic operator algebra $\mathscr{A}$. A vector x is noncyclic for $\mathscr{A}$ if $\mathscr{A}x$ is not dense in X. These results are summarized in Theorem 5. Parts (i) and (iii) of Theorem 5 also are found in [5, Theorem 2].

THEOREM 5. *Let N be the set of noncyclic vectors of a strictly cyclic operator algebra $\mathscr{A}$,*

(i) *if $x \notin N$, then x is a strictly cyclic vector for $\mathscr{A}$,*

(ii) *N is invariant under $\mathscr{A}$,*

(iii) *N is closed in X,*

(iv) *N is the unique proper maximal invariant subset of $\mathscr{A}$,*

(v) *if N is not linear, then $N + N = X$, where $N + N = \{x + y: x, y \in N\}$.*

Proof. (i) If $x \notin N$, then $\overline{\mathscr{A}x} = X$ and thus by Lemma 1 since $\mathscr{A}x$ is invariant under $\mathscr{A}$, we have $\mathscr{A}x = X$ and x is strictly cyclic. (ii) Assume that $x \in N$ and $A \in \mathscr{A}$. Then $\mathscr{A}Ax \subset \mathscr{A}x$ and consequently $\mathscr{A}Ax \neq X$. That is, $Ax \in N$ for each A in $\mathscr{A}$ which proves (ii). (iii) By (ii) $\mathscr{A}N \subset N$. Since $\mathscr{A}$ has a strictly cyclic vector, we know by Lemma 1 that $\bar{N}$ contains no strictly cyclic vector for $\mathscr{A}$. Thus by (i) $\bar{N}$ contains only noncyclic vectors for $\mathscr{A}$, which says that N is closed. (iv) By (ii) N is invariant under $\mathscr{A}$. By hypothesis $\mathscr{A}$ has a strictly cyclic vector so that $N \neq X$. These two observations essentially prove (iv) since an element x of a proper invariant subset of $\mathscr{A}$ is necessarily an element of N. (v) If N is nonlinear, then since N is homogeneous, we know that $N \neq N + N$. Therefore, since $N + N$ is invariant under $\mathscr{A}$(by (ii) we know that $N + N = X$ by (iv)).

To see that there exist strictly cyclic operator algebras for which N is linear and those for which N is nonlinear let us reconsider Example 1 of [2].

EXAMPLE. Let X be a Banach space, $\dim X \geqq 2$ and let $x_0 \in X$, $x_0 \neq 0$. Let each of x^* and y^* be a continuous linear functional on X such that $x^*(x_0) = y^*(x_0) = 1$. For each x in X define A_x by

$$A_x y = x^*(x)[y - y^*(y)x_0] + y^*(y)x$$

and let $\mathscr{A} = \{A_x : x \in X\}$.

It was observed in [2] that $\mathscr{A}$ is a strictly cyclic operator algebra with strictly cyclic, separating vector x_0.

A simple argument shows that a vector y_0 of X is cyclic (and hence by Theorem 5 strictly cyclic) if and only if $y^*(y_0) \neq 0$ and $x^*(y_0) \neq 0$. Thus the set N of noncyclic vectors coincides with $\ker y^* \cup \ker x^*$. Consequently, N is linear if x^* and y^* are dependent and nonlinear otherwise.

References

1. W. B. Arveson, *A density theorem for operator algebras*, Duke Math. J., **34** (1967), 635–647.

2. M. R. Embry, *Strictly cyclic operator algebras on a Banach space*, Pacific J. Math., **45** (1973), 443–452.

3. D. A. Herrero, *Algebras de Operadores Transitivas que Contienen una Subalgebra de Multiplicidad Estricta Finita*, Revista de la Union Mat. Argentina, (to appear).

4. ———, *Operator algebras of finite strict multiplicity*, Indiana Univ. Math. J., **22** (1972), 13–24.

5. D. A. Herrero and A. Lambert, *On Strictly Cyclic Algebras, P-Algebras and Reflexive Operators*, preprint.

6. A. Lambert, *Strictly cyclic operator algebras*, Pacific J. Math., **39** (1971), 717–726.

Received July 14, 1972 and in revised form August 25, 1972.

UNIVERSITY OF NORTH CAROLINA AT CHARLOTTE

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

CONGRUENCE LATTICES OF SEMILATTICES

RALPH FREESE AND J. B. NATION

The main result of this paper is that the class of congruence lattices of semilattices satisfies no nontrivial lattice identities. It is also shown that the class of subalgebra lattices of semilattices satisfies no nontrivial lattice identities. As a consequence it is shown that if $\mathscr{V}$ is a semigroup variety all of whose congruence lattices satisfy some fixed nontrivial lattice identity, then all the members of $\mathscr{V}$ are groups with exponent dividing a fixed finite number.

Given a variety (equational class) $\mathscr{K}$ of algebras, among the interesting questions we can ask about the members of $\mathscr{K}$ is the following: does there exist a lattice identity δ such that for each algebra $A \in \mathscr{K}$, the congruence lattice $\Theta(A)$ satisfies δ? In the case that $\mathscr{K}$ has distributive congruences, many strong conclusions can be drawn about the algebras of $\mathscr{K}$ [1, 2, 7]. In the case that $\mathscr{K}$ has permutable congruences or modular congruences, there is reason to hope that some similar results may be obtainable [4, 8].

A standard method of proving that a class of lattices satisfies no nontrivial lattice identities is to show that all partition lattices (lattices of equivalence relations) are contained as sublattices. The lattices of congruences of semilattices, however, are known to be pseudo-complemented [9]. It follows that the partition lattice on a three-element set (the five-element two-dimensional lattice) is not isomorphic to a sublattice of the congruence lattice of a semilattice, and in fact is not a homomorphic image of a sublattice of the congruence lattice of a finite semilattice. Nonetheless we shall show in this paper that the congruence lattices of semilattices satisfy no nontrivial lattice identities. This solves Problem 6 of [10]. Using a theorem of T. Evans [6], we also show that if $\mathscr{V}$ is a variety of semigroups all of whose congruence lattices satisfy some fixed nontrivial lattice identity, then all the members of $\mathscr{V}$ are groups with exponent dividing a fixed finite number.

In § 1 we give definitions and a few basic facts about the congruences of semilattices. In § 2 we prove our main theorem, and in § 3 we apply it to obtain the corollary about varieties of semigroups.

1. A semilattice is a commutative idempotent semigroup. We may impose a partial ordering on a semilattice S by defining

$$x \leqq y \quad \text{if} \quad xy = x\,.$$

Under this ordering, any two elements $x, y \in S$ have a greatest lower bound, namely their product xy. S is called a meet semilattice. It may be that x and y have a least upper bound $w \in S$; if so, we define

$$x + y = w \,.$$

Thus $+$ is a partial operation on S, and $x + y$ is called the join of x and y. If S is finite, and if x and y have a common upper bound, then $x + y$ exists and

$$x + y = \Pi\{z \in S : z \geqq x \text{ and } z \geqq y\} \,.$$

The least element of a semilattice, if it exists, is denoted by 0; the greatest element, if it exists, by 1.

A dual ideal of a semilattice S is a set $D \subseteqq S$ satisfying (i) $d_1, d_2 \in D$ implies $d_1 d_2 \in D$, and (ii) $x \geqq d \in D$ implies $x \in D$. We will denote the principal dual ideal above x by $1/x$, i.e.,

$$1/x = \{z \in S : z \geqq x\} \,.$$

For reference we note that if $x + y$ is defined, then

$$1/x \cap 1/y = 1/x + y \,.$$

If S and T are semilattices, then $S \times T$ will denote the (external) direct product of S and T. We shall use round symbols ($\cap$, $\cup$) for set operations, and sharp symbols ($\wedge$, $\vee$) for lattice operations.

The following theorem is basic to the study of semilattice congruences.

THEOREM 1. [9] *Let $\underset{\sim}{2}$ denote the two-element semilattice. If S is any semilattice and D is a dual ideal of S, then the mapping $f: S \to \underset{\sim}{2}$ defined by*

$$f(x) = \begin{cases} 1 & \text{if} \quad x \in D \\ 0 & \text{if} \quad x \notin D \end{cases}$$

is a homomorphism. Thus $\underset{\sim}{2}$ is the only subdirectly irreducible semilattice, and the dual of $\Theta(S)$ is a point lattice ($\Theta(S)$ is a copoint lattice).

In the rest of this section we note some easily provable facts about the congruence lattice of a semilattice S.

(1) Suppose $\Theta(S)$ is atomic, and let A be the set of atoms of $\Theta(S)$. Let x^* denote the pseudo-complement of x. Then if $a \in A$, a^* is a coatom of $\Theta(S)$, and 0 is a unique irredundant meet of $\{a^* : a \in A\}$. Hence S is a unique subdirect product of $|A|$ copies of $\underset{\sim}{2}$, but of no fewer.

It is not hard to show that if θ covers 0 in $\Theta(S)$ then θ covers 0 in $\Pi(S)$, the partition lattice on S. From this and Theorem 1 it follows that

(2) $\Theta(S)$ is semimodular and if $\Theta(S)$ is finite and c is the number of coatoms of $\Theta(S)$ then

$$\dim \Theta(S) = c = |S| - 1 .$$

(3) $\Theta(S)$ is relatively pseudo-complemented [9].

A lattice L is called locally distributive if the quotient sublattices u_a/a is distributive for all $a \in L$, where u_a is the join of the elements covering a. In a compactly generated lattice, local distributivity is equivalent to the conjunction of semimodularity and relatively pseudo-complementation [3]. Hence

(4) $\Theta(S)$ is locally distributive.

The problem of characterizing all lattices isomorphic to congruence lattices of semilattices remains open. The above conditions are not sufficient, even in the finite case.

2. In this section we prove the main result of this paper.

THEOREM 2. *Let δ be a nontrivial lattice identity. Then there exists a finite semilattice $S(\delta)$ such that δ fails in the congruence lattice $\Theta(S(\delta))$.*

The theorem is an immediate consequence of Lemmas 1 and 4 to be proven below.

LEMMA 1. *Let S be a finite meet semilattice, and let $\mathscr{S}(S)$ be the lattice of (partial) join-subalgebras of S, with $0 \in S$ considered as a distinguished element. Then the congruence lattice $\Theta(S)$ is dually isomorphic to $\mathscr{S}(S)$.*

A partial join subalgebra of S is a subset containing 0 and closed under joins, whenever they exist.

Proof. The dual atoms of $\Theta(S)$ are the partitions $\psi_d = (1/d)(S - 1/d)$ for $0 \neq d \in S$. On the other hand, the atoms of $\mathscr{S}(S)$ are the subalgebras $\xi_d = \{0, d\}$ for $0 \neq d \in S$. We want to show that the mapping $\psi_d \rightarrow \xi_d$ induces a dual isomorphism of $\Theta(S)$ onto $\mathscr{S}(S)$. Since $\Theta(S)$ is a copoint lattice and $\mathscr{S}(S)$ is a point lattice, it is sufficient to show that their closure operations are duals under the mapping, i.e., that

$$\psi_c \geqq \psi_{d_1} \wedge \cdots \wedge \psi_{d_k} \quad \text{if and only if} \quad \xi_c \leqq \xi_{d_1} \vee \cdots \vee \xi_{d_k} .$$

This is equivalent to

$$\psi_c \geqq \psi_{d_1} \wedge \cdots \wedge \psi_{d_k} \quad \text{if and only if} \quad c \in \langle d_1, \cdots, d_k \rangle$$

where $\langle d_1, \cdots, d_k \rangle$ denotes the join subalgebra generated by $\{d_1, \cdots, d_k\}$. Notice that the equivalence classes of $\psi_{d_1} \wedge \cdots \wedge \psi_{d_k}$ are

$$\left(\bigcap_{i \in I} 1/d_i - \bigcup_{j \in I^c} 1/d_j \right)$$

for $I \subseteqq \{1, \cdots, k\}$. If $\psi_c \geqq \psi_{d_1} \wedge \cdots \wedge \psi_{d_k}$ then each of these classes is contained in either $1/c$ or $S - 1/c$. Considered the $\psi_{d_1} \wedge \cdots \wedge \psi_{d_k}$ – class which contains c. Then c is the least element of that class, and thus

$$c = \sum_{i \in I} d_i \quad \text{for some} \quad I \subseteqq \{1, \cdots, k\} .$$

Hence $c \in \langle d_1, \cdots, d_k \rangle$.

Conversely, if $c \in \langle d_1, \cdots, d_k \rangle$, then $c = \sum_{i \in I} d_i$ for some $I \subseteqq \{1, \cdots, k\}$. Thus the congruence $\bigwedge_{i \in I} \psi_{d_i}$ has one class equal to $1/c$ and the rest contained in $S - 1/c$. Hence

$$\psi_c \geqq \bigwedge_{i \in I} \psi_{d_i} \geqq \psi_{d_1} \wedge \cdots \wedge \psi_{d_k} .$$

This completes the proof of Lemma 1.

Suppose $\sigma \leqq \tau$ is a nontrivial lattice identity, i.e., $\sigma \leqq \tau$ does not hold in a free lattice. Then we construct a finite semilattice $S(\sigma)$ (depending only on σ) such that $\sigma \leqq \tau$ fails in $\mathscr{S}(S(\sigma))$. Combined with Lemma 1, this will prove Theorem 2.

Let $X = \{x, y, z, \cdots\}$ be a countable set, and let $FL(X)$ donote the free lattice on X. For each element $\sigma \in FL(X)$ we will define a finite semilattice $S(\sigma)$. First of all we write each $\sigma \in FL(X)$ in canonical form. Then we define

$$\begin{aligned} S(x) &= \underset{\sim}{2} \qquad \text{for } x \in X \\ S(\sigma_1 \vee \sigma_2) &= S(\sigma_1) \times S(\sigma_2) \\ S(\sigma_1 \wedge \sigma_2) &= S(\sigma_1) \times S(\sigma_2) - \Gamma \end{aligned}$$

where

$$\Gamma = 1/(1, 0) \cup 1/(0, 1) - \{(1, 1)\} .$$

Let us look more carefully at the construction. If $S(\sigma_1)$ and $S(\sigma_2)$ are lattices, then $S(\sigma_1) \times S(\sigma_2) - \Gamma$ is meet-closed and has a unit element; hence it is a lattice. It follows by induction that $S(\sigma)$ is a lattice for each $\sigma \in FL(X)$. We need to know how to compute joins in $S(\sigma)$. In $S(\sigma_1 \vee \sigma_2)$ joins are of course taken componentwise. In $S(\sigma_1 \wedge \sigma_2)$ we have

$$(*)\qquad (r_1, r_2) + (s_1, s_2) = \begin{cases} (r_1 + s_1, r_2 + s_2) & \text{if} \quad r_1 + s_1 \neq 1 \quad \text{and} \quad r_2 + s_2 \neq 1 \\ (1, 1) & \text{if} \quad r_1 + s_1 = 1 \quad \text{or} \quad r_2 + s_2 = 1 \,. \end{cases}$$

In any $S(\sigma)$ let us denote $(1, 1)$ by 1.

For each $\sigma \in FL(X)$ we now define a homomorphism φ_σ of $FL(X)$ into $\mathscr{S}(S(\sigma))$. We do this by associating with each $y \in X$ a join-subalgebra $\varphi_\sigma(y)$ of $S(\sigma)$, and extending this map to a homomorphism in the (unique) natural way. Once again we proceed inductively, with $\sigma \in FL(X)$ written in canonical form. For $y \in X$ we set

$$\varphi_x(y) = \begin{cases} S(x) & \text{if} \quad y = x \\ \{0\} & \text{if} \quad y \neq x \end{cases}$$

$$\varphi_{\sigma_1 \vee \sigma_2}(y) = \{(r_1, r_2) \colon r_1 \in \varphi_{\sigma_1}(y),\ r_2 \in \varphi_{\sigma_2}(y)\}$$

$$\varphi_{\sigma_1 \wedge \sigma_2}(y) = \{(r_1, r_2) \colon r_1 \in \varphi_{\sigma_1}(y) - \{1\},\ r_2 \in \varphi_{\sigma_2}(y) - \{1\}\} \cup \Lambda(\varphi_{\sigma_1}(y), \varphi_{\sigma_2}(y))$$

where

$$\Lambda(A, B) = \begin{cases} \varnothing & \text{if} \quad 1 \notin A \quad \text{and} \quad 1 \notin B \\ \{1\} & \text{if} \quad 1 \in A \quad \text{or} \quad 1 \in B \,. \end{cases}$$

Our computations will be based upon the following lemma.

LEMMA 2. *If* $\rho \in FL(X)$, *then*
(i) $\varphi_{\sigma_1 \vee \sigma_2}(\rho) = \varphi_{\sigma_1}(\rho) \times \varphi_{\sigma_2}(\rho)$
(ii) $\varphi_{\sigma_1 \wedge \sigma_2}(\rho) - \{1\} = \{(r, s) \in \varphi_{\sigma_1}(\rho) \times \varphi_{\sigma_2}(\rho) \colon r \neq 1$ *and* $s \neq 1\}$.

Proof. We prove (ii); the proof of (i) is similar but easier. We proceed by induction on the length of ρ. For $\rho = y \in X$ the lemma is immediate from the definitions. Now note that since $0 \in T$ for every $T \in \mathscr{S}(S(\sigma))$, we have

$$T_1 \vee T_2 = \{t_1 + t_2 \colon t_1 \in T_1,\ t_2 \in T_2\} \,.$$

Hence if $\rho = \rho_1 \vee \rho_2$, then by (*) we have

$$\begin{aligned} \varphi_{\sigma_1 \wedge \sigma_2}(\rho) - \{1\} &= \varphi_{\sigma_1 \wedge \sigma_2}(\rho_1) \vee \varphi_{\sigma_1 \wedge \sigma_2}(\rho_2) - \{1\} \\ &= \{(r_1, s_1) + (r_2, s_2) \colon (r_1, s_1) \in \varphi_{\sigma_1 \wedge \sigma_2}(\rho_1)\,, \\ &\qquad (r_2, s_2) \in \varphi_{\sigma_1 \wedge \sigma_1}(\rho_2),\ r_1 + r_2 \neq 1,\ s_1 + s_2 \neq 1\} \,. \end{aligned}$$

By the inductive hypothesis we have

$$(r_i, s_i) \in \varphi_{\sigma_1 \wedge \sigma_2}(\rho_i) - \{1\} = \varphi_{\sigma_1}(\rho_i) - \{1\} \times \varphi_{\sigma_2}(\rho_i) - \{1\}$$

for $i = 1, 2$ and hence

$$\varphi_{\sigma_1\wedge\sigma_2}(\rho) - \{1\} = \{(r, s) \in \varphi_{\sigma_1}(\rho) \times \varphi_{\sigma_2}(\rho): r \neq 1 \quad \text{and} \quad s \neq 1\} .$$

On the other hand, if $\rho = \rho_1 \wedge \rho_2$, then

$$\varphi_{\sigma_1\wedge\sigma_2}(\rho) - \{1\} = \varphi_{\sigma_1\wedge\sigma_2}(\rho_1) - \{1\} \cap \varphi_{\sigma_1\wedge\sigma_2}(\rho_2) - \{1\}$$

and the conclusion of the lemma follows.

LEMMA 3. *If* $\rho \in FL(X)$, *then* $1 \in \varphi_\rho(\rho)$.

Proof. As usual we proceed by induction on the length of ρ. If $\rho = y \in X$ the lemma follows from the definitions. If $\rho = \rho_1 \vee \rho_2$, then $1 \in \varphi_{\rho_i}(\rho_i)$ ($i = 1, 2$), and thus by Lemma 2 (i) we have

$$(1, 0) \in \varphi_\rho(\rho_1) \quad \text{and} \quad (0, 1) \in \varphi_\rho(\rho_2)$$

from which it follows that $1 \in \varphi_\rho(\rho_1) \vee \varphi_\rho(\rho_2) = \varphi_\rho(\rho)$.

If $\rho = \rho_1 \wedge \rho_2$, we can again assume $1 \in \varphi_{\rho_i}(\rho_i)$ for $i = 1, 2$. We need to show that $1 \in \varphi_{\rho_1\wedge\rho_2}(\rho_i)$. We prove a slightly stronger statement: if $1 \in \varphi_{\rho_1}(\sigma)$, then $1 \in \varphi_{\rho_1\wedge\rho_2}(\sigma)$. If $\sigma = y \in X$ this is immediate. Suppose $\sigma = \sigma_1 \vee \sigma_2$, then $1 \in \varphi_{\rho_1}(\sigma_1) \vee \varphi_{\rho_1}(\sigma_2)$ and hence $1 = t_1 + t_2$, where $t_i \in \varphi_{\rho_1}(\sigma_i)$. If $t_1 \neq 1$, $t_2 \neq 1$, then by Lemma 2 (ii) we have

$$1 = (t_1, 0) + (t_2, 0) \in \varphi_{\rho_1\wedge\rho_2}(\sigma_1) \vee \varphi_{\rho_1\wedge\rho_2}(\sigma_2) = \varphi_{\rho_1\wedge\rho_2}(\sigma) .$$

If $t_i = 1$ for some i then by induction $1 \in \varphi_{\rho_1}(\sigma_i)$ implies

$$1 \in \varphi_{\rho_1\wedge\rho_2}(\sigma_i) \subseteqq \varphi_{\rho_1\wedge\rho_2}(\sigma) .$$

Suppose $\sigma = \sigma_1 \wedge \sigma_2$. Then $1 \in \varphi_{\rho_1}(\sigma_1) \cap \varphi_{\rho_1}(\sigma_2)$. By induction $1 \in \varphi_{\rho_1\wedge\rho_2}(\sigma_i)$ for $i = 1, 2$ and we are done.

LEMMA 4. *If* $\sigma \nleqq \tau$ *in* $FL(X)$, *then* $1 \notin \varphi_\sigma(\tau)$.

Assume we have proven Lemma 4. Then Lemmas 3 and 4 combine to yield: $1 \in \varphi_\sigma(\tau)$ if and only if $\sigma \leqq \tau$ in $FL(X)$. Hence $\varphi_\sigma(\sigma) \subseteqq \varphi_\sigma(\tau)$ if and only if $\sigma \leqq \tau$ in $FL(X)$, and Theorem 2 follows.

Proof of Lemma 4. Suppose the lemma is false. Let σ be a word of minimum length such that $1 \in \varphi_\sigma(\tau')$ for some τ' such that $\sigma \nleqq \tau'$ in $FL(X)$. Let τ be of minimal length such that $\sigma \nleqq \tau$ and $1 \in \varphi_\sigma(\tau)$. We will show that these conditions lead to a contradiction. The cases $\sigma \in X$, $\sigma = \sigma_1 \vee \sigma_2$, $\sigma = \sigma_1 \wedge \sigma_2$, and $\tau \in X$ or $\tau = \tau_1 \wedge \tau_2$ are easy to handle. Let us assume, then, that $\sigma = \sigma_1 \wedge \sigma_2$ and $\tau = \tau_1 \vee \tau_2$. Then since $\sigma \nleqq \tau$ we have

$$\sigma \nleqq \tau_1 \quad \text{and} \quad \sigma \nleqq \tau_2 \quad \text{and} \quad \sigma_1 \nleqq \tau \quad \text{and} \quad \sigma_2 \nleqq \tau .$$

Since $1 \in \varphi_\sigma(\tau) = \varphi_\sigma(\tau_1) \vee \varphi_\sigma(\tau_2)$, there exist $t_i \in \varphi_\sigma(\tau_i)$ such that $t_1 + t_2 = 1$. If $t_i = 1$ for some i then by the minimal length of τ we have $\sigma \leqq \tau_i$, a contradiction. Thus $t_i \neq 1$ and by Lemma 2 (ii) we can write $t_i = (r_i, s_i)$ where $r_i \in \varphi_{\sigma_1}(\tau_i)$ and $s_i \in \varphi_{\tau_2}(\tau_i)$. Now either $r_1 + r_2 = 1$ in $S(\sigma_1)$, which means $1 \in \varphi_{\sigma_1}(\tau_1) \vee \varphi_{\sigma_1}(\tau_2) = \varphi_{\sigma_1}(\tau)$ and hence $\sigma_1 \leqq \tau$, or $s_1 + s_2 = 1$ and $\sigma_2 \leqq \tau$. Both these statements are contradictions.

Since the semilattices $S(\sigma)$ constructed above are in fact lattices, they are join semilattices. Thus, the above proof shows that any nontrivial lattice identity fails in the subalgebra lattice of some finite semilattice.

Now the congruence lattices of lattices satisfy every nontrivial lattice identity, while those of semilattices satisfy no identity. It is reasonable then to ask if there is some "natural" restricted class $\mathscr{K}$ of semilattices such that the congruence lattices of semilattices in $\mathscr{K}$ satisfy some lattice identity.

One such class is known [5]. A simple argument based on Theorem 1 shows that $\Theta(S)$ is nonmodular if and only if S contains a pair of noncomparable elements with a common upper bound. Hence if S is finite $\Theta(S)$ is either nonmodular, or else it is isomorphic to the Boolean algebra of subsets of some set.

Ot the other hand, the semilattices $S(\sigma)$ constructed in § 2 are in fact lattices; in particular, the join of every pair of elements is defined. It follows from Theorem 1 that $S(\sigma)$ can be imbedded as a join semilattice into a Boolean algebra $B(\sigma)$. Considering $B(\sigma)$ as a meet semilattice, we see that every nontrivial lattice identity fails in the (semilattice) congruence of some finite Boolean algebra.

3. We can now prove an interesting corollary about varieties of semigroups. Let R denote the two-element semigroup with multiplication law $xy = y$; L the two-element semigroup with multiplication law $xy = x$; and C the two-element semigroup with constant multiplication. The following theorem is due to T. Evans [6].

THEOREM 3. *The atoms of the lattice of varieties of semigroups are the varieties generated by R, L, C, $\underset{\sim}{2}$ (the variety of all semilattices), and the cyclic groups of prime order. If a nontrivial variety of semigroups does not contain R, L, C, or $\underset{\sim}{2}$, then it is a subvariety of $\mathscr{B}_n$, the variety of groups of exponent dividing n, for some finite n.*

Now if T is a semigroup in the variety generated by R, L, or C, then $\Theta(T)$ is just the partition lattice on T. Hence Theorems 2 and 3 combine to give the following corollary.

COROLLARY. *If $\mathscr{V}$ is a semigroup variety all of whose congruence* *lattices satisfy some fixed nontrivial lattice identity, then $\mathscr{V}$ is a subvariety of $\mathscr{B}_n$ for some finite n.*

REFERENCES

1. K. A. Baker, *Equational bases for finite algebras,* Notices Amer. Math. Soc., **19** (1972), A-44.

2. ———, *Primitive satisfaction and equational problems for lattices and other algebras,* to appear.

3. P. Crawley and R. P. Dilworth, *Algebraic Theory of Lattices,* Prentice-Hall, Englewood Cliffs, N. J., 1973.

4. Alan Day, *A characterization of modularity for congruence lattices of algebras,* Canad. Math. Bull., **12** (1969), 167-173.

5. R. A. Dean and R. H. Oehmke, *Idempotent semigroups with distributive right congruence lattices,* Pacific J. Math., **14** (1964), 1187-1209.

6. Trevor Evans, *The lattice of semigroup varieties,* Semigroup Forum **2** (1971), 1-43.

7. B. Jónsson, *Algebras whose congruence lattices are distributive,* Math. Scand., **21** (1967), 110-121.

8. A. Malcev, *On the general theory of algebraic systems,* Mat. Sb. (N. S.), **35** (77) (1954), 3-20.

9. D. Papert, *Congruence relations in semi-lattices,* London Math. Soc., **39** (1964), 723-729.

10. E. T. Schmidt, *Kongruenzrelationen Algebraischer Strukturen,* Veb Deutscher Verlag, Berlin, 1969.

Received July 10, 1972.

CALIFORNIA INSTITUTE OF TECHNOLOGY
AND
UNIVERSITY OF HAWAII

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

A NONASSOCIATIVE EXTENSION OF THE CLASS OF DISTRIBUTIVE LATTICES

E. FRIED AND G. GRÄTZER

Let $Z = \{0, 1, 2\}$ and define two binary operations $\wedge$ and $\vee$ on Z as follows: $0 \wedge 1 = 0, 0 \vee 1 = 1, 1 \wedge 2 = 1, 1 \vee 2 = 2, 2 \wedge 0 = 2, 2 \vee 0 = 2$, both operations are idempotent and commutative. This paper deals with the equational class $\boldsymbol{Z}$ generated by the algebra $\langle Z; \wedge, \vee\rangle$. The class $\boldsymbol{Z}$ contains the class of all distributive lattices and $\boldsymbol{Z}$ is a subclass of the class of weakly associative lattices (*trellis*, T-lattice) in the sense of E. Fried and H. Skala.

The purpose of this paper is to prove that $\boldsymbol{Z}$ shares the most important properties of the class of distributive lattices.

A *tournament* $\langle T; <\rangle$ is a set T with a binary relation $<$ such that for all $a, b \in T$ exactly one of $a = b, a < b$, and $b < a$ holds. Equivalently, a tournament is a directed graph without loops such that exactly one directed edge connects any two distinct points. Just as chains (linearly ordered sets) can be turned into lattices we can define meet and join on a tournament $\langle T; <\rangle$ by the rule:

$$\text{if } x < y, \text{ then } x = x \wedge y = y \wedge x \text{ and } y = x \vee y = y \vee x\,,$$
$$\text{and } x = x \wedge x = x \vee x \text{ for all } x.$$

Since for all $x, y \in T, x \neq y$, we have $x < y$ or $y < x$ the above rule defines $\wedge$ and $\vee$ on T.

Of course, the algebra $\langle T; \wedge, \vee\rangle$ we constructed is not a lattice: neither $\wedge$ nor $\vee$ is associative unless $\langle T; <\rangle$ is a chain, that is, $<$ is transitive. However, as it was observed in E. Fried [5], the two operations are idempotent, commutative; the absorption identities hold and also a weak form of the associative identities.

The smallest example of a nontransitive tournament is the three-element cycle $\langle\{0, 1, 2\}; <\rangle$ in which $0 < 1, 1 < 2$, and $2 < 0$. In the corresponding algebra Z neither $\wedge$ nor $\vee$ is associative.

Z plays the same role for tournaments as the two-element lattice does for distributive lattices. A tournament (algebra) $\langle T; \wedge, \vee\rangle$ is not a chain if and only if it contains Z as a subalgebra.

In this paper we investigate the equational class $\boldsymbol{Z}$ generated by the algebra Z. Observe that $C_2 = \langle\{0, 1\}; \wedge, \vee\rangle$ is a subalgebra of Z, in fact, it is a two-element chain. Therefore, $\boldsymbol{Z}$ contains as a subclass the class $\boldsymbol{D}$ of all distributive lattices. (Indeed, $\boldsymbol{D}$ is generated by C_2.)

The results of this paper can be summarized as follows: many of the most important properties of $\boldsymbol{D}$ generalize to $\boldsymbol{Z}$, and, in fact, $\boldsymbol{Z}$ is the only equational class (other than $\boldsymbol{D}$) generated by tournaments to which these results generalize.

In Part I, we discuss congruences in, and identities of $\boldsymbol{Z}$. Section 2 contains some preliminary results and some important concepts, including the proper form of distributivity for tournaments. In §3 the minimal congruence relation $\Theta(a, b)$ is described in $\boldsymbol{Z}$ and is applied to show that the Congruence Extension Property and the Amalgamation Property hold for $\boldsymbol{Z}$. In §4 it is shown that the result of §3 characterizes the class $\boldsymbol{Z}$. This is applied in §5 to find a finite set of identities (in fact, two) characterizing the class $\boldsymbol{Z}$[1]. Part II contains the structure theorems. In §6 we describe the structure of finite algebras in $\boldsymbol{Z}$: they are all of the form $D \times Z^k$, where D is a uniquely determined distributive lattice. Section 7 gives the structure of free algebras over $\boldsymbol{Z}$: the free algebra on n generators is of the form $F_D(n) \times Z^{k_n}$, where $F_D(n)$ is the free distributive lattice on n generators and $k_n = 3^{n-1} - 2^n + 1$. We prove in §8 that every algebra in $\boldsymbol{Z}$ can be embedded in an injective one. The injectives in $\boldsymbol{D}$ are known to be the complete Boolean lattices. The injectives in $\boldsymbol{Z}$ are the extensions of Z by complete Boolean lattices.

EXAMPLES. An "evaluation" of elements of a set A is a map φ of A into another set S, equipped with a binary relation $<$, meaning "better than". We say that b is better than $a (a, b \in A)$ if $a\varphi < b\varphi$. If we want to be able to compare any two elements of A, then we have to assume that $\langle S; < \rangle$ is a tournament.

Evaluating a sample $\langle a_1, \cdots, a_n \rangle$ of elements of A we get an "evaluation vector": $\langle a_1\varphi, \cdots, a_n\varphi \rangle$. The study of the equational class generated by $\langle S; \wedge, \vee \rangle$ is the investigation of the algebra of the evaluation vectors. Thus $\boldsymbol{Z}$ is the "algebra" of the evaluation vectors over Z.

Given a set X we can consider the set $P(X)$ of all partitions $\langle X_0, X_1 \rangle$ of X into two sets. If $\langle X_0, X_1 \rangle, \langle Y_0, Y_1 \rangle \in P(X)$ we can set $\langle X_0, X_1 \rangle \leqq \langle Y_0, Y_1 \rangle$ if and only if $X_0 \subseteqq Y_0$. This makes $P(X)$ into a distributive lattice. Any distributive lattice is a sublattice (up to isomorphism) of some $P(X)$.

Now consider the set $Z(X)$ of all partitions of X into three subsets $\langle X_0, X_1, X_2 \rangle$. For $\langle X_0, X_1, X_2 \rangle, \langle Y_0, Y_1, Y_2 \rangle \in Z(X)$ we declare $\langle X_0, X_1,$

[1] The results of this paper were announced in the Notices of the American Mathematical Society 18 (1971), 402 and 548. Independently, in 1971 K. Baker announced in a lecture a general result, namely that every equational class of finite type in which the algebras have distributive congruence lattices and which is generated by a finite algebra can be defined by a finite set of identities. Our result in §5 is a very special case of Baker's result. Of course, the general method of Baker yields more complicated identities for $\boldsymbol{Z}$.

$X_2\rangle \leqq \langle Y_0, Y_1, Y_2\rangle$ if and only if $X_0 \subseteqq Y_0 \cup Y_1$, $X_1 \subseteqq Y_1 \cup Y_2$, and $X_2 \subseteqq Y_2 \cup Y_0$. Again, $Z(X) \in \boldsymbol{Z}$, and any member of $\boldsymbol{Z}$ will be (up to isomorphism) a subalgebra of some $Z(X)$. Observe that $Z(X)$ contains $P(X)$ as a subalgebra under the correspondence $\langle X_0, X_1\rangle \to \langle X_1, X_0, \varnothing\rangle$.

PART I. Congruences and Identities.

2. Preliminary results. An algebra $\langle A; \wedge, \vee\rangle$ is called a *weakly associative lattice* (*WA*-lattice) if it satisfies the following set of identities

$$x \wedge x = x \quad\text{and}\quad x \vee x = x \qquad \text{(idempotency)}; \tag{1}$$

$$x \wedge y = y \wedge x \quad\text{and}\quad x \vee y = y \vee x \qquad \text{(commutativity)}; \tag{2}$$

$$x \wedge (x \vee y) = x \quad\text{and}\quad x \vee (x \wedge y) = x \qquad \text{(absorption identities)}; \tag{3}$$

$$\begin{aligned} &((x \wedge z) \vee (y \wedge z)) \vee z = z \quad\text{and}\\ &((x \vee z) \wedge (y \vee z)) \wedge z = z \qquad \text{(weak associativity)}. \end{aligned} \tag{4}$$

This axiom system was discovered independently by E. Fried [5] (he called these *T-lattices*) and H. M. Skala [16] (she called them *trellis*).

(1)–(4) are not independent. (3) implies (1), and (4) and (1) imply (3). Observe, that the first identity (and, similarly, the second identity) of (4) can be written in the form

$$((x \wedge z) \vee (y \wedge z)) \vee z = (x \wedge z) \vee ((y \wedge z) \vee z)$$

which justifies the name weak associativity.

It is easy to see that in a *WA*-lattice the polynomial $p(x, y, z) = ((x \wedge y) \vee (y \wedge z)) \vee (z \wedge x)$ satisfies the identities

$$x = p(x, x, y) = p(x, y, x) = p(y, x, x)$$

implying (B. Jónsson [13]) that

LEMMA 1. *The congruence lattice of a WA-lattice is distributive.*

If A and B are *WA*-lattices, Θ a congruence relation of A, Φ a congruence relation of B, then we can define a congruence relation $\Theta \times \Phi$ on $A \times B$: $\langle a, b\rangle \equiv \langle a_1, b_1\rangle(\Theta \times \Phi)$ if $a \equiv a_1(\Theta)$ and $b \equiv b_1(\Phi)$. Let $C(D)$ denote the congruence lattice of D. Lemma 1 is known to imply

COROLLARY. *Every congruence relation of $A \times B$ is of the form*

$\Theta \times \Phi$ with $\Theta \in C(A)$ and $\Phi \in C(B)$. Therefore, $C(A \times B) \cong C(A) \times C(B)$.

Combining Lemma 1 with another result of B. Jónsson [13] we get the crucial

LEMMA 2. *Let A be a finite WA-lattice and let K be the equational class generated by A. Then every algebra in K is isomorphic to a subalgebra of an algebra of the form $A_1^{I_1} \times \cdots \times A_n^{I_n}$, where $I, \cdots, I_n$ are arbitrary sets and $A_1, \cdots, A_n$ are homomorphic images of subalgebras of A.*

In a WA-lattice A we can define

$$x \leqq y \text{ if and only if } x = x \wedge y \text{ if and only if } y = x \vee y .$$

The equivalence of the second and third clauses follows from the absorption identities. Observe that $x \leqq x$, and $x \leqq y$ and $y \leqq x$ imply $x = y$. Also, $x \leqq x \vee y$, $y \leqq x \vee y$, and it follows from (4) that $x \leqq t$ and $y \leqq t$ imply $x \vee y \leqq t$; these can be summarized by stating that $x \vee y$ is the least upper bound of x and y. Dually, $x \wedge y$ is the greatest lower bound of x and y. These properties give an alternative definition of WA-lattices in terms of $\leqq$ (E. Fried [5] and M. H. Skala [16]).

We conclude from this immediately, that any tournament is a WA-lattice. Furthermore, since a homomorphic image of a tournament A is isomorphic to a subalgebra of A we conclude from Lemma 2:

LEMMA 3. *Let A be a finite tournament. Then the equational class K generated by A consists of subalgebras of direct powers of A. In particular, every subdirectly irreducible member of K is a subalgebra of A.*

Applying this to Z and to the equational class $\boldsymbol{Z}$ it generates we conclude that every member of $\boldsymbol{Z}$ is isomorphic to a subalgebra of some Z^I. The subdirectly irreducible algebras in $\boldsymbol{Z}$ are Z and C_2. Thus $\boldsymbol{Z}$ contains $\boldsymbol{D}$, in fact, $\boldsymbol{Z}$ covers $\boldsymbol{D}$.

Given an algebra A and $a, b \in A$ there is a smallest congruence relation Θ under which $a \equiv b(\Theta)$. This congruence relation is denoted by $\Theta(a, b)$; it is called a *principal* congruence relation. Principal congruences of distributive lattices are described in G. Grätzer and E. T. Schmidt [10] and G. Grätzer [8]:

LEMMA 4. *Let L be a distributive lattice, $a, b, c, d \in L$, $a \leqq b$, and $c \leqq d$. Then the following conditions are equivalent:*

(i) $c \equiv d(\Theta(a, b))$;

(ii) $c = (a \vee c) \wedge d$ *and* $d = (b \vee c) \wedge d$.

(iii) $a \wedge c = a \wedge d$ *and* $b \vee c = b \vee d$.

The most important result of this paper, namely Theorem 2, is patterned after Lemma 4.

Lemma 4 implies that any distributive lattice L has the property that $c \equiv d(\Theta(a, b))$ can be decided in the sublattice generated by a, b, c, and d. This property has an important consequence by A. Day [4]:

LEMMA 5. *Let* $\boldsymbol{K}$ *be an equational class of algebras with the property that for any* $A \in \boldsymbol{K}$ *and* $a, b, c, d \in A$, $c \equiv d(\Theta(a, b))$ *can be decided in the subalgebra generated by* a, b, c, *and* d. *Then* $\boldsymbol{K}$ *has the Congruence Extension Property, that is, if* $A, B \in \boldsymbol{K}$, A *a subalgebra of* B *and if* Θ *is a congruence relation on* A, *then there is a congruence relation* Φ *on* B *such that* Φ *restricted to* A *is* Θ.

Another property of distributive lattices we need to generalize is the uniqueness of relative complements.

Let $\boldsymbol{T}$ denote the equation class generated by all tournaments.

LEMMA 6. *The distributive law*

$$x \wedge (y \vee z) = ((x \wedge y) \vee (x \wedge z)) \wedge (y \vee z) \tag{5}$$

holds in $\boldsymbol{T}$.

Proof. Let A be a tournament, $x, y, z \in A$. If two of x, y, and z are equal, then (5) holds since it is true in lattices. If $\{x, y, z\}$ is a chain, again, (5) is trivial. So we can assume that $\{x, y, z\}$ is isomorphic to $\boldsymbol{Z}$. Since (5) is symmetric in y and z, we can assume that $y < z$. Therefore, $y < z < x < y$. In this case, $x \wedge (y \vee z) = x \wedge z = z$ and $((x \wedge y) \vee (x \wedge z)) \wedge (y \vee z) = (x \vee z) \wedge z = z$, so (5) holds. Thus (5) holds for all algebras generating $\boldsymbol{T}$, so it holds for $\boldsymbol{T}$.

LEMMA 7. *Let* A *be a WA-lattice satisfying* (5). *Then for* $a, b, c \in A$, $a \wedge b = a \wedge c$ *and* $a \vee b = a \vee c$ *imply that* $b = c$.

Proof.

$b = b \wedge (a \vee b)$	by (3)
$= b \wedge (a \vee c)$	since $a \vee b = a \vee c$
$= ((b \wedge a) \vee (b \wedge c)) \wedge (a \vee c)$	by (5)
$= ((c \wedge a) \vee (b \wedge c)) \wedge (a \vee b)$	since $b \wedge a = c \wedge a$ and $a \vee c = a \vee b$
$= c \wedge (a \vee b)$	by (5)
$= c \wedge (a \vee c)$	since $a \vee b = a \vee c$
$= c$	by (3) ,

which was to be proved.

It should be pointed out that, unlike in lattices, (5) is not self-dual. The independence of (5) and its dual is shown in [6].

In conclusion we mention that a list of identities describing $\boldsymbol{T}$ was given in [9].

3. Principal congruences. In this section we state and verify the analogue of Lemma 4 for $\boldsymbol{Z}$. To facilitate the discussion we introduce some notation. We define five polynomials in the variables x_1, x_2, x_3, x_4, and x_5:

(6) $\quad p_1 = x_2 \wedge x_5, p_2 = x_2 \vee x_4, p_3 = (x_3 \vee x_5) \vee x_4,$ and $p_4 = x_5$

(7) $$p = (((x_1 \vee p_1) \vee p_2) \wedge p_3) \wedge p_4 .$$

THEOREM 1. *Let $A \in \boldsymbol{Z}$, and let $a, b, c, d \in A$, $a \leqq b$, and $c \leqq d$. Then the following conditions are equivalent*:

(i) $c \equiv d(\Theta(a, b))$;

(ii) $c = p(a, a, b, c, d)$ *and* $d = p(b, a, b, c, d)$;

(iii) $a \wedge (c \wedge b) = a \wedge (d \wedge b)$ *and* $(a \vee c) \vee b = (a \vee d) \vee b$.

REMARK. If A is a lattice, then $p_1 \leqq x_2 \leqq p_2$; similarly, $p_3 \geqq x_5 = p_4$. Therefore, $p = (x_1 \vee (p_1 \vee p_2)) \wedge (p_3 \wedge p_4) = (x_1 \vee p_2) \wedge x_5 = (x_1 \vee (x_2 \wedge x_5)) \wedge x_5$, reducing the first half of (ii) to the first half of Lemma 4. (ii). The second half of (ii) can be handled similarly. As for (ii), $\wedge$ and $\vee$ are associative if A is a lattice, and so $a \wedge (c \wedge b) = b \wedge c$, $a \wedge (d \wedge b) = a \wedge d$, and so on, yielding Lemma 4. (iii). Observe the different placing of the parentheses in the two equations in (iii).

Proof. (i) *implies* (ii). We prove this implication in several steps.

(α) A is isomorphic to $C_2 = \{0, 1\}$. Since C_2 is a lattice a reference to Lemma 4 settles the matter. Or, equivalently, check the implication for $a = 0$, $b = 1$ and $c = 0$, $d = 1$, or $c = d = 0$, or $c = d = 1$, and for $a = b = 1$ and $c = d = 0$ or $c = d = 1$ (seven cases).

(β) A is (isomorphic to) $Z = \{0, 1, 2\}$. If $\{a, b, c, d\} \neq Z$, then we proceed as under (α). If $a = b$, then we must have $c = d$, thus we can assume that $a \neq b$. By symmetry, we can assume that $a = 0$, $b = 1$. Since Z is simple, $\langle c, d\rangle$ could be $\langle 2, 2\rangle$, $\langle 1, 2\rangle$, or $\langle 2, 0\rangle$ (all other pairs contradict that $c \leqq d$ or that $\{a, b, c, d\} = Z$). Therefore, it is sufficient to check the implication in Z^3 for $a = \langle 0, 0, 0\rangle$, $b = \langle 1, 1, 1\rangle$, $c = \langle 2, 1, 2\rangle$, and $d = \langle 2, 2, 0\rangle$. Compute:

$$\begin{aligned} p_1 &= a \wedge d = \langle 2, 2, 0\rangle,\ p_2 = a \vee c = \langle 0, 1, 0\rangle,\ p_3 = (b \vee d) \vee c \\ &= \langle 2, 2, 1\rangle \vee \langle 2, 1, 2\rangle = \langle 2, 2, 2\rangle,\ p_4 = \langle 2, 2, 0\rangle,\ p(a, a, b, c, d) \\ &= (((a \vee p_1) \vee p_2) \wedge p_3) \wedge p_4 \\ &= (((\langle 0, 0, 0\rangle \vee \langle 2, 2, 0\rangle) \vee \langle 0, 1, 0\rangle) \wedge \langle 2, 2, 2\rangle) \wedge \langle 2, 2, 0\rangle \\ &= ((\langle 0, 0, 0\rangle \vee \langle 0, 1, 0\rangle) \wedge \langle 2, 2, 2\rangle) \wedge \langle 2, 2, 0\rangle \\ &= (\langle 0, 1, 0\rangle \wedge \langle 2, 2, 2\rangle) \wedge \langle 2, 2, 0\rangle \\ &= \langle 2, 1, 2\rangle \wedge \langle 2, 2, 0\rangle = \langle 2, 1, 2\rangle = c\ , \end{aligned}$$

and similarly, $p(b, a, b, c, d) = d$.

(γ) Assume the implication to hold for the algebras $A_1, \cdots, A_n$, and let B be a subalgebra of $A_1 \times \cdots \times A_n$. Then the implication holds in B. Indeed, let $a, b, c, d \in B$, $a \leqq b$, $c \leqq d$ and let $c \equiv d(\Theta(a, b))$ in B. By a result of A. I. Malcev (see Theorem 10.3 of [7]) there is a sequence of elements $z_0 = c, z_1, \cdots, z_m = d$ of B, and unary algebraic functions $p_0, \cdots, p_{n-1}$ of B such that $\{p_i(a), p_i(b)\} = \{z_i, z_{i+1}\}$ for $i = 0, 1, \cdots, m-1$.

For an element u of B let $u^{(i)}$ denote the ith component of u, that is, $u = \langle u^{(i)}, \cdots, u^{(n)}\rangle$. A unary algebraic function is of the form $q(u_1, \cdots, x, \cdots, u_k)$, where $u_1, \cdots, u_k \in B$ and p is a polynomial. So we can define $p^{(i)}$ a unary algebraic function on A_i by $q(u_1^{(i)}, \cdots, x, \cdots, u_k^{(i)})$.

Using the sequence of elements of A_i: $z_0^{(i)}, z_1^{(i)}, \cdots, z_m^{(i)}$ of B_i, and the unary algebraic functions: $p_0^{(i)}, \cdots, p_{n-1}^{(i)}$, we conclude that

$$a^{(i)} \equiv b^{(i)}(\Theta(c^{(i)}, d^{(i)})) \text{ in } A_i\ .$$

Thus, by assumption,

$$\begin{aligned} c^{(i)} &= p(a^{(i)}, a^{(i)}, b^{(i)}, c^{(i)}, d^{(i)}) \quad \text{and} \\ d^{(i)} &= p(b^{(i)}, a^{(i)}, b^{(i)}, c^{(i)}, d^{(i)})\ , \end{aligned}$$

for $i = 1, \cdots, n$. Hence, $c = p(a, a, b, c, d)$ and $d = p(b, a, b, c, d)$, which was to be proved.

Now we are ready to prove the implication. Let $A \in \boldsymbol{Z}$, $a, b, c, d \in A$, $a \leqq b$, $c \leqq d$, and $c \equiv d(\Theta(a, b))$. Invoking Malcev's result used above we can assume that A is finitely generated. Since $\boldsymbol{Z}$ is generated by Z, an n-generated algebra can have no more than 3^{3^n} elements, hence it is finite. Thus A is finite. By Lemma 3, A can be embedded in some Z^k. By (α) and (β) the implication holds in Z, hence by (γ) it holds in A, completing the proof.

(ii) *implies* (iii). This implication takes the form of a universal Horn sentence (see, for instance, [7], §46), therefore, it holds in $\boldsymbol{Z}$ if and only if it holds in Z. In Z, if $a = b$, then the assumption implies

$c = d$, hence the conclusion obviously holds. If $a \neq b$ $(a < b)$, then the functions $a \wedge (x \wedge b)$ and $(a \vee x) \vee b$ are constants $(a \wedge (x \wedge b) = a$ and $(a \vee x) \vee b = b)$, so the conclusion is obvious.

(iii) *implies* (i). Let $\Theta = \Theta(a, b)$ and let $[x]\Theta$ denote the congruence class of A containing x. Then

$$[a \wedge (c \wedge b)]\Theta = [a \wedge (d \wedge b)]\Theta ,$$

and so, using $[a]\Theta = [b]\Theta$, and $x \wedge (y \wedge x) = y \wedge x$, we obtain, in turn:

$$[a]\Theta \wedge ([c]\Theta \wedge [b]\Theta) = [a]\Theta \wedge ([d]\Theta \wedge [b]\Theta)$$

$$[a]\Theta \wedge [c]\Theta = [a]\Theta \wedge [d]\Theta . \tag{8}$$

Similarly,

$$[a]\Theta \vee [c]\Theta = [a]\Theta \vee [d]\Theta . \tag{9}$$

Applying Lemma 7 to A/Θ, (8) and (9) imply that $[c]\Theta = [d]\Theta$, that is, $c \equiv d(\Theta(a, b))$, completing the proof of Theorem 1.

Observe, that Theorem 1 implies that the assumptions of Lemma 5 are satisfied in $\boldsymbol{Z}$. Thus,

COROLLARY 1. *$\boldsymbol{Z}$ has the Congruence Extension Property.*

A class K of algebras is said to have the Amalgamation Property if for any $A, B_1, B_2 \in K$, and embeddings $f_i: A \to B_i$, $i = 1, 2$, there is a $C \in K$ and embeddings $g_i: B_i \to C$, $i = 1, 2$, such that $f_1 g_1 = f_2 g_2$. For a general discussion of the Amalgamation Property see B. Jónsson [12].

COROLLARY 2. *$\boldsymbol{Z}$ has the Amalgamation Property.*

Proof. By Theorem 13.16 of [8] it is sufficient to prove that for given $A, B_1, B_2 \in \boldsymbol{Z}$, embeddings $f_i: A \to B_i$, $i = 1, 2$, and $a, b \in B_1$, $a \neq b$, there exist homomorphisms $g_i: B_i \to \boldsymbol{Z}$, $i = 1, 2$, such that $f_1 g_1 = f_2 g_2$ and $a g_1 \neq b g_1$. By Lemma 3, there is a homomorphism $g_1: B_1 \to Z$ satisfying $a g_1 = b g_1$. Let Θ be the congruence relation of A induced by g_1. By Corollary 1, there exists a congruence relation $\bar{\Theta}$ on B_2 satisfying $\bar{\Theta}_A = \Theta$. Let g_2' be the natural homomorphism of B_2 onto $B_2/\bar{\Theta}$. By Lemma 3 again, there is a homomorphism $g_2'': B_2/\bar{\Theta} \to Z$. We define $g_2 = g_2' g_2''$. Obviously, $f_1 g_1 = f_2 g_2$, concluding the proof.

In closing this section, we mention that the polynomial p, which plays a central role in Theorem 1, was found using free algebras. A free algebra was used also to discover the identity (5) in order to get Lemma 7.

Alternate forms of p such as $p = (((x_1 \vee q_1) \wedge q_2) \vee q_3) \wedge q_4$ or any of the other possibilities can also be found using the free algebra technique.

4. The characterization theorem. We started out in our research trying to find a finite set of identities characterizing $\boldsymbol{Z}$. Since we believed that the equivalence of Lemma 4. (i) and (ii) is characteristic of $\boldsymbol{D}$ we wanted to find the analogous result for $\boldsymbol{Z}$ hoping that it would characterize $\boldsymbol{Z}$. The next step would then be to find a set of identities based on which the analogous result for $\boldsymbol{Z}$ can be proved. As we shall see in §5, this runs into some problems. The situation was saved by Theorem 1. (iii) and by the fact that Theorem 1. (iii) can also be used to characterize $\boldsymbol{Z}$. Since this is the result needed in §5 we omit the original theorem and prove only the latter one.

THEOREM 2. *Let $\boldsymbol{K}$ be an equational class of WA-lattices in which for any $A \in \boldsymbol{K}$, $a, b, c, d \in A$, $a \leqq b, c \leqq d$, and $c \equiv d(\Theta(a, b))$ imply that $a \wedge (c \wedge b) = a \wedge (d \wedge b)$ and $(a \vee c) \vee b = (a \vee d) \vee b$. Then $\boldsymbol{K} \subseteqq \boldsymbol{Z}$.*

Let A be a subdirectly irreducible algebra in $\boldsymbol{K}$. We shall prove that $A \cong C_2$ or $A \cong Z$. This obviously implies that $\boldsymbol{K} \subseteqq \boldsymbol{Z}$.

If $|A| = 2$, then $A \cong C_2$ since A is a WA-lattice. Thus we can assume that $|A| > 2$.

Since A is a subdirectly irreducible algebra with more than two elements, A has a congruence relation $\Phi \neq \omega$ with the property that $\Phi \leqq \Theta$ for any congruence relation Θ of A with $\Theta \neq \omega$. Since $\Phi \neq \omega$ there is a congruence class G of Φ of more than one element.

We claim that there are elements $a, b, c \in A$ such that $a, b \in G$, $a < b$, and $c < a$ or $b < c$. To prove this take $x, y \in G$, $x \neq y$. Obviously, $x \wedge y \in G$ and $x \neq x \wedge y$ or $y \neq x \wedge y$. Set $a' = x \wedge y$ and $b' = x$ or $b' = y$ so that $a' \neq b'$. If a' and b' do not satisfy the requirements with some $c' \in A$, then for all $d \in A$ we have $a' \wedge d = a'$ and $b' \vee d = b'$, that is, $a' < d < b'$ for any $d \in A, d \neq a', d \neq b'$. In this case set $c = a', a = d$, and $b = b'$.

So we can assume that we have

$$c < a < b, a, b \in G, \tag{10}$$

since the other case, $a < b < c$ can be proceeded with similarly (dually).

By (10) $c \neq a$, so $\Theta(c, a) \neq \omega$. By the definition of Φ, we have $\Phi \leqq \Theta(c, a)$, and by the definition of G and (10) we must have

$$a \equiv b(\Theta(c, a)) .$$

We apply the hypothesis of Theorem 2 to this congruence. We obtain:

$$(c \vee a) \vee a = (c \vee b) \vee a ,$$

that is, $a = (c \vee b) \vee a$, in other words,

$$c < b \leqq a \,. \tag{11}$$

We claim that

$$c = c \vee b \,. \tag{12}$$

Assume, to the contrary, that $c < c \vee b$. Then $\Theta(c, c \vee b) \neq \omega$, and so as above we get

$$a \equiv b(\Theta(c, c \vee b)) \,.$$

Applying the hypothesis of Theorem 2 to this congruence we obtain:

$$c \wedge (a \wedge (c \vee b)) = c \wedge (b \wedge (c \vee b)) \,. \tag{13}$$

By (11), $a \wedge (c \vee b) = a$ and, by (10), $c \wedge a = c$ so (13) yields $c = c \wedge b$, or, equivalently,

$$c \leqq b \,. \tag{14}$$

On the other hand, by (11), $c \vee b \leqq a$; combining this with (14) we obtain $b \leqq a$, contradicting (10). This verifies (12).

(10) and (12) jointly mean that $\{a, b, c\}$ is a subalgebra of A and $\{a, b, c\}$ is isomorphic to Z.

We claim that $A = \{a, b, c\}$.

Assume to the contrary that there is an element $d \in A$ such that $d \notin \{a, b, c\}$. We claim that d can be chosen to be comparable to one of a, b, and c. Indeed, if there is no such d then for an arbitrary $e \in A$, $a \wedge e = a$, since $a \wedge e < a$ implies that $a \wedge e = c$ and so $c < e$. Similarly, $a \vee e = a$, implying that $a = e$, a contradiction. Thus, by reason of symmetry and duality we can assume that there is an element $d \in A$ satisfying

$$d \notin \{a, b, c\} \quad \text{and} \quad d < a \,. \tag{15}$$

Since $a, b \in G$ and $b < c < a$ we conclude that $c \in G$. Thus $d \neq a$ implies the congruence

$$b \equiv c(\Theta(d, a)) \,.$$

Therefore,

$$d \wedge (b \wedge a) = d \wedge (c \wedge a) \,. \tag{16}$$

But by (10) $b \wedge a = a$ and $c \wedge a = c$; by (15), $d \wedge a = d$, hence (16) yields: $d = d \wedge c$. Since $c \neq d$ this means that $d < c$. So we get the congruence

$$a \equiv b(\Theta(d, c)) \,,$$

which implies that

$$(d \vee a) \vee c = (d \vee b) \vee c . \tag{17}$$

But $d \vee a = a, a \vee c = c$, hence (17) gives

$$a = (d \vee b) \vee c . \tag{18}$$

Observe that $d \leqq c$ and $b \leqq c$. Therefore, $d \vee b \leqq c$, and so $(d \vee b) \vee c = c$, contradicting (18) and $a \neq c$.

This contradiction shows that $A = \{a, b, c\}$, that is, $A \cong Z$, which completes the proof of Theorem 2.

5. Identities for *Z*. We want to find a finite set Σ of identities characterizing Z. This set Σ should express that Z is a class of WA-lattices in which minimal congruences can be described by Theorem 1. It is easy to find identities which imply that the relation given by Theorem 1. (ii) is reflexive, symmetric, and has the Substitution Property for $\wedge$ and $\vee$. However, transitivity takes the form "if Theorem 1. (ii) holds for c, d, and for c_1, d_1, and $d = c_1$, then it holds for c, d_1" which we could not turn into an identity.

The trick is to find identities that prove that $\Theta(a, b)$ is in some sense the transitive extension of the relation given by Theorem 1. (ii). Then to show that this implies that Theorem 1. (iii) can be used to describe $\Theta(a, b)$.

We need some notation. We shall use $p, p_1, \cdots, p_4$ of (6) and (7) without references. Two 4-ary polynomials derived from p will be used often:

$$q_1 = p(x_1, x_1, x_1 \vee x_2, x_3, x_3 \vee x_4) \tag{19}$$

$$q_2 = p(x_1 \vee x_2, x_1, x_1 \vee x_2, x_3, x_3 \vee x_4) . \tag{20}$$

Finally, for the polynomials t_1, t_2, t_3, and t_4, let $R(t_1, t_2, t_3, t_4)$ denote the identities

$$t_3 = q_1(t_1, t_2, t_3, t_4) \quad \text{and} \quad t_3 \vee t_4 = q_2(t_1, t_2, t_3, t_3 \vee t_4) . \tag{21}$$

Σ consists of three sets of identities. Σ_1 is a set of identities for WA-lattices (for instance, (1)–(4)) and one more identity

$$((x \vee y) \vee (x \wedge z)) \vee (x \vee z) = (x \vee (y \vee z)) \vee z . \tag{22}$$

Σ_2 is the following eight identities:

$$R(x_1, x_2, q_i \wedge x_5, q_j \wedge x_5), \{i, j\} = \{1, 2\} \tag{23}$$

$$R(x_1, x_2, q_i \vee x_5, q_j \vee x_5), \{i, j\} = \{1, 2\} . \tag{24}$$

Σ_3 consists of two identities

$$x_1 \wedge (q_1 \wedge (x_1 \vee x_2)) = x_1 \wedge (q_2 \wedge (x_1 \vee x_2)) \tag{25}$$

$$(x_1 \vee q_1) \vee (x_1 \vee x_2) = (x_1 \vee q_2) \vee (x_1 \vee x_2) \,. \tag{26}$$

THEOREM 3. *$\Sigma = \Sigma_1 \cup \Sigma_2 \cup \Sigma_3$ defines the class $\boldsymbol{Z}$.*

Proof. First we have to see that Σ is satisfied in $\boldsymbol{Z}$. Σ_1 is obviously satisfied excepting (22).

It is sufficient to verify (22) in Z. Let $a, b, c \in Z$. If $|\{a, b, c\}| \leqq 2$, then they form a sublattice, in which (22) becomes $(a \vee b) \vee (a \vee c) = a \vee (b \vee c)$, a triviality. Thus we can assume that $Z = \{a, b, c\}$. If $a = 0, b = 1, c = 2$, then $((0 \vee 1) \vee (0 \wedge 2)) \vee (0 \vee 2) = (1 \vee 2) \vee 0 = 2 \vee 0 = 0$ and $((2 \vee 1) \vee 0) \vee 2 = (2 \vee 0) \vee 2 = 0 \vee 2 = 0$; if $a = 1, b = 0, c = 2$, then $((1 \vee 0) \vee (1 \wedge 2)) \vee (1 \vee 2) = ((1 \vee 1) \vee 1) \vee 2 = 1 \vee 2 = 2$ and $((2 \vee 0) \vee 1) \vee 2 = (0 \vee 1) \vee 2 = 1 \vee 2 = 2$. All the other substitutions agree with one of these two (up to automorphism) showing (22) in Z. If $x_1 \equiv x_1 \vee x_2$, then by (19), and (20), $q_1 \equiv q_2$ and so $q_1 \wedge x_5 \equiv q_2 \wedge x_5$. In other words,

$$q_1 \wedge x_5 \equiv q_2 \wedge x_5(\Theta(x_1, x_1 \vee x_2)) \,,$$

or

$$q_1 \wedge x_5 \equiv (q_1 \wedge x_5) \vee (q_2 \wedge x_5)(\Theta(x_1, x_1 \vee x_2)) \,.$$

Applying Theorem 1. (ii) to this congruence we obtain

$$q_1 \wedge x_5 = q_1(x_1, x_1 \vee x_2, q_1 \wedge x_5, (q_1 \wedge x_5) \vee (q_2 \wedge x_5))$$
$$(q_1 \wedge x_5) \vee (q_2 \wedge x_5) = q_2(x_1, x_1 \vee x_2, q_1 \wedge x_5, (q_1 \wedge x_5) \vee (q_2 \wedge x_5)) \,.$$

By (21), these two are written in the form $R(x_1, x_1 \vee x_2, q_1 \wedge x_5, q_2 \wedge x_5)$. The other six identities under (23) and (24) are similarly proved. Finally, since $q_1 \equiv q_2(\Theta(x_1, x_1 \vee x_2))$, an application of Theorem 1. (iii) proves (25) and (26).

Now let K be the class of all algebras satisfying Σ. By what we have proved above, $\boldsymbol{Z} \subseteqq K$.

Let $A \in K$, $a, b \in A$, and $a \leqq b$. We define a binary relation Φ on A: $c \equiv d(\Phi)$ if and only if there exists a sequence $c = r_0, r_1, \cdots, r_n = d$ of elements of A such that, for all $i = 0, \cdots, n-1$, r_i and r_{i+1} are comparable and $R(a, b, r_i \wedge r_{i+1}, r_i \vee r_{i+1})$.

We claim that Φ is a congruence relation, in fact, $\Phi = \Theta(a, b)$.

Φ is obviously symmetric and transitive. Next we show that Φ is reflexive, in other words, for all $c \in A$, $R(a, b, c, c)$. By (19)–(21), this means that

$$p(a, a, b, c, c) = p(b, a, b, c, c) = c\ . \tag{27}$$

Using (6) we compute: $p_1 = a \wedge c$, $p_2 = a \vee c$, $p_3 = (b \vee c) \vee c = b \vee c$ (by (3)), $p_4 = c$, and so

$$\begin{aligned} p(a, a, b, c, c) &= (((a \vee (a \wedge c)) \vee (a \vee c)) \wedge (b \vee c)) \wedge c && \text{by (3)} \\ &= ((a \vee c) \wedge (b \vee c)) \wedge c = c && \text{by (4)}\ . \end{aligned}$$

For the second half of (27) compute: $p_1 = a \wedge c$, $p_2 = a \vee c$, $p_3 = (b \vee c) \vee c = b \vee c$, $p_4 = c$ and so

$$\begin{aligned} &p(b, a, b, c, c) \\ &= (((b \vee (a \wedge c)) \vee (a \vee c)) \wedge (b \vee c)) \wedge c && \text{since } a \leqq b \\ &= ((((a \vee b) \vee (a \wedge c)) \vee (a \vee c)) \wedge (b \vee c)) \wedge c \\ & && \text{apply (22) with } x = a,\ y = b, \text{ and } z = c \\ &= (((a \vee (b \vee c)) \vee c) \wedge (b \vee c)) \wedge c \\ & && \text{use the second half of (4) with } x = \\ & && a \vee (b \vee c),\ y = b, \text{ and } z = c \\ &= c\ . \end{aligned}$$

To show the Substitution Property for $\wedge$, let $c \equiv d(\Phi)$ with the sequence $r_0, \cdots, r_n$ and let $e \in A$. Consider the sequence $e \wedge c = e \wedge r_0$, $(e \wedge r_0) \vee (e \wedge r_1)$, $e \wedge r_1$, $((e \wedge r_1) \vee (e \wedge r_2)$, $e \wedge r_2, \cdots, e \wedge r_n = e \wedge d$. For any given i, $0 \leqq i < n$, either $r_i \leqq r_{i+1}$ or $r_{i+1} \leqq r_i$. Let us assume that $r_i \leqq r_{i+1}$ (if $r_{i+1} \leqq r_i$ we proceed similarly). By the definition of Φ, we have $R(a, b, r_i, r_{i+1})$. By the definition of R, this means that

$$r_i = q_1(a, b, r_i, r_{i+1})$$

and

$$r_{i+1} = q_2(a, b, r_i, r_{i+1})\ .$$

By (23),

$$R(a, b, q_1 \wedge e, q_2 \wedge e) \quad \text{and} \quad R(a, b, q_2 \wedge e, q_1 \wedge e)\ ,$$

that is,

$$R(a, b, r_i \wedge e, r_{i+1} \wedge e) \quad \text{and} \quad R(a, b, r_{i+1} \wedge e, r_i \wedge e)\ .$$

Therefore, by the definition of R:

$$R(a, b, r_i \wedge e, (r_i \wedge e) \vee (r_{i+1} \wedge e))$$

and

$$R(a, b, r_{i+1} \wedge e, (r_i \wedge e) \vee (r_{i+1} \wedge e))\ ,$$

showing $c \wedge e \equiv d \wedge e(\Phi)$.

Using (24) rather than (23) we prove that $c \vee e \equiv d \vee e(\Phi)$.

Thus Φ is a congruence relation.

Observe that $p(t, a, b, a, b) = (t \vee a) \wedge b$. Thus $p(a, a, b, a, b) = a$ and $p(b, a, b, a, b) = b$, proving $a \equiv b(\Phi)$.

Finally, if $a \equiv b(\Theta)$ for any congruence relation Θ, then $\Phi \leqq \Theta$, thus $\Phi = \Theta(a, b)$.

Now let $c \equiv d(\Theta(a, b))$, $c = r_0, \cdots, r_n = d$ as given in the definition of Φ. For a given i, then $r_i = q_1(a, b, r_i \wedge r_{i+1}, r_i \vee r_{i+1})$ and $r_{i+1} = q_2(a, b, r_i \wedge r_{i+1}, r_i \vee r_{i+1})$. Substituting these into (25) and (26) we obtain the crucial equations:

$$a \wedge (r_i \wedge b) = a \wedge (r_{i+1} \wedge b)$$
$$(a \vee r_i) \vee b = (a \vee r_{i+1}) \vee b$$

for all $i = 0, \cdots, n - 1$. Thus

$$a \wedge (c \wedge b) = a \wedge (d \wedge b)$$

and

$$(a \vee c) \vee b = (a \vee d) \vee b \, .$$

In other words, we have shown that $c \equiv d(\Theta(a, b))$ implies the two previous equations, which is the hypothesis of Theorem 2.

Therefore, by Theorem 2, $\boldsymbol{K} \subseteqq \boldsymbol{Z}$. Combining this with $\boldsymbol{Z} \subseteqq \boldsymbol{K}$ we conclude that $\boldsymbol{K} = \boldsymbol{Z}$, completing the proof of Theorem 3.

It should be noted that it is much easier to prove that $\boldsymbol{Z}$ can be characterized by a finite set of identities. The proof given above actually exhibits one such set.

No more than five variables were used in the identities in Σ, hence,

COROLLARY 1. *An algebra $\langle A; \wedge, \vee \rangle$ belongs to $\boldsymbol{Z}$ if and only if every subalgebra of $\langle A; \wedge, \vee \rangle$ generated by five elements belongs to $\boldsymbol{Z}$.*

It is easily seen, that in the Corollary, "five" cannot be replaced by "three". We do not know whether "four" would do.

Since two identities of an idempotent class can always be substituted by one, the finite set Σ of Theorem 3 can be reduced to three identities. R. Padmanabhan [15] has shown that the three identities can be replaced by two.

COROLLARY 2. *There exist two identities characterizing $\boldsymbol{Z}$.*

PART II. Structure Theorems.

6. Finite algebras. The main result of this section is the following.

THEOREM 4. *Every finite algebra A of $\mathbf{Z}$ has a representation of the form*

$$A \cong D \times Z^k ,$$

where D is a (finite) distributive lattice and k a nonnegative integer. In this representation k is unique and D is unique up to isomorphism. In fact, D is a maximal homomorphic image of A in $\mathbf{D}$.

This result is based on three lemmas.

LEMMA 8. *Let the algebra A be a subdirect product of the algebras $A_1, \cdots, A_n$. Let us assume that there exists a family $p_\lambda, \lambda \in \Lambda$ of polynomials satisfying the following two conditions:*

(i) *$p_\lambda(x, y, y, \cdots, y) = x$ holds in all A_i;*

(ii) *for $a, b \in A_i$ there exist $a_1, a_2, \cdots \in A_i$ and $\lambda \in \Lambda$ such that $p_\lambda(a, a_1, a_2, \cdots) = b$.*

Let, further, A be a subdirect product of $A_1, \cdots, A_n$ with the property that for each $i, 1 \leqq i \leqq n$, there is an element $\langle c_1, \cdots, c_n \rangle \in A$ such that $\langle c_1, \cdots, c_{i-1}, a, c_{i+1}, \cdots, c_n \rangle \in A$ for all $a \in A_i$. Then A is the direct product of $A_1, \cdots, A_n$.

Proof. For $n = 1$ the statement is obvious. Let us assume that it has been proved for all $k < n$. Let A and $A_1, \cdots, A_n$ be given as in the lemma. Let B be the algebra we get from A by omitting the first component of each element of A. Obviously, B is a subdirect product of $A_2, \cdots, A_n$ and this subdirect product satisfies all the hypotheses of Lemma 8 (the element of B chosen for $i, 2 \leqq i \leqq n$ is the element of A chosen for i with its first component omitted). Therefore, by induction hypothesis, $B = A_2 \times \cdots \times A_n$.

It is also clear that A is a subdirect product of A_1 and B, and (using the hypothesis for $A, A_1, \cdots, A_n$ and $i = 1$) there is an element $d \in B$ such that $\langle c, d \rangle \in A$ for all $c \in A_1$. Now take an arbitrary $\langle a, b \rangle \in A_1 \times B$. Since A is a subdirect product of A_1 and B, there exist $e \in A_1$ such that $\langle e, b \rangle \in A$. By (ii), there exist $\lambda \in \Lambda$ and $a_1, a_2, \cdots \in A$, such that $p_\lambda(e, a_1, a_2, \cdots) = a$. Thus, $\langle e, b \rangle, \langle a_1, d \rangle, \langle a_2, d \rangle, \cdots \in A$ and so (using (ii)):

$$\begin{aligned} p_\lambda(\langle e, b \rangle, \langle a_1, d \rangle, \langle a_2, d \rangle, \cdots) &= \langle p_\lambda(e, a_1, a_2, \cdots), p_\lambda(b, d, d, \cdots) \rangle \\ &= \langle a, b \rangle \end{aligned}$$

is also in A, proving $A = A_1 \times B$. Thus,

$$A = A_1 \times A_2 \times \cdots \times A_n ,$$

completing the proof of the lemma.

LEMMA 9. *Let us assume that for the algebras $A_1, \cdots, A_n$ the polynomials $p_\lambda, \lambda \in \Lambda$ exist satisfying* (i) *and* (ii) *of Lemma* 8. *In addition, let us assume that for each $a, b, c \in A_i$, $b \neq c$ there is a polynomial g satisfying $g(a, b, c) = a$ for which $g(x, y, y) = y$ holds in $A_1, \cdots, A_n$. Then any subdirect product of A of $A_1, \cdots, A_n$ is isomorphic to a direct product of some of the $A_1, \cdots, A_n$.*

Proof. Again, we proceed by induction and the case $n = 1$ is obvious. For $1 \leqq i \leqq n$, consider the homomorphism $\varphi_i: A \rightarrow A^{(i)}$ which is the map omitting the ith component. If, for some i, φ_i is an isomorphism, then A is isomorphic to a subdirect product of $A_1, \cdots, A_{i-1}, A_{i+1}, \cdots, A_n$, and by the induction hypothesis, the conclusions of Lemma 9 holds for A. So we can assume that no φ_i is an isomorphism.

Now we show that $A, A_1, \cdots, A_n$ satisfy the conditions of Lemma 8. We have assumed the existence of the $p_\lambda, \lambda \in \Lambda$.

Choose an $i, 1 \leqq i \leqq n$. We want to prove that there exists a $\langle c_1, \cdots, c_n \rangle \in A$ such that $\langle c_1, \cdots, c_{i-1}, a, c_{i+1}, \cdots, c_n \rangle \in A$ for all $a \in A_i$.

To simplify the notation let $i = 1$. Since φ_1 is not an isomorphism there are elements $\bar{c}, \bar{d} \in A$ such that $\bar{c} \neq \bar{d}$ and $\bar{c}\varphi_1 = \bar{d}\varphi_1$. In other words, $\langle c, c_2, \cdots, c_n \rangle, \langle d, c_2, \cdots, c_n \rangle \in A$ for some $c_2 \in A_2, \cdots, c_n \in A_n$ and $c, d \in A_1, c \neq d$.

For an arbitrary $a \in A_1$, there are $a_2 \in A_2, \cdots, a_n \in A_n$ such that $\langle a, a_2, \cdots, a_n \rangle \in A$, since A is a subdirect product. Choose a polynomial g satisfying $g(a, c, d) = a$ (and, of course, $g(x, y, y) = y$). Then $g(\langle a, a_2, \cdots, a_n \rangle, \langle c, c_2, \cdots, c_n \rangle, \langle d, c_2, \cdots, c_n \rangle) = \langle a, c_2, \cdots, c_n \rangle$ is in A, verifying the condition. Thus, by Lemma 8, $A = A_1 \times \cdots \times A_n$.

LEMMA 10. *Any finite subdirect power of Z is isomorphic to some direct power of Z.*

Proof. We shall verify that the hypotheses of Lemmas 8 and 9 are satisfied in Z. Let $\Lambda = \{1, 2\}$, $p_1 = ((x \vee y) \vee z) \wedge x$ and $p_2 = ((x \wedge y) \wedge z) \vee x$. It is obvious that (i) of Lemma 8 holds. Let, say, $a = 0$. Then $p_1(0, 0, 0) = 0$, $p_2(0, 2, 1) = 1$, and $p_1(0, 1, 2) = 2$, verifying (ii) of Lemma 8.

We also select $g_1 = (x \vee y) \wedge z$ and $g_2 = (x \wedge z) \vee y$. Obviously, $g_i(x, y, y) = y, i = 1, 2$. If $b \neq c$, then $b < c$ or $c < b$. In the first case let $b = 0$ and $c = 1$; then $g_1(a, b, c) = a$ for $a = 0, 1$ and $g_2(a, b, c) =$

a for $a = 2$. In the second case, $g_1(a, b, c) = a$ for $a = 2$ and $g_2(a, b, c) = a$ for $a = 0$ and 1. This completes the proof of Lemma 10.

Now we are ready to prove Theorem 4. Let A be a finite algebra in $\boldsymbol{Z}$. The only subdirectly irreducible members of $\boldsymbol{Z}$ are C_2 and Z therefore A is a subdirect product of two algebras D and E, where D is a subdirect power of C_2 and E is a subdirect power of Z. Obviously, D is a distributive lattice. By Lemma 10, $E = Z^k$ for some integer k. Thus we have proved that A is (isomorphic to) a subalgebra of $D \times Z^k$. We prove that, in fact, $A = D \times Z^k$.

Let 1 be the greatest element of D and $a \in E = Z^k$. We show that $\langle 1, a\rangle \in A$. Indeed, since A is a subdirect product of D and E there are elements $b \in E$ and $d \in D$ satisfying

$$\langle 1, b\rangle \quad \text{and} \quad \langle d, a\rangle \in A .$$

Define $e \in E$ by the rule:

$$\begin{aligned} e(i) &= a(i) \quad \text{if} \quad b(i) \leqq a(i) \\ e(i) &> b(i) \quad \text{if} \quad b(i) > a(i) . \end{aligned}$$

Note that $e(i)$, the ith component of e, is in Z so the condition $e(i) > b(i)$ uniquely determines $e(i)$. Choose an $f \in D$ such that $\langle f, e\rangle \in A$. Then

$$(\langle 1, b\rangle \vee \langle f, e\rangle) \vee \langle d, a\rangle \in A .$$

This element is obviously of the form $\langle 1, g\rangle$, and

$$g(i) = (b(i) \vee e(i)) \vee a(i) = e(i) \vee a(i) = a(i) .$$

Thus $g = a$, proving $\langle 1, a\rangle \in A$.

Now take an arbitrary $d \in D$ and $a \in E$. Then $\langle d, b\rangle \in A$ for some $b \in E$. For $a, b \in E$ let us construct $e \in E$ as follows:

$$\begin{aligned} &e(i) = a(i) &&\text{if} \quad a(i) = b(i) \\ &e(i) \neq a(i) \quad \text{and} \quad b(i) &&\text{if} \quad a(i) \neq b(i) . \end{aligned}$$

Then $\langle 1, a\rangle$ and $\langle 1, e\rangle \in A$ and so

$$(\langle d, b\rangle \wedge \langle 1, e\rangle) \wedge \langle 1, a\rangle = \langle d, g\rangle \in A ,$$

where $g(i) = (b(i) \wedge e(i)) \wedge a(i)$. If $a(i) = b(i)$, then $e(i) = a(i)$, and so $g(i) = a(i)$. If $a(i) < b(i)$, then (we are in Z) $b(i) < e(i) < a(i)$ and so $g(i) = (b(i) \wedge e(i)) \wedge a(i) = b(i) \wedge a(i) = a(i)$. Finally, if $b(i) < a(i)$, then $a(i) < e(i) < b(i)$, hence $g(i) = (b(i) \wedge e(i)) \wedge a(i) = e(i) \wedge a(i) = a(i)$, proving $g = a$ and $\langle d, a\rangle \in A$. This completes the proof of the first part of Theorem 4.

To prove the uniqueness of D we show that D is a maximal homomorphic image of A in $\boldsymbol{D}$. It is obvious that D is a homomorphic image of A in $\boldsymbol{D}$. Let Θ be an arbitrary congruence relation on A such that $A/\Theta \in \boldsymbol{D}$. By the Corollary to Lemma 1, $\Theta = \Phi \times \Theta_1 \times \cdots \times \Theta_k$ and

$$A/\Theta \cong D/\Theta \times Z_1/\Theta_1 \times \cdots \times Z_k/\Theta_k\,, \quad \text{where} \quad Z_1 = \cdots = Z_k = Z\,.$$

Since $Z/\Theta_i \in \boldsymbol{D}$ only if $|Z/\Theta_i| = 1$, we conclude that $A/\Theta \cong D/\Theta$, proving that D is a maximal homomorphic image. This implies the uniqueness of D up to isomorphism. Knowing that D is unique, it obviously follows that k is unique. This concludes the proof of Theorem 4.

COROLLARY. *The congruence lattice $C(A)$ of any finite algebra A in $\boldsymbol{Z}$ is a finite Boolean lattice.*

Proof. Indeed, if $A = D \times Z^k$, then $C(A) \cong C(D) \times C(Z)^k = C(D) \times C_2^k$, and $C(D)$ is known to be Boolean.

7. Free algebras. The following results describe the structure of free algebras over $\boldsymbol{Z}$ in terms of the free algebra over $\boldsymbol{D}$.

THEOREM 5. *Let $F_{\boldsymbol{D}}(n)$ and $F_{\boldsymbol{Z}}(n)$ denote the free algebra on n generators over $\boldsymbol{D}$ and $\boldsymbol{Z}$, respectively. Then*

$$F_{\boldsymbol{Z}}(n) \cong F_{\boldsymbol{D}}(n) \times Z^{k_n}\,,$$

where $k_n = 3^{n-1} - 2^n + 1$.

Proof[2]. Let $F = F_{\boldsymbol{Z}}(n)$ and $D = F_{\boldsymbol{D}}(n)$, and let $X = \{x_1, \cdots, x_n\}$ be a set of free generators of F. Obviously, $F_{\boldsymbol{Z}}(n)$ is a subalgebra of Z^{3^n}, hence finite. Thus by Theorem 4,

$$F \cong D \times Z^k\,,$$

for some nonnegative integer k. By the corollaries to Lemma 1 and Theorem 4, k is the number of congruence relations Θ of F satisfying $F/\Theta \cong Z$.

Let φ_1 and φ_2 be homomorphisms of F onto Z inducing the same congruence relation Θ. Then $X\varphi_1 = X\varphi_2 = Z$ and φ_i, $i = 1, 2$, partitions X into X_0^i, X_1^i, X_2^i by setting $X_j^i = j\varphi_i^{-1}$. Since these partitions are the restrictions of the Θ-classes to X, they agree. It is easily seen that for $a \in X_0^i$ and $b \in X_1^i$ the fact that $a\varphi_i < b\varphi_i$ is expressed by $a \vee b \equiv b(\Theta)$. Therefore, for some automorphisms α of Z, we have $\varphi_1 = \varphi_2\alpha$. Since the converse is obvious, we conclude that k equals

[2] We would like to thank R. Quackenbush for a considerable simplification of the original proof.

the number of maps of X onto Z up to automorphisms of Z, or equivalently, all maps φ of X onto Z satisfying $x_1\varphi = 0$. There are altogether 3^{n-1} maps of $\{x_2, \cdots, x_{n-1}\}$ into Z. Of these, 2^{n-1} does not have 1 in the image and 2^{n-1} does not have 2 in the image, the overlap being one map (the constant 0 map). Therefore, $k = 3^{n-1} - 2\cdot 2^{n-1} + 1 = 3^{n-1} - 2^n + 1$, as claimed.

We can apply Theorem 5 to describe all finite projective algebras in $\boldsymbol{Z}$.

COROLLARY. *A finite algebra A is projective in $\boldsymbol{Z}$ if and only if it is isomorphic to some $P \times Z^k$ where k is a nonnegative integer and P is projective in $\boldsymbol{D}$.*

REMARK. By R. Balbes [2] (see also G. Grätzer and B. Wolk [11]) a finite distributive lattice is projective in $\boldsymbol{D}$ if and only if the join of any two meet irreducible elements is again meet irreducible.

Proof. It is well-known that A is projective if and only if it is a retract (idempotent endomorphic image) of a free algebra. Firstly, let $A = P \times Z^k$ where P is projective in $\boldsymbol{D}$. Choose an integer n such that P is a retract on $F_D(n)$ and $k < k_n$. Then, obviously, A is a retract of $F_Z(n)$. Conversely, let A be a retract of some $F_Z(n)$. By Theorem 4, $A \cong D \times Z^k$. Since D is a retract of A, we conclude that D is a retract of $F_Z(n)$. By the Corollary of Lemma 1, the retraction must collapse all copies of Z, hence P is a retract of $F_D(n)$, showing that P is projective in $\boldsymbol{D}$. This concludes the proof.

8. **Injective algebras.** The algebra I of $\boldsymbol{Z}$ is called *injective* (see, for instance, [8], §13) if for any $A, B \in \boldsymbol{Z}$, A a subalgebra B, any homomorphism $\varphi: A \to I$ can be extended to a homomorphism of B into I.

THEOREM 6. *Z is injective in $\boldsymbol{Z}$. Any direct power of Z is injective in $\boldsymbol{Z}$ and, therefore, every algebra can be embedded in an injective. An algebra is injective if and only if it is isomorphic to the extension of Z by a complete Boolean algebra.*

Proof. Rather than giving a direct proof of these results we shall employ a trick from [14] and then use a result of [3] to get the last statement of Theorem 6, which implies the other two.

Let us denote by $\hat{Z}$ the algebra Z with three new nullary operations: 0, 1, and 2. Let $\hat{\boldsymbol{Z}}$ denote the equational class generated by $\hat{Z}$. Just as in Lemma 2, every algebra in $\hat{\boldsymbol{Z}}$ can be embedded in a direct power of $\hat{Z}$.

$\hat{\boldsymbol{Z}}$ is generated by a finite simple algebra $\hat{Z}$ with no subalgebras and so by a result of A. Day [3], the injectives in $\hat{\boldsymbol{D}}$ are exactly the algebras $\hat{Z}[B]$ where B is a complete Boolean algebra (for this concept see [7], §22).

Therefore, it suffices to prove the following statement:

An algebra A is injective in $\boldsymbol{Z}$ if and only if 0, 1, and 2 can be interpreted on A so that the resulting algebra $\hat{A}$ belongs to $\hat{\boldsymbol{Z}}$ and $\hat{A}$ is injective in $\hat{\boldsymbol{Z}}$.

Indeed, if A is injective in $\boldsymbol{Z}$, then for some set J there is a homomorphism (in fact, a retraction) φ of Z^J onto A. We can interpret $0, 1, 2$ on Z^J as on $(\hat{Z})^J$, and then on A by 0φ, 1φ, and 2φ. This makes $\hat{A}$ a homomorphic image of $(\hat{Z})^J \in \hat{\boldsymbol{Z}}$, and so $\hat{A} \in \hat{\boldsymbol{Z}}$. Since $\hat{A}$ is a retract of $(\hat{Z})^J$, it is injective in $\hat{\boldsymbol{Z}}$. The converse is obvious. This completes the proof of Theorem 6.

It follows from Theorem 6 and B. Banasehewski [1] that every algebra A in $\boldsymbol{Z}$ has an injective hull uniquely determined up to isomorphism (leaving A fixed).

References

1. B. Banaschewski, *Injectivity and essential extensions in equational classes of algebras*, Proceedings of the Conference on Universal Algebra, October, 1969, (Queen's University, Kingston, 1970), 131–147.

2. R. Balbes, *Projective and injective distributive lattices*, Pacific J. Math., **21** (1967), 405–420.

3. A. Day, *Injectivity in congruence distributive equational classes*, Thesis, McMaster University, 1970.

4. ———, *A note on the congruence extension property*, Algebra Universalis, **1** (1971), 234–235.

5. E. Fried, *Tournaments and nonassociative lattices*, Ann. Univ. Sci. Budapest Eötvös Sect. Math., **13** (1970), 151–164.

6. E. Fried and G. Grätzer, *On some classes of nonassociative lattices*, To appear.

7. G. Grätzer, *Universal Algebra*, The University Series in Higher Mathematics, D. Van Nostrand Co., Princeton, N. J., 1968.

8. ———, *Lattice Theory, First Concepts and Distributive Lattices*, W. H. Freeman and Co., San Francisco, Calif., 1971.

9. G. Grätzer and H. Lakser, *Identities for equational classes generated by tournaments*, Notices Amer. Math. Soc., **18** (1971), 794.

10. G. Grätzer and E. T. Schmidt, *Ideals and congruence relations in lattices*, Acta Math. Acad. Sci. Hungarí, **9** (1958), 137–175.

11. G. Grätzer and B. Wolk, *Finite projective distributive lattices*, Canad. Math. Bull., **13** (1970), 139–140.

12. B. Jónsson, *Extensions of relational structures*, Symposium on the Theory of Models, North-Holland Publ. Co., Amsterdam, 1965, 146–157.

13. ———, *Algebras whose congruence lattices are distributive*, Math. Scand., **21** (1967), 110–121.

14. J. Kagan and R. Quackenbush, *Self-dual closure algebras*, Manuscript.

15. R. Padmanabhan, *Equational theory of idempotent algebras*, Algebra Universalis, **2** (1972), 57–61.

16. H. L. Skala, *Trellis theory*, Algebra Universalis, **1** (1971), 218–233.

Received July 19, 1972 and in revised form October 15, 1972.

University of Manitoba

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

ON NUMERICAL RANGES OF ELEMENTS OF LOCALLY m-CONVEX ALGEBRAS

J. R. GILES AND D. O. KOEHLER

The concept of numerical range is extended from normed algebras to locally m-convex algebras. It is shown that the approximating relations between the numerical range and the spectrum of an element are preserved in the generalization. The set of elements with bounded numerical range is characterized and the relation between boundedness of the spectrum and of the numerical range is discussed. The Vidav-Palmer theory is generalized to give a characterization of b^*-algebras by numerical range.

In a complex unital Banach algebra the numerical range of an element is a set of complex numbers which can be used to approximate the spectrum of an element. In a complex locally m-convex algebra with identity, for each element we define a set of numerical ranges and establish similar approximation to the spectrum of the element. In a normed algebra the spectrum and the numerical range of each element are bounded sets, but in a locally m-convex algebra the spectrum and the numerical ranges of an element may be unbounded. For a locally m-convex algebra with identity we characterize those elements with a bounded numerical range as an important normed subalgebra, and we discuss the relation between boundedness of the spectrum and the numerical ranges. In the normed algebra theory the study of hermitian elements, those with real numerical range, has led to the important Vidav-Palmer theory characterizing unital B^*-algebras among unital Banach algebras. We generalize the results of this theory to a characterization of b^*-algebras by numerical range.

We would like to thank Dr. T. Husain for the valuable discussions we have had with him on this subject. We would also like to express our appreciation to the referee for his valuable suggestions.

1. The numerical ranges of an element. For a complex normed algebra $(A, \|\cdot\|)$ with identity 1 where $\|1\| = 1$, i.e., a *complex unital normed algebra*, we define the set

$$D(A, \|\cdot\|; 1) \equiv \{f \in A' : f(1) = 1 \quad \text{and} \quad \|f\| = 1\} .$$

For each $a \in A$ we define the *numerical range of* a as the set

$$V(A, \|\cdot\|; a) \equiv \{f(a) : f \in D(A, \|\cdot\|; 1)\} ,$$

and the *numerical radius of* a as

$$v(A, \|\cdot\|; a) \equiv \sup\{|\lambda|: \lambda \in V(A, \|\cdot\|; a)\} .$$

The set $D(A, \|\cdot\|; 1)$ is a convex weak * compact subset of A' and the numerical range $V(A, \|\cdot\|; a)$ is also a convex compact subset of the complex numbers, [2, p. 16]. The properties and applications of normed algebra numerical ranges have been studied extensively and the main results are conveniently presented by F. F. Bonsall and J. Duncan in [2].

Locally m-convex algebras, i.e., l.m.c. algebras, are examined in some detail by E. A. Michael in [4]. We call a l.m.c. algebra with identity a *unital l.m.c. algebra*. It is our aim to extend the concept of numerical range from complex unital normed algebras to complex unital l.m.c. algebras. It is sufficient for our purpose to note that, for a given l.m.c. algebra A with identity 1 there exists a separating family of submultiplicative semi-norms $\{p_\alpha\}$ on A which generates the topology and is such that $p_\alpha(1) = 1$ for all α, [3, p. 7]. Given such an algebra, we denote by $P(A)$ the class of all such families of semi-norms on A, and by $(A, \{p_\alpha\})$ the algebra A with a particular family of semi-norms $\{p_\alpha\} \in P(A)$.

Given $(A, \{p_\alpha\})$, for each α let N_α denote the nullspace of p_α, A_α denote the quotient space A/N_α, and $\|\cdot\|_\alpha$ denote the norm on A_α defined by $\|x + N_\alpha\|_\alpha = p_\alpha(x)$. For each α consider the natural linear mapping $x \mapsto x_\alpha \equiv x + N_\alpha$ of A onto A_α. We note that 1_α is the identity in A_α and that $\|1_\alpha\|_\alpha = 1$ for each α. Michael has given the significant result that A is isomorphic to a subalgebra of the product of the normed algebras $(A_\alpha, \|\cdot\|_\alpha)$, Proposition 2.7, [4, p. 13]. Using this characterization of l.m.c. algebras we are able to generalize much of the numerical range theory for normed algebras directly to a theory of numerical range for l.m.c. algebras.

Given $(A, \{p_\alpha\})$, we define the set

$$D_\alpha(A, p_\alpha; 1) \equiv \{f \in A' : f(1) = 1 \text{ and } |f(x)| \leqq p_\alpha(x) \text{ for all } x \in A\} ,$$

and we write

$$D(A, \{p_\alpha\}; 1) \equiv \bigcup_\alpha \{D_\alpha(A, p_\alpha; 1)\} .$$

For each $a \in A$ we write

$$V_\alpha(A, p_\alpha; a) \equiv \{f(a) : f \in D_\alpha(A, p_\alpha; 1)\} ,$$

and define the *numerical range of* a as the set

$$V(A, \{p_\alpha\}; a) \equiv \bigcup_\alpha \{V_\alpha(A, p_\alpha; a)\} .$$

To each linear functional f on (A, p_α) which annihilates N_α, we can define the linear functional F on A_α by $F(x_\alpha) = f(x)$, and to each linear functional F on A_α we can define the linear functional f on (A, p_α) by $f(x) = F(x_\alpha)$. Consequently, from the definition of the norm in A_α we see that $D_\alpha(A, p_\alpha; 1)$ is isomorphic to $D(A_\alpha, \|\cdot\|_\alpha; 1_\alpha)$, and for $a \in A$

$$V_\alpha(A, p_\alpha; a) = V(A_\alpha, \|\cdot\|_\alpha; a_\alpha) .$$

Hence, we have the numerical range of a characterized by the normed algebra numerical ranges of the a_α in that

$$V(A, \{p_\alpha\}; a) = \bigcup_\alpha \{V(A_\alpha, \|\cdot\|_\alpha; a_\alpha)\} .$$

Both $D(A, \{p_\alpha\}; 1)$ and $V(A, \{p_\alpha\}; a)$ depend upon the particular family of semi-norms $\{p_\alpha\} \in P(A)$ chosen to associate with A. It is clear that when $\{p_\alpha\}$ is a directed family, $D(A, \{p_\alpha\}; 1)$ is a convex subset of A' and the numerical range $V(A, \{p_\alpha\}; a)$ is a convex subset of the complex numbers.

For each $a \in A$ we write

$$v_\alpha(A, p_\alpha; a) \equiv \sup\{|\lambda| : \lambda \in V_\alpha(A, p_\alpha; a)\} ,$$

and we define the *numerical radius of* a as

$$v(A, \{p_\alpha\}; a) \equiv \sup\{|\lambda| : \lambda \in V(A, \{p_\alpha\}; a)\} .$$

We note that $v_\alpha(A, p_\alpha; a) \leqq p_\alpha(a)$ for each α, and we allow $v(A, \{p_\alpha\}; a) = \infty$. We have that

$$\begin{aligned} v(A, \{p_\alpha\}; a) &= \sup_\alpha v_\alpha(A, p_\alpha; a) \\ &= \sup_\alpha v(A_\alpha, \|\cdot\|_\alpha; a_\alpha) . \end{aligned}$$

It is clear that the numerical range and the numerical radius have the following properties. For $a \in A$ and λ, μ complex

$$V(A, \{p_\alpha\}; \lambda a + \mu) = \lambda V(A, \{p_\alpha\}; a) + \mu$$

and

$$v(A, \{p_\alpha\}; \lambda a + \mu) \leqq |\lambda| v(A, \{p_\alpha\}; a) + |\mu| ,$$

and for $a, b \in A$

$$V(A, \{p_\alpha\}; a + b) \subseteqq V(A, \{p_\alpha\}; a) + V(A, \{p_\alpha\}; b)$$

and

$$v(A, \{p_\alpha\}; a + b) \leqq v(A, \{p_\alpha\}; a) + v(A, \{p_\alpha\}; b) .$$

2. **The numerical ranges and the spectrum.** In a unital Banach

algebra the numerical range of an element approximates its spectrum. We now establish similar approximating relations between the numerical ranges and the spectrum of an element in a complete unital l.m.c. algebra.

We recall that, given an algebra A with identity, for each $a \in A$, the *spectrum of* a is defined as the set

$$\sigma(A; a) \equiv \{\lambda : a - \lambda \text{ is not invertible}\} .$$

THEOREM 1. *Let A be a complete unital l.m.c. algebra. Given $(A, \{p_\alpha\})$, for each $a \in A$*

$$\sigma(A, a) \subseteqq V(A, \{p_\alpha\}; a) .$$

Proof. For each α let $\bar{A}_\alpha$ denote the completion of A_α. We have from Corollary 5.3(a), [4, p. 22] that

$$\sigma(A; a) = \bigcup_\alpha \sigma(\bar{A}_\alpha; a_\alpha) .$$

But from Theorem 2.6, [2, p. 19] we have that

$$\sigma(\bar{A}_\alpha; a) \subseteqq V(\bar{A}_\alpha, \|\cdot\|_\alpha; a_\alpha) ,$$

and from Theorem 2.4, [2, p. 16] that

$$V(\bar{A}_\alpha, \|\cdot\|_\alpha; a_\alpha) = V(A_\alpha, \|\cdot\|_\alpha; a_\alpha)$$

so it follows that

$$\begin{aligned} \sigma(A; a) &\subseteqq \bigcup_\alpha \{V(A_\alpha, \|\cdot\|_\alpha; a_\alpha)\} \\ &= V(A, \{p_\alpha\}; a) . \end{aligned}$$

THEOREM 2. *Let A be a complete unital l.m.c. algebra. For each $a \in A$*

$$\operatorname{co} \sigma(A; a) \subseteqq \bigcap \{V(A, \{p_\alpha\}; a) : \{p_\alpha\} \in P(A)\} \subseteqq \overline{\operatorname{co}}\, \sigma(A; a) .$$

Proof. From Theorem 1 we have that

$$\operatorname{co} \sigma(A; a) \subseteqq \bigcap \{V(A, \{p_\alpha\}, a); \{p_\alpha\} \in P(A)\} .$$

If $\overline{\operatorname{co}}\, \sigma(A; a)$ is not all the complex plane then, for any $\lambda \notin \overline{\operatorname{co}}\, \sigma(A; a)$ there exists an open disc D_λ center λ such that D_λ can be strictly separated from $\overline{\operatorname{co}}\, \sigma(A; a)$ by a straight line L. Since

$$\sigma(A; a) = \bigcup_\alpha \sigma(\bar{A}_\alpha; a_\alpha) ,$$

D_λ is strictly separated from $\sigma(\bar{A}_\alpha; a_\alpha)$ for any α, by the straight line L. However, for each α, $\sigma(\bar{A}_\alpha; a_\alpha)$ is a compact set so there exists an open disc $D_\alpha \supseteqq \sigma(\bar{A}_\alpha; a_\alpha)$ which is strictly separated from D_λ by the same straight line L. We have from [2, p. 23] that, for each α, there exists a norm $\|\cdot\|'_\alpha$ equivalent to $\|\cdot\|_\alpha$ on $\bar{A}_\alpha$ such that

$$\sigma(\bar{A}_\alpha; a_\alpha) \subseteqq V(\bar{A}_\alpha, \|\cdot\|'_\alpha; a_\alpha) \subseteqq D_\alpha .$$

Now for each α,

$$V(\bar{A}_\alpha, \|\cdot\|'_\alpha; a_\alpha) = V(A, \|\cdot\|'_\alpha; a_\alpha) .$$

Defining the semi-norm p'_α on A by

$$p'_\alpha(x) = \|x_\alpha\|'_\alpha ,$$

it is clear that the family $\{p'_\alpha\} \in P(A)$, and

$$V(A, \{p'_\alpha\}; a) = \bigcup_\alpha \{V(A_\alpha, \|\cdot\|'_\alpha; a_\alpha)\} .$$

So D_λ is strictly separated from $V(A, \{p'_\alpha\}, a)$ by the straight line L. It follows that D_λ is strictly separated from $\bigcap\{V(A, \{p_\alpha\}; a)\colon \{p_\alpha\} \in P(A)\}$, and this implies that

$$\bigcap\{V(A, \{p_\alpha\}; a)\colon \{p_\alpha\} \in P(A)\} \subseteqq \overline{\mathrm{co}}\, \sigma(A; a) .$$

3. Elements with bounded numerical range. We now establish an important set of inequalities which are generalizations of an inequality from the normed algebra theory, and we use them to characterize elements with bounded numerical range.

LEMMA 1. *Let A be a unital l.m.c. algebra. Given $(A, \{p_\alpha\})$, for $a \in A$ and each α*

$$v(A, \{p_\alpha\}, a) \geqq \frac{1}{e} p_\alpha(a) .$$

Proof. From Theorem 4.1, [2, p. 34] we have, for each α

$$v(A_\alpha, \|\cdot\|_\alpha; a_\alpha) \geqq \frac{1}{e}\|a_\alpha\|_\alpha .$$

So

$$\begin{aligned} v(A, \{p_\alpha\}, a) &= \sup_\alpha v(A_\alpha, \|\cdot\|_\alpha; a_\alpha) \\ &\geqq \frac{1}{e}\|a_\alpha\|_\alpha = \frac{1}{e} p_\alpha(a) , \end{aligned}$$

for each α.

From the fact that every $\{p_\alpha\} \in P(A)$ is a separating family we can make the following deduction.

COROLLARY 1. *If for a given* $a \in A$, *there exists an* $(A, \{p_\alpha\})$ *such that* $V(A, \{p_\alpha\}; a) = \{0\}$ *then* $a = 0$.

We can also make a statement about elements with bounded numerical range.

COROLLARY 2. *Given* $(A, \{p_\alpha\})$, *if for* $a \in A$, $V(A, \{p_\alpha\}; a)$ *is bounded then* $\sup_\alpha p_\alpha(a) < \infty$.

For the characterization of the set of elements with bounded numerical range we also use the following lemma.

LEMMA 2. *Let* A *be a unital l.m.c. algebra. Given* $(A, \{p_\alpha\})$ *we have, for each* $a \in A$

$$\sup \operatorname{Re} V(A, \{p_\alpha\}; a) = \left.\begin{matrix} \inf\limits_{\lambda>0} \\ \lim\limits_{\lambda\to 0+} \end{matrix}\right\} \frac{1}{\lambda}\left\{\sup_\alpha p_\alpha(1 + \lambda a) - 1\right\}.$$

Proof. It is clear that for any $f \in D(A, \{p_\alpha\}; 1)$ and $\lambda > 0$

$$\operatorname{Re} f(a) \leqq \frac{1}{\lambda}\left\{\sup_\alpha p_\alpha(1 + \lambda a) - 1\right\}$$

and therefore,

$$\sup \operatorname{Re} V(A, \{p_\alpha\}, a) \leqq \inf_{\lambda>0} \frac{1}{\lambda}\left\{\sup_\alpha p_\alpha(1 + \lambda a) - 1\right\}. \tag{1}$$

It follows that the result holds when $V(A, \{p_\alpha\}; a)$ is unbounded. We consider the case when $V(A, \{p_\alpha\}; a)$ is bounded, and write for every α

$$\tau_\alpha \equiv \sup \operatorname{Re} V_\alpha(A_\alpha, \|\cdot\|_\alpha; a_\alpha),$$

and

$$\tau \equiv \sup \operatorname{Re} V(A, \{p_\alpha\}; a).$$

Now by [2, p. 18], for every α

$$\frac{1}{\lambda}\{\| 1_\alpha + \lambda a_\alpha \|_\alpha - 1\} \leqq (1 - \lambda\tau_\alpha)^{-1}\{\tau_\alpha + \lambda \| a_\alpha^2 \|_\alpha\},$$

when $0 < \lambda < \| a_\alpha \|_\alpha^{-1}$. Since $V(A, \{p_\alpha\}; a)$ is bounded we have from Corollary 2 to Lemma 1 that there exists an $M > 0$ such that $M \geqq$

$\sup_\alpha p_\alpha(a)$. Then, for every α

$$\frac{1}{\lambda}\{p_\alpha(1+\lambda a)-1\} \leqq (1-\lambda\tau)^{-1}\{\tau+\lambda M^2\}$$

when $0<\lambda<1/M$. Therefore,

$$\lim_{\lambda\to 0+}\frac{1}{\lambda}\left\{\sup_\alpha p_\alpha(1+\lambda a)-1\right\} \leqq \tau .$$

Together with inequality (1), this completes the proof.

Let A be a unital l.m.c. algebra. Given $(A, \{p_\alpha\})$ we can define the subalgebra

$$B \equiv \left\{x \in A\colon \sup_\alpha p_\alpha(x) < \infty\right\} .$$

Now $p(x) \equiv \sup_\alpha p_\alpha(x)$ is a norm for B since $\{p_\alpha\}$ is a separating family and we note that $1 \in B$ and $p(1) = 1$.

It can be seen from the proof of Theorem 2.3, [1, p. 32], that if A is a complete unital l.m.c. algebra then given $(A, \{p_\alpha\})$, the normed subalgebra (B, p) is complete. However, an examination of the sequence $\{x_n\}$ where $x_n = \{1, 2, 3, \cdots, n, \cdots, n, \cdots\}$, in the algebra A of Example 2 below, shows that there exists an incomplete unital l.m.c. algebra A with $\{p_\alpha\} \in P(A)$ such that (B, p) is complete.

Given $(A, \{p_\alpha\})$, we can characterize elements with bounded numerical range as elements of (B, p).

THEOREM 3. *If A is a unital l.m.c. algebra then given $(A, \{p_\alpha\})$,*

$$B = \{x \in A\colon V(A, \{p_\alpha\}; x) \text{ is bounded}\}$$

and when $\{p_\alpha\}$ is a directed family, for every $a \in B$

$$\overline{V(A, \{p_\alpha\}; a)} = V(B, p; a) .$$

Proof. If, for a given $a \in A$, $V(A, \{p_\alpha\}; a)$ is bounded then Corollary 2 to Lemma 1 implies that $a \in B$. If $a \in B$ then $\sup_\alpha p_\alpha(1+\lambda a) = p(1+\lambda a)$ for all λ, so

$$\begin{aligned}\sup \operatorname{Re} V(A, \{p_\alpha\}; a) &= \left.\begin{matrix}\inf\limits_{\lambda>0}\\ \lim\limits_{\lambda\to 0+}\end{matrix}\right\}\frac{1}{\lambda}\{p(1+\lambda a)-1\}\\ &= \sup \operatorname{Re} V(B, p; a) .\end{aligned}$$

Hence, since $V(B, p; \lambda a) = \lambda V(B, p; a)$ and $V(A, \{p_\alpha\}; \lambda a) = \lambda V(A, \{p_\alpha\}; a)$ for all λ complex, $|\lambda| = 1$, we deduce that every $a \in B$ has bounded numerical range $V(A, \{p_\alpha\}; a)$. When $\{p_\alpha\}$ is a directed family, both numerical ranges are convex sets so we deduce from the Krein-Milman Theorem that for every $a \in B$

$$\overline{V(A, \{p_\alpha\}; a)} = V(B, p; a) .$$

The following result relates boundedness of the spectrum to boundedness of the numerical range.

THEOREM 4. *Let A be a complete unital l.m.c. algebra. For any $a \in A$, $\sigma(A; a)$ is bounded if and only if there exists an $(A, \{p_\alpha\})$ such that $V(A, \{p_\alpha\}; a)$ is bounded.*

Proof. If for $a \in A$ there exists an $(A, \{p_\alpha\})$ such that $V(A, \{p_\alpha\}; a)$ is bounded, then it follows from Theorem 1 that $\sigma(A; a)$ is bounded.

Conversely, consider $a \in A$ with $\sigma(A; a)$ bounded. There exists a disc D in the complex plane such that $\sigma(A; a) \subseteqq D$. Now $\sigma(A; a) = \bigcup_\alpha \sigma(\bar{A}_\alpha; a_\alpha)$. From [2, p. 23], for each α there exists a norm $\|\cdot\|'_\alpha$ equivalent to $\|\cdot\|_\alpha$ on $\bar{A}_\alpha$ such that

$$\sigma(\bar{A}_\alpha; a_\alpha) \subseteqq V(\bar{A}_\alpha; \|\cdot\|'_\alpha; a_\alpha) \subseteqq D .$$

For each α, defining the semi-norm p'_α on A by

$$p'_\alpha(x) = \|x_\alpha\|'_\alpha ,$$

the family $\{p'_\alpha\} \in P(A)$ and

$$\sigma(A; a) \subseteqq V(A, \{p'_\alpha\}; a) = \bigcup_\alpha V(A_\alpha, \|\cdot\|'_\alpha, a_\alpha) \subseteqq D .$$

Further to the relation between boundedness of the spectrum and the numerical range given in Theorem 4, the following example shows that there exist l.m.c. algebras A where $\sigma(A; a)$ is bounded for a given $a \in A$ but where there exists $(A, \{p_\alpha\})$ such that $V(A, \{p_\alpha\}; a)$ is unbounded.

EXAMPLE 1. Let A be the algebra of all sequences of complex numbers, $x \equiv \{\lambda_1, \lambda_2, \cdots, \lambda_n, \cdots\}$ with pointwise definition of addition and multiplication by a scalar, but with convolution multiplication and with unit $1 \equiv \{1, 0, \cdots, 0, \cdots\}$. A sequence of submultiplicative semi-norms $\{p_n\}$ is defined on A by

$$p_n(x) = \sum_{k=1}^{n} |\lambda_k| ,$$

and the sequence satisfies $p_n(1) = 1$ for all n, and is separating. Consider $a \in A$ such that $\lambda_n \nrightarrow 0$. Then $p_n(a) \to \infty$, so by Lemma 1 $V(A, \{p_n\}; a)$ is unbounded. But $\sigma(A; a) = \{\lambda_1\}$, which is bounded.

It is worth noting that a complete unital l.m.c. algebra with a bounded numerical range property has the following property.

THEOREM 5. *If A is a complete unital l.m.c. algebra where there exists an $(A, \{p_\alpha\})$ such that $V(A, \{p_\alpha\}; a)$ is bounded for all $a \in A$, then $\sigma(A; a)$ is compact for all $a \in A$.*

Proof. We note that $B = A$ and so $\sigma(A; a) = \sigma(B; a)$, for every $a \in A$. Since A is complete it follows that (B, p) is complete and so $\sigma(B; a)$ is compact for every $a \in A$.

However, the algebra A of Example 1 has $\sigma(A; a)$ compact for all $a \in A$ but there exists an $(A, \{p_\alpha\})$ such that $V(A, \{p_\alpha\}; a)$ is not bounded for all $a \in A$.

It is known that there exist non-normable l.m.c. algebras where $\sigma(A; a)$ is compact for all $a \in A$, [4, p. 80]. The following example gives the further information that there exist non-normable l.m.c. algebras A where, for a certain $(A, \{p_\alpha\})$, $V(A, \{p_\alpha\}; a)$ is bounded for all $a \in A$.

EXAMPLE 2. Let A be the algebra l^∞ of all bounded sequences of complex numbers, $x \equiv \{\lambda_1, \lambda_2, \cdots, \lambda_n, \cdots\}$, with pointwise definition of the algebra operations and with unit $1 \equiv \{1, 1, \cdots, 1, \cdots\}$. A sequence of submultiplicative semi-norms $\{p_n\}$ is defined on A by

$$p_n(x) = |\lambda_n| ,$$

and the sequence satisfies $p_n(1) = 1$ for all n, and is separating. Now p, defined by

$$p(x) \equiv \sup_n \{p_n(x)\}$$

is the usual l^∞-norm on A, so $B = A$ and from Theorem 3, $V(A, \{p_n\}; a)$ is bounded for all $a \in A$. However, it is clear that $(A, \{p_n\})$ is non-normable.

4. A characterization of b^*-algebras by numerical range. A *l.m.c. * algebra* A is a l.m.c. algebra with a continuous involution *. We let $S(A)$ denote the set $\{x \in A: x = x^*\}$, the selfadjoint elements of A. A *b^*-algebra* A is a complete l.m.c. * algebra where there exists a family $\{p_\alpha\} \in P(A)$ such that $p_\alpha(x^*x) = p_\alpha(x)^2$ for all $x \in A$ and every α, [1, p. 31].

In a unital normed algebra $(A, \|\cdot\|)$, the *set of hermitian elements* $H(A, \|\cdot\|)$ is the set of elements a with real numerical range $V(A, \|\cdot\|; a)$. For a unital l.m.c. algebra A, given $(A, \{p_\alpha\})$ we define the *set of hermitian elements* $H(A, \{p_\alpha\})$ as the set of elements a with real numerical range $V(A, \{p_\alpha\}; a)$. It is clear from the definition of the numerical range in $(A, \{p_\alpha\})$ that $a \in H(A, \{p_\alpha\})$ if and only if $a_\alpha \in H(A_\alpha, \|\cdot\|_\alpha)$ for all α.

One of the outstanding successes of the normed algebra numerical range theory is the Vidav-Palmer Theorem [2, p. 65] which characterizes unital B^* algebras as unital Banach algebras which have an hermitian decomposition. We now consider a generalization of this work to the characterization of unital b^*-algebras amongst the complete unital l.m.c. algebras.

We need the following property of the hermitian elements.

LEMMA 3. *Let A be a complete unital l.m.c. algebra. Given $(A, \{p_\alpha\})$, the set $H(A, \{p_\alpha\})$ is closed.*

Proof. Consider h a cluster point of $H(A, \{p_\alpha\})$. Then, for each α, h_α is a cluster point of $H(A_\alpha, \|\cdot\|_\alpha)$. But by Lemma 7, [5, p. 198], $H(A_\alpha, \|\cdot\|_\alpha)$ is closed in $(A_\alpha, \|\cdot\|_\alpha)$, so $h_\alpha \in H(A_\alpha, \|\cdot\|_\alpha)$. Since $V(A, \{p_\alpha\}; h) = \bigcup_\alpha V(A_\alpha, \|\cdot\|_\alpha; h_\alpha)$ we have that $h \in H(A, \{p_\alpha\})$.

THEOREM 6. *Let A be a complete unital l.m.c. algebra. Given $(A, \{p_\alpha\})$, the following statements are equivalent.*

(i) *$A = H(A, \{p_\alpha\}) + iH(A, \{p_\alpha\})$, a direct sum,*

(ii) *There is an involution $*$ on A such that A is a l.m.c. $*$ algebra where $S(A) = H(A, \{p_\alpha\})$,*

(iii) *There is an involution $*$ on A such that A is $*$ isomorphic to a $*$ subalgebra of a product of B^*-algebras $(\bar{A}_\alpha, \|\cdot\|_\alpha)$,*

(iv) *There is an involution $*$ on A such that A is a b^*-algebra,*

(v) *There is an involution on B such that (B, p) is a dense B^*-algebra.*

Proof. (i) $\Rightarrow$ (ii) Since $A = H(A, \{p_\alpha\}) + iH(A, \{p_\alpha\})$, we define the involution $*$ on A as follows: for $x = h + ik$ where $h, k \in H(A, \{p_\alpha\})$ put $x^* = h - ik$. We need to show that $*$ is continuous on A. Now for every α, $A_\alpha = H(A_\alpha, \|\cdot\|_\alpha) + iH(A_\alpha, \|\cdot\|_\alpha)$ and $*$ induces an involution $*$ on A_α where for $x_\alpha = h_\alpha + ik$ we have $x_\alpha^* = h_\alpha - ik_\alpha$. But from Lemma 5.8, [2, p. 50], since $A_\alpha \subseteqq J(\bar{A}_\alpha)$, we have for every α, that $*$ is continuous on A_α and since $p_\alpha(x) = \|x_\alpha\|_\alpha$ for all $x \in A$, $*$ is continuous on A. It is clear that with this involution $*$, $S(A) = H(A, \{p_\alpha\})$.

(ii) $\Rightarrow$ (iii) Since $H(A, \{p_\alpha\}) = S(A)$ we have $A = H(A, \{p_\alpha\}) + iH(A, \{p_\alpha\})$ and so, for every α, $A_\alpha = H(A_\alpha, \|\cdot\|_\alpha) + iH(A_\alpha, \|\cdot\|_\alpha)$. But by Theorem 8.2, [2, p. 74], A_α is a pre-B^*-algebra and so $\bar{A}_\alpha$ is a B^*-algebra for every α. Our result follows from Michael's characterization of l.m.c. algebras, Proposition 2.7, [4, p. 13].

(iii) $\Rightarrow$ (iv) For every α, since $(\bar{A}_\alpha, \|\cdot\|_\alpha)$ is a B^*-algebra and $p_\alpha(x) = \|x_\alpha\|_\alpha$ for all $x \in A$, we have $p_\alpha(x^*x) = p_\alpha(x)^2$ for all $x \in A$; that is, A is a b^*-algebra.

(iv) $\Rightarrow$ (v) This is proved as Theorem 2.3, [1, p. 32].

(v) $\Rightarrow$ (i) Since (B, p) is a unital B^*-algebra, $B = H(B, p) + iH(B, p)$. For any $h, k \in H(B, p)$, we have from Theorem 3, that $h, k \in H(A, \{p_\alpha\})$. But then for every α, $h_\alpha, k_\alpha \in H(A_\alpha, \|\cdot\|_\alpha)$ and since $\|x_\alpha\|_\alpha = p_\alpha(x)$ for all $x \in A$, we have from inequality (1), [2, p. 50] that

$$p_\alpha(h) \leqq e p_\alpha(h + ik) .$$

This inequality implies that for any net $\{h_\gamma + ik_\gamma\}$ in B convergent to x in A, both $\{h_\gamma\}$ and $\{k_\gamma\}$ converge to say h and k. But by Lemma 3 the set $H(A, \{p_\alpha\})$ is closed in A so $h, k \in H(A, \{p_\alpha\})$ and $x = h + ik$. Since B is dense in A, we have $A = H(A, \{p_\alpha\}) + iH(A, \{p_\alpha\})$.

It should be noted that this theorem gives in (v) $\Rightarrow$ (iv), a converse to Theorem 2.3, [1, p. 32], and in (iv) $\Rightarrow$ (iii), a simpler proof for Theorem 2.4, [1, p. 32], by using numerical range techniques.

The following is an application of Theorem 6 and is a generalization of Theorem 7.6, [2, p. 71].

THEOREM 7. *Let Ω be a locally compact Hausdorff space and let $A \equiv \mathscr{C}(\Omega)$ be the algebra of all complex continuous functions on Ω. If A is an F-algebra under the compact-open topology, then any l.m.c. topology generated by a family of semi-norms $\{p_\alpha\}$ such that $p_\alpha(f) = p_\alpha(|f|)$ for all $f \in A$ and $p_\alpha(1) = 1$ for all α, under which A is an F-algebra, is the compact-open topology.*

Proof. We can introduce the exponential function in A, $\exp a = 1 + \sum_{n=1}^{\infty} (1/n!) a^n$, and it is clear that $(\exp a)_\alpha = \exp a_\alpha$, for every α. Now if g is a real continuous function on Ω then for real λ and for each α

$$\begin{aligned} \|\exp(i\lambda g_\alpha)\|_\alpha &= \|(\exp i\lambda g)_\alpha\|_\alpha \\ &= p_\alpha(\exp i\lambda g) \\ &= p_\alpha(|\exp i\lambda g|) \\ &= p_\alpha(1) = 1 . \end{aligned}$$

Therefore, by Lemma 5.2, [2, p. 46], $g_\alpha \in H(A_\alpha, \|\cdot\|_\alpha)$ for every α, and so $g \in H(A, \{p_\alpha\})$. We have then $A = H(A, \{p_\alpha\}) + iH(A, \{p_\alpha\})$ and by Theorem 6 we conclude that A is a b^*-algebra. Now, by Theorem 4.2, [1, p. 36], A with the compact-open topology is also a b^*-algebra. So by Theorem 3.7, [1, p. 35], the l.m.c. topology generated by $\{p_\alpha\}$ is the compact-open topology.

In the above theorem we note that A with the compact-open topology and A with the l.m.c. topology generated by the semi-norms, must both be F-algebras; one of these being an F-algebra is not sufficient. We are indebted to the referee for the following examples which illustrate this point.

EXAMPLE 3. Let $\Omega \equiv [0, 1]$ with the usual topology and let $\{K_\alpha\}$ be the set of compact countable subsets of Ω. A family of semi-norms $\{p_\alpha\}$ defined on A by

$$p_\alpha(x) = \sup_{x \in K_\alpha} \{|f(x)|\} ,$$

satisfies the conditions of the theorem except that A with this topology is not an F-algebra, [4, Example 3.8, p. 19]. However, A with the compact-open topology is a Banach algebra, so it is clear that the topology generated by the family of semi-norms is not the compact-open topology.

EXAMPLE 4. Let Ω be the set of ordinal numbers smaller than the first uncountable ordinal, with the order topology. With norm p defined on A by

$$p(x) = \sup_{x \in \Omega} \{|f(x)|\} ,$$

A is a Banach algebra. However, A with the compact-open topology is not an F-algebra, [4, Example 3.7, p. 19], so it is clear that the norm topology on A is not the compact-open topology.

Note added in proof. We are indebted to Dr. R. T. Moore for pointing out, in connection with Example 2, that the following result can be deduced from Theorem 3 by the Open Mapping Theorem.

THEOREM. *A unital F-algebra A is normable if and only if there exists an $(A, \{p_\alpha\})$ such that $V(A, \{p_\alpha\}; a)$ is bounded for all $a \in A$.*

REFERENCES

1. C. Apostol, *b*-algebras and their representation*, J. London Math. Soc., 2nd Series, **3** (1971), 30–38.
2. F. F. Bonsall and J. Duncan, *Numerical Range of Operators on Normed Spaces and of Elements of Normed Algebras*, London Math. Soc. Lecture Note Series 2, Cambridge, 1971.
3. R. M. Brooks, *On locally m-convex *-algebras*, Pacific J. Math, **23** (1967), 5–23.
4. E. A. Michael, *Locally multiplicatively convex topological algebras*, Amer. Math. Soc. Mem., **11** (1952).
5. B. Sims, *A characterization of Banach-star-algebras by numerical range*, Bull. Austral. Math. Soc., **4** (1971), 193–200.

Received August 21, 1972 and in revised form September 28, 1972. The work of the first author was partially supported under a visiting postdoctoral fellowship at McMaster University, Ontario. The work of the second author was supported under a visiting senior lectureship at the University of Newcastle, N.S.W.

UNIVERSITY OF NEWCASTLE
AND
MIAMI UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

ON DOMINANT AND CODOMINANT DIMENSION OF $QF-3$ RINGS

DAVID A. HILL

In this paper the concept of codominant dimension is defined and studied for modules over a ring. When the ring R is artinian, a left R module M has codominant dimension at least n in case there exists a projective resolution

$$P_n \longrightarrow P_{n-1} \longrightarrow \cdots \longrightarrow P_1 \longrightarrow M \longrightarrow 0$$

with P_i injective. It is proved that every left R-module has the above property if and only if R has dominant dimension at least n. The concept of codominant dimension is also used to study semi-perfect $QF-3$ rings.

Let R be an associative ring with an identity 1. Denote by ${}_{\Re}R$ (resp. $R_{\Re}$) the left (resp. right) R-module R. Using the terminology of [5], we have the following definitions:

(1) R is left $QF-3$, if R has a faithful projective injective left ideal.

(2) R is left $QF-3^+$ if the injective hull $E({}_{\Re}R)$ is projective.

(3) R is left $QF-3'$ if $E({}_{\Re}R)$ is torsionless, i.e., there exists a set A such that $E(R) \leqq \prod_A R$.

In general (1) $\Rightarrow$ (3). For perfect rings the three conditions are equivalent for left and right $QF-3$ rings. (See [5].)

The dominant dimension of a left (resp. right) R-module M, denoted by dom. dim $({}_{\Re}M)$ (resp. dom. dim $(M_{\Re})$) is at least n, if there exists an exact sequence

$$0 \longrightarrow M \longrightarrow X_1 \longrightarrow \cdots \longrightarrow X_n$$

of left (resp. right) R-module where each X_i is torsionless and injective for $i = 1, \cdots, n$. See [3] for details.

Note that this says when dom. dim $({}_{\Re}R) \geqq 1$ and R is left-artinian that $E(Re_i)$ for $i = 1, \cdots, n$ is projective where $\{e_i\}$, $i = 1, \cdots, n$ is a complete set of orthogonal idempotents, and that each X_i is projective.

We define codominant dimension as follows:

Let M be a left R-module. The codom. dim of M is at least n in case there exists an exact sequence

$$P_n \longrightarrow P_{n-1} \longrightarrow \cdots \longrightarrow P_1 \longrightarrow M \longrightarrow 0$$

where P_i is torsionless and injective for $i = 1, \cdots, n$.

Following the notation of [3], we say that if such an exact sequence exists for $1 \leqq i \leqq n$, but no such sequence exists for $1 \leqq i \leqq n+1$, then codom. dim $({}_{\mathfrak{R}}M) = n$. If such a sequence exists for all n then codom. dim $({}_{\mathfrak{R}}M) = \infty$. If no such sequence exists codom. dim $({}_{\mathfrak{R}}M) = 0$.

An R-module U is defined to be a cogenerator if for any module M we can embed it in a product of copies of U. We have:

LEMMA. *Let U, V be left injective cogenerators then the* codom. dim (U) = codom. dim (V).

The proof follows easily from properties of injective cogenerators and shall omit it.

Let U be a left injective cogenerator. If the codom. dim $(U) = n$, we say that R has l. codom. dim $({}_{\mathfrak{R}}R) = n$. In a similar manner one defines r. codom. dim $(R_{\mathfrak{R}})$. Note that if ${}_{\mathfrak{R}}R$ is artinian, products of projectives are projective and direct sums of injectives are injective. Hence l. codom. dim $({}_{\mathfrak{R}}R) = n$ is equivalent to the existence of a resolution

$$P_n \longrightarrow P_{n-1} \longrightarrow \cdots \longrightarrow P_1 \longrightarrow U \longrightarrow 0$$

where P_i is projective and injective and $U = E(S_1) \oplus \cdots \oplus E(S_n)$ where S_i: $i = 1, \cdots, n$ is a copy of each simple left R-module.

In §1 we characterize semi-perfect $QF-3^+$ rings in terms of their finitely generated projective, injectives.

In §2 we show that l. dom. dim $({}_{\mathfrak{R}}R)$ and l. codom. dim $({}_{\mathfrak{R}}R)$ are the same for artinian rings. Hence, if R is artinian $QF-3$ then the l.-dom. dim (r-dom. dim) l. codom. dim (r-codom. dim) are the same.

For notation we use J to donote the Jacobson radical, and $R^{(A)}(R^A)$ denotes a direct sum (resp. direct product) of A-copies of R. Also $E(M)$ will be used to denote the injective hull of an R-module M and $P(M)$ will denote the projective cover of M when M has a projective cover. For a left R-module M, we let $\ell_{\mathfrak{R}}(M) = \{x \in R \mid x \cdot M = 0\}$, and $r_{\mathfrak{M}}(I) = \{x \in M \mid I \cdot x = 0\}$ where $I \subseteqq R$. We will use $T(M)$ to denote $M/J(M)$ where $J(M)$ is the Jacobson radical of M.

1. $QF-3$ **Rings.** Recall that if ${}_{\mathfrak{R}}R$ is noetherian $rt \cdot QF - 3 \Leftrightarrow rt \cdot QF - 3^+$. (See [1] and [6].)

To begin with we shall prove that under those hypotheses

$$rt \cdot QF - 3^+ \Longleftrightarrow rt \cdot QF - 3' .$$

PROPOSITION 1.1. *Let ${}_{\mathfrak{R}}R$ be noetherian. If $E(R_{\mathfrak{R}})$ is torsionless then $E(R_{\mathfrak{R}})$ is projective.*

Proof. Given that $0 \to E \xrightarrow{\theta} R^A$ is monic, where A is an indexing set. We show that there exists a finite number of R_α's, $\alpha \in A$ say $R_{\alpha_i}, \cdots, R_{\alpha_m}$ such that $\pi\theta|_R = \tilde{\theta}$ where π is the projection $R^A \to \bigoplus \sum_{i=1}^m R_{\alpha_i}$ is monic. Let S be the set of all finite intersections of right ideals $\{K_\alpha\}_{\alpha \in A}$ where $K_\alpha = \ker(\pi_\alpha \circ \theta|_R)$. Note that $\bigcap_{i=1}^n K_{\alpha_i}$ induces a natural embedding of

$$0 \longrightarrow R\Big/\bigcap_{i=1}^n K_{\alpha_i} \longrightarrow R^{(n)} .$$

Thus $R/\bigcap_{i=1}^n K_{\alpha_i}$ is torsionless. Hence by [2, Thm. I, p. 350]

$$\bigcap_{i=1}^n K_{\alpha_i} = r_{\mathfrak{R}}\ell_{\mathfrak{R}}\left(\bigcap_{i=1}^n K_{\alpha_i}\right) .$$

Now since ${}_{\mathfrak{R}}R$ noetherian, the set $\{\ell_{\mathfrak{R}}(\bigcap_{i=1}^n K_{\alpha_i})\}$ has a maximal element $\ell_{\mathfrak{R}}(\bigcap_{i=1}^m K_{\alpha_i})$ where $\bigcap_{i=1}^n K_{\alpha_i} \in S$. Thus $r_{\mathfrak{R}}\ell_{\mathfrak{R}}(\bigcap_{i=1}^m K_{\alpha_i}) = \bigcap_{i=1}^m K_{\alpha_i}$ is a minimal right ideal in S. But then $x \in \bigcap_{i=1}^m K_{\alpha_i} \Rightarrow x \in \bigcap_{\alpha \in A} K_\alpha$. Thus $\bigcap_{i=1}^m K_{\alpha_i} = 0$. This implies that $\tilde{\theta}$ is monic. But then $\pi\theta$ is monic since $\ker(\pi\theta) \cap R \neq 0$ if $\ker(\pi\theta) \neq 0$. This shows E is projective.

We next show that $QF-3^+ \Rightarrow QF-3$ for semi-perfect rings.

First we need the following lemma.

LEMMA 1.2. *Let K be finitely generated. Suppose there exists an exact sequence*

$$0 \longrightarrow K \longrightarrow E_1 \longrightarrow \cdots \longrightarrow E_n$$

where $E(K) = E_1$, $E_{i+1} = E(E_i)$ for $1 \leqq i \leqq n-1$ and each E_i is projective. Then $E_1, \cdots, E_n$ are all finitely generated.

Proof. This follows easily from the proof of [4, Lemma 1].

PROPOSITION 1.3. *Suppose R is semi-perfect. If R is left $QF-3^+$ then R is left $QF-3$.*

Proof. By Lemma 1.2 $E(R)$ is finitely generated. Since R is semi-perfect $E(R) \cong \bigoplus \sum_{i=1}^n Re_i$, where each e_i is an indecomposable idempotent.

Let $Re_1, \cdots, Re_k$ be a subset of $Re_1, \cdots, Re_n$, where the set $\{Re_1, \cdots, Re_k\}$ is a complete set of isomorphism classes of $\{Re_1, \cdots, Re_n\}$. Then $U = Re_1 \oplus \cdots \oplus Re_k$ is a minimal projective injective.

Now we come to the main theorem of this section.

THEOREM 1.4. *Let R be semi-perfect. The following are equivalent*:

(a) *R is left QF* $- 3^+$.

(b) $E({}_{\Re}R)$ *is finitely generated and every finitely generated left injective has an injective projective cover.*

(c) *Every finitely generated left projective has a projective injective hull.*

Proof. (b) $\Rightarrow$ (a): Consider

$$P(E(R)) \longrightarrow E(R) \longrightarrow 0 .$$

Embed $R \xrightarrow{i_{\Re}} E(R)$ then by the projectivity of R there exists a map $\theta'\colon R \to P(E(R))$ such that θ' is monic.

Consider the following diagram:

$$\begin{tikzcd} 0 \arrow[r] & R \arrow[r, "i_{\Re}"] \arrow[d, "\theta'"'] & E(R) \arrow[dl, dashed, "\theta''"] \\ & P(E(R)) . & \end{tikzcd}$$

Here $\theta''(r) = \theta'(r)$ for all $r \in R$. Also θ'' is monic. The injectivity of $E(R)$ forces $E(R)$ to be a direct summand of $P(E(R))$, hence projective.

(a) $\Leftrightarrow$ (c): Consider $R^{(n)}$, $R^{(n)} \leqq E(R)^{(n)}$. Thus $E(P) \leqq E(R)^n$, where $P \oplus P' = R^{(n)}$, as a direct summand. Hence $E(P)$ is projective. The converse is trivial.

(a) $\Rightarrow$ (b): By Lemma 1.2 $E(R)$ is finitely generated.

Consider $P(E) \xrightarrow{\theta} E \to 0$ where $P(E)$ is finitely generated injective. Let $R^{(n)} \xrightarrow{\rho} E \to 0$. Combining the above maps we have the following diagrams:

$$\begin{tikzcd} 0 \arrow[r] & R^{(n)} \arrow[r, "i_{\Re}^{(n)}"] \arrow[d, "\rho"'] & E(R)^{(n)} \arrow[dl, dashed, "\rho'"] \\ & E . & \end{tikzcd}$$

So we have ρ' epic and $\rho' \circ i_{\Re}^{(n)} = \rho$. Further we have

$$\begin{tikzcd} & E(R)^{(n)} \arrow[dl, dashed, "\rho''"'] \arrow[d, "\rho'"] & \\ P(E) \arrow[r, "\theta"] & E \arrow[r] & 0 \end{tikzcd}$$

Noting that ρ'' is epic and $P(E)$ is projective, $P(E)$ is a direct summand of $E(R)^{(n)}$. Hence injective.

A ring is perfect in case every module has a projective cover. We show that $QF - 3^+$ rings can be characterized in terms of the projective cover of $E({}_{\mathfrak{R}}R)$.

THEOREM 1.5. *Let R be perfect. Then every indecomposable summand of $P(E({}_{\mathfrak{R}}R))$ is injective if and only if R is left $QF-3^{+}$.*

Proof. $\Rightarrow$ Consider the following diagram:

$$\begin{array}{ccccc} & & {}_{\mathfrak{R}}R & & \\ & \swarrow^{f} & \downarrow i & & \\ P(E({}_{\mathfrak{R}}R)) & \xrightarrow{\pi} & E({}_{\mathfrak{R}}R) & \longrightarrow & 0\,. \end{array}$$

Here i is a monomorphism and π is epic. Since R is projective there exists on f such that $\pi f = i$. Clearly f is monic. Since R is perfect $P(E({}_{\mathfrak{R}}R)) \cong \sum_{\alpha \in A} Re_{\alpha}$, where e_{α} are primitive idempotents of R. Now $\operatorname{Im}(f)$ is contained in $\sum_{\alpha=1}^{n} Re_{\alpha}$, for n a positive integer, since ${}_{\mathfrak{R}}R$ is cyclic.Thus using the hypothesis, $E({}_{\mathfrak{R}}R)$ is projective and R is left $QF-3^{+}$. $\Leftarrow$ This is trivial.

2. Codominant dimension of rings. We begin with a lemma which holds the key to the main results of this section.

LEMMA 2.1. *Let R be a ring. The following conditions are equivalent.*

(1) *For every projective left R-module P, there exists an exact sequence*

$$0 \longrightarrow P \longrightarrow E_1 \longrightarrow \cdots \longrightarrow E_n$$

where E_i, $1 \leqq i \leqq n$, are injective and projective.

(2) *For every injective left R-module Q, there exists an exact sequence*

$$P_n \longrightarrow P_{n-1} \longrightarrow \cdots \longrightarrow P_1 \longrightarrow Q \longrightarrow 0$$

where P_i, $1 \leqq i \leqq n$, are injective and projective.

Proof. (1) $\Rightarrow$ (2). For $n = 1$ a modification for the proof of Theorem 1.4 will suffice. We assume the lemma is true for the nth case and prove the $n+1$ case. So consider the following exact sequences.

(1) $$0 \longrightarrow P_{n+1} \xrightarrow{J_1} E_1 \xrightarrow{J_2} E_2 \longrightarrow \cdots \xrightarrow{J_{n+1}} E_{n+1}$$

(2) $$P_{n+1} \xrightarrow{\theta_1} P_n \xrightarrow{i_n} \cdots \longrightarrow P_1 \xrightarrow{i_1} Q \longrightarrow 0\,.$$

Here Q is an arbitrary injective module and

$$P_1, \cdots, P_n, E_1, \cdots, E_{n+1}$$

are both projective and injective and P_{n+1} is projective.

Also E_k is the injective hull of Cok (J_k).

Denote by K the image of θ_1. Using the injectivity of P_n, there is a map θ_2: $E_1 \to P_n$ such $\theta_2 J_1 = i_{n+1}\theta_1$ where i_{n+1} is the embedding of K into P_n. The injectivity of P_{n-1} and the exact sequence $0 \to E_1/P_{n+1} \to E_2$ induce a map θ_3: $E_2 \to P_{n-1}$ which one can easily check has the property $\theta_3 J_2 = i_n \theta_2$.

In like manner we can define θ_k: $E_{k-1} \to P_{n+2-k}$ such that

$$\theta_k J_{k-1} = i_{n+3-k}\theta_{k-1}\ , \quad k = 2, \cdots, n+2\ .$$

This information is summed up in the following diagram:

$$\begin{tikzcd}
0 \arrow[r] & P_{n+1} \arrow[r] \arrow[d, "\theta_1"] & E_1 \arrow[r, "J_2"] \arrow[d, "\theta_2"] \arrow[dl, dashed, "h_{n+1}"'] & E_2 \arrow[d, "\theta_3"] \arrow[dl, dashed, "h_n"'] & \cdots E_{n+1} \arrow[r, "J_{n+1}"] \arrow[d, "\theta_{n+1}"] & E_{n+1} \arrow[d, "\theta_{n+2}"] \arrow[dl, dashed, "h_1"'] & \\
0 \arrow[r] & K \arrow[r, "i_{n+1}"] \arrow[d] & P_n \arrow[r, "i_n"] & P_{n-1} \cdots & P_1 \arrow[r, "i_1"] & Q \arrow[r] & 0\ . \\
& 0 & & & & &
\end{tikzcd}$$

Having constructed θ_{n+2}, the projectivity of E_{n+1} induces a map h_1: $E_{n+1} \to P_1$ such $i_1 h_1 = \theta_{n+2}$. Now consider the map $h_1 J_{n+1} - \theta_{n+1}$: $E_n \to P_1$. We have $i_1(h_1 J_{n+1} - \theta_{n+1}) = \theta_{n+2} J_{n+1} - i_1\theta_{n+1} = 0$. So Im $(h_1 J_{n+1} - \theta_{n+1}) \leqq$ ker (i_1).

Now consider the following diagram:

$$\begin{tikzcd}
& E_n \arrow[dl, dashed] \arrow[d, "h_1 J_{n+1} - \theta_{n+1}"] & \\
P_2 \arrow[r] & \mathrm{Im}\,(i_2) \arrow[r] & 0\ .
\end{tikzcd}$$

We can construct h_2 using the projectivity of E_n. By a similar argument we can show that Im $(h_2 J_n - \theta_n) \leqq$ ker (i_2). By a recursive argument we can construct $h_k J_{n+2-k} - \theta_{n+2-k}$ for $k = 1, \cdots, n$ in like manner. In particular we have $h_n J_2 - \theta_2$: $E_1 \to P_n$ where Im $(h_n J_2 - \theta_2) \leqq K$. We need only show equality to complete the proof. Let $k \in K$. Then there exists an $x \in P_{n+1}$ such that $\theta_1(x) = k$. Thus $(h_n J_2 - \theta_2)(J_1(-x)) = \theta_2 J_1(x) = \theta_1(x) = k$. Thus $h_n J_2 - \theta_2$ maps on to K. The proof (2) $\Rightarrow$ (1) is similar. This completes the proof.

Noting that for left artinian rings products of projectives are projective, and direct sums of injectives are injective one can easily show that dom. dim $(R) \geqq n$ implies dom. dim. $(P) \geqq n$ for all projective P.

Likewise letting $I = \bigoplus \sum E_\alpha(S_\alpha)$ be the minimal injective cogenerator of R, we find that codom. dim$(I) \geqq n$ implies codom. dim $(Q) \geqq n$ for all injectives Q. Thus we have:

THEOREM 2.2. *Let R be left artinian then the following are equivalent*:

(1) *The* $\inf\{m \in Z \mid$ dom. dim $(P) = m$ *for all P projectives*$\} = n$.

(2) *The* $\inf\{m \in Z \mid$ dom. dim $(Q) = m$ *for all Q injectives*$\} = n$.

(3) l. dom. dim $({}_{\mathfrak{R}}R) = n$.

(4) l. codom. dim $({}_{\mathfrak{R}}R) = n$.

If no such n exists we say l. dom. dim $(R) = \infty$

Proof. (3) $\Rightarrow$ (1), (4) $\Rightarrow$ (2) by our previous discussion. (1) $\Rightarrow$ (3): There exists a projective module P such dom. dim $(P) = n$.

Now $P \cong \bigoplus \sum_{\mathscr{A}} Re_\alpha$, $\{e_\alpha\}$ primitive idempotents such that for some e_β dom. dim $(Re_\beta) < n+1$ where $e_\beta \in \{e_\alpha\}$. Since $Re_\beta < R$, $n+1 >$ dom. dim $(R) \geqq n$. This yields the desired result. (2) $\Rightarrow$ (4) is similar. (1) $\Rightarrow$ (2): By Lemma 2.1 $\inf\{m \in Z \mid$ codom. dim $(Q) = m\} \geqq n$. If inf of the above set is strictly greater than n, another application of the lemma forces $\inf\{m \in Z \mid m =$ dom. dim (P), P projective$\} > n$ which is impossible. (2) $\Rightarrow$ (1) is similar.

Let R be left artinian and both left and right $QF-3$. Then by [4, Thm. 10] l. dom. dim $({}_{\mathfrak{R}}R) =$ r. dom. dim $(R_{\mathfrak{R}})$. Thus in view of 2.2 we have:

PROPOSITION 2.3. *Let ${}_{\mathfrak{R}}R$ be artinian and $QF-3$. Then* l. domdim $({}_{\mathfrak{R}}R) =$ r. domdin $(R_{\mathfrak{R}}) =$ l. codomdin $({}_{\mathfrak{R}}R) =$ r. codomdim $(R_{\mathfrak{R}}) = n$.

Acknowledgement. The author wishes to thank the referee for his proof to Theorem 1.5 which is simpler than the author's original version.

REFERENCES

1. J. P. Jans, *Projective injective modules*, Pacific J. Math., **9** (1959), 1103–1108.
2. T. Kato, *Duality of cyclic modules*, Tohoku Math. J., **14** (1967), 349–356.
3. ———, *Rings of dominant dimension* $\geqq 1$, Proc. Japan Acad., **44** (1968), 579–584.
4. B. J. Muller, *Dominant dimension of semi-primary rings*, J. reine angew. Math., **232** (1968), 173–179.
5. H. Tachikawa, *On left QF—3 rings*, Pacific J. Math., **31** (1970), 255–268.
6. ———, *Lectures on QF—3 and QF—1 Rings*, Carleton Mathematical Lecture Notes No. 1, July, 1972.

Received February 8, 1972 and in revised form Junuary 3, 1973.

UNIVERSITY OF WESTERN AUSTRALIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

ROUND AND PFISTER FORMS OVER $R(t)$

J. S. HSIA AND ROBERT P. JOHNSON

An anisotropic quadratic form ϕ is called round if $\phi \cong a\phi$ whenever ϕ represents $a \neq 0$. All round forms over $R(t)$ are completely determined. Connections with Pfister's strongly multiplicative forms and with the reduced algebraic K-theory groups k_n of Milnor are studied.

The concept of a round form was introduced by Witt (see [5] and [8]) to give new simple proofs of results of Pfister on the structure of the Witt ring over fields. In a previous paper [3] we determined all round forms over a global field. In this paper we completely determine all round forms over $\boldsymbol{R}(t)$, the field of rational functions in one variable over the reals.

We now describe our main results.

Let ϕ be an anisotropic form of dimension > 1 over $\boldsymbol{R}(t)$. Then ϕ is round if and only if $\phi \cong (n \times (1, f)) \oplus (1, fg)$ for some $f, g \in \boldsymbol{R}(t)$ such that f is a product of distinct linear factors and g is a product of irreducible quadratic factors. Our proof gives a method of computing f and g, which are essentially unique (see 2.5 and 2.6). We study a generalization of a round form, called a group form, over $\boldsymbol{R}(t)$ and measure how far group forms are from being round (see [3] for group forms over global fields).

In the last section we show that a form of dimension $2^n (n \geqq 2)$ is a Pfister form if and only if it is a round form of determinant one. Such a form can be written uniquely as $2^{n-1} \times (1, f)$ for some $f \in \boldsymbol{R}[t]$ which is $\pm$ a product of distinct monic linear factors. From this and a theorem of Elman and Lam we see that every element of $k_n\boldsymbol{R}(t)$ can be written uniquely as $l(-1)^{n-1}l(-f)$ with f as above.

1. Preliminaries. We will consider only quadratic forms (often simply called "forms") over a field F of characteristic $\neq 2$. We write $\phi \oplus \psi$ for the orthogonal sum and $\phi \otimes \psi$ for the tensor product of quadratic forms [5, p. 8]. We call ϕ *hyperbolic* if $\phi \cong m \times (1, -1)$, i.e., ϕ is a direct sum of hyperbolic planes.

Define $\dot{D}\phi = \{a \in \dot{F} \mid \phi \text{ represents } a\}$ and $G\phi = \{a \in \dot{F} \mid a\phi \cong \phi\}$ where $\dot{F} = F - \{0\}$. An anisotropic form ϕ is called *round* if and only if $\dot{D}\phi = G\phi$ (or equivalently $\dot{D}\phi \subseteqq G\phi$); an isotropic form is called round if and only if it is hyperbolic [5, p. 22]. A form ϕ is called a *Pfister form* if $\phi \cong (1, a_1) \otimes \cdots \otimes (1, a_n) (a_i \in \dot{F})$.

We will frequently refer to [4] for results on quadratic forms over $F = \boldsymbol{R}(t)$. The valuations of F which are trivial on $\boldsymbol{R}$ are of three types: if the prime element is $t-\alpha(\alpha\in\boldsymbol{R})$, the valuation is called *real*; if the prime element is an irreducible quadratic polynomial it is called *complex*; if the prime element is t^{-1} it is called *infinite*. A *spot* is an equivalence class of valuations [7]. If p is a real or infinite spot then the completion F_p of F at p is isomorphic to $\boldsymbol{R}((\pi))$ (a *real series field*) where π is a prime element. If p is complex, $F_p\cong \boldsymbol{C}((\pi))$ is called a *complex series field*. See [4] for results on quadratic forms over series fields.

If ϕ is a quadratic form over $\boldsymbol{R}(t)$ and if $\alpha\in\boldsymbol{R}$, we define "ϕ at α" to be the quadratic form over $\boldsymbol{R}$ obtained by replacing t by α in the matrix of ϕ. Thus ϕ at α is well-defined for almost all $\alpha\in\boldsymbol{R}$. The following result is Proposition 2.1 of [4] and is due to Witt.

1.1. *A nonsingular quadratic form of dimension $\geqq 3$ over $\boldsymbol{R}(t)$ is isotropic if and only if for almost all $\alpha\in\boldsymbol{R}$, the form at α is isotropic over $\boldsymbol{R}$. Thus if ϕ is a quadratic form of dimension $\geqq 2$ over $\boldsymbol{R}(t)$ and if $0\neq f(t)\in\boldsymbol{R}(t)$, then ϕ represents $f(t)\Leftrightarrow$ for almost all $\alpha\in\boldsymbol{R}$, ϕ at α represents $f(\alpha)$.*

If we write $\phi\cong(a_1,\cdots,a_n)$ over a field F then $\det\phi=a_1\cdots a_n$ modulo $\dot{F}^2$. When $F=\boldsymbol{R}(t)$ we assume $\det\phi$ is written as $\pm$ a product of distinct monic irreducible polynomials.

The following result generalizes Proposition 2.2 of [4].

1.2. *Let ϕ,ψ be quadratic forms over $\boldsymbol{R}(t)$. If $\phi\cong\psi$ at α for almost all $\alpha\in\boldsymbol{R}$ and if $\det\phi$, $\det\psi$ have the same irreducible quadratic factors, then $\phi\cong\psi$.*

Proof. Clear for $\dim\phi=1$. We assume this result is true whenever $\dim\phi<n$ and prove it for $\dim\phi=n>1$. Let ϕ represent $a\neq 0$. Then $\phi\oplus(-a)$ is isotropic so by 1.1, $\psi\oplus(-a)$ is isotropic. Thus ψ represents a. Write $\phi\cong(a)\oplus\phi_1$ and $\psi\cong(a)\oplus\psi_1$ and apply the induction hypothesis.

1.3. *Let $f(t)\in\boldsymbol{R}[t]$ and $\alpha\in\boldsymbol{R}$ with $f(\alpha)\neq 0$. Then $(f(t))\cong(f(\alpha))$ (one-dimensional quadratic forms) over the completion of $\boldsymbol{R}(t)$ at the spot with prime element $t-\alpha$.*

Proof. Write $f(t)=a_0+a_1(t-\alpha)+\cdots+a_n(t-\alpha)^n$ and apply the Local Square Theorem [7, 63: 1a], noting $f(\alpha)=a_0$.

2. Round forms over $\boldsymbol{R}(t)$.

We will need the following result, which determines all round forms over a series field.

2.1. *Let ϕ be an anisotropic quadratic form over a real or complex series field F.*

(a) *If F is complex, then ϕ is round $\Leftrightarrow \phi$ represents* 1.

(b) *Let F be a real series field. Then F is pythagorean and formally real. So if* $\dim \phi$ *is odd, ϕ is round $\Leftrightarrow \phi \cong (1, \cdots, 1)$. If* $\dim \phi = 2m$ *is even then ϕ is round $\Leftrightarrow \phi \cong m \times (1, 1)$ or $m \times (1, \pm \pi)$.*

Proof. (a) By [4, 1.2], $\dim \phi \leqq 2$ whenever ϕ is anisotropic over a complex series field. Now apply [5, 2.4].

(b) It follows easily from the Local Square Theorem [7, 63: 1a] that F is pythagorean. Now apply [5, 2.4] and [4, 1.6].

Now let F be a field of characteristic $\neq 2$ and let Ω be a set of discrete or archimedian spots on F (see [7] for terminology). We say that (F, Ω) satisfies the *Weak Hasse-Minkowski Theorem* if whenever σ and τ are quadratic forms over F with $\sigma_p \cong \tau_p$ for all $p \in \Omega$, then $\sigma \cong \tau$ (σ_p denotes the form σ viewed over the completion F_p of F at p).

2.2. *Let (F, Ω) satisfy the Weak Hasse-Minkowski Theorem. Let ϕ be anisotropic over F. Then ϕ is round $\Leftrightarrow$ for all $p \in \Omega$,*

(1) *ϕ_p is round*

or (2) *ϕ_p is isotropic and ϕ'_p (the anisotropic part of ϕ_p) is round and universal.*

Proof. ($\Rightarrow$): Assume ϕ is round. Let $p \in \Omega$. We first assume ϕ_p is anisotropic and show ϕ_p is round. Let $b \in \dot{D}(\phi_p)$. Approximate b by $a \in \dot{D}\phi$. By the Local Square Theorem, we can obtain $a \in b\dot{F}_p^2$. Thus $\phi \cong a\phi \Rightarrow \phi_p \cong b\phi_p$ so ϕ_p is round.

Now assume ϕ_p is isotropic and not hyperbolic. Write $\phi_p = \phi'_p \oplus H$ with H hyperbolic. We will show $\phi'_p \cong b\phi'_p$ for all $b \in \dot{F}_p$ and so (2) holds. Now ϕ_p represents b so we find that $\phi_p \cong b\phi_p$ by the argument of the preceding paragraph. Thus $\phi'_p \cong b\phi'_p$.

($\Leftarrow$): Let $a \in \dot{D}\phi$. Applying (1) or (2), we have $\phi_p \cong a\phi_p$ for all $p \in \Omega$. By the Weak Hasse-Minkowski Theorem, $\phi \cong a\phi$, so ϕ is round.

EXAMPLES 2.3. The Weak Hasse-Minkowski Theorem holds in the following cases:

(1) Let $F = K(t)$ where K is an arbitrary field of characteristic $\neq 2$ and let Ω be the set of all spots on F that are trivial on K. Using [6, Theorem 5.3] one can show that (F, Ω) satisfies the Weak Hasse-Minkowski Theorem.

(2) Let F be a global field and let Ω be the set of all non-trivial spots on F. We have the following precise results in this case [3, 2.4]: let ϕ be an anisotropic form over F and let $\dim \phi > 2$. Then ϕ is round if and only if: (1) $\dim \phi \equiv 0 \bmod 4$, (2) at all real

spots (if there are any) ϕ is hyperbolic or positive definite, and (3) $\det \phi = 1$. We note that the *Strong Hasse-Minkowski Theorem* holds for (F, Ω), i.e., if a form ϕ is isotropic for all $p \in \Omega$ then ϕ is isotropic.

(3) Cassels, Ellison, and Pfister (J. Number Theory, 3 (1971), p. 147) have recently shown that the Strong Hasse-Minkowski Theorem fails for $F = K(t)$ where $K = \boldsymbol{R}(x)$ (x, t independent indeterminants over $\boldsymbol{R}$) though the weak theorem holds as we have mentioned in (1).

The next two results determine all round forms over $\boldsymbol{R}(t)$.

2.4. *There is no odd-dimensional round form over* $\boldsymbol{R}(t)$ *except the form* $\phi = (1)$.

Proof. Note that $\boldsymbol{R}(t)$ is non-pythogorean since $t^2 + 1$ is not a square. Now apply [5, 2.4].

THEOREM 2.5. *Let* ϕ *be an anisotropic form of dimension* $2m$ *over* $\boldsymbol{R}(t)$. *Then the following are equivalent*:

(1) ϕ *is round.*

(2) $\phi \cong ((m-1) \times (1, f)) \oplus (1, fg)$ *for some* $f, g \in \boldsymbol{R}[t]$ *such that* f *is a product of distinct linear factors and* f *or* $-f$ *is monic, and* g *is a product of monic irreducible quadratic factors* (*we allow* $f = 1$ *or* -1 *and allow* $g = 1$).

(3) *For almost all* $\alpha \in \boldsymbol{R}$, ϕ *at* α *is hyperbolic or positive definite.*

(4) ϕ_p *is round for all real or infinite spots* p *on* $\boldsymbol{R}(t)$.

Proof. (1) $\Leftrightarrow$ (4) follows from 2.2 since there is no universal anisotropic form over a real series field. We will show (2) $\Rightarrow$ (4) $\Rightarrow$ (3) $\Rightarrow$ (2). (2) $\Rightarrow$ (4) follows from 2.1 and 1.3.

(4) $\Rightarrow$ (3): Assume (4). Write $\phi \cong (f_1(t), \cdots, f_{2m}(t))$ with the $f_i(t) \in \boldsymbol{R}[t]$. Let $\alpha \in \boldsymbol{R}$ such that $f_i(\alpha) \neq 0$ for all i. Let p be the real spot with prime element $t - \alpha$. By 1.3, $\phi_p \cong (f_1(\alpha), \cdots, f_{2m}(\alpha))$. By 2.1, $\phi_p \cong m \times (1, 1)$ or $m \times (1, -1)$. So by [4, 1.6], ϕ at α is $\cong m \times (1, 1)$ or $m \times (1, -1)$.

(3) $\Rightarrow$ (2): Write $\phi \cong (f_1, \cdots, f_{2m})$ with the $f_i \in \boldsymbol{R}[t]$. Let S be the set of all $a \in \boldsymbol{R}$ such that $f_i(a) = 0$ for some i. Write $S = \{a_1, \cdots, a_k\}$ with $a_1 < a_2 < \cdots < a_k$. If I is any of the intervals $(-\infty, a_1)$, (a_1, a_2), $\cdots$, (a_k, ∞) then ϕ at α is hyperbolic for all $\alpha \in I$ or is positive definite for all $\alpha \in I$. The idea now is to merge together adjacent intervals if ϕ at α looks the same in the adjacent intervals. If ϕ at α is positive definite (respectively, hyperbolic) for almost all $\alpha \in \boldsymbol{R}$ then we let $f = 1$ (respectively, -1). Otherwise, there is an ordered subset $\{b_1 < b_2 < \cdots < b_j\}$ of S such that if J is any of the intervals $(-\infty, b_1)$, (b_1, b_2), $\cdots$, (b_j, ∞) then ϕ at α is hyperbolic for almost all

$\alpha \in J$ or is positive definite for almost all $\alpha \in J$, and such that whenever ϕ is hyperbolic in one of these intervals then it is positive definite in the adjacent intervals. Now let $f = (t - b_1) \cdots (t - b_j)$ if ϕ at α is positive definite for almost all $\alpha > b_j$, and let $f = -(t - b_1) \cdots (t - b_j)$ otherwise. Let g be the product of all the (monic) irreducible quadratic factors of det ϕ. Then by 1.2, $\phi \cong ((m - 1) \times (1, f)) \oplus (1, fg)$.

REMARK 2.6. (1). Part (2) of the above theorem gives us a *canonical form* for an anisotropic round form of even dimension over $\boldsymbol{R}(t)$, i.e., f and g are uniquely determined. This fact follows easily from 1.2. The proof of (3) $\Rightarrow$ (2) gives us a constructive method of finding f and g (provided we know the decomposition of the f_i into irreducible factors).

(2) Part (3) of the theorem provides us with the easiest way to check whether a given anisotropic form ϕ of even dimension over $\boldsymbol{R}(t)$ is round. If $\phi \cong (f_1, \cdots, f_{2m})$ with the $f_i \in \boldsymbol{R}(t)$ and if $\{a_1 < a_2 < \cdots < a_k\}$ is the ordered set of all real roots of the f_i's, we need only compute ϕ at α for one value of α in each of the intervals $(-\infty, a_1)$, (a_1, a_2), $\cdots$, (a_k, ∞).

As in [3], we call a quadratic form ϕ over a field F a *group form* if $\dot{D}\phi$ is a subgroup of $\dot{F}$. Every round form is clearly a group form. We now briefly investigate group forms over $\boldsymbol{R}(t)$.

2.7. *Let F be a field with a set Ω of discrete or archimedean spots on F. Assume (F, Ω) satisfies the Strong Hasse-Minkowski Theorem (local isotropy implies isotropy). Then a quadratic form ϕ over F is a group form $\Leftrightarrow \phi_p$ is a group form for all $p \in \Omega$.*

Proof. ($\Rightarrow$): See the proof of 3.2 of [3]. ($\Leftarrow$): Let $a, b \in \dot{D}\phi$. Then $ab \in \dot{D}\phi_p$ for all $p \in \Omega$ so $ab \in \dot{D}\phi$.

By [4, 2.3] and [7, 42:11], $\boldsymbol{R}(t)$ satisfies the Strong Hasse-Minkowski Theorem with respect to the set of all real and complex spots. Thus by 2.7 and 1.1, we have:

2.8. *Let ϕ be a quadratic form over $\boldsymbol{R}(t)$. Then ϕ is a group form $\Leftrightarrow \phi$ represents 1. If $\dim \phi \geqq 2$ then ϕ is a group form $\Leftrightarrow \phi$ at α represents 1 for almost all $\alpha \in \boldsymbol{R}$.*

If ϕ is an anisotropic group form over any field then ϕ is round $\Leftrightarrow$ the factor group $\dot{D}\phi / G\phi = 1$. Thus this factor group measures how far an anisotropic group form is from being round. We now investigate this factor group.

2.9. *Let ϕ be a group form over $\boldsymbol{R}(t)$ and assume ϕ is not round.*

Then $\dot{D}\phi/G\phi$ is infinite unless $\phi \cong (m \times (1, -1)) \oplus (1, -g)$ where $m \geqq 1$ and g is a product of monic irreducible quadratic factors. In this latter case $\dot{D}\phi/G\phi = 1$.

Proof. (1) We first assume $\dim \phi$ is odd and > 1. Clearly $G\phi = \dot{F}^2$. If f is any monic irreducible quadratic polynomial over $\boldsymbol{R}$, then $f \in \dot{D}\phi$ by 1.1. Thus $\dot{D}\phi/G\phi$ is infinite.

(2) Now assume $\dim \phi$ is even and ϕ is anisotropic. Then there is an interval $I = (a, b)$ such that if $\alpha \in I$, then ϕ at α is $\cong (m \times (1)) \oplus (n \times (-1))$ for fixed positive integers m, n with $m \neq n$ (to see this, apply (3) of 2.5 and (2) of 2.6). Let $a < x < y < b$ and define $f_{xy}(t) = (t - x)(t - y) \in \boldsymbol{R}[t]$. Then $f_{xy}(\alpha) > 0$ if $\alpha \notin I$ so $f_{xy}(t) \in \dot{D}\phi$ by 1.1. Let $y < y_1 < b$, so that $f_{xy_1}(t) \in \dot{D}\phi$ also. Let $h(t) = f_{xy}(t) \div f_{xy_1}(t)$. Then $h(t) \notin G\phi$ by 1.2 since $h(\alpha) < 0$ for $y < \alpha < y_1$. It is now clear that if we choose an infinite sequence of numbers $y < y_1 < y_2 < \cdots < b$ then we obtain an infinite number of distinct cosets of $G\phi$ in $\dot{D}\phi$.

(3) Let $\dim \phi$ be even and let ϕ be isotropic (but not hyperbolic), and assume that ϕ at α is non-hyperbolic for infinitely many $\alpha \in \boldsymbol{R}$. Then there is an open interval I such that for all $\alpha \in I$, ϕ at α is isotropic but not hyperbolic. Thus by the proof of (2) above, $\dot{D}\phi/G\phi$ is infinite.

(4) Finally, assume $\dim \phi$ is even and ϕ is isotropic (but not hyperbolic), and assume that ϕ at α is hyperbolic for almost all $\alpha \in \boldsymbol{R}$. Then by 1.2, $\phi \cong (m \times (1, -1)) \oplus (1, -g)$ where g is a product of monic irreducible quadratic factors. By 1.1, $\dot{D}\phi = \dot{F}$ (where $F = \boldsymbol{R}(t)$. Now $G\phi = G(1, -g) = \dot{F}$ by 1.2 so $\dot{D}\phi/G\phi = 1$.

3. Pfister forms and k_n over $\boldsymbol{R}(t)$. We first consider Pfister forms over $\boldsymbol{R}(t)$.

3.1. *Let ϕ be a quadratic form over $\boldsymbol{R}(t)$ with $\dim \phi = 2^n (n \geqq 2)$. Then the following are equivalent*:

(1) *ϕ is a Pfister form.*

(2) *$\phi \cong 2^{n-1} \times (1, f)$ for some $f \in \boldsymbol{R}[t]$ which is $\pm$ a product of distinct monic linear factors (we allow $f = \pm 1$).*

(3) *ϕ is round and $\det \phi = 1$.*

Proof. (1) $\Rightarrow$ (3) is clear. (3) $\Rightarrow$ (2) by 2.5 (if ϕ is isotropic, let $f = -1$). (2) $\Rightarrow$ (1) is clear.

In (2), f is uniquely determined by ϕ (see 2.6).

We now consider, for the field $F = \boldsymbol{R}(t)$, the algebraic K-groups

$k_nF = K_nF/2K_nF$ of Milnor [6]. k_n is generated additively by the elements $l(c_1)\cdots l(c_n)(c_i \in \dot{F})$. We have $l(-a_1)\cdots l(-a_n)=l(-b_1)\cdots l(-b_n)\Leftrightarrow (1, a_1)\otimes\cdots\otimes(1, a_n)\cong(1, b_1)\otimes\cdots\otimes(1, b_n)$ [2, Main theorem 3.2].

Let $n > 1$. By 3.1 and [2, 3.2], every element of k_nF can be written uniquely in the form $l(-1)^{n-1}l(-f)$ for some $f \in F$ which is $\pm$ a product of distinct monic linear factors or is ± 1. Thus k_nF is isomorphic to the subgroup of $\dot{F}/\dot{F}^2$ consisting of the square classes of products of linear polynomials (note that $l(-1)^{n-1}l(-f)+l(-1)^{n-1}l(-g)=l(-1)^{n-1}l(fg)$). Furthermore, there is a natural isomorphism s_n of k_n onto I^n/I^{n+1} where I is the ideal of the even-dimensional forms of the Witt ring $W(F)$ [2, 6.1].

REMARK 3.2. By [6, 2.3], for $n \geqq 1$ and for any field E there is an isomorphism $K_nE(t)\cong K_nE\oplus(\bigoplus K_{n-1}E[t]/(\pi))$ where the second direct sum extends over all nonzero prime ideals (π) of $E[t]$. Now let $E=\boldsymbol{R}$ and let $n \geqq 2$. The above isomorphism induces an isomorphism $k_n\boldsymbol{R}(t)\cong k_n\boldsymbol{R}\oplus(\bigoplus k_{n-1}\boldsymbol{R}[t]/(\pi))$ where the second direct sum extends over all the polynomials $\pi = t-\alpha$, $\alpha\in\boldsymbol{R}$ (note that k_{n-1} of the complex numbers is 0). Now $k_n\boldsymbol{R}$ and $k_{n-1}\boldsymbol{R}$ are groups of order 2 by [6, 1.6] or [2, 3.2]. Thus there is an isomorphism $k_n\boldsymbol{R}(t)\cong \boldsymbol{Z}_2\oplus(\bigoplus_{\boldsymbol{R}}\boldsymbol{Z}_2)$. This isomorphism is given explicitly as follows: $l(-1)^{n-1}l(-f)$(where f is $\pm$ a product of distinct monic linear factors) maps to $a\oplus(\bigoplus a_\alpha)(\alpha\in\boldsymbol{R})$ where a is 0 if and only if f is monic, and a_α is 1 if and only if $t-\alpha$ divides f.

REMARK 3.3. Let us briefly see what happens when we let our field F be a global field and let $n \geqq 3$. Then we have:

(1) Every Pfister form of dimension 2^n over F is isometric to a form $2^{n-1}\times(1, a)$ for some $a\in\dot{F}$. Also $2^{n-1}\times(1, a)\cong 2^{n-1}\times(1, b)\Leftrightarrow ab\in\dot{F}_p^2$ for all real spots p on F. These facts follow easily from the Weak Hasse-Minkowski Theorem.

(2) By (1) and by [2, Main Theorem 3.2], we see that every element of k_nF can be written as $l(-1)^{n-1}l(-a)$ for some $a\in\dot{F}$, and $l(-1)^{n-1}l(-a)=l(-1)^{n-1}l(-b)\Leftrightarrow ab\in\dot{F}_p^2$ for all p real. Thus $k_nF\cong\bigoplus k_nF_p$ where the direct sum extends over all real spots p (note that $k_nF_p=\boldsymbol{Z}_2$). This fact was first proved by Tate (see appendix of [6]). Elman and Lam [1] gave a simple proof (using the Strong Hasse-Minkowski Theorem) which does not depend on [2].

(3) There are round forms ϕ over F of dimension 2^n (with $\det\phi=1$) which are not Pfister forms [3, 2.6].

Added in proof. In connection with Example 2.3(3), we point out here that, without using elliptic curves theory, examples of rational function fields which do not satisfy the Strong Hasse-Minkowski Theorem can be found in the article: "*On the Hasse Principle for Quadratic Forms*", P.A.M.S., 39 (1973).

The results in § 2 have been generalized recently by R. Elman in his article: "*Rund forms over real algebraic function fields in one variable*" (to appear). Instead of using the local-global method as we have done, Elman's approach is entirely different; he uses the algebraic theory of Pfister forms.

References

1. R. Elman and T. Y. Lam, *Determination of $k_n(n \geqq 3)$ for global fields*, Proc. Amer. Math. Soc., **31** (1972), 427–428.
2. ———, *Pfister forms and K-theory of fields*, J. Algebra, **23** (1972), 181–213.
3. J. S. Hsia and R. P. Johnson, *Round and group quadratic forms over global fields*, J. Number Theory, Vol. **5**, No. 5 (1973), 356–366.
4. J. T. Knight, *Quadratic forms over $R(t)$*, Proc. Cambridge Philos. Soc., **62** (1966), 197–205.
5. F. Lorenz, *Quadratische Formen über Körpern*, Lecture Notes in Math., **130**, Springer-Verlag, Berlin-Heidelberg-New York, 1970.
6. J. Milnor, *Algebraic K-theory and quadratic forms*, Invent. Math., **9** (1970), 318–344.
7. O. T. O'Meara, *Introduction to Quadratic Forms*, Springer-Verlag, 1963.
8. W. Scharlau, *Quadratic. Forms*, Queen's papers on pure and applied math., no. 22, Queen's Univ., Kingston, Ontario, 1969.

Received August 14, 1972. Research of the first author was supported in part by the National Science Foundation under grant NSF GP-23656.

OHIO STATE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

EQUALLY PARTITIONED GROUPS

I. M. ISAACS

It is proved that the only finite groups which can be partitioned by subgroups of equal orders are the p-groups of exponent p. The connection between equally partitioned groups and Sperner spaces is discussed. It is also proved that finite groups partitioned by pairwise permutable subgroups are abelian.

1. Let G be a group and let Π be a collection of proper subgroups of G. Then Π is said to *partition* G if every nonidentity element of G is contained in exactly one $H \in \Pi$. If G is a p-group of exponent p and $|G| > p$, we may let Π be the set of cyclic subgroups of G. Then Π is a partition consisting of subgroups of equal finite orders. Our main result is that the p-groups of exponent p are the only finite groups which can be equally partitioned.

The methods of proof in this paper depend strongly on the finiteness of the group and give no information about which infinite groups can be partitioned by subgroups of equal finite orders.

I began to consider equally partitioned groups after attending a lecture by Prof. A. Barlotti on Sperner spaces. Examples of these geometric objects (which generalize affine spaces) are provided by such groups. In fact the Sperner spaces which arise from finite equally partitioned groups are exactly those which Barlotti and Cofman [2] call translation spaces. This will be discussed further in § 3.

2. Only finite groups will be considered. A great deal is known about partitioned groups. (We mention in particular the papers [1] and [5].) Our theorem, however, is much more elementary and does not depend on the deeper results.

The following easy lemma (which appears in [1]) is crucial to the study of partitioned groups.

LEMMA 1. *Let G be partitioned by Π and let $x, y \in G - \{1\}$ with $xy = yx$. Suppose x and y lie in different elements of Π. Then x and y have equal prime orders.*

Proof. Suppose $o(x) < o(y)$. Then $(xy)^{o(x)} = y^{o(x)} \neq 1$. Let $y \in H \in \Pi$ then $(xy)^{o(x)} \in H$ and hence $xy \in H$. Thus $x \in H$, a contradiction. Therefore $o(x) = o(y)$. Similarly, $o(x^n) = o(y) = o(x)$ for positive integers $n < o(x)$. It follows that $o(x)$ is prime.

LEMMA 2. *Let G be equally partitioned by Π and let $X \subseteqq G$ be* *a subset,* $X \nsubseteq \{1\}$. *Then there exists* $H \in \Pi$ *such that* H *contains no conjugate of* X.

Proof. Suppose that the lemma is false and for each $H \in \Pi$, choose X_H conjugate to X with $X_H \subseteq H$. Let $N_H = N_H(X_H)$ so that H contains at least $|H : N_H|$ conjugates of X. Let $N = N_G(X)$. Then if $|G| = g$ and $|H| = h$ for $H \in \Pi$, we have

$$|G : N| = |G : N_G(X_H)| \leqq |G : N_H| = |G : H||H : N_H|$$

and hence

$$|H : N_H| \geqq h|G : N|/g .$$

Now $|G : N|$ is the number of conjugates of X in G and thus

$$\begin{aligned} |G : N| &\geqq \textstyle\sum_{H \in \Pi} |H : N_H| \\ &\geqq |\Pi||G : N| h/g . \end{aligned}$$

However, $|\Pi| = (g - 1)/(h - 1) > g/h$ and this yields a contradiction.

NOTE. It follows from Lemma 2 that if G is equally partitioned by Π, then no element of Π can contain a full Sylow p-subgroup of G for any $p \mid |G|$. Otherwise, every $H \in \Pi$ would contain an S_p subgroup, violating the lemma.

LEMMA 3. *Let G be equally partitioned. Then every element of G has prime order.*

Proof. Suppose that $x \in G$ has composite order and let $\mathscr{K}$ be the conjugacy class of x. Let Π be the given partition. By Lemma 2, there exists $H \in \Pi$ with $H \cap \mathscr{K} = \phi$. By Lemma 1, no element of H centralizes any element of $\mathscr{K}$. Thus H acts semi-regularly on $\mathscr{K}$ and hence $|H| \mid |\mathscr{K}|$.

Now pick $K \in \Pi$ with $x \in K$. Then K acts semi-regularly by conjugation on $\mathscr{K} - K$ so that $|K| \mid |(\mathscr{K} - K)|$. Since $|H| = |K|$, we conclude that $|K| \mid |\mathscr{K} \cap K|$. This is a contradiction because $0 < |\mathscr{K} \cap K| < |K|$.

The next two results are routine applications of standard facts. We include them for completeness.

LEMMA 4. *Suppose G has a nontrivial normal p-subgroup where p is the largest prime divisor of* $|G|$. *Assume that every element of G has prime order and let* $P \in \mathrm{Syl}_p(G)$. *Then either* $P = G$ *or* $|G : P|$ *is prime and* $P \lhd G$.

Proof. Let $1 \neq U \lhd G$ where U is a p-group. Now G can contain no subgroup, W, of order qr where q and r are (possibly equal) primes different from p. This is so since otherwise $C_U(w) = 1$ for all $1 \neq w \in W$ and this forces W to be cyclic (Satz V. 8. 15b of [3]).

There is nothing to prove if $P = G$ so suppose $P < G$ and let q be the smallest prime divisor of $|G|$. Let $Q \in \mathrm{Syl}_p(G)$. Then $|Q| = q$ and thus G has a normal q-complement, M.

If $M = P$, the proof is complete. Suppose that $P < M$. Then Q normalizes some $R \in \mathrm{Syl}_r(M)$ for $r \neq p$. Thus $|R| = r$ and $|QR| = qr$, a contradiction.

COROLLARY 5. *Assume that every element of G has prime order. Let $P \in \mathrm{Syl}_p(G)$, where p is the largest prime divisor of $|G|$. Then P is a T. I. set (i.e., $P \cap P^x = 1$ for all $x \notin N(P)$).*

Proof. Assume that the corollary is false and let $1 < D = P \cap P^x$ where $P \neq P^x$ and $|D|$ is maximal. Then $N_G(D) = N$ does not have a unique Sylow p-subgroup. This violates Lemma 4 as applied to N.

THEOREM 6. *Let G be equally partitioned. Then G is a p-group of exponent p.*

Proof. Let p be the largest prime divisor of $|G|$ and let $P \in \mathrm{Syl}_p(G)$. By Lemmas 3 and 4, $N(P) = PC$ where either $C = 1$ or $|C| = q$, a prime. By Corollary 5, P is a T. I. set.

We establish some notation. Let $|G| = g$, $|P| = p^b$ and $|C| = c$. Let Π be the given partition and let $|H| = h$ for all $H \in \Pi$. Let p^a be the p-part of h.

Since P is a T. I. set, it follows that $P \cap U \in \mathrm{Syl}_p(U)$ for all subgroups $U \subseteqq G$ with $P \cap U \neq 1$. Thus $|P \cap H| = p^a$ for all $H \in \Pi$ such that $P \cap H \neq 1$. Since $P = \bigcup_{H \in \Pi}(P \cap H)$, it follows that $(p^a - 1) \mid (p^b - 1)$. We can also conclude from the fact that P is a T. I. set that G contains exactly $g(p^b - 1)/p^b c$ elements of order p.

Now by Lemma 2, we may choose $H \in \Pi$ with $H \cap C^g = 1$ for all $g \in G$. Let $P_0 \in \mathrm{Syl}_p(H)$. We may assume that $P_0 \subseteqq P$. Since P is a T. I. set, $N_H(P_0) \subseteqq N_G(P) = PC$. It follows that $N_H(P_0) = P_0 C_0$ where $C_0 \subseteqq C^g$ for some g. Thus $C_0 = 1$ and $P_0 = N_H(P_0)$. By Sylow's Theorem it follows that $h/p^a \equiv 1 \bmod p$.

Let $K \in \Pi$ and let $P_1 \in \mathrm{Syl}_p(K)$. Reasoning as above, we conclude that $N_K(P_1) = P_1 C_1$ where $C_1 \subseteqq C^x$ for some x. Thus $h/(p^a |C_1|) \equiv 1 \bmod p$ and hence $|C_1| \equiv 1 \bmod p$. However, $|C_1| = 1$ or q where q is a prime $< p$. It follows that $C_1 = 1$ and thus every $K \in \Pi$ has self-normalizing Sylow p-subgroups.

Since the Sylow p-pubgroups of $K \in \Pi$ are T. I. sets, it follows that each such K contains exactly $h(p^a - 1)/p^a$ elements of order p. Since $|\Pi| = (g-1)/(h-1)$, this yields

$$g(p^b - 1)/p^b c = (g-1)h(p^a - 1)/(h-1)p^a \ . \tag{1}$$

Since $g/h < (g-1)/(h-1)$, we conclude from (1) that

$$1/c > (p^b - 1)/p^b c > (p^a - 1)/p^a = 1 - 1/p^a \geqq 1/2$$

and thus $c = 1$. Now (1) yields

$$(g-1)h(p^a - 1)p^b = (h-1)g(p^b - 1)p^a \ . \tag{2}$$

Since $((g-1), gp^a) = 1$ and $(p^b - 1)/(p^a - 1)$ is an integer, we obtain

$$gp^a \mid hp^b \ .$$

The p-parts of gp^a and hp^b are equal and $h \mid g$. It follows that $hp^b \mid gp^a$ and thus

$$hp^b = gp^a \ . \tag{3}$$

Combining this with (2) yields

$$(h-1)(p^b - 1) = (g-1)(p^a - 1) \tag{4}$$

and subtracting (4) from (3), one obtains

$$h + p^b = g + p^a \ .$$

Since $h \mid g$ and $h < g$, we have

$$g/2 \leqq g - h = p^b - p^a < p^b \ .$$

Since $p^b \mid g$, we conclude that $p^b = g$ and the result follows.

NOTE. Once it was established that $c = 1$, above, the proof could have been finished using Frobenius' Theorem, ([3], Hauptsatz V. 7. 6). Since P is a self-normalizing T. I. set, Frobenius' Theorem yields a normal p-complement, U, for G. Also $C_U(x) = 1$ for all $1 \neq x \in P$. If $U \neq 1$, it follows from the fact that P has exponent p that $|P| = p$. A contradiction now results by applying the note following Lemma 2.

3. In this section we discuss the connection between Sperner spaces and equally partitioned groups.

DEFINITION. ([4].) A *Sperner space* is a set, S, of "points" and a collection, $\mathscr{L}$, of proper finite subsets of S, called "lines" such that

(a) every two points determine a unique line,

(b) all lines have equal numbers of points,

(c) an equivalence relation (called "parallelism") is defined on $\mathscr{L}$ and

(d) for each $x \in S$, there is exactly one line which contains x in each parallel class.

If G is a group which is equally partitioned by Π, we may define a Sperner space by taking $S = G$, $\mathscr{L} = \{Hx \mid H \in \Pi, x \in G\}$ and setting $(Hx) \| (Ky)$ if and only if $H = K$. It is routine to check that this does define a Sperner space. We denote this space by $S(G, \Pi)$.

Given a Sperner space, $(S, \mathscr{L})$, we consider the groups, $G(S, \mathscr{L})$, consisting of all those collineations of S which map each line to a line parallel to itself. Since no two distinct parallel lines of $(S, \mathscr{L})$ can intersect (by condition (d)), it follows that if $g \in G(S, \mathscr{L})$ fixes a point, $x \in S$, then g fixes every line through x. It now follows easily that only the identity of $G(S, \mathscr{L})$ fixes two points of S.

Let $G_0(S, \mathscr{L}) = \{1\} \cup \{g \in G(S, \mathscr{L}) \mid g$ fixes no points of $S\}$. In [2], Barlotti and Cofman call a Sperner space $(S, \mathscr{L})$ a *translation space* if $G_0(S, \mathscr{L})$ is a group which is transitive on S. If S is finite, it follows from Frobenius' Theorem ([3], Satz. V. 8. 2 (a)) that $(S, \mathscr{L})$ is a translation space if and only if $G(S, \mathscr{L})$ is transitive on S. If (G, Π) is a finite equally partitioned group and $(S, \mathscr{L}) = S(G, \Pi)$, then $G(S, \mathscr{L})$ contains right multiplications by elements of G and hence is transitive. It follows that $S(G, \Pi)$ is a translation space and $G_0(S, \mathscr{L})$ is the group of right multiplications.

We claim that if $(S, \mathscr{L})$ is any finite translation space then $(S, \mathscr{L}) \cong S(G, \Pi)$ for some equally partitioned group (G, Π). Let $G = G_0(S, \mathscr{L})$ and choose a point $e \in S$. For $l \in \mathscr{L}$, let H_l be the (setwise) stabilizer of l in G and let $\Pi = \{H_l \mid l \in \mathscr{L}$ and $e \in l\}$. If $e, x \in l$ and $g \in G$ with $eg = x$, then $x \in l \cap lg$ and thus $l = lg$ and $g \in H_l$. It follows that H_l is transitive on l and $|H_l| = |l|$. Therefore, all $H \in \Pi$ have equal order. If $H, K \in \Pi$ with $H \neq K$, then $H \cap K$ fixes e and hence $H \cap K = 1$. Also

$$\begin{aligned}|G| = |S| &= 1 + \Sigma\{|l| - 1 \mid l \in \mathscr{L} \text{ and } e \in l\}\\ &= 1 + \Sigma\{|H| - 1 \mid H \in \Pi\} = |\bigcup \Pi|\end{aligned}$$

and thus Π is a partition for G.

To see that $S(G, \Pi) \cong (S, \mathscr{L})$, define $\theta : G \to S$ by $\theta(g) = eg$. It is routine to show that θ is an isomorphism of Sperner spaces.

One further remark on the correspondence between finite translation spaces and finite equally partitioned groups is in order. If (G, Π) and (G_1, Π_1) are two equally partitioned groups such that $S(G, \Pi) \cong S(G_1, \Pi_1)$, then $G \cong G_1$ and this group isomorphism can be chosen so as to carry Π to Π_1. This follows since $G \cong G_0(S(G, \Pi))$ and under this (natural) isomorphism, Π corresponds exactly to the set of stabilizers of the lines through 1.

Let $(S, \mathscr{L})$ be a finite translation space. By Theorem 6, $|S| = p^b$ for some prime, p, and $|l| = p^a$ for $l \in \mathscr{L}$. Also, $(p^a - 1) \mid (p^b - 1)$ and as is well known, this forces $a \mid b$. We may define the *dimension* of $(S, \mathscr{L})$ to be b/a.

Let $q = p^a$ and let $K = \mathrm{GF}(q)$. Let V be a vectorspace of dimension n over K and let Π be the set of one-dimensional subspaces of V. Then Π equally partitions V and of course $S(V, \Pi)$ is an affine space of dimension n. This suggests the question of which translation spaces, $(S, \mathscr{L})$, correspond to abelian equally partitioned groups. These are *not* necessarily affine although they do satisfy the following condition:

(*) Let $l, m \in \mathscr{L}$ with $l \cap m \neq \varnothing$. Let $x \in l$ and $y \in m$. Let $l' \| l$ with $y \in l'$ and $m' \| m$ with $x \in m'$. Then $l' \cap m' \neq \varnothing$.

It is easy to see that $S(G, \Pi)$ satisfies (*) if and only if for every $H, K \in \Pi$ and every $h \in H$ and $k \in K$ we have $Hk \cap Kh \neq \varnothing$. This condition is clearly satisfied if G is abelian since then $hk \in Hk \cap Kh$. In the next section we prove that only in abelian groups does this condition hold.

4. We begin with the following lemma.

LEMMA 7. *Let $H, K \subseteqq G$. Then $HK = KH$ if and only if for every $h \in H$ and $k \in K$ we have $Hk \cap Kh \neq 1$.*

Proof. Suppose $HK = KH$. Let $h \in H$ and $k \in K$. Then $kh^{-1} \in KH = HK$ and $kh^{-1} = h_1^{-1}k_1$ for some $h_1 \in H$ and $k_1 \in K$. Thus $h_1k = k_1h \in Hk \cap Kh$.

Conversely, let $x \in KH$. Write $x = kh^{-1}$ for some $k \in K$ and $h \in H$. Now choose $k_1h = h_1k \in Kh \cap Hk$ so that $k_1 \in K$ and $h_1 \in H$. Then $x = kh^{-1} = h_1^{-1}k_1 \in HK$ and $KH \subseteqq HK$. The reverse inclusion follows symmetrically and the proof is complete.

The main result of this section is the following.

THEOREM 8. *Let G be a finite group partitioned by Π. Assume that $HK = KH$ for all $H, K \in \Pi$. Then G is an elementary abelian p-group.*

Note that we do not assume that all elements of Π have equal order. Theorem 8 and Lemma 7 prove the claim made at the end of § 3. To prove Theorem 8, we strengthen it somewhat and use induction.

THEOREM 9. *Let G be finite and partitioned by Π. Suppose $A \in \Pi$ and $AH = HA$ for all $H \in \Pi$. Then $A \lhd G$.*

Proof. We use induction on $|G|$. If $A < L < G$, then L is partitioned by $\Pi_0 = \{H \cap L \mid H \in \Pi\}$. If $H \in \Pi$, then AH is a group and $AH \cap L = A(H \cap L)$. Thus $A(H \cap L) = (H \cap L)A$ and by induction $A \lhd L$. Let $N = N(A)$. If $H \in \Pi$ and $AH < G$, it follows that $A \lhd AH$ and $H \subseteq N$.

Assume $N < G$ and let $\Pi_1 = \{H \in \Pi \mid H \nsubseteq N\}$. Then $HA = G$ for all $H \in \Pi_1$ and hence $|H| = |G : A|$ for these H. Also for $H \in \Pi_1$, we have $N = A(N \cap H)$ and thus $|N \cap H| = |N : A|$.

Now

$$G - N = \bigcup \{H - (H \cap N) \mid H \in \Pi_1\}$$

and since this union is disjoint, we obtain

$$|G| - |N| = |\Pi_1| (|G : A| - |N : A|) .$$

Solving this yields $|\Pi_1| = |A|$.

Now

$$\begin{aligned} |\bigcup \Pi_1| &= 1 + |\Pi_1| (|G : A| - 1) \\ &= 1 + |G| - |A| . \end{aligned}$$

It follows that $\Pi = \Pi_1 \cup \{A\}$ and every element of $G - A$ lies in some $H \in \Pi_1$.

Let $g \in G$. To show that $A^g = A$, it suffices to show that $A^g \cap H = 1$ for all $H \in \Pi_1$. Choose $H \in \Pi_1$. Since $G = AH$, we may write $g = ah$ for some $a \in A$ and $h \in H$. Then

$$A^g \cap H = A^h \cap H = (A \cap H)^h = 1$$

and the proof is complete.

Proof of Theorem 8. By Theorem 9 we have $H \lhd G$ for all $H \in \Pi$. Therefore, if $H, K \in \Pi$, $H \neq K$ we have $K \subseteq C(H)$ and hence $G = H \cup C(H)$. Since $H < G$, we have $G = C(H)$ and $H \subseteq Z(G)$. It follows that G is abelian. The result now follows by Lemma 1.

5. In this section we discuss a class of examples of equally partitioned groups. Since every p-group of exponent p is equally partitioned by its cyclic subgroups, it is interesting to look for examples of groups partitioned by subgroups of order $q = p^a > p$. The elementary abelian groups of order q^n have this property. Nonabelian examples are provided by the next result if $p > 2$.

THEOREM 10. *Let* $n \leqq p$ and $q = p^e$. *Then the Sylow p-subgroups of* GL(n, q) *are partitioned by abelian subgroups of order* q.

NOTE. If $n > p$, then the Sylow p-subgroups of $\mathrm{GL}(n, q)$ do not have exponent p and hence cannot be equally partitioned.

Proof of Theorem 10. Let $K = GF(q)$ and let A be the space of strictly upper triangular $n \times n$ matrices over K. Then $P = \{I + a \mid a \in A\}$ is a Sylow p-subgroup of $\mathrm{GL}(n, q)$. For $a \in A$, let $M_a(t) = \exp(at)$ for $t \in K$. This is well defined since $(at)^n = 0$ and $n \leqq p$. Since $M_a(s)M_a(t) = M_a(s + t)$, we conclude that $P_a = \{M_a(t) \mid t \in K\}$ in an abelian subgroup of P.

We will show that if $a, b \in A$ and $\exp(a) = \exp(b)$, then $a = b$. It will follow that $|P_a| = q$ if $a \neq 0$ and that $P_a \cap P_b = 1$ unless $b = at$ for some $t \in K$; in which case $P_a = P_b$. Taking $\Pi = \{P_a \mid 0 \neq a \in A\}$ we have $|\Pi| = (|A| - 1)/(q - 1)$ and

$$|\bigcup \Pi| = |\Pi|(q - 1) + 1 = |A| = |P|$$

as desired.

Suppose then that $\exp(a) = \exp(b)$. For $m \in \boldsymbol{Z}$, $\exp(ma) = \exp(a)^m$ and thus $\exp(at) = \exp(bt)$ for all $t \in \mathrm{GF}(p)$. Let x be an indeterminate and let $E(x) = \exp(ax) - \exp(bx)$. Then $E(x)$ is a matrix with polynomial entries of degree $< p$. Since $E(t) = 0$ for all $t \in \mathrm{GF}(p)$, it follows that $E(x)$ is identically 0. Comparing coefficients of x yields $a = b$ and the proof is complete.

We close with the following question: Does there exist a group partitioned by subgroups of equal order not all of which are abelian?

REFERENCES

1. R. Baer, *Partitionen endicher Gruppen*, Math. Zeit., **75** (1961), 333–372.
2. A. Barlotti and J. Cofman, *Finite Sperner spaces constructed from projective and affine spaces*, to appear.
3. B. Huppert, *Endliche Gruppen I*, Springer-Verlag, Berlin, 1967.
4. E. Sperner, *Affine Raume mit schwacher Inzidenz und zugehörige algebraische Structuren*, J. Reine Angew. Math., **204** (1960), 205–215.
5. M. Suzuki, *On a finite group with a partition*, Arch. Math., **12** (1961), 241–254.

Received July 20, 1972. Research partially supported by NSF grant GP-32813X.

UNIVERSITY OF WISCONSIN-MADISON

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

HYPERPOLYNOMIAL APPROXIMATION OF SOLUTIONS OF NONLINEAR INTEGRO-DIFFERENTIAL EQUATIONS

A. G. KARTSATOS AND E. B. SAFF

Consider the integro-differential equation

$$(*)\qquad U(x) \equiv x' + A(t, x) + \int_a^t F(t, s, x(s))ds = T(t), t \in [a, b]$$

subject to the initial condition

$$(**)\qquad x(a) = h .$$

Then a problem in approximation theory is whether a solution $x(t)$ of ((*), (**)) can be approximated, uniformly on $[a, b]$, by a sequence of polynomials P_n, which satisfy (**) and minimize the expression $\|T(\cdot) - U(P_n)\|$, where $\|\cdot\|$ is a certain norm. It is shown here that such a sequence of minimizing polynomials, or, more generally, hyperpolynomials, exists with respect to the L_p-norm ($1 < p \leqq \infty$) and converges to $x(t)$, uniformly on $[a, b]$, under the mere assumption of existence and uniqueness of $x(t)$.

The results of this paper are intimately related to those of Stein [11], who studied the approximation of solutions of scalar linear integro-differential equations of the form

$$(1)\qquad W(x) \equiv L(x) - \int_a^b h(t, s)x(s)ds = f(t) ,$$

($L(x) \equiv x^{(m)}(t) + f_1(t)x^{(m-1)}(t) + \cdots + f_m(t)x(t)$) subject to the two-point boundary conditions:

$$(2)\qquad W_i(x) \equiv A_i(x) + B_i(x) + \int_a^b V_i(t)x(t)dt = 0 , \quad i = 1, 2, \cdots, m$$

where $A_i(u) \equiv \sum_{k=1}^m a_{ik}u^{(k-1)}(a)$, $B_i(u) \equiv \sum_{k=1}^m b_{ik}u^{(k-1)}(b)$. Namely, he showed that under certain condition on L, h, f, if $x(t)$ is the unique solution of (1), which satisfies the linearly independent boundary conditions (2), then for every $n \geqq 2m - 1$ there exists a unique polynomial p_n of degree at most n, which satisfies (2) and best approximates the solution of (1) with respect to the L_p-norm ($1 \leqq p < \infty$). He then considered the convergence of the sequences $\{p_n^{(k)}\}$, $k = 1, 2, \cdots, m - 1$ to the solution $x(t)$ and its derivatives up to the order $m - 1$ respectively. Extension of these results were also made for trigonometric polynomials, or linear combinations of orthonormal functions. The present paper extends the results of Stein and has points of contact with the rest of the papers in the references.

1. Preliminaries. Let $R = (-\infty, +\infty)$. For the system ((*), (**)) we assume the following: $A(t, u)$ is an m-vector of functions defined and continuous on $[a, b] X R^m$. $F(t, s, u)$ is an m-vector of functions defined and continuous on the set $S \equiv \{(t, s, u) \in [a, b] X [a, b] X R^m; s \leqq t\}$. $T(t)$ is an m-vector of functions defined and continuous on $[a, b]$.

Let B_k, $k = 1, m$, be the Banach space of all k-vectors of continuous functions on $[a, b]$ with norm

$$\|f\|_{B_k} = \sup_{t \in [a,b]} \|f(t)\| ,$$

where, for a vector $u \in R^k$, $\|u\| = \max_{1 \leqq i \leqq k} |u_i|$. By B'_k we denote the Banach space of all functions $f \in B_k$ which are continuously differentiable on $[a, b]$. The norm now is

$$\|f\|_{B'_k} = \max_{i=0,1} \{\|f^{(i)}\|_{B_k}\} .$$

A sequence $\{g_n\}$ of functions in B'_1 is said to be linearly independent if every finite number of the g_n's is linearly independent on $[a, b]$. A linearly independent sequence $\{g_n\}$ is said to be a *d-sequence* if the set of all finite linear combinations of the g_n's is dense in B'_1. For each $i = 1, 2, \cdots, m$ let $\{g_{n,i}\}_{n=1}^{\infty}$ be a fixed d-sequence in B'_1. We assume without loss of generality that $g_{1,i}(a) \neq 0$, $i = 1, 2, \cdots, m$. By a *hyperpolynomial* of degree at most j we mean a function p of the form

$$p = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{bmatrix} = \begin{bmatrix} c_{1,1}g_{1,1} + c_{2,1}g_{2,1} + \cdots + c_{j,1}g_{j,1} \\ c_{1,2}g_{1,2} + c_{2,2}g_{2,2} + \cdots + c_{j,2}g_{j,2} \\ \vdots \quad\quad \vdots \quad\quad \vdots \\ c_{1,m}g_{1,m} + c_{2,m}g_{2,m} + \cdots + c_{j,m}g_{j,m} \end{bmatrix} .$$

By Π_n we denote the set of all hyperpolynomials of degree at most n which satisfy the initial condition (**). For a function $f \in B_m$ we put

$$\|f\|_p = \left[\int_a^b \|f(t)\|^p dt\right]^{1/p} , \quad 1 \leqq p < +\infty .$$

We also make use of the symbol $\|f\|_\infty$ instead of $\|f\|_{B_m}$.

2. Main results.

THEOREM 1. *Let $1 < p \leqq \infty$ and suppose that the system ((*), (**)) has a unique*[1] *solution $x(t)$ defined on $[a, b]$. Then for each n suffi-*

[1] Uniqueness means that any solution of ((*), (**)) which is defined on a subinterval $[a, c]$ of $[a, b]$ must coincide with $x(t)$.

ciently large there exists a hyperpolynomial $Q_n \in \Pi_n$ *such that*

$$\| T - U(Q_n) \|_p = \inf_{P \in \Pi_n} \| T - U(P) \|_p . \tag{3}$$

Furthermore, the sequence $Q_n(t)$ *converges uniformly to* $x(t)$ *on* $[a, b]$. *For the case* $p = \infty$ *we have, in addition, that the sequence* $Q'_n(t)$ *converges uniformly to* $x'(t)$ *on* $[a, b]$.

The proof requires the following lemmas:

LEMMA 1. *The set of all hyperpolynomials is dense in* B'_m.

Proof. Obvious.

LEMMA 2. *Let* $f \in B'_m$ *satisfy* $(**)$. *Then there exists a sequence of hyperpolynomials* $p_n \in \Pi_n$, $n = 1, 2, \cdots$, *such that*

$$\lim_{n \to \infty} \| f - p_n \|_{B'_m} = 0 . \tag{4}$$

Proof. By Lemma 1 there exists a sequence $\{q_n\}$ of hyperpolynomials such that

$$\lim_{n \to \infty} \| f - q_n \|_{B'_m} = 0 . \tag{5}$$

We can (and do) assume that each q_n is of degree at most n, respectively, where $n = 1, 2, \cdots$.

Put $d_n \equiv h - q_n(a)$ and let $d_{n,i}$ be the ith component of d_n. Set

$$s_n(t) \equiv \begin{bmatrix} c_{n,1} g_{1,1}(t) \\ c_{n,2} g_{1,2}(t) \\ \vdots \\ c_{n,m} g_{1,m}(t) \end{bmatrix} ,$$

where $c_{n,i} \equiv d_{n,i}/g_{1,i}(a)$. Since

$$\| d_n \| = \| h - q_n(a) \| = \| f(a) - q_n(a) \| \longrightarrow 0 \text{ as } n \longrightarrow \infty ,$$

it follows that

$$\lim_{n \to \infty} c_{n,i} = 0, \quad \text{for each} \quad i = 1, 2, \cdots, m .$$

Hence

$$\| s_n \|_{B'_m} \longrightarrow 0 \quad \text{as} \quad n \longrightarrow \infty . \tag{6}$$

Now define $p_n(t) \equiv q_n(t) + s_n(t)$. Then

$$p_n(a) = q_n(a) + s_n(a) = q_n(a) + d_n = h ,$$

and so $p_n \in \Pi_n$ for each $n = 1, 2, \cdots$. From (5) and (6) it follows that

$$\lim_{n\to\infty} \| f - p_n \|_{B'_m} = 0 .$$

LEMMA 3. *Let*

$$\mu_{n,p} \equiv \inf_{P \in \Pi_n} \| T - U(P) \|_p .$$

Then $\mu_{n,p} \to 0$ *as* $n \to \infty$.

Proof. It suffices to show that $\mu_{n,\infty} \to 0$ as $n \to \infty$. By Lemma 2 there exists a sequence $p_n \in \Pi_n$, $n = 1, 2, \cdots$, such that

$$\lim_{n\to\infty} \| x - p_n \|_{B'_m} = 0 .$$

Since $x(t)$ satisfies (*) we deduce that

$$(7) \quad \begin{aligned} \mu_{n,\infty} &\leqq \| T - U(p_n) \|_\infty \leqq \| x' - p'_n \|_\infty + \| A(\cdot, x) - A(\cdot, p_n) \|_\infty \\ &+ (b - a) \max_{a \leqq s \leqq t \leqq b} \| F(t, s, x(s)) - F(t, s, p_n(s)) \| . \end{aligned}$$

Obviously $\| x'_n - p'_n \|_\infty \leqq \| x_n - p_n \|_{B'_m} \to 0$ as $n \to \infty$. Also from the uniform convergence of the p_n to x and the continuity of the functions A and F it follows that the last two terms in the right-hand member of (7) tend to zero as $n \to \infty$. This proves Lemma 3.

LEMMA 4. *If* $P_n \in \Pi_n$ *is a sequence of hyperpolynomials such that*

$$(8) \qquad \lim_{n\to\infty} \| T - U(P_n) \|_p = 0 , \qquad 1 < p \leqq \infty ,$$

then the $P_n(t)$ *converge uniformly to* $x(t)$ *on* $[a, b]$. *For the case* $p = \infty$ *we have, in addition, that the derivatives* $P'_n(t)$ *converge uniformly to* $x'(t)$ *on* $[a, b]$.

Proof. The proof is similar, but not identical, to that of [2, Thm. 3, p. 17]. We shall sketch the argument for the real line only.

Let M be a constant such that $|x(t)| < M$ for all $t \in [a, b]$. Note that $|h| = |x(a)| < M$. Set $\mathscr{R} \equiv [a, b]X[-M, M]$. Since the norms $\| U(P_n) \|_p$ are uniformly bounded, and the functions $A(t, u)$ and $F(t, s, u)$ are continuous, there exist constants K_1 and K_2 such that

$$\int_a^b | U(P_n)(t) - A(t, u) |^p dt \leqq K_1^p,\ u \in [-M, M] ,$$

$$| F(t, s, u) | \leqq K_2 , \qquad a \leqq s \leqq t \leqq b,\ u \in [-M, M] .$$

Let $K \equiv K_1 + K_2(b - a)^{1+1/p}$, and consider the curves C_1: $u = h + K(t - a)^{1/q}$, C_2: $u = h - K(t - a)^{1/q}$, where q satisfies the equation $1/p + 1/q = 1$. Let t_i^*, $a < t_i^* \leqq b$, $i = 1, 2$, be the abscissa of the second point of intersection of the curve C_i with the boundary of the rectangle $\mathscr{R}$. Put $t^* \equiv \min(t_1^*, t_2^*)$. We shall show that for each n there holds

$$|P_n(t)| \leqq M,\ t \in [a, t^*]\ . \tag{9}$$

Let t_n be the abscissa of the first point to the right of a at which the graph of $P_n(t)$ intersects the boundary of $\mathscr{R}$. Integrating the equation

$$P_n'(t) = U(P_n)(t) - A(t, P_n(t)) - \int_a^t F(t, s, P_n(s))ds \tag{10}$$

from a to t_n, we deduce that

$$\begin{aligned}
|P_n(t_n) - h| &\leqq \int_a^{t_n} |U(P_n)(t) - A(t, P_n(t))|dt + \int_a^{t_n}\int_a^t |F(t, s, P_n(s))|dsdt \\
&\leqq \left[\int_a^{t_n} |U(P_n)(t) - A(t, P_n(t))|^p dt\right]^{1/p} (t_n - a)^{1/q} \\
&\quad + K_2(b - a)(t_n - a) \\
&\leqq K_1(t_n - a)^{1/q} + K_2(b - a)^{1+1/p}(t_n - a)^{1/q} = K(t_n - a)^{1/q}\ .
\end{aligned}$$

Hence the point $(t_n, P_n(t_n))$ lies between the curves C_1 and C_2. Thus $t_n \geqq t^*$, which proves (9).

It also follows from integrating the equation (10) that the sequence $P_n(t)$ is equicontinuous on $[a, t^*]$. Therefore, by Ascoli's Theorem, each subsequence of the $P_n(t)$ possesses a subsequence which converges uniformly on $[a, t^*]$. Suppose that $y(t)$ is the uniform limit on $[a, t^*]$ of the subsequence $P_k(t)$. From (8) and Hölder's inequality it follows that

$$\lim_{k\to\infty} \int_a^t U(P_k)(\tau)d\tau = \int_a^t T(\tau)d\tau\ , \quad t \in [a, b]\ . \tag{11}$$

Taking the limit as $k \to \infty$ in the equation

$$P_k(t) - h = \int_a^t U(P_k)(\tau)d\tau - \int_a^t A(\tau, P_k(\tau))d\tau - \int_a^t\int_a^\tau F(\tau, s, P_k(s))dsd\tau\ ,$$

we deduce from (11) and the continuity of the functions A and F that

$$y(t) - h = \int_a^t T(\tau)d\tau - \int_a^t A(\tau, y(\tau))d\tau - \int_a^t\int_a^\tau F(\tau, s, y(s))dsd\tau\ ,$$

for $t \in [a, t^*]$. Thus $y(t)$ satisfies the system ((*), (**)) on $[a, t^*]$ and so must equal $x(t)$ on this interval. Since $y(t)$ was an arbitrarily chosen limit function, the original sequence $P_n(t)$ must converge to $x(t)$ uniformly on $[a, t^*]$.

Considering the fact that the proof given above carries over under the more general hypothesis that the initial values of the $P_n(t)$ converge to the corresponding initial value of $x(t)$, one can show, as in the proof of [2, Thm. 3, p. 17], that the sequence $P_n(t)$ converges to $x(t)$ uniformly on $[a, b]$.

For the case $p = \infty$ it follows immediately from equation (10) that $\lim_{n\to\infty} P'_n(t) = x'(t)$ uniformly on $[a, b]$.

Proof of Theorem 1. It is clear from Lemmas 3 and 4 that if the minimizing hyperpolynomials Q_n exist, then they have the asserted convergence properties.

We first show that if Q_k does not exist, then there is a hyperpolynomial $P_k \in \Pi_k$ such that

$$\|T - U(P_k)\|_p < \mu_{k,p} + 1/k\ , \tag{12}$$

and

$$\|P_k\|_\infty > k\ . \tag{13}$$

If this were not the case, there exists a sequence of hyperpolynomials $\pi_j \in \Pi_k$ such that

$$\|T - U(\pi_j)\|_p \longrightarrow \mu_{k,p} \quad \text{as} \quad j \longrightarrow \infty\ , \tag{14}$$

and

$$\|\pi_j\|_\infty \leqq k\ , \quad \forall j\ .$$

It is not difficult to show that the set $\{\pi \in \Pi_k \mid \|\pi\|_\infty \leqq k\}$ is compact in the B'_m norm. Hence there is a subsequence of the π_j which converges in the B'_m norm to a hyperpolynomial $\pi_0 \in \Pi_k$. From (14) and the continuity of the functions A and F it follows that

$$\|T - U(\pi_0)\|_p = \mu_{k,p}\ ,$$

which is a contradiction.

Now suppose that there is an increasing sequence of positive integers k such that Q_k does not exist. Then there is a sequence of hyperpolynomials $P_k \in \Pi_k$ which satisfy (12) and (13). For this sequence we have

$$\|T - U(P_k)\|_p \longrightarrow 0 \quad \text{as} \quad k \longrightarrow \infty\ , \tag{15}$$

and

$$\|P_k\|_\infty \longrightarrow \infty \quad \text{as} \quad k \longrightarrow \infty\ .$$

But from (15) and Lemma 4 we also have $\|P_k\|_\infty \to \|x\|_\infty$ as $k \to \infty$, which is a contradiction.

Hence Q_n exists for n sufficiently large. This completes the proof of Theorem 1.

To prove the existence and convergence of best L_1 approximating hyperpolynomials we impose Lipschitz conditions on the functions A, F.

THEOREM 2. *Suppose that*

$$\|A(t, u) - A(t, v)\| \leqq \lambda_1 \|u - v\|, \ (t, u, v) \in [a, b] X R^m X R^m ,$$
$$\|F(t, s, u) - F(t, s, v)\| \leqq \lambda_2 \|u - v\|, \ (t, s, u, v) \in S X R^m ,$$

where λ_1, λ_2 *are fixed positive constants. Let the system* ((*), (**)) *have the unique solution* $x(t)$ *on* $[a, b]$. *Then for each* n *sufficiently large there exists a hyperpolynomial* $Q_n \in \Pi_n$ *such that*

$$\| T - U(Q_n) \|_1 = \inf_{P \in \Pi_n} \| T - U(P) \|_1 .$$

Furthermore, the sequence $Q_n(t)$ *converges uniformly to* $x(t)$ *on* $[a, b]$.

The proof relies on the following analogue of Lemma 4:

LEMMA 5. *If* $P_n \in \Pi_n$ *is a sequence of hyperpolynomials such that* $\lim_{n\to\infty} \| T - U(P_n) \|_1 = 0$, *then the* $P_n(t)$ *converge uniformly to* $x(t)$ *on* $[a, b]$.

Proof. Clearly,

$$\begin{aligned} \|x(t) - P_n(t)\| &\leqq \int_a^t \| T(\tau) - U(P_n)(\tau) \| d\tau + \int_a^t \|A(\tau, x(\tau)) - A(\tau, P_n(\tau))\| d\tau \\ &\quad + \int_a^t \int_a^\tau \| F(\tau, s, x(s)) - F(\tau, s, P_n(s)) \| ds d\tau \\ &\leqq \| T - U(P_n) \|_1 + \lambda_1 \int_a^t \| x(\tau) - P_n(\tau) \| d\tau \\ &\quad + \lambda_2 (b - a) \int_a^t \| x(\tau) - P_n(\tau) \| d\tau . \end{aligned}$$

From Gronwall's inequality we deduce that

$$\|x(t) - P_n(t)\| \leqq \| T - U(P_n) \|_1 \exp [(\lambda_1 + \lambda_2(b - a))(b - a)] .$$

Thus $\|x - P_n\|_\infty \to 0$ as $n \to \infty$.

Proof of Theorem 2. It follows from Lemmas 3 and 5 that if the minimizing hyperpolynomials exist, then they converge uniformly to $x(t)$ on $[a, b]$. To establish existence one argues as in the proof of Theorem 1.

REMARKS. Let A, F satisfy the conditions of Theorem 2 and, for

$1 \leqq p < \infty$, let $Q_n \in \Pi_n$ denote L_p-norm-minimizing hyperpolynomials. Concerning the degree of convergence of the Q_n to x it can be shown, by use of Hölder's inequality and Gronwall's inequality, that

$$||x - Q_n||_\infty \leqq \mu_{n,p}(b-a)^{(p-1)/p} \exp\left[(\lambda_1 + \lambda_2(b-a))(b-a)\right] .$$

Also if the functions $T(t) - U(Q_n)(t)$ satisfy a Lipschitz condition on $[a, b]$ uniformly w.r.t. n, the sequence $Q'_n(t)$ converges uniformly to $x'(t)$ on $[a, b]$. The proof of this fact follows from Theorem 5 in [13].

The results of this paper can be extended to integro-differential equations with Fredholm integrals of the form

$$W(x) = x' + A(t, x) + \int_a^b F(t, s, x(s))ds = T(t) .$$

It would be of interest to obtain similar results for equations of the type (*) under linearly independent boundary conditions of the form:

$$Bx(a) + Cx(b) + \int_a^b V(t)x(t)dt = h ,$$

where B, C are constant $m \times m$ matrices and V is a continuous $m \times m$ matrix-valued function on $[a, b]$.

REFERENCES

1. A. Bacopoulos and A. G. Kartsatos, *On polynomials approximating the solutions of nonlinear differential equations*, Pacific J. Math., (to appear).

2. W. A. Coppel, *Stability and Asymptotic Behavior of Differential Equations*, D. C. Heath and Co., Boston, 1965.

3. M. S. Henry, *Best approximate solutions of nonlinear differential equations*, J. Approx. Theory, **3** (1970), 59-65.

4. M. S. Henry and F. Max Stein, *On the uniform convergence of a best approximate solution of the Riccati matrix equation*, Rend. Circ. Mat. Palermo, **18** (1969), 307-312.

5. ———, *An L_q approximate solution of the Riccati matrix equation*, J. Approx. Theory, **2** (1969), 237-240.

6. ———, *An approximate solution of the Riccati matrix equation*, Proc. Amer. Math. Soc., **25** (1970), 8-12.

7. R. G. Huffstutler and F. M. Stein, *The best approximate solution of the extended Riccati equation using the L_q norm*, Rend. Circ. Mat. Palermo, **16** (1967), 373-381.

8. ———, *The approximate solution of certain nonlinear differential equations*, Proc. Amer. Math. Soc., **19** (1968), 998-1002.

9. ———, *The approximate solution of $y' = F(x, y)$*, Pacific J. Math., **24** (1968), 283-289.

10. E. N. Oberg, *The approximate solution of integral equations*, Bull. Amer. Math. Soc., **41** (1935), 276-284.

11. F. M. Stein, *The approximate solution of integro-differential equations*, Doctoral Dissertation, State Univ. of Iowa, 1955.

12. F. M. Stein and R. G. Huffstutler, *The approximate solution of Riccati's equation*, J. SIAN Numer. Anal., **3** (1966), 425-434.

13. F. M. Stein and K. F. Klopfenstein, *Approximate solutions of a system of differential equations*, J. Approx. Theory, **1** (1968), 279–292.

Received July 21, 1972. The research of the second author was supported, in part, by NSF Grant GF-19275.

UNIVERSITY OF SOUTH FLORIDA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

ON ELEMENTARY IDEALS OF θ-CURVES IN THE 3-SPHERE AND 2-LINKS IN THE 4-SPHERE

SHIN'ICHI KINOSHITA

Let L be a polyhedron in an n-sphere $S^n(n \geqq 3)$ that does not separate S^n. A topological invariant of the position of L in S^n can be introduced as follows: Let l be an integral $(n-2)$-cycle on L. For each nonnegative integer d, the dth elementary ideal $E_d(l)$ is associated to l on L in S^n. If l and l' are homologous on L, then $E_d(l)$ is equal to $E_d(l')$. Now the collection of $E_d(l)$ for all possible l is a topological invariant of L in S^n.

In this paper the following two cases of $E_d(l)$ are considered: (1) l is a 1-cycle on a θ-curve L in S^3, and (2) l is a 2-cycle on a 2-link L in S^4, i.e., the union of two disjoint 2-spheres in S^4, where each of two 2-spheres is trivially imbedded in S^4.

The dth elementary ideal $E_d(l)$ of l on L is defined as follows (more precisely see [3]): Let G be the fundamental group $\pi(S^n - L)$ and H the multiplicative infinite cyclic group generated by t. Let ψ be a homomorphism of G into H defined by

$$g^{\psi} = t^{\operatorname{link}(g,l)} ,$$

where link (g, l) is the linking number between g and l. Using Fox's free differential calculus, we associate to ψ the dth elementary ideal E_d of the group G, evaluated in the group ring JH of H over integers. This dth elementary ideal E_d depends only on G and ψ, and hence it depends only on the position of l on L in S^n. We shall denote it by $E_d(l)$.

In this paper we shall prove the following two theorems.

THEOREM 1. *Let $f(t)$ be an integral polynomial with $f(1) = 1$. Then there exists a θ-curve L_f in S^3, and an integral 1-cycle l on L_f such that*

$$\begin{cases} E_0(l) = E_1(l) = (0) , \\ E_2(l) = (f(t)) \quad \text{and} \\ E_d(l) = (1) , \quad \text{if} \quad d > 2 . \end{cases}$$

THEOREM 2. *Let $f(t)$ be an integral polynomial with $f(1) = 1$. Then there exists a 2-link L_f in S^4, and an integral 2-cycle l on L_f such that*

(1) *each component of L_f is a trivially imbedded 2-sphere in S^4, and that*

(2) *we have*

$$\begin{cases} E_0(l) = E_1(l) = (0)\,, \\ E_2(l) = (f(t)) \quad \textit{and} \\ E_d(l) = (1)\,, \quad \textit{if} \quad d > 2\,. \end{cases}$$

COROLLARY. *Let $f(t)$ be an integral polynomial with $f(1) = 1$. Then there exists an oriented 2-link L_f in S^4 such that*

(1) *each component of L_f is a trivial 2-sphere in S^4, and that*

(2) *the* dth *elementary ideal of L_f, in the usual sense and in the reduced form, is as follows*:

$$\begin{cases} E_0(L_f) = E_1(L_f) = (0)\,, \\ E_2(L_f) = (f(t)) \quad \textit{and} \\ E_d(L_f) = (1)\,, \quad \textit{if} \quad d > 2\,. \end{cases}$$

REMARK. This kind of example was first considered in [1].

The construction of these two examples are closely related. They are also closely related to the construction of 2-spheres in S^4 in [2].

1. Let P be the family of all integral polynomials $f(t)$ which can be expressed in the following form:

$$\begin{aligned} t^{-(\varepsilon_1+\cdots+\varepsilon_n)}(1 - t^{\delta_1}) + t^{-(\varepsilon_2+\cdots+\varepsilon_n)}(1 - t^{\delta_2}) \\ + \cdots + t^{-\varepsilon_n}(1 - t^{\delta_n}) + 1\,, \end{aligned} \tag{1}$$

where $\varepsilon_i = \pm 1$ and $\delta_i = \varepsilon_i$ or $\delta_i = 0$ for $i = 1, 2, \cdots, n$. We assume that $1 \in P$.

LEMMA. *We have $f(t) \in P$, if and only if $f(1) = 1$.*

Proof. If $f(t) \in P$, then clearly we have $f(1) = 1$. Suppose that $f(1) = 1$. Then we have

$$\begin{aligned} f(t) - 1 &= (1 - t)(a_m t^m + \cdots + a_0) \\ &\quad - (1 - t)(b_m t^m + \cdots + b_0) \\ &= (1 - t)(a_m t^m + \cdots + a_0) \\ &\quad + (1 - t^{-1})(b_m t^{m+1} + \cdots + b_0 t)\,, \end{aligned}$$

where $a_i, b_i \geqq 0$ for $i = 1, 2, \cdots, n$. This means that $f(t)$ with $f(1) = 1$ can be obtained from 1 by applying a finite number of operation:

$$g(t) \longrightarrow g(t) + t^p(1 - t^\delta)\,,$$

where $p \geqq 0$ and $\delta = \pm 1$.

We assume $1 \in P$. Hence we should prove that if $f(t) \in P$, then $f(t) + t^p(1 - t^\delta) \in P$. Suppose that $f(t)$ has form (1). Now let

$$p = -(\varepsilon'_1 + \cdots + \varepsilon'_k + \varepsilon'_{k+1} + \cdots + \varepsilon'_{k+n}),$$

where $\varepsilon'_{k+i} = \varepsilon_i$ for $i = 1, 2, \cdots, n$ and let

$$\delta'_1 = \delta,\ \delta'_2 = \cdots = \delta'_k = 0 \quad \text{and} \quad \delta'_{k+i} = \delta_i$$

for $i = 1, 2, \cdots, n$. Then clearly we have

$$t^{-(\varepsilon'_1 + \cdots + \varepsilon'_k + \varepsilon'_{k+1} + \cdots + \varepsilon'_{k+n})}(1 - t^{\delta'_1}) + \cdots + t^{\varepsilon'_{k+n}}(1 - t^{\delta'_{k+n}}) = t^p(1 - t^\delta) + f(t).$$

Hence the proof is complete.

2. Let $f(t)$ be an integral polynomial with $f(1) = 1$. Suppose that $f(t)$ is expressed as (1). Now we construct a 1-dimensional polyhedron K_f in $E^3 (\subset S^3)$ as follows: The left-most side of K_f is shown in Fig. 1. Then for each i ($i = 1, \cdots, n$) we add step by step one of the four figures in Fig. 2. This depends on values of ε_i and

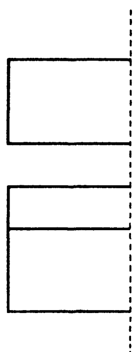

Fig. 1.

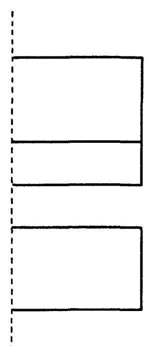

Fig. 3.

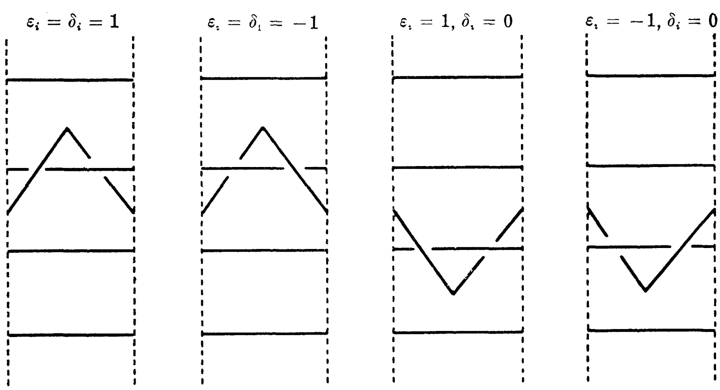

Fig. 2.

δ_i as in Fig. 2. The right-most side of K_f is shown in Fig. 3.

Now we give a presentation of the fundamental group of $E^3 - K_f$ (and that of $S^3 - K_f$, too). We use the Wirtinger presentation. If $a_0, \cdots, a_n, c_0, \cdots, c_m, d_0, \cdots, d_{m'}$ $(m + m' = n)$ are paths in Fig. 4, and

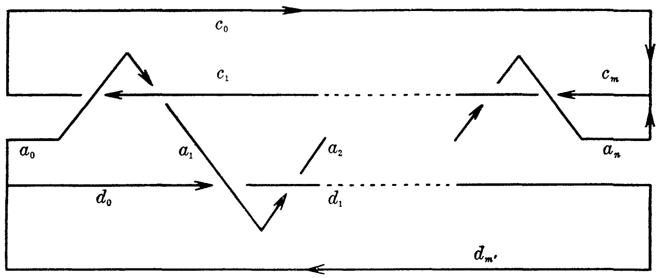

Fig. 4.

also, as usual, the paths which represent elements of the fundamental group in question, then the presentation is given as follows:

Generators: $\begin{cases} a_0, \cdots, a_n, \\ c_0, \cdots, c_m, \\ d_0, \cdots, d_{m'} (m + m' = n) . \end{cases}$

Relations:

(i) If $\varepsilon_i = 1, \delta_i = 1$, then

$$\begin{cases} c_{j-1} = a_{i-1} c_j a_{i-1}^{-1} , \\ a_i = c_j a_{i-1} c_j^{-1} , \end{cases}$$

(ii) If $\varepsilon_i = -1, \delta_i = -1$, then

$$\begin{cases} c_j = a_i c_{j-1} a_i^{-1} , \\ a_{i-1} = c_{j-1} a_i c_{j-1}^{-1} , \end{cases}$$

(iii) If $\varepsilon_i = 1, \delta_i = 0$, then

$$\begin{cases} d_j = a_{i-1} d_{j-1} a_{i-1}^{-1} , \\ a_i = d_j a_{i-1} d_j^{-1} , \end{cases}$$

(iv) If $\varepsilon_i = -1, \delta_i = 0$, then

$$\begin{cases} a_{i-1} = d_{j-1} a_i d_{j-1}^{-1} , \\ d_{j-1} = a_i d_j a_i^{-1} , \end{cases}$$

for each $i = 1, \cdots, n$, and

$$c_0 c_m^{-1} a_n = 1 .$$

3. Let k_f be a 1-cycle on K_f such that

$$\begin{cases} \operatorname{link}(a_i, k_f) = 0\,, & \text{for} \quad i = 0, 1, \cdots, n\,, \\ \operatorname{link}(c_i, k_f) = 1\,, & \text{for} \quad i = 0, 1, \cdots, m\,, \\ \operatorname{link}(d_i, k_f) = 1\,, & \text{for} \quad i = 0, 1, \cdots, m'\,. \end{cases}$$

We consider the elementary ideals of k_f on K_f in S^3. For each pair a_{i-1} and a_i the corresponding two rows in the Alexander matrix are elementary equivalent to the following:

(1) If $\varepsilon_i = \delta_i$, then

$$\begin{array}{cccc} a_{i-1} & a_i & c_{j-1} & c_j \end{array}$$
$$\begin{bmatrix} 1 - t^{\varepsilon_i} & 0 & -1 & 1 \\ t^{\varepsilon_i} & -1 & 0 & 0 \end{bmatrix}.$$

(2) If $\delta_i = 0$, then

$$\begin{array}{cccc} a_{i-1} & a_i & d_{j-1} & d_j \end{array}$$
$$\begin{bmatrix} 1 - t^{\varepsilon_i} & 0 & 1 & -1 \\ t^{\varepsilon_i} & -1 & 0 & 0 \end{bmatrix}.$$

From the last relation we have the following entries to the Alexander matrix.

$$\begin{array}{ccc} a_n & c_0 & c_m \\ [1 & 1 & -1] \end{array}$$

Hence we have matrix (*) as an Alexander matrix of k_f on K_f in S^3. Matrix (*) is elementary equivalent to (**). Note that we add a suitable number of rows of zeros. Matrix (**) can be reduced to (***) by elementary operations. Now it is easy to see that

$$(*) \quad \left(\begin{array}{c|c|c}
\begin{matrix} t^{\varepsilon_1} & -1 & & \\ & \ddots & \ddots & \\ & & t^{\varepsilon_n} & -1 \end{matrix} & 0 & 0 \\ \hline
\begin{matrix} & & & \;\;1 \end{matrix} & \begin{matrix} 1 & & -1 \end{matrix} & 0 \\ \hline
1 - t^{\varepsilon_i} & \begin{matrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{matrix} & 0 \\ \hline
1 - t^{\varepsilon_i} & 0 & \begin{matrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \end{matrix}
\end{array}\right) \begin{array}{l} \\ \\ \\ \\ \\ \varepsilon_i = \delta_i \\ \\ \\ \delta_i = 0 \end{array}.$$

$$(**)\quad \left(\begin{array}{ccccc|cc|c}
t^{\varepsilon_1} & -1 & & & & & & \\
 & \ddots & \ddots & & & & 0 & 0 \\
 & & t^{\varepsilon_n} & -1 & & & & \\
\hline
1-t^{\delta_1} & \cdots & 1-t^{\delta_n} & 1 & & & 0 & 0 \\
\hline
\ddots & & & & & \ddots & & \\
\cdots & 1-t^{\delta_i} & \cdots & & & \begin{Bmatrix} -1 & 1 \\ 0 & 0 \end{Bmatrix} & \cdots & 0 \\
 & & \ddots & & & & \ddots &
\end{array}\right) \begin{Bmatrix} \varepsilon_i = \gamma_i \\ \delta_i = 0 \end{Bmatrix}.$$

$$(***)\quad \left(\begin{array}{cccc|c|c}
t^{\varepsilon_1} & -1 & & & & \\
 & \ddots & \ddots & & 0 & 0 \\
 & & t^{\varepsilon_n} & -1 & & \\
\hline
1-t^{\delta_1} & \cdots & 1-t^{\delta_n} & 1 & 0 & 0
\end{array}\right).$$

$$\begin{cases} E_0(k_f) = E_1(k_f) = (0)\,, \\ E_2(k_f) = (t^{-(\varepsilon_1+\cdots+\varepsilon_n)}(1 - t^{\delta}1) + \cdots \\ \qquad\qquad + t^{-\varepsilon_n}(1 - t^{\delta}n) + 1) = (f(t))\,, \\ E_d(k_f) = (1)\,, \quad \text{if} \quad d > 2\,. \end{cases}$$

4. *Proof of Theorem* 1. Let $f(t)$ with $f(1) = 1$ be given. First construct K_f in S^3 and k_f on K_f as in 2 and 3. The construction of the corresponding θ-curve L_f is shown in Fig. 5. The 1-cycle l_f on

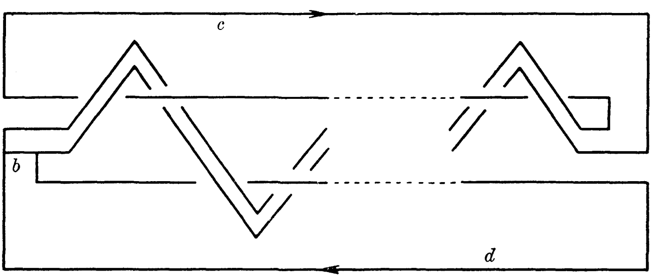

Fig. 5.

L_f has coefficient 1 on the oriented arc c and on the oriented arc d, respectively, and coefficient 0 on the arc b. It is easy to see that

$\pi(S^3 - L_f)$ is isomorphic to $\pi(S^3 - K_f)$ and $E_d(l_f) = E_d(k_f)$ for every nonnegative integer d.

REMARK. It is proved in [3] that if l is a 1-cycle on a θ-curve L in S^3, then we have

$$\begin{cases} E_0(l) = E_1(l) = (0)\,, & \text{and} \\ (E_d(l))^\circ = (1)\,, & \text{if} \quad d \geqq 2\,, \end{cases}$$

where $\circ$ is a trivializer (i.e., the operation to let $t = 1$ in $E_d(l)(t)$).

5. *Proof of Theorem* 2. Let $f(t)$ with $f(1) = 1$ be given. First construct K_f in S^3 and k_f on K_f as in 2 and 3. Then construct the corresponding two arcs C and D in E^3_+ as in Fig. 6, where

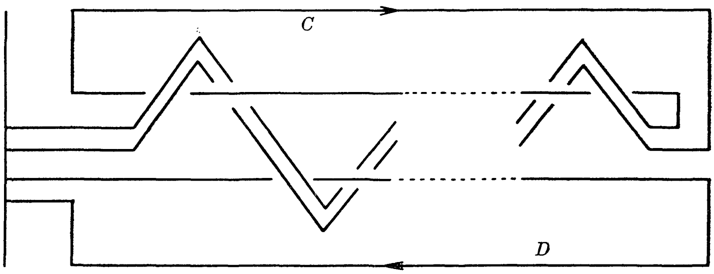

Fig. 6.

$$E^3_+ = \{(x_1, x_2, x_3) \mid x_1 \geqq 0\}\,.$$

Then the usual construction of the spinning of these arcs around the plane

$$\{(x_1, x_2, x_3, x_4) \mid x_1 = 0,\, x_4 = 0\}$$

gives rise to a 2-link L_f in S^4.

Now the arc C represents a trivial knot in E^3_+. A part of the step to see this is shown in Fig. 7. From this it follows that the 2-sphere S^2_C, which is the result of spinning C, is trivial in S^4. Clearly the same is true for the 2-sphere S^2_D, the result of spinning D.

We have

$$\pi(S^3 - K_f) \cong \pi(E^3_+ - C \cup D) \cong \pi(S^4 - L_f)\,,$$

and to find a 2-cycle l_f on L_f that corresponds to k_f on K_f is easy. Then we have

$$E_d(k_f) = E_d(l_f)$$

for every $d \geqq 0$. Hence the proof is complete.

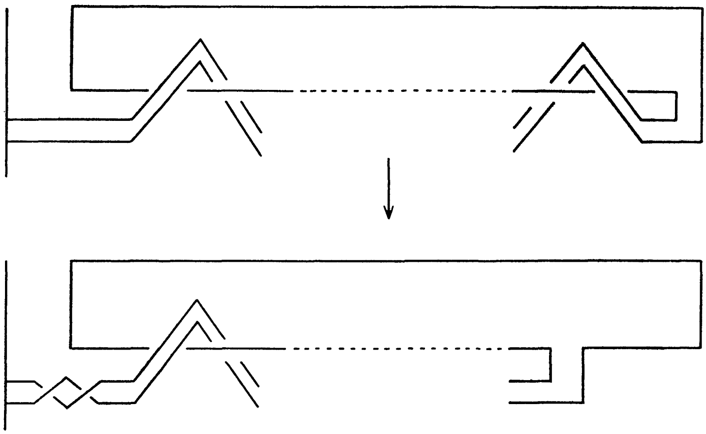

Fig. 7.

Proof of Corollary. We have $L_f = S_C^2 \cup S_D^2$ in S^4 in the example above. Then $l_f = l_c + l_d$, where l_c and l_d are fundamental cycles of S_C^2 and S_D^2, respectively. This completes the proof.

REMARK. Let L be a 2-link in S^4. Then it is known that for each 2-cycle l on L we always have

$$\begin{cases} E_0(l) = E_1(l) = 0 \,, \\ (E_d(l))^\circ = (1)\,, \quad \text{if} \quad d \geqq 2 \,, \end{cases}$$

where $\circ$ is a trivializer. (See [3] and [4].)

REFERENCES

1. E. H. Van Kampen, *Zur Isotopie zweidimensionaler Flaechen im* R^4, Abh. Math. Sem. Univ. Hamburg, **6** (1927), 216.
2. S. Kinoshita, *On the Alexander polynomials of* 2-*spheres in a* 4-*sphere*, Ann. of Math., **74** (1961), 518-531.
3. ———, *On elementary ideals of polyhedra in the* 3-*sphere*, to appear in the Pacific J. Math., **42** (1972), 89-98.
4. Y. Shinohara and D. W. Sumners, *Homology invariants of cyclic coverings with application to links*, Trans. Amer. Math. Soc., **163** (1972), 101-121.

Received August 3, 1972. The author of this paper is partially supported by NSF Grant GP-19964.

FLORIDA STATE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

CONVERGENCE OF BAIRE MEASURES

R. B. KIRK

Assume that there are no measurable cardinals. Then E. Granirer has proved that if a net $\{m_i\}$ of finite Baire measures on a completely regular Hausdorff space converges weakly to a finite Baire measure m, then $\{m_i\}$ converges to m uniformly on each uniformly bounded, equicontinuous subset of C^b, the space of bounded continuous functions. In this paper a relatively simple proof of Granirer's theorem is given based on a recent result of the author. The same method is used to prove the following analogue of Granirer's theorem. Let $\{m_i\}$ be a net of Baire measures on X each having compact support in the realcompactification of the underlying space X, and assume that $\int_X f dm_i \to \int_X f dm$ for every continuous function f on X where m is a Baire measure having compact support in the realcompactification of X. Then $\{m_i\}$ converges to m uniformly on each pointwise bounded, equicontinuous subset of C, the space of continuous functions on X. (The situation in the presence of measurable cardinals is also treated.)

In what follows, X will denote a completely regular Hausdorff space, C will denote the linear space of all continuous real-valued functions on X and C^b will denote the subspace of C consisting of all the uniformly bounded functions in C. The *Baire algebra* is the smallest σ-algebra on X with respect to which each of the functions in C is measurable. (Equivalently, it is the σ-algebra generated by the zero sets in X.) The linear space of all signed Baire measures on X with finite variation is denoted by M_σ, and the set of nonnegative elements in M_σ (i.e., the set of finite Baire measures) is denoted by M_σ^+. The space M_σ and C^b may be paired in the sense of Bourbaki by the bilinear form $\langle m, f\rangle = \int_X f dm = \int_X f dm^+ - \int_X f dm^-$ for all $m \in M_\sigma$ and all $f \in C^b$. By the *weak topology* on M_σ, will we mean the topology $\sigma(M_\sigma, C^b)$.

Let νX denote the realcompactification of X. (See [2], p. 116.) A Baire measure m on X is said to have *compact support in the realcompactification of* X if there is a compact set $G \subset \nu X$ such that for every zero set Z in νX with $G \subset Z$, it follows that $m(X \cap Z) = m(X)$. Let M_c denote the subspace of M_σ consisting of those elements whose total variations have compact support in the realcompactification of X. The set of nonnegative elements of M_c is denoted by M_c^+. It is not hard to verify that if $m \in M_c^+$, then $C \subset L^1(m)$. Hence the spaces M_c and C may be paired in the sense of Bourbaki by the bilinear form $\langle m, f\rangle = \int_X f dm = \int_X f dm^+ - \int_X f dm^-$ for all $m \in M_c$ and all $f \in C$. By the *weak topology* on M_c, we will mean the topology $\sigma(M_c, C)$.

Let B be a subset of C. Then B is *pointwise bounded* if for every $x \in X$, $\sup\{|f(x)| : f \in B\} < \infty$. It is said to be *uniformly bounded*, if $\sup\{\|f\|_X : f \in B\} < \infty$ where $\|f\|_X = \sup\{|f(x)| : x \in X\}$. (Of course, if B is uniformly bounded, then $B \subset C^b$.) The set B is *equicontinuous* (or *locally equicontinuous*) if for every $x \in X$ and for every positive number ε, there is a neighborhood U of x such that for all $y \in U$ and all $f \in B$, $|f(x) - f(y)| \leqq \varepsilon$. Let $\mathscr{E}$ denote the family of all pointwise bounded, equicontinuous subsets of C; and let $\mathscr{E}^b$ denote the family of all uniformly bounded, equicontinuous subsets of C^b. It is clear that if $B \in \mathscr{E}^b$, then B is a $\sigma(C^b, M_\sigma)$-bounded and that $C^b = \bigcup\{B : B \in \mathscr{E}^b\}$. Hence it follows that the topology e^b of uniform convergence on the sets in $\mathscr{E}^b$ is a locally convex topology on M_σ which is compatible with the pair (M_σ, C^b). (See [7], p. 255.) It is also the case that if $B \in \mathscr{E}$, then B is a $\sigma(C, M_c)$-bounded subset of C. (This fact is proved in Proposition 2.2 below.) Since $C = \bigcup\{B : B \in \mathscr{E}\}$, it follows that the topology e of uniform convergence on the sets in $\mathscr{E}$ is a locally convex topology on M_c compatible with the pair (M_c, C).

Recall that a set Y has a *measurable cardinal* if there is a probability measure defined on the algebra of all subsets of Y which is zero on all singleton sets. Otherwise, Y is said to have a *nonmeasurable cardinal*. It is consistent with the standard axiomatic treatments of set theory to assume that all sets have nonmeasurable cardinals. It is also known that if the continuum hypothesis holds, then the continuum has a nonmeasurable cardinal. It is not known whether or not the statement that there are no measurable cardinals is independent of the axioms of set theory.

The completely regular Hausdorff space X is a *D-space* if whenever d is a continuous pseudometric on X and Y is a d-discrete subset of X, then Y has a nonmeasurable cardinal. The concept of a D-space was introduced by Granirer in [3]. From the remarks made above about measurable cardinals, it is clearly consistent with the usual axioms of set theory to assume that every completely regular Hausdorff space is a D-space. The following result is proved by Granirer. (See [3], Theorem 2.)

THEOREM A. *Let X be a completely regular Hausdorff space. Then X is a D-space if and only if whenever $\{m_i\}$ is a net in M_σ^+ which converges weakly to $m \in M_\sigma$, then $\{m_i\}$ converges to m for the topology e^b.*

We will present a relatively simple proof of this theorem based on Theorem 1.1 below which was recently obtained by the author. (In fact, our main result, Theorem 1.5, is somewhat stronger than Theorem A.) The advantage of our method is that it allows the analysis to be carried out for nets of measures with finite support and reduces the measure theory needed to a minimum. The same method yields a proof of the following result which we believe to be new.

THEOREM B. *Let X be completely regular Hausdorff. Then the following hold.*

1. *If X is a D-space, then whenever $\{m_i\}$ is a net in M_c^+ which converges weakly to m in M_c, it follows that $\{m_i\}$ converges to m for the topology e.*

2. *Assume the continuum hypothesis. If X is not a D-space, then there is a net $\{m_i\}$ in M_c^+ which converges weakly to some m in M_c such that $\{m_i\}$ is not convergent for the topology e.*

1. **Weak convergence in M_σ.** Let L denote the subspace of M_σ consisting of those elements whose total variations have finite support. Hence $m \in L$ if and only if there is a finite set $A \subset X$ such that $m(B) = 0$ for every Baire set B disjoint from A. Every element $m \in L$ has a unique extension to a finite signed measure $\bar{m}$ on the algebra of subsets of X. For each $m \in L$, let ξ be the real-valued function defined by $\xi(x) = \bar{m}(\{x\})$ for all $x \in X$. In this way the space L may be identified with the space of all real-valued functions on X which vanish on the complement of a finite subset of X. We will use this representation of L throughout the paper. For notational purposes, we will use ξ to denote a generic element of L. The restriction to L of the bilinear form pairing M_σ and C^b is given by $\langle \xi, f \rangle = \Sigma\{\xi(x)f(x): x \in X\}$ for all $\xi \in L$ and all $f \in C^b$. The set of non-negative functions in L will be denoted by L^+.

A Baire measure m on X is said to be *separable* if for every continuous pseudometric d on X, there is a d-closed set $Z \subset X$ such that $m(X - Z) = 0$ and such that Z is d-separable. (Since every d-closed set is a zero set in X, it follows that $m(X - Z)$ is defined.) An arbitrary element of M is separable if its total-variation is separable. Let M_s denote the subspace of M_σ consisting of the separable elements of M_σ. The space M_s was first introduced by Dudley in [1]. It can be shown that X is a D-space if and only if $M_s = M_\sigma$. (Indeed, if X is a D-space, then $M_s = M_\sigma$ is a consequence of Theorem III, p. 137 in [8]. On the other hand, if X is not a D-space, then there is a continuous pseudometric d on X and a d-discrete set $Y \subset X$ such that Y has a measurable cardinal. If μ is a nontrivial measure on the subsets of Y, let m be defined by $m(B) = \mu(B \cap Y)$ for every Baire set B in X. It is clear that $m \in M_\sigma$. However, m is not d-separable as can easily be seen so that $m \notin M_s$.) Hence, again it is consistent with the axioms of set theory to assume that $M_s = M_\sigma$ for all completely regular Hausdorff spaces. The following result was proved by the author in [6].

THEOREM 1.1. *Let X be a completely regular Hausdorff space and let M_s be equipped with the topology e^b of uniform convergence on the uniformly bounded, equicontinuous subsets of C^b. Then the following hold.*

1. *M_s is complete.*
2. *L is dense in M_s.*
3. *The dual space of M_s is C^b.*

We will require several results from the theory of measures on a topological space which we will now review briefly. (The reader is referred to [9] for further details.) Recall that a Baire measure m is *τ-additive* if whenever $\{Z_i: i \in I\}$ is a downward directed system of zero sets in X with $\bigcap \{Z_i: i \in I\} = \emptyset$, then $m(Z_i) \to 0$. (The family $\{Z_i: i \in I\}$ is *downward directed* if for each pair $i_1, i_2 \in I$, there is $i_3 \in I$ such that $Z_{i_3} \subset Z_{i_1} \cap Z_{i_2}$.) Equivalently, m is net-additive if for each upward directed system $\{U_i: i \in I\}$ of cozero sets (complements of zero sets) in X with $\bigcup \{U_i: i \in I\} = X$, then $m(U_i) \to m(X)$. The *support* of a Baire measure m is the set $\operatorname{supp} m = \bigcap \{Z: Z$ is a zero set in X and $m(X) = m(Z)\}$. If $\operatorname{supp} m = \emptyset$, then m is said to be *entirely without support*. The following result is proved in [5].

THEOREM 1.2. *Let m be a Baire measure on X. If m is not net-additive, then there is a Baire measure m' on X such that $0 < m' \leqq m$ and such that m' is entirely without support.*

If d is a continuous pseudometric on X, define an equivalence relation on X by $x \equiv y$ if $d(x, y) = 0$; and let X^* denote the set of equivalence classes. For $\bar{x}, \bar{y} \in X^*$, define $d^*(\bar{x}, \bar{y}) = d(x, y)$. Then (X^*, d^*) is a metric space which we will call the *metric space associated with d*. Let $Q: X \to X^*$ be the quotient map. Since Q is continuous, it follows that $Q^{-1}[B]$ is a Baire set in X whenever B is a Baire set in X^*. If m is a Baire measure on X, define $\bar{m}(B) = m(Q^{-1}[B])$ for every Baire set in X^*. Then $\bar{m}$ is a Baire measure on X^*. The following lemma is a consequence of Theorem 28 and Remark 4, p. 175 of Varadarajan in [9]. However, since the proof given below is essentially different, we will include it for the sake of completeness.

LEMMA 1.3. *Let d be a continuous pseudometric on X, and let*

m be a separable Baire measure on X. If $\{U_i: i \in I\}$ is a cover of X by d-open sets and if ε is an arbitrary positive number, then there is a finite set $\{i_1, \cdots, i_n\} \subset I$ such that $m(X - \bigcup_{k=1}^n U_{i_k}) \leqq \varepsilon$.

Proof. Let (X^*, d^*) be the metric space associated with d, and let $\bar{m}$ be the Baire measure on X^* corresponding to m. It will be sufficient to prove that $\bar{m}$ is net-additive on X^*. Indeed, assume that $\bar{m}$ is net-additive. Since U_i is d-open, $\bar{U}_i = Q[U_i]$ is open in X^*; and hence it is a cozero set in X^*. The family of all finite unions of the sets in $\{\bar{U}_i: i \in I\}$ is then an upward directed family of cozero sets whose union is X^*. But then there is a finite set $\{i_1, \cdots, i_n\} \subset I$ such that $m(X - \bigcup_{k=1}^n U_{i_k}) = \bar{m}(X^* - \bigcup_{k=1}^n \bar{U}_{i_k}) \leqq \varepsilon$ since $\bar{m}$ is net-additive.

We will now show that $\bar{m}$ is net-additive. If this is not the case, then by Theorem 1.2 there is a Baire measure μ on X^* such that $0 < \mu \leqq \bar{m}$ and such that μ is entirely without support. Then there is a separable Baire measure m_0 on X such that $m_0 \leqq m$ and such that $\bar{m}_0 = \mu$. Indeed, let $E = \{f \in C^b: f = f^* \circ Q$ for some $f^* \in C^b(X^*)\}$; and define $\varphi(f) = \int_{X^*} f^* d\mu$ for each $f \in E$ where $f = f^* \circ Q$. Then φ^* is a linear functional on the linear space E. Furthermore, φ^* is majorized on E by the subadditive functional p defined on C^b by $p(f) = \int_X f^+ dm$ for all $f \in C^b$. Hence by the Hahn-Banach theorem, there is a linear functional φ on C^b which extends φ^* and which is majorized by p on C^b. It is not difficult to verify that φ is nonnegative and satisfies the integral property. (A nonnegative functional φ on C^b satisfies the *integral property* if for every decreasing sequence $\{f_n\} \subset C^b$ such that $f_n \downarrow 0$ pointwise, it follows that $\varphi(f_n) \downarrow 0$.) It follows by the Alexandrov representation theorem (see Theorems 1.2 and 1.5 in [5]) that there is a Baire measure m_0 on X such that $\varphi(f) = \int_X f dm_0$ for all $f \in C^b$. It is clear that $m_0 \leqq m$ and that $\bar{m}_0 = \mu$ as claimed. (Note that since m is separable and since $m_0 \leqq m$, it follows that m_0 is also separable.)

Since $\bar{m}_0$ is entirely without support in X^*, there is for each $\bar{x} \in X^*$ an open set $U_{\bar{x}}$ in X^* with $\bar{m}_0(U_{\bar{x}}) = 0$. Since $\{U_{\bar{x}}: \bar{x} \in X^*\}$ is an open cover of X^* and since X^* is paracompact (being a metric space), there is a partition of unity $\{f_j^*: j \in J\}$ subordinate to the cover $\{U_{\bar{x}}: x \in X^*\}$. For each finite set $\tau \subset J$, define $f_\tau = \sum \{f_j^* \circ Q: j \in \tau\}$. Then $\{f_\tau\}$ is easily seen to be uniformly bounded and equicontinuous. Since the net $\{f_\tau\}$ converges to 1 pointwise, hence by Proposition 9.2 in [6], $\int_X f_\tau dm_0 \to m_0(X)$. On the other hand, since f_j^* has its support in $U_{\bar{x}}$ for some $\bar{x} \in X^*$, it follows that $\int_X f_j^* \circ Q dm_0 = \int_{X^*} f_j^* d\bar{m}_0 = 0$.

Thus $\int_X f_\tau dm_0 = 0$ for all τ. Thus $m_0(X) = \lim \int_X f_\tau dm_0 = 0$. This contradicts the fact that $m_0(X) = \bar{m}_0(X^*) = \mu(X^*) > 0$. The proof is complete.

We remark here that Lemma 1.3 is the only result from the theory of measures in a topological space which will be required in proof of Theorem 1.5 (the main result in this section). This theorem is somewhat stronger than Theorem A. A proof of Theorem A itself can be based on a result of Marczewski and Sikorski ([8], Theorem III) without reference to Lemma 1.3. (This result of Marczewski and Sikorski is also used by Granirer in his proof of Theorem B.)

For $\xi \in L$ and $W \subset X$, define the element $(\xi)_W \in L$ by $(\xi)_W(x) = \xi(x)$ for $x \in W$ and $(\xi)_W(x) = 0$ for $x \in X - W$. (That is, $(\xi)_W = \xi \cdot \mathscr{X}_W$ where $\mathscr{X}_W$ is the characteristic function of the set W. We can now prove the following.

PROPOSITION 1.4. *Let X be a completely regular Hausdorff space, and let $\{\xi_i: i \in I\}$ be a net in L^+. Assume that $\{\xi_i\}$ converges to $m \in M_s$ in the $\sigma(M_s, C^b)$-sense. Then $\{\xi_i\}$ converges to m in the e^b-sense.*

Proof. We will show that $\{\xi_i\}$ is an e^b-Cauchy net. The result will then be immediate from Theorem 1.1. Assume without loss of generality that $m \neq 0$. Fix a set $B \in \mathscr{E}^b$ and a positive number ε. For $x, y \in X$, define $d(x, y) = \sup\{|f(x) - f(y)|: f \subset B\}$. Then it is easily verified that since $B \in \mathscr{E}^b$, d is a continuous pseudometric on X. Since the net $\{\langle \xi_i, 1\rangle\}$ converges to $\langle m, 1\rangle$, we may assume without loss of generality that $P = \sup\{|\langle \xi_i, 1\rangle|: i \in I\}$ is finite. Let $M = \sup\{\|f\|_X: f \in B\}$ which is also finite since B is uniformly bounded.

Since d is continuous, there is for each $x \in X$ a d-open set U_x such that $d(x, y) \leqq \varepsilon P^{-1}$ for all $y \in U_x$. In particular, we then have,

$$(1) \qquad |f(x) - f(y)| \leqq \varepsilon P^{-1} \quad \text{for all} \quad y \in U_x, f \in B\,.$$

By Lemma 1.3 there is a finite set $\{x_1, \cdots, x_n\} \subset X$ such that $m(X - U) \leqq \varepsilon M^{-1}$ where $U = \bigcup_{k=1}^n U_{x_k}$. (Note that each U_x is a cozero set in X so that U is also a cozero set.) Since m is regular, there is a zero set Z in X such that $Z \subset U$ and such that $m(U - Z) \leqq \varepsilon M^{-1}$. Let $f_0 \in C^b$ be such that $0 \leqq f_0 \leqq 1, f_0 = 1$ on $X - U$ and $f_0 = 0$ on Z. Since $\{\xi_i\}$ converges to m weakly, there is an $i_1 \in I$ such that if $i \leqq i_1$, then $|\langle \xi_i - m, f_0\rangle| \leqq \varepsilon M^{-1}$. Since $\xi_i \geqq 0$, we have for $i \geqq i_1$,

$$\begin{aligned} 0 &\leqq \langle (\xi_i)_{X-U}, 1\rangle \leqq \langle \xi_i, f_0\rangle \\ &\leqq \langle \xi_i - m, f_0\rangle + \int_{X-Z} f_0 dm \leqq \varepsilon M^{-1} + m(X - Z) \\ &\leqq \varepsilon M^{-1} + m(X - U) + m(X - Z) \leqq 3\varepsilon M^{-1}\,. \end{aligned}$$

Thus we have demonstrated the following inequality which we note for future reference.

$$(2) \qquad \langle (\xi_i)_{X-U}, 1\rangle \leqq 3\varepsilon M^{-1}, \quad \text{for all} \quad i \geqq i_1 .$$

The set of vectors $K = \{(f(x_1), \cdots, f(x_n)) : f \in B\}$ is a totally-bounded set in R^n. Hence there is a finite set $A \subset B$ such that the set $K_A = \{(f(x_1), \cdots, f(x_n)) : f \in A\}$ is an εP^{-1}-net for K. Since $\{\xi_i\}$ is weakly convergent and since A is finite, there is an $i_2 \in I$ such that if $i, j \geqq i_2$, then

$$(3) \qquad |\langle \xi_i - \xi_j, f\rangle| \leqq \varepsilon, \quad \text{for all} \quad f \in A .$$

Let $i_0 \in I$ be greater than both i_1, i_2. Fix $i, j \geqq i_0$ and let $f \in B$. Choose $f^* \in A$ such that $|f(x_k) - f^*(x_k)| \leqq \varepsilon P^{-1}$ for all $k = 1, \cdots, n$. We then have by (2) and (3) that

$$\begin{aligned} \langle \xi_i - \xi_j, f\rangle &= \langle \xi_i - \xi_j, f^*\rangle + \langle \xi_i - \xi_j, f - f^*\rangle \\ &\leqq \varepsilon + \langle (\xi_i - \xi_j)_{X-U}, |f - f^*|\rangle + \langle (\xi_i - \xi_j)_U, |f - f^*|\rangle \\ &\leqq \varepsilon + \langle (\xi_i)_{X-U}, |f| + |f^*|\rangle + \langle (\xi_j)_{X-U}, |f| + |f^*|\rangle \\ &\qquad + \langle (\xi_i - \xi_j)_U, f - f^*\rangle \\ &\leqq \varepsilon + 2M\langle (\xi_i)_{X-U}, 1\rangle + 2M\langle (\xi_j)_{X-U}, 1\rangle \\ &\qquad + \langle (\xi_i - \xi_j)_U, f - f^*\rangle \\ &\leqq \varepsilon + 2M(3\varepsilon M^{-1}) + 2M(3\varepsilon M^{-1}) + \langle (\xi_i - \xi_j)_U, f - f^*\rangle \\ &\leqq 13\varepsilon + \langle (\xi_i - \xi_j)_U, f - f^*\rangle . \end{aligned}$$

Hence we have shown that,

$$(4) \quad \langle \xi_i - \xi_j, f\rangle \leqq 13\varepsilon + \langle (\xi_i - \xi_j)_U, f - f^*\rangle, \quad \text{for all} \quad i, j \geqq i_0 .$$

Let $U_k = U_{x_k}$ for $k = 1, \cdots, n$ and let $U_0 = \varnothing$. By (1) and the fact that $|f(x_k) - f^*(x_k)| \leqq \varepsilon P^{-1}$ for all $k = 1, \cdots, n$, we have for $i \geqq i_0$ that

$$\begin{aligned} &\langle (\xi_i)_U, |f - f^*|\rangle \\ &\quad = \sum_{k=1}^{n} \langle (\xi_i)_{U_k - U_{k-1}}, |f - f^*|\rangle \\ &\quad \leqq \sum_{k=1}^{n} \sum_{x \in U_k - U_{k-1}} \xi_i(x)\, |f(x) - f^*(x)| \\ &\quad \leqq \sum_{k=1}^{n} \sum_{x \in U_k - U_{k-1}} \xi_i(x)\{|f(x) - f(x_k)| + |f(x_k) - f^*(x_k)| \\ &\qquad + |f^*(x_k) - f^*(x)|\} \\ &\quad \leqq 3\varepsilon P^{-1} \sum_{k=1}^{n} \sum_{x \in U_k - U_{k-1}} \xi_i(x) \leqq 3\varepsilon P^{-1}\langle \xi_i, 1\rangle \leqq 3\varepsilon . \end{aligned}$$

That is, we have

$$(5) \qquad \langle (\xi_i)_U, f - f^* \rangle \leqq 3\varepsilon, \quad \text{for all} \quad i \geqq i_0 .$$

Combining (4) and (5), we obtain that for all $i, j \geqq i_0$ and all $f \in B$,

$$\langle \xi_i - \xi_j, f \rangle \leqq 13\varepsilon + \langle (\xi_i)_U, |f - f^*|\rangle + \langle (\xi_j)_U, |f - f^*|\rangle \leqq 19\varepsilon .$$

Since ε and B were arbitrary, it now follows that $\{\xi_i\}$ is an e^b-Cauchy net. The proof is complete.

THEOREM 1.5. *Let X be completely regular Hausdorff. Then the weak topology and the e^b-topology are identical on M_s^+.*

Proof. It is sufficient to show that if G is an e^b-closed set in M_s^+, then G is weakly closed in M_s^+. But this is immediate from Proposition 1.4 and the fact that L^+ is weakly dense in M_s^+. The proof is complete.

Theorem A now follows easily. Indeed, if X is a D-space, then $M_s = M_\sigma$ as noted above; and Theorem A reduces to Theorem 1.5. If X is not a D-space, then there is a Baire measure m with $m \in M_\sigma - M_s$. Since L^+ is weakly dense in M_σ, there is a net $\{\xi_i\}$ in L^+ which converges weakly to m. However, $\{\xi_i\}$ will not converge in the e^b-sense since otherwise $\{\xi_i\}$ would be an e^b-Cauchy net which would imply by Theorem 1.1 that $m \in M_s$.

In [3] Granirer proves the following as an application of Theorem A. It is also an immediate consequence of Theorem 1.1.

THEOREM 1.6. *Let X be a completely regular Hausdorff space. Then X is a D-space if and only if every uniformly bounded, equicontinuous subset of C^b is relatively $\sigma(M_\sigma, C^b)$-compact.*

Proof. If X is a D-space, then $M_\sigma = M_s$. Since by Theorem 1.1 the dual of M_σ with the topology e^b is C^b, it follows by the Banach-Alaoglu theorem that $B^{\circ\circ}$ is $\sigma(C^b, M_\sigma)$-compact whenever $B \in \mathscr{E}^b$. (Of course, $B^{\circ\circ}$ denotes the bipolar of B for the pair.) On the other hand, if X is not a D-space, then by Theorem 1.1, M_s is a proper closed subspace of M_σ for the topology e^b. Hence by the Hahn-Banach theorem, the dual space of M_σ for this topology is strictly larger than C^b. This implies by the Mackey-Arens theorem that there is a $B \in \mathscr{E}^b$ such that $B^{\circ\circ}$ is not $\sigma(C^b, M_\sigma)$-compact. But, as is easily verified, $B^{\circ\circ} \in \mathscr{E}^b$. This completes the proof.

2. Weak convergence in M_c. The following is proved in [6], Theorem 4.4. (The essence of the theorem is due to Hewitt in [4].)

THEOREM 2.1. *The order dual of C is isomorphic as a Riesz* *space to* M_c. *The isomorphism is given by* $\varphi \leftrightarrow m$ *where* $\varphi(f) = \int_X f dm$ *for all* $f \in C$. *In particular,* $C \subset L^1(m)$ *for all* $m \in M_c^+$.

We now prove the following as promised in the introduction.

PROPOSITION 2.2. *If* $B \in \mathscr{E}$, *then* B *is a* $\sigma(C, M_c)$*-bounded subset of* C.

Proof. Fix $m \in M_c^+$. It is sufficient to show that $\left\{\int_X |f|\, dm : f \in B\right\}$ is bounded. If this is not so, then there is a sequence $\{f_n\} \in B$ such that $\int_X |f_n|\, dm \to +\infty$. For each $n \in N$, define $g_n = \sup\{|f_k| : k = 1, \cdots, n\}$ and $g = \sup\{|f_k| : k \in N\}$. Then g is a real-valued, continuous function. Indeed, it is clear that g is real-valued since B is pointwise bounded. In order to see that g is continuous, fix $x \in X$ and $\varepsilon > 0$. Let U be a neighborhood of x such that $|f(x) - f(y)| \leqq \varepsilon/3$ for all $y \in U$ and all $f \in B$. We now claim that $|g(x) - g(y)| \leqq \varepsilon$ for all $y \in U$. Indeed, fix $y \in U$, and choose $k \in N$ so large that $|g(x) - g_k(x)| \leqq \varepsilon/3$ and $|g(y) - g_k(y)| \leqq \varepsilon/3$. Then there are $i, j \in \{1, \cdots, k\}$ such that $g_k(x) = |f_i(x)|$ and $g_k(y) = |f_j(y)|$. Hence we have that,

$$\begin{aligned}
&|g(x) - g(y)| \\
&\quad \leqq |g(x) - g_k(x)| + |g_k(x) - g_k(y)| + |g_k(y) - g(y)| \\
&\quad \leqq 2\varepsilon/3 + ||f_i|(x) - |f_j|(y)| \\
&\quad \leqq 2\varepsilon/3 + \max\{||f_i|(x) - |f_i|(y)|, ||f_j|(x) - |f_j|(y)|\} \leqq \varepsilon .
\end{aligned}$$

The proof is complete.

Define $M_{sc} = M_s \cap M_c$. If X is a D-space, then $M_{sc} = M_c$. On the other hand, if X is not a D-space, then for some continuous pseudometric d on X, there is a d-closed subset $Z \subset X$ with a measurable cardinal. It is known that if the continuum hypothesis holds and if Z has a measurable cardinal, then there is a probability measure on the algebra of all subsets of Z which is zero on all singletons and which assumes only the values 0 or 1. From this it follows that there is a point in νX such that the valuation functional on C corresponding to this point is represented (according to Theorem 2.1) by a nonseparable element of M_c. That is, M_{sc} is a proper subspace of M_c. In summary then, it follows that if the continuum hypothesis holds, then X is a D-space if and only if $M_{sc} = M_c$. The following result is proved in [6].

THEOREM 2.3. *Let* X *be completely regular Hausdorff, and let* M_{sc} *be equipped with the topology* e *of uniform convergence on the pointwise* *bounded, equicontinuous subsets of C. Then the following hold.*

1. *M_{sc} is complete.*
2. *The dual space of M_{sc} is C.*
3. *L is dense in M_{sc}.*

If X itself is realcompact, then obviously $M_{sc} = M_c$. Hence we have the following.

PROPOSITION 2.4. *Let X be realcompact, and let $\{\xi_i: i \in I\}$ be a net in L^+. If $\{\xi_i\}$ converges to $m \in M_c$ for the topology $\sigma(M_c, C)$, then $\{\xi_i\}$ converges to m for the topology e.*

Proof. We will show that $\{\xi_i\}$ is an e-Cauchy net. The result will then follow immediately from Theorem 2.3. Assume without loss of generality that $m \neq 0$. Fix a set $B \subset \mathscr{E}$ and a positive number ε. For $x, y \in X$, define $d(x, y) = \sup\{|f(x) - f(y)|: f \in B\}$. Since $B \in \mathscr{E}$, it follows that d is a continuous pseudometric on X. Let G be the support of m which is a compact subset of X by assumption. Let $M = \sup\{\|f\|_G: f \in B\}$ which is finite since $B \in \mathscr{E}$. For all $x \in X$, define $h(x) = d(x, G) = \inf\{d(x, y): y \in G\}$. Then h is an element of C. Since $\{\xi_i\}$ converges weakly to m, there is an $i_1 \in I$ such that $|\langle \xi_i - m, h\rangle| \leqq \varepsilon$ for all $i \geqq i_1$. But $h = 0$ on G so that $\langle m, h> = 0$. Hence we have that,

$$(1) \qquad |\langle \xi_i, h\rangle| \leqq \varepsilon, \quad \text{for all} \quad i \geqq i_1 .$$

Since the net $\{\langle \xi_i, 1\rangle\}$ converges to $\langle m, 1\rangle$, we may assume that $P = \sup\{|\langle \xi_i, 1\rangle|: i \in I\}$ is finite.

For each $x \in G$, let U_x be a cozero set neighborhood of x such that $|f(x) - f(y)| \leqq \varepsilon P^{-1}$ for all $y \in U_x$ and all $f \in B$. Since G is compact, there is a finite cover $\{U_{x_1}, \cdots, U_{x_n}\}$ of G. Define $U = U_{x_1} \cup \cdots \cup U_{x_n}$. The set of vectors $K = \{(f(x_1), \cdots, f(x_n)): f \in B\}$ is a totolly bounded subset of R^n. Let A be a finite subset of B such that the set $K_A = \{(f(x_1), \cdots, f(x_n)): f \in A\}$ is an εP^{-1}-net for K.

Since $\{\xi_i\}$ is weakly convergent and since A is finite, there is an $i_2 \in I$ such that,

$$(2) \qquad |\langle \xi_i - \xi_j, f\rangle| \leqq \varepsilon, \quad \text{for all} \quad i, j \geqq i_2 \quad \text{and all} \quad f \in A .$$

Finally, as in the proof of Proposition 1.4, there is an $i_3 \in I$ such that,

$$(3) \qquad \langle (\xi_i)_{X-U}, 1\rangle \leqq \varepsilon M^{-1}, \quad \text{for all} \quad i \geqq i_3 .$$

Now let $i_0 \in I$ be greater than i_1, i_2, and i_3. Fix $i, j \geqq i_0$ and let $f \in B$. Choose $f^* \in A$ such that $|f(x_k) - f^*(x_k)| \leqq \varepsilon P^{-1}$ for $k = 1, \cdots, n$. We then have from (2) that,

$$(4)\quad \begin{aligned}\langle \xi_i - \xi_j, f\rangle &= \langle \xi_i - \xi_j, f^*\rangle + \langle \xi_i - \xi_j, f - f^*\rangle \\ &\leqq \varepsilon + \langle (\xi_i - \xi_j)_{X-U}, |f - f^*|\rangle + \langle (\xi_i - \xi_j)_U, |f - f^*|\rangle .\end{aligned}$$

However, for $i \geqq i_0$, letting $U_k = U_{x_k}$ for $k = 1, \cdots, n$ and $U_0 = \varnothing$,

$$\begin{aligned}&\langle (\xi_i)_U, |f - f^*|\rangle \\ &= \sum_{k=1}^{n} \sum_{x \in U_k - U_{k-1}} \xi_i(x)\, |f(x) - f^*(x)| \\ &\leqq \sum_{k=1}^{n} \sum_{x \in U_k - U_{k-1}} \xi_i(x)\{|f(x) - f(x_k)| + |f(x_k) - f^*(x_k)| + |f^*(x_k) - f^*(x)|\} \\ &\leqq 3\varepsilon P^{-1} \sum_{k=1}^{n} \sum_{x \in U_k - U_{k-1}} \xi_i(x) \leqq 3\varepsilon P^{-1}\langle \xi_i, 1\rangle \leqq 3\varepsilon .\end{aligned}$$

Thus we have shown that,

$$(5)\qquad \langle (\xi_i)_U, |f - f^*|\rangle \leqq 3\varepsilon, \quad \text{for all} \quad i \geqq i_0 .$$

Note that if $f \in B$, then $|f| \leqq h + M$. Hence for $i \geqq i$, we have from (1) and (3), that

$$\begin{aligned}&\langle (\xi_i)_{X-U}, |f - f^*|\rangle \\ &\leqq 2\langle (\xi_i)_{X-U}, h + M\rangle \leqq 2\{\langle \xi_i, h\rangle + M\langle (\xi_i)_{X-U}, 1\rangle\} \\ &\leqq 2\{\varepsilon + M\varepsilon M^{-1}\} \leqq 4\varepsilon .\end{aligned}$$

Thus we have shown that,

$$(6)\qquad \langle (\xi_i)_{X-U}, |f - f^*| > \leqq 4\varepsilon, \quad \text{for all } i \geqq i_0 .$$

Combining (4), (5), and (6), we have for $i, j \geqq i_0$ that,

$$\begin{aligned}\langle \xi_i - \xi_j, f\rangle \leqq \varepsilon &+ \langle (\xi_i)_{X-U}, |f - f^*|\rangle + \langle (\xi_j)_{X-U}, |f - f^*|\rangle \\ &+ \langle (\xi_i)_U, |f - f^*|\rangle + |\langle (\xi_j)_{X-U}, |f - f^*|\rangle \leqq 15\varepsilon .\end{aligned}$$

It follows that $\{\xi_i\}$ is an e-Cauchy net as claimed. The proof is complete.

PROPOSITION 2.5. *Let X be a D-space. Then every continuous pseudometric on X has a (unique) extension to a continuous pseudometric on νX.*

Proof. Let $\tilde{X}$ denote the completion of X for the finest uniform structure on X compatible with the topology on X. Denote this structure by $\mathscr{U}^f$. Then every continuous pseudometric on X has a unique extension to $\tilde{X}$ since the set of all such pseudometrics is a gauge for this uniformity. The proof will be complete if we show that $\tilde{X} = \nu X$. But since every continuous real-valued function on X is $\mathscr{U}^f$-uniformly continuous, it follows that X is C-embedded in $\tilde{X}$.

Hence $\nu\tilde{X} = \nu X$ (by [2], Theorem 8.6). If we can show that $\tilde{X}$ is a D-space, then by Shirota's theorem ([2], p. 229), it will follow that $\tilde{X} = \nu\tilde{X}$; and the proof will be complete.

Assume that $\tilde{X}$ is not a D-space. Then there is a continuous pseudometric $\tilde{d}$ on $\tilde{X}$ and a $\tilde{d}$-closed, discrete subset $\tilde{Z}$ of $\tilde{X}$ which has a measurable cardinal. Let d denote the restriction of $\tilde{d}$ to X. For each $x \in \tilde{Z}$, define $0 < \alpha(x) = \inf\{\tilde{d}(x, y): y \in \tilde{Z}$ and $x \neq y\}$. Since X is dense in $\bar{X}$, there is for each point $x \in \tilde{Z}$ a point $\psi(x) \in X$ such that $\tilde{d}(x, \psi(x)) \leqq \alpha(x)/3$. (Such a function exists by the axiom of choice.) Then the set $Z = \{\psi(x): x \in \tilde{Z}\}$ is a d-discrete subset of X. Since ψ is clearly one-to-one, Z also has a measurable cardinal. But this contradicts the assumption that X is a D-space. The proof is complete.

We note that the fact $\bar{X}$ is a D-space in the above proof is a special case of Remark 2, p. 11 in [3]. For $f \in C$, let $\bar{f}$ denote the unique continuous extension of f to νX. If B is a subset of C, let $\bar{B} = \{\bar{f}: f \in B\}$. We then have the following.

PROPOSITION 2.6. *Let X be a D-space. If B is a pointwise bounded and equicontinuous subset of $C(X)$, then $\bar{B}$ is a pointwise bounded and equicontinuous subset of $C(\nu X)$.*

Proof. For each pair $x, y \in X$, define $d(x, y) = \sup\{|f(x) - f(y)|: f \in B\}$. Since B is pointwise bounded and equicontinuous on X, it follows that d is a continuous pseudometric on X. By Proposition 2.5 there is an unique continuous extension $\tilde{d}$ of d to νX. It then follows that for all $x, y \in \nu X$ and for all $f \in B$, $|\bar{f}(x) - \bar{f}(y)| \leqq \tilde{d}(x, y)$. But this implies that $\bar{B}$ is equicontinuous and pointwise bounded on νX. The proof is complete.

THEOREM 2.7. *Let X be a D-space, and let $\{m_i\}$ be a net in M_c^+. If $\{m_i\}$ converges weakly to $m \in M_c$, then $\{m_i\}$ converges to m for the topology e.*

Proof. Since L^+ is weakly dense in M_c, it is sufficient to show that if $\{\xi_i\}$ is a net in L^+ which converges weakly to $m \in M_c$, then $\{\xi_i\}$ converges for the topology e. Hence fix $B \in \mathscr{E}$. For $f \in C$, let $\bar{f}$ be its extension to νX; and let $\bar{B} = \{\bar{f}: f \in B\}$ as above. For each $m \in M_c(X)$ and for each $f \in C$, define $\varphi(\bar{f}) = \int_X f dm$. Then by Theorem 2.1, there is an $\bar{m} \in M_c(\nu X)$ such that $\bar{\varphi}(\bar{f}) = \int_{\nu X} \bar{f} d\bar{m}$ for all $f \in C(X)$. Since $\{\xi_i\}$ converges to m for the $\sigma(M_c(X), C(X))$ topology, it follows that $\{\bar{\xi}_i\}$ converges to $\bar{m}$ for the $\sigma(M_c(\nu X), C(\nu X))$ topology. Since B is pointwise bounded and equicontinuous on X, it follows by Proposition 2.6 that $\bar{B}$ is pointwise bounded and equicontinuous on νX. Since νX is realcompact, it follows from Proposition 2.4 that $\{\bar{\xi}_i\}$ converges to $\bar{m}$ uniformly over $\bar{B}$. But it is then immediate that $\{\xi_i\}$ converges to m uniformly over B. The proof is complete.

Theorem B now follows easily. Indeed, if X is a D-space, then it reduces to Theorem 2.7. On the other hand, assume that X is not a D-space. As we have noted above, if the continuum hypothesis holds, it follows that M_{sc} is a proper subspace of M_c. Let $m \in M_c^+ - M_{sc}^+$. Since L^+ is weakly dense in M_c^+, there is a net $\{\xi_i\}$ in L^+ which converges weakly to m. However, by Theorem 2.3, M_{sc} is complete for the topology e so that $\{\xi_i\}$ does not converge for the topology e. The proof is complete.

We can also prove the following analogue of Theorem 1.6 (Granirer's Theorem 1).

THEOREM 2.8. *Let X be completely regular Hausdorff. Then the following hold.*

1. *If X is a D-space, then every pointwise bounded, equicontinuous subset of C is relatively $\sigma(C, M_c)$-compact.*

2. *Assume the continuum hypothesis. If X is not a D-space, then there is a pointwise bounded, equicontinuous subset of C which is not relatively $\sigma(C, M_c)$-compact.*

Proof. 1. If X is a D-space, $M_{sc} = M_c$. Hence $B^{\circ\circ}$ is $\sigma(C, M_c)$-compact for every $B \in \mathscr{E}$ by Theorem 2.3 and the Banach-Alaoglu theorem.

2. If X is not a D-space, then the continuum hypothesis implies that M_{sc} is a proper subspace of M_c. By Theorem 2.3, M_{sc} is a closed subspace for the topology e. It follows by the Hahn-Banach theorem that the dual space of M_c for the topology e is then strictly larger than C. Hence by the Mackey-Arens theorem, there is a $B \in \mathscr{E}$ for which $B^{\circ\circ}$ is not $\sigma(C, M_c)$-compact. But, as is easily verified, if $B \in \mathscr{E}$, then $B^{\circ\circ} \in \mathscr{E}$. The proof is complete.

REFERENCES

1. R. M. Dudley, *Convergence of Baire measures*, Studia Math., **27** (1966), 251–268.

2. L. Gillman and M. Jerison, *Rings of Continuous Functions*, Van Nostrand, 1955.

3. E. Granirer, *On Baire measures on D-topological spaces*, Fund. Math., **60** (1967), 1–22.

4. E. Hewitt, *Linear functionals on spaces of continuous functions*, Fund. Math., **37** (1950), 161–189.

5. R. B. Kirk, *Measures in topological spaces and B-compactness*, Proc. Acad. Sci., Amsterdam, **A** (1969), 172–183.

6. R. B. Kirk, *Complete topologies on spaces of Baire measures*, Trans. Amer. Math. Soc., (to appear).
7. G. Köthe, *Topological Vector Spaces*, I Springer-Verlag, Inc., New York, 1969.
8. E. Marczewski and R. Sikorski, *Measures in non-separable metric spaces*, Colloq. Math., **I** (1948) 133-139.
9. V. S. Varadarajan, *Measures on topological spaces*, (Russian, Mat. Sbornik N. S., **55** (1961), 33-100. Translated in Amer. Math. Soc. Translations (2), **78** (1965), 161-228.

Received July 7, 1972. The author wishes to thank the Graduate School of Southern Illinois University for supporting released time during the summer of 1972 when this work was completed.

SOUTHERN ILLINOIS UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

THE SEIFERT AND VAN KAMPEN THEOREM VIA REGULAR COVERING SPACES

R. J. KNILL

The Seifert and Van Kampen theorem has lately been phrased as the solution to a universal mapping problem. There is given here an analogous theorem for regular covering spaces, regarded as principal bundles with discrete structure groups. The universal covering space of a union of two spaces is built up from the universal covering spaces of the two subspaces by an application of the associated bundle and clutching constructions. When all spaces are semi-locally simply connected, the Seifert and Van Kampen theorem is a consequence.

The technique of building a covering space piece by piece was effectively exploited by Neuwirth [10] to construct nonsimply connected covering spaces. We give an alternative approach to constructing a regular covering space of a base B which is either the union of open sets B_1 and B_2 with connected intersection B_0, or which is an adjunction space $B_1 \cup_f B_2$, where f glues a closed subspace B_0 of B_1 to B_2. We assume that all spaces are connected, and that there is given a regular covering space ξ_i of B_i for $i = 0, 1$, and 2, together with morphisms of covering spaces $\xi_0 \to \xi_i$, $i = 1, 2$. By regarding a regular covering space as a principal bundle with discrete structure group and applying the associated bundle and clutching constructions, we obtain a regular covering space ξ of B as the pushout of $\xi_1 \leftarrow \xi_0 \to \xi_2$. The structure group of ξ is the pushout of the structure groups of ξ_0, ξ_1, and ξ_2. One may obtain in this fashion the universal covering space of B from universal covering spaces of B_0, B_1, and B_2, or one may obtain the universal abelian covering space (i.e., that one having the maximal possible abelian structure group) from the universal abelian covering spaces of B_0, B_1, and B_2. The proof is by universal mapping arguments. In contrast to Neuwirth [10], the Seifert and Van Kampen theorem, under the hypotheses that all base spaces are locally connected and semi-locally simply connected, is a corollary. It is interesting that local homotopy conditions in a neighborhood of B_0, such as those assumed by Van Kampen and others ([15], [11], and [2]), turn out to be unnecessary, provided the space B_1 is paracompact. On the other hand, the Van Kampen formulae may apply in cases where B_0 is not connected or one of the B_i does not have a universal covering space, neither case being included in our results.

The only difficult arguments involve the validity of the clutching construction [14], [1], and [7] for $B_1 \cup_f B_2$. That is to say, one must establish the existence of a local product structure for any bundle formed by clutching. In the alternative case of open B_1 and B_2, with $B = B_1 \cup B_2$, the clutching construction easily yields a locally trivial object. Thus the person interested in the simplest route to a form of Seifert and Van Kampen will consider only the case of open B_1 and B_2, with $B = B_1 \cup B_2$.

The author wishes to acknowledge that the formulation in terms of universal properties was suggested by the referee.

1. Universality of the associated bundle. If ξ is a principal bundle with structure group G, and $u: G \to K$ is a continuous group homomorphism, then G acts on K on the left via u, and the associated bundle construction yields a new bundle with structure group K. This associated bundle is the solution to a certain universal mapping problem in the category of all principal bundles.

All spaces are assumed to be regular. Recall [14] that a principal bundle ξ consists of *bundle space* E_ξ, a *base space* B_ξ, a *projection* $p_\xi: E_\xi \to B_\xi$. There is also given a topological *structure group* G_ξ which acts on E_ξ freely from the right and such that p_ξ is equivalent map of E_ξ onto the space of orbits of G_ξ. It is assumed that p_ξ is *locally trivial*. This means that there is an open covering $\{V_i\}_{i \in I}$ of B_ξ by sets called *coordinate neighborhoods*, and for each $i \in I$ there exists a map (= continuous function) s_i: $V_i \to E_\xi$. This map is assumed to be a *section* of ξ *over* V_i, that is, for any point b of V_i, $p_\xi(s_i(b)) = b$. Local triviality is the condition that the function S_i defined below is a homeomorphism.

$$S_i: V_i \times G_\xi \longrightarrow p_\xi^{-1}(V_i) , \qquad S_i(b, g) = s_i(b) \cdot g .$$

Here $s_i(b) \cdot g$ denotes the right translate of $s_i(b)$ by an element g of G. The functions S_i are called *coordinate functions*. The coordinate function determined by any section of ξ is a homeomorphism.

LEMMA 1.1. *If ξ is a principal bundle and f and h are maps of a space X into E_ξ such that $p_\xi^\circ f = p_\xi^\circ h$, then there is a unique map $t: X \to G_\xi$ such that for any point x of X, equation* (1) *holds*

$$f(x) = h(x) \cdot t(x) . \tag{1}$$

Proof. For any point x of X, equation (1) determines $t(x)$ uniquely, by the freeness of the action of G_ξ on E_ξ. The map t is continuous by local triviality of p_ξ.

DEFINITION 1.2. Let ξ and ξ' be principal bundles. A *morphism*

$h: \xi \longrightarrow \xi'$ is a map of E_ξ into $E_{\xi'}$ such that there exists a function $h_G: G_\xi \longrightarrow G_{\xi'}$, which for any point x in E_ξ, and element g of G_ξ, satisfies equation (2)

$$(2) \qquad h(x \cdot g) = h(x) \cdot h_G(g) .$$

By the lemma, h_G is unique and continuous. It is an easy consequence of (2) that h_G is a homomorphism. Note that (2) also implies that h maps a fibre $p_\xi^{-1}(b)$ into a fibre $p_{\xi'}^{-1}(h_B(b))$. Let $h_B: B \longrightarrow B'$ be the unique and continuous function such that for any point x of E_ξ, equation (3) holds.

$$(3) \qquad p_{\xi'}(h(x)) = h_B(p_\xi(x)) .$$

Let ξ be a principal bundle, let K be a topological group, and let $u: G_\xi \longrightarrow K$ be a continuous homomorphism. Then G_ξ acts on K on the left via u, and the "weakly" associated bundle with fibre K will be denoted by α_u. See [14, §§ 8.7, 9.1]. The bundle space of α_u is usually denoted by $E \times_G K$, where $E = E_\xi$. It is formed as the quotient space of $E \times K$ by the relation which identifies a point (x, k) with $(x \cdot g, u(g^{-1})k)$, for every element g of G_ξ. The equivalence class of (x, k) is denoted $\langle x, k \rangle$. The action of K on $E \times_G K$ is defined by the rule $\langle x, k_1 \rangle \cdot k_2 = \langle x, k_1 k_2 \rangle$. The base space of α_u is that of ξ, $B_{\alpha_u} = B_\xi$. The projection is defined by the rule $p_{\alpha_u}\langle x, k \rangle = p_\xi(x)$. There is a natural map $u^\#: E \longrightarrow E \times_G K$ defined by the rule $u^\#(x) = \langle x, e \rangle$, where e is the identity element of K. If u is understood, then it will be convenient to write $\xi \times_G K$ for α_u.

THEOREM 1.3. *Let ξ be a principal bundle and let $u: G_\xi \longrightarrow K$ be a continuous homomorphism of topological groups. Then $u^\#$ is a morphism of principal bundles such that $u^\#_G = u$. If $h: \xi \longrightarrow \xi'$ is a morphism of principal bundles and $v: K \longrightarrow G_{\xi'}$ is a continuous homomorphism such that $h_G = v \circ u$, then there exists a unique morphism*

$$(h, v)^\#: \xi \times {}_G K \longrightarrow \xi'$$

such that (4) *and* (5) *hold.*

$$(4) \qquad h = (h, v)^\# \circ u^\#$$

$$(5) \qquad v = (h, v)^\#_G .$$

Proof. For x in E_ξ and g in G_ξ, $u^\#(x \cdot g) = \langle x \cdot g, e \rangle = \langle x \cdot g, e \rangle = \langle x, u(g) \rangle = \langle x, e \rangle \cdot u(g) = u^\#(x) \cdot u(g)$. This proves that $u^\#$ is a morphism and $u^\#_G = u$. The conditions (4) and (5) on $(h, v)^\#$ are equivalent by (2) to defining

$$(h, v)^\# \langle x, k \rangle = h(x) \cdot v(k) , \qquad \langle x, k \rangle \in E \times {}_G K .$$

The uniqueness of $(h, v)^{\#}$ follows also from this rule, so the proof is complete.

Roughly that theorem says that $h\colon \xi \to \xi'$ factors through α_u if and only if h_G factors through K, and the former factorization is determined by the latter. In this sense the associated bundle is the solution to a universal mapping problem.

2. The clutching construction. The clutching construction ([14], p. 97) is extended. There are given spaces B_0, B_1, and B_2 such that B_0 is a subspace of B_1 with suitable properties relative to bundles. There is given a map $f\colon B_0 \to B_2$ by means of which B_1 is attached to B_2 to form $B = B_1 \cup_f B_2$ ([3], p. 127 f). There is given a common structure group K for principal bundles ξ_0, ξ_1, and ξ_2 with respective base spaces B_0, B_1, and B_2 and there are given morphisms $j\colon \xi_0 \to \xi_1$ and $h\colon \xi_0 \to \xi_2$ such that $j_B\colon B_0 \to B_1$ is the inclusion map and $h_B = f$, and such that j_G and h_G are the identity homomorphism of K. If E_i is the bundle space of ξ_i for $i = 0, 1$, and 2, and let $E = E_1 \cup_h E_2$. The projection p of E onto B is defined by functoriality of the attaching construction, and likewise by functoriality K acts on E freely from the right. In order to conclude that $\xi = (E, p, B, K)$ is a principal bundle it suffices to prove that the projection is locally trivial. We write $\xi_1 \cup_h \xi_2$ for ξ.

DEFINITION 2.1. For any ξ_1 and ξ_2 as above, and for a closed subset A of B_1, the set, $\mathscr{G}(A, \xi_1)$, of *germs* of sections of ξ_1 *over* A is defined as follows. An element of $\mathscr{G}(A, \xi_1)$ is an equivalence class of sections of ξ_1 over neighborhoods of A, where two such sections, s and t, are defined to be equivalent if for some open neighborhood V of A both s and t are defined throughout V and their restrictions to V are equal. For a closed subset A of B_0, likewise there is defined $\mathscr{G}(A, \xi_0)$, and restriction of sections induces a function

$$j^*\colon \mathscr{G}(A, \xi_1) \longrightarrow \mathscr{G}(A, \xi_0)\ .$$

Let ξ_2 be a principal bundle, and let $f\colon B_0 \to B_{\xi_2}$ be a map. Let $f^*(\xi)$ be the induced principal bundle ($f^{-1}(\xi)$ in [14]) with base B_0, and let $\bar{f}\colon f^*(\xi_2) \to \xi_2$ be the canonical morphism. Recall that $\bar{f}_B = f$, $\bar{f}_G$ is the identity homomorphism, $\bar{f}_G\colon G_{f^*(\xi_2)} = G_{\xi_2}$, and it is easy to see that

$$\begin{tikzcd}
& \xi_1 \arrow[dr, "h'"] & \\
\xi_0 \arrow[ur, "j"] \arrow[dr, "h"'] & & \xi \\
& \xi_2 \arrow[ur, "j'"'] &
\end{tikzcd}$$

DIAGRAM I

$$\begin{tikzcd}
& G_{\xi_1} \arrow[dr, "h'_G = u_1"] & \\
G_{\xi_0} \arrow[ur, "j_G"] \arrow[dr, "h_G"'] & & G_\xi = K \\
& G_{\xi_2} \arrow[ur, "j'_G = u_2"'] &
\end{tikzcd}$$

DIAGRAM II

$$\begin{tikzcd}
 & & E_1 \arrow[d, "u_1^\#"] \arrow[ddrr] & & \\
 & & E_1 \times_G K \arrow[dr, "H'"] & & \\
E_0 \arrow[r, "u_0^\#"] \arrow[uurr, "j"] \arrow[ddrr, "h"'] \arrow[ur, phantom, "\text{(i)}"] \arrow[dr, phantom, "\text{(ii)}"] & E_0 \times_G K \arrow[ur, "J"] \arrow[dr, "H"'] & & E & E' \\
 & & E_2 \times_G K \arrow[ur, "J'"'] & & \\
 & & E_2 \arrow[u, "u_2^\#"'] \arrow[uurr] & &
\end{tikzcd}$$

DIAGRAM III

these two conditions characterize $f^*(\xi_2)$. Recall further that for any section $s\colon V_2 \to E_{\xi_2}$, where V_2 is a subset of B_{ξ_2}, there exists a unique section denoted $f^*(s)\colon f^{-1}(V_2) \to E_{f^*(\xi_2)}$ such that $s \circ f = \bar{f}^0 f^*(s)$.

THEOREM 2.2. *Let ξ_0, ξ_1, ξ_2, f, j, h, and ξ be as above. Let*

$$h'\colon E_1 \longrightarrow E_1 \cup_h E_2$$

$$j'\colon E_2 \longrightarrow E_1 \cup_h E_2$$

be the canonical maps. Under any of the conditions (2A), (2B), *or* (2C) *stated below, ξ is a principal bundle and h' and j' are morphisms of principal bundles which fill in the pushout diagram for j and h in the category of principal bundles* (diagram I).

(2A) *B_1 and B_2 are open subspaces of a common space $B = B_1 \cup B_2$, and $B_0 = B_1 \cap B_2$. The inclusion of B_0 into B_2 is $h_B = f$.*

(2B) *B_1 and B_2 are closed subspaces of a common space $B = B_1 \cup B_2$, and $B_0 = B_1 \cap B_2$. The inclusion of B_0 into B_2 is $h_B = f$. Further for each point b of B_0 the restriction function of $\mathscr{G}(b, \xi_1)$ to $\mathscr{G}(b, \xi_0)$ is onto.*

(2C) *B_0 is a closed subspace of B_1, and $h_B = f$ is arbitrary. Further for every closed subset A of B_0, the restriction function of $\mathscr{G}(A, \xi_1)$ to $\mathscr{G}(A, \xi_0)$ is onto.* (Cf. 4.2 *and* 4.3).

Proof. For properties of the attaching construction see [3, pp. 127–129]. The canonical maps h' and j' are the restrictions to E_1 and E_2, resp., of the quotient projection of the free union, $E_1 + E_2$, onto $E_1 \cup_h E_2$. Assuming that ξ is a principal bundle, then h' and j' fill in the pushout diagram ([9], p. 10) for j and h, since they do so as maps in the category of topological spaces. It remains to show that p is locally trivial. In case (2A), the coordinate neighborhoods for p_{ξ_1} and p_{ξ_2} are trivially coordinate neighborhoods for p. In case (2B), the coordinate neighborhoods for p_{ξ_1} and p_{ξ_2} in the complements of B_0 in B_1 and B_2, resp., are trivially coordinate neighborhoods for p. For an arbitrary point b of B_1, let V_2 be an open neighborhood of b relative to B_2, and let s_2 be a section of ξ_2 over V_2. Then $f^*(s_2)$ is the restriction of s_2 to a section, s_0, of ξ_0 over $V_2 \cap B_0$. By hypothesis, the germ of s_0 over b extends to a germ of a section s_1 of ξ_1 over a neighborhood V_1 of b relative to B_1. By cutting down V_1 and V_2, if necessary, we may assume that s_1, s_0, and s_2 all are defined on $V_1 \cap V_2$ and all agree there. Then $s_1 \cup s_2$ is a section of ξ over $V_1 \cup V_2$, and the coordinate function defined by $s_1 \cup s_2$ is a homeomorphism, by the functoriality of the attaching construction. In case (2C), the coordinate neighborhoods for p_{ξ_1} in the complement of B_0 in B_1 are trivially coordinate neighborhoods for p. It remains to find coordinate neighborhoods of the points b of B_2. So, let s_2 be a section of ξ_2 over an open neighborhood, V_2, of b. Let A_2 be a closed neighborhood of b contained in V_2, and let $A = f^{-1}(A_2)$. Again $f^*(s_2)$ is a section, s_0, of ξ_0 in a neighborhood of A, and by hypothesis, the germ of s_0 over A extends to a germ of a section, s_1, of ξ_1 over a neighborhood, V_1, of A relative to B_1. By cutting down V_1 and V_2, if necessary, we may assume that s_1 and s_0 are both defined on $V_1 \cap B_0$ and both agree there. Then $s_1 \cup_f s_2$ is a section of ξ (defined by functoriality of the attaching construction), the coordinate function defined by it being a homeomorphism by functoriality of the attaching construction. This completes the proof of theorem.

3. Universal covering spaces. The Seifert and Van Kampen theorem for regular covering spaces is the statement that pushouts exist in the category of regular covering spaces provided suitable conditions are satisfied by the base spaces. The construction does not generalize to locally compact principal fibre bundles since it would rely on the existence of pushouts in the category of locally compact structure groups.

A *regular covering space* of a connected base space B is a principal bundle with base B and a discrete structure group. The bundle space of a regular covering space is not assumed to be connected. The *universal covering space* of B (if it exists) is the regular covering space such that for a fixed element x_0 of E_ξ, ξ has the *universality* property: for any regular covering space ξ', for any map $f\colon B \to B_{\xi'}$, and for any point x' of $E_{\xi'}$ such that $p_{\xi'}(x') = f(p_\xi(x_0))$, there exists a unique morphism $f^-\colon \xi \to \xi'$ such that $(f^-)_B = f$, and $f^-(x_0) = x'$. It is not hard to conclude that the universal covering space of a given base space B is unique up to isomorphism, that its bundle space is connected, that the universality property does not depend upon the choice of fixed element x_0, and that it suffices to verify the universality property relative to covering spaces ξ' with the same base B and for f equal to the identity of B. It is also known that if a regular covering space of a base B has a path connected simply connected bundle space, then it is the universal covering space of B, and its structure group is naturally isomorphic to the fundamental group of B. There exists an example ([13], p. 84 and [6]) of a connected, locally path connected metric space B such that B has a universal covering space which is not simply connected.

We now consider circumstances similar to those of paragraph 2. There are given spaces B_0, B_1, and B_2 such that B_0 is a subspace of B_1. There is given a map, f, of B_0 into B_2, by means of which B_1 is attached to B_2 to form $B = B_1 \cup_f B_2$. There are given regular covering spaces ξ_0, ξ_1, and ξ_2 with respective base spaces B_0, B_1, and B_2. *No assumption is made that the structure groups are isomorphic.* There are given morphisms, j and h, of ξ_0 to ξ_1, and of ξ_0 to ξ_2, respectively, such that j_B is the inclusion map of B_0 into B_1, and such that $h_B = f$. Let E_i be the bundle space of ξ_i, for $i = 0, 1,$ and 2. There exists a discrete group, K, and there exist morphisms, u_1 and u_2, of G_{ξ_1} and G_{ξ_2}, respectively, into K filling in the pushout diagram for j_G and h_G (diagram II). Let $u_0 = u_1 \circ j_G = u_2 \circ h_G$. Using these homomorphisms define the associated bundles $\xi_i \times_G K = \alpha_{u_i}$ for $i = 0, 1,$ and 2. Then j and h induce morphisms

$$J\colon E_0 \times_G K \longrightarrow E_1 \times_G K\,, \qquad J(\langle x, k\rangle) = \langle j(x), k\rangle$$

$$H\colon E_0 \times_G K \longrightarrow E_2 \times_G K\,, \qquad H(\langle x, k\rangle) = \langle h(x), k\rangle\,.$$

Evidently $J_G = H_G =$ the identity homomorphism of K. Then define

$$E = (E_1 \times_G K) \bigcup\nolimits_H (E_2 \times_G K)$$

$$\xi = (\xi_1 \times_G K) \bigcup\nolimits_H (\xi_2 \times_G K)\,.$$

There are natural maps H' and J' induced by the clutching construction filling in a pushout diagram of spaces (the inside cell of diagram III). Let $h' = H' \circ u_1^\sharp$ and $j = J' \circ u_2^\sharp$. Then diagram I is a pushout diagram, provided that ξ is a principal bundle. The theorem will be that this will be so in cases (3A), (3B), and (3C).

(3A) B_1 and B_2 are open subspaces of a common space $B = B_1 \cup B_2$, and $B_0 = B_1 \cap B_2$. The inclusion of B_0 into B_2 is $h_B = f$.

(3B) B_1 and B_2 are closed subspaces of a common space $B = B_1 \cup B_2$, and $B_0 = B_1 \cap B_2$. The inclusion of B_0 into B_2 is $h_B = f$.

(3C) B_0 is a closed subspace of B_1, B_1 is paracompact, and $h_B = f$ is arbitrary.

LEMMA 3.1. *Case* (3B) *implies case* (2B), *and case* (3C) *implies* *case* (2C), *for the principal bundles* $\xi_i \times_G K$ *over* B_i, $i = 0, 1, 2$.

Proof. In case (3B) note that since K is discrete, then for a point b of B_0 any two sections of $\xi_0 \times_G K$ which are defined over neighborhoods of b must have the same germ over b. It follows that the restriction function of $\mathscr{G}(b, \xi_1 \times_G K)$ to $\mathscr{G}(b, \xi_0 \times_G K)$ is an isomorphism, and case (2B) holds. In case (3C), we regard all regular covering spaces to be sheaves of sets. Let A be any closed subset of B_0. By a standard theorem of sheaf theory ([4], p. 150), any section over A extends to a section over a neighborhood of A, and since two sections must agree over an open set, then it follows that the restriction function of $\mathscr{G}(A, \xi_1 \times_G K)$ to $\mathscr{G}(A, \xi_0 \times_G K)$ is an isomorphism and case (2C) holds.

THEOREM 3.2. *Let* ξ_0, ξ_1, ξ_2, f, j, h, *and* ξ *be as above. Under any of the conditions* (3A), (3B), *or* (3C), *diagram I is a pushout diagram in the category of regular covering spaces, and the induced diagram II of homomorphisms of structure groups is a pushout diagram in the category of groups.*

Proof. Cells (i) and (ii) of diagram III are commutative as an application of (1.3). Let $h' = H' \circ u_1^\#$, and $j = J' \circ u_2^\#$. Since H'_G and J'_G both are the identity homomorphism of K, then

$$h'_G = (u_1^\#)_G = u_1 \,,$$

and

$$j'_G = u_2 \,.$$

We show that h' and j' fill in the pushout diagram for j and h in the category of regular covering spaces. Suppose that for $i = 0, 1$, and 2, there are given morphisms $l_i \colon \xi_i \longrightarrow \xi'$ such that $l_0 = j \circ l_1 = h \circ l_2$. Since then

$$(l_0)_G = (l_1)_G \circ j_G = (l_2)_G \circ h_G$$

then there exists a unique homomorphism $v \colon K \longrightarrow G_{\xi'}$ such that

$$v \circ h'_G = l_1 \,,$$

and

$$v \circ j'_G = l_2 \,.$$

For each $i = 0, 1$ or 2, l_i induces a unique morphism

$$(l_i, v)^\# \colon \xi_i \times_G K \longrightarrow \xi'$$

such that

$$(l_i, v)^{\#}_G = v\ , \quad \text{and} \quad (l_i, v)\circ u_i^{\#} = l_i\ .$$

It follows that

$$(l_1, v)^{\#}\circ J = (l_0, v)^{\#} = (l_2, v)^{\#}\circ H\ .$$

Since H' and J' fill in the pushout diagram for J and H then $(l_1, v)^{\#}$ and $(l_2, v)^{\#}$ induce a unique morphism $l\colon \xi \longrightarrow \xi'$ such that

$$(l_1, v)^{\#} = l\circ H'\ , \quad \text{and} \quad (l_2, v)^{\#} = l\circ J'\ .$$

By its construction, l satisfies

$$l\circ h' = l_1\ , \quad \text{and} \quad l\circ j' = l_2$$

and l is the unique morphism which does so. Since l_1 and l_2 were arbitrary such that $l_1\circ j = l_2\circ h$, this completes the proof that h' and j' fill in the pushout diagram for j and h.

The Corollary 3.3 is the analogue for universal covering spaces of the Seifert and Van Kampen theorem.

COROLLARY 3.3. *Under the conditions of* 3.2, *if ξ_i is the universal covering space of B_i for $i = 0, 1,$ and 2, then ξ is the universal covering space of B, and the diagram II of induced homomorphisms of structure groups is a pushout diagram in the category of groups.*

Proof. There exist morphisms $j\colon \xi_0 \longrightarrow \xi_1$, and $h\colon \xi_0 \longrightarrow \xi_2$ such that $j_B =$ the inclusion of B_0 into B_1, and $h_B = f$. Then there exist pushout diagrams in the category of covering spaces and in the category of groups. (See illustration I.) Let ξ' be any regular covering space of B, and suppose there are given points x of E_ξ, such that $p_\xi(x) = p_{\xi'}(x')$. We must find a morphism $l\colon \xi \longrightarrow \xi'$ such that $l(x) = x'$, and show that such a morphism is unique, in order to complete the proof. We may assume that x is so chosen that there is a point x_0 in E_{ξ_0} such that $h'(j(x_0)) = x$. For $i = 0, 1,$ and 2, let $l_i\colon \xi_i \longrightarrow \xi'$ be the unique morphism such that $(l_i)_B$ is the natural map of B_i into B defined by the attaching construction, and such that $l_0(x_0) = x'$, $l_1(j(x_0)) = x'$, and $l_2(h(x_0)) = x'$. Since they agree at one point, x_0, the maps l_0, $l_1\circ j$, and $l_2\circ h$ are all equal. Let l be the unique morphism such that $l\circ h' = l_1$, and $l\circ j' = l_2$. This completes the proof.

COMMENT 3.4. Theorem 3.2 here stated and proven for all principal bundles with structure groups in the category of (discrete) groups, could be stated and be valid with no change in proof for any full subcategory of the category of principal bundles provided the corresponding category of structure groups had pushouts. The notion of universal principal bundle in that category would make sense provided the structure groups were discrete. For example, by taking the corresponding category of structure groups to be the category of abelian groups, one arrives at the notion of universal *abelian* regular covering space, and Theorems 3.2 and 3.3 in terms of *abelian* regular covering spaces would remain valid.

4. **Extension properties of germs.** The extension properties of germs are developed for principal bundles in general, there being interest in the validity of the clutching construction. There is a trade-off in hypotheses to be made, it being necessary to strengthen hypotheses on the base spaces in order to admit less stringent conditions on the structure group.

Throughout this section ξ_1 is a principal bundle with base B_1 and structure group K, and B_0 is a closed subspace of B_1. Let $\xi_0 = \xi_1 | B_0 = j^*(\xi_1)$ be the restriction of ξ_1 to B_0, where j is the inclusion of B_0 into B_1. Recall that for a closed subset A of B_1, one says of A that it has the *neighborhood extension property* in B_1 relative to K provided that for any map f of A into K, there exists an extension of f to a map of a neighborhood of A into K.

THEOREM 4.1. *If every closed subset of B_1 has the neighborhood extension property in B_1 relative to K, and if B_1 is paracompact, then case* (2C) *holds.*

Proof. The proof is a straightforward modification of standard arguments which conclude a global property from the corresponding local property; see [14], p. 55, or [5], Theorem 2.7. To begin, let V_0 be a closed relative neighborhood of A in B_0, and let s_0 be any section of ξ_0 over V_0. It will be shown that there exists an extension of s_0 to a section s_1 of ξ_1 over a closed neighborhood, V_1, of V_0 in B_1. That would suffice to prove the theorem. Let $\{U_i\}_{i \in I}$ be a locally finite open covering of B_1 by coordinate neighborhoods, and for each index i, let s_i be a section of ξ_1 over U_i. Let $\{V_i\}_{i \in I}$ be a closed covering of B_1 such that for each index i, V_i is contained in U_i. For any subset J of I, let

$$V_J = V_0 \cup (\bigcup \{V_i : i \in J, \text{ and } V_0 \cap V_i \neq \emptyset\}) .$$

Let $\mathscr{S}$ be the set of all pairs (J, s), where J is a subset of I, and s is an extension of s_0 to a section of ξ_1 over a closed set, N_s, which is a relative neighborhood of V_0 in V_J. $\mathscr{S}$ is partially ordered by the relation defined as $(J, s) \leqq (J', s')$ if $J \subset J'$ and s' extends s. By the Hausdorff maximal principal, there exists a maximal chain, $\mathscr{C}$, contained in $\mathscr{S}$. If we let

$$J' = \bigcup \{J: (J, s) \in \mathscr{C}\} ,$$

and

$$s' = \bigcup \{s: (J, s) \in \mathscr{C}\} ,$$

then (J', s') is the maximum of $\mathscr{C}$. If $J' = I$, then $N_{s'}$ would be a neighborhood of V_0 in B_1, and the theorem would be proven. Suppose that there were an index i not contained in J'. Then $V_i \cap V_0 \neq \emptyset$, for otherwise $(J' \cup \{i\}, s')$ would be an element of $\mathscr{S}$ greater than (J', s'). Let

$$t: N_{s'} \cap V_i \longrightarrow K$$

be the map defined by the equation

$$s'(x) = s_i(x) \cdot t(x) , \qquad x \in N_{s'} \cap V_i .$$

By the neighborhood extension property hypothesis, there is an extension, t', of t to a map of a closed neighborhood, M, of $N_{s'} \cap V_i$ into K. Then *extend* s' to a section s'' over $N_{s'} \cup (M \cap V_i)$ by letting, for any point x of $M \cap V_i$,

$$s''(x) = s_i(x) \cdot t'(x) .$$

Then $(J' \cup \{i\}, s'')$ is an element of $\mathscr{S}$ greater than (J', s'), contrary to the maximality of $\mathscr{C}$ and of (J', s''). It follows that $J' = I$. As previously observed, this proves the theorem.

COROLLARY 4.2. *If B_1 is a paracompact space, and K is a Lie group, then case* (2C) *holds.*

Proof. Since K is topologically complete and an ANR, then [8] every closed subset of B_1 has the neighborhood extension property in B_1 relative to K, and so 4.1 implies 4.2.

If the structure group, K, is not a Lie group, all is not lost.

THEOREM 4.3. *If ξ_1 has the homotopy lifting property and B_0 is a neighborhood deformation retract of B_1, and in particular if the inclusion of B_0 into B_1 is a cofibration, then case* (2C) *holds.*

Proof. This straightforward application of the homotopy lifting property is left to the reader.

References

1. M. Atiyah, *K-Theory Lectures*, Fall 1964, Harvard University, Cambridge.
2. R. Brown, *Groupoids and Van Kampen's theorem*, Proc. London Math. Soc., (3) **17** (1967), 385–401.
3. J. Dugundji, *Topology*, Boston, 1966.
4. R. Godement, *Topologie algebrique et theorie des faisceaux*, Hermann, Paris, 1958.
5. Albrecht Dold, *Partitions of unity in the theory of fibrations*, Ann. of Math., **78** (1963), 223–255.
6. H. B. Griffiths, *The fundamental group of two spaces with a common point*, Quarterly J. of Math., Oxford (2nd) Series, **5** (1954), 175–190.
7. Dale Husemoller, *Fibre Bundles*, McGraw-Hill, New York, 1966.
8. Ernest Michael, *Some extension theorems for constructing functions*, Pacific J. Math., **3** (1953), 789–806.
9. B. Mitchell, *Theory of Categories*, Academic Press, New York, 1965.
10. L. P. Neuwirth, *Knot Groups*, Annals of Mathematics Studies Number 56, Princeton, 1965.
11. P. Olum, *Nonabelian cohomology and Van Kampen's theorem*, Ann. of Math., (2) **68** (1958), 658–668.
12. H. Siefert and W. Threllfall, *Lehrbuck der Topologie*, Berlin, 1934.
13. E. Spanier, *Algebraic Topology*, New York, 1966.
14. N. Steenrod, *The Topology of Fibre Bundles*, Princeton University Press, Princeton, 1957.
15. E. R. Van Kampen, *On the connection between the fundamental groups of some related spaces*, Amer. J. Math., **55** (1933), 261–267.

Received June 23, 1972 and in revised form May 15, 1973. Research supported in part by NSF Grant # GP 29707.

Tulane University

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

HOMOMORPHISMS OF MATRIX RINGS INTO MATRIX RINGS

AMOS KOVACS

Let $V_n(R_n)$ be the universal ring with respect to embeddings of the matrix ring R_n into $n \times n$ matrix rings over commutative rings. A construction and a representation is given for this ring. As a main tool in the construction, it is proved that every R homomorphism of R_n, R a commutative ring, is the restriction of an inner automorphism of U_n, for some $U \supseteq R$. Using this, a necessary and sufficient condition for n^2 matrices in R_n to be matrix units is given.

1. Introduction and notations. All rings to be considered in this paper, except those denoted specifically as matrix rings, will be commutative rings with unit. All homomorphisms are unitary. The unit of a subring coincides with the unit of its over-ring.

Denote by R_n the ring of $n \times n$ matrices over a ring R. Let $\eta: R \to S$ be a ring homomorphism then η induces a homomorphism $\eta_n: R_n \to S_n$ given by: $\eta_n(r_{ij}) = (\eta(r_{ij}))$. If $A \in R_n$, $(A)_{ij}$ will denote the (i, j)th entry of A. The identity element and the standard matrix units of all matrix rings will be denoted by I and $\{E_{ij}\}$ respectively.

Let A be an R algebra. It was proved by Amitsur ([1], Theorem 2) that there exists a commutative R algebra $V_m^R(A)$, and a map $\rho: A \to (V_m^R(A))_m$ which is universal for homomorphisms of A into $m \times m$ matrix rings over commutative rings, i.e.;

(1) For every $\tau: A \to H_m$, with H a commutative R algebra, there exists a homomorphism $\eta: V_m^R(A) \to H$ such that the following diagram is commutative;

$$\begin{tikzcd} A \arrow[r, "\rho"] \arrow[dr, "\tau"'] & (V_m^R(A))_m \arrow[d, "\eta_m"] \\ & H_m \end{tikzcd}$$

(2) $V_m^R(A)$ is generated over R by the entries $\{[\rho(a)]_{ij} \mid a \in A\}$.

Properties (1) and (2) determine $V_m^R(A)$ up to isomorphism and ρ up to a multiple by an isomorphism of $V_m^R(A)$.

In this paper we will give an explicit construction for the ring $V_m^R(R_n)$. The case $n = m$ will be treated separately. We start with investigating the nature of R-homomorphisms of R_n into itself.

2. On the automorphisms of matrix rings.

LEMMA 1. *Let C be a subdirectly irreducible commutative ring. C can be embedded in a local ring Q which is the complete ring of quotients of C. For noetherian C we may take $C = Q$.*

Proof. Let $\mathfrak{M}$ be the set zero divisors in C, then, by [6] $\mathfrak{M}$ is a maximal ideal in C. Let Q be the local ring $C_{\mathfrak{m}}$ and let f be the canonical homomorphism $f\colon C \to C_{\mathfrak{m}}$.

The elements of $C - \mathfrak{M}$ are not zero divisors in C, hence f is an injection. Furthermore, the elements of $C - \mathfrak{M}$ are exactly the regular elements of C, and so Q is the complete ring of quotients of C.

Now, if C is noetherian we have, by [3], $\mathfrak{M} = \mathfrak{N}(C)$-the nil radical of C, and therefore, $J(C) \subseteqq \mathfrak{M} = \mathfrak{N}(C) \subseteqq J(C)$ where $J(C)$ denotes the Jacobson radical of C. Hence $\mathfrak{M} = J(C)$ and being maximal it is the unique maximal ideal of C. Consequently C is local and $C = C_{\mathfrak{m}} = Q$.

THEOREM 2. *Let C and Q be as in the lemma; $\{E_{ij} \mid 1 \leqq i, j \leqq n\}$ be the set of the standard matrix units in C_n, and $\{F_{ij} \mid i \leqq i, j \leqq n\}$ another set of matrix units in C_n, then there exists a matrix $A \in C_n$, invertible in Q_n such that:*

$$1 \leqq i, j \leqq n \qquad E_{ij}A = AF_{ij} \, .$$

If C is noetherian A is invertible in C_n.

Proof. By definition $\sum_{\nu=1}^{n} F_{\nu\nu} = I$ hence $\sum_{\nu=1}^{n} (F_{\nu\nu})_{11} = (\sum_{\nu=1}^{n} F_{\nu\nu})_{11} = 1$. Now C is subdirectly irreducible, so that the zero divisors in C form an ideal. Consequently not all of the $(F_{\nu\nu})_{11}$ are zero divisors and we have some $1 \leqq \nu \leqq n$ such that $(F_{\nu\nu})_{11}$ is regular. Without loss of generality we may assume that $\alpha = (F_{11})_{11}$ is regular. Put now

$$A = \sum_{\nu=1}^{n} E_{\nu 1}F_{1\nu} \qquad B' = \sum_{\nu=1}^{n} F_{\nu 1}E_{1\nu}$$

than for all i and j we have $E_{ij}A = E_{i1}F_{1j} = AF_{ij}$ and also

$$AB' = \Big(\sum_{\nu=1}^{n} E_{\nu 1}F_{1\nu}\Big)\Big(\sum_{\mu=1}^{n} F_{\mu 1}E_{1\mu}\Big) = \sum_{\nu=1}^{n} E_{\nu 1}F_{11}E_{1\nu} = \alpha I \, .$$

Now α is regular in C and hence invertible in Q, thus $B = \alpha^{-1}B'$ is the inverse of A in Q_n. If C is noetherian then by Lemma 1 $C = Q$ and $B = A^{-1} \in C_n$.

Note that in a local ring the noninvertible elements form an ideal so that the proof of Theorem 2 can be easily modified to give an elementary proof of the following well-known (e.g. [2]) theorem: "Let C be a local ring then every C-homomorphism of C_n is an inner automorphism".

THEOREM 3. *Let C be a commutative ring with unit, $E_{ij}, F_{ij} \in C_n$ as above, then there exists a commutative ring U containing C and an invertible matrix $A \in U_n$ such that $E_{ij}A = AF_{ij}$ for all $1 \leqq i, j \leqq n$.*

Proof. C may be represented as a subdirect product of subdirectly irreducible rings ([4], Theorem 1, p. 219). There exists therefore a set of subdirectly irreducible rings with unit, $\{C^\gamma \mid \gamma \in \Gamma\}$ such that $C \subseteqq \prod_{\gamma \in \Gamma} C^\gamma \subseteqq \prod_{\gamma \in \Gamma} Q^\gamma = U$ where Q^γ is the complete ring of quotients of C^γ. Hence $C_n \subseteqq (\Pi C^\gamma)_n \subseteqq (\Pi Q^\gamma)_n = U_n$. Let $\pi^\gamma: \Pi Q^\gamma \to Q^\gamma$ be the canonical projection. Put $E^\gamma_{ij} = \pi^\gamma_n(E_{ij})$, $F^\gamma_{ij} = \pi^\gamma_n(F_{ij})$ then, by definition E^γ_{ij} are the standard matrix units in C^γ, F^γ_{ij} are another set of matrix units. By Theorem 2 it follows that there are invertible matrices $A^\gamma \in Q^\gamma_n$ such that $E^\gamma_{ij}A^\gamma = A^\gamma F^\gamma_{ij}$. Let $A \in (\Pi Q^\gamma)_n = U_n$ be defined by $((A)_{\nu\mu})_{(\gamma)} = (A^\gamma)_{\nu\mu}$, namely for every $\gamma \in \Gamma$ $\pi^\gamma_n(A) = A^\gamma$. Clearly A is invertible in U_m, its inverse being given by $((A^{-1})_{\nu\mu})_{(\gamma)} = ((A^\gamma)^{-1})_{\nu\mu}$. Clearly A satisfies $E_{ij}A = AF_{ij}$ for all i and j.

COROLLARY 4. (1) *For a given ring C there exists a ring $U \supseteqq C$ such that every C-homomorphism $\eta: C_n \to C_n$ can be extended to an inner automorphism of U_n.*

(2) *Given η, the ring U of* (1) *may be chosen so that the inner automorphism will be given by a matrix of determinant* 1.

Proof. (1) follows immediately from Theorem 3 by taking $F_{ij} = \eta(E_{ij})$ then, by the theorem we have a ring U and a matrix $A \in U_n$ such that $A^{-1}BA = \eta(B)$, $B \in C_n$.

(2) For a fixed η we adjoin to U the nth root of $a^{-1} = \det(A^{-1})$ and replace A by $a^{-1/n}A$.

REMARK. The ring U of Theorem 3 is not uniquely determined. For example one may take $U = \Pi C_{\mathfrak{m}}$ where the product is taken over all maximal ideals in C. This ring will have the same property.

COROLLARY 5. *Let C be a commutative ring with unit, $\{F_{ij} \mid 1 \leqq i, j \leqq n\}$ a set of matrix units in C_n then* $\mathrm{Cent}_{C_n}(F_{ij})$*-the centralizer of all the F_{ij} in C_n — is C, and every element in C_n may be written in a unique way as $\sum_{ij} c_{ij}F_{ij}$, $c_{ij} \in C$.*

Proof. By Theorem 3 $F_{ij} = AE_{ij}A^{-1}$ where A is in some larger matrix ring $U_n \supseteqq C_n$. Hence; $\mathrm{Cent}_{C_n}(F_{ij}) = \mathrm{Cent}_{U_n}(F_{ij}) \cap C_n = A\ \mathrm{Cent}_{U_n}(E_{ij})A^{-1} \cap C_n = U \cap C_n = C$. The proof of the second part of the corollary is classical (e.g. [4], Proposition 6, p. 52).

3. The ring $V_n^R(R_n)$. Let R be a fixed ring with unit. All rings henceforth will be R-algebras and all homomorphisms R-homomorphisms. We shall write $V_n(A)$ for $V_n^R(A)$. We now proceed to give an explicit construction for the ring $V_n(R_n)$.

Let $\{x_{ij} \mid 1 \leqq i, j \leqq n\}$ be n^2 commutative indeterminates over R, let $R[x_{ij}]$ denote the ring of polynominals in the x_{ij} over R. Denote by $D \subseteqq R[x_{ij}]$ the ideal generated in $R[x_{ij}]$ by the polynomial $\det(x_{ij})-1$, and put $K = R[x_{ij}]/D$. Clearly we may take $R \subseteqq K$. Put $\xi_{ij} = x_{ij} + D \in K$, and $\Xi = (\xi_{ij}) \in K_n$. The matrix Ξ is invertible in K_n, and its inverse is given by $\Xi^{-1} = \mathrm{adj}\Xi = (\hat{\Xi}_{ji})$ where $\hat{\Xi}_{ij}$ is the algebraic complement of ξ_{ij} in Ξ.

Let S be the subalgebra of K generated over R by the n^4 elements $\{\xi_{ij}\hat{\Xi}_{kl} \mid 1 \leqq i, j, k, l \leqq n\}$. ($S$ contains the unit element of K for $\sum_{i=1}^n \xi_{1i}\hat{\Xi}_{1i} = \det(\xi_{ij}) = 1$.)

Define a map $\rho\colon R_n \to K_n$ by $\rho(B) = \Xi B \Xi^{-1}$. ρ is clearly a unitary R homomorphism, furthermore we have:

$$\begin{aligned}[\rho(E_{ij})]_{kl} &= (\Xi E_{ij}\Xi^{-1})_{kl} = \sum_{\mu}\sum_{\nu}(\Xi)_{k\mu}(E_{ij})_{\mu\nu}(\Xi^{-1})_{\nu l} \\ &= (\Xi)_{ki}(\Xi^{-1})_{jl} = \xi_{ki}\hat{\Xi}_{lj} \in S\end{aligned}$$

so that $\rho(R_n) \subseteqq S_n$ and we may regard ρ as a map from R_n to S_n and note that the entries $\{[\rho(A)]_{ij} \mid A \in R_n\}$ generate S. For this ring S and the homomorphism ρ we prove;

THEOREM 6. *S is the universal ring $V_n(R_n)$ and ρ is the canonical embedding of R_n in $S_n = (V_n(R_n))_n$.*

Proof. We have seen that S is generated by the appropriate elements so that all that remains to be shown is that every homomorphism $\tau\colon R_n \to C_n$ factors through ρ.

As C is an R algebra, we have the natural homomorphism $i\colon R \to C$. Denote $i(r) = r'$ $i_n(B) = B'$ for all $r \in R$, $B \in R_n$.

Let E_{ij} be the standard matrix units of R_n than E'_{ij} are the standard matrix units of C_n and $\tau(E_{ij}) = F_{ij}$ are a set of matrix units in C_n. By Theorem 3 there exists an R algebra $U \supseteqq C$ and an invertible matrix $(a_{ij}) = A \in U_n$ such that $F_{ij} = AE'_{ij}A^{-1}$; furthermore, by Corollary 4 (2) we may suppose that $\det(A) = 1$. We clearly have for all $B \in R_n$ $\tau(B) = AB'A^{-1}$.

Define $\eta^\circ: R[x_{ij}] \to U$ by $\eta^\circ(x_{ij}) = a_{ij}$, then $\eta^\circ(\det(x_{ij}) - 1) = \det(\eta^\circ(x_{ij})) - 1 = \det(A) - 1 = 0$, and therefore η° induces a homomorphism $\bar{\eta}: K = R[x_{ij}]/D \to U$ such that $\bar{\eta}(\xi_{ij}) = a_{ij}$, and we have the map $\bar{\eta}_n: K_n \to U_n$ for which $\bar{\eta}_n(\Xi) = A$, $\bar{\eta}_n(\Xi^{-1}) = A^{-1}$ and $\bar{\eta}_n(B) = B'$ for $B \in R_n$. For all $B \in R_n$ we have:

$\bar{\eta}_n\rho(B) = \bar{\eta}_n(\Xi B \Xi^{-1}) = \bar{\eta}_n(\Xi)\bar{\eta}_n(B)\bar{\eta}_n(\Xi^{-1}) = AB'A^{-1} = \tau(B)$ so that $\bar{\eta}_n\rho = \tau$. Let η be the restriction of $\bar{\eta}$ to S, then we have: $\rho(R_n) \subseteqq S_n$ and $\eta_n(S_n) \subseteqq C_n$. The last inclusion follows from the fact that for $B \in R_n$ $\eta([\rho(B)]_{ij} = [\eta_n\rho(B)]_{ij} = [\tau(B)]_{ij} \in C$ and since S is generated by the elements $[\rho(B)]_{ij}$, $\eta(S) \subseteqq C$. Consequently the following diagram is well defined and commutative.

$$\begin{tikzcd} R_n \arrow[r, "\rho"] \arrow[dr, "\tau"'] & S_n \arrow[d, "\eta_a"] \\ & C_n \end{tikzcd}$$

which completes the proof.

COROLLARY 7. (1) *If R is an integral domain (in particular $R = F$ a field) then so is $V_n(R_n)$.*

(2) *If R is noetherian so is $V_n(R_n)$.*

Proof. (1) If R is a domain then so is $R[x_{ij}]$ and the polynomial $\det(x_{ij}) - 1$ is prime in $R[x_{ij}]$. Hence, D is a prime ideal and $K = R[x_{ij}]/D$ is a domain. $V_n(R_n) = S \subseteqq K$ is hence also a domain.

(2) $S = V_n(R_n)$ is finitely generated over R (see also [1]).

4. **An alternative representation of $V_n(R_n)$.** $V_n(R_n)$ was shown to be generated by n^4 elements $\xi_{ki}\hat{\Xi}_{lj}$. We aim now to describe the ring in terms of these generators and their relations. To this end we begin with a ring $R[z_{kl}^{ij}]$ with z_{kl}^{ij} n^4 commutative indeterminates over R. The elements z_{kl}^{ij} are to represent the generators $\xi_{ki}\hat{\Xi}_{lj}$ and so they must satisfy the relations arising from the commutativity of the ξ_{ij} and from the fact that $\Xi\Xi^{-1} = \Xi^{-1}\Xi = I$, namely;

(1) $z_{kl}^{ij} z_{tr}^{sq} = z_{kr}^{iq} z_{tl}^{sj}$ for all i, j, k, l, s, q, t, r

(2) $\sum_{i=1}^n z_{kl}^{ii} = \delta_{kl}$ for all k, l

(3) $\sum_{k=1}^n z_{kk}^{ij} = \delta_{ij}$ for all i, j.

In fact we will show that the generators of $V_n(R_n)$ satisfy no other relations except these. This will be done by showing that $R[z_{kl}^{ij}]$ modulo those relations is again the universal ring $V_n(R_n)$.

We begin with conditions for matrices in a matrix ring to be matrix units.

THEOREM 8. *Let $\{F^{ij} \mid 1 \leqq i, j \leqq n\}$ be a set of n^2 matrices in a matrix ring C_n. The $\{F^{ij}\}$ are a set of matrix units in C_n if and only if they satisfy the following conditions*:

(1) $(F^{ij})_{kl}(F^{sq})_{tr} = (F^{iq})_{kr}(F^{sj})_{tl}$ for all i, j, k, l, s, q, t, r

(2) $\sum_{i=1}^n (F^{ii})_{kl} = \delta_{kl}$ for all k, l

(3) $\sum_{k=1}^n (F^{ij})_{kk} = \delta_{ij}$ for all i, j.

Proof. Suppose the F^{ij} are matrix units in C_n, then, by Theorem 3 we have $F^{ij} = AE_{ij}A^{-1}$ where $A \in U_n$, for some $U \supseteqq C$. Put $A = (a_{ij})$ and $A^{-1} = (a'_{ij})$ and evaluate the left side of (1);

$$\begin{aligned}(F^{ij})_{kl}(F^{sq})_{tr} &= (AE_{ij}A^{-1})_{kl}(AE_{sq}A^{-1})_{tr} \\ &= \Big(\sum_{\nu\mu} a_{k\nu}(E_{ij})_{\nu\mu}a'_{\mu l}\Big)\Big(\sum_{\pi\sigma} a_{t\pi}(E_{sq})_{\pi\sigma}a'_{\sigma r}\Big) \\ &= (a_{ki}a'_{jl})(a_{ts}a'_{qr}) .\end{aligned}$$

While the right side of (1) gives:

$$\begin{aligned}(F^{iq})_{kr}(F^{sj})_{tl} &= (AE_{iq}A^{-1})_{kr}(AE_{sj}A^{-1})_{tl} \\ &= \Big(\sum_{\nu\mu} a_{k\nu}(E_{iq})_{\nu\mu}a'_{\mu r}\Big)\Big(\sum_{\pi\sigma} a_{t\pi}(E_{sj})_{\pi\sigma}a'_{\sigma l}\Big) \\ &= (a_{ki}a'_{qr})(a_{ts}a'_{jl}) .\end{aligned}$$

Which, by the commutativity in U proves (1). To prove (2) we have only to notice that $\sum_{i=1}^n (F^{ii})_{kl} = (\sum_{i=1}^n F^{ii})_{kl} = (I)_{kl} = \delta_{kl}$. Condition (3) states that $tr(F^{ij}) = \delta_{ij}$, now for $i \neq j$ we have

$$tr(F^{ij}) = tr(F^{ii}F^{ij}) = tr(F^{ij}F^{ii}) = tr(0) = 0 = \delta_{ij}$$

while for $i = j$ we have, using (1) and (2)

$$\begin{aligned}tr(F^{ii}) &= 1\cdot\sum_{k=1}^n (F^{ii})_{kk} = \Big(\sum_{l=1}^n (F^{ll})_{ii}\Big)\Big(\sum_{k=1}^n (F^{ii})_{kk}\Big) \\ &= \sum_{l,k} (F^{li})_{ik}(F^{il})_{ki} = \sum_{l} (F^{ll})_{ii} = 1 = \delta_{ii} .\end{aligned}$$

Conversely, suppose the F^{ij} satisfy conditions 1, 2, 3 than

$$\begin{aligned}[(F^{\nu\mu})(F^{\pi\sigma})]_{kl} &= \sum_{t=1}^n (F^{\nu\mu})_{kt}(F^{\pi\sigma})_{tl} = \sum_{t=1}^n (F^{\nu\sigma})_{kl}(F^{\pi\mu})_{tt} \\ &= (F^{\nu\sigma})_{kl}\sum_{t=1}^n (F^{\pi\mu})_{tt} = (F^{\nu\sigma})_{kl}\delta_{\pi\mu} = (\delta_{\pi\mu}F^{\nu\sigma})_{kl} .\end{aligned}$$

This being true for all k, l we have $F^{\nu\mu}F^{\pi\sigma} = \delta_{\pi\mu}F^{\nu\sigma}$. We also have $(\sum_{i=1}^n F^{ii})_{kl} = \sum_{i=1}^n (F^{ii})_{kl} = \delta_{kl} = (I)_{kl}$ hence $\sum_{i=1}^n F^{ii} = I$ which concludes the proof that the F^{ij} are indeed matrix units. (A similar result on orthogonal idempotents was obtained in [7].)

We are finally in the position to present $V_n(R_n)$ in terms of generators and relations. We start as before with the ring $R[z_{kl}^{ij}]$ and let J be the ideal generated by all the polynomials of the forms: $z_{kl}^{ij}z_{tr}^{sq} - z_{kr}^{iq}z_{tl}^{sj}$, $\sum_{i=1}^{n} z_{kl}^{ii} - \delta_{kl}$, $\sum_{k=1}^{n} z_{kk}^{ij} - \delta_{ij}$ where the indices range over all possible combinations. We denote by ζ_{kl}^{ij} the class $z_{kl}^{ij} + J$ in the quotient ring $S' = R[x_{kl}^{ij}]/J = R[\zeta_{kl}^{ij}]$. We define now a homomorphism $\rho' \colon R_n \to S'_n$ by $[\rho'(A)]_{kl} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\zeta_{kl}^{ij}$.

The relations imposed on the ζ_{kl}^{ij} and Theorem 8 imply that the matrices $\rho'(E_{ij}) = (\zeta_{kl}^{ij})$ are clearly a set of matrix units in S' so that evidently ρ' is indeed an R-homomorphism.

THEOREM 9. *With these notations $S' = V_n(R_n)$ and ρ' is the canonical embedding of R_n in $S'_n = (V_n(R_n))_n$.*

Proof. We note that $[\rho'(E_{ij})]_{kl} = \zeta_{kl}^{ij}$ so that S' is generated by the appropriate elements. All that remains to be shown is the universal property of $(S'\rho')$. Let $\tau \colon R_n \to C_n$ be a homomophism, $F^{ij} = \tau(E_{ij})$ are then a set of matrix units in C_n. Define $\tilde{\eta} \colon R[z_{kl}^{ij}] \to C_n$ by $\tilde{\eta}(z_{kl}^{ij}) = [\tau(E_{ij})]_{kl} = (F^{ij})_{kl}$, by Theorem 8 $J \subseteqq \operatorname{Ker} \tilde{\eta}$ and so $\tilde{\eta}$ induces a map $\eta \colon S' \to C$ for which $\eta(\zeta_{kl}^{ij}) = [\tau(E_{ij})]_{kl}$. It remains to show that $\eta_n\rho' = \tau$, and clearly it is enough to demonstrate the equality on the generators E_{ij} of R_n, indeed $[\eta_n\rho'(E_{ij})]_{kl} = \eta[[\rho'(E_{ij})]_{kl}] = \eta(\zeta_{kl}^{ij}) = [\tau(E_{ij})]_{kl}$. This being true for all k, l and for all i, j we have $\eta_n\rho' = \tau$ as required.

REMARK. By the uniqueness of $V_n(R_n)$ S and S' should be isomorphic. The isomorphism is given by the correspondence $\theta \colon \zeta_{kl}^{ij} \to \xi_{ki}\hat{\Xi}_{lj}$ and $\rho = \theta\rho'$.

5. Embedding in matrix rings of different order. In this section we investigate the homomorphisms of a matrix ring into matrix rings of higher orders. In particular we give a description of the ring $V_m(R_n)$ for all n and m.

If n and m are integers such that n/m we have an injection $\delta \colon C_n \to C_m$ which places an $n \times n$ matrix m/n-times along the diagonal of an $m \times m$ matrix. The combined map $R_n \xrightarrow{i_n} C_n \xrightarrow{\delta} C_m$ will be denoted by δ'.

Our first result is elementary:

LEMMA 10. *There exists a unitary R homomorphism $\tau \colon R_n \to C_m$ if and only if n/m.*

Proof. If n/m we have exhibited such an homomorphism, namely δ'. Conversely, suppose there exists a $\tau \colon R_n \to C_m$. Let $\mathfrak{M}$ be a maximal ideal in C. We have the induced R homomorphism $\bar{\tau}: R_n \xrightarrow{\tau} C_m \xrightarrow{\pi_m} (C/\mathfrak{M})_m = K_m$, therefore, without loss of generality we may assume $C = K$ a field. Let E_{ij} be the standard matrix units of R_n, then $\tau(E_{ij}) = f_{ij}$ are n^2 matrix units in K_m. We have $K_n \cong \sum_{ij} f_{ij} K \subseteqq K_m$ and $K_m \cong K_n \otimes_K \mathrm{Cent}_{K_m}(K_n)$ by taking dimension over k we clearly have n/m.

Consequently we can assert that $V_m(R_n) \neq \{0\}$ if and only if m/n, which we shall assume henceforth. We would like now to generalize Theorem 3 to the case of n^2 matrix units in a matrix ring of order m. Turning to subdirectly irreducible components does not seem to be very helpful and so we localize. Our next result is again not new. It was proved for example by Knus [5] in a more general setting. Our proof is rather elementary except for the use of the classical Skolem-Noether theorem.

THEOREM 11. *Let $(C, \mathfrak{M})$ be a local ring. $\{e_{ij}\}$ and $\{f_{ij}\}$ two sets of n^2 matrix units in C_m (n/m), there exists an invertible matrix $A \in C_m$ such that $f_{ij} = Ae_{ij}A^{-1}$ for $1 \leqq i, j \leqq n$.*

Proof. $\{\pi_m(e_{ij})\}$ and $\{\pi_m(f_{ij})\}$ are two sets of n^2 matrix units in a matrix ring $(C/\mathfrak{M})_m$ over a field. By the Skolem-Noether theorem there exists an invertible matrix $\bar{y} \in (C/\mathfrak{M})_m$ such that $\pi_m(f_{ij}) = \bar{y}\pi_m(e_{ij})\bar{y}^{-1}$. Let y be a matrix in C_m with $\pi_m(y) = \bar{y}$. $0 \neq \det(\pi_m(y)) = \pi(\det(y)) \in C/\mathfrak{M}$, therefore, $\det(y) \notin \mathfrak{M}$ is invertible in C and y is an invertible matrix in C_m. Put $A = \sum_{\nu=1}^{n} f_{\nu 1} y e_{1\nu}$ then we have

$$(1) \qquad 1 \leqq i,\ j \leqq n \qquad f_{ij}A = f_{i1} y e_{1j} = Ae_{ij}$$

furthermore:

$$\begin{aligned} \pi_m(A)\bar{y}^{-1} &= \sum_{\nu=1}^{n} \pi_m(f_{\nu 1})\,\bar{y}\pi_m(e_{1\nu})\bar{y}^{-1} = \sum_{\nu=1}^{n} \pi_m(f_{\nu 1})\pi_m(f_{1\nu}) \\ &= \sum_{\nu=1}^{n} \pi_m(f_{\nu\nu}) = \pi_m(I) = I\,, \end{aligned}$$

so $\pi_m(A)$ is invertible in $(C/\mathfrak{M})_m$ and hence, as above, A is invertible in C_m which, by (1), completes the proof of the theorem.

The next theorem can be proved with the help of Theorem 11 in the same way that we proved Theorem 3.

THEOREM 12. *Let C be a commutative ring with unit. $\{e_{ij}\}$ and $\{f_{ij}\}$ two sets of n^2 matrix units in C_m (n/m). There exists a commutative ring $U \supseteqq C$ and an invertible matrix $A \in U_m$ such that $e_{ij} = Af_{ij}A^{-1}$ for all $1 \leqq i,\ j \leqq n$. The ring U is independent of* *the* $\{e_{ij}\}$ *and* $\{f_{ij}\}$. *For fixed* $\{e_{ij}\}$ *we may choose the ring* U *and the matrix* A *in such a way that* $\det(A) = 1$.

We can now show that $V_m(R_n)$ is a subalgebra of $V_m(R_m)$, more precisely:

THEOREM 13. *Let* $\rho: R_m \to (V_m^R(R_m))_m$ *and* $\delta: R_n \to R_m$ *be the canonical maps. Then* $V_m^R(R_n)$ *is the subalgebra* S, *generated in* $V_m^R(R_m)$ *by the entries* $\{[\rho\delta(B)_{ij} \mid B \in R_n\}$ *and* $\rho\delta$ *is the corresponding canonical embedding.*

Proof. The condition on the generators of S is fulfilled by definition so we have only to show the factoring property for maps. Let $\tau: R_n \to C_m$ be a homomorphism, by the usual reasoning we have a ring $U \supseteqq C$ and a matrix $A \in U_m$ such that for $B \in R_n$ $\tau(B) = A\delta'(B)A^{-1}$ where δ' is as in Lemma 10. Define $\tau': R_m \to U_m$ by $\tau'(D) = AD'A^{-1}$, we have the following commutative diagram:

$$\begin{tikzcd} R_n \arrow[r, "\delta"] \arrow[d, "\tau"'] & R_m \arrow[r, "\rho"] \arrow[d, "\tau'"'] & (V_m(R_m))_m \arrow[dl, dashed, "\bar{\eta}_m"] \\ C_m \arrow[r, hook] & U_m & \end{tikzcd}$$

The square on the left is commutative by the definitions, while $\bar{\eta}$ and the commutativity of the triangle are given by the universality of $V_m(R_m)$.

Define η to be the restriction of $\bar{\eta}$ to S. Then, for the generators of S we have $\eta([\rho\delta(B)]_{ij}) = [\eta_m\rho\delta(B)]_{ij} = [\tau(B)]_{ij} \in C$. Therefore, $\eta(S) \subseteqq C$ and the diagram

$$\begin{tikzcd} R_n \arrow[r, "\rho\delta"] \arrow[dr, "\tau"'] & S_m \arrow[d, "\eta_m"] \\ & C_m \end{tikzcd}$$

is well defined and commutative, which concludes the proof of the theorem.

REMARK. Obviously Corollary 7 may now be formulated for the ring $V_m^R(R_n)$.

REFERENCES

1. S. A. Amitsur, *Embedding in matrix rings*, Pacific J. Math., **36** (1971), 21-29.
2. L. N. Childs and F. R. DeMeyer, *On automorphisms of separable algebras*, Pacific J. Math., **23** (1967), 25-34.
3. N. Divinsky, *Commutative subdirectly irreducible rings*, Proc. Amer. Math. Soc., **8** (1957), 642-648.
4. J. Jacobson, *Structure of Rings*, Amer. Math. Soc. Colloquium Publications, V. **23** 1964.
5. M. A. Knus, *Sur le théorème de Skolem-Noether et sur les dérivations des algébras d'Azumaya*, C. R. Acad. Sc. Paris, **270** (1970), Série A, 637-639.
6. N. McCoy, *Subdirectly irreducible commutative rings*, Duke Math. J., **12** (1945), 381-387.
7. F. S. Nowlan, *A note on primitive idempotent elements of a total matric algebra*, Bull. Amer. Math. Soc., **37** (1931), 854-856.

Received July 5, 1972. This paper was written while the author was doing his Ph. D. thesis at the Hebrew University of Jerusalem, under the supervision of Professor S. A. Amitsur, to whom he wishes to express his warm thanks.

THE HEBREW UNIVERSITY OF JERUSALEM

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

$\widetilde{HD}$-MINIMAL BUT NO HD-MINIMAL

YOUNG K. KWON

Let U_{HD}^k (resp. $U_{\widetilde{HD}}^k$) be the class of Riemannian n-manifolds ($n \geqq 2$) on which there exist k non-proportional HD-minimal (resp. $\widetilde{HD}$-minimal) functions. The purpose of the present paper is to construct a Riemannian n-manifold $n \geqq 3$ which carries a unique (up to constant factors) $\widetilde{HD}$-minimal function but no HD-minimal functions. Thus the inclusion relation

$$U_{HD}^1 \subset U_{\widetilde{HD}}^1$$

is strict for $n \geqq 3$. By welding k copies of this Riemannian n-manifold, it is then established that the inclusion relation

$$U_{HD}^k \subset U_{\widetilde{HD}}^k$$

is strict for all $k \geqq 1$ and $n \geqq 3$. The problem still remains open for $n = 2$.

1. An HD-function (harmonic and Dirichlet-finite) ω on a Riemannian n-manifold M is called HD-minimal on M if ω is positive on M and every HD-function ω' with $0 < \omega' \leqq \omega$ reduces to a constant multiple of ω on M. Let $\{\omega_n\}$ be a sequence of positive HD-functions on M. If the sequence $\{\omega_n\}$ decreases on M, the limit function is harmonic on M by Harnack's inequality. Such a harmonic function is called an $\widetilde{HD}$-function on M, and $\widetilde{HD}$-minimality can be defined as in the case of HD-minimal functions.

These functions were introduced by Constantinescu and Cornea [1] and systematically studied by Nakai [6]. In particular the following characterization by Nakai is important (loc. cit., cf. also Kwon-Sario [5]):

(i) A Riemannian n-manifold M carries an HD-minimal function ω if and only if the Royden harmonic boundary Δ_M of M contains a point p, isolated in Δ_M. In this case $\omega(p) > 0$ and $\omega \equiv 0$ on $\Delta_M - \{p\}$.

(ii) A Riemannian n-manifold M carries an $\widetilde{HD}$-minimal function ω if and only if the Royden harmonic boundary Δ_M of M has a point p of positive harmonic measure. These are corresponded such that $\limsup_{x \in M, x \to p} \omega(x) > 0$ and $\limsup_{x \in M, x \to q} \omega(x) = 0$ for almost all $q \in \Delta_M - \{p\}$ with respect to a harmonic measure on Δ_M.

Since an isolated point of Δ_M has a positive harmonic measure, the above characterization yields the inclusion

$$U_{HD}^k \subset U_{\widetilde{HD}}^k$$

for all $k \geqq 1$.

For the notation and terminology we refer the reader to the monograph by Sario-Nakai [7].

2. Let $n \geqq 3$. Denote by M_0 the punctured Euclidean n-space $R^n - 0$ with the Riemannian metric tensor

$$g_{ij}(x) = |x|^{-4}(1 + |x|^{n-2})^{4/(n-2)}\delta_{ij}\,, \qquad 1 \leqq i,\ j \leqq n$$

where $|x| = [\sum_{i=1}^n (x^i)^2]^{1/2}$ for $x = (x^1, x^2, \cdots, x^n) \in M_0$.

For each pair (m, l) of positive integers m, l, set

$$H_{ml} = \{8^k x \in M_0 \mid |x| = 1 \text{ and } x^1 \geqq 0\}\,,$$

where $k = 2^{m-1}(2l-1) - 1$, and $ax = (ax^1, ax^2, \cdots, ax^n)$ for $x = (x^1, x^2, \cdots, x^n) \in M_0$ and real a. Let M_0' be the slit manifold obtained from M_0 by deleting all the closed hemispheres H_{ml}. Take a sequence $\{M_0'(l)\}_1^\infty$ of copies of M_0'. For each fixed $m \geqq 1$ and subsequently for fixed $j \geqq 0$ and $1 \leqq i \leqq 2^{m-1}$, connect $M_0'(i + 2^m j)$, crosswise along all the hemispheres $H_{ml}(l \geqq 1)$, with $M_0'(i + 2^{m-1} + 2^m j)$.

The resulting Riemannian n-manifold N is an infinitely sheeted covering manifold of M_0. Let π: $N \to M_0$ be the natural projection.

The following result is essential to our problem (Kwon [4]):

THEOREM 1. *A function $u(x)$ is harmonic on N if and only if $[1 + |\pi(x)|^{2-n}]u(x)$ is Δ_e-harmonic (harmonic with respect to the Euclidean structure) on N. In particular every bounded harmonic function $u(x)$ on the submanifold*

$$G = \left\{x \in N \,\middle|\, |\pi(x)| > \frac{1}{3}\right\}$$

is constant on $\pi^{-1}(x)$ for each $x \in M_0$ whenever it continuously vanishes on

$$\partial G = \left\{x \in N \,\middle|\, |\pi(x)| = \frac{1}{3}\right\}.$$

3. For each integer $l \geqq 1$, consider the subset of N:

$$N_l = [M_0'(l)] \cup \left[\bigcup_{i \neq l} G_i\right]$$

where

$$G_i = \left\{x \in M_0'(i) \,\middle|\, |\pi(x)| > \frac{1}{3}\right\}.$$

It is obvious that

$$G = \bigcup_{i=1}^{\infty} G_i$$

and the Riemannian n-manifold G is an infinitely sheeted covering manifold of the annulus $\{x \in M_0 \mid 1/3 < |x| < \infty\}$.

We are now ready to state our main result:

THEOREM 2. *The Riemannian n-manifold G $(n \geqq 3)$ carries a unique (up to constant factors) $\widetilde{HD}$-minimal function but no HD-minimal functions. Thus the inclusion*

$$U^1_{HD} \subset U^1_{\widetilde{HD}}$$

is strict for Riemannian manifolds of $\dim \geqq 3$.

The proof will be given in 4 − 5.

4. For $m \geqq 1$ construct $u_m \in HBD(N_m)$, the class of bounded HD-functions on N_m, such that $0 \leqq u_m \leqq 1$ on N, $u_m \equiv 0$ on $\bigcup_{i=1}^{m-1} [M_0'(i) - G_i]$, and $u_m \equiv 1$ on $\bigcup_{i=m+1}^{\infty} [M_0'(i) - G_i]$. Clearly $u_m \geqq u_{m+1}$ on N and therefore the sequence $\{u_m\}$ converges to an $\widetilde{HD}$-function u on G, uniformly on compact subsets of G. It is easy to see that $0 \leqq u < 1$ on G and $u \mid N - G \equiv 0$. Since

$$u_m(x) \geqq \frac{|\pi(x)|^{n-2} - 3^{2-n}}{|\pi(x)|^{n-2} + 1}$$

on G by maximum principle and Theorem 1, it follows that $0 < u < 1$ on G. Note that $\lim_{|\pi(x)| \to \infty} u_m(x) = 1$.

We claim that the function u is $\widetilde{HD}$-minimal on G. In fact, let $v \in \widetilde{HD}(G)$ be such that $0 < v \leqq u$ on G. In view of

$$0 \leqq \limsup_{x \in G, x \to y} v(x) \leqq \limsup_{x \in G, x \to y} u(x) = 0$$

for all $y \in \partial G$, v can be continuously extended to N by setting $v \equiv 0$ on $N - G$. By Theorem 1 v attains the same value at all the points in N which lie over the same point in M_0. Thus we may assume that u, v are bounded harmonic functions on $\pi(G) = \{\pi(x) \mid x \in G\}$ such that u, $v \equiv 0$ on $\pi(\partial G)$.

Again by Theorem 1, $(1 + |x|^{2-n})v(x)$ is Δ_e-harmonic on $\pi(G)$. In view of the fact that Δ_e-harmonicity is invariant by the Kelvin transformation, the function

$$\frac{1}{3^{n-2}|x|^{n-2}}(1 + 3^{2(n-2)}|x|^{n-2})v\left(\frac{x}{9|x|^2}\right)$$

is Δ_e-harmonic on M_0 for $0 < |x| < 1/3$ and continuously vanishes for

$|x| = 1/3$. Therefore, there exists a constant $a \geqq 0$ such that

$$v\left(\frac{x}{9|x|^2}\right) = \frac{3^{n-2}a}{1 + 3^{2(n-2)}|x|^{n-2}}$$

on M_0 for $0 < |x| < 1/3$ (cf., e.g. Helms [3, p. 81]). Thus

$$\lim_{x \to 0} v\left(\frac{x}{9|x|^2}\right) = 3^{n-2}a$$

exists and $v = 3^{n-2}au$ on G, as desired.

5. Suppose that there exists another $\widetilde{HD}$-minimal function ω on G. Choose a point $q \in \Delta_{M,G}$, the Royden harmonic boundary of G, such that q has a positive harmonic measure and

$$\limsup_{x \in G, x \to q'} \omega(x) = 0$$

for almost all $q' \in \Delta_{M,G} - \{q\}$ relative to a harmonic measure for G. Let $j: G^* \to \bar{G} \subset N^*$ be the subjective continuous mapping such that $j \mid G$ is the identity mapping and $f(x) = f(j(x))$ for all $x \in G^*$, the Royden compactification of G, and $f \in M(N)$, the Royden algebra of N. Here $\bar{G}$ is the closure of G in N^*. Note that a Borel set $E \subset \partial G$ has a positive harmonic measure if and only if $j^{-1}(E)$ has a positive harmonic measure (cf. Sario-Nakai [7, p. 192]). Therefore, $j(q) \notin \partial G$ and $\partial G \subset j(\Delta_{M,G})$.

For each $m \geqq 1$, $u_m(q) = u_m(j(q)) = 1$ since $j(q) \in \overline{\partial G} - \partial G$. Thus it is not difficult to see that $0 < \omega \leqq \beta u_m$ on G, where

$$\beta = \limsup_{x \in G, x \to q} \omega(x) > 0 .$$

Therefore, $0 < \omega \leqq \beta u$ on G and ω is a constant multiple of u on G as in 4.

It remains to show that u is not HD-minimal on G. If it were, u would have a finite Dirichlet integral. But u has a continuous extension to $G \cup \partial G$ with $u \mid \partial G \equiv 0$. Then by Theorem 1 u must attain the same value at all the points in G which lie over the same point in $\pi(G)$, a contradiction.

This completes the proof of Theorem 2.

6. Let G' be the Riemannian n-manifold obtained from G by deleting two disjoint closed subsets B, C, where

$$B = \left\{x \in M_0'(1) \mid |x| = \frac{9}{24} \text{ and } x^1 \geqq 0\right\},$$

$$C = \left\{x \in M_0'(1) \mid |x| = \frac{11}{24} \text{ and } x^1 \geqq 0\right\}.$$

For each $k \geqq 2$ take k copies $G_1, G_2, \cdots, G_k$ of G', and identify, crosswise, B_i with C_{i+1} for $1 \leqq i \leqq m$. Here we set $C_{n+1} = C_1$. Then it is easy to see that the resulting Riemannian n-manifold $G^{(k)}$ has exactly k non-proportional $\widetilde{HD}$-minimal functions but no HD-minimal functions.

COROLLARY. *For all* $k \geqq 1$ *the strict inclusion*

$$U^k_{HD} < U^k_{\widetilde{HD}}$$

holds for Riemannian manifolds of dim $\geqq 3$.

7. For the sake of completeness we shall sketch a proof of Theorem 1. In view of the simple relation

$$\Delta u = |x|^{n+2}(1 + |x|^{n-2})^{-(n+2)/(n-2)} \cdot \Delta_e[(1 + |\pi x|^{2-n})u] ,$$

it suffices to show the latter half.

For each integer $k \geqq 0$ let U_k be a component of the open set

$$\{x \in N \mid 2^{3k-1} < |\pi(x)| < 2^{3k+1}\} ,$$

and S_k a compact subset of U_k which lie over the set

$$\{x \in M_0 \mid |x| = 2^{3k}\} .$$

Since U_k is a magnification of U_0 and the Δ_e-harmonicity is invariant under a magnification, it is not difficult to see that there exists a constant q, $0 < q < 1$, such that

$$|u(x)| \leqq q \cdot \sup\{|u(x)| \mid x \in U_k\}$$

on S_k for any harmonic function u on U_k which changes sign on S_k. Note that q is independent of k.

Let u be a harmonic function on G such that $|u| \leqq 1$ and it continuously vanishes on ∂G. For each $m \geqq 1$, denote by π_m the cover transformation of G which interchanges the sheets of G: the points in $G \cap M_0'(i + 2^m j)$ are interchanged with points, with the same projection, in $M_0'(i + 2^{m-1} + 2^m j)$ for $j \geqq 0$ and $1 \leqq i \leqq 2^{m-1}$. Define v_m on G by

$$v_m(x) = \frac{1}{2}[u(x) - u(\pi_m(x))] .$$

Clearly v_m is harmonic on G, $|v_m| \leqq 1$, and v_m changes sign on S_k, $k = 2^{m-1}(2l - 1) - 1$. Therefore,

$$\max\{|v_m(x)| \mid x \in S_k\} \leqq q$$

for all $l \geqq 1$. By induction on l, we derive that $|v_m| \leqq q^l$ on $S_{k'}$, where $k' = 2^{m-1} - 1$. Letting $l \to \infty$, we conclude that $v_m \equiv 0$ on G, as desired.

REFERENCES

1. C. Constantinescu and A. Cornea, *Über den idealen Rand und einige seiner Anwendungen bei der Klassifikation der Riemannschen Flächen*, Nagoya Math. J., **13** (1958), 169–233.
2. ———, *Ideale Ränder Riemannscher Flächen*, Springer, 1963, 244.
3. L. Helms, *Introduction to Potential Theory*, Wiley-Interscience, 1969, 282.
4. Y. K. Kwon, *Strict inclusion $O_{HB} < O_{HD}$ for all dimensions*, (to appear).
5. Y. K. Kwon and L. Sario, *Harmonic functions on a subregion of a Riemannian manifold*, J. Ind. Math. Soc., **35** (1971), 135–161.
6. M. Nakai, *A measure on the harmonic boundary of a Riemann surface*, Nagoya Math. J., **17** (1960), 181–218.
7. L. Sario and M. Nakai, *Classification Theory of Riemann Surfaces*, Springer, 1970, 446.

Received August 21, 1972 and in revised form January 17, 1973.

UNIVERSITY OF TEXAS

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

ON THE RENEWAL FUNCTION WHEN SOME OF THE MEAN RENEWAL LIFETIMES ARE INFINITE

MAKOTO MAEJIMA

Let $\{X_i,\ i = 1, 2, \cdots\}$ be a sequence of independent and nonnegative random variables with the distribution function $F_i(x)$. Some of $\int_0^\infty x dF_i(x)$ may be infinite. Let $H(t)$ be the renewal function. The main object of this note is to show that in order to have the asymptotic relation $H(t)/t \sim 1/L(t)$ as $t \to \infty$, it is necessary and sufficient that $\mu(t) \sim L(t)$ as $t \to \infty$, where $L(t)$ is a function of slow growth and $\mu(t) = \lim_{n\to\infty} (1/n) \sum_{i=1}^n \mu_i(t)$, $\mu_i(t)$ being $\int_0^t [1 - F_i(x)]dx$, is supposed to exist uniformly in t.

Let $H(t)$ be the renewal function for a renewal process, that is, a sequence $\{X_i,\ i = 1, 2, \cdots\}$ of nonnegative, independent and identically distributed random variables. Namely $H(t) = EN(t) = E[\sup\{n; S_n \leqq t\}]$, where $S_n = \sum_{i=1}^n X_i$. Smith [3] has studied the limiting behaviors of $H(t)/t$ for the case in which $EX_i = \infty$.

We now consider an extended renewal process in which X_i, $i = 1, 2, \cdots$ may not be identically distributed. We also in this case use the similar notations S_n and $N(t)$, and we may also define $H(t)$ in the similar manner under the condition that S_n has no finite limit point. The main object of this note is to give a generalization of a result of Smith to our extended case.

2. Some lemmas.

We begin with some lemmas for an extended renewal process with the finite mean lifetimes.

Let $\{X_i, i = 1, 2, \cdots\}$ be a sequence of independent and nonnegative random variables with $0 < EX_i = \mu_i < \infty$ and let $F_i(x)$ be the distribution function of X_i.

LEMMA 1. *Suppose that*

$$\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mu_i > 0 \tag{2.1}$$

exists and that

$$\lim_{A\to\infty} \int_A^\infty x dF_i(x) = 0$$

holds uniformly with respect to i. Then we have $EN^\alpha(t) < \infty$ for each $t > 0$, for $\alpha = 1, 2, \cdots$.

This lemma was first proved by Kawata [2] for $\alpha = 1$, and Hatori [1] showed it for any positive integer α.

LEMMA 2. *Suppose that $EN(t)$ and $EN^2(t)$ are finite and that* (2.1) *is true. Then we have for every t*

$$ES_{N(t)+1} = \mu(H(t) + 1) + \sum_{n=1}^{\infty} n\varepsilon_n \Pr\{N(t) + 1 = n\},$$

where ε_n is defined by

$$\varepsilon_n = \frac{1}{n}\sum_{i=1}^{n} \mu_i - \mu$$

which converges to zero as $n \to \infty$.

Proof. Letting

$$\begin{aligned} Z_n &= 1, \text{ if } n \leqq N(t) + 1, \\ &= 0, \text{ otherwise,} \end{aligned}$$

we have

$$ES_{N(t)+1} = E\sum_{n=1}^{N(t)+1} X_n = E\sum_{n=1}^{\infty} X_n Z_n. \tag{2.2}$$

Since

$$\begin{aligned} \{Z_n = 0\} &= \{N(t) + 1 < n\} = \bigcup_{k=1}^{n-1}\{N(t) + 1 = k\} \\ &= \{X_1 > t\} \cup \bigcup_{k=2}^{n-1}\{(X_1 + \cdots + X_{k-1} \leqq t) \cap (X_1 + \cdots + X_k > t)\}, \end{aligned}$$

Z_n is independent of X_n. Thus, noticing the nonnegativeness of X_n, we see that (2.2) is

$$\sum_{n=1}^{\infty} EX_n Z_n = \sum_{n=1}^{\infty} EX_n EZ_n = \sum_{n=1}^{\infty} \mu_n \Pr\{N(t) + 1 \geqq n\},$$

which turns out to be

$$\begin{aligned} ES_{N(t)+1} &= \sum_{n=1}^{\infty} (\mu + n\varepsilon_n - (n-1)\varepsilon_{n-1}) \Pr\{N(t) + 1 \geqq n\} \\ &= \mu(H(t) + 1) + \sum_{n=1}^{\infty} (n\varepsilon_n - (n-1)\varepsilon_{n-1}) \Pr\{N(t) + 1 \geqq n\}. \end{aligned}$$

Since

$$\sum_{n=1}^{\infty} |n\varepsilon_n \Pr\{N(t) + 1 \geqq n\}| \leqq \sup_n |\varepsilon_n| (EN^2(t) + 2) < \infty$$

by the finiteness of $EN^2(t)$, we may rewrite

$$\sum_{n=1}^{\infty} (n\varepsilon_n - (n-1)\varepsilon_{n-1}) \Pr\{N(t)+1 \geqq n\} = \sum_{n=1}^{\infty} n\varepsilon_n \Pr\{N(t)+1 = n\} ,$$

so that

$$ES_{N(t)+1} = \mu(H(t)+1) + \sum_{n=1}^{\infty} n\varepsilon_n \Pr\{N(t)+1 = n\} ,$$

which is the conclusion.

3. A theorem. We return to the case where X_i may have the infinite mean renewal lifetimes. Let $L(t)$ be a function of slow growth, that is, for every fixed $c > 0$, $L(ct)/L(t) \to 1$ as $t \to \infty$. We shall show the following theorem which is an extension of a result due to Smith ([3], Theorem 1, (i), $\nu = 1$) to the case of nonidentically distributed random variables.

THEOREM. *Let $\{X_i, i = 1, 2, \cdots\}$ be a sequence of independent and nonnegative random variables with the distribution function $F_i(x)$. Suppose that*

$$\mu(t) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mu_i(t) > 0 \tag{3.1}$$

exists uniformly in $0 < t < \infty$, where

$$\mu_i(t) = \int_0^t [1 - F_i(x)]dx .$$

Then the necessary and sufficient condition for the validity of the asymptotic relation

$$\frac{H(t)}{t} \sim \frac{1}{L(t)} , \quad as \quad t \to \infty , \tag{3.2}$$

where $L(t)$ is a function of slow growth, is that

$$\mu(t) \sim L(t) , \quad as \quad t \to \infty . \tag{3.3}$$

Before proving the theorem we shall show some lemmas.

We now define a new renewal process $\{X_i^*\}$ for a fixed positive number t^* by putting

$$X_i^* = X_i , \quad \text{if } X_i \leqq t^* ,$$
$$= t^* , \quad \text{otherwise} .$$

We note that $EX_i^* = \mu_i(t^*)$ is finite. For the new variables X_i^*, we define S_n^*, $N^*(t)$ and $H^*(t)$ in obvious ways. Then we may easily verify the conditions of Lemma 1 for a fixed t^* and the following lemma is immediate.

LEMMA 3. *Suppose that* (3.1) *exists for* t^*. *Then* $E\{N^*(t)\}^\alpha < \infty$ *for* $\alpha = 1, 2, \cdots$.

The next two lemmas play essential roles in the proof of Theorem.

LEMMA 4. *Suppose that* (3.1) *exists uniformly in* t. *Then we have*

$$\liminf_{t\to\infty} \frac{H(t)\mu(t)}{t} \geqq 1 .$$

Proof. We consider X_i^* defined above. Since $EN^*(t)$ and $E\{N^*(t)\}^2$ are finite by Lemma 3, we have that for all t,

$$t < \mu(t^*)(H^*(t) + 1) + \sum_{n=1}^{\infty} n\varepsilon_n(t^*) \Pr\{N^*(t) + 1 = n\} \tag{3.4}$$

by Lemma 2 and noting $t < S_{N(t)+1}$, where $\varepsilon_n(t^*)$ is defined by

$$\frac{1}{n}\sum_{i=1}^{n} \mu_i(t^*) = \mu(t^*) + \varepsilon_n(t^*) .$$

Now (3.4) holds for $t = t^*$, in particular. Thus we have

$$t^* < \mu(t^*)(H^*(t^*) + 1) + \sum_{n=1}^{\infty} n\varepsilon_n(t^*) \Pr\{N^*(t^*) + 1 = n\} . \tag{3.5}$$

Next, we estimate of the order of $\varepsilon_n(t)$ as $t \to \infty$. Since the function $1 - F_i(x)$ decreases to zero as $x \to \infty$, so does $\mu_i(t)/t$ as $t \to \infty$. In view of the assumption that (3.1) exists uniformly in t, it follows that, for any $\varepsilon > 0$, there exists a constant N independent of t such that

$$\left|\mu(t) - \frac{1}{n}\sum_{i=1}^{n} \mu_i(t)\right| < \varepsilon , \quad \text{for} \quad n \geqq N . \tag{3.6}$$

Then we have

$$\begin{aligned} \frac{1}{t}\,|\,\mu(t)\,| &\leqq \frac{1}{t}\left|\mu(t) - \frac{1}{N}\sum_{i=1}^{N} \mu_i(t)\right| + \frac{1}{t}\cdot\frac{1}{N}\sum_{i=1}^{N} |\,\mu_i(t)\,| \\ &< \frac{\varepsilon}{t} + \varepsilon < 2\varepsilon \end{aligned}$$

for sufficiently large t, taking into account the fact that $\mu_i(t)/t \to 0$ as $t \to \infty$. Thus, we have for sufficiently large t

$$\frac{1}{t}\,|\,\varepsilon_n(t)\,| = \frac{1}{t}\left|\mu(t) - \frac{1}{n}\sum_{i=1}^{n}\mu_i(t)\right| < \frac{\varepsilon}{N^2}\,, \tag{3.7}$$

for the fixed N and for all $n \leqq N$. Therefore we have, for large t^*, from (3.6) and (3.7)

$$\begin{aligned}
&\sum_{n=1}^{\infty} n\varepsilon_n(t^*) \Pr\{N^*(t^*) + 1 = n\} \\
&< \frac{\varepsilon}{N^2} t^* \sum_{n=1}^{N} n \Pr\{N^*(t^*) + 1 = n\} + \varepsilon \sum_{n=N+1}^{\infty} n \Pr\{N^*(t^*) + 1 = n\} \\
&< \varepsilon(t^* + H^*(t^*) + 1)\,.
\end{aligned} \tag{3.8}$$

Now we shall show that

$$\limsup_{t^*\to\infty} \frac{H^*(t^*)}{t^*} < \infty\,. \tag{3.9}$$

In order to show this, we define new truncated random variables $X_{i,A}$ for some constant A by putting

$$\begin{aligned}
X_{i,A} &= X_i\,, \quad \text{if} \quad X_i \leqq A\,, \\
&= A\,, \quad \text{otherwise}\,.
\end{aligned}$$

Clearly $EX_{i,A} = \mu_i(A)$ is finite and by the elementary renewal theorem for an extended renewal process, we have that, if $H_A(t)$ is the renewal function associated with $\{X_{i,A}\}$, then

$$\lim_{t\to\infty} \frac{H_A(t)}{t} = \frac{1}{\mu(A)}\,.$$

(For details, see Kawata [2].) (3.9) follows from the remark that $H^*(t^*) \leqq H_A(t^*)$ for $t^* \geqq A$. Since ε is arbitrary in (3.8), we have from (3.8)

$$\lim_{t^*\to\infty} \frac{1}{t^*}\left|\sum_{n=1}^{\infty} n\varepsilon_n(t^*) \Pr\{N^*(t^*) + 1 = n\}\right| = 0\,. \tag{3.10}$$

Therefore, from (3.5)

$$\liminf_{t^*\to\infty} \frac{1}{t^*}\mu(t^*)(H^*(t^*) + 1) \geqq 1\,.$$

On the other hand, we have

$$\Pr\{S_n^* \leqq t^*\} = \Pr\{S_n \leqq t^*\}\,,$$

for $n = 2, 3, \cdots$, and

$$\Pr\{S_1^* \leqq t^*\} = 1\,.$$

Thus

$$H^*(t^*) = H(t^*) + \Pr\{X_1 > t^*\}$$

and so

$$\liminf_{t\to\infty} \frac{1}{t} \mu(t)(H(t) + \Pr\{X_1 > t\} + 1) \geqq 1 .$$

Noticing that $\mu(t)/t \to 0$ as $t \to \infty$, we have the conclusion of the lemma.

LEMMA 5. *Under the same conditions as in Lemma 4, we have for arbitrary* $\delta > 0$

$$\limsup_{t\to\infty} \frac{H(t)\mu(\delta t)}{t} \leqq 1 + \delta .$$

Proof. Take $\delta > 0$ arbitrarily and let $\hat{X}_n$ represent new variables truncated according to the rule

$$\begin{aligned} \hat{X}_n &= X_n , \quad \text{if} \quad X_n \leqq \delta t^* , \\ &= \delta t^* , \quad \text{otherwise} . \end{aligned}$$

It is clear that $E\hat{X}_n = \mu_n(\delta t^*) < \infty$. Then, noting that $t \geqq S_{N(t)+1} - X_{N(t)+1}$, we have, by Lemma 2,

$$\begin{aligned} t &\geqq \mu(\delta t^*)(\hat{H}(t) + 1) + \sum_{n=1}^{\infty} n\varepsilon_n(\delta t^*) \Pr\{\hat{N}(t) + 1 = n\} - E\hat{X}_{\hat{N}(t)+1} \\ &\geqq \mu(\delta t^*)(\hat{H}(t) + 1) + \sum_{n=1}^{\infty} n\varepsilon_n(\delta t^*) \Pr\{\hat{N}(t) + 1 = n\} - \delta t^* , \end{aligned} \tag{3.11}$$

where $\hat{N}(t)$ and $\hat{H}(t)$ are defined in the renewal process associated with the new truncated variables $\{\hat{X}_n\}$. Since (3.11) holds for $t = t^*$, in particular, we have

$$(1 + \delta)t^* \geqq \mu(\delta t^*)(\hat{H}(t^*) + 1) + \sum_{n=1}^{\infty} n\varepsilon_n(\delta t^*) \Pr\{\hat{N}(t^*) + 1 = n\} .$$

The same arguments as in the proof of Lemma 4 yield that

$$\lim_{t^*\to\infty} \frac{1}{t^*} \left| \sum_{n=1}^{\infty} n\varepsilon_n(\delta t^*) \Pr\{\hat{N}(t^*) + 1 = n\} \right| = 0 \tag{3.12}$$

for the fixed $\delta > 0$. Noting that

$$\hat{H}(t^*) \geqq H(t^*) ,$$

we have the required result.

We now turn to the proof of the theorem.

Proof of Theorem. We first assume that

$$\frac{H(t)}{t} \sim \frac{1}{L(t)}\,, \quad \text{as} \quad t \to \infty\,.$$

By Lemma 4 we have

$$\liminf_{t\to\infty} \frac{\mu(t)}{L(t)} \geqq 1\,, \tag{3.13}$$

and by Lemma 5, for any $\delta > 0$,

$$\limsup_{t\to\infty} \frac{\mu(\delta t)}{L(t)} \leqq 1 + \delta\,.$$

Writing δt for t, and using the fact that $L(t/\delta) \sim L(t)$ as $t \to \infty$, we have

$$\limsup_{t\to\infty} \frac{\mu(t)}{L(t)} \leqq 1 + \delta\,.$$

Since δ can be arbitrarily small, we, taking into account (3.13), conclude the necessity part.

Furthermore, in view of the assumption $\mu(t)$ is a function of slow growth, it follows by Lemma 5 that

$$\limsup_{t\to\infty} \frac{H(t)\mu(t)}{t} < 1 + \delta\,.$$

Since δ is arbitrary, Lemma 4 gives the sufficiency part.

When $\lim_{t\to\infty} \mu(t) = \infty$, we can relax slightly the condition of the uniform existence of $\mu(t)$ in the following way.

COROLLARY. *Suppose that*

$$\mu(t) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mu_i(t) > 0$$

exists for all t, (not necessarily uniformly), and that there exists a constant K, independent of t, such that

$$\left| \mu(t) - \frac{1}{n} \sum_{i=1}^{n} \mu_i(t) \right| = |\,\varepsilon_n(t)\,| < K$$

for $n \geqq N$, N being some finite positive integer. If $\lim_{t\to\infty}\mu(t) = \infty$, *then the necessary and sufficient condition for the validity of the asymptotic relation* (3.2) *is* (3.3).

Proof. In the proof of theorem, the condition relaxed has been used only in order to show (3.10) and (3.12). Thus, it suffices to show that (3.10) holds under the conditions of this corollary.

Now, we have

$$\frac{1}{t}\,|\mu(t)| \leqq \frac{1}{t}\left|\mu(t) - \frac{1}{N}\sum_{i=1}^{N}\mu_i(t)\right| + \frac{1}{t}\cdot\frac{1}{N}\sum_{i=1}^{N}|\mu_i(t)|$$
$$< \frac{K}{t} + \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\mu_i(t)}{t}\right| ,$$

and so $|\mu(t)|/t$ can be arbitrarily small for the sufficiently large t. Thus, $\varepsilon_n(t) = o(t)$ for all $n \leqq N$. Therefore, we have

$$\left|\sum_{n=1}^{\infty} n\varepsilon_n(t^*)\Pr\{N^*(t^*)+1=n\}\right|$$
$$< o(t^*)\sum_{n=1}^{N} n\Pr\{N^*(t^*)+1=n\} + K\sum_{n=N+1}^{\infty} n\Pr\{N^*(t^*)+1=n\}$$
$$< o(t^*)N^2 + K(H^*(t^*)+1) .$$

Now we shall show under the condition that $\mu(t)\to\infty$, that

$$\lim_{t^*\to\infty}\frac{H^*(t^*)}{t^*} = 0 .$$

As in the proof of the previous theorem, we have

$$\limsup_{t^*\to\infty}\frac{H^*(t^*)}{t^*} \leqq \lim_{t^*\to\infty}\frac{H_A(t^*)}{t^*} = \frac{1}{\mu(A)} .$$

Since A is arbitrary, this shows that

$$\lim_{t^*\to\infty}\frac{H^*(t^*)}{t^*} = 0 ,$$

and (3.10) holds.

The author wishes to express his sincere appreciation to Professor Tatsuo Kawata of Keio University for continuing guidances and encouragements.

References

1. H. Hatori, *A note on a renewal theorem*, Kodai Math. Sem. Rep., **12** (1960), 28–37.
2. T. Kawata, *A renewal theorem*, J. Math. Soc. Japan, **8** (1956), 118–126.
3. W. L. Smith, *A note on the renewal function when the mean renewal lifetime is infinite*, J. Roy. Stat. Soc., **B23** (1961), 230–237.

Received March 3, 1973.

Keio University

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No 1, 1973

COHOMOLOGICAL DIMENSION OF DISCRETE MODULES OVER PROFINITE GROUPS

JUAN JOSÉ MARTÍNEZ

The main purpose of this note is to show that the finiteness of the cohomological dimension of a discrete module is closely related to the finiteness of its injective dimension. Moreover, a sufficient condition for the finiteness of the cohomological dimension is given. Both results are proved making a heavy use of the theory of cohomological triviality for finite groups.

The reader is referred to [3] for a treatment of profinite cohomology.

Throughout this note, G is a profinite group. As usual, the cohomology of G is denoted by $H(G,\ )$.

Recall that, if A is a discrete G-module, the infimum of the (set of) nonnegative integers r such that $H^n(S, A) = 0$, for any integer $n > r$ and any closed subgroup S of G, is called the *cohomological dimension* of A, and is denoted by $cd(G, A)$. If S is a closed subgroup of G, $H^n(S, A) \cong \varinjlim H^n(V, A)$, where V runs through all open subgroups of G containing S [3, Chap. I, Proposition 8, p. I–9]. Hence, if $H^n(V, A) = 0$ for every open subgroup V of G, then $H^n(S, A) = 0$ for every closed subgroup S of G.

In this paper, a discrete module is called *injective* only when it is injective in the corresponding category of discrete modules. If A is injective, it is well-known that $cd(G, A) = 0$, because, for instance, A is V-injective for all open subgroups V of G. Finally, recall that the *injective dimension* of A, denoted by $id(G, A)$, is the least length of an injective resolution of A.

The connection between cohomologically trivial modules over finite groups [2, Chap. IX, § 3, p. 148] and discrete modules of cohomological dimension zero over profinite groups was observed, and used, by Tate in his duality theory for profinite cohomology [3, Annexe au Chapitre I, p. I–79]. Tate's observation is quoted, for future reference, in the following.

LEMMA 1. *Let A be a discrete G-module. Then, $cd(G, A) = 0$ if, and only if, for every open, normal subgroup U of G, the G/U-module A^U is cohomologically trivial.*

Proof. See [3, Annexe au Chapitre I, Lemme 1, p. I–82]. Notice that G/U is a finite group, because G is compact and U is open.

The Nakayama-Tate criterion for cohomological triviality takes the following form, in the cohomology theory of profinite groups.

PROPOSITION 2. *Let A be a discrete G-module. If there exists a positive integer q such that $H^q(V, A) = H^{q+1}(V, A) = 0$ for all open subgroups V of G, then $cd(G, A) < q$.*

Proof. Since A embeds in an injective, whose cohomological dimension is zero, by repeated applications of dimension-shifting it suffices to consider the case $q = 1$. Let U be an open, normal subgroup of G. If V is any subgroup of G containing U, the Hochschild-Serre spectral sequence of the V/U-module A^U yields the exact sequence for low degrees

$$\begin{aligned} 0 \longrightarrow H^1(V/U, A^U) \longrightarrow H^1(V, A) \longrightarrow H^1(U, A)^{V/U} \\ \longrightarrow H^2(V/U, A^U) \longrightarrow H^2(V, A) \ . \end{aligned}$$

Since U is open, so is V, and thus, $H^1(U, A) = H^1(V, A) = H^2(V, A) = 0$. Therefore, $H^1(V/U, A^U) = H^2(V/U, A^U) = 0$, and applying the Nakayama-Tate criterion [2, Chap. IX, Théorème 8, p. 152], the G/U-module A^U is cohomologically trivial. By (1), the proof is complete.

The main result of this paper can be stated as follows.

THEOREM 3. *Let A be a discrete G-module, and let q be a positive integer. Then, $id(G, A) \leqq q$ if, and only if, $cd(G, A) \leqq q$ and $H^q(U, A)$ is a divisible abelian group for every open, normal subgroup U of G.*

Proof. Assume the assertion true for $q - 1$, with $q > 1$. If $id(G, A) \leqq q$, A has an injective resolution of length $\leqq q$, say

$$0 \longrightarrow A \xrightarrow{e} X_0 \xrightarrow{d_0} X_1 \longrightarrow \cdots \longrightarrow X_{q-1} \xrightarrow{d_{q-1}} X_q \longrightarrow 0 \ .$$

If $B = \operatorname{Coker} e$ and $f\colon X_0 \to B$ is the canonical morphism, the sequence of discrete G-modules

$$0 \longrightarrow A \xrightarrow{e} X_0 \xrightarrow{f} B \longrightarrow 0$$

is exact. Since $cd(G, X_0) = 0$ (injectivity of X_0), from the corresponding cohomology sequence it follows that

$$H^n(S, B) \cong H^{n+1}(S, A)$$

for any positive integer n and any closed subgroup S of G. Therefore, it is enough to prove that $cd(G, B) \leqq q - 1$, and that $H^{q-1}(U, B)$ is divisible for all open, normal subgroups U of G. By the induction hypothesis, this follows from showing that $id(G, B) \leqq q - 1$. In fact, if $e'\colon B \to X_1$ is the morphism induced by $d_0\colon X_0 \to X_1$, then $\operatorname{Ker} e' = 0$ and $\operatorname{Im} e' = \operatorname{Im} d_0$. Thus, the sequence

$$0 \longrightarrow B \xrightarrow{e'} X_1 \xrightarrow{d_1} X_2 \longrightarrow \cdots \longrightarrow X_{q-1} \xrightarrow{d_{q-1}} X_q \longrightarrow 0$$

is exact.

Reciprocally, if $cd(G, A) \leqq q$, let

$$0 \longrightarrow A \xrightarrow{g} Q \xrightarrow{h} C \longrightarrow 0$$

be an exact sequence of discrete G-modules, with Q injective. Then, $cd(G, C) \leqq q-1$, because

$$H^n(S, C) \cong H^{n+1}(S, A)$$

for all positive integers n and all closed subgroups S of G. By the same reason, if $H^q(U, A)$ is divisible for every open, normal subgroup U of G, then so is $H^{q-1}(U, C)$. Hence, by induction, C admits an injective resolution of length $\leqq q-1$, say

$$0 \longrightarrow C \xrightarrow{i} Y_0 \xrightarrow{d_0} Y_1 \longrightarrow \cdots \longrightarrow Y_{q-2} \xrightarrow{d_{q-2}} Y_{q-1} \longrightarrow 0 .$$

Since $\operatorname{Ker} ih = \operatorname{Ker} h$ and $\operatorname{Im} ih = \operatorname{Im} i$, the sequence

$$0 \longrightarrow A \xrightarrow{g} Q \xrightarrow{ih} Y_0 \xrightarrow{d_0} Y_1 \longrightarrow \cdots \longrightarrow Y_{q-2} \xrightarrow{d_{q-2}} Y_{q-1} \longrightarrow 0$$

is exact, and so $id(G, A) \leqq q$.

It remains to prove the assertion for $q = 1$.

Let

$$0 \longrightarrow A \longrightarrow X_0 \longrightarrow X_1 \longrightarrow 0$$

be an exact sequence of discrete G-modules, where X_0 and X_1 are injectives. Since $cd(G, X_0) = cd(G, X_1) = 0$, passing to cohomology it follows that $cd(G, A) \leqq 1$, and that the connecting operator $\partial_S\colon X_1^S \to H^1(S, A)$ is an epimorphism for all closed subgroups S of G. But, if D is any injective, discrete G-module and U is any open, normal subgroup of G, it is easy to check that D^U is an injective G/U-module, whence [2, Chap. IX, Lemme 7, p. 153] implies D^U is divisible. Therefore, as the image of a divisible group, $H^1(U, A)$ is divisible for all open, normal subgroups U of G.

Reciprocally, suppose $cd(G, A) \leqq 1$, and let

$$0 \longrightarrow A \longrightarrow Y_0 \longrightarrow Y_1 \longrightarrow 0$$

be an exact sequence of discrete G-modules, with Y_0 injective. Since $cd(G, Y_0) = 0$, taking cohomology it follows that $cd(G, Y_1) = 0$, and that the sequence of abelian groups

$$Y_0^S \longrightarrow Y_1^S \xrightarrow{\partial_S} H^1(S, A) \longrightarrow 0$$

is exact for all closed subgroups S of G. If U is an open, normal subgroup of G, $\operatorname{Ker}\partial_U$ is divisible, because so is Y_0^U. Therefore, if $\operatorname{Im}\partial_U = H^1(U, A)$ is divisible, then $\operatorname{Dom}\partial_U = Y_1^U$ is also divisible, and the proof is complete applying to Y_1 the following.

PROPOSITION 4. *Let A be a discrete G-module. If $cd(G, A) = 0$, and A^U is a divisible abelian group for every open, normal subgroup U of G, then A is injective.*

Proof. Recall that the category of discrete G-modules has injective envelopes for each of its objects. Since $(\boldsymbol{Z}[G/U])_U$, where U runs through all open, normal subgroups of G, is a family of generators, this result can be obtained by using a general theorem from category theory, due to Mitchell [1, Chap. III, Theorem 3.2, p. 89].

Let $f\colon A \to Q$ be an injective envelope of A (in the category of discrete G-modules). If $C = \operatorname{Coker} f$ and $g\colon Q \to C$ is the canonical morphism, the sequence of discrete G-modules

$$0 \longrightarrow A \xrightarrow{f} Q \xrightarrow{g} C \longrightarrow 0$$

is exact. Thus, if U is an open, normal subgroup of G, the sequence of G/U-modules

$$0 \longrightarrow A^U \xrightarrow{f^U} Q^U \xrightarrow{g^U} C^U \longrightarrow 0$$

is exact, because $cd(G, A) = 0$. Since Q^U is an injective G/U-module and $R \cap \operatorname{Im} f^U = R \cap \operatorname{Im} f$ for any sub-G/U-module R of Q^U (because, regarding R as a G-module, U operates trivially on R), $f^U\colon A^U \to Q^U$ is an injective envelope of A^U (in the category of G/U-modules). On the other hand, since $cd(G, A) = 0$, A^U is a cohomologically trivial G/U-module, by (1). Thus, A^U is G/U-injective [2, Chap. IX, Théorème 10, p. 154], and hence, $C^U = 0$ [1, Chap. III, Proposition 2.5, p. 88]. Since $C = \bigcup C^U$, $C = 0$, whence the result.

COROLLARY 5. *Let A be a discrete G-module, and let r be a nonnegative integer. If $cd(G, A) \leqq r$, then $id(G, A) \leqq r + 1$.*

Proof. Take $q = r + 1$ in (3).

This result can be applied to profinite groups of finite dimension, as follows.

COROLLARY 6. *Let r be a nonnegative integer. The following statements are true:*

(i) *If p is a prime number and $cd_p(G) \leqq r$, then $id(G, A) \leqq r + 1$ for all discrete G-modules A which are p-primary abelian groups.*

(ii) *If* $cd(G) \leqq r$, *then* $id(G, A) \leqq r + 1$ *for all discrete G-modules A which are torsion abelian groups.*

(iii) *If* $scd(G) \leqq r$, *then* $id(G, A) \leqq r + 1$ *for all discrete G-modules A.*

(iv) *If* $cd(G) \leqq r$, *then* $id(G, A) \leqq r + 2$ *for all discrete G-modules A.*

Proof. Applying [3, Chap. I, Proposition 14, p. I–20] and [3, Chap. I, Proposition 11, p. I–17], the following three equivalences are clear:

(i) $cd_p(G) \leqq r$ if, and only if, $cd(G, A) \leqq r$ for all p-primary, discrete G-modules A.

(ii) $cd(G) \leqq r$ if, and only if, $cd(G, A) \leqq r$ for all torsion, discrete G-modules A.

(iii) $scd(G) \leqq r$ if, and only if, $cd(G, A) \leqq r$ for all discrete G-modules A.

Finally, (6, iv) is clear by [3, Chap. I, Proposition 13, p. 1–19].

References

1. B. Mitchell, *Theory of Categories*, Pure and applied mathematics 17, Academic Press, New York, 1965.

2. J-P. Serre, *Corps Locaux*, Actualités scientifiques et industrielles 1296, Hermann, Paris, 1962.

3. ———, *Cohomologie Galoisiene*, Lecture notes in mathematics 5, Springer-Verlag, Berlin, 1965.

Received August 14, 1972 and in revised form December 18, 1972.

UNIVERSIDAD DE BUENOS AIRES

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

SEMIPERFECT RINGS WITH ABELIAN GROUP OF UNITS

W. K. NICHOLSON

In 1963, Gilmer characterized all finite commutative rings with a cyclic group of units and, in 1967, Eldridge and Fischer generalized these results to rings with minimum condition. In the present paper these results are extended to semiperfect rings and generalizations of the three theorems are obtained. It is shown that a semiperfect ring with cyclic group of units is finite and is either commutative or is the direct sum of a commutative ring and the 2×2 upper triangular matrix ring over the field of two elements. Let R be semiperfect with an abelian group of units. It is shown that R is finite if either the group of units is finite or the group of units is finitely generated and the Jacobson radical is nil.

The proofs of all these results depend on our main theorem: The structure of a semiperfect ring R with an abelian group of units is described completely up to the structure of commutative local rings. (That is commutative rings with a unique maximal ideal.) The groups of units of these local rings are shown to be direct factors of the group of units of R.

1. Preliminaries. Throughout this paper we assume that all rings are associative and have an identity and that all modules are unital. If R is a ring we denote its group of units by R^* and its Jacobson radical by $J(R)$. The ring of residues of the integers modulo n will be denoted by $\mathscr{Z}_n$. The following notions will be referred to several times below.

DEFINITION 1. Let $R_1, R_2, \cdots, R_n$ be rings and, if $i \neq j$, let X_{ij} be an $R_i - R_j$ bimodule. We define the *semidirect sum* $[R_i, X_{ij}]$ to be the ring of all $n \times n$ "matrices" (x_{ij}) where $x_{ii} \in R_i$ for each i and $x_{ij} \in X_{ij}$ for all $i \neq j$. These are added componentwise and we define the product $(x_{ij})(y_{ij}) = (z_{ij})$ as follows:

$$z_{ii} = x_{ii}y_{ii} \qquad \text{for all} \quad i = 1, 2, \cdots, n ,$$
$$z_{ij} = x_{ii}y_{ij} + x_{ij}y_{jj} \quad \text{for all} \quad i \neq j .$$

It is easy to verify that $[R_i, X_{ij}]$ is an associative ring. If the bimodules X_{ij} are all zero the semidirect sum $[R_i, X_{ij}]$ reduces to the usual direct sum $R_1 \oplus R_2 \oplus \cdots \oplus R_n$. More generally, we have that $[0, X_{ij}]$ is an ideal of $[R_i, X_{ij}]$ which squares to zero, and the quotient ring is isomorphic to $R_1 \oplus R_2 \oplus \cdots \oplus R_n$. Clearly the direct sum of two of these semidirect sums is again a semidirect sum.

DEFINITION 2. Let R be a ring. A left R-module X is said to be *G-unital* if $ux = x$ for all $u \in R^*$ and all $x \in X$. A bimodule is called G-unital if it is G-unital as a left and as a right module.

Clearly 0 is a G-unital module and submodules and quotient modules of such modules are again of the same type. If X is a G-unital R-module we have $x + x = 0$ for every $x \in X$ since $-1 \in R^*$. In other words X is an elementary abelian 2-group.

PROPOSITION 1. *Let $R_1, R_2, \cdots, R_n$ be rings and let X_{ij} be an $R_i - R_j$ bimodule for all $i \neq j$. The semidirect sum $[R_i, X_{ij}]$ has the following properties*:

(1) $[R_i, X_{ij}]^* = \{(x_{ij}) \mid x_{ii} \in R_i^*$ *for each* $i\}$.

(2) $[R_i, X_{ij}]^*$ *is abelian if and only if each* R_i^* *is abelian and each* X_{ij} *is G-unital.*

(3) *If* $[R_i, X_{ij}]^*$ *is abelian it is isomorphic to the direct product of all the multiplicative groups* R_i^* *and all the additive groups* X_{ij}.

Proof. (1) If (x_{ij}) is given and $x_{ii} \in R_i^*$ for each i it is easy to verify that $(x_{ij})^{-1} = (y_{ij})$ where $y_{ii} = x_{ii}^{-1}$ for each i and $y_{ij} = -x_{ii}^{-1}x_{ij}x_{jj}^{-1}$ for all $i \neq j$. The converse is clear.

(2) Suppose $[R_i, X_{ij}]^*$ is abelian. Then (1) and the definition of multiplication in $[R_i, X_{ij}]$ imply that each R_i^* is abelian. Now choose arbitrary elements $x_{ii}, y_{ii} \in R_i^*$ for each i and $x_{ij}, y_{ij} \in X_{ij}$ for all $i \neq j$. The units (x_{ij}) and (y_{ij}) commute so, for all $i \neq j$:

$$x_{ii}y_{ij} + x_{ij}y_{jj} = y_{ii}x_{ij} + y_{ij}x_{jj} .$$

If $x_{ij} = 0$ and $x_{jj} = 1$ this shows $x_{ii}y_{ij} = y_{ij}$. Similarly $x_{ij}y_{jj} = x_{ij}$ and it follows that each X_{ij} is G-unital.

Conversely: If $(x_{ij}), (y_{ij}) \in [R_i, X_{ij}]^*$ then, using (1), $x_{ii}y_{ii} = y_{ii}x_{ii}$. Furthermore, since the X_{ij} are G-unital, we have

$$x_{ii}y_{ij} + x_{ij}y_{jj} = y_{ij} + x_{ij} = x_{ij} + y_{ij} = y_{ii}x_{ij} + y_{ij}x_{jj}$$

for all $i \neq j$ and it follows that $[R_i, X_{ij}]^*$ is abelian.

(3) This follows easily from the definition of multiplication in $[R_i, X_{ij}]$ and the fact that each X_{ij} is G-unital.

PROPOSITION 2. *Let R be a local ring.* ($R/J(R)$ *a divisor ring.*)

(1) *If the group of units of R is abelian then R is commutative.*

(2) *If R possesses a nonzero G-unital module then* $R/J(R) \cong \mathscr{Z}_2$.

(3) *If* $R/J(R) \cong \mathscr{Z}_2$ *the G-unital R-modules are precisely the (additive) elementary abelian 2-groups.*

Proof. (1) If $a, b \in J(R)$ then $1 + a, 1 + b \in R^*$ so $(1 + a)(1 + b) =$

$(1 + b)(1 + a)$. This implies $ab = ba$. If $a \in J(R)$ and $u \in R^*$ then $(1 + a)u = u(1 + a)$ so $au = ua$.

(2) Let $X \neq 0$ be a G-unital left R-module. If $x \in X$ and $a \in J(R)$ then $(1 + a)x = x$ so $ax = 0$. Hence X is a G-unital $R/J(R)$ module so we can assume R is a division ring. But then if $0 \neq r \in R$ we have that $(1 - r)x = 0$ for every $x \in X$. Since $X \neq 0$ this means $1 - r$ is not a unit so $r = 1$. Hence $R \cong \mathscr{Z}_2$.

(3) If X is an elementary abelian 2-group then it is a vector space over $\mathscr{Z}_2$. Since $R/J(R) \cong \mathscr{Z}_2$, R acts on X as follows: If $x \in X$ and $r \in R$ we have $rx = x$ if $r \in R^*$ and $rx = 0$ if $r \in J(R)$. Clearly, X is G-unital. Conversely: Every G-unital module is an elementary 2-group since $-1 \in R^*$ and the action is as described.

In the next section we shall use these results to characterize the semiperfect rings R where R^* is abelian.

2. The main theorem. Throughout this section R will denote a semiperfect ring with R^* abelian. It is well known [2, Th. 20, p. 159] that R is semiperfect if and only if we can write $1 = e_1 + e_2 + \cdots + e_n$ where the e_i are orthogonal local idempotents. Hence each of the rings e_iRe_i is local and, if $i \neq j$, the ring e_iRe_j is an $e_iRe_i - e_jRe_j$ bimodule. For the moment let $[e_iRe_j]$ denote the set of all $n \times n$ "matrices" (x_{ij}) with the (i, j) entry x_{ij} drawn from e_iRe_j. This is a ring if ordinary matrix operations are used.

Define a map $\phi: R \rightarrow [e_iRe_j]$ by $\phi(r) = (e_ire_j)$ for each $r \in R$. Then ϕ is clearly a homomorphism of additive groups and it is a ring homomorphism since the (i, j) entry of $\phi(r)\phi(s)$ is

$$\sum_k (e_ire_k)(e_kse_j) = e_ir(e_1 + e_2 + \cdots + e_n)se_j = e_irse_j .$$

Moreover ϕ is one-to-one. Indeed, if $\phi(r) = 0$ then $e_ire_j = 0$ for all i, j and so $r = \sum_{i,j} e_ire_j = 0$. Finally ϕ is onto. For if $(e_ir_{ij}e_j) \in [e_iRe_j]$ is given let $r = \sum_{i,j} e_ir_{ij}e_j$. It is easy to check that $\phi(r) = (e_ir_{ij}e_j)$. Hence ϕ is a ring isomorphism and so we have represented R as a generalized matrix ring. Our aim is to show that it is a semidirect sum.

LEMMA 1. *If* $e_i \neq e_j$ *and* $e_j \neq e_k$ *then* $e_iRe_jRe_k = 0$.

Proof. Let $x \in e_iRe_j$ and $y \in e_jRe_k$. Then $x^2 = 0$ (since $e_ie_j = 0$) so $1 + x$ is a unit. Similarly $1 + y$ is a unit and so, since R^* is abelian, $xy = yx$. But $x = e_ix$ and $y = e_jy$ so that $xy = e_ixy = e_iyx = e_ie_jyx = 0$.

It now follows easily that the multiplication in $[e_iRe_j]$ is that of the semidirect sum. Indeed if $(x_{ij}), (y_{ij}) \in [e_iRe_j]$ and we write $(x_{ij})(y_{ij}) = (z_{ij})$, then using the lemma:

$$z_{ii} = \sum_k x_{ik}y_{ki} = x_{ii}y_{ii} \quad \text{for all} \quad i = 1, 2, \cdots, n ,$$

$$z_{ij} = \sum_k x_{ik}y_{kj} = x_{ii}y_{ji} + x_{ij}y_{jj} \quad \text{for all} \quad i \neq j .$$

Hence, in the notation of § 1, $R \cong [e_iRe_i, e_iRe_j]$. But then Proposition 1 shows that each $(e_iRe_i)^*$ is abelian and each e_iRe_j is a G-unital bimodule. Since each e_iRe_i is local, it is commutative by Proposition 2. Furthermore, e_i is central if and only if $e_iRe_j = 0 = e_jRe_i$ for all $j \neq i$. It follows that either e_i is central or e_iRe_i possesses a nonzero G-unital module. In the latter case $e_iRe_i/J(e_iRe_i) \cong \mathscr{Z}_2$ by Proposition 2. This proves the "only if" part of the following theorem; the rest follows from Propositions 1 and 2.

THEOREM 1. *Let R be a semiperfect ring. The group of units of R is abelian if and only if $R \cong T \oplus S$ where T is zero or a direct sum of commutative local rings R_i and S is zero or $S \cong [L_i, X_{ij}]$. Here each L_i is a commutative local ring with $L_i/J(L_i) \cong \mathscr{Z}_2$ and each X_{ij} is a G-unital $L_i - L_j$ bimodule. Moreover*:

(1) *The bimodules X_{ij} can be chosen to be arbitrary elementary abelian 2-group where the action of L_i is defined as follows: If $r \in L_i$ and $x \in X_{ij}$ set $rx = x$ if $r \in L_i^*$ and $rx = 0$ if $r \in J(L_i)$.*

(2) *The group of units of R is isomorphic to the direct product of all the groups R_i^*, all the groups L_i^* and all the (additive) groups X_{ij}.*

This characterizes R completely up to the structure of the commutative local rings involved. The groups of units of these local rings inherit many properties from R^* by (2) and often this leads to a complete characterization. This will be exemplified in § 3 below in the case where R^* is assumed to be cyclic. Also, each of these local rings is a homomorphic image of R so they inherit many ring-theoretic conditions which could be imposed on R, for example the descending chain condition.

An immediate consequence of Theorem 1 is that if R is semiperfect and R^* is abelian then $R/J(R)$ is a finite direct sum of fields. Of course this result follows from structure theory.

The next result is a generalization (in the case where R has an identity) of a theorem of Eldridge and Fischer ([3], Th. 1, p. 244).

COROLLARY 1. *Let R be a semiperfect ring with abelian group of units. Then R is commutative if either of the following conditions is satisfied*:

(1) $2x = 0$ *in* R *implies* $x = 0$
(2) R^* *has no direct factor each element of which has order* 2.

Proof. Let $[L_i, X_{ij}]$ be the semidirect sum appearing in the decomposition of R. If $x \in X_{ij}$ then $2x = 0$ since X_{ij} is G-unital so each $X_{ij} = 0$ if condition (1) holds. If condition (2) holds each $X_{ij} = 0$ by (2) of the theorem. The result follows.

COROLLARY 2. *Let R be a semiperfect ring. If the group of units of R is abelian and finite then R is finite.*

Proof. By (2) of Theorem 1 each of the local rings appearing in the decomposition of R has a finite group of units and each of the bimodules appearing in the semidirect sum is finite. But if L is a local ring and L^* is finite then $J(L)$ is finite since $1 + J(L) \subseteqq L^*$ and $L/J(L)$ is finite since $[L/J(L)]^* \cong L^*/(1 + J(L))$. This implies L is finite and the result follows.

A natural question is whether a semiperfect ring R must be finite when R^* is assumed to be abelian and finitely generated. The answer is yes if R^* is cyclic (see Theorem 2 below) or if $J(R)$ is nil. The next result will be useful in both cases.

LEMMA 2. *Let R be a commutative local ring. If R^* is finitely generated then R is noetherian, $R/J(R)$ is a finite field and $J(R)^n/J(R)^{n+1}$ is a finite ring for each n.*

Proof. If $A_1 \subseteqq A_2 \subseteqq \cdots$ is a chain of (proper) ideals of R we have the chain $1 + A_1 \subseteqq 1 + A_2 \subseteqq \cdots$ of subgroups of R^*. It follows that R is noetherian. We have $[R/J(R)]^* \cong R^*/1 + J(R)$ so the field $R/J(R)$ has a finitely generated group of units. Hence it is finite. Finally, $J(R)^n/J(R)^{n+1}$ is a vector space over $R/J(R)$ and is finite dimensional since R is noetherian.

We can now prove the following result which generalizes another result of Eldridge and Fischer ([3], Th. 2, p. 245).

PROPOSITION 3. *Let R be a semiperfect ring with R^* abelian. If R^* is finitely generated and $J(R)$ is nil then R is finite.*

Proof. Decompose R as in Theorem 1 and let L be one of the commutative local rings which appear. Then L^* is finitely generated by (2) of Theorem 1 and $J(L)$ is nil ($L = eRe$ for some $e^2 = e \in R$). Since L is noetherian by Lemma 2, write $J(L) = La_1 + La_2 + \cdots + La_n$ where the $a_i \in J(L)$. Hence, if $m \geqq 1$ we have $J(L)^m = \sum Lx_1^{k_1}a_2^{k_2} \cdots a_n^{k_n}$

where the sum is taken over all $k_i \geqq 0$ satisfying $k_1 + k_2 + \cdots + k_n = m$. Since $J(L)$ is nil this implies that $J(L)$ is nilpotent. But then Lemma 2 implies that $J(L)$, and hence L, is finite.

It remains to show that if $[L_i, X_{ij}]$ is the semidirect sum appearing in the decomposition of R then each X_{ij} is finite. But each X_{ij} is finitely generated as an additive group and so, since it is a vector space over $\mathscr{Z}_2$, it is finite. This completes the proof.

We remark that the hypothesis that $J(R)$ is nil was used only to show that $J(L)$ is nilpotent.

3. Cyclic groups of units. Gilmer [4] has characterized all finite commutative rings with a cyclic group of units and Eldridge and Fischer [3] have extended these results to artinian rings. In order to cover the semiperfect case we need the following negative result.

PROPOSITION 4. *If R is a commutative local ring the group of units of R is not infinite cyclic.*

Proof. Assume, on the contrary, that R^* is infinite cyclic. Then the characteristic of R is two since $(-1)^2 = 1$. By Lemma 2 R is noetherian and $R/J(R)$ is finite. Hence, $J(R)^2 \neq J(R)$. But the additive group $J(R)/J(R)^2$ is cyclic since it is naturally isomorphic to the multiplicative group $1 + J(R)/1 + J(R)^2$. Since the characteristic is two, it follows that $J(R)/J(R)^2$ has two elements. But $J(R)/J(R)^2$ is a vector space over the field $R/J(R)$ so $R/J(R) \cong \mathscr{Z}_2$.

We now claim that R is an integral domain. If not let P be any prime ideal of R. Then $P \neq 0$ so $(R/P)^* \cong R^*/1 + P$ is finite cyclic. Since R/P is local it follows that R/P is a finite integral domain and hence that P is maximal. Hence every prime ideal is maximal and so ([6], p. 203) R is artinian. But then R is finite by Proposition 3, a contradiction. Hence R is an integral domain.

Now let u be a generator of $R^* = 1 + J(R)$. Write $u = 1 + a$ and $u^{-1} = 1 + b$ where $a, b \in J(R)$. Then $1 + a + a^2$ is a unit so $1 + a + a^2 = (1 + a)^k$ for some $k \in \mathscr{Z}$. It is easy to check that $k = 0, 1, 2$ are impossible. Suppose $k \geqq 3$. Then we have

$$1 + a + a^2 = 1 + \gamma a + \delta a^2 + a^t u$$

where $3 \leqq t \leqq k$, u is a unit and γ and δ are each either 0 or 1. Since R is a domain it is easy to check that each of these possibilities for γ and δ lead to a contradiction. Hence we must have $1+a+a^2 = u^{-l}$ for some $l \geqq 1$. But $1 + a + a^2 = 1 + u + u^2$ and $1 + b + b^2 = 1 + u^{-1} + u^{-2}$ as is easily verified. Hence

$$1 + b + b^2 = u^{-2}(1 + a + a^2) = u^{-2}u^{-l} = (1 + b)^{l+2} .$$

This leads to a contradiction just as before and so completes the proof.

We can now obtain a generalization of another result of Eldridge and Fischer ([3], Th. 3, p. 248) and, in so doing, obtain a much easier proof of that result.

THEOREM 2. *Let R be a semiperfect ring with cyclic group of units. Then R is finite and is either commutative or is isomorphic to the direct sum of a commutative ring and the ring of 2×2 upper triangular matrices over $\mathscr{Z}_2$.*

Proof. Decompose R as in Theorem 1. By (2) of Theorem 1, each of the local rings appearing has a cyclic group of units and, by Proposition 4, it is finite. It follows that each of these local rings is finite. Now let $[L_i, X_{ij}]$ be the semidirect sum appearing. Each X_{ij} is cyclic as an additive group and so, since it is an elementary abelian 2-group, it has two elements. Hence R is finite. Furthermore, the fact that all X_{ij} are direct factors of R^* means that at most one is nonzero. If all are zero then $[L_i, X_{ij}]$ is commutative so R is commutative. If, without loss of generality, $X_{12} \neq 0$, we have

$$[L_i, X_{ij}] = \begin{pmatrix} L_1 & X_{12} \\ 0 & L_2 \end{pmatrix} \oplus L_3 \oplus \cdots \oplus L_n$$

where $\begin{pmatrix} L_1 & X_{12} \\ 0 & L_2 \end{pmatrix}$ is a semidirect sum.

Moreover, each L_i^* has odd order so L_i has characteristic two. But then, if $a \in J(L_i)$, there exists an odd integer n such that $1 = (1 + a)^n = 1 + a + a^2 r$ where $r \in L_i$. Hence $a(1 + ar) = 0$ so $a = 0$. This means $J(L_i) = 0$ so $L_i \cong \mathscr{Z}_2$. In particular, $\begin{pmatrix} L_1 & X_{12} \\ 0 & L_2 \end{pmatrix}$ is isomorphic to the ring of 2×2 upper triangular matrices over $\mathscr{Z}_2$. This completes the proof.

This theorem completely characterizes the semiperfect rings with a cyclic group of units since the finite commutative local rings of this type have been characterized by Gilmer [4] and later by Ayoub [1] and Pearson and Schneider [5]. Gilmer cited the ring of 2×2 upper triangular matrices over $\mathscr{Z}_2$ as an example of a finite noncommutative ring with cyclic group of units. Theorem 2 shows that this is essentially the only such semiperfect ring.

REFERENCES

1. C. W. Ayoub, *On finite primary rings and their groups of units*, Compositio Math., **21** (1969), 247-252.
2. E. A. Behrens, *Ring Theory*, Academic Press, New York, 1972.
3. K. E. Eldridge and I. Fischer, *D. C. C. rings with a cyclic group of units*, Duke Math. J., **34** (1967), 243-248.
4. R. W. Gilmer, *Finite rings with a cyclic multiplicative group of units*, Amer. J. Math., **85** (1963), 447-452.
5. K. R. Pearson and J. E. Schneider, *Rings with a cyclic group of units*, J. of Algebra, **16** (1970), 243-251.
6. D. Zariski and P. Samuel, *Commutative Algebra*, Vol. I, Van Nostrand, Princeton, N. J., 1958.

Received August 4, 1972.

THE UNIVERSITY OF CALGARY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

THREE THEOREMS ON IMBEDDED PRIME DIVISORS OF PRINCIPAL IDEALS

L. J. RATLIFF, JR.

Let B be a finitely generated integral domain over a Noetherian domain A. The first theorem shows that there are only finitely many imbedded prime divisors of principal ideals in B if and only if this holds in A. The second theorem gives a necessary and sufficient condition in order that only finitely many height one prime ideals in A ramify in B, when A is locally factorial. The third theorem characterizes local domains which contain infinitely many imbedded prime divisors of principal ideals.

1. Introduction. For convenience of description, let $\mathscr{I}(C)$ denote the set of imbedded prime divisors of principal ideals generated by regular elements in a ring C. Then a desirable property of a Noetherian domain A is for $\mathscr{I}(A)$ to be a finite set, for this implies that there exist nonzero elements $a \in A$ such that no principal ideal in A_a has an imbedded prime divisor, hence $A_a = A_a^{(1)}$, and so A_a has certain other nice properties. Theorem 2.4 shows that this desirable property is inherited by finitely generated extension domains of A, and its Corollary (2.8) extends this result to the case A contains nonzero zero divisors. Corollary 2.9 shows that for a large class of Noetherian rings A, $\mathscr{I}(B)$ is a finite set (where B is a certain type of finitely generated extension ring of A).

In § 3, (3.1) gives two characterizations of a local domain R such that $\mathscr{I}(R)$ is an infinite set. That $\mathscr{I}(R)$ can be an infinite set is closely related to some open problems on unmixed and quasi-unmixed local domains, and characterizations of such R are therefore important. Corollary 3.4 extends these characterizations to certain local rings, and then an example is given of a local domain L such that: $L^{(1)}$ is not a finite L-algebra; $\mathscr{I}(L)$ is a finite set; and, $\mathscr{I}(L^*)$ is an infinite set, where L^* is the completion of L. This example also gives some information concerning the open problem of whether an integrally closed local domain must be unmixed.

Section 4 is concerned with unramification of height one prime ideals. Such unramification is of importance in a numbe of problems in local ring theory, for example in the purity of the branch locus [5, (41.1)]. Theorem 4.4 gives two necessary and sufficient conditions in order that only finitely many height one prime ideals in a locally factorial Noetherian domain A ramify in a finite separably generated extension domain B of A. The paper is closed with some corollaries to (4.4), among these are that the conditions are satisfied if either A is a Dedekind domain (4.6) or B is a flat A-algebra (4.8).

2. Finite extension rings. All rings in this paper are assumed to be commutative rings with a unit, and the undefined terminology in the paper is, in general, the same as that in [5].

We begin with the following definition.

DEFINITION 2.1. For a ring A, let $\mathscr{I}(A) = \{p \in \operatorname{Spec} A;\ p$ is an imbedded prime divisor of a principal ideal generated by a regular element in $A\}$.

That $\mathscr{I}(A)$ must be infinite in many cases when A contains nonzero zero divisors follows from the following remark. It is because of this remark that we shall usually only consider rings with no imbedded prime divisors of zero in this paper.

REMARK 2.2. (Cf. [6, Lemma 6].) Let q be a prime divisor of zero in a Noetherian ring A, and let b be a regular element in A such that $(q, b)A \neq A$. Then each minimal prime divisor of $(q, b)A$ is a prime divisor of bA.

Proof. This follows immediately from [6, Lemma 6] and its proof.

The following lemma sets the stage for an easy proof of the first theorem in this paper, and provides information needed throughout the paper. Although the lemma could be adapted to the case A contains nonzero zero divisors, it is stated only for the integral domain case, since this case is sufficient for our purposes in this paper. The following known fact is needed for the proof of (2.3): If B is a Noetherian ring which is a flat extension of a Noetherian ring A and I is an ideal in A, then a prime ideal P in B is a prime divisor (resp., minimal prime divisor) of IB if and only if P is a prime divisor (resp., minimal prime divisor) of pB, for some prime divisor (resp., minimal prime divisor) p of I, in which case $P \cap A = p$ [5, (18.11)]. Also, implicitly used in the proof (and frequently throughout this paper) is the fact that if $p \in \mathscr{I}(A)$ and a is a regular element in p, then p is an imbedded prime divisor of aA [5, (12.6)] (so the intersection of infinitely many such p consists of zero divisors).

LEMMA 2.3. *Let A be a Noetherian domain.*

(1) *For each nonzero b in A, $\mathscr{I}(A_b) = \{pA_b;\ b \notin p \in \mathscr{I}(A)\}$, and $\mathscr{I}(A) = \{P \cap A;\ P \in \mathscr{I}(A_b)\} \cup \{p;\ p$ is an imbedded prime divisor of $bA\}$.*

(2) *For each* $n \geqq 1$, $\mathscr{I}(A_n) = \{pA_n; p \in \mathscr{I}(A)\}$, *where* $A_n = A[X_1, \cdots, X_n]$ *with the* X_i *indeterminates.*

(3) *If* A' *is a finite free integral extension ring of* A, *then* $\mathscr{I}(A') = \{p' \in \operatorname{Spec} A'; p' \cap A \in \mathscr{I}(A)\}$.

(4) *If* A^* *is a flat* A*-algebra, then* $\mathscr{I}(A^*) \supseteqq \{p^* \in \operatorname{Spec} A^*; p^*$ *is a prime divisor of* pA^*, *for some* $p \in \mathscr{I}(A)\}$.

(5) $\mathscr{I}(A)$ *is a finite set if and only if* $A^{(1)} = \bigcap\{A_p; p \in \operatorname{Spec} A$ *and height* $p = 1\} \subseteqq A_b$, *for some nonzero* b *in* A.

(6) *If* B *is a finitely generated algebraic extension domain of* A, *then* $\mathscr{I}(B)$ *is a finite set if and only if* $\mathscr{I}(A)$ *is a finite set.*

Proof. (1) and (2) are straightforward, (3) and (4) follow from [5, (18.11)], and (5) is given in [7, Lemma 5.15(8)] (see (3.3) below). It clearly suffices to prove (6) in the case $B = A[b]$, for some $b \in B$. Then there exists a nonzero $a \in A$ such that $A_a[b] = B_a$ is a free integral extension domain of A_a. Therefore, it follows from (1) and (3) that $\mathscr{I}(B_a)$ is a finite set if and only if $\mathscr{I}(A)$ is, hence, by (1), $\mathscr{I}(B)$ is a finite set if and only if $\mathscr{I}(A)$ is a finite set.

THEOREM 2.4. *Let* B *be a finitely generated integral domain over a Noetherian domain* A. *Then* $\mathscr{I}(B)$ *is a finite set if and only if* $\mathscr{I}(A)$ *is a finite set.*

Proof. There are elements $X_1, \cdots, X_n$ in B which are algebraically independent over A such that B is a (finite) algebraic extension domain of $A_n = A[X_1, \cdots, X_n]$. Therefore, the conclusion follows from (2.3)(2) and (6).

To generalize (2.4) to the case where A contains nonzero divisors of zero, the following two lemmas are needed.

LEMMA 2.5. *Let* q *be a minimal prime ideal in a Noetherian ring* A. *Then there exists* $a \in A$, $\notin q$ *such that, for all* $P \in \operatorname{Spec} A$ *such that* $q \subseteqq P$ *and* $a \notin P$, $P \in \mathscr{I}(A)$ *if and only if* $P/q \in \mathscr{I}(A/q)$.

Proof. If $b \in A$, $\notin q$, then it clearly suffices to prove the lemma for A_b instead of A. Hence it may be assumed that q is nilpotent. The lemma now readily follows from [3, IV. (6.10.6)]. Specifically, the referenced result is essentially local, and passing from the language of preschemes to the language of commutative rings we find that it asserts that there exists $a \in A$, $\notin q$ such that, for all $P \in \operatorname{Spec} A$ such that $q \subseteqq P$ and $a \notin P$, altitude $A_P =$ altitude $(A/q)_{P/q}$ and Prof $A_P =$ Prof $(A/q)_{P/q}$ (since q is nilpotent), where Prof R is the length of a maximal R-sequence with R a local ring. The lemma follows from this, since $P \in \mathscr{I}(A)$ if and only if Prof $A_P = 1$ and height $P > 1$.

LEMMA 2.6. *Let A be a Noetherian ring such that each prime divisor of zero in A is minimal. Then $\mathscr{I}(A)$ is a finite set if and only if, for each minimal prime ideal q in A, $\mathscr{I}(A/q)$ is a finite set.*

Proof. Let $q_1, \cdots, q_g$ be the minimal prime ideals in A, and, for $i = 1, \cdots, g$, let a_i be as in (2.5) for q_i. Therefore, for all $P \in \operatorname{Spec} A$ such that $q_i \subsetneqq P$ and $a_i \notin P$, $P \in \mathscr{I}(A)$ if and only if $P/q_i \in \mathscr{I}(A/q_i)$. Now, if $\mathscr{I}(A)$ is an infinite set, then there is an i such that $q_i = \bigcap \{P; q_i \subset P \in \mathscr{I}(A)\}$. Fix such an i, and let $q = q_i$ and $a = a_i$. Then $qA_a = \bigcap \{PA_a; qA_a \subset PA_a \in \mathscr{I}(A_a)\}$. Hence (2.5) implies $\mathscr{I}(A/q)$ is an infinite set, since $\mathscr{I}((A/q)_{a+q})$ is. Conversely, if $\mathscr{I}(A/q_i)$ is an infinite set, for some $i = 1, \cdots, g$, then fix such an i and let $q = q_i$ and $a = a_i$. Then $(0) = \bigcap \{P/q, P/q \in \mathscr{I}(A/q)\}$, so $(0)(A/q)_{a+q} = \bigcap \{(P/q)_{a+q}; (P/q)_{a+q} \in \mathscr{I}((A/q)_{a+q})\}$, hence (2.5) implies $\mathscr{I}(A)$ is an infinite set.

COROLLARY 2.7. *With A as in* (2.6), *$\mathscr{I}(A)$ is a finite set if and only if $\mathscr{I}(A/(\operatorname{Rad} A))$ is a finite set.*

Proof. Clear by (2.6).

COROLLARY 2.8. *Let B be a finitely generated ring over a Noetherian ring A, and assume that all prime divisors of zero in A and in B are minimal and that, for each prime divisor q of zero in B, $q \cap A$ is a prime divisor of zero. Then $\mathscr{I}(A)$ is a finite set if and only if $\mathscr{I}(B)$ is a finite set.*

Proof. Let $q_1, \cdots, q_g$ be the prime divisors of zero in B. Then, by (2.6) and (2.4), $\mathscr{I}(B)$ is a finite set if and only if $\mathscr{I}(B/q_i)$ is, for all i, if and only if $\mathscr{I}(A/(q_i \cap A))$ is, for all i, if and only if $\mathscr{I}(A)$ is a finite set.

The corollary shows that such finite extension rings of a large class of Noetherian rings have only finitely many imbedded prime divisors of principal ideals. Specifically, the following corollary holds.

COROLLARY 2.9. *Let A and B be as in* (2.8). *Then $\mathscr{I}(B)$ is finite in each of the following cases:*

(1) *A is locally factorial* (4.1).
(2) *A is integrally closed.*
(3) *A is pseudogeometric* [5, p. 131].
(4) *A is Japanese* [2, 0. (23.1.1)].
(5) *The integral closure of A is a finite A-algebra.*
(6) *A is locally Macaulay.*

(7) *A is excellent* [3, IV, (7.8.2)].
(8) *Altitude* $A \leqq 1$.
(9) *A is semi-local and altitude* $A \leqq 2$.
(10) *A is an analytically unramified semi-local ring.*
(11) *A is an unmixed semi-local domain.*
(12) *A is a local ring whose completion has no imbedded prime divisor of zero.*

Proof. By (2.8) it may be assumed, where appropriate, that $\operatorname{Rad} A = (0)$. Then it is well known that in each case $\mathscr{I}(A)$ is a finite set (for (11), see (3.5)(2) below and (2.3)(5), and for (12), see (3.4) below), hence the conclusion follows from (2.8).

One further result which is related to (2.4) and which gives a sharper conclusion when $\mathscr{I}(A)$ is finite is given in the following proposition.

PROPOSITION 2.10. *Let A and B be as in* (2.4), *and assume* $\mathscr{I}(A)$ *is a finite set. Then there exists a nonzero a in A such that* $B_a^{(1)}$ *is a finite* B_a*-algebra.*

Proof. It is known [5, (14.4)] that there exist $X_1, \cdots, X_n$ in B which are algebraically independent over A and a nonzero c in A such that B_c is a (finite) integral extension domain of $A[1/c, X_1, \cdots, X_n]$. Also, since $\mathscr{I}(A)$ is a finite set, there is a nonzero d in A such that $A^{(1)} \subseteqq A_d$ (2.3)(5). Then $A_d^{(1)} = A_d$ [7, Corollary 5.9(2)], so, with $a = cd$, $A_a^{(1)} = A_a$ [7, Corollary 5.9(2)], hence $D^{(1)} = D$ [7, Lemma 5.11(2)], where $D = A[1/a, X_1, \cdots, X_n]$. Therefore, since B_a is a finite integral extension demain of D, $B_a^{(1)}$ is a finite B_a-algebra [3, IV. (5.10.17)].

3. The local domain case. With A and B as in (2.8), if A is semi-local of altitude at most two, then $\mathscr{I}(B)$ is a finite set (2.9)(9). This leads to the question: Can $\mathscr{I}(R)$ be an infinite set when R is a local ring? The answer is well known to be "yes", if there is an imbedded prime divisor q of zero in R such that depth $q > 1$ (2.2). Quite recently the answer was shown to be "yes" even when R is a local domain. Specifically, in [1, Proposition 3.5] an example was constructed of a local domain R such that altitude $R = 3$ and R contains infinitely many height two prime ideals P such that R_P is not Macaulay (hence PR_P, and so P also, is an imbedded prime divisor of each nonzero principal ideal that is contained in it).

The following theorem gives two characterizations of a local domain R such that $\mathscr{I}(R)$ is an infinite set.

THEOREM 3.1. *The following statements are equivalent for a* *local domain* (R, M):

(1) $\mathscr{S}(R)$ *is an infinite set.*

(2) *For no nonzero* $b \in R$ *is* $R^{(1)} \subseteqq R_b$.

(3) *There exists an imbedded prime divisor* q *of zero in the* M-*adic completion* R^* *of* R *such that* $q = \bigcap \{p^* \in \mathscr{S}(R^*); q \subset p^*$ *and* $p^* \cap R \in \mathscr{S}(R)\}$.

Proof. The equivalence of (1) and (2) has already been noted in (2.3)(5). For the equivalence of (1) and (3), assume first that (1) holds. For each $p \in \mathscr{S}(R)$, let p^* be a minimal prime divisor of pR^*, and let $\mathscr{S}^* = \{p^*; p \in \mathscr{S}(R)\}$ (so $\mathscr{S}^* \subseteqq \mathscr{S}(R^*)$ (2.3)(4)). Then, for each $p^* \in \mathscr{S}^*$, $R^*_{p^*}$ is not Macaulay, so $I \subseteqq p^*$, where I is the radical ideal in R^* which defines the non-Macaulay locus of R^* (the Macaulay locus of R^* (that is, the set $\{p^* \in \operatorname{Spec} R^*; R^*_{p^*}$ is Macaulay$\}$) is open (in the Zariski topology on $\operatorname{Spec} R^*$) [3, IV. (5.11.8)]). Now infinitely many p^* must contain some (minimal) prime divisor of I, say q, and then $q \subseteqq \bigcap \{p^*; q \subseteqq p^* \in \mathscr{S}^*\}$. Since R^* is Noetherian, it follows that $q = \bigcap \{p^*; q \subset p^* \in \mathscr{S}^*\}$ and q is a prime divisor of zero. Since R^*_q is not Macaulay, height $q > 0$, so (3) holds.

Conversely, if (3) holds, then let $\mathscr{S} = \{p^* \cap R; q \subset p^* \in \mathscr{S}(R^*)$ and $p^* \cap R \in \mathscr{S}(R)\}$, so $\mathscr{S} \subseteqq \mathscr{S}(R)$ and (1) holds if $\mathscr{S}$ is an infinite set. Since q is a prime divisor of zero, there are infinitely many p^*, so (1) holds if, for each $p^* \cap R \in \mathscr{S}$, p^* is a prime divisor of $(p^* \cap R)R^*$. But this is true by [5, (18.11)], since $p^* \in \mathscr{S}(R^*)$ and $p^* \cap R \in \mathscr{S}(R)$.

REMARKS 3.2. (1) Possibly the theorem remains true if the set in (3) is replaced by $\{p^* \in \operatorname{Spec} R^*; q \subset p^*$ and $p^* \cap R \in \mathscr{S}(R)\}$. At least the author knows of no case where this last set does not work. Of course, this last set always works in the case that infinitely many of the p^* satisfy: height $p^* =$ height $q + 1$ (2.2).

(2) The proof that (1) $\Rightarrow$ (3) shows that, if $\mathscr{S}(R^*)$ is an infinite set, then R^* has an imbedded prime divisor of zero. The converse is not true, as is seen if altitude $R = 2$.

(3) If altitude $R = 2$ and R^* has an imbedded prime divisor of zero, then $R^{(1)}$ is not a finite R-algebra (3.5)(1) but $R^{(1)} \subseteqq R_b$, for each nonzero $b \in M$. For an example of this, see [1, Proposition 3.3].

(4) It is easy to see that if altitude $R = 3$, then $\mathscr{S}(R)$ is an infinite set if and only if the Macaulay locus of R is not open. For, the Macaulay locus of R is open if and only if $\mathscr{S}(R)$ is a finite set, when altitude $R = 3$.

To generalize (3.1) to the case where R contains nonzero zero divisors, the following lemma is needed.

LEMMA 3.3. *Let* A *be a Noetherian ring which has no imbedded* *prime divisors of zero, and let altitude* $A > 0$. *Then* $\mathscr{I}(A)$ *is a finite set if and only if* $A^{(1)} = \bigcap\{A_{(p)}; p \in \operatorname{Spec} A, \text{ height } p = 1,$ *and* p *contains a regular element*$\} \subseteqq A_b$, *for some regular element* $b \in A$, *where* $A_{(p)} = \{a/c; a \in A$ *and* c *is a regular element in* A *and not in* $p\}$.

Proof. If $\mathscr{I}(A)$ is a finite set, then, since altitude $A > 0$, let b be a regular element in $\bigcap\{P; P \in \mathscr{I}(A)\}$, so $\mathscr{I}(A_b)$ is empty. Hence $A_b =$ [8, Corollary 2.18(2)] $A_b^{(1)} = (A^{(1)})_b \supseteqq A^{(1)}$. Conversely, if $A^{(1)} \subseteqq A_b$, then $(A^{(1)})_b = A_b \subseteqq A_b^{(1)} = (A^{(1)})_b$, so $A_b = A_b^{(1)}$, hence $b \in \bigcap\{P; P \in \mathscr{I}(A)\}$, by [8, Corollary 2.18(2)]. Thus $\mathscr{I}(A)$ is a finite set, since b is regular and bA has only finitely many prime divisors.

COROLLARY 3.4. *Let* R *be a local ring which has no imbedded prime divisors of zero, let altitude* $R > 0$, *and let* R^* *be the completion of* R. *Then the following statements are equivalent:*

(1) $\mathscr{I}(R)$ *is an infinite set.*

(2) *For no regular element* $b \in R$ *is* $R^{(1)} \subseteqq R_b$.

(3) *There exists an imbedded prime divisor* q^* *of zero in* R^* *such that* $q^* = \bigcap\{P \in \mathscr{I}(R^*); q^* \subset P$ *and* $P \cap R \in \mathscr{I}(R)\}$.

Proof. (1) $\Leftrightarrow$ (2) is given by (3.3).

$\mathscr{I}(R)$ is an infinite set if and only if $\mathscr{I}(R/q)$ is, for some (minimal) prime divisor q of zero (2.6), and $\mathscr{I}(R/q)$ is an infinite set if and only if there is an imbedded prime divisor $q^{*\prime} = q^*/qR^*$ of zero in R^*/qR^* (equivalently, there exists an imbedded prime divisor q^* of zero in R^* [5, (18.11)]) such that $q^{*\prime} = \bigcap\{P' \in \mathscr{I}(R^*/qR^*); q^{*\prime} \subset P'$ and $P' \cap (R/q) \in \mathscr{I}(R/q)\}$ (3.1). By (2.5), there exists $a \in R$, $\notin q$ such that, for all $Q \in \operatorname{Spec} R$ such that $q \subseteqq Q$ and $a \notin Q$, $Q \in \mathscr{I}(R)$ if and only if $Q/q \in \mathscr{I}(R/q)$. Since $a + q$ is a regular element in R/q, $q^{*\prime} = \bigcap\{P' \in \mathscr{I}(R^*/qR^*); q^{*\prime} \subset P', P' \cap (R/q) \in \mathscr{I}(R/q)$, and $a + q \notin P' \cap (R/q)\}$; so, by (2.5), $\mathscr{I}(R)$ is an infinite set if and only if there is an imbedded prime divisor q^* of zero in R^* such that $q^* = \bigcap\{P; P' = P/qR^* \in \mathscr{I}(R^*/qR^*), q^* \subset P$, and $P \cap R \in \mathscr{I}(R)\}$. Since P' is a prime divisor of $(P' \cap (R/q))(R^*/qR^*) = ((P \cap R)/q)(R^*/qR^*) = ((P \cap R)R^*)/qR^*$ (since $P' \in \mathscr{I}(R^*/qR^*)$ and $P' \cap (R/q) \in \mathscr{I}(R/q)$), P is a prime divisor of $(P \cap R)R^*$, hence $P \in \mathscr{I}(R^*)$ (2.3)(4).

Example 3.6 below shows somewhat more than (3.2)(3). For the example the following information is needed.

REMARKS 3.5. Let (R, M) be a local domain.

(1) [3, IV. (7.2.3)]. $R^{(1)}$ is a finite R-algebra if and only if the following condition holds: If $q \subset p$ are prime ideals in the M-adic

completion R^* of R such that q is a prime divisor of zero and height $p \cap R \geqq 2$, then height $p/q \geqq 2$.

(2) [7, Lemma 5.11(1)]. If R is unmixed, then $R^{(1)}$ is a finite R-algebra.

If $R^{(1)}$ is a finite R-algebra, then clearly $R^{(1)} \subseteqq R_b$, for some nonzero b in R. On the other hand, if $R^{(1)}$ is not a finite R-algebra, then there is at least one prime ideal p in R^* such that p contains a prime divisor q of zero such that height $p/q = 1 <$ height $p \cap R$ (3.5) (1), and necessarily $p \cap R \in \mathscr{I}(R)$ (since $p \in \mathscr{I}(R^*)$ (2.2), so p is a prime divisor of bR^*, for each nonzero $b \in p \cap R$, hence $p \cap R$ is a prime divisor of each such bR [5, (18.11)]). Since there is one such p, it might be asked if there are necessarily infinitely many such if depth $q > 1$ (so $p \neq M^*$). The answer is "no". In fact, the example below shows the stronger result that $\mathscr{I}(R)$ can be finite in this case.

EXAMPLE 3.6. *A local domain L such that $L^{(1)}$ is not a finite L-algebra and each imbedded prime divisor of zero in the completion of L has depth greater than one, but $L^{(1)} \subseteqq L_b$, for some nonzero b in L (hence $\mathscr{I}(L)$ is a finite set* (2.3)(5)). *Further, $\mathscr{I}(L^*)$ is an infinite set.* Let (R, M) be a local domain whose completion (R^*, M^*) has an imbedded prime divisor of zero, and assume $\mathscr{I}(R)$ is a finite set. (For example, let $R = A$ in [1, Proposition 3.3].) Let Q be an M-primary ideal, let $\mathscr{R} = \mathscr{R}(R, Q) = R[tQ, u]$ be the Rees ring of R with respect to Q (t is an indeterminate, $u = 1/t$, and $tQ = \{tm; m \in Q\}$), and let $\mathscr{R}^* = \mathscr{R}(R^*, Q^*)$, where $Q^* = QR^*$. Let $\mathscr{M} = (tQ, M, u)\mathscr{R}$ and $\mathscr{M}^* = (tQ^*, M^*, u)\mathscr{R}^*$ be the maximal homogeneous ideals in $\mathscr{R}$ and $\mathscr{R}^*$, respectively. Then $L = \mathscr{R}_{\mathscr{M}}$ is a dense subspace of $L' = \mathscr{R}^*_{\mathscr{M}^*}$ [7, Lemma 3.2], so, by (3.5)(1), to show that $L^{(1)}$ is not a finite L-algebra, it suffices to prove that there exist prime ideals $q' \subset p'$ in L' such that q' is an imbedded prime divisor of zero and height $p'/q' = 1 <$ height $p' \cap L$ (since this will then be reproduced in the completion L^* of L and L'). Also, since the prime divisors of zero in L' are the ideals $(qR^*[t, u] \cap \mathscr{R}^*)L'$ with q a prime divisor of zero in R^* [9, Theorem 1.5], there are no depth one prime divisors of zero in L' (since $qR^*[t, u] \cap \mathscr{R}^* \subset M^*R^*[t, u] \cap \mathscr{R}^* \subset \mathscr{M}^*$), so there will be none in L^* (since, for each prime ideal P in L', each prime divisor of PL^* has depth equal to depth P, by [5, (36.5) and Exercise p. 135] applied to L'/P (that L'/P satisfies the second chain condition for prime ideals follows from [7, Corollary 2.9], since $R^*/(P \cap R^*)$ does [5, (34.4)])). Let q be an imbedded prime divisor of zero in R^* and let $q^* = qR^*[t, u] \cap \mathscr{R}^*$, so q^* is an imbedded prime divisor of zero in $\mathscr{R}^*$. Let $I = (q^*, u)\mathscr{R}^*$. Then height $q + 1 \leqq$ height $I =$ (say) $h <$ height $\mathscr{M}^*$ (since $\mathscr{R}^*/q^* \cong \mathscr{R}(R^*/q, (Q^* + q)/q)$ [9, Theorem 2.1] and depth $q \geqq 1$, so depth $q^* \geqq 2$), and I is homogeneous, so there is a height h prime divisor p^* of I which is contained $\mathscr{M}^*$. Since $u\mathscr{R}^* \cap R^* = Q^*$ is M^*-primary, $\mathscr{R}_{p^*\cap\mathscr{R}}$ is a dense subspace of $\mathscr{R}^*_{p^*}$ [7, Lemma 3.2] (so height $p^* \cap \mathscr{R} =$ height $p^* = h >$ height $p^*/q^* = 1$), and $p^* \cap \mathscr{R} \subset \mathscr{M}$. However, $\mathscr{I}(L)$ is a finite set, since $\mathscr{I}(\mathscr{R})$ is, by (2.4). Finally, $\mathscr{I}(L^*)$ is an infinite set, since $\mathscr{I}(L')$ is by (2.2) applied to q^*L'.

REMARK 3.7. One reason for using the ring $L = \mathscr{R}_{\mathscr{M}}$ instead of $P = R[X]_{(M,X)}$ is that for the ring L it *is not* necessary to assume that $R^{(1)}$ is not a finite R-algebra to prove that $L^{(1)}$ is not a finite L-algebra, whereas for P it is necessary to assume this to prove that $P^{(1)}$ is not a finite P-algebra (but this does hold for $R = A$ in [1, Proposition 3.3]). A more important reason is it is an open problem if R must be unmixed when R is integrally closed. Related to this problem, the rings L (as Q varies) show that, if R is integrally closed and not unmixed, then there does not exist an M-primary ideal Q such that $L^{(1)}$ is a finite L-algebra.

4. Unramification of height one prime ideals. To obtain the main result of this section, the following two definitions and lemma are needed.

DEFINITION 4.1. A ring A is *locally factorial* in case A_P is a unique factorization domain, for all prime ideals P in A.

LEMMA 4.2. *Let B be a Noetherian integral domain which contains a locally factorial Noetherian domain A, and let P be a prime ideal in B such that P is a prime divisor of pB, for some height one prime ideal p in A. Then the following statements hold:*

(1) *P is a prime divisor of qB, for all height one prime ideals $q \subseteqq P \cap A$.*

(2) *P is an imbedded prime divisor of pB if and only if $P \in \mathscr{I}(B)$.*

(3) *P is a minimal prime divisor of pB if and only if height $P = 1$.*

Proof. Since, for each height one prime ideal $q \subseteqq P \cap A$, $qA_{P\cap A}$ is principal and $A_{P\cap A} \subseteqq B_P$, (1) and (2) hold, and (3) also follows from this and the Principal Ideal Theorem.

DEFINITION 4.3. (Cf. [5, pp. 144–145].) Let $A \subset B$ be rings, let P be a prime ideal in B, and let $p = P \cap A$. Then P is *unramified over A* in case $PB_P = pB_P$ and B/P is separably generated over A/p (that is, the quotient field of B/P is a separable extension field of the quotient field of A/p). If q is a prime ideal in A, then qB is *unramified over* A in case, for every prime divisor Q of qB, Q is unramified over A.

The following theorem is a considerable generalization of [4, Theorem 5] (see (4.6) below).

THEOREM 4.4. *Let B be a finite separably generated extension domain of a locally factorial Noetherian domain A. Then the following statements are equivalent, where $\mathscr{S} = \{p \in \operatorname{Spec} A;$ height $p = 1\}$:*

(1) *pB is unramified over A, for all except finitely many $p \in \mathscr{S}$.*

(2) *If P is a height one prime ideal in B, then height $P \cap A \leqq 1$, and only finitely many of the ideals pB, where $p \in \mathscr{S}$, have imbedded prime divisors.*

(3) *If P is a prime divisor of pB, where $p \in \mathscr{S}$, then $P \cap A = p$.*

Proof. (1) $\Rightarrow$ (2). By (1), with $p \in \mathscr{S}$, only finitely many pB have imbedded prime divisors. Also, if P is a height one prime ideal in B such that height $P \cap A > 1$, then clearly P is a prime divisor of pB, for all $p \in \mathscr{S}$ such that $p \subset P \cap A$. Since there are infinitely many such p, (1) implies that P is unramified over at least one such p, hence $P \cap A = p$; contradiction. Therefore, height $P \cap A \leqq 1$, as desired.

(2) $\Rightarrow$ (3). Let $p \in \mathscr{S}$ and let P be a prime divisor of pB. If P is an imbedded prime divisor of pB, then $P \cap A = p$, by (2) and (4.2)(1). If P is a minimal prime divisor of pB, the height $P = 1$ (4.2)(3), so (2) implies that $P \cap A = p$.

(3) $\Rightarrow$ (1). Since B is a finite separably generated extension domain of A, there exists a separating transcendence basis $X_1, \cdots, X_n$ in B of B over A and a nonzero element $a \in A$ such that B_a is integrally dependent on $D = A[1/a, X_1, \cdots, X_n]$ [5, (39.11)]. Then D is locally factorial, since A is. Let D' and B' be the integral closures of D and B in the quotient field of B, respectively, so $B'_a = D'$. Since B_a is a finite separable algebraic extension of D, let $b \in B_a$ such that $D[b]$ and B_a have the same quotient field, and let d be the discriminant of the minimum polynomial of b over the quotient field of D (so $d \in D$, since D is integrally closed). Then $D'_d = D_d)' = D_d[b]$ [5, (10.17) and (10.18)], so every prime ideal in D'_d is unramified over D_d [5, (38.9)]. Let $\mathscr{J}$ be the set of $p \in \mathscr{S}$ such that: (i) $a \in p$; or (ii) $d \in pD$; or (iii) pB has an imbedded prime divisor. That $\mathscr{J}$ is a finite set follows from: (i) A is Noetherian; (ii) D is Noetherian and pD is a height one prime ideal; and (iii) if P is an imbedded prime divisor of pB, then $P \in \mathscr{I}(B)$, by (4.2)(2), and $P \cap A = p$, by (3), hence there are only finitely many such P by (2.9)(1). It will now be shown that if $p \in \mathscr{S}, \notin \mathscr{J}$, then pB is unramified over A.

For this, note first that pD is a prime ideal, $pD \cap A = p$, and D/pD is a pure transcendental extension of $(A/p)_{\bar{a}}$, where $\bar{a} = a$ (modulo p). Therefore, since every prime ideal in D'_d is unramified over D_d, pD'_d is unramified over A ($p \in \mathscr{S}$, $\notin \mathscr{J}$).

Since $p \notin \mathscr{J}$, pB has no imbedded prime divisors, so if P is a prime divisor of pB, then height $P = 1$ (4.2)(3). Also, $a \notin P$, since $P \cap A = p$, by (3). Further, $d \notin PB_a$; for B_a is integrally dependent on D, hence height $PB_a \cap D = 1$, since D is integrally closed, and so $pD = PB_a \cap D$ and $d \notin pD$. Therefore, $B_P = (B_{ad})_{PB_{ad}}$ and $D_d[b] \subseteqq B_{ad} \subseteqq D'_d = D_d[b]$. Hence, since pD'_d is unramified over A, PB_P is unramified over A, and so pB is unramified over A.

REMARKS 4.5. (1) In (2) of (4.4), the condition: "Height $P \cap A \leqq 1$, for all height one prime ideals P in B"; is equivalent to: "Height $QB > 1$, for all prime ideals Q in A such that height $Q > 1$. Also, the condition: "Only finitely many of the ideals pB have imbedded prime divisors", is equivalent to: "If P is an imbedded prime divisor of pB, then $P \cap A = p$", by (4.2)(1) and (2) and (2.9)(1).

(2) A regular domain satisfies the conditions on A in (4.4).

(3) The proof of (4.4) shows that (1) $\Rightarrow$ (2) $\Rightarrow$ (3) for an arbitrary Noetherian domain which contains A, and (3) $\Rightarrow$ (1) with $\mathscr{C}$ replacing B in the statements, where either $\mathscr{C} = B'$ or $\mathscr{C}$ is a Noetherian ring such that $B \subseteqq \mathscr{C} \subseteqq B'$. (That $\mathscr{J}(\mathscr{C})$ is finite follows from: $a\mathscr{C}$ and $d\mathscr{C}_a$ have only finitely many imbedded prime divisors and $\mathscr{C}_{ad} = D'_d$.)

COROLLARY 4.6. (Cf. [4, Theorem 5].) *If B is a finite separably generated extension domain of a Dedekind domain A, then* (1)–(3) *of* (4.4) *hold.*

Proof. A is a locally factorial Noetherian domain, and (3) in (4.4) is satisfied.

COROLLARY 4.7. *Let A and B be as in* (4.4), *and assume that A is a regular local domain of altitude two. If $MB = B$, where M is the maximal ideal in A, then* (1)–(3) *of* (4.4) *hold.*

Proof. The proof is the same as the proof of (4.6).

COROLLARY 4.8. *Let A and B be as in* (4.4), *and assume that B is a flat A-algebra. Then* (1)–(3) *of* (4.4) *hold.*

Proof. (3) in (4.4) is satisfied.

COROLLARY 4.9. *Let A and B be as in* (4.4), *and assume* (1)–(3)

hold. Then B_P is a regular local ring, for all except finitely many height one prime ideals P in B.

Proof. Let P be a height one prime ideal in B. If $P \cap A \neq (0)$, then height $P \cap A = 1$ (4.4)(3), and only finitely many such P fail to satisfy B_P is a regular local ring (4.4)(1). If $P \cap A = (0)$, then let $S = A \sim (0)$, let $F = A_S$, and let $C = B_S$. Then D_S is a finite separably generated extension domain of the field F, hence it is well known [5, (36.6)] that the integral closure C' of C is a finite C-algebra. Therefore, only finitely many height one prime ideals Q in C fail to satisfy C_Q is a regular local ring (namely those Q which contain $C : C'$). Since PC is a height one prime ideal, for all such P (height $P = 1$ and $P \cap A = (0)$), the conclusion follows.

COROLLARY 4.10. *Let A and B be as in* (4.4), *let $\mathscr{C}$ be a ring such that either $\mathscr{C} = B'$, the integral closure of B, or $\mathscr{C}$ is Noetherian and $B \subseteqq \mathscr{C} \subseteqq B'$, and let $\mathscr{H} = \{P \in \operatorname{Spec} \mathscr{C};$ height $P = 1 \geqq$ height $P \cap A\}$. Then $\mathscr{C}_P$ is a regular local ring, for all except finitely many $P \in \mathscr{H}$.*

Proof. Only finitely many $P \in \mathscr{H}$ such that $P \cap A = (0)$ fail to satisfy $\mathscr{C}_P$ is a regular local ring, as in the proof of (4.9). Also, by the proof that (3) $\Rightarrow$ (1) in (4.4), for all $P \in \mathscr{H}$ such that $P \cap A \neq (0)$ and $a \notin P$ and $d \notin P\mathscr{C}_a$, $\mathscr{C}_P$ is a regular local ring. The conclusion follows from this.

REFERENCES

1. D. Ferrand and M. Raynaud, *Fibres formelles d'un anneau local Noethérian*, Ann. Sci. École Norm. Sup., **3** (1970), 295–311.
2. A. Grothendieck, *Elements de Géométrie Algebrique*, IV (Première Partie), Paris: Institut des Hautes Etudes Scientifiques, 1964.
3. ———, *Elements de Géométrie Algebrique*, IV (Seconde Partie), Paris: Institut des Hautes Etudes Scientifiques, 1965.
4. M. Nagata, *A general theory of algebraic geometry over Dedekind domains, II*, Amer. J. Math., **80** (1958), 382–420.
5. ———, *Local Rings*, New York, Interscience, 1962.
6. D. G. Northcott and D. Rees, *Principal systems*, Quart. J. Math. Oxford Ser., **8** (1957), 119–127.
7. L. J. Ratliff, Jr., *On quasi-unmixed local domains, the altitude formula, and the chain condition for prime ideals (II)*, Amer. J. Math., **92** (1970), 99–144.
8. ———, *On prime divisors of the integral closure of a principal ideal*, J. Reine Angew. Math., forthcoming.
9. D. Rees, *A note on form rings and ideals*, Mathematika, **4** (1957), 51–60.

Received July 5, 1972. Work on this paper was supported in part by the National Science Foundation Grant NSF-28939.

UNIVERSITY OF CALIFORNIA, RIVERSIDE

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

SOME COMMUTANTS IN $B(c)$ WHICH ARE ALMOST MATRICES

B. E. RHOADES AND A. WILANSKY

We determine necessary and sufficient conditions for two linear operators in $B(c)$ to commute. Specializing one of the operators to be a conservative triangular matrix we determine that most such operators have commutants consisting of almost matrices of a special form.

Almost matrices were developed in [10] for reasons not related to this paper, but they find application here in that the commutants in $B(c)$ of certain matrices must be almost matrices.

Let c denote the space of convergent sequences, $B(c)$ the algebra of all bounded linear operators over c, e the sequence of all ones, and e^k the coordinate sequences with a one in the kth position and zeros elsewhere. If $T \in B(c)$, then one can define continuous linear functionals χ and χ_i by $\chi(T) = \lim Te - \sum_k \lim (Te^k)$ and $\chi_i(T) = (Te)_i - \sum_k (Te^k)_i$, $i = 1, 2, \cdots$. (See, e.g. [9, p. 241].) It is known [1, p. 8] that any $T \in B(c)$ has the representation $T = v \otimes \lim + B$, where B is the matrix representation of the restriction of T to c_0, the subspace of null sequences, v is the bounded sequence $v = \{\chi_i(T)\}$, and $v \otimes \lim x = (\lim x)v$ for each $x \in c$.

The second adjoint of T (see, e.g. [1, p. 8] or [10, p. 357]) has the matrix representation

$$(*) \qquad T'' = \begin{pmatrix} \chi(T) & b_1 & b_2 & \cdot & \cdot & \cdot \\ \chi_1(T) & b_{11} & b_{12} & \cdot & \cdot & \cdot \\ \chi_2(T) & b_{21} & b_{22} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

where the b_i's occur in the representation of $\lim \circ T \in c'$ as $(\lim \circ T)(x) = \lim (Tx) = (T) \lim x + \sum_k b_k x_k$; namely, $b_i = \lim Te^i$. With the use of $(*)$ it is easy to describe the commutant of any $Q \in B(c)$.

THEOREM 1. *Let $Q = u \otimes \lim + A \in B(c)$. Then* Com (Q) *in* $B(c) = \{T = v \otimes \lim + B \in B(c)$: *$T$ satisfies* (1)–(3)$\}$, *where*

$$(1) \qquad u_n \chi(T) + \sum_{k=1}^{\infty} a_{nk} v_k = v_n \chi(Q) + \sum_{k=1}^{\infty} b_{nk} u_k \,; \qquad n = 1, 2, \cdots$$

$$(2) \qquad u_n b_k + \sum_{j=1}^{\infty} a_{nj} b_{jk} = v_n a_k + \sum_{j=1}^{\infty} b_{nj} a_{jk} \,; \qquad n, k = 1, 2, \cdots$$

$$(3) \qquad \sum_{k=1}^{\infty} a_k v_k = \sum_{k=1}^{\infty} b_k u_k ,$$

and where $a_k = \lim Q(e^k)$, $b_k = \lim T(e^k)$.

To prove Theorem 1, use the representation $(*)$ for T'' and Q'' and then equate the corresponding terms in the products $T''Q''$ and $Q''T''$. For example, (1) is obtained by equating $(Q''T'')_{n1}$ and $(T''Q'')_{n1}$. When Q is a matrix A, each $u_n = 0$ and each $a_k = \lim_n a_{nk}$. The following result is an immediate consequence of Theorem 1.

COROLLARY 1. *Let* A *be a conservative matrix,* $T \in B(c)$. *Then* $A \leftrightarrow T$ *if and only if*

$$(4) \qquad Av = \chi(A)v$$

$$(5) \qquad \sum_{j=1}^{\infty} a_{nj} b_{jk} = v_n a_k + \sum_{j=1}^{\infty} b_{nj} a_{jk} ; \qquad n, k = 1, 2, \cdots$$

$$(6) \qquad a \perp v, \text{ where } a = \{a_n\} .$$

A conservative matrix A is called multiplicative if $\lim_A x = \chi(A) \lim x$ for each $x \in c$; i.e., if each $a_k = 0$.

COROLLARY 2. *Let* A *be a conservative multiplicative matrix. Then* $A \leftrightarrow T$ *if and only if* A *satisfies* (4) *and*

$$(7) \qquad B \longleftrightarrow A .$$

If A is multiplicative, then each $a_k = 0$ and condition (5) of Corollary 1 reduces to (7) of Corollary 2. Since $a = 0$, (6) holds automatically.

THEOREM 2. *Let* A *be a conservative matrix. Then* $A \leftrightarrow v \otimes \lim$ *if and only if*

$$(8) \qquad (\lim x)Av = (\lim_A x)v \text{ for each } x \in c .$$

To establish (8) note that $A(v \otimes \lim)(x) = A(\lim x)v = (\lim x)Av$, and $(v \otimes \lim)(Ax) = (\lim Ax)v = (\lim_A x)v$.

COROLLARY 3. *Let* A *be a conservative multiplicative matrix. Then* $A \leftrightarrow u \otimes \lim$ *if and only if* A *satisfies* (4).

COROLLARY 4. *Let* A *be a conservative multiplicative matrix. Then* $A \leftrightarrow T$ *if and only if* $A \leftrightarrow v \otimes \lim$ *and* $A \leftrightarrow B$.

For $T \in B(c)$, T is called an almost matrix if $v \in c$. A matrix A

is called triangular if $a_{nk} = 0$ for each $k > n$. We shall now examine some triangular matrices whose commutants consist of almost matrices.

THEOREM 3. *Let A be a conservative triangular matrix with $a_{nn} \neq \chi(A)$ for $n > 1$. Consider the conditions*

$$\sum_{k=1}^{n} a_{nk} = \chi(A) \text{ for } n > 1 \tag{9}$$

$$T \leftrightarrow A \text{ implies } T \text{ is an almost matrix with } v = \lambda e . \tag{10}$$

Then $(9) \Rightarrow (10)$. *If, in addition,* $\lambda \neq 0$, *then* $(10) \Rightarrow (9)$.

To prove that $(9) \Rightarrow (10)$, suppose $T \leftrightarrow A$. From (4) of Corollary 1,

$$\sum_{k=1}^{n} a_{nk} v_k = \chi(A) v_n = \left(\sum_{k=1}^{n} a_{nk}\right) v_n , \qquad n > 1 .$$

We may rewrite the equation in the form $\sum_{k=1}^{n} (v_k - v_n) a_{nk} = 0$, which, along with the hypothesis $a_{nn} \neq \chi(A)$ for $n > 1$, yields $v_n = v_1$, for $n > 1$.

For $n > 1$, $(T''A'')_{n+1,1} = \lambda \chi(A)$ and $(A''T'')_{n+1,1} = \lambda \sum_{k=1}^{n} a_{nk}$. Thus, if $\lambda \neq 0$, $\chi(A) = \sum_{k=1}^{n} a_{nk}$.

The result stated at the end of paragraph 2 in the next section shows that the condition $\lambda \neq 0$ is necessary for (10) to imply (9).

The identity matrix shows that the restriction $a_{nn} \neq \chi(A)$ for $n > 1$ cannot be removed.

COROLLARY 5. *Let A be a conservative triangular matrix with $\sum_{k=1}^{n} a_{nk} = \chi(A)$ for $n > 1$ and $a_{nn} \neq \chi(A)$ for each n. Then $T \leftrightarrow A$ implies T is a matrix.*

From Theorem 3, $v_n = v_1$. From (4) with $n = 1$ we get $a_{11} v_1 = \chi(A) v_1$. Since $a_{11} \neq \chi(A)$, $v_1 = 0$ and A is a matrix.

Applications. 1. Let C denote the Casàro matrix of order 1. Then Theorem 3 of [7] follows immediately from Theorem 3 of this paper.

2. Endl [2], Hausdorff [4], Jakimovski [5] (see [11, p. 190]) and Leininger [6] have defined summability methods which are generalizations of the Hausdorff methods. The $(H, \lambda_n; \mu_n)$ transform of [5] is defined by a triangular matrix $H = (h_{nk})$ with entries $h_{nn} = \mu_n$, $h_{nk} = (-1)^{n-k} \lambda_{k+1} \cdots \lambda_n [\mu_k, \cdots, \mu_n]$, $k < n$, where

$$[\mu_k, \cdots, \mu_n] = \sum_{i=k}^{n} \frac{\mu_i}{(\lambda_i - \lambda_k) \cdots (\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1}) \cdots (\lambda_i - \lambda_n)} ,$$

$\{\mu_n\}$ is a real or complex sequence, and $\{\lambda_n\}$ satisfies $0 \leqq \lambda_0 < \lambda_1 < \cdots < \lambda_n < \cdots$, $\lim_n \lambda_n = \infty$ and $\sum_i \lambda_i^{-1} = \infty$. If $\lambda_n = n$, $n \geqq 0$, then $(H, \lambda_n; \mu_n)$ reduces to the ordinary Hausdorff transformations.

[4] is a special case of [5] with $\lambda_0 = 0$. [2] is the special case of [5] with $\lambda_n = n + \alpha$.

Each conservative method $(H, \lambda_n; \mu_n)$ with distinct diagonal entries and $\lambda_0 = 0$ satisfies the conditions of Theorem 3. Thus, if $T \leftrightarrow (H, \lambda_n; \mu_n)$; T is an almost matrix with $v = \lambda e$. If, in addition, $(H, \lambda_n; \mu_n)$ satisfies condition (1) of [7], then $T \leftrightarrow (H, \lambda_n; \mu_n)$ implies that B is a generalized Hausdorff matrix of the same type as $(H, \lambda_n; \mu_n)$.

If $\lambda_0 > 0$, then (9) of Theorem 3 is not satisfied. However, $\lim_n \sum_k h_{nk} = \mu_0$, and one can establish the following: Let $(H, \lambda_n; \mu_n)$ be a multiplicative generalized Hausdorff matrix with $\lambda_0 > 0$ and $\mu_n \neq \mu_0$ for all $n > 0$. Then Com $(H, \lambda_n; \mu_n)$ in Γ = Com $(H, \lambda_n; \mu_n)$ in $B(c)$.

The commutant question for the matrices of [6] remains open.

3. Let A be the shift, i.e., $a_{n+1,n} = 1$, $a_{nk} = 0$ otherwise. Then Theorem 1.1 of [8] follows from Corollary 5.

4. Let A be any regular Nörlund method with $p_n > 0$ for all n. (A matrix A is said to be regular if $\lim_A x = \lim x$ for each $x \in c$.) Then, by Theorem 3, if $T \leftrightarrow A$ then T is an almost matrix with $v = \lambda e$.

5. A triangle is a triangular matrix with each $a_{nn} \neq 0$. A factorable triangular matrix has entries of the form $a_{nk} = c_k d_n$, $k \leqq n$. Let A be a regular factorable triangle with all row sums one. By Theorem 3, if $T \leftrightarrow A$, then T is an almost matrix with $v = \lambda e$. This result holds, in particular, for the weighted mean methods (see [3, p. 57]).

THEOREM 4. *Let A be a conservative triangular matrix with $\sum_{k=1}^n a_{nk} = \chi(A)$ for each n, and $a_{nn} \neq \chi(A)$ for $n > 1$. Then the following are equivalent:*

(i) *A is multiplicative.*

(ii) *$T \leftrightarrow A$ if and only if there exists a scalar $\lambda \neq 0$ such that $T = \lambda e \otimes \lim + B$, where $B \leftrightarrow A$.*

(i) $\Rightarrow$ (ii). Suppose $T \leftrightarrow A$. By Corollary 2 we have (4) and $B \leftrightarrow A$. The hypotheses then allow us to use Theorem 3. Suppose now that T has the indicated form. Since $v = \lambda e$ and $\sum_{k=1}^n a_{nk} = \chi(A)$ for each n, A satisfies (4). By Corollary 2, $A \leftrightarrow T$.

(ii) $\Rightarrow$ (i). Using Corollary 4 and Theorem 2 we have (8). Set $x = e^k$ to get $a_k = 0$ for each k, since $\lambda \neq 0$. Thus A is multiplicative.

Note that the condition $\lambda \neq 0$ is not used in the proof of (i) $\Rightarrow$ (ii). However, it is necessary for the converse. For, let H denote the Hausdorff matrix generated by $\mu_n = n(n+1)^{-1}$, K the compact Hausdorff matrix generated by $\{1, 0, 0, \cdots\}$. Then, since $H = I - C$; where C is the Cesàro matrix of order 1, $A \leftrightarrow H$ if and only if $A \leftrightarrow C$. But $K \leftrightarrow C$. Therefore, $K \leftrightarrow H$ and K is not multiplicative.

The condition $\sum_{k=1}^{n} a_{nk} = \chi(A)$ for each n cannot be removed. For example, let A be the matrix defined by $a_{11} = 1$, $a_{2n+1,2n-1} = 1$, $a_{2n,2n} = (n+1)/n$, $n = 1, 2, \cdots$, $a_{nk} = 0$ otherwise. Let T be the operator with $v_{2n-1} = 1$, $v_{2n} = 0$, and B a diagonal matrix with $b_{2n,2n} = 1$, $b_{2n-1,2n-1} = 0$. Then $T \in B(c)$, A is regular, $a_{nn} \neq 1 = \chi(A)$ for any n, and $A \leftrightarrow T$, but T is not an almost matrix.

COROLLARY 6. *Let A satisfy the hypotheses of Theorem 4 with $\chi(A) = 1$. Then the following are equivalent:*

(i) *A is regular.*

(ii) *$T \leftrightarrow A$ if and only if there exists a scalar $\lambda \neq 0$ such that $T = \lambda e \otimes \lim + B$, where $B \leftrightarrow A$.*

In Theorem 4 merely observe that the conditions A multiplicative and $\chi(A) = 1$ imply A is regular.

A natural question to ask is whether there exist matrices whose commutant in $B(c)$ not only contains almost matrices different from those with $v = \lambda e$, but also such that Com (A) in $B(c)$ is included in the set of almost matrices. The answer is yes, as the following example illustrates.

Let v be a positive nonconstant convergent sequence with $v_n \neq 0$ for any n, $\lim_n v_n \neq 0$, $v_n/v_{n-1} \leqq 1$ for all n, and $\lim_n v_{n+1}/v_n = 1$. Let A be the matrix defined by $a_{11} = 1$, $a_{n,n-1} = v_n/v_{n-1}$, $n > 1$, $a_{nk} = 0$ otherwise. We wish to show that $A \leftrightarrow T = v \otimes \lim + B$, where $B \leftrightarrow A$. From Corollary 2 we need to verify (4) and (7).

To verify (4) for $n = 1$, $a_{11}v_1 = v_1 = \chi(A)v_1$. For $n > 1$, $A_n(v) = a_{n,n-1}v_{n-1} = v_n = \chi(A)v_n$.

It remains to determine those matrices B which commute with A. It is not difficult, using the techniques of [7], to show that Com (A) in Δ = Com (A) in Γ.

We shall now show that Com $(A) = \{f(A): f$ is analytic in $D = \{z: |z| \leqq 1\}\}$.

For convenience set $\alpha_n = v_{n+1}/v_n$. Suppose $B \leftrightarrow A$. Equating $(BA)_{n,k-1}$ and $(AB)_{n,k-1}$ we get, for $k > 2$,

$$b_{nk} = \frac{\alpha_{n-1}\alpha_{n-2} \cdots \alpha_{n-k+2}}{\alpha_{k-1} \cdots \alpha_2} b_{n-k+2,2} \, .$$

Thus we may write

$$b_{n,n-k} = \alpha_{n-1}\alpha_{n-2} \cdots \alpha_{n-k}\lambda_k \, , \qquad 1 \leqq k \leqq n-2 \, , \tag{11}$$

where $\lambda_k = b_{k+2,2}/\alpha_{k+1} \cdots \alpha_2$, $k \geqq 1$.

For $r = 1, 2, \cdots$,

$$(A^r)_{n,n-k} = \begin{cases} 1 , & n-k=1,\ k=1 \\ \alpha_1\alpha_2 \cdots \alpha_{n-1} , & n-k=1<n \leqq r+1 \\ \alpha_{n-1} \cdots \alpha_{n-r} , & r=k \\ 0 , & \text{otherwise} . \end{cases}$$

Note that for $n-k>1$, the only nonzero entries of A^r occur on the rth diagonal. Thus for any n, there exists only one nonzero element in row n. With λ_0 any arbitrary scalar, and for any fixed n, k with $n-k>1$, $\sum_{j=0}^{\infty} \lambda_j (A^j)_{n,n-k}$ has at most two nonzero terms. One is $\lambda_k (A^k)_{n,n-k}$ and the other is $\lambda_0 \delta^n_{n-k}$. Therefore,

$$\sum_{j=0}^{\infty} \lambda_j (A^j)_{n,n-k} = \left(\sum_{j=0}^{\infty} \lambda_j A^j\right)_{n,n-k} = (f(A))_{n,n-k} .$$

For $n-k=1$, $n>1$,

$$\sum_{j=0}^{\infty} \lambda_j (A^j)_{n1} = \sum_{j=n}^{\infty} \lambda_j (\alpha_1\alpha_2 \cdots \alpha_{n-1}) = (f(A))_{n1} .$$

For $n-k=1$, $n=1$,

$$\sum_{j=0}^{\infty} \lambda_j (A^j)_{11} = \sum_{j=0}^{\infty} \lambda_j = (f(A))_{11} ,$$

assuming $\sum_j \lambda_j$ converges, so that $B = f(A)$.

Using (11), we may write $\lambda_k = b_{n,n-k}/\alpha_{n-1}\alpha_{n-2} \cdots \alpha_{n-k}$; since $\alpha_1 \cdots \alpha_n = u_{n+1}/u_1$, we have

$$\sum_{k=1}^{n} |\lambda_k| = \sum_{k=1}^{n} \left| \frac{u_{n-k}}{u_n} b_{n,n-k} \right| = \frac{1}{u_n} \sum_{k=1}^{n} u_k | b_{nk} | .$$

Since $\| B \| < \infty$ and $\{u_n\}$ is bounded away from zero, $f(z) = \sum_j \lambda_j z^j$ is analytic in D.

Conversely, if B has the form $f(A)$ for some f analytic in D, then clearly B commutes with A.

We conclude with a few remarks concerning conull matrices. A conservative matrix is conull if $\chi(A) = 0$. From (4) of Corollary 1, $Av = 0$. Therefore, Com (A) in $B(c) = \{T \in B(c)\colon v \in \text{null space of } A\}$. If A is a triangle, then $v = 0$ and Com (A) in $B(c) =$ Com (A) in Γ. If A is triangular, with only a finite number of zeros on the main diagonal, then $v \in$ linear span $(e_1, e_2, \cdots, e_n)$, where n is the largest integer for which $a_{nn} = 0$. Of course, if A is the zero matrix, then Com (A) in $B(c) = B(c)$.

REFERENCES

1. H. I. Brown, D. R. Kerr, Jr., and H. H. Stratton, *The structure of* $B[c]$ *and extensions of the concept of conull matrix*, Proc. Amer. Math. Soc., **22** (1969), 7–14; **35** (1972), 515–518.

2. K. Endl, *Untersuchen über Momentprobleme bei Verfahren von Hausdorffschen Typus*, Math. Ann., **139** (1960), 403–432.

3. G. H. Hardy, *Divergent Series*, Oxford, 1949.

4. F. Hausdorff, *Summationsmethoden und Momentfolgen* II, Math. Zeit., **9** (1921), 280–299

5. A. Jakimovski, *The product of summability methods; new classes of transformations and their properties*, I, II, Technical (Scientific) Note No. 2, Contract No. AF61 (052)-187 (1959), Hebrew University.

6. C. W. Leininger, *Some properties of a generalized Hausdorff mean*, Proc. Amer. Math. Soc., **20** (1969), 88–96.

7. B. E. Rhoades, *Commutants of some Hausdorff matrices*, Pacific J. Math., **42** (1972), 715–719.

8. N. K. Sharma, *Spectral aspects of summability methods*, Ph. D. Dissertation, Indiana University, Bloomington, Indiana, (1971).

9. A. Wilansky, *Topological divisors of zero and Tauberian theorems*, Trans. Amer. Math. Soc., **113** (1964), 240–251.

10. ———, *Subalgebras of* $B(X)$, Proc. Amer. Math. Soc., **29** (1971), 355–360.

11. K. Zeller and W. Beekmann, *Theorie der Limitierungs-verfahren*, Springer-Verlag, Berlin-Heidelberg-New York, 1970.

Received August 30, 1972.

INDIANA UNIVERSITY
AND
LEHIGH UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

CROSS-SECTIONS OF DECOMPOSITIONS

J. P. RILEY

The following question was raised by R. H. Bing: "Is it true that if G is a monotone decomposition of E^3 into straight line intervals and one-point sets, then E^3/G is homeomorphic to E^3?" In his paper "Point-like decompositions of E^3" he described a possible counter example. This example has the interesting property that it has many tame cross-sections, but if its decomposition space is homeomorphic to E^3, its set of nondegenerate elements would have to form a wild Cantor set. This suggests that it would be interesting to study the connection between the embedding of a cross-section and the embedding of the set of nondegenerate elements in the decomposition space.

1. Introduction. Most of the terminology and notation used in this paper is standard. The reader is referred to [1], [3], [4], and [6].

If S is a 2-sphere in E^3, then by Int S we will mean the bounded component of E^3-S and by Ext S, the unbounded component.

Let G be an upper semi-continuous decomposition of E^3 and let H be the set of all nondegenerate elements of G. We will say that a set $R \subset E^3$ is a cross-section of G if (i) $R \cap h$ is a singleton for each $h \in H$, and (ii) the natural map P restricted to R is homeomorphism onto $\overline{P(H)}$. We note that cross-sections exist only for certain decompositions. A simple example may be constructed as follows: Let $a_n = 1/n$, for $n = 1, 2, \cdots$ and let $b_n = -1/n$ for $n = 1, 2, \cdots$. Let the set of nondegenerate elements of our decomposition consist of the closed interval from $(0, 1, 0)$ to $(0, -1, 0)$, the closed interval from $(a_n, 1/2, 0)$ to $(a_n, 1, 0)$ for each positive integer n, and the closed interval from $(b_n, -1/2, 0)$ to $(b_n, -1, 0)$ for each positive integer n.

II. Cross-sections of decompositions. The following question naturally arises: How are the embeddings of a cross-section R and $\overline{P(H)}$ related when E^3/G is homeomorphic to E^3? We will give some partial results to this question.

THEOREM 1. *Let G be an upper semi-continuous decomposition of E^3 into points and straight line intervals pointing in only a countable number of directions whose lengths are bounded away from zero such that $P(H)$ is a compact 0-dimensional set. If there exists a cross-section C of G then C is tame.*

Proof. In the special case where the elements of H point in only one direction, we can easily show the tameness by a modification of the proof of Theorem 2 of [7].

Suppose that $H = \bigcup_{n=1}^{\infty} H_n$ where the elements of H_n are all parallel and if $h_1 \in H_i$ and $h_2 \in H_j$ where $j \neq i$ then h_1 is not parallel to h_2. Let C_n be the set of all points $c \in C$ such that $c \in h$ for some $h \in H_n$. Let G_n be the upper semi-continuous decomposition of E^3 whose only nondegenerate elements are the elements of H_n and let P_n be the natural map. Then E^3/G_n is homeomorphic to E^3 and $P_n(H_n)$ is tame in E^3/G_n. So by the special case C_n is tame and by Corollary 2 to Theorem 3 of [7], C is tame.

The following two lemmas will be stated without proof. Their proofs are similar to that of Lemma A of [7] and use standard techniques. Lemma B is similar to Theorem 2.3 of [3].

LEMMA A. *Let G be an upper semi-continuous decomposition of E^3 such that $P(H)$ is a compact 0-dimensional set. Let $h \in H$ and suppose that there exist 2-spheres S_1 and S_2 such that $h \subset \operatorname{Int} S_1 \cap \operatorname{Int} S_2$ and $(S_1 \cup S_2) \cap (\cup H) = \emptyset$. Then there exists a 2-spheres S such that $h \subset \operatorname{Int} S$, $S \cup \operatorname{Int} S \subset S_1 \cup \operatorname{Int} S_1$, and if $k \in H$ then $k \subset \operatorname{Int} S$ iff $k \subset \operatorname{Int} S_1 \cap \operatorname{Int} S_2$.*

LEMMA B. Let *$S_1, S_2, \cdots, S_n$ be a finite collection of 2-sphere whose interiors cover $\cup H$ and which miss $\cup H$. Then there exists a finite collection of 2-spheres $R_1, R_2, \cdots, R_n$ such that $R_1 = S_1$, $(R_i \cup \operatorname{Int} R_i) \cap (R_j \cup \operatorname{Int} R_j) = \emptyset$ if $i \neq j$, and $h \subset \operatorname{Int} R_i$ iff $h \subset \operatorname{Int} S_i$ and $h \cap \operatorname{Int} S_j = \emptyset$ for $j < i$.*

THEOREM 2. *Let C be a wild Cantor set in E^3 with the property that if x and y are distinct points of C, then there exist disjoint 2-spheres S_1 and S_2 such that $(S_1 \cup S_2) \cap C = \emptyset$, $x \in \operatorname{Int} S_1 \cap \operatorname{Ext} S_2$ and $y \in \operatorname{Int} S_2 \cap \operatorname{Ext} S_1$. Then there exists a monotone decomposition G of E^3 such that C is a cross-section for G, E^3/G is homeomorphic to E^3 and $P(\bar{H})$ is tame.*

Proof. Let C be a wild Cantor set in E^3 with the required property. For each $x \in C$ we choose a 2-sphere $S_1(x)$ as follows:

Let $N_1(x)$ be a 2-sphere of radius 1/2, centered at x. Let $C_1(x) = \{t \in C \mid t \notin \operatorname{Int} N_1(x)\}$. Then for each $y \in C_1(x)$ choose disjoint 2-spheres $S(y)$ and $R(y)$ such that $(S(y) \cup R(y)) \cap C = \emptyset$, $x \in \operatorname{Int} S(y) \cap \operatorname{Ext} R(y)$, and $y \in \operatorname{Int} R(y) \cap \operatorname{Ext} S(y)$. Now choose a set $y_1, y_2, \cdots, y_n$ of elements of $C_1(x)$ such that $\{\operatorname{Int} R(y_1), \operatorname{Int} R(y_2), \cdots, \operatorname{Int} R(y_n)\}$ covers $C_1(x)$. We now apply Lemma A to get a 2-sphere $S_1(x)$ such that $x \in \operatorname{Int} S_1(x)$, $S_1(x) \cap C = \emptyset$, $C_1(x) \subset \operatorname{Ext} S_1(x)$ and $S_1(x) \subset S(y_i) \cup \operatorname{Int} S(y_i)$ for $i = 1, 2, \cdots, n$. Therefore, there exists a finite collection of points $x_1, x_2, \cdots,$

$x_{m(1)}$ of C such that $C \subset \text{Int } S_1(x_1) \cup \text{Int } S_1(x_2) \cup \cdots \cup \text{Int } S_1(x_{m(1)})$. We replace $\mathscr{S}_1 = \{S_1(x_1), S_1(x_2), \cdots, S_1(x_{m(1)})\}$ by another collection of 2-spheres $\mathscr{T}_1 = \{T_{11}, T_{12}, \cdots, T_{1n(1)}\}$ satisfying the conclusions of Lemma B with respect to $\mathscr{S}_1$.

We will now proceed to construct a sequence $\mathscr{T}_1, \mathscr{T}_2, \mathscr{T}_3, \cdots$ of finite covers of C. Suppose that $\mathscr{T}_{k-1}$ has been chosen. For each point $x \in C$ we choose a 2-sphere $N_k(x)$ centered at x with radius $1/2^k$. We then proceed to choose $\mathscr{T}_k$ by the same process as in the construction of $\mathscr{T}_1$. We note that if $y_1, y_2 \in T_{kj} \cap C$ then $d(y_1, y_2) < 1/2^{k-1}$ since $T_{jk} \cap C \subset N_k(x)$ for some $x \in C$. Now for $x \in C$ we define h_x to be $\bigcap_{k=1}^{\infty} (T_{ki} \cup \text{Int } T_{ki})$ where T_{ki} is the 2-sphere in T_k whose interior contains x. Let G be the decomposition of E^3 whose only nondegenerate elements are the nondegenerate elements of $\{h_x \mid x \in C\}$. It follows easily that G is upper semi-continuous and it is clear that C is a cross-section for G. A theorem of Harrold [5] shows that E^3/G is homeomorphic to E^3 and from the criteria of [3], we see that $\overline{P(H)}$ is tame.

The Cantor set constructed in [2] is an example of a wild Cantor set satisfying the hypothesis of Theorem 2.

We can note that if C is a wild Cantor set in E^3 which does not satisfy the condition of Theorem 2, also, if C is a cross-section of a decomposition G whose decomposition space is homeomorphic to E^3 then $P(H_G)$ is a wild Cantor set which does not satisfy the condition of Theorem 2.

References

1. S. Armentrout, *Monotone decompositions of E^3*, Topology Seminar Wisconsin, 1965, Princeton (1966), 1-25.
2. R. H. Bing, *A homeomorphism between the 3-sphere and the sum of two solid horned spheres*, Annals of Mathematics, **56** (1952), 354-362.
3. ———, *Tame Cantor sets in E^3*, Pacific J. Math., **11** (1961), 435-446.
4. ———, *Decompositions of E^3*, Topology of 3-manifolds and Related Topics, Prentice Hall (1962), 5-21.
5. O. G. Harrold, Jr., *A sufficient condition that a monotone image of the three-sphere be a topological three-sphere*, Proc. Amer. Math. Soc., **9** (1958), 846-850.
6. R. L. Moore, *Foundations of Point Set Theory*, (new ed.), Amer. Math. Soc. Colloquium Publications, **13** (1962).
7. J. P. Riley, *Decompositions of E^3 and the tameness of their sets of nondegenerate elements*, Duke Math. J., **38** No. 2, (June, 1971).

Received May 16, 1972 and in revised form September 31, 1972.

UNIVERSITY OF DELAWARE

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

A CHARACTERIZATION OF THE MACKEY UNIFORMITY $m(L^\infty, L^1)$ FOR FINITE MEASURES

K. D. STROYAN

Let μ be a finite positive measure on a σ-algebra $\mathscr{M}$ over a set X. As usual $L^\infty(\mu)$ denotes the space of μ-essentially bounded measurable functions and $L^1(\mu)$ denotes the space of μ-integrable functions. In this article we use nonstandard analysis to give a simple description of the Mackey uniformity $m(L^\infty, L^1)$. The Mackey uniformity is the finest locally convex linear uniformity on L^∞ for which each continuous linear functional has an L^1 representation. The famous theorem of Mackey-Arens says it is given by uniform convergence on the weakly compact subsets of L^1.

Our description is simply this: Let p be a seminorm on L^∞. Then p is Mackey continuous if and only if whenever g is a finitely bounded element of the nonstandard extension $^*L^\infty$ which is infinitesimal, except possibly on a set of infinitesimal internal measure, then $p(g)$ is infinitesimal.

For the reader who is unfamiliar with nonstandard analysis we remark that ψ is L^∞-norm continuous at f if and only if whenever g is a finitely bounded element of the nonstandard extension which is infinitesimally close to f, except possibly on a set of zero internal measure, then $\psi(g)$ is infinitesimally close to $\psi(f)$. (This follows easily from Robinson's treatment of metric spaces.) We write $f \overset{n}{=} g$ if ($\|f - g\|_\infty$ is finite and) $f(x)$ is infinitesimally close to $g(x), f(x) \approx g(x)$, except possibly on a set of measure zero and say f is a norm-infinitesimal from g.

This characterization uses the idea of a linear infinitesimal relation which generalizes the nonstandard treatment of metric and uniform spaces given by Robinson [6], Luxemburg [3], and Machover and Hirschfeld [5]. The generalization first appeared in the authors dissertation in the context of bounded holomorphic functions, see Stroyan [7] and Luxemburg and Stroyan [4]. The reader is referred to the references [3, 4, 5, 6] for an introduction to standard analysis which we shall not give.

We say a measure λ is μ-continuous if for every $\varepsilon \in \boldsymbol{R}^+$ there is a $\delta \in \boldsymbol{R}^+$ so that whenever $\mu(E) < \delta$ for $E \in \mathscr{M}$, then $|\lambda(E)| < \varepsilon$.

In the nonstandard model an internal measure (or *-measure) λ is called *μ-S-continuous* if $\mu(E) \approx 0$, for $E \in {}^*\mathscr{M}$, implies $\lambda(E) \approx 0$. A function $f \in {}^*L^1$ is μ-S-continuous if the *-measure $\lambda(E) = \int_E f(x) d\mu(x)$ is μ-S-continuous. This is equivalent to saying that for every standard

$\varepsilon \in {}^{\sigma}\boldsymbol{R}^{+} = \{{}^{*}r : r \in \boldsymbol{R}^{+}\}$ there exists a standard $\delta \in {}^{\sigma}\boldsymbol{R}^{+}$ so that if $\mu(E) < \delta$, then $|\lambda(E)| < \varepsilon$.

LEMMA 1. *If $K \subseteqq L^1(\mu)$ is weakly compact, then K is L^1-norm bounded and uniformly μ-continuous, that is, for every $\varepsilon \in \boldsymbol{R}^{+}$, there exists a $\delta \in \boldsymbol{R}^{+}$ so that $\mu(E) < \delta$ implies $\left|\int_E k(x)d\mu(x)\right| < \varepsilon$ for any $k \in K$.*

Proof. This standard result can be found, for example, in Dunford and Schwartz [1].

We wish to point out here that K is uniformly μ-continuous if and only if each member of ${}^{*}K$ is μ-S-continuous. To see the equivalence of these conditions observe that uniform continuity is expressed by the formal sentence

$$(\forall \varepsilon \in \boldsymbol{R}^{+})(\exists \delta \in \boldsymbol{R}^{+})(\forall k \in K)(\forall E \in \mathscr{M})\left[\mu(E) < \delta \Rightarrow \left|\int_E k d\mu\right| < \varepsilon\right].$$

By Leibniz principle (that 'whatever' is true or false for the standard model is true or false for the nonstandard or ideal one) we have the equivalent sentence in the nonstandard model

$$(\forall \varepsilon \in {}^{*}\boldsymbol{R}^{+})(\exists \delta \in {}^{*}\boldsymbol{R}^{+})(\forall k \in {}^{*}K)(\forall E \in {}^{*}\mathscr{M})\left[\mu(E) < \delta \Rightarrow \left|\int_E k d\mu\right| < \varepsilon\right].$$

If K is standard and uniformly μ-continuous and $E_0 \in {}^{*}\mathscr{M}$ has infinitesimal μ-measure, take $\varepsilon_0 \in {}^{\sigma}\boldsymbol{R}^{+}$, a standard positive tolerance, and apply the $\varepsilon - \delta$ formula in the standard model to that particular ε_0. That is, there is a standard δ_0, etc. Now shift the particular sentence

$$(\forall k \in K)(\forall E \in \mathscr{M})\left[\mu(E) < \delta_0 \Rightarrow \left|\int_E k d\mu\right| < \varepsilon_0\right]$$

to the nonstandard model (put *'s on K and $\mathscr{M}$). Since $\mu(E_0) < \delta$, the integral is less than an arbitrary standard positive ε_0, hence infinitesimal.

Conversely, if each member of ${}^{*}K$ is μ-S-continuous and $\varepsilon_0 \in {}^{\sigma}\boldsymbol{R}^{+}$ is given, then taking $\delta \approx 0$ we see that the formula

$$(\exists \delta \in {}^{*}\boldsymbol{R}^{+})(\forall k \in {}^{*}K)(\forall E \in {}^{*}\mathscr{M})\left[\mu(E) < \delta \Rightarrow \left|\int_E k d\mu\right| < \varepsilon_0\right]$$

holds in the nonstandard model. But this formula has a standard interpretation (without the *'s) which amounts to uniform μ-continuity for that particular ε_0. Since ε_0 was an arbitrary standard tolerance we are done.

Another simple nonstandard reformulation is as follows.

LEMMA 2. *Let $\mathscr{F}$ be a family of functions from a set Y into $\boldsymbol{C}$. Let Σ be a collection of subsets of Y. The uniformity of uniform convergence on the sets of Σ is characterized by the infinitesimal relation on $^*\mathscr{F}$ given by "$f \overset{\Sigma}{=} g$ if and only if $f(s) \approx g(s)$ for all $s \in \bigcup [^*S: S \in \Sigma]$". More precisely, the entourages of that uniformity are exactly those subsets U of $\mathscr{F} \times \mathscr{F}$ for which $^*U \supseteqq \{(f, g): f \overset{\Sigma}{=} g\}$ and $\{(f, g): f \overset{\Sigma}{=} g\} = \bigcap [^*U: U$ is an entourage in the standard model$]$.*

Proof. The seminorms $\sup [|f(s) - g(s)|: s \in S]$ characterize this uniformity and Luxemburg [3] has shown that the monad of the uniformity given by $\{(f, g): \lambda(f, g) \approx 0$ for all standard semimetrics λ in a gauge of a uniformity$\}$ characterizes the uniformity (in an enlargement).

If $f(s) \approx g(s)$ for $s \in \bigcup (^*S: S \in \Sigma)$ then the set $\{|f(s) - g(s)|: s \in {}^*S\}$ is a bounded internal set. In fact, since it contains only infinitesimals it is bounded by every positively finite number and since that is an external set it actually has infinitesimal bounds. This means that the standard semimetrics are infinitesimal and the converse is clear.

We apply Lemma 2 to the Mackey uniformity to see that in $^*L^\infty$ the Mackey infinitesimals are characterized by

$$\text{"}f \overset{m}{=} g \text{ if and only if } \int_X f(x)k(x)d\mu(x) \approx \int_X g(x)k(x)d\mu(x)$$
$$\text{for every weakly compact } k \in \mathrm{cpt}_w\,(^*L^1)\text{"}\,.$$

The weakly compact points of $^*L^1$ are given by

$$\mathrm{cpt}_w\,(^*L^1) = \bigcup\,[^*K: K \text{ is a weakly compact subset of } L^1]\,.$$

Now we apply Lemma 1 to see that the weakly compact points of $^*L^1$ are norm finitely bounded and μ-S-continuous. This observation makes it clear that the infinitesimal relation:

$$\text{"}f \overset{M}{=} g \text{ if and only if } \|f - g\|_\infty \text{ is finite and } f(x) \approx g(x)$$
$$\text{except on a set of infinitesimal internal measure"}\,,$$

is finer than the Mackey infinitesimals. This is because if k is merely a L^1-norm finite μ-S-continuous $^*L^1$ function and $f(x) \approx g(x)$ except on E with $\mu(E) \approx 0$, then

$$\left|\int_X (f(x) - g(x))k(x)d\mu(x)\right| \leqq \left|\int_{X\setminus E} (f - g)kd\mu\right| + \|f - g\|\left|\int_E |k d\mu\right|$$

and both terms on the right side are infinitesimal so that

$$\int_X fkd\mu \approx \int_X gkd\mu\,.$$

The relation $\overset{M}{=}$ is not the monad of a uniformity [3] but it is close enough to $\overset{m}{=}$ to recapture it.

Suppose now that $\varphi: L^\infty \to C$ is a standard linear functional satisfying the continuity requirement that whenever $f \overset{M}{=} g$ in $^*L^\infty$, then $\varphi(f) \approx \varphi(g)$. We wish to draw two immediate conclusions from this. First, φ is norm-continuous since $f \overset{n}{=} g$ implies $f \overset{M}{=} g$. Second, φ induces a μ-S-continuous standard measure on X via

$$\Phi(E) = \varphi(\chi_E) , \quad \text{where} \quad \chi_E(x) = \begin{cases} 0, & x \notin E \\ 1, & x \in E . \end{cases}$$

The next result says Φ is countably additive.

LEMMA 3. *If λ is a μ-S-continuous finite internal measure, then $\Lambda(E) = \mathrm{st}\,(\lambda(^*E))$, for $E \in \mathscr{M}$, is countably additive.*

Proof. Λ is finitely additive by the additivity of st. Given $\varepsilon \in {}^\sigma\boldsymbol{R}^+$ there is a $\delta \in {}^\sigma\boldsymbol{R}^+$ so that if $\mu(E) < \delta$, then $|\lambda(E)| < \varepsilon$. Now take a partition E_k of X. The sum $\Sigma\mu(E_k)$ converges so given δ there is an l so that $\sum_{k=l}^\infty \mu(E_k) < \delta$, hence $|\lambda(\bigcup_{k \geqq l} E_k)| < \varepsilon$ and

$$\left| \Lambda(\bigcup E_k) - \sum_{k=1}^{l} \Lambda(E_k) \right| \leqq \varepsilon .$$

Now we can apply the Radon-Nikodym theorem to get an L^1 representation for φ. Therefore, $\overset{M}{=}$-continuous standard linear functionals are in L^1, or in other words, $\overset{M}{=}$-continuity is compatible with the dual pair $\langle L^\infty, L^1 \rangle$.

Let ψ be an arbitrary (linear or not) standard functional on $^*L^\infty$ which satisfies $\overset{M}{=}$-continuity: $f \overset{M}{=} g$ implies $\psi(f) \approx \psi(g)$. Define a uniformity on L^∞ by the semimetrics

$$|\psi(f) - \psi(g)|$$

for ψ standard and $\overset{M}{=}$-continuous.

LEMMA 4. *If $\|f - g\|_\infty$ is finite, then*

$$f \overset{M}{=} g \textit{ if and only if } \|f - g\|_1 \approx 0 .$$

Proof. If $f \overset{M}{=} g$, then $f(x) \approx g(x)$ except on E with $\mu(E) \approx 0$, so

$$\int |f - g|\, d\mu \leqq \int_{X \setminus E} |f - g|\, d\mu + \int_E \|f - g\|_\infty\, d\mu \approx 0 .$$

Conversely, suppose $\int |f - g|\, d\mu \approx 0$. For each $n \in {}^*N$ define the internal sequence

$$\varepsilon_n = \mu\{x: |f(x) - g(x)| > 1/n\}\,.$$

We know that for standard $n \in {}^\sigma N$, $\varepsilon_n \approx 0$ and Robinson's infinitesimal sequence lemma ([6], Theorem 3.3.20 or [4], Theorem 8.1.4) says ε_n is infinitesimal out to some infinite subscript, so $f \overset{M}{=} g$.

Fix a standard functional ψ, we will show that there exists a sequence ε_n so that

$$\bigcup\,[F(n, \varepsilon_n): n \in N] \subseteqq \{(f, g): |\psi(f) - \psi(g)| < 1\}$$

where

$$F(n, \varepsilon) = \{(f, g): \|f - g\|_\infty < n \quad \text{and} \quad \|f - g\|_1 < \varepsilon\}\,.$$

Take $n \in N$, since $\overset{M}{=}$ agrees with L^1-infinitesimals on the set $\{(f, g): \|f - g\|_\infty < n\}$ we know $(\exists \varepsilon \in {}^*\boldsymbol{R}^+)[F(n, \varepsilon) \subseteqq \{(f, g): |\psi(f) - \psi(g)| < 1\}]$ holds in the nonstandard model by taking $\varepsilon \approx 0$. Therefore, the same sentence holds in the standard model, so select such an ε and call it ε_n.

Sets of the form $\bigcup\,[F(n, \varepsilon_n): n \in N]$ generate a standard linear uniformity finer than that generated by the ψ's. This is the finest uniformity agreeing with the L^1-norm on L^∞-norm bounded sets.

Finally, consider the collection of standard $\overset{M}{=}$-continuous seminorms p on L^∞, that is, if $f \overset{M}{=} g$ then $p(f - g) \approx 0$. These generate the finest locally convex linear uniformity whose monad contains $\overset{M}{=}$. This is the Mackey topology since any Mackey-continuous seminorm is $\overset{M}{=}$-continuous and every $\overset{M}{=}$-continuous linear functional has an L^1 representation.

We have shown:

THEOREM. *The Mackey uniformity $m(L^\infty, L^1)$ for a finite measure is characterized by the infinitesimal relation on ${}^*L^\infty$ given by:*

"$f \overset{M}{=} g$ if and only if $\|f - g\|_\infty$ is finite and $f(x)$ is infinitesimally close to $g(x)$, except possibility on a set of infinitesimal measure".

Precisely, a standard seminorm $p: L^\infty \longrightarrow \boldsymbol{R}^+$ is Mackey continuous if and only if whenever $f \overset{M}{=} g$, then $p(f - g) \approx 0$.

REFERENCES

1. N. Dunford and J. T. Schwartz, *Linear Operators Part I*, Pure and Applied Math., Interscience, New York, 1958.
2. W. A. J. Luxemburg, editor, *Applications of Model Theory to Algebra, Analysis, and Probability*, Holt, Rinehart and Winston, New York, 1969.
3. ———, *A General Theory of Monads*, in [**2**].
4. W. A. J. Luxemburg and K. D. Stroyan, *Lecture Notes on Nonstandard Analysis* to appear.
5. M. Machover and J. Hirschfeld, *Lectures on Non-Standard Analysis*, Springer-Verlag lecture notes 94, Berlin, 1969.
6. A. Robinson, *Non-Standard Analysis*, Studies in Logic and the Foudations of Math., North Holland, Amsterdam, 1966.
7. K. D. Stroyan, *Infinitesimal relations on the space of bounded holomorphic functions*, (to appear).

Received August 21, 1972. The author was supported by NSF Grant GP-24182.

UNIVERSITY OF IOWA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

THE SCHOLZ-BRAUER PROBLEM ON ADDITION CHAINS

EDWARD G. THURBER

An addition chain for a positive integer n is a set $1 = a_0 < a_1 < \cdots < a_r = n$ of integers such that every element a_i is the sum $a_j + a_k$ of two preceding members (not necessarily distinct) of the set. The smallest length r for which an addition chain for n exists is denoted by $l(n)$. Let $\lambda(n) = [\log_2 n]$, and let $\nu(n)$ denote the number of ones in the binary representation of n. The purpose of this paper is to show how to establish the result that if $\nu(n) \geqq 9$ then $l(n) \geqq \lambda(n) + 4$. This is the $m = 3$ case of the conjecture that if $\nu(n) \geqq 2^m + 1$ then $l(n) \geqq \lambda(n) + m + 1$ for which cases $m = 0, 1, 2$ have previously been estabished. The fact that the conjecture is true for $m = 3$ leads to the theorem that $n = 2^m(23) + 7$ for $m \geqq 5$ is an infinite class of integers for which $l(2n) = l(n)$. The paper concludes with this result.

An addition chain for a positive integer n is a set $1 = a_0 < a_1 < a_2 < \cdots < a_r = n$ of integers such that every element a_i is the sum $a_j + a_k$ of two preceding members (not necessarily distinct) of the set. The smallest length r for which an addition chain for n exists is denoted by $l(n)$. Let $\lambda(n) = [\log_2 n]$, and let $\nu(n)$ denote the number of ones in the binary representation of n. Step i of an addition chain is $a_i = a_j + a_k$ for some $k \leqq j < i$. Since $a_i \leqq 2a_j \leqq 2a_{i-1}$, either $\lambda(a_i) = \lambda(a_{i-1})$ or $\lambda(a_i) = \lambda(a_{i-1}) + 1$. Step i is called a small step in the former case and a big step in the latter case. Since $a_i \leqq 2a_{i-1}$, a member of the chain must occur in each of the half-open intervals $[2^k, 2^{k+1})$ for $0 \leqq k \leqq \lambda(n)$. Every time a step takes the chain from one interval to the next it is a big step; otherwise, it is a small step. There are $\lambda(n)$ big steps in the chain, and the remaining steps are small steps. If $N(a_i)$ represents the number of small steps in the chain to a_i, then the length r of the chain may be expressed as $r = \lambda(n) + N(n)$.

A conjecture which is equivalent to one made by K. B. Stolarsky [10] states that if $\nu(n) \geqq 2^m + 1$, then $l(n) \geqq \lambda(n) + m + 1$. That is to say if $\nu(n) \geqq 2^m + 1$, then there are at least $m + 1$ small steps in any chain for n. The conjecture is true for $m = 0, 1, 2$. These results may be found in [8] with the case $m = 2$ being part of D. E. Knuth's Theorem C. The primary purpose of this paper is to show how to establish the conjecture for $m = 3$ and to show this case leads to the result that there is an infinite class of integers for which $l(2n) = l(n)$.

If a_j and a_k are two integers written in binary notation and placed one on top of the other in order to add or subtract, the resultant figure is called a configuration and is designated by a_j/a_k. The configuration is divided up into slots numbered from left to right. If $a_j = 101100111$ and $a_k = 10101110$, then a_j/a_k is as follows:

$$\begin{array}{rl} & 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9 \\ a_j = & 1\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 1 \\ a_k = & \ 1\ 0\ 1\ 0\ 1\ 1\ 1\ 0\ . \end{array}$$

The slot numbers are written above. Slot 4 is called a 1/1 slot, slot 9 is a 1/0 slot etc. Two lemmas which involve integers written in their binary notation are the following:

LEMMA 1. *If $a_i = a_j + a_k$ and if c represents the number of carries in $a_j + a_k$, then $\nu(a_i) = \nu(a_j) + \nu(a_k) - c$.*

LEMMA 2. *If $a_t = a_j - a_k$ and there are s 1/1 slots in a_j/a_k and a one appears in a_t p times under either a 1/1 slot or a 0/0 slot, then $\nu(a_t) = \nu(a_j) - s + p$.*

Two further lemmas will now be given which involve numbers in an addition chain.

LEMMA 3. *If a_j and a_k are two members of an addition chain and if $\lambda(a_j) = \lambda(a_k) + m (m \geqq 0)$ and $2^m a_k < a_j$, then $N(a_j) \geqq N(a_k) + 1$.*

Proof. Since $\lambda(a_j) = \lambda(a_k) + m$, there are precisely m big steps from a_k to a_j in the chain, but $2^m a_k < a_j$ implies that there are at least $m + 1$ steps in the chain from a_k to a_j; hence, at least one of them is a small step.

LEMMA 4. *If a_j and a_k are two members of an addition chain and if $\lambda(a_j) = \lambda(a_k) + m (m \geqq 2)$ and $a_j > 2^{m-1}a_k + 2^{m-2}a_k$, then $N(a_j) \geqq N(a_k) + 1$ unless $a_j = 2^{m-1}a_{k+1}$.*

Proof. Suppose that there are no small steps from a_k to a_j. Assume that there is at least one t such that $2 \leqq t \leqq m$ and $a_{k+t} \neq 2a_{k+t-1}$. Then $a_{k+t} \leqq a_{k+t-1} + a_{k+t-2} \leqq 2^{t-1}a_k + 2^{t-2}a_k$ which implies that $a_{k+m} = a_{k+t+(m-t)} \leqq 2^{m-t}a_{k+t} \leqq 2^{m-t}(2^{t-1}a_k + 2^{t-2}a_k) = 2^{m-1}a_k + 2^{m-2}a_k < a_j$. Thus, $a_{k+m} < a_j$ which implies that there is at least one small step from a_k to a_j which is a contradiction. Therefore, if there are no small steps from a_k to a_j, then $a_{k+t} = 2a_{k+t-1}$ for $2 \leqq t \leqq m$ which implies that $a_j = 2^{m-1}a_{k+1}$. It follows that if $a_j \neq 2^{m-1}a_{k+1}$, then $N(a_j) \geqq N(a_k) + 1$.

Knuth's Theorem C [8] along with the four previous lemmas will be much used in the work that follows. The statement of Theorem

C follows with the integers being expressed in binary form.

THEOREM C. *If $\nu(n) \geqq 4$, then $l(n) \geqq \lambda(n) + 3$ except when $\nu(n) = 4$ and n has one of the four following forms*: (A) $n = 1\cdots d\cdots 1\cdots 1\cdots d\cdots 1\cdots$ *where d indicates the number of zeros between the first and second one and between the third and fourth one.* (B) $n = 1\cdots d\cdots 1\cdots 1\cdots e\cdots 1\cdots$ *where d and e again indicate zeros and $e = d - 1$.* (C) $n = 1001\cdots 11\cdots$. (D) $n = 10000111\cdots$. *In these four cases $l(n) = \lambda(n) + 2$.*

The $m = 3$ case of the conjecture will now be stated as a theorem, and the method of proof will be described.

THEOREM 1. *If $\nu(n) \geqq 9$, then $l(n) \geqq \lambda(n) + 4$.*

Proof. Let $1 = a_0 < a_1 < \cdots < a_r = n$ be an addition chain for an integer n for which $\nu(n) \geqq 9$. Let a_i denote the first member of the chain for which $\nu(a_i) \geqq 9$. Then $a_i = a_j + a_k$ where $k < j$ since if $k = j$, then $a_i = 2a_j$ which would mean that $\nu(a_i) = \nu(a_j)$. Thus, a_j and a_k are distinct members of the chain, and since $\nu(a_j) \leqq 8$ and $\nu(a_k) \leqq 8$, it follows from Lemma 1 that $9 \leqq \nu(a_i) \leqq 16$. Each of the eight cases for $\nu(a_i)$ must be considered, and for each of these cases the possibilities for $\nu(a_j)$ and $\nu(a_k)$ must be considered. For convenience the various cases will be listed as ordered triples $(\nu(a_i), \nu(a_j), \nu(a_k))$. There are 120 cases altogether. The case (9, 5, 4) will be considered first.

By Lemma 1 $c = 0$ for (9, 5, 4), and the only possibility for a_j/a_k is:

$$\begin{aligned} a_j &= 1\cdots\cdots \\ +a_k &= \cdots 1\cdots \\ a_i &= 1\cdots\cdots\,. \end{aligned}$$

As can be seen $\lambda(a_i) = \lambda(a_j)$ and, thus, there is at least one small step from a_j to a_i. Case $m = 2$ of the conjecture implies that $N(a_j) \geqq 3$ since $\nu(a_j) = 5$. Thus, $N(n) \geqq N(a_i) \geqq N(a_j) + 1 \geqq 4$.

Case (9, 4, 5) is virtually the same as (9, 5, 4) except that it is $N(a_k)$ which is greater than or equal to 3. Since $N(a_j) \geqq N(a_k)$, it follows as before that $N(n) \geqq 4$.

The 34 additional cases for which $c = 0$ are handled in the same manner as these cases.

For $c = 1$ there are 28 cases for $(\nu(a_i), \nu(a_j), \nu(a_k))$. Since $a_i \leqq 2a_j$, either $\lambda(a_i) = \lambda(a_j)$ or $\lambda(a_i) = \lambda(a_j) + 1$. If $\lambda(a_i) = \lambda(a_j)$, then as in the cases where $c = 0$ it may be concluded that $N(n) \geqq 4$. If $\lambda(a_i) = \lambda(a_j) + 1$, then with $c = 1$ the only possibility for a_j/a_k is:

$$\begin{aligned} a_j &= 1\cdots \\ +a_k &= \ 1\cdots \\ a_i &= 10\cdots\,. \end{aligned}$$

As previously noted a_j and a_k are distinct members of the chain, and since $\lambda(a_j) = \lambda(a_k)$ it follows that $N(a_j) \geqq N(a_k) + 1$. For those cases where $\nu(a_k) \geqq 5$, $N(n) \geqq N(a_j) \geqq N(a_k) + 1 \geqq 4$. When $\nu(a_k) \leqq 4$, some further work is necessary.

The cases where $3 \leqq \nu(a_k) \leqq 4$ shall first be considered. By Lemma 1 $\nu(a_j) \geqq 6$ since $c = 1$. $a_j \neq 2a_k$ since $\nu(a_j) \neq \nu(a_k)$, and it follows that either $a_j = a_m + a_s$ where $s \leqq m$ and $a_m \neq a_k$ or $a_j = a_k + a_t$ where $t < k$. Suppose $a_j = a_m + a_s$ where $a_m \neq a_k$. Since $a_j \leqq 2a_m$, the possibilities on the number line are:

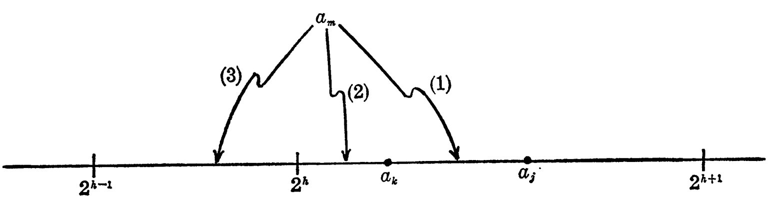

FIGURE 1

In case (1) $N(a_j) \geqq N(a_k) + 2 \geqq 4$ since $\nu(a_k) \geqq 3$. In case (2) $N(a_m) \geqq 2$ for if $N(a_m) \leqq 1$, then $1 = a_0 < a_1 < \cdots < a_m < a_j$ is an addition chain for a_j with less than three small steps contradicting the fact that $\nu(a_j) \geqq 5$ implies $N(a_j) \geqq 3$. Thus, $N(a_j) \geqq N(a_m) + 2 \geqq 4$. In case (3) similar reasoning shows that $N(a_m) \geqq 3$, and, consequently, $N(a_j) \geqq N(a_m) + 1 \geqq 4$. In all three cases $N(n) \geqq N(a_j) \geqq 4$.

Suppose $a_j = a_k + a_t$ where $t < k < j$. Then $a_t = a_j - a_k$. Since $c = 1$ there is only one 1/1 slot in a_j/a_k. When a_j/a_k is considered from a subtraction point of view, it follows from Lemma 2 that $\nu(a_t) \geqq 5$ which means that $N(a_t) \geqq 3$. Thus, $N(n) \geqq N(a_j) \geqq N(a_k) + 1 \geqq N(a_t) + 1 \geqq 4$.

All cases for $c = 1$ have been dispensed with except (9, 8, 2). In this case $\nu(a_k) = 2$ implies $N(a_k) \geqq 1$. If $N(a_k) = 1$, then it may be concluded that all members of the chain preceding a_k have two or less ones in their binary representation. Thus, $\nu(a_{k+1}) \leqq 4$ and $\nu(a_{k+2)}) \leqq 6$. Since $\lambda(a_j) = \lambda(a_k)$, this means that $N(n) \geqq N(a_j) \geqq N(a_k) + 3 \geqq 4$. If $N(a_k) \geqq 2$, then $N(n) \geqq 4$ in the same manner as when $3 \leqq \nu(a_k) \leqq 4$.

For $c = 2$ the cases where $\nu(a_j) \geqq 5$, $\nu(a_k) \geqq 5$, and $\nu(a_j) \neq \nu(a_k)$ are handled rather easily. As with the $c = 1$ cases it may be supposed that $\lambda(a_i) = \lambda(a_j) + 1$. If $\lambda(a_j) = \lambda(a_k)$, then $N(n) \geqq N(a_j) \geqq N(a_k) + 1 \geqq 4$. Thus, it may be supposed that $\lambda(a_j) > \lambda(a_k)$, and the only possibility for a_j/a_k with $c = 2$ is:

$$\begin{aligned} a_j &= 11\cdots \\ +\,a_k &= 1\cdots \\ a_i &= 100\cdots . \end{aligned}$$

If $a_j = a_m + a_s$ where $s \leqq m < j$ and $a_m \neq a_k$, then there are three possibilities on the number line:

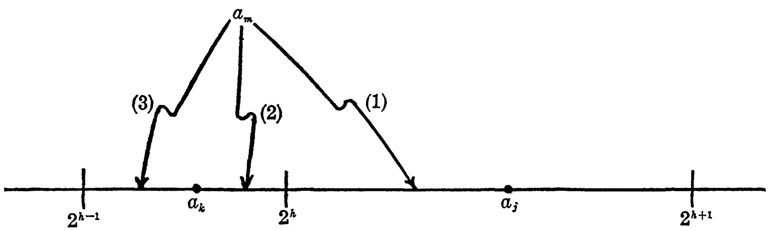

FIGURE 2

In cases (1) and (2) $N(n) \geqq N(a_j) \geqq N(a_k) + 1 \geqq 4$ since $\nu(a_k) \geqq 5$. In case (3) $N(a_m) \geqq 3$ or else $1 = a_0 < a_1 < \cdots < a_m < a_j$ is a chain for a_j with less than three small steps which contradicts $\nu(a_j) \geqq 5$. Thus, $N(n) \geqq N(a_j) \geqq N(a_m) + 1 \geqq 4$. If $a_j = a_k + a_t$, then $a_t = a_j - a_k$. Since $c = 2$, there can be no more 1/1 slots in a_j/a_k, and since $\nu(a_j) \neq \nu(a_k)$, $a_j \neq 2a_k$ which means that a_k and a_t are distinct members of the chain. a_j/a_k then looks as follows:

$$\begin{aligned} a_j &= 11\cdots \\ -\,a_k &= 1\cdots \\ a_t &= 1\cdots . \end{aligned}$$

By Lemma 2 $\nu(a_t) \geqq 5$ since $\nu(a_j) \geqq 5$. Since $\lambda(a_k) = \lambda(a_t)$, $N(n) \geqq N(a_k) \geqq N(a_t) + 1 \geqq 4$.

There are 12 cases for which $c = 2$, $\nu(a_j) \geqq 5$, $\nu(a_k) \geqq 5$, and $\nu(a_j) \neq \nu(a_k)$. Thus, 76 of the 120 cases for $(\nu(a_i), \nu(a_j), \nu(a_k))$ have been dispensed with so far. In (10, 6, 6), (12, 7, 7), and (14, 8, 8) $\nu(a_j) = \nu(a_k)$, and it is possible that $a_j = 2a_k$. This means that $a_k = a_t$; hence, a_k and a_t are not distinct members of the chain. Thus, the statement that $N(a_k) \geqq N(a_t) + 1$ cannot be made as with the other cases where $c = 2$ and $\nu(a_j) \geqq 5$ and $\nu(a_k) \geqq 5$. Some additional concepts need to be discussed at this point which make it possible to dispense with cases such as these.

Let $l_8(n)$ denote the minimal length of an addition chain for an integer n all of whose members have eight or less ones in their binary representation. A list of propositions concerning $l_8(n)$ will now be given. The proof of one of these propositions will then be given. The proofs of the others are similar.

PROPOSITION 1. *If* $\nu(n) = 7$ *and* $n = 111\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 2. *If* $\nu(n) = 8$ *and* $n = 111\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$ *unless* $n = 1111\cdots 1111\cdots$.

PROPOSITION 3. *If* $\nu(n) = 7$ *and* $n = 110\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$ *unless* $n = 11001\cdots 1111\cdots$.

PROPOSITION 4. *If* $\nu(n) = 8$ *and* $n = 110\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$ *unless*

$$n = 11\cdots d\cdots 11\cdots 11\cdots e\cdots 11\cdots \quad \textit{where } e = d \textit{ or } e = d - 1\,.$$

(Note: The d and e again stand for d and e zeros respectively between the ones.)

PROPOSITION 5. *If* $\nu(n) = 6$ *and* $n = 111\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$ *unless* $n = 111\cdots 111\cdots$, $111001011\cdots$, $1111\cdots 1001\cdots$, $1111\cdots 101\cdots$, *or* $1111\cdots 11\cdots$.

PROPOSITION 6. *If* $\nu(n) = 7$ *and* $n = 10111\cdots 01\cdots 01\cdots 01\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 7. *If* $\nu(n) = 8$ *and* $n = 1011111\cdots 01\cdots 01\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 8. *If* $\nu(n) = 8$ *and*

$$\begin{aligned} n = 10111\cdots &01\cdots 01\cdots 0011\cdots\,, \\ &01\cdots 0011\cdots 01\cdots\,, \\ &0011\cdots 01\cdots 01\cdots\,, \end{aligned}$$

then $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 9. *If* $\nu(n) = 8$ *and*

$$\begin{aligned} n = 1011\cdots &01\cdots 01\cdots 00111\cdots\,, \\ &01\cdots 00111\cdots 01\cdots\,, \\ &00111\cdots 01\cdots 01\cdots\,, \end{aligned}$$

then $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 10. *If* $\nu(n) = 8$ *and* $n = 1010111\cdots 01\cdots 01\cdots 01\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 11. *If* $\nu(n) = 8$ *and* $n = 1011011\cdots 01\cdots 01\cdots 01$, *then* $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 12. *If* $\nu(n) = 6$ *and* $n = 11\cdots01\cdots01\cdots01\cdots01\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 13. *If* $\nu(n) = 7$ *and* $n = 1011\cdots01\cdots01\cdots01\cdots01\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 14. *If* $\nu(n) = 8$ *and* $n = 101111\cdots01\cdots01\cdots01\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 15. *If* $\nu(n)=8$ *and* $n=101011\cdots01\cdots01\cdots01\cdots01\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 16. *If* $\nu(n) = 8$ *and*

$$\begin{aligned} n = 1011\cdots01\cdots01\cdots01\cdots0011\cdots\ , \\ 01\cdots01\cdots0011\cdots01\cdots\ , \\ 01\cdots0011\cdots01\cdots01\cdots\ , \\ 0011\cdots01\cdots01\cdots01\cdots\ , \end{aligned}$$

then $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 17. *If* $\nu(n) = 8$ *and* $n = 10111\cdots01\cdots01\cdots01\cdots01\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 18. *If* $\nu(n) = 8$ *and* $n = 1011\cdots01\cdots01\cdots01\cdots01\cdots01\cdots$, *then* $l_8(n) \geqq \lambda(n) + 4$.

PROPOSITION 19. *If* $\nu(n) = 7$ *and* $n = 1011100\cdots111$, *then* $l_8(n) \geqq \lambda(n) + 4$.

Proof. (Prop. 1) Let $1 = a_0 < a_1 < \cdots < a_r = n$ be an addition chain for n where $\nu(n) = 7$ and $n = 111\cdots$. It shall be assumed that all members of the chain have eight or less ones in their binary representation. Let a_i denote the first member of the chain for which $\nu(a_i) = 7$ and $a_i = 111\cdots$. $a_i = a_j + a_k$ for some $k \leqq j < i$. In fact $k < j$ for if $a_j = a_k$ then $a_i = 2a_j$ which would mean that $\nu(a_j) = 7$ and $a_j = 111\cdots$ contradicting the fact that a_i was chosen as the first member of the chain having these properties. Thus, a_j and a_k are distinct members of the chain and $1 \leqq \nu(a_j), \nu(a_k) \leqq 8$. The 49 cases for $(\nu(a_j), \nu(a_k))$ must be considered.

$a_i \leqq 2a_j$ implies that $\lambda(a_i) = \lambda(a_j)$ or $\lambda(a_i) = \lambda(a_j) + 1$. If $\nu(a_k) \geqq 5$, it may be assumed that $\lambda(a_i) = \lambda(a_j) + 1$; otherwise, $N(n) \geqq N(a_i) \geqq N(a_j) + 1 \geqq N(a_k) + 1 \geqq 4$. However, if $\lambda(a_i) = \lambda(a_j) + 1$, the only way to obtain $a_i = 111\cdots$ is if a_j/a_k is as follows:

$$\begin{aligned} a_j &= \ 111\cdots \\ +a_k &= \ \ 11\cdots\cdot \\ a_i &= 111\cdots\cdot\ . \end{aligned}$$

The arrows indicate that at least three carries are needed with this configuration. As can be seen $\lambda(a_j) = \lambda(a_k)$, and it follows that $N(n) \geqq N(a_j) \geqq N(a_k) + 1 \geqq 4$ for all cases where $\nu(a_k) \geqq 5$. If $\nu(a_k) \leqq 4$ and $\nu(a_j) \geqq 5$, then the configuration still holds, and all cases where $\nu(a_j) = 7$ may be dispensed with since $a_j = 111\cdots$ again contradicts the "firstness" of a_i. The cases (8, 1), (6, 3), (6, 2), (6, 1), (5, 4), (5, 3), and (5, 2) all have less than three carries in $a_j + a_k$ by Lemma 1 while at least three carries are needed in the configuration. In case (8, 2) only two carries are possible while three are needed. In (8, 3) it may be assumed as with case (10, 8, 3) of Theorem 1 that $a_j = a_k + a_t$ (see Figure 1). $a_t = a_j - a_k$, and by Lemma 2 $\nu(a_t) \geqq 5$ which implies that $N(n) \geqq N(a_j) \geqq N(a_k) + 1 \geqq N(a_t) + 1 \geqq 4$. In (8, 4) it may be assumed that a_k is one of the four special types in Theorem C; otherwise, $N(a_k) \geqq 3$ which implies $N(n) \geqq N(a_j) \geqq N(a_k) + 1 \geqq 4$. Since $a_k = 11\cdots$, this means that $a_k = 11\cdots11\cdots$. As in (8, 3) it may be assumed that $a_j = a_k + a_t$, and as with (8, 3) $\nu(a_t) \geqq 5$ unless there are four 1/1 slots in a_j/a_k. By Lemma 1 $c = 5$ in $a_j + a_k$, and the only way to meet all of these requirements is if a_j/a_k is as follows:

$$\begin{aligned} a_j &= \ \ 11111\cdots1\cdots1\cdots1\cdots \\ +\ a_k &= \ \ 11011\cdots0\cdots0\cdots0\cdots \\ a_i &= 111010\cdots1\cdots1\cdots1\cdots \end{aligned} \quad \text{implies} \quad \begin{aligned} a_j &= 11111\cdots1\cdots1\cdots1\cdots \\ -\ a_k &= 11011\cdots0\cdots0\cdots0\cdots \\ a_t &= \quad\ 100\cdots1\cdots1\cdots1\cdots\ . \end{aligned}$$

$\lambda(a_k) = \lambda(a_t) + 2$ while $2^2 a_t < a_k$, and so by Lemma 3 $N(a_k) \geqq N(a_t) + 1 \geqq 3$. Thus, $N(n) \geqq N(a_j) \geqq N(a_k) + 1 \geqq 4$. In (6, 4) $c = 3$ by Lemma 1. Therefore, a_j/a_k must be:

$$\begin{aligned} a_j &= \ \ 111\cdots1\cdots0\cdots \\ +a_k &= \ \ 111\cdots0\cdots1\cdots \\ a_i &= 1110\cdots1\cdots1\cdots\ . \end{aligned}$$

By Theorem C $N(a_k) \geqq 3$; hence, $N(n) \geqq N(a_j) \geqq N(a_k) + 1 \geqq 4$.

The only remaining cases to be considered are (4, 4), (4, 3), and (3,4). $\lambda(a_i) = \lambda(a_j) + 1$ is not possible since at least three carries are needed while these cases by Lemma 1 have less than two. When either $\nu(a_j) = 4$ or $\nu(a_k) = 4$, it may be assumed that a_j and a_k are what shall be called "special fours" meaning that they are one of the types in Theorem C. Otherwise, $N(n) \geqq N(a_i) \geqq N(a_j) + 1 \geqq 4$ since it may be assumed that $\lambda(a_i) = \lambda(a_j)$. In (4, 4) the possible configurations a_j/a_k for obtaining $a_i = 111\cdots$ with $c = 1$ are:

$$\begin{array}{lrl@{\qquad}lrl}
(1) & a_j = & 1001\cdots & (2) & a_j = & 101\cdots01\cdots \\
 & +a_k = & 101\cdots & & +a_k = & 10\cdots01\cdots \\
 & a_i = & 1110\cdots & & a_i = & 111\cdots10\cdots \\
(3) & a_j = & 100\cdots01\cdots & (4) & a_j = & 11\cdots \\
 & +a_k = & 11\cdots01\cdots & & +a_k = & 00\cdots \\
 & a_i = & 111\cdots10\cdots & & a_i = & 111\cdots\,.
\end{array}$$

In (1), (2), and (3) either $a_j = a_m + a_s$ where $a_m \neq a_k$ or $a_j = a_k + a_t$. If $a_j = a_m + a_s$, then $N(a_j) \geqq 3$ by reasoning similar to that used in (9, 6, 5) of Theorem 1 (see Figure 2). Thus, $N(n) \geqq N(a_i) \geqq N(a_j) + 1 \geqq 4$. It shall be assumed then that $a_j = a_k + a_t$. In (1) there are two possibilities for $a_t = a_j - a_k$:

$$\begin{array}{lrl@{\qquad}lrl}
(a) & a_j = & 1001\cdots & (b) & a_j = & 1001\cdots \\
 & -a_k = & 101\cdots & & -a_k = & 101\cdots \\
 & a_t = & 100\cdots & & a_t = & 11\cdots\,.
\end{array}$$

Since $c = 1$, there can be no further 1/1 slots in a_j/a_k. Thus, in (a) $\nu(a_t) \geqq 3$ by Lemma 2, and since $\lambda(a_k) = \lambda(a_t)$ and $a_t \neq a_k$, this means $N(n) \geqq N(a_i) \geqq N(a_j) + 1 \geqq N(a_k) + 1 \geqq N(a_t) + 2 \geqq 4$. In (b) $\nu(a_t) \geqq 5$ by Lemma 2, and, so, $N(n) \geqq N(a_i) \geqq N(a_j) + 1 \geqq N(a_t) + 1 \geqq 4$. (2) may be dispensed with in the same manner as (1) part (a) while in (3) since a_k is a "special four" a_j/a_k becomes:

$$\begin{array}{rl}
a_j = & 100\cdots010\cdots \\
-a_k = & 11\cdots011\cdots \\
a_t = & \cdots\cdots111\cdots\,.
\end{array}$$

By Lemma 2 $\nu(a_t) \geqq 5$; hence, $N(n) \geqq 4$ as in (1) part (b).

In (4) it may be assumed that the first two digits in a_k are ones; otherwise, $\lambda(a_j) = \lambda(a_k) + m$ for some positive integer m while $2^m a_k < a_j$. By Lemma 3 this would mean $N(a_j) \geqq N(a_k) + 1 \geqq 3$, and, hence, $N(n) \geqq 4$. Since a_j and a_k both start with two ones and are "special fours", they must both have the form $11\cdots11\cdots$, but in this event it is not possible to have $c = 1$ in $a_j + a_k$.

In (4, 3) and (3, 4) $c = 0$ which means that there are no 1/1 slots in a_j/a_k. The possibilities for a_j/a_k are the following:

$$\begin{array}{lrl@{\quad}lrl@{\quad}lrl@{\quad}lrl}
(1) & a_j = & 101\cdots & (2) & a_j = & 100\cdots & (3) & a_j = & 110\cdots & (4) & a_j = & 111\cdots \\
 & +a_k = & 10\cdots & & +a_k = & 11\cdots & & +a_k = & 1\cdots & & +a_k = & 000\cdots \\
 & a_i = & 111\cdots & & a_i = & 111\cdots & & a_i = & 111\cdots & & a_i = & 111\cdots\,.
\end{array}$$

In (1) $N(n) \geqq 4$ for both (4, 3) and (3, 4) by the same reasoning used in (4, 4) with configuration (1) part (a). The remaining configurations will now be discussed for (4,3).

In (2) $a_t = a_j - a_k$ and by Lemma 2 $\nu(a_t) \geqq 4$. Thus, $N(a_t) \geqq 3$ which implies $N(n) \geqq 4$ unless a_t is a "special four". Since $a_k = 11\cdots$, it may be assumed that a_t also starts with two ones by the same reasoning that was used for a_k in (4, 4) configuration (4). Thus, $a_t = 11\cdots11\cdots$. Since there can be no ones under a 0/0 slot in a_j/a_k (otherwise $\nu(a_t) \geqq 5$), there are only two possibilities for a_j/a_k:

$$\begin{array}{rrl} \text{(a)} & a_j = & 1001\cdots101\cdots \\ & -a_k = & \ 110\cdots010\cdots \\ & a_t = & \ \ 11\cdots011\cdots \end{array} \qquad \begin{array}{rrl} \text{(b)} & a_j = & 1000111\cdots \\ & -a_k = & \ 111000\cdots \\ & a_t = & \ \ \ \ 1111\cdots . \end{array}$$

In (a) $N(a_k) \geqq 3$ by arguments used before unless $a_k = a_t + a_u$ for some $u \leqq t < k$. If a_k/a_t is examined, it may be seen that $\nu(a_u) \geqq 4$, $\lambda(a_t) = \lambda(a_u)$ and $a_u \neq a_t$. Thus, $N(a_t) \geqq N(a_u) + 1 \geqq 3$ which implies $N(n) \geqq 4$. In (b) a_j is not a "special four" and, so, $N(n) \geqq N(a_i) \geqq N(a_j) + 1 \geqq 4$.

In (3) $a_j = 11\cdots11\cdots$ since a_j is a "special four". As in configuration (4) of (4, 4) it may be assumed that a_k starts with two ones. a_j/a_k is then:

$$\begin{aligned} a_j &= 1100\cdots11\cdots \\ +a_k &= \ \ \ 11\cdots00\cdots \\ a_i &= 1111\cdots11\cdots . \end{aligned}$$

As can be seen $a_j > 2a_k + a_k$, and, so, by Lemma 4 $N(a_j) \geqq N(a_k) + 1 \geqq 3$ unless $a_j = 2a_{k+1}$. Since $\nu(a_{k+1}) = 4$ and $\lambda(a_{k+1}) = \lambda(a_k) + 1$, it follows as before that $N(a_{k+1}) \geqq 3$ unless $a_{k+1} = a_k + a_t$ for some $t \leqq k$. From a_j/a_k and the fact that $a_j = 2a_{k+1}$ it may be determined that a_{k+1}/a_k is as follows:

$$\begin{aligned} a_{k+1} &= 1100\cdots11\cdots \\ -a_k &= \ 11\cdots00\cdots\cdot \\ a_t &= \ 1\cdots\cdots1\cdots\cdot . \end{aligned}$$

By Lemma 2 $\nu(a_t) \geqq 3$. Thus, $N(n) \geqq N(a_i) \geqq N(a_j) + 1 \geqq N(a_k) + 1 \geqq N(a_t) + 2 \geqq 4$.

In (4) $a_j = 1111\cdots$ since a_j is a "special four", and since $\nu(a_k) = 3$, it follows that $\lambda(a_j) = \lambda(a_k) + m$ for some positive integer m while $2^m a_k < a_j$. By Lemma 3 $N(a_j) \geqq N(a_k) + 1 \geqq 3$ which implies $N(n) \geqq 4$. Configurations (2), (3), and (4) will now be discussed for (3, 4).

In (2) it may again be assumed that $a_j = a_k + a_t$, and $a_k = 11\cdots11\cdots$ since a_k is a "special four". By Lemma 2 $\nu(a_t) \geqq 3$, and a one can occur in a_t at most once under a 0/0 slot in a_j/a_k or else

$\nu(a_t) \geqq 5$. The possibilities for a_j/a_k are:

$$\begin{array}{rrlrrl}
\text{(a)} & a_j = & 10000\cdots & \text{(b)} & a_j = & 100\cdots100\cdots \\
 & -a_k = & \ 1111\cdots & & -a_k = & \ 11\cdots011\cdots \\
 & a_t = & \quad\ \ 1\cdots & & a_t = & \ \ 1\cdots001\cdots \\
\text{(c)} & a_j = & 100000\cdots & \text{(d)} & a_j = & 100\cdots1000\cdots \\
 & -a_k = & \ 11011\cdots & & -a_k = & \ 11\cdots0011\cdots \\
 & a_t = & \quad\ 101\cdots & & a_t = & \ \ 1\cdots0101\cdots .
\end{array}$$

In (a) and (b) $\nu(a_t) = 3$, and no matter where the remaining ones in a_t are placed the conditions of Lemma 3 will apply. In (d) $\nu(a_t) = 4$, and, so it may be assumed that a_t is a "special four" in which case a_t must start as $a_t = 10\cdots$. Thus, the conditions of Lemma 3 also apply to (c) and (d), and in all four cases $N(a_k) \geqq N(a_t) + 1 \geqq 3$ which implies that $N(n) \geqq 4$.

In (3) it may again be assumed as in configuration (4) of (4, 4) that the first two digits of a_k are ones, and since a_k is a "special four", this means that $a_k = 11\cdots11\cdots$. As in (4, 3) configuration (3) it may also be assumed that $a_j = 2a_{k+1}$ and that $a_{k+1} = a_k + a_t$ for some $t \leqq k$. These facts together with a_j/a_k determine a_{k+1}/a_k:

$$\begin{array}{rlcrl}
a_j = & 1100\cdots00\cdots & \text{implies} & a_{k+1} = & 1100\cdots00\cdots \\
+a_k = & \quad 11\cdots11\cdots & & -a_k = & \ 11\cdots110\cdots \\
a_i = & 1111\cdots11\cdots & & a_t = & \ 1\cdots\cdots10\cdots .
\end{array}$$

No matter where the other one in a_{k+1} is placed, it can be seen that $\nu(a_t) \geqq 3$, $\lambda(a_k) = \lambda(a_t)$ and $a_t \neq a_k$. Thus, $N(a_k) \geqq N(a_t) + 1 \geqq 3$ which implies $N(n) \geqq 4$.

In (4) a_k is a "special four", and the conditions of Lemma 3 will apply unless $a_k = 111\cdots$. $\lambda(a_j) = \lambda(a_k) + m$ for some $m \geqq 2$ while $a_j > 2^{m-1}a_k + 2^{m-2}a_k$, and, so, by Lemma 4 $N(a_j) \geqq N(a_k) + 1 \geqq 3$ unless $a_j = 2^{m-1}a_{k+1}$. As before it may be assumed that $a_{k+1} = a_k + a_t$ for some $t \leqq k$, and these facts together with a_j/a_k determine a_{k+1}/a_k:

$$\begin{array}{rlcrl}
a_j = & 111\cdots0000\cdots & \text{implies} & a_{k+1} = & 11100\cdots \\
+a_k = & \qquad\quad 1111\cdots & & -a_k = & \ 1111\cdots \\
a_i = & 111\cdots1111\cdots & & a_t = & \ 1101\cdots
\end{array}$$

$N(a_k) \geqq N(a_t) + 1 \geqq 3$; hence, $N(n) \geqq 4$.

In all 49 cases it has been shown that $N(n) \geqq 4$, and, so, it may be concluded that if $\nu(a_i) = 7$ and $a_i = 111\cdots$, then $l_8(n) \geqq \lambda(n) + 4$.

In Proposition 2 a_i denotes the first member of the chain for which $\nu(a_i) = 8$, $a_i = 111\cdots$ but $a_i \neq 1111\cdots1111\cdots$. The proof is then carried out in the same manner as the proof of Proposition 1. The proofs of the remaining propositions are similar, and as each one is proved it may be used in the proof of the next one. Propositions 1 to 5 are extremely helpful in the proofs of the remaining propositions and in that part of the proof of Theorem 1 that remains. We shall now return to the proof of Theorem 1 to demonstrate how the propositions are used. As an example of the remaining cases (9, 7, 7) will be examined.

To recall a_i is the first member of an addition chain for n for which $\nu(a_i) \geqq 9$. $a_i = a_j + a_k$ where $\nu(a_j) \leqq 8$ and $\nu(a_k) \leqq 8$. The propositions concerning $l_8(n)$ are applicable to a_j and a_k and all other members of the chain preceding a_i. As in (9, 6, 5) it may be assumed in (9, 7, 7) that $\lambda(a_i) = \lambda(a_j) + 1$ and $\lambda(a_j) > \lambda(a_k)$. Also if $\lambda(a_j) = \lambda(a_k) + m$, it may be assumed that $a_j \geqq 2^m a_k$ or else by Lemma 3 $N(a_j) \geqq N(a_k) + 1 \geqq 4$ which implies $N(n) \geqq 4$. In (9, 7, 7) $c = 5$, and the possibilities for a_j/a_k are now listed. These possibilities are the ways to proceed from left to right to the first 1/1 slot in a_j/a_k without exceeding five carries and with the previously mentioned restrictions kept in mind.

$$
\begin{array}{llll}
(1) & a_j = 11\cdots\cdot & (2) & a_j = 101\cdots \\
& +a_k = 11\cdots & & +a_k = 11\cdots \\
& a_i = 10\cdots\cdots & & a_i = 100\cdots\cdot
\end{array}
\qquad
\begin{array}{ll}
(3) & a_j = 111\cdots\cdots \\
& +a_k = 111\cdots \\
& a_i = 100\cdots\cdots\cdot
\end{array}
$$

$$
\begin{array}{llllll}
(4) & a_j = 1011\cdots\cdot & (5) & a_j = 1001\cdots & (6) & a_j = 1101\cdots \\
& +a_k = 1011\cdots & & +a_k = 111\cdots & & +a_k = 11\cdots \\
& a_i = 1000\cdots\cdots & & a_i = 1000\cdots\cdot & & a_i = 1000\cdots\cdot
\end{array}
$$

$$
\begin{array}{llllll}
(7) & a_j = 1111\cdots\cdots\cdot & (8) & a_j = 10101\cdots & (9) & a_j = 10011\cdots \\
& +a_k = 1111\cdots & & +a_k = 1011\cdots & & +a_k = 1101\cdots \\
& a_i = 1000\cdots\cdots\cdots & & a_i = 100000\cdots & & a_i = 100000\cdots
\end{array}
$$

$$
\begin{array}{llllll}
(10) & a_j = 10001\cdots & (11) & a_j = 11001\cdots & (12) & a_j = 111010\cdots \\
& +a_k = 1111\cdots & & +a_k = 111\cdots & & +a_k = 111\cdots \\
& a_i = 100000\cdots & & a_i = 100000\cdots & & a_i = 1000001\cdots
\end{array}
$$

$$
\begin{array}{ll}
(13) & a_j = 111110000\cdots \\
& +a_k = 11111\cdots \\
& a_i = 1000001111\cdots\ .
\end{array}
$$

In configurations (3), (5), (6), (7), (9), (10), (11), (12), and (13) Propositions 1 and 3 imply that either $N(a_j) \geqq 4$ or $N(a_k) \geqq 4$. In either event this means that $N(n) \geqq N(a_j) \geqq 4$. In (1) $N(n) \geqq 4$ in the same manner unless a_j and a_k both have the binary form $11001\cdots1111\cdots$, but in this event it is impossible to arrange a_j/a_k so that $c = 5$. In (2) it may be assumed that $a_k = 11001\cdots1111\cdots$ and that $a_j = a_k + a_t$ for some $t \leqq k$ (see Figure 2). Since $c = 5$ there can be at most two more 1/1 slots in a_j/a_k. There are two possibilities for a_j/a_k:

$$\begin{array}{lrl} \text{(a)} & a_j = & 101\cdots\cdots\cdots 00\cdots \\ & -a_k = & 11001\cdots 1111\cdots \\ & a_t = & 10\cdots\cdots\cdots 01\cdots \end{array} \qquad \begin{array}{lrl} \text{(b)} & a_j = & 101\cdots\cdots\cdots 00\cdots \\ & -a_k = & 11001\cdots 1111\cdots \\ & a_t = & 1\cdots\cdots\cdots 01\cdots . \end{array}$$

In (a) it is impossible to have two further 1/1 slots in a_j/a_k with zeros under them. Thus, $\nu(a_t) \geqq 5$ by Lemma 2, and since $\lambda(a_k) = \lambda(a_t)$ and $a_t \neq a_k$, $N(n) \geqq N(a_k) \geqq N(a_t) + 1 \geqq 4$. Configuration (b) can be filled out a little further by realizing that the 1 can occur under the 1/1 slot only if a_j/a_k is as follows:

$$\begin{aligned} a_j &= 10100\cdots\cdots 00\cdots \\ -a_k &= 11001\cdots 1111\cdots \\ a_t &= 111\cdots\cdots 01\cdots . \end{aligned}$$

It is impossible to have zeros in a_t under any further 1/1 slots in a_j/a_k, and, so, by Lemma 2 $\nu(a_t) \geqq 9$ which contradicts the fact that a_i is the first member of the chain for which $\nu(a_i) \geqq 9$. In (4) it may again be assumed that $a_j = a_k + a_t$ for some $t \leqq k$. $c = 5$ implies that there is one more 1/1 slot in a_j/a_k; hence, $\nu(a_t) \geqq 5$ by Lemma 2. It is evident that $\lambda(a_k) = \lambda(a_t)$, and if $a_t \neq a_k$, $N(n) \geqq N(a_k) \geqq N(a_t) + 1 \geqq 4$. It is possible in this case, however, that $a_t = a_k$ which means $a_j = 2a_k$. With $c = 5$ the configuration would be:

$$\begin{aligned} a_j &= 101110\cdots 10\cdots 10\cdots 10\cdots \\ +a_k &= 101110\cdots 1\cdots 01\cdots 01\cdots \\ a_i &= 1000101\cdots 11\cdots 11\cdots 11\cdots . \end{aligned}$$

By Proposition 6 $N(n) \geqq N(a_j) \geqq 4$. In (8) it is not possible that $a_j = 2a_k$, and, so, $N(n) \geqq 4$ as in (4) when $a_t \neq a_k$. This concludes the proof of (9, 7, 7).

The proof of the remaining cases is similar. Once Theorem 1 is established it follows that the propositions concerning $l_8(n)$ are true in general. That is $l(n)$ may be used in the statements of all of the propositions instead of $l_8(n)$. The reason for this is that if an integer with more than eight ones in its binary representation does occur in one of the chains then by Theorem 1 there are at least four small steps in the chain up to that integer which means that $N(n) \geqq 4$. In particular Proposition 19 may be restated to say that if $\nu(n) = 7$ and $n = 1011100\cdots 111$, then $l(n) \geqq \lambda(n) + 4$. This leads to the result that there exists an infinite class of integers for which $l(2n) = l(n)$. This is the essence of the following theorem.

THEOREM 2. *If* $n = 2^m(23) + 7$ *where* $m \geqq 5$, *then* $l(2n) = l(n) = m + 8$.

Proof. n has the binary form $n = 1011100\cdots111$, and by the restatement of Proposition 19 $l(n) \geqq \lambda(n) + 4$. On the other hand,

$$1, 2, 3, 4, 7, 14, 21, 23, 2(23), \cdots, 2^m(23), 2^m(23) + 7 = n$$

is a chain for n with only four small steps. Thus, $l(n) = \lambda(n) + 4$.

$2n = 2^{m+1}(23) + 14 = 1011100\cdots1110$. $\nu(2n) = 7$ implies that $l(2n) \geqq \lambda(2n) + 3$ while

$$1, 2, 4, 5, 9, 14, 23, 2(23), \cdots, 2^{m+1}(23), 2^{m+1}(23) + 14 = 2n$$

is a chain for $2n$ with only three small steps. Thus, $l(2n) = \lambda(2n) + 3$.

Since $\lambda(2n) = \lambda(n) + 1 = m + 5$, it follows that $l(2n) = \lambda(2n) + 3 = \lambda(n) + 4 = l(n) = m + 8$.

More details of the proofs of the Propositions and Theorem 1 are available in [12] and in private manuscripts.

References

1. R. E. Bellman, *Advanced problem* 5125, Amer. Math. Monthly, **70** (1963), 765.
2. A. T. Brauer, *On addition chains*, Amer. Math. Soc., **45** (1939), 736–739.
3. P. Erdös, *Remarks on number theory III on addition chains*, Acta Arith. VI (1960), 77-81.
4. A. A. Gioia, M. V. Subbarao, and M. Sugunumma, *The Scholz-Brauer problem in addition chains*, Duke Math. J., **29** (1962), 481–487.
5. W. Hansen, *Zum Scholz-Brauerchen problem*, J. Reine Angew. Math., **202** (1959), 129–136.
6. A. M. Il'in, *On additive number chains*, Problemy Kibernet, **13** (1965), 245–248.
7. H. Kato, *On addition chains*, Ph. D. Dissertation, University of Southern California, June, 1970.
8. D. E. Knuth, *The Art of Computer Programming*, 2 Addison-Wesley, Reading, Mass., (1969), 398–422.
9. A. Scholz, *Jahresbericht*, Deutsche Math.-Verein., **47** (1937), 41.
10. K. B. Stolarsky, *A lower bound for the Scholz-Brauer problem*, Canad. J. Math., **21** (1969), 675–683.
11. E. G. Straus, *Addition chains of vectors*, Amer. Math. Monthly, **71** (1964), 806–808.
12. E. G. Thurber, *The Scholz-Brauer problem on addition chains*, Ph. D. Dissertation, University of Southern California, September, 1971.
13. W. R. Utz, *A note on the Scholz-Brauer problem in addition chains*, Proc. Amer. Math. Soc., **4** (1953), 462–463.
14. C. T. Whyburn, *A note on addition chains*, Proc. Amer. Math. Soc., **16** (1965), 1134.

Received January 1, 1972.

BIOLA COLLEGE

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

SUBMANIFOLDS OF ACYCLIC 3-MANIFOLDS

JOŽE VRABEC

It is proved that, from the viewpoint of "geometric" homology theory, an arbitrary embedding of a closed surface S in any 3-manifold with trivial first homology group looks exactly like the standard embedding of S in the euclidean 3-space. A consequence: every compact subset of a 3-manifold with trivial first homology group can be embedded in a homology 3-sphere. Necessary and sufficient (homological) conditions are given for a compact 3-manifold to be embeddable in some acyclic 3-manifold (or in some homology 3-sphere).

1. Definitions and preliminaries.

Manifolds. We work in the PL category. Each manifold is supposed to have a fixed PL structure. If M is a manifold, then by a submanifold of M or by a surface, simple closed curve, arc, etc., in M we always mean a respective object contained in M as a subpolyhedron (in the chosen PL structure of M). All maps are assumed to be PL. Our manifolds are never automatically assumed to be without boundary, compact, connected, or orientable. However, by a *surface* we mean a compact, connected, orientable 2-manifold. A *cube with n handles* is a 3-manifold homeomorphic to a regular neighborhood of a connected finite linear graph of Euler characteristic $1-n$ in E^3.

We denote the interior of a manifold M by $\operatorname{int} M$ and the boundary by $\operatorname{Bd} M$. However, if M is oriented, then by ∂M we denote the manifold $\operatorname{Bd} M$ oriented coherently with M. The symbol ∂ also denotes the boundary in the homological sense. Let M be an oriented manifold and P a codimension 0 submanifold of M. Whenever we talk of P as an oriented manifold, we assume that P has the orientation inherited from M, unless explicitly stated otherwise. If M is an oriented manifold, then M with the opposite orientation is sometimes denoted by $-M$.

Homology. All homology and cohomology groups, cycles, chains, etc., have integer coefficients. If z_1, z_2 are n-cycles in a space X, then $z_1 \sim z_2$ means "z_1 is homologous to z_2". A compact oriented n-submanifold N of an m-manifold M generates a uniquely determined PL n-chain in M. This chain is a cycle if and only if N is a closed manifold. We shall make no distinction in notation between N and the n-chain it represents. If M is a manifold of dimension at least 2, then every element of $H_1(M)$ can be represented by an oriented closed 1-manifold in int M. If M is a 3-manifold, if $J \subset M$ is a closed oriented 1-manifold, and if $J \sim 0$ in M, then there exists a compact oriented 2-manifold F in M such that $J = \partial F$.

If X, Y are spaces and $f: X \to Y$ a map, then by f_* we denote the homomorphism $H_1(X) \to H_1(Y)$ induced by f.

Let S be an oriented 2-manifold and x, y either two 1-cycles in S or two elements of $H_1(S)$. By $\mathrm{sc}(x, y)$ we denote the (integral) intersection number of x and y. The following is well-known.

LEMMA 1.1. *Let M be an oriented 3-manifold and let J, K be closed oriented 1-manifolds in ∂M. If $J \sim K \sim 0$ in M, then* $\mathrm{sc}(J, K) = 0$.

A polyhedron X is *acyclic* if it is connected and has $H_n(X) = 0$ for $n > 0$. We will call X *1-acyclic* if it is connected and has $H_1(X) = 0$. Note that any 1-acyclic manifold W is orientable. The reason is that $\pi_1(W)$ contains no subgroups of index 2; a subgroup of index 2 would contain the commutator subgroup of $\pi_1(W)$, but the commutator subgroup is the whole $\pi_1(W)$ since $H_1(W) = 0$. A *homology n-sphere* is an n-manifold whose homology is isomorphic to the homology of the n-sphere. An n-manifold will be called *subacyclic* if it can be embedded in an acyclic n-manifold.

2-Manifolds. We give the definition of oriented piping (in dimension 2). Let S be a 2-manifold and $J, K \subset S$ two disjoint oriented simple closed curves. Let $A \subset S$ be an arc from a point $x \in J$ to a point $y \in K$; let int $A \subset \operatorname{int} S - (J \cup K)$. Take a regular neighborhood N of A in S. The intersection $N \cap J$ is a small arc $J_0 \subset J$ containing x in its interior. Similarly, $N \cap K$ is an arc $K_0 \subset K$ with $y \in \operatorname{int} K_0$. Let D be the closure of the component of $N - (J \cup K)$ which contains int A. Then D is a disk and Bd D consists of J_0, K_0, and two "long" arcs in Bd N. Suppose that D can be oriented coherently with both J_0 and K_0. Then the simple closed curve $L = (J \cup K \cup \operatorname{Bd} D) - \operatorname{int}(J_0 \cup K_0)$ can be oriented so that it induces in $J - \operatorname{int} J_0$ the same orientation as J and in $K - \operatorname{int} K_0$ the same orientation as K. If this is the case, we say that the oriented simple closed curve L is obtained from $J \cup K$ by *piping along* A (or that L is obtained by *piping J to K* or by *piping together J and K*). If we think of J, K, L as 1-cycles and of D as a 2-chain, then $J + K - L = \partial D$. Hence $L \sim J + K$. The following lemma is obvious.

LEMMA 1.2. *If S is an oriented surface, then any two components J and K of ∂S can be piped together along any properly embedded arc $A \subset S$ which joins J and K.*

For a compact 2-manifold S we define the *genus* of S to be the sum of genera of the components of S.

Groups. If G and H are groups, then $G \approx H$ means "G is isomorphic to H". Since we deal only with abelian groups we use the term "free group" in the meaning "free abelian group". Let G be a free (abelian) group. We will call $x \in G$ a *basic* element of G if x is a member of some basis of G. Using standard facts we can prove that x is basic if and only if the subgroup of G generated by x is a nonzero direct summand of G, or, if and only if x is not equal to ny for any integer $n > 1$ and any $y \in G$.

Matrices. If $\boldsymbol{A}$ is any matrix, let $\boldsymbol{A}'$ denote the transposed of $\boldsymbol{A}$. For any positive integer n we denote by $\boldsymbol{I}_n$ and $\boldsymbol{O}_n$ the identity and zero $n \times n$ matrices, respectively. If $n = 2m$, let $\boldsymbol{J}_n$ be the matrix

$$\boldsymbol{J}_n = \begin{bmatrix} \boldsymbol{O}_m & \boldsymbol{I}_m \\ -\boldsymbol{I}_m & \boldsymbol{O}_m \end{bmatrix}.$$

For any two integers i, j let δ_{ij} denote the Kronecker symbol: $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise.

2. Surfaces in 1-acyclic 3-manifolds. The main result of this section is the following theorem, which is (together with 2.13 and 2.14) an extension of Theorem 32.3 in [2]. Note that if S is a closed 2-manifold in the interior of a 1-acyclic 3-manifold W, then S separates W. The reason is that every simple closed curve in W bounds modulo 2 in W and has therefore zero intersection number modulo 2 with S. Also, since W is orientable and S separates W, S is necessarily orientable.

THEOREM 2.1. *Let W be a 1-acyclic 3-manifold and S a closed surface of genus g in* int W. *Denote by U and V the closures of the two components of $W - S$. Then there exist oriented simple closed curves $J_1, \cdots, J_g, K_1, \cdots, K_g$ in S such that*

(1) *J_i and K_i intersect transversely at a single point, for each i, and $J_i \cap J_j = J_i \cap K_j = K_i \cap K_j = \emptyset$ if $i \neq j$;*

(2) *$J_i \sim 0$ in U and $K_i \sim 0$ in V ($i = 1, \cdots, g$);*

(3) *the homology classes of $J_1, \cdots, J_g$ form a free basis of $H_1(V)$ and the homology classes of $K_1, \cdots, K_g$ form a free basis of $H_1(U)$.*

The situation described by this theorem reminds us of the standard embedding of S in E^3; in fact, the only difference is that in the latter case we can choose $J_1, \cdots, J_g, K_1, \cdots, K_g$ so that each J_i bounds a disk (not only an orientable surface) in U and each K_i bounds a disk in V.

We postpone the proof of 2.1, which will occupy most of this section, and first prove two consequences of 2.1.

THEOREM 2.2. *Let W be a 1-acyclic 3-manifold and S a closed surface of genus g in* $\operatorname{int} W$. *Denote by U and V the closures of the two components of $W - S$. Let V' be a cube with g handles. Then there exists a homeomorphism $h\colon \operatorname{Bd} V' \to S$ such that*

(1) *the 3-manifold $W' = V' \cup_h U$ is 1-acyclic;*

(2) *if J is a closed oriented 1-manifold in S, then $J \sim 0$ in V if and only if $h^{-1}(J) \sim 0$ in V'.*

Proof. Assume 2.1. Think of V' as embedded in E^3; let $S' = \operatorname{Bd} V'$ and $U' = E^3 - \operatorname{int} V'$. Let J_i, K_i $(i = 1, \cdots, g)$ be oriented simple closed curves in S satisfying the conclusions of 2.1. Let $J_i', K_i' \subset S'$ have analogous meaning (with respect to U' and V'). Then there exists a homeomorphism $h\colon S' \to S$ which maps each J_i' onto J_i and each K_i' onto K_i (not necessarily in an orientation preserving way). Let $W' = V' \cup_h U$.

As is well-known, (1) of 2.1 implies that the homology classes of $J_1, \cdots, J_g, K_1, \cdots, K_g$ form a basis of $H_1(S)$. The homology classes of $K_1, \cdots, K_g$ belong to the kernel of $H_1(S) \to H_1(V)$; it follows from 2.1 (3) that no nontrivial linear combination of the J_i is homologous to 0 in V. Therefore, a 1-cycle in S bounds in V if and only if it is homologous in S to a linear combination of $K_1, \cdots, K_g$. Similarly, a 1-cycle in S' bounds in V' if and only if it is homologous in S' to a linear combination of $K_1', \cdots, K_g'$. Therefore, (2) of 2.2 follows directly from the choice of h.

To prove (1) of 2.2 choose an arbitrary $x \in H_1(W')$. We have to show that $x = 0$. Since S is connected and separates W', x can be represented by a sum $z_1 + z_2$ where z_1 is a 1-cycle in U and z_2 is a 1-cycle in V'. By 2.1 (3), z_1 is homologous in U to a linear combination of $K_1, \cdots, K_g$ and z_2 is homologous in V' to a linear combination of $J_1', \cdots, J_g'$. Since the sewing map h was chosen so that each $J_i' \sim 0$ in U and each $K_i \sim 0$ in V', $z_1 + z_2 \sim 0$ in W'.

THEOREM 2.3. *If C is a compact subset of a 1-acyclic 3-manifold W, then C can be embedded in a homology 3-sphere (and thus also in an acyclic 3-manifold unless C is itself a homology 3-sphere).*

Proof. We may assume that $C \subset \operatorname{int} W$. Cover C by a compact connected 3-submanifold M of $\operatorname{int} W$. Take a boundary component S

of M. Denote by U and V the closures of the components of $W - S$; let U be the one which contains M. By 2.2 we can replace V by a cube with handles, V', in such a way that $W' = U \cup V'$ is still 1-acyclic. If we perform a similar surgery along each boundary component of M, we end up with M embedded in a closed 1-acyclic 3-manifold, Σ say. It follows from Poincaré duality that Σ is a homology 3-sphere. If C is not a closed 3-manifold, then there is a point $p \in \Sigma - C$ and hence C lies in the acyclic 3-manifold $\Sigma - p$.

Before we start proving Theorem 2.1 we establish some homological properties of surfaces. Let S be a closed oriented surface and let $a_1, \cdots, a_n \in H_1(S)$. The *intersection number matrix* or the *sc-matrix* of the ordered n-tuple $(a_1, \cdots, a_n)$ is the $n \times n$ matrix $\boldsymbol{A} = (a_{ij})$, where $a_{ij} = \mathrm{sc}\,(a_i, a_j)$. Obviously $\boldsymbol{A}$ is skew-symmetric. The following lemma is proved by a straightforward computation.

LEMMA 2.4. *Let S be a closed oriented surface and $a_1, \cdots, a_m, b_1, \cdots, b_n \in H_1(S)$. Let $\boldsymbol{A}$ be the sc-matrix of $(a_1, \cdots, a_m)$ and $\boldsymbol{B}$ the sc-matrix of $(b_1, \cdots, b_n)$. Suppose that there exists an $m \times n$ matrix $\boldsymbol{T}$ with integer entries such that the column vector $(a_1, \cdots, a_n)'$ is the product of $\boldsymbol{T}$ with the column vector $(b_1, \cdots, b_n)'$. Then $\boldsymbol{A} = \boldsymbol{TBT}'$.*

LEMMA 2.5. *Let S be a closed oriented surface of genus g. Let $a_1, \cdots, a_{2g} \in H_1(S)$ and let $\boldsymbol{A}$ be the* sc-*matrix of $(a_1, \cdots, a_{2g})$. Then $\{a_1, \cdots, a_{2g}\}$ is a basis of $H_1(S)$ if and only if* $\det \boldsymbol{A} = 1$.

Proof. It is well-known that $H_1(S)$ is free of rank $2g$ and that it has a basis $\{b_1, \cdots, b_{2g}\}$ whose sc-matrix is $\boldsymbol{J}_{2g}$. There exists a $2g \times 2g$ matrix $\boldsymbol{T}$ with integer entries such that $(a_1, \cdots, a_{2g})'$ is the product of $\boldsymbol{T}$ with $(b_1, \cdots, b_{2g})'$. From 2.4 we obtain $\det \boldsymbol{A} = (\det \boldsymbol{T})^2$. Obviously $\{a_1, \cdots, a_{2g}\}$ is a basis of $H_1(S)$ if and only if $\boldsymbol{T}$ has an inverse with integer entries, and this is true if and only if $\det \boldsymbol{T} = \pm 1$. The lemma follows.

COROLLARY 2.6. *Let S be a closed surface of genus g. Let A be a subgroup of $H_1(S)$ such that* $\mathrm{sc}\,(x, y) = 0$ *for any $x, y \in A$. Then the rank of A is at most g.*

Proof. Let r be the rank of A. There exists a basis $\{a_1, \cdots, a_r\}$ of A, a basis $\{b_1, \cdots, b_{2g}\}$ of $H_1(S)$, and positive integers $k_1, \cdots, k_r$ such that $a_i = k_i b_i$ $(i = 1, \cdots, r)$. Obviously $\mathrm{sc}\,(b_i, b_j) = 0$ if $i, j \leqq r$. Therefore $\boldsymbol{B}$, the sc-matrix of $(b_1, \cdots, b_{2g})$, contains a zero $r \times r$ block. If $r > g$, then $\det \boldsymbol{B} = 0$; but this is impossible by 2.5.

The next proposition is an algebraic version of Theorem 2.1.

PROPOSITION 2.7. *Let W be a 1-acyclic 3-manifold and S a closed surface of genus g in* $\operatorname{int} W$. *Denote by U and V the closures of the components of $W - S$ and by $i: S \to U$, $j: S \to V$ the inclusions. Let $A = \operatorname{Ker} i_*$, $B = \operatorname{Ker} j_*$. Then*

(1) *$H_1(S) = A \oplus B$ and either of A, B has rank g;*

(2) *$i_* \mid B: B \to H_1(U)$ and $j_* \mid A: A \to H_1(V)$ are isomorphisms;*

(3) *if $x, y \in H_1(S)$ are either both in A or both in B, then* $\operatorname{sc}(x, y) = 0$.

Proof. Consider the Mayer-Vietoris sequence of $(W; U, V)$:

$$\cdots \longrightarrow H_1(S) \xrightarrow{\alpha} H_1(U) \oplus H_1(V) \longrightarrow H_1(W) \longrightarrow \cdots .$$

Since $H_1(W) = 0$, α is an epimorphism. We will show that it is also one-to-one. Recall that α is defined by $\alpha(x) = (i_*(x), -j_*(x))$. Take an $x \in \operatorname{Ker} \alpha = A \cap B$. Represent x by a closed oriented 1-manifold $J \subset S$. Then J bounds compact, oriented, properly embedded 2-manifolds $G' \subset U$ and $G'' \subset V$. $G = G' \cup G''$ is a closed orientable 2-manifold in $\operatorname{int} W$. Let $G_1, \cdots, G_n$ be the components of G and let $G'_r = G' \cap G_r$, $G''_r = G'' \cap G_r$, $J_r = J \cap G_r = \partial G'_r = \partial G''_r$ $(r = 1, \cdots, n)$. To prove that $x = 0$ it suffices to show that each J_r bounds a compact oriented 2-submanifold of S.

G_r separates W. Let M be the union of G_r and a component of $\operatorname{int} W - G_r$. Orient M so that $\partial M = G'_r \cup (-G''_r)$. Let $M' = M \cap U$, $F = M \cap S$. Then $\operatorname{Bd} M' = F \cup G'_r$. If we orient F so that $\partial M' = (-F) \cup G'_r$, then $\partial F = \partial G'_r = J_r$. We have thus shown that $\operatorname{Ker} \alpha = 0$ and therefore α is an isomorphism. This proves (2) and the first part of (1) of our proposition. Obviously 1.1 implies (3), and (3) together with 2.6 imply the second part of (1).

Now we start proving Theorem 2.1. In the first step we will choose the homology classes for the simple closed curves which we want to construct: a_i will be the homology class of J_i, b_i of K_i. We work with a surface in limbo.

LEMMA 2.8. *Let S be a closed oriented surface of genus g. Suppose that the group $H_1(S)$ is represented as a direct sum $A \oplus B$ so that* $\operatorname{sc}(x, y) = 0$ *for any two elements $x, y \in H_1(S)$ which lie either both in A or both in B. Then there exist bases $\{a_1, \cdots, a_g\}$ of A and $\{b_1, \cdots, b_g\}$ of B such that* $\operatorname{sc}(a_i, b_j) = \delta_{ij}$ *for each i and j.*

ADDENDUM 2.9. *Let $0 \leqq r \leqq g$ and $0 \leqq s \leqq g$. Suppose that $\{a_1, \cdots, a_r\}$ is a basis of a direct summand of A and $\{b_1, \cdots, b_s\}$ is a basis of a direct summand of B such that* $\operatorname{sc}(a_i, b_j) = \delta_{ij}$ $(i = 1, \cdots, r;$

$j = 1, \cdots, s$). *Then we can find* $a_{r+1}, \cdots, a_g, b_{s+1}, \cdots, b_g$ *such that* $a_1, \cdots, a_g, b_1, \cdots, b_g$ *satisfy the conclusion of* 2.8.

Proofs of 2.8 *and* 2.9. First note that 2.6 implies that A and B have rank g. Assume that $r \geqq s$. If $r < g$, choose any elements $a'_{r+1}, \cdots, a'_g$ such that $\{a_1, \cdots, a_r, a'_{r+1}, \cdots, a'_g\}$ is a basis of A. Then set

$$a_i = a'_i - \sum_{k=1}^{s} \mathrm{sc}\,(a'_i, b_k)a_k\ , \quad i = r+1, \cdots, g$$

(if $s = 0$, let $a_i = a'_i$). Obviously $a_1, \cdots, a_g$ again form a basis of A and $\mathrm{sc}\,(a_i, b_j) = \delta_{ij}$ for $1 \leqq i \leqq g$, $1 \leqq j \leqq s$.

Let us first consider the case $s = 0$. Choose an arbitrary basis $\{b'_1, \cdots, b'_g\}$ of B. Let $\boldsymbol{C}$ be the sc-matrix of $(a_1, \cdots, a_g, b'_1, \cdots, b'_g)$. Then

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{O}_g & \boldsymbol{D} \\ -\boldsymbol{D}' & \boldsymbol{O}_g \end{bmatrix}$$

where $\boldsymbol{D}$ is the $g \times g$ matrix whose (i, j)-entry is $\mathrm{sc}\,(a_i, b'_j)$. By 2.5, $\det \boldsymbol{C} = (\det \boldsymbol{D})^2 = 1$, therefore, $\boldsymbol{D}'$ has an inverse $\boldsymbol{U} = (u_{ij})$ with integer entries. Put

2.10 $$b_i = \sum_{j=1}^{g} u_{ij} b'_j \quad (i = 1, \cdots, g)\ .$$

Let

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{I}_g & \boldsymbol{O}_g \\ \boldsymbol{O}_g & \boldsymbol{U} \end{bmatrix}.$$

Then, by 2.4, the sc-matrix of $(a_1, \cdots, a_g, b_1, \cdots, b_g)$ is $\boldsymbol{TCT}' = \boldsymbol{J}_{2g}$. This is what we wished to have.

If $s > 0$ we work in the same way except that we do not start with an arbitrary basis $\{b'_1, \cdots, b'_g\}$ of B. Choose $b''_{s+1}, \cdots, b''_g \in B$ such that $b_1, \cdots, b_s, b''_{s+1}, \cdots, b''_g$ form a basis of B. Then set $b'_i = b_i$ for $i = 1, \cdots, s$ and

$$b'_i = b''_i - \sum_{k=1}^{s} \mathrm{sc}\,(a_k, b''_i)b_k \quad \text{for} \quad i = s+1, \cdots, g\ .$$

Then $\{b'_1, \cdots, b'_g\}$ is a basis of B and $\mathrm{sc}\,(a_i, b'_j) = \delta_{ij}$ unless $i, j > s$. This means that the matrices $\boldsymbol{D}$ and $\boldsymbol{U}$, defined as above, have the form

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{I}_s & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{E} \end{bmatrix}, \qquad \boldsymbol{U} = \begin{bmatrix} \boldsymbol{I}_s & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{V} \end{bmatrix},$$

where $\boldsymbol{E}$ is a $(g-s) \times (g-s)$ matrix and $\boldsymbol{V} = (\boldsymbol{E}')^{-1}$. Therefore, the defining formula 2.10 yields the a priori given b_i for $i = 1, \cdots, s$.

Lemma 2.8 and its Addendum are proved.

We have chosen the homology classes of our future simple closed curves J_i and K_i. Now we will show that the chosen homology classes can really be represented by simple closed curves.

PROPOSITION 2.11. *Let S be a closed surface and $x \in H_1(S)$. Then there exist an oriented simple closed curve $J \subset S$ and a positive integer n such that x is the homology class of the 1-cycle nJ.*

This proposition obviously follows from.

LEMMA 2.12. *Let S be a closed surface and $K \subset S$ a closed oriented 1-manifold. Then there exists a sequence $K^{(1)}, K^{(2)}, \cdots, K^{(m)}$ of closed oriented 1-submanifolds of S such that*

(1) $K^{(1)} = K$;

(2) *$K^{(i+1)}$ is obtained from $K^{(i)}$ either by omitting a component of $K^{(i)}$ which separates S or by piping together two components of $K^{(i)}$ $(i = 1, \cdots, m-1)$;*

(3) *any two components of $K^{(m)}$ are homologous in S.*

Proof. We use induction on the number of components of K. If K is connected, then there is nothing to prove. Suppose that 2.12 is true if K has less than n components $(n \geqq 2)$. Choose a closed oriented 1-manifold $K \subset S$ which has n components, say $K_1, \cdots, K_n$. Denote by T the 2-manifold obtained by cutting S along K, and let $p: T \to S$ be the corresponding identification map. Let L_{r1}, L_{r2} be the two components of $p^{-1}(K_r)$ $(r = 1, \cdots, n)$; orient them so that p maps each of them onto K_r in an orientation preserving way.

Case 1. Suppose that T has a component T_0 with connected boundary; let for instance $\mathrm{Bd}\, T_0 = L_{ri}$. Then K_r separates S. By induction hypothesis, 2.12 holds for $K' = K - K_r$. It obviously follows that 2.12 holds for K.

Case 2. Suppose that T has a component T_0 which has more than two boundary components. Then T_0 can be oriented so that two of its boundary components, say L_{ri} and L_{sj}, are oriented coherently with T_0. By 1.2, L_{ri} and L_{sj} can be piped together along any properly embedded arc $A \subset T_0$. We claim that $r \neq s$. Indeed, S is obtained from T by sewing each L_{k1} to L_{k2} by an orientation preserving homeomorphism and hence, if L_{k1} and L_{k2} lie in the same component of T and if T is given any orientation, one of L_{k1}, L_{k2} is oriented coherently and the other incoherently with T. It follows that we can pipe K_r to K_s along the arc $p(A)$. Denote by K_r' the oriented simple closed curve obtained by this piping. By induction hypothesis, 2.12 holds for $K' = (K - K_r - K_s) \cup K'_r$, hence it holds for K.

Case 3. Finally we consider the case when each component of T is bounded by exactly two simple closed curves. If some component of T can be oriented coherently with both its boundary components, then we prove as in Case 2 that 2.12 holds for K. Suppose that no component of T can be oriented coherently with its boundary. Obviously T has n components, say $T_1, \cdots, T_n$. Let $S_i = p(T_i)$ $(i = 1, \cdots, n)$. Since $n > 1$, no S_i is a closed surface and therefore each S_i is bounded by two components of K. We may assume that the numbering has been chosen so that $\operatorname{Bd} S_i = K_i \cup K_{i+1}$ $(i = 1, \cdots, n - 1)$ and $\operatorname{Bd} S_n = K_n \cup K_1$. Since no T_i can be oriented coherently with its boundary, the same holds for S_i. This means that each S_i can be oriented so that $\partial S_i = (-K_i) \cup K_{i+1}$ $(i = 1, \cdots, n - 1)$. It follows that $K_1 \sim K_2 \sim \cdots \sim K_n$ in S. This concludes the proof of 2.12.

Theorem 2.1 follows from 2.7 and the following proposition.

PROPOSITION 2.13. *Let S be a closed oriented surface of genus g. Suppose that $H_1(S)$ is represented as a direct sum $A \oplus B$ so that* $\operatorname{sc}(x, y) = 0$ *for any $x, y \in H_1(S)$ that lie either both in A or both in B. Then there exist oriented simple closed curves J_i, K_i in S $(i = 1, \cdots, g)$ such that*

(1) *for each i, $J_i \cap K_i$ is a single point and* $\operatorname{sc}(J_i, K_i) = 1$; *if $i \neq j$, then $J_i \cap J_j = J_i \cap K_j = K_i \cap K_j = \varnothing$;*

(2) *the homology classes of $J_1, \cdots, J_g$ form a basis of A and the homology classes of $K_1, \cdots, K_g$ form a basis of B.*

ADDENDUM 2.14. *Suppose that we are given elements $a_1, \cdots, a_{r'} \in A$, $b_1, \cdots, b_{s'} \in B$ $(0 \leqq r' \leqq g,\ 0 \leqq s' \leqq g)$ and oriented simple closed curves $J_1, \cdots, J_r$, $K_1, \cdots, K_s \subset S$ $(0 \leqq r \leqq r',\ 0 \leqq s \leqq s')$ such that the following conditions are satisfied:*

(i) *if $i \leqq \min(r, s)$, $J_i \cap K_i$ is a single point; if $i \neq j$, then $J_i \cap J_j = \varnothing$, $J_i \cap K_j = \varnothing$, $K_i \cap K_j = \varnothing$ (each of these equalities is satisfied for all pairs i, j for which it makes sense);*

(ii) *a_i is the homology class of J_i and b_j is the homology class of K_j $(i = 1, \cdots, r;\ j = 1, \cdots, s)$;*

(iii) *$\{a_1, \cdots, a_{r'}\}$ is a basis of a direct summand of A and $\{b_1, \cdots, b_{s'}\}$ is a basis of a direct summand of B;*

(iv) $\operatorname{sc}(a_i, b_j) = \delta_{ij}$ $(i = 1, \cdots, r';\ j = 1, \cdots, s')$. *Then there exist oriented simple closed curves $J_{r+1}, \cdots, J_g$, $K_{s+1}, \cdots, K_g$ such that J_i represents a_i $(i = r + 1, \cdots, r')$, K_j represents b_j $(j = s + 1, \cdots, s')$, and $J_1, \cdots, J_g$, $K_1, \cdots, K_g$ satisfy the conclusions of* 2.13.

REMARK. It is not difficult to show that if $J_1, \cdots, J_r$ are disjoint oriented simple closed curves in S such that $S - (J_1 \cup \cdots \cup J_r)$ is connected, then the homology classes of these curves freely generate a direct summand of $H_1(S)$. Therefore, if $r' = r$ and $s' = s$, we can replace the condition (iii) above by: $a_i \in A$, $b_i \in B$, and $S - (J_1 \cup \cdots \cup J_r)$, $S - (K_1 \cup \cdots \cup K_s)$ are connected.

In the proof of 2.13 and 2.14 we shall need the following three lemmas. The proofs of 2.15 and 2.17 are easy and we omit them.

LEMMA 2.15. *Let S be a surface, $J \subset S$ an oriented simple closed curve, and $L \subset S$ an oriented closed* 1-*manifold. Orient S so that $n = \mathrm{sc}\,(J, L) \geqq 0$. Then L is homologous to an oriented closed* 1-*manifold $K \subset S$ such that $J \cap K$ contains exactly n points.*

The K of 2.15 may have to have more components than L. On the other hand, the following lemma is valid.

LEMMA 2.16. *Let S be an oriented surface, $J \subset \mathrm{int}\, S$ an oriented simple closed curve, and $L \subset \mathrm{int}\, S$ an oriented closed* 1-*manifold such that* $\mathrm{sc}\,(J, L) = 1$. *Then L is homologous to an oriented simple closed curve $K \subset S$ such that $J \cap K$ contains exactly* 1 *point.*

Proof. By 2.15 we may assume that $J \cap L$ has only one point. We will prove the lemma by induction on the number of components of L. If this number is 1, we can take $K = L$. Suppose that 2.16 is true if L has at most n components ($n \geqq 1$). Take an L with $n + 1$ components, say $L_0, L_1, \cdots, L_n$; let L_0 be the component which intersects J.

Denote by T the 2-manifold obtained by cutting S along all components of L and let $p: T \to S$ be the corresponding identification map. Let L_i', L_i'' be the two boundary components of T composing $p^{-1}(L_i)$ ($i = 0, 1, \cdots, n$). Orient L_i' and L_i'' so that p maps each of them onto L_i in an orientation preserving way. Obviously $p^{-1}(J)$ is an arc connecting L_0' and L_0''. Hence L_0'' and L_0' lie in the same component, T_0 say, of T. Clearly, $\mathrm{Bd}\, T_0$ intersects $p^{-1}(L - L_0)$. Changing the notation, if necessary, we can assume that $L_1' \subset \mathrm{Bd}\, T_0$. Orient T_0 coherently with L_1'. Then one of L_0', L_0'' is oriented coherently with T_0 and the other incoherently. Assume that L_0' is oriented coherently with T_0. Let $A \subset T_0$ be a properly embedded arc which misses $p^{-1}(J)$ and joins L_0' to L_1'. By 1.2 we can pipe L_0' to L_1' along A. It follows that in S we can pipe L_0 to L_1 along the arc $p(A)$, whose interior misses $J \cup L$. This piping changes L to a closed oriented 1-manifold, homologous to L, which still intersects J at a single point and has only n components. Therefore, the induction hypothesis implies that 2.16 holds for L.

LEMMA 2.17. *Let S be a closed surface of genus $g > 0$ and let $J, K \subset S$ be two simple closed curves crossing each other at a single point. Denote by T the surface obtained by cutting S along J and K and let $p: T \to S$ be the corresponding identification map. Let S' be the closed surface obtained by attaching a disk to T along the boundary curve $p^{-1}(J \cup K)$ of T; let $k: T \to S'$ be the inclusion. Then k_* is an isomorphism and $p_*k_*^{-1}: H_1(S') \to H_1(S)$ maps $H_1(S')$ isomorphically onto the direct summand of $H_1(S)$ which consists of homology classes of 1-cycles that have zero intersection numbers with both J and K. Moreover, if we orient S and S' so that p preserves orientation, then $p_*k_*^{-1}$ preserves intersection numbers.*

Proofs of 2.13 *and* 2.14. By 2.8 and 2.9 we can assume that $r' = s' = g$. We also assume that $r \geqq s$.

The proof is by induction on the genus of S. If this genus is 0, there is nothing to prove. Suppose that 2.13 and 2.14 are true if the genus of S is less than $g (g > 0)$ and consider a situation with the genus of S equal to g.

If $r > 0$, then we already have J_1. If $r = 0$, choose any oriented simple closed curve representing a_1 (it follows from 2.11 that one such exists) and call it J_1. If $s > 0$ (and hence $r > 0$ by our assumption), then we already have K_1. Suppose that $s = 0$. Represent b_1 by a closed oriented 1-manifold L. By 2.15 we can assume that $L \cap (J_2 \cup \cdots \cup J_r) = \varnothing$. Applying 2.16 to $S - (J_2 \cup \cdots \cup J_r)$, J_1, and L we can find an oriented simple closed curve $K_1 \sim L$ such that $J_1 \cap K_1$ is a single point and $K_1 \cap (J_2 \cup \cdots \cup J_r) = \varnothing$.

We can therefore assume that we already have a "good" pair J_1, K_1, either preassigned or constructed as described above. Define T, p, S', and k as in the statement of 2.17, with J_1 and K_1 taking the roles of J and K, respectively. It follows from 2.17 that S', $g' = g - 1$, $A' = k_*p_*^{-1}(A)$, $B' = k_*p_*^{-1}(B)$, $a_i' = k_*p_*^{-1}(a_i)$ and $b_i' = k_*p_*^{-1}(b_i)$ $(i = 2, \cdots, g)$, $J_i' = kp^{-1}(J_i)$ and $K_j' = kp^{-1}(K_j)$ $(i = 2, \cdots, r;$ $j = 2, \cdots, s)$ satisfy the hypotheses of 2.13 and 2.14. By induction hypothesis we can represent each a_i' by an oriented simple closed curve $J_i' \subset k(T) \subset S'$ and each b_j' by an oriented simple closed curve $K_j' \subset k(T) \subset S'$ $(i = r + 1, \cdots, g;\ j = s + 1, \cdots, g)$ such that $J_2', \cdots, J_g', K_2', \cdots, K_g'$ satisfy (1) of 2.13. Let $J_i = pk^{-1}(J_i')$, $K_j = pk^{-1}(K_j')$ $(i = r + 1, \cdots, g;\ j = s + 1, \cdots, g)$. Then $J_1, \cdots, J_g$, $K_1, \cdots, K_g$ satisfy the conclusions of 2.13 and 2.14.

We conclude this section with a proof of the following theorem.

THEOREM 2.18. *Let U be a cube with g handles. Denote* Bd U *by S and let $i: S \to U$ be the inclusion. Let $\{a_1, \cdots, a_g\}$ be any basis of* Ker i_*. *Then we can represent each a_r by an oriented simple* *closed curve $J_r \subset S$ $(r = 1, \cdots, g)$ such that $J_1, \cdots, J_g$ bound disjoint disks in U.*

This result is implicitly contained in pp. 296–299 of [2]. But perhaps it is worth while stating and proving it explicitly. Let us first consider the following weaker lemma.

LEMMA 2.19. *Let U, S, and i be as in 2.18. Then every $x \in \operatorname{Ker} i_*$ can be represented by a 1-cycle nJ where n is a positive integer and $J \subset S$ is an oriented simple closed curve which bounds a disk in U.*

Proof. We can assume that $x \neq 0$. Let $\{K_1, \cdots, K_g\}$ be a collection of oriented simple closed curves in S which bound disjoint disks in U and whose union does not separate S. Then the homology classes of $K_1, \cdots, K_g$ form a basis of $\operatorname{Ker} i_*$. Therefore, there exist integers $n_1, \cdots, n_g$ such that the 1-cycle $n_1K_1 + \cdots + n_gK_g$ represents x. This obviously implies that x can be represented by an oriented closed 1-manifold K such that the components of K bound disjoint disks in U. Therefore, 2.19 easily follows from 2.12 and the following obvious lemma.

LEMMA 2.20. *Let U be a 3-manifold and let L_1, L_2 be oriented simple closed curves in $\operatorname{Bd} U$ bounding disjoint properly embedded disks E_1 and E_2, respectively, in U. Suppose that L_1 can be piped to L_2 along an arc $A \subset \operatorname{Bd} U$ and let L be the simple closed curve obtained by this piping. Let N be a neighborhood of A in U containing the "pipe" $L - (L_1 \cup L_2)$. Then L bounds a properly embedded disk $E \subset U$ which is contained in $E_1 \cup E_2 \cup N$.*

Proof of 2.18. Suppose that 2.18 is false and take the smallest g for which 2.18 fails. By 2.19, $g > 1$.

Embed U into E^3. Let $V = E^3 - \operatorname{int} U$ and let $j: S \to V$ be the inclusion. Choose an orientation for S. It follows from 2.7, 2.8, and 2.9 that there exists a basis $\{b_1, \cdots, b_g\}$ of $\operatorname{Ker} j_*$ such that $\operatorname{sc}(a_r, b_s) = \delta_{rs}$ $(r, s = 1, \cdots, g)$. By 2.19 we can represent a_1 by an oriented simple closed curve $J_1 \subset S$ which bounds a properly embedded disk $D_1 \subset U$. Obviously $U - D_1$ is connected.

Represent a_r by an oriented closed 1-manifold $K_r \subset S$ and b_r by an oriented closed 1-manifold $L_r \subset S$ $(r = 2, \cdots, g)$; by 2.15 we may assume that $J_1 \cap K_r = J_1 \cap L_r = \varnothing$. Choose compact, oriented, properly embedded 2-manifolds $F_r \subset U$, $G_r \subset V$ such that $\partial F_r = K_r$, $\partial G_r = L_r$. If F_r intersects D_1, we can put F_r in general position with D_1, remove the part of F_r which lies in a regular neighborhood of D_1, and then patch the resultant holes in F_r by disjoint disks "parallel" to D_1. In this manner we can replace F_r by another compact, oriented, properly embedded 2-manifold in U such that it is bounded by K_r and misses D_1. Therefore, we will assume that the originally chosen $F_2, \cdots, F_g$ were already disjoint from D_1.

Choose a regular neighborhood N of D_1 in E^3 and let $U' = U - \operatorname{int} N$, $V' = V \cup N$, $S' = \operatorname{Bd} U' = \operatorname{Bd} V'$. Let i' and j' be the inclusions of S' in U' and V', respectively. U' is again a cube with handles ([2], 6.2). Let $T = S \cap S'$. Then $S' - \operatorname{int} T$ consists of two disjoint disks, which we denote by D_1' and D_1''. Let $J_1' = \operatorname{Bd} D_1'$, $J_1'' = \operatorname{Bd} D_1''$. Orient J_1' and J_1'' so that $J_1' \sim J_1 \sim J_1''$ in S. Let $a_r', b_r' \in H_1(S')$ be the homology classes of K_r, L_r, respectively ($r = 2, \cdots, g$). Since K_r bounds F_r in U' and L_r bounds G_r in V' we have $a_r' \in \operatorname{Ker} i_*'$, $b_r' \in \operatorname{Ker} j_*'$. If we give S' the orientation which on T agrees with the chosen orientation of S, then $\operatorname{sc}(a_r', b_s') = \operatorname{sc}(K_r, L_s) = \delta_{rs}$. Thus it follows from 2.5 that the a_r' and b_r' form a basis of $H_1(S')$ and therefore, $\{a_2', \cdots, a_g'\}$ is a basis of $\operatorname{Ker} i_*'$.

By supposition, 2.18 is true for cubes with $g - 1$ handles. Thus there exist oriented simple closed curves $J_2', \cdots, J_g' \subset S'$ and disjoint properly embedded disks $D_2', \cdots, D_g'$ in U' such that for each r the following are true:

(a) J_r' is in the homology class a_r' and hence $J_r' \sim K_r$ in S';

(b) $J_r' = \operatorname{Bd} D_r'$.

Without loss of generality we can assume

(c) $J_r' \subset \operatorname{int} T$.

Note that $T - (J_2' \cup \cdots \cup J_g')$ is connected.

It is easy to see that (a) above implies that $K_r \sim J_r' + n_r' J_1' + n_r'' J_1''$ in T for some integers n_r', n_r''. Hence $K_r \sim J_r' + n_r J_1$ in S, where $n_r = n_r' + n_r''$. We will therefore try to replace each $J_r' + n_r J_1$ by a homologous oriented simple closed curve bounding a disk in U.

Suppose that $n_2 \neq 0$. Let for instance $n_2 > 0$. We can show, by the same argument as twice before, that it is possible to pipe J_2' to J_1' along an arc whose interior misses $J_1' \cup J_1'' \cup J_2' \cup \cdots \cup J_g'$. (If $n_2 < 0$, we pipe J_2' to $-J_1'$.) By this piping we obtain an oriented simple closed curve J_2''; 2.20 implies that J_2'' bounds a properly embedded disk $D_2'' \subset U$ which is disjoint from $D_1, D_3', \cdots, D_g'$. We replace J_2' by J_2'' and D_2' by D_2''. Now we have $K_2 \sim J_2'' + m_2 J_1$, where $|m_2| = |n_2| - 1$. It should now be clear how to finish the proof of 2.18 by induction on the number $|n_2| + \cdots + |n_g|$.

3. Compact 3-submanifolds of acyclic 3-manifolds.

In this section we will prove the following two theorems.

THEOREM 3.1. *A compact connected* 3-*manifold* M *whose boundary* *has m components ($m > 0$) is subacyclic if and only if it satisfies the following conditions* (1), (2), *and either* (3′) *or* (3″):

(1) *M is orientable;*

(2) *$H_1(M)$ is free;*

(3′) *$H_2(M)$ is free of rank $m - 1$;*

(3″) *$H_1(\mathrm{Bd}\, M) \to H_1(M)$ is onto.*

THEOREM 3.2. *Let M be a compact, connected, subacyclic 3-manifold and J a closed oriented 1-manifold lying in a boundary component S of M. Let F be an oriented surface and $h\colon \partial F \to J$ an orientation preserving homeomorphism. Then the polyhedron $P = F \cup_h M$ can be embedded in an acyclic 3-manifold if and only if J satisfies one of the following two conditions.*

(1) *The homology class of J in M is a basic element of $H_1(M)$.*

(2) *There exist compact 2-submanifolds $G, H \subset S$ such that $G \cup H = S$, $G \cap H = J$, and there exists an orientation of S such that $\partial G = -\partial H = J$.*

The proof of 3.1 in one direction is quite easy. Suppose that M lies in an open acyclic 3-manifold W. Then M is orientable. Let V be the closure of $W - M$. The Mayer-Vietoris sequence of $(W; M, V)$ contains the following subsequence

$$0 \longrightarrow H_1(\mathrm{Bd}\, M) \longrightarrow H_1(M) \oplus H_1(V) \longrightarrow 0 .$$

It follows that $H_1(M)$ is free and that $H_1(\mathrm{Bd}\, M) \to H_1(M)$ is onto.

The other direction of 3.1 will be proved by induction on m. First we show that the conditions (3′) and (3″) of 3.1 are equivalent.

LEMMA 3.3. *Let M be a compact connected 3-manifold with m boundary components and suppose that* (1) *and* (2) *of* 3.1 *are satisfied. Then* (3′) *and* (3″) *of* 3.1 *are equivalent and they imply that $H_2(M, \mathrm{Bd}\, M) \approx H_1(M)$ and that the following sequence is split exact:*

$$0 \longrightarrow H_2(M, \mathrm{Bd}\, M) \xrightarrow{\partial_*} H_1(\mathrm{Bd}\, M) \xrightarrow{i_*} H_1(M) \longrightarrow 0$$

(here ∂_* and i_* are the homomorphisms from the homology sequence of the pair $(M, \mathrm{Bd}\, M)$).

Proof. Since $H_1(M)$ and $H_2(M)$ are free, duality and the Universal Coefficient Theorem yield the following two relations

$$H_1(M, \mathrm{Bd}\, M) \approx H^2(M) \approx H_2(M), \quad H_2(M, \mathrm{Bd}\, M) \approx H^1(M) \approx H_1(M) .$$

Consider the exact sequence for the reduced homology of the pair $(M, \mathrm{Bd}\, M)$:

$$\cdots \longrightarrow H_2(\operatorname{Bd} M) \longrightarrow H_2(M) \longrightarrow H_2(M, \operatorname{Bd} M) \longrightarrow$$
$$\longrightarrow H_1(\operatorname{Bd} M) \longrightarrow H_1(M) \longrightarrow H_1(M, \operatorname{Bd} M) \longrightarrow \tilde{H}_0(\operatorname{Bd} M) \longrightarrow 0\,.$$

Suppose that $H_1(\operatorname{Bd} M) \to H_1(M)$ is onto. Then $\tilde{H}_0(\operatorname{Bd} M) \approx H_1(M, \operatorname{Bd} M) \approx H_2(M)$ and hence $H_2(M)$ is free of rank $m - 1$. Now suppose that $H_2(M)$ is free of rank $m-1$. Then $H_1(M, \operatorname{Bd} M) \to \tilde{H}_0(\operatorname{Bd} M)$ is an epimorphism of free groups of the same rank and thus it is actually an isomorphism. It follows that $H_1(\operatorname{Bd} M) \to H_1(M)$ is onto.

We conclude the proof of 3.3 by showing that (3′) of 3.1 implies that $\partial_*: H_2(M, \operatorname{Bd} M) \to H_1(\operatorname{Bd} M)$ is one-to-one. It suffices to show that $H_2(\operatorname{Bd} M) \to H_2(M)$ is onto, and this follows from the fact that the image of $H_2(\operatorname{Bd} M) \to H_2(M)$ is free of rank $m - 1$ (this is true for any 3-manifold M which has exactly m compact orientable boundary components) and that $H_2(M, \operatorname{Bd} M)$ is torsion free.

Now we start proving the remaining direction of 3.1.

LEMMA 3.4. *Let M be a compact connected 3-manifold having precisely m boundary components and satisfying* (1), (2), *and* (3′) *of* 3.1. *Suppose that there exists an oriented simple closed curve $K \subset \operatorname{Bd} M$ such that the homology class of K in M is a basic element of $H_1(M)$. Then M can be embedded in a compact connected 3-manifold M' which has again m boundary components, again satisfies* (1), (2), *and* (3′) *of* 3.1, *and whose boundary has smaller genus than* $\operatorname{Bd} M$.

Proof. Denote by S the component of $\operatorname{Bd} M$ which contains K and let A be a regular neighborhood of K in S. Let M' be the 3-manifold obtained by attaching a 2-handle H to M along A. Since K does not separate S, M' has exactly m boundary components and $\operatorname{Bd} M'$ has smaller genus than $\operatorname{Bd} M$. Obviously M' is compact, connected, and orientable. By considering the Mayer-Vietoris sequence of $(M'; M, H)$ for reduced homology we can prove that $H_1(M')$ is free and $H_2(M') \approx H_2(M)$.

LEMMA 3.5. *Theorem* 3.1 *is valid for* $m = 1$.

Proof. Suppose that this is false. Among all 3-manifolds M which are counterexamples to 3.1 for $m = 1$ choose one whose boundary has the smallest genus. Because of 2.3, $H_1(M)$ is nontrivial. Choose a basic element $x \in H_1(M)$. It follows from 3.1 (3″) and 2.11 that x can be represented by a simple closed curve $K \subset \operatorname{Bd} M$. But then 3.4 yields a 3-manifold M' which is a counterexample to 3.1 for $m = 1$ and whose boundary has smaller genus than $\operatorname{Bd} M$. This contradicts our choice of M.

LEMMA 3.6. *Theorem* 3.1 *is valid for* $m = 2$.

Proof. Suppose that the lemma is false. Choose a 3-manifold M which is a counterexample to 3.1 for $m = 2$ and whose boundary has the smallest possible genus. Our plan is to find a simple closed curve $K \subset \operatorname{Bd} M$ representing a basic element of $H_1(M)$; as in 3.5 this will lead to a contradiction.

Let S' and S'' be the two components of $\operatorname{Bd} M$, let g' be the genus of S' and g'' the genus of S'', and let $i': S' \to M$, $i'': S'' \to M$, $i: \operatorname{Bd} M \to M$ be inclusions. By 3.3, $H_1(M)$ has rank $g' + g''$ and therefore $g' + g'' > 0$.

SUBLEMMA 1. $\operatorname{Ker} i'_* = \operatorname{Ker} i''_* = 0$.

Proof. Suppose that e.g. $\operatorname{Ker} i'_* \neq 0$. Since $H_1(M)$ is free, $\operatorname{Ker} i'_*$ is a direct summand of $H_1(S')$. Therefore, it follows from 2.11 that there exists a nonseparating oriented simple closed curve $J \subset S'$ such that $J \sim 0$ in M. Let $K \subset S'$ be a simple closed curve intersecting J transversely at exactly one point. Choose an orientation for M, orient $\operatorname{Bd} M$ coherently with M, and then orient K so that $\operatorname{sc}(J, K) = 1$.

We claim that K represents a basic element of $H_1(M)$. Suppose that for some oriented closed 1-manifold $L \subset M$ and for some positive integer n, K is homologous to nL in M. Since M satisfies (3″) of 3.1 we can assume that $L \subset \operatorname{Bd} M$. Then $K - nL$ is a 1-cycle in $\operatorname{Bd} M$, homologous to 0 in M. By 1.1, $1 - n \operatorname{sc}(J, L) = \operatorname{sc}(J, K - nL) = 0$. Hence $n = 1$ and consequently K represents a basic element of $H_1(M)$. As we know, this leads to a contradiction and hence our supposition above must be wrong. Sublemma 1 is proved.

Identify $H_1(\operatorname{Bd} M)$ with $H_1(S') \oplus H_1(S'')$ and let $p': H_1(\operatorname{Bd} M) \to H_1(S')$, $p'': H_1(\operatorname{Bd} M) \to H_1(S'')$ be natural projections.

SUBLEMMA 2. The compositions

$$p'\partial_*: H_2(M, \operatorname{Bd} M) \longrightarrow H_1(S'), \qquad p''\partial_*: H_2(M, \operatorname{Bd} M) \longrightarrow H_1(S'')$$

are monomorphisms and hence $g' = g''$.

Proof. Let $x \in H_2(M, \operatorname{Bd} M)$ be such that $p'\partial_*(x) = 0$. Then $i''_* p''\partial_*(x) = i'_* p'\partial_*(x) + i''_* p''\partial_*(x) = i_*\partial_*(x) = 0$. This equality and Sublemma 1 imply that $p''\partial_*(x) = 0$. Therefore, $\partial_*(x) = 0$ and hence, by 3.3, $x = 0$. Similarly we show that $p''\partial_*$ is one-to-one.

By 3.3, $H_2(M, \operatorname{Bd} M)$ has rank $g' + g''$. Since $p'\partial_*$ is one-to-one, $g' + g'' \leqq 2g'$; similarly, $g' + g'' \leqq 2g''$. Hence $g' = g''$.

Denote $g' = g''$ by g.

SUBLEMMA 3. There exist oriented simple closed curves $J, K \subset \operatorname{Bd} M$, one lying in S' and the other in S'' and neither homologous to 0 in $\operatorname{Bd} M$, and there exists a positive integer r such that $J + rK \sim 0$ in M.

Proof. Choose a basis $\{a'_1, \cdots, a'_{2g}\}$ of $H_1(S')$, a basis $\{b'_1, \cdots, b'_{2g}\}$ of $E' = p'\partial_* H_2(M, \operatorname{Bd} M)$, and positive integers $n'_1, \cdots, n'_{2g}$ such that $b'_j = n'_j a'_j$ $(j = 1, \cdots, 2g)$. Let n' be the greatest common divisor of $n'_1, \cdots, n'_{2g}$.

Suppose that $n' = 1$. Then E' contains a basic element of $H_1(S')$. Indeed; suppose that no element of E' is basic for $H_1(S')$. Then there exists a prime q such that $E' \subset qH_1(S')$ (see e.g. [1], 5.1.1). Since $n' = 1$ there is an s $(1 \leqq s \leqq 2g)$ such that n'_s is not divisible by q. Then, as a'_s is a basic element of $H_1(S')$, $b'_s = n'_s a'_s$ is not equal to qx for any $x \in H_1(S')$ and this contradicts our previous conclusion. Thus there really exists a $b' \in E'$ which is a basic element of $H_1(S')$. Let $b'' = p''\partial_*(p'\partial_*)^{-1}(b') \in H_1(S'')$. By Sublemma 2, $b'' \neq 0$. It follows from 2.11 that there exist oriented simple closed curves $J \subset S'$, $K \subset S''$ and a positive integer r such that J represents b' and rK represents b''. Since $(b', b'') = \partial_*(p'\partial_*)^{-1}(b') \in \operatorname{Ker} i_*$, $J + rK \sim 0$ in M. Thus Sublemma 3 is true in this case.

Now suppose that $n' > 1$. For each j let $b''_j = p''\partial_*(p'\partial_*)^{-1}(b'_j) \in H_1(S'')$; let a''_j be the basic element of $H_1(S'')$ and n''_j the positive integer such that $b''_j = n''_j a''_j$. Let n'' be the greatest common divisor of $n''_1, \cdots, n''_{2g}$. If $n'' = 1$ we show as above that Sublemma 3 is valid. Suppose that $n'' > 1$. We will show that this leads to a contradiction. Choose a prime divisor q of n''. By 2.5 the determinant of the intersection number matrix of $(a'_1, \cdots, a'_{2g})$ is equal to 1. Therefore, there exists an entry of this matrix, say $\operatorname{sc}(a'_s, a'_t)$, which is not divisible by q. Note that each pair (b'_j, b''_j) lies in $\operatorname{Ker} i_*$. Therefore, 1.1 implies that $\operatorname{sc}((b'_s, b''_s), (b'_t, b''_t)) = 0$ and hence $\operatorname{sc}(b'_s, b'_t) = -\operatorname{sc}(b''_s, b''_t)$. Since the number $n'_s n'_t \operatorname{sc}(a'_s, a'_t) = -n''_s n''_t \operatorname{sc}(a''_s, a''_t)$ is divisible by q and $\operatorname{sc}(a'_s, a'_t)$ is not, one of n'_s, n'_t, say n'_s, is divisible by q. Let for instance $n'_s = qk'$, $n''_s = qk''$. Put $a' = k'a'_s$, $a'' = k''a''_s$. Then the basic element $(p'\partial_*)^{-1}(b'_s)$ of $H_2(M, \operatorname{Bd} M)$ is mapped by ∂_* to $(b'_s, b''_s) = q(a', a'')$. This contradicts the fact that ∂_* embeds $H_2(M, \operatorname{Bd} M)$ as a direct summand into $H_1(\operatorname{Bd} M)$. Sublemma 3 is proved.

We conclude the proof of 3.6 with

SUBLEMMA 4. K represents a basic element of $H_1(M)$.

Proof. Assume that $J \subset S'$, $K \subset S''$. Let $u \in H_1(S')$, $v \in H_1(S'')$ be the homology classes of J, K, respectively; let $x = \partial_*^{-1}(u, rv) \in$

$H_2(M, \operatorname{Bd} M)$. Since u is a basic element of $H_1(S')$, x is basic for $H_2(M, \operatorname{Bd} M)$. Let $H_2(M, \operatorname{Bd} M) = A \oplus B$ where A is the subgroup generated by x. Let A' be the subgroup of $H_1(S')$ generated by u, A'' the subgroup of $H_1(S'')$ generated by v, B' the smallest direct summand of $H_1(S')$ containing $p'\partial_*(B)$, and B'' the smallest direct summand of $H_1(S'')$ containing $p''\partial_*(B)$. Then $H_1(S') = A' \oplus B'$ and $H_1(S'') = A'' \oplus B''$.

We have to show that $i''_*(v) = i_*(0, v)$ is a basic element of $H_1(M)$. It follows from Sublemma 1 that $i''_*(v) \neq 0$. Suppose that there exists an integer $n > 1$ and an element $z \in H_1(M)$ such that $i''_*(v) = nz$. It follows from 3.3 that there exist elements $a \in H_1(S')$, $b \in H_1(S'')$, $y \in H_2(M, \operatorname{Bd} M)$ such that $z = i_*(a, b)$ and $(na, nb - v) = \partial_*(y)$. Let $a = \alpha u + a_0$, $b = \beta v + b_0$, $y = \lambda x + y_0$, where α, β, λ are integers and $a_0 \in B'$, $b_0 \in B''$, $y_0 \in B$. Then we have $\partial_*(y) = \lambda\partial_*(x) + \partial_*(y_0)$ or $(na, nb - v) = \lambda(u, rv) + \partial_*(y_0)$. Applying on both sides of this equation the natural projection $H_1(\operatorname{Bd} M) \to A'$ we obtain $n\alpha u = \lambda u$; projecting to A'' yields $(n\beta - 1)v = \lambda rv$. The former equation implies that λ is divisible by n, which contradicts the latter equation. Sublemma 4 and Lemma 3.6 are proved.

The following lemma is a special case of Theorem 3.2.

LEMMA 3.7. *Let M' be a compact, connected, subacyclic 3-manifold, S a boundary component of M', and $A \subset S$ a separating annulus. Let M be the 3-manifold obtained by attaching a 2-handle to M' along A. Then M is subacyclic.*

Proof. Embed M' in a homology 3-sphere Σ. Let U be the closure of the component of $\Sigma - S$ which intersects M'. Denote by V the 3-manifold obtained by attaching a 2-handle H to U along A. Then there exists a natural embedding of M into V and therefore in order to prove our lemma it suffices to show that V is subacyclic. Obviously V is orientable. Since $H_2(U) = 0$ (by the already proved part of 3.1), the following is a section of the Mayer-Vietoris sequence of $(V; U, H)$ for reduced homology:

$$0 \longrightarrow H_2(V) \longrightarrow H_1(A) \longrightarrow H_1(U) \longrightarrow H_1(V) \longrightarrow 0 .$$

As A separates S the homomorphism $H_1(A) \to H_1(U)$ is trivial. Hence $H_2(V) \approx H_1(A)$ and $H_1(V) \approx H_1(U)$. Since $\operatorname{Bd} V$ has two components it follows from 3.6 that V is subacyclic.

We conclude the proof of 3.1 by proving

LEMMA 3.8. *Suppose that* 3.1 *holds for $m < k$ $(k > 2)$. Then* 3.1 *is true for $m = k$.*

Proof. Let M be a compact, connected 3-manifold whose boundary has k components and which satisfies conditions (1), (2), and (3′) of 3.1 (with $m = k$). Let J be a properly embedded arc in M whose endpoints lie in different components of $\mathrm{Bd}\, M$. Let N be a regular neighborhood of J in M and let M' be the closure of $M - N$. Then M' is a compact, connected, orientable 3-manifold with $k - 1$ boundary components. Let S be the component of $\mathrm{Bd}\, M'$ which intersects N. We can think of N as a 2-handle attached to M' along the annulus $A = M' \cap N$, which separates S. By considering the Mayer-Vietoris sequence of $(M; M', N)$ for reduced homology we can prove that M' satisfies (2) and (3′) of 3.1 (with $m = k - 1$). By the hypothesis of the lemma this implies that M' is subacyclic. Hence, by 3.7, M is subacyclic.

Proof of 3.2. We consider all possible situations with respect to the homological properties of J in M. We divide these situations in two larger groups. First we consider

Case 1. Suppose that $J \sim 0$ in S. In this case there exists a unique pair of compact 2-submanifolds of S, say G, H, such that $G \cup H = S$, $G \cap H = J$. Indeed; choose a point $x_0 \in S - J$ and let G be the closure of the set of all points in $S - J$ that can be reached from x_0 by some arc in S which misses J or crosses J an even number of times; let $H = S - \mathrm{int}\, G$. If at least one of G, H is such that it cannot be oriented coherently with J, then the polyhedron P cannot be embedded in any acyclic 3-manifold. Suppose that e.g. G cannot be oriented coherently with J. Then $F \cup G \subset P$ is a non-orientable closed 2-manifold and therefore, as observed in the beginning of § 2, $F \cup G$ cannot be embedded in any acyclic 3-manifold.

Now suppose that both G and H can be oriented coherently with J. Then there exists an orientation for S such that, for the induced orientations in G and H, $J = \partial G = -\partial H$. Give S this orientation. Let $G_1, \cdots, G_n$ be the components of G. For $i = 1, \cdots, n$ do the following. Let $K_i = h^{-1}(\mathrm{Bd}\, G_i) \subset \mathrm{Bd}\, F$. Choose a disk with holes $G'_i \subset G_i$ such that $\mathrm{Bd}\, G'_i = J_i \cup \mathrm{Bd}\, G_i$ where J_i is a simple closed curve in $\mathrm{int}\, G_i$. Similarly choose a disk with holes $F'_i \subset F$ such that $\mathrm{Bd}\, F'_i = K_i \cup K'_i$ where K'_i is a simple closed curve in $\mathrm{int}\, F$; let $F'_1, \cdots, F'_n$ be pairwise disjoint. Orient J_i coherently with G'_i and K'_i coherently with F'_i.

Let $C = S \times I$, where $I = [0, 1]$, be an outer collar of M on S and let $M' = M \cup C$, $S' = S \times 1 \subset \mathrm{Bd}\, M'$ ($S \times 0$ is identified with S in the natural way). Since F'_i is homeomorphic to G'_i there exists a proper embedding h'_i of F'_i into $G'_i \times I \subset C$ such that $h'_i \mid K_i = h \mid K_i$ and $h'_i(K'_i) = J'_i = J_i \times 1 \subset S'$. In particular, choose a function $f_i \colon F'_i \rightarrow$

I such that $f_i(K_i) = 0$, $f_i(K_i') = 1$, $f_i(\operatorname{int} F_i') = \operatorname{int} I$; extend $h \mid K_i$ to a homeomorphism $h_i\colon F_i' \to G_i'$ and then set $h_i'(x) = (h_i(x), f_i(x)) \in G_i' \times I$ $(x \in F_i')$.

Let $H' = (H \cup \bigcup G_i') \times 1 \subset S'$, $G' = S' - \operatorname{int} H'$, $J' = \bigcup J_i'$. Orient S' and J' so that the natural homeomorphisms $S' \to S$, $J_i' \to J_i$ preserve orientations. Then $J' = -\partial G' = \partial H'$. Denote by F' the closure of $F - \bigcup F_i'$ and let $h'\colon \partial F' \to J'$ be defined by $h' \mid K_i' = h_i' \mid K_i'$ $(i = 1, \cdots, n)$. Then h' is an orientation reversing homeomorphism. There is an obvious embedding of P into $P' = F' \cup_{h'} M'$. Thus, if we can prove that P' can be embedded in some acyclic 3-manifold W, then P can be embedded in W. Note that each component of G' has connected boundary and that this implies that H' is connected. This means that we have reduced our problem to the case when one of G, H, say G, is connected. If we apply the procedure described above to this situation, we reduce the problem to the case when J is a separating simple closed curve in S.

Let us therefore assume that $J \subset S$ is a separating simple closed curve. Let A be a regular neighborhood of J in S. Denote by M' the 3-manifold obtained by attaching a 2-handle to M along A. Obviously P can be embedded in M'. By 3.7, M' is subacyclic and therefore P can be embedded in some acyclic 3-manifold.

Case 2. Suppose that J is not homologous to 0 in S.

If $J \sim 0$ in M, then P certainly cannot be embedded in any acyclic 3-manifold. Suppose that there exists an embedding $i\colon P \to W$ where W is an open acyclic 3-manifold. Let U be the closure of the component of $W - i(S)$ which contains $i(\operatorname{int} M)$ and let V be the closure of the other component of $W - i(S)$. Then $i(J)$ bounds in both U and V. But this contradicts 2.7 (1).

Now suppose that the homology class of J in M is equal to kx for some integer $k > 1$ and some nonzero $x \in H_1(M)$. In this case the image of x under $H_1(M) \to H_1(P)$ is a nonzero element of order k in $H_1(P)$. Since 3.1 implies that a compact subpolyhedron of an acyclic 3-manifold has free first homology group, P cannot be embedded in any acyclic 3-manifold.

Finally suppose that J represents a basic element of $H_1(M)$. Consider the Mayer-Vietoris sequence of $(P; M, F)$ for reduced homology:

$$0 \longrightarrow H_2(M) \longrightarrow H_2(P) \longrightarrow H_1(J) \xrightarrow{\alpha} H_1(M) \oplus H_1(F)$$
$$\xrightarrow{\beta} H_1(P) \longrightarrow \tilde{H}_0(J) \longrightarrow 0\ .$$

It is not difficult to prove that α embeds $H_1(J)$ as a direct summand into $H_1(M) \oplus H_1(F)$. Therefore, the homomorphism $H_2(P) \to H_1(J)$ is trivial and, as $H_1(M)$ and $H_1(F)$ are free, the image of β is free. It follows that $H_2(P) \approx H_2(M)$ and that $H_1(P) \approx \operatorname{Im} \beta \oplus \tilde{H}_0(J)$ is free.

Identify F with $F \times 0 \subset F \times I$. Give $\operatorname{Bd}(F \times I)$ the orientation induced by the orientation of F and choose an orientation for S. Extend h to an orientation reversing embedding $g\colon (\operatorname{Bd} F) \times I \to S$ and construct $V = (F \times I) \cup_g M$. Then V is a compact, connected, orientable 3-manifold containing P, and P is a deformation retract of V. It follows that $H_2(V) \approx H_2(M)$ and that $H_1(V)$ is free. Suppose that $\operatorname{Bd} M$ has m components. Then, by 3.1, $H_2(V) \approx H_2(M)$ is free of rank $m - 1$. This implies that $\operatorname{Bd} V$ has at most m components. On the other hand, $\operatorname{Bd} V$ has at least as many components as $\operatorname{Bd} M$. Thus V has exactly m boundary components. Now it follows from 3.1 that V is subacyclic. This concludes the proof of 3.2.

REFERENCES

1. P. J. Hilton and S. Wylie, *Homology Theory: An Introduction to Algebraic Topology,* Cambridge Univ. Press, Cambridge, 1960.
2. C. D. Papakyriakopoulos, *A reduction of the Poincaré conjecture to group theoretic conjectures,* Ann. of Math., (2) **77** (1963), 250-305.

Received July 17, 1972. Section 2 of this paper contains some of the results of the author's Ph. D. thesis written at the University of Wisconsin under the supervision of Professor J. W. Cannon. The author thanks the referee for his comments and suggestions.

UNIVERSITY OF LJUBLJANA, LJUBLJANA, YUGOSLAVIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

ADJOINT BOUNDARY VALUE PROBLEMS FOR COMPACTIFIED SINGULAR DIFFERENTIAL OPERATORS

PHILIP W. WALKER

This paper is concerned with differential operators and their adjoints induced in the Hilbert space $\mathscr{L}^2(w)$ by an operator $(1/w)l$ where l is an nth order singular differential operator and w is a weight. It is shown that weights may be chosen and boundary conditions may be imposed so that the structure of these operators is similar to that of regular differential operators.

1. Preliminaries. Throughout l will denote an operator of the form,

$$l(y) = y^{(n)} + \sum_{k=1}^{n} p_k y^{(n-k)} , \tag{1.1}$$

where each p_k is an $(n-k)$ times continuously differentiable complex valued function on an interval (a, b). *We allow $a = -\infty$ and/or $b = \infty$.* The formal adjoint of l will be denoted by l^+. Hence

$$l^+(y) = (-1)^n y^{(n)} + \sum_{k=1}^{n} (-1)^{n-k} (\bar{p}_k y)^{(n-k)}$$

for all n times differentiable y on (a, b).

If y is an $(n-1)$ times differentiable function then $k(y)$ will denote the vector valued function, column $(y, y', \cdots, y^{(n-1)})$, and if each of $y_1, y_2, \cdots, y_n$ is an $(n-1)$ times differentiable function then $K(y_1, \cdots, y_n)$ will denote the matrix valued function whose (i, j) entry is $y_j^{(i-1)}$ for $1 \leqq i, j \leqq n$.

$\mathscr{C}$ will denote the complex numbers, the space of all complex $n \times 1$ column vectors will be denoted by $\mathscr{C}^n$, and the space of all complex $n \times n$ matrices will be denoted by $\mathscr{M}^n$. If M is a matrix then M^* will denote its conjugate transpose.

DEFINITION 1.1. Let $(\varphi_1, \cdots, \varphi_n)$ be a sequence of linearly independent solutions to

$$l(y) = 0 \quad \text{on} \quad (a, b) . \tag{1.2}$$

The statement that $(\theta_1, \cdots, \theta_n)$ is the adjoint of $(\varphi_1, \cdots, \varphi_n)$ means that θ_k is the complex conjugate of the (k, n) entry of the matrix $[K(\varphi_1, \cdots, \varphi_n)]^{-1}$ for $k = 1, 2, \cdots, n$.

We shall make use of the following facts concerning adjoints of fundamental systems of solutions to Eq. (1.2).

LEMMA 1.2. *Let* $(\theta_1, \cdots, \theta_n)$ *be the adjoint of* $(\varphi_1, \cdots, \varphi_n)$, *a sequence of linearly independent solutions of Eq.* (1.2). *Let* $t_0 \in (a, b)$ *and let* $f\colon (a, b) \to \mathscr{C}$ *be Lebesgue integrable on every compact subinterval of* (a, b). *It follows that*

$$l(y) = f \text{ a.e. } \quad on \quad (a, b) \tag{1.3}$$

if and only if

$$\begin{aligned} k(y)(t) = K(\varphi_1, \cdots, \varphi_n)(t)\{&[K(\varphi_1, \cdots, \varphi_n)(t_0)]^{-1}k(y)(t_0) \\ &+ \int_{t_0}^{t} [(\theta_1, \theta_2, \cdots, \theta_n)(s)]^* f(s)ds\} \end{aligned} \tag{1.4}$$

for all t *in* (a, b).

This follows from consideration of the standard vector matrix formulation of Eq. (1.3) and from Eq. (3.2), p. 75 of [1].

LEMMA 1.3. *Let* φ_k *and* θ_k, $k = 1, \cdots, n$, *be as in Lemma* 1.2. *It follows that* $(\theta_1, \cdots, \theta_n)$ *is a linearly independent sequence of solutions to*

$$l^+(y) = 0 \quad on \quad (a, b)\ . \tag{1.5}$$

See problem 19, p. 101 of [1] and Theorem 5, p. 38 of [5]. Note that in the latter reference the formal adjoint differential operator is defined without taking complex conjugates. The same is true in Ref. [2] wherein on p. 69 in Corollary 3.8.2c we find justification for

LEMMA 1.4. *Let* φ_k *and* θ_k *be as in Lemma* 1.2. *Then*

$$[K(\theta_1, \cdots, \theta_n)]^* P[K(\varphi_1, \cdots, \varphi_n)] \equiv I \ on \ (a, b)\ .$$

Where I *is the* $n \times n$ *identity matrix and* P *is the concomitant matrix of* l.

See §3.7 of [2] or p. 285 of [1] (therein denoted B) for the definition of P.

By a weight we mean a positive real valued continuous function. If w is a weight on (a, b) then $\mathscr{L}^2(w)$ will denote the Hilbert space of all (equivalence classes of) Lebesgue measurable $f\colon (a, b) \to \mathscr{C}$ satisfying $\int_a^b |f|^2 w < \infty$. If $f, g \in \mathscr{L}^2(w)$ then $\langle f, g\rangle = \int_a^b f\bar{g}w$, and $||f|| = \sqrt{\langle f, f\rangle}$.

DEFINITION 1.5. The statement that w is a compactifying weight for l means that all solutions to Eq. (1.2) and the all solutions to Eq. (1.5) lie in $\mathscr{L}^2(w)$.

Since the solution spaces of Eqs. (1.2) and (1.3) are finite dimensional spaces of continuous functions it follows that every operator l has many compactifying weights. The reason for the terminology is that operators induced in $\mathscr{L}^2(w)$ by $(1/w)l$ and certain boundary conditions will have compact inverses.

The study of operators with a compactifying weight is in some sense complementary to the study of those with an l-admissible weight considered in [7].

2. Solutions of the eigenvalue equation. Our first theorem shows that solutions to differential equations with a compactifying weight behave in a manner similar to solutions of second order self-adjoint equations of the limit-circle type. (See §2 p. 225 of [1].)

THEOREM 2.1. *Let w be a compactifying weight for l. If $f \in \mathscr{L}^2(w)$ and $\lambda \in \mathscr{C}$ (γ may be real, even zero) then every solution to*

$$l(y) = w(\lambda y + f) \text{ a.e. } \quad on \quad (a, b) \tag{2.1}$$

lies in $\mathscr{L}^2(w)$.

Proof. Suppose that y satisfies Eq. (2.1). Let $t_0 \in (a, b)$ and let $(\varphi_1, \cdots, \varphi_n)$ and $(\theta_1, \cdots, \theta_n)$ be as in Lemma 1.2. Inspection of the first components of vector Eq. (1.4) shows that

$$y(t) = \varphi(t) + \sum_{k=1}^{n} \varphi_k(t) \int_{t_0}^{t} \overline{\theta_k(s)}(f(s) + \lambda y(s))w(s)ds$$

for all $t \in$ (a, b) where φ is a solution to Eq. (1.2). Thus for $t_0 \leqq t < b$ it follows from the Cauchy-Schwartz inequality that

$$|y(t)| \leqq |\varphi(t)| + \sum_{k=1}^{n} |\varphi_k(t)| \Big\{ ||\theta_k|| \cdot ||f|| + |\lambda| \left(\int_{t_0}^{t} |\theta_k(s)|^2 w(s)ds \right)^{1/2} \left(\int_{t_0}^{t} |y(s)|^2 w(s)ds \right)^{1/2} \Big\} .$$

Thus

$$|y(t)| \leqq |u(t)| + g(t) \left\{ \int_{t_0}^{t} |y(s)|^2 w(s)ds \right\}^{1/2}$$

for $t_0 \leqq t < b$ where $u = |\varphi| + \sum_{k=1}^{n} |\varphi_k| \cdot ||\theta_k|| \cdot ||f||$ and

$$g = \sum_{k=1}^{n} |\varphi_k| \cdot |\lambda| \cdot ||\theta_k|| .$$

Note that each of u and g is in $\mathscr{L}^2(w)$. Applying Theorem 1 of [4] with $G(t) = t^2$, and $\alpha(t) = \beta(t) \equiv 1/2$, we have that

$$\int_{t_0}^{t} |y(s)|^2 w(s) ds \leqq c \int_{t_0}^{t} |u(s)|^2 w(s) ds$$

for $t_0 \leqq t < b$ where $c = 2\exp(2\|g\|)$. A similar argument shows that for $a < \tau \leqq t_0$,

$$\int_{\tau}^{t_0} |y(s)|^2 w(s) ds \leqq c \int_{\tau}^{t_0} |u(s)|^2 w(s) ds .$$

Thus $y \in \mathscr{L}^2(w)$.

The next theorem provides a method for specifying initial conditions for the solutions of Eq. (2.1) at the endpoints of the interval (a, b).

THEOREM 2.2. *Let w be a compactifying weight for l, let $f \in \mathscr{L}^2(w)$ and let $\lambda \in \mathscr{C}$. Let $(\varphi_1, \cdots, \varphi_n)$ be a linearly independent sequence of solutions of Eq.* (1.2) *and let $\boldsymbol{Y} = \boldsymbol{K}(\varphi_1, \cdots, \varphi_n)$. If y is a solution to Eq.* (2.1) *then*

$$\lim_{t\to a} \boldsymbol{Y}^{-1}(t)\boldsymbol{k}(y)(t) \quad \text{and} \quad \lim_{t\to b} \boldsymbol{Y}^{-1}(t)\boldsymbol{k}(y)(t)$$

exist and are finite. Moreover, if $\boldsymbol{c} \in \mathscr{C}^n$ then there is exactly one solution y of Eq. (2.1) *satisfying*

$$\lim_{t\to a} \boldsymbol{Y}^{-1}(t)\boldsymbol{k}(y)(t) = \boldsymbol{c} , \tag{2.2}$$

and there is exactly one solution y of Eq. (2.1) *satisfying*

$$\lim_{t\to b} \boldsymbol{Y}^{-1}(t)\boldsymbol{k}(y)(t) = \boldsymbol{c} .$$

Proof. Let $(\theta_1, \cdots, \theta_n)$ be the adjoint of $(\varphi_1, \cdots, \varphi_n)$ and let $t_0 \in (a, b)$. From Eq. (1.4) it follows that if y satisfies Eq. (2.1) then

$$\begin{aligned} \boldsymbol{Y}^{-1}(t)\boldsymbol{k}(y)(t) &= \boldsymbol{Y}^{-1}(t_0)\boldsymbol{k}(y)(t_0) \\ &\quad + \int_{t_0}^{t} w(s)(f(s) + \lambda y(s))[(\theta_1, \cdots, \theta_n)(s).]^* ds \end{aligned}$$

for all t in (a, b). Since each of $\theta_1, \cdots, \theta_n$, f, and y (by Theorem 2.1) is in $\mathscr{L}^2(w)$ it follows that the limits indicated exist, and that Eq. (2.2) will be satisfied if and only if

$$\boldsymbol{k}(y)(t_0) = \boldsymbol{Y}(t_0)\left\{\boldsymbol{c} - \int_{t_0}^{a} w(s)(f(s) + \lambda y(s))[(\theta_1, \cdots (\theta_n)(s)]^* ds\right\} . \tag{2.3}$$

This is just a standard initial condition for solutions of Eq. (2.1); hence there is exactly one solution satisfying Eq. (2.3). The proof of the last assertion of the theorem is analogous.

3. Maximal and minimal operators. For each operator l and each compactifying weight w, D denotes the set of all functions y in $\mathscr{L}^2(w)$ which have (on each compact subinterval of (a, b)) an absolutely continuous $(n-1)$st derivative and which have the property that $(1/w)l(y)$ is in $\mathscr{L}^2(w)$. L (the maximal operator) denotes the restriction of $(1/w)l$ to D. D^+ and L^+ are defined in the same may with l replaced by l^+.

Let $(\varphi_1, \cdots, \varphi_n)$ be a linearly independent sequence of solutions to Eq. (1.2), and let $Y = \boldsymbol{K}(\varphi_1, \cdots, \varphi_n)$. D_0 denotes the set of all y in D satisfying

$$\lim_{t\to a} Y^{-1}(t)\boldsymbol{k}(y)(t) = 0 = \lim_{t\to b} Y^{-1}(t)\boldsymbol{k}(y)(t)\ . \tag{3.1}$$

Note that D_0 is independent of the fundamental system $(\varphi_1, \cdots, \varphi_n)$ which is used. (See Theorem 2.3 p. 70 of [1].) L_0 (the minimal operator) denotes the restriction of L to D_0. D_0^+ and L_0^+ are defined in the same way with Eq. (1.2), D, and L replaced respectively by Eq. (1.5), D^+, and L^+.

The main result of this section is presented in the following theorem. It is of interest to note that we are able, in the case of a compactifying weight, to deleniate the minimal operator through the boundary conditions (3.1); whereas in earlier treatments of similar problems, even with symmetric operators with maximal deficiency indices, (see §17 of [6] and §XIII. 2 of [3]). The minimal operator has been viewed less succinctly as the closure of what would correspond to the restriction of our operator L to function with compact support interior to (a, b). (See Corollary 3.5.)

THEOREM 3.1. *Let w be a compactifying weight for l. Then L_0 is a densely defined operator on $\mathscr{L}^2(w)$,*

$$L_0^* = L^+ \quad \textit{and} \quad (L^+)^* = L_0\ ,$$

where $*$ denotes the adjoint operator in $\mathscr{L}^2(w)$.

The proof of this theorem will require the following lemmas, some of which were motivated by the material in §17.3 of [6].

LEMMA 3.2. *Let w be a compactifying weight for l and let $f \in \mathscr{L}^2(w)$. There is exactly one solution y to*

$$l(y) = wf \text{ a.e. } \textit{on } (a, b) \tag{3.2}$$

lying in D_0 if and only if f is orthogonal to all solutions of $l^+(y) = 0$ on (a, b). Also $\mathscr{L}^2(w)$ is the orthogonal direct sum of range of L_0 and the null space of L^+.

Proof. Using the notation of Theorem 2.2, let y be the solution of Eq. (3.2) satisfying

$$\lim_{t\to a} \boldsymbol{Y}^{-1}(t)\boldsymbol{k}(y)(t) = \mathbf{0} .$$

By Theorem 2.1, y is in $\mathscr{L}^2(w)$. Let $\lambda = 0$, and $\boldsymbol{c} = \mathbf{0}$ in Eq. (2.3); multiplying both sides of this equation by $\boldsymbol{Y}^{-1}(t_0)$, and taking the limit as $t_0 \to b$ we see that y will also satisfy

$$\lim_{t\to b} \boldsymbol{Y}^{-1}(t)\boldsymbol{k}(y)(t) = \mathbf{0} ,$$

hence be in D_0, if and only if

$$\mathbf{0} = \text{column}\,(\langle f, \theta_1\rangle, \cdots, \langle f, \theta_n\rangle) .$$

In view of Lemma 1.3 the first assertion is proved. Since the null space of L^+ is of finite dimension, the Hilbert space $\mathscr{L}^2(w)$ is the orthogonal direct sum of it and its orthogonal complement. We have shown that this orthogonal complement is the range of L_0.

Lemma 1.4 and Theorem 2.2 allow us to give a particularly simple expression for Lagrange's identity. Note that if w is a compactifying weight for l then it is also a compactifying weight for l^+. Hence by Theorem 2.2 the vectors $\boldsymbol{z}_a$ and $\boldsymbol{z}_b$ defined below do exist.

LEMMA 3.3. *Let w be a compactifying weight for l. Let $(\varphi_1, \cdots, \varphi_n)$ be a linearly independent sequence of solutions to Eq.* (1.2) *and let $(\theta_1, \cdots, \theta_n)$ be the adjoint of this sequence. For each $y \in D$ and $z \in D^+$ let*

$$\boldsymbol{y}_a = \lim_{t\to a} [\boldsymbol{K}(\varphi_1, \cdots, \varphi_n)(t)]^{-1}\boldsymbol{k}(y)(t)$$

and

$$\boldsymbol{z}_a = \lim_{t\to a} [\boldsymbol{K}(\theta_1, \cdots, \theta_n)(t)]^{-1}\boldsymbol{k}(z)(t) ,$$

and let $\boldsymbol{y}_b$ and $\boldsymbol{z}_b$ be defined in the same way taking the limits at b rather than at a.

It follows that if $y \in D$ and $z \in D^+$ then

$$\langle Ly, z\rangle - \langle y, L^+z\rangle = \boldsymbol{z}_b^*\boldsymbol{y}_b - \boldsymbol{z}_a^*\boldsymbol{y}_a .$$

Proof. If $a < \alpha < \beta < b$ then

$$\begin{aligned}\int_\alpha^\beta \left(\frac{1}{w}\right) l(y)\bar{z}w &- \int_\alpha^\beta \left(\frac{1}{w}\right)\overline{l^+(z)}yw \\ &= \int_\alpha^\beta [l(y)\bar{z} - y\overline{l^+(z)}] \\ &= \{[\boldsymbol{k}(z)]^*\boldsymbol{P}\boldsymbol{k}(y)\}\big|_\alpha^\beta\end{aligned}$$

where P is the concomitant matrix for l. (See pp. 86 and 285 of [1].) In view of Lemma 1.4 this last expression is the same as

$$\{[[\boldsymbol{K}(\theta_1, \cdots, \theta_n)]^{-1}\boldsymbol{k}(z)]^*[\boldsymbol{K}(\varphi_1, \cdots, \varphi_n)]^{-1}\boldsymbol{k}(y)\}|_\alpha^\beta .$$

The conclusion to the lemma then follows by taking limits as $\beta \to b$ and $\alpha \to a$.

LEMMA 3.4. *If the hypotheses of Lemma* 3.3 *are satisfied and each of* $\boldsymbol{c}_1$ *and* $\boldsymbol{c}_2$ *is in* $\mathscr{C}^n$ *then there is a* $y \in D$ *satisfying*

$$\boldsymbol{y}_a = \boldsymbol{c}_1 \quad \textit{and} \quad \boldsymbol{y}_b = \boldsymbol{c}_2$$

and there is a $z \in D$ *satisfying*

$$\boldsymbol{z}_a = \boldsymbol{c}_1 \quad \textit{and} \quad \boldsymbol{z}_b = \boldsymbol{c}_2 .$$

Proof. We shall show that there is a $u \in D$ such that $\boldsymbol{u}_a = \boldsymbol{c}_1$ and $\boldsymbol{u}_b = \boldsymbol{0}$. A similar argument would show that there is a $\boldsymbol{v} \in D$ such that $\boldsymbol{v}_a = 0$ and $\boldsymbol{v}_b = \boldsymbol{c}_2$; then $y = u + v$ will satisfy the conclusion to the lemma.

Let z_j be the solution to $l^+(y) = 0$ on (a, b) that

$$\boldsymbol{z}_{ja} = \boldsymbol{e}_{ij}$$

for $j = 1,2, \cdots, n$ where $\boldsymbol{e}_{ij}$ is the $n \times 1$ matrix with (i, j) entry 1 if $i = j$ and 0 otherwise. Since

$$[\boldsymbol{K}(\theta_1, \cdots, \theta_n)]^{-1}\boldsymbol{K}(z_1, \cdots, z_n)$$

has the limit $\boldsymbol{I}$ (the $n \times n$ identity matrix) at a, it follows that $\boldsymbol{K}(z_1, \cdots, z_n)$ is nonsingular at some (hence all points) point in (a, b). Thus $z_1, \cdots, z_n$ are linearly independent and their Gram determinate (computed with respect to the inner product of $\mathscr{L}^2(w)$) is nonzero. In view of this fact we may let f be the linear combination of $z_1, \cdots, z_n$ such that

$$\text{column}\,(\langle f, z_1\rangle, \cdots, \langle f, z_n\rangle) = -\,\boldsymbol{c}_1 .$$

By Theorem 2.2 we may let u be the element in D such that $Lu = f$ and $\boldsymbol{y}_b = 0$. By Lemma 3.3 it follows that

$$\langle f, z_j\rangle = \langle Lu, z_j\rangle = \langle u, L^+z_j\rangle - \boldsymbol{z}_{ja}^*\boldsymbol{u}_a ,$$

and since $L^+z_j = 0$ for $j = 1, 2, \cdots, n$ and $\boldsymbol{z}_{ja} = \boldsymbol{e}_{ij}$ we have that $\boldsymbol{u}_a = \boldsymbol{c}_1$. The argument for the existence of the $z \in D^+$ is similar.

Proof of Theorem 3.1. That D_0 is dense in $\mathscr{L}^2(w)$ follows from the fact that D_0 contains all n times continuously differentiable functions with compact support interior to (a, b).

For the remainder of the proof we will adopt the notation of Lemma 3.3.

If $y \in D_0$ and $z \in D^+$ then $\boldsymbol{y}_a = \boldsymbol{0} = \boldsymbol{y}_b$ hence by Lemma 3.3,

$$\langle L_0 y, z\rangle = \langle Ly, z\rangle = \langle y, L^+ z\rangle .$$

Thus $L^+ \subseteqq L_0^*$.

Suppose that $g \in L_0^*$. Let $h = L_0^* g$ and let z be any element of D^+ satisfying $l^+(z) = wh$ a.e. on (a, b). (See Theorem 2.1.) If $y \in D_0$ it follows from Lemma 3.3 that $\langle y, h\rangle = \langle y, L^+ z\rangle = \langle L_0 y, z\rangle$ and it follows from the definition of the adjoint operator that $\langle y, h\rangle = \langle y, L_0^* g\rangle = \langle L_0 y, g\rangle$. Hence $\langle L_0 y, z - g\rangle = 0$ for all $y \in D_0$. From Lemma 3.2 we have that $\mathscr{L}^2(w)$ is the orthogonal direct sum of the range of L_0 and the null space of L^+. Thus $z - g$ (after modification on a set of measure zero) is in the null space of L^+. In particular $z - g \in D^+$ and since $z \in D^+$ it follows that $g \in D^+$. Since $L^+ g = L^+ z$ and $L^+ z = h = L_0^* g$ it follows that $L_0^* \subseteqq L^+$. Hence the fact that $L_0^* = L^+$ has been established.

From $L_0^* = L^+$ we have that $L_0^{**} = (L^+)^*$ and since $A \subseteqq A^{**}$ for any densely defined operator A it follows that $L_0 \subseteqq (L^+)^*$.

Applying the part of Theorem 3.1 that has been proved with l replaced by l^+ we find that $(L_0^+)^* = L$. Since $L_0^+ \subseteqq L^+$ implying $(L^+)^* \subseteqq (L_0^+)^*$ it follows that $(L^+)^* \subseteqq L$. Thus if $y \in (L^+)^*$ then $y \in D$ and $(L^+)^* y = Ly$; and if $z \in D^+$, by definition of adjoint, we have

$$\langle y, L^+ z\rangle = \langle (L^+)^* y, z\rangle$$

or

$$\langle y, L^+ z\rangle = \langle Ly, z\rangle .$$

On the other hand, by Lemma 3.3 it follows that

$$\langle Ly, z\rangle = \langle y, L^+ z\rangle + \boldsymbol{z}_b^* \boldsymbol{y}_b - \boldsymbol{z}_a^* \boldsymbol{y}_a .$$

Thus $\boldsymbol{z}_b^* \boldsymbol{y}_b - \boldsymbol{z}_a^* \boldsymbol{y}_a = 0$ for all $z \in D^+$. Since by Lemma 3.4 there is a $z \in D^+$ such that $\boldsymbol{z}_a$ and $\boldsymbol{z}_b$ have any preassigned values it follows that $\boldsymbol{y}_a = \boldsymbol{y}_b = 0$. Since we already have $y \in D$ and $(L^+)^* y = Ly$ it follows that $y \in D_0$ and $(L^+)^* y = L_0 y$. Thus $(L^+)^* \subseteqq L_0$. This completes the proof of the fact that $(L^+)^* = L_0$.

COROLLARY 3.5. *The operator L_0 is closed in $\mathscr{L}^2(w)$.*

Proof. The adjoint of any densely defined operator is closed and by Theorem 3.1,

$$L_0^{**} = (L^+)^* = L_0 .$$

4. Intermediate operators and their adjoints. In this section we shall consider operators which lie between the maximal and minimal operators and their adjoints. We shall continue to use the notation developed in §3 and assume that all our operators are based on an nth order operator l with a compactifying weight. Furthermore, all vectors $\boldsymbol{y}_a$, $\boldsymbol{y}_b$, z_a, and z_b are to be formed using an arbitrary but fixed sequence $(\varphi_1, \cdots, \varphi_n)$ (of linearly independent solutions to $l(y) = 0$) and its adjoint. (See Lemma 3.3.)

If each of $\boldsymbol{M}$ and $\boldsymbol{N}$ is in $\mathscr{M}^n$ and $\boldsymbol{B}$ is the $n \times 2n$ matrix $(\boldsymbol{M}: \boldsymbol{N})$ then $D_{\boldsymbol{B}}$ will denote the set of all $y \in D$ such that

$$\boldsymbol{M}\boldsymbol{y}_a + \boldsymbol{N}\boldsymbol{y}_b = \boldsymbol{0} \tag{4.1}$$

and $D_{\boldsymbol{B}}^+$ will denote the set of all $z \in D^+$ such that

$$z_a^* = \boldsymbol{c}^*\boldsymbol{M} \quad \text{and} \quad z_b^* = -\boldsymbol{c}^*\boldsymbol{N} \quad \text{for some} \quad \boldsymbol{c} \in \mathscr{C}^n . \tag{4.2}$$

$L_{\boldsymbol{B}}$ and $L_{\boldsymbol{B}}^+$ will denote the restrictions of L and L^+ to $D_{\boldsymbol{B}}$ and $D_{\boldsymbol{B}}^+$ respectively.

The following theorem shows that the boundary conditions 4.1 and 4.2 deleneate mutually adjoint operators in $\mathscr{L}^2(w)$.

THEOREM 4.1. *If each of $\boldsymbol{M}$ and $\boldsymbol{N}$ is in $\mathscr{M}^n$ then $(L_{\boldsymbol{B}})^* = L_{\boldsymbol{B}}^+$ and $(L_{\boldsymbol{B}}^+)^* = L_{\boldsymbol{B}}$.*

Proof. By Lemma 3.3, if $y \in D_{\boldsymbol{B}}$ and $z \in D_{\boldsymbol{B}}^+$ then

$$\langle L_{\boldsymbol{B}}y, z\rangle - \langle y, L_{\boldsymbol{B}}^+z\rangle = z_b^*\boldsymbol{y}_b - z_a^*\boldsymbol{y}_a , \tag{4.3}$$

and from 4.1 and 4.2 it follows that the right hand side of this equation is zero. Thus $L_{\boldsymbol{B}}^+ \subseteqq L_{\boldsymbol{B}}^*$.

By its definition we have that $L_0 \subseteqq L_{\boldsymbol{B}}$, hence $L_{\boldsymbol{B}}^* \subseteqq L_0^*$ so from Theorem 3.1 we have that $L_{\boldsymbol{B}}^* \subseteqq L^+$. Thus $L_{\boldsymbol{B}}^*z = L^+z$ for all z in the domain of $L_{\boldsymbol{B}}$. Suppose now that z is in the domain of $L_{\boldsymbol{B}}^*$. Then, by definition of adjoint $\langle Ly, z\rangle = \langle L_{\boldsymbol{B}}y, z\rangle = \langle y, L_{\boldsymbol{B}}^*, z\rangle = \langle y, L^+z\rangle$, for all y in $D_{\boldsymbol{B}}$. On the other hand, by Lemma 3.3 we have that

$$\langle Ly, z\rangle - \langle y, L^+z\rangle = z_b^*\boldsymbol{y}_b - z_a^*\boldsymbol{y}_a .$$

Hence $z_b^*\boldsymbol{y}_b - z_a^*\boldsymbol{y}_a = 0$ for all $y \in D_{\boldsymbol{B}}$.
Or the vector $\begin{bmatrix} z_a \\ -z_b \end{bmatrix}$ is orthogonal in $\mathscr{C}^{2n}$ (with respect to the standard inner product) to the subspace of all vectors $\begin{bmatrix} \boldsymbol{y}_a \\ \boldsymbol{y}_b \end{bmatrix}$ such that $y \in D_{\boldsymbol{B}}$. We denote this subspace by V. In view of Lemma 3.4 V is the set of all vectors $\begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix}$ such that

$$Mu + Nv = 0 .$$

Therefore, another way to view V is that it is the orthogonal complement in $\mathscr{C}^{2n}$ of the column space of $\begin{bmatrix} M^* \\ N^* \end{bmatrix}$. Hence $\begin{bmatrix} z_a \\ -z_b \end{bmatrix}$ must be in this column space or

$$\begin{bmatrix} z_a \\ -z_b \end{bmatrix} = \begin{bmatrix} M^* \\ N^* \end{bmatrix} c \text{ for some } c \in \mathscr{C}^n .$$

Thus condition 4.2 is satisfied and $z \in D_B^+$. We have shown then that $L_B^* z = L^+ z$ for all z in the domain of L_B^* and that this domain is a subset of D_B^+. Thus $L_B^* \subseteqq L_B^+$. This completes the proof of the first assertion of the theorem.

Again conditions 4.1 and 4.2 imply that the right hand side of equation 4.3 is zero when $y \in D_B$ and $z \in D_B^+$. Thus $L_B \subseteqq (L_B^+)^*$. Also from its definition $L_0^+ \subseteqq L_B^+$, hence $(L_B^+)^* \subseteqq (L_0^+)^*$; so by Theorem 3.1 applied to l^+ we have that $(L_B^+)^* \subseteqq L$. If y is in the domain of $(L_B^+)^*$ then

$$\langle Ly, z\rangle - \langle (L_B^+)^* y, z\rangle = \langle y, L_B^+ z\rangle = \langle y, L^+ z\rangle$$

for all $z \in D_B^+$ and from Lemma 3.3

$$\langle Ly, z\rangle - \langle y, L^+ z\rangle = z_b^* y_b - z_a^* y_a .$$

Thus $z_b^* y_b - z_a^* y_a = 0$ for all $z \in D_B^+$ or the vector $\begin{bmatrix} y_a \\ y_b \end{bmatrix}$ is orthogonal in $\mathscr{C}^{2n}$ to the subspace of all vectors $\begin{bmatrix} z_a \\ -z_b \end{bmatrix}$ such that $z \in D_B^+$. We denote this subspace by W. Again by Lemma 3.4 we conclude that W is the set of all vectors $\begin{bmatrix} u \\ v \end{bmatrix}$ such that

$$u^* = c^* M \quad \text{and} \quad v^* = c^* N$$

for some $c \in \mathscr{C}^n$ or that W is the column space of the matrix $\begin{bmatrix} M^* \\ N^* \end{bmatrix}$. Since $\begin{bmatrix} y_a \\ y_b \end{bmatrix}$ is orthogonal to W we have that $My_a + Ny_b = 0$. Thus $y \in D_B$ and we have completed the argument that $(D_B^+)^* \subseteqq D_B$, and from $(L_B^+)^* \subseteqq L$ we have that $(L_B^+)^* \subseteqq L_B$. Thus $(L_B^+)^* = L_B$.

The next theorem shows that boundary conditions of the type 4.2 can be expressed by conditions of the type 4.1 and conversely.

THEOREM 4.2. *Suppose that* $M, N \in \mathscr{M}^n$ *and that* m *is the column rank of* $\begin{bmatrix} M^* \\ -N^* \end{bmatrix}$. *Let* D *be a* $2n \times (2n - m)$ *matrix whose columns form a basis in* $\mathscr{C}^{2n}$ *for the orthogonal complement of the column space of* $\begin{bmatrix} M^* \\ -N^* \end{bmatrix}$ *and let* P *and* Q *be the* $n \times (2n - m)$ *matrices such* *that* $\boldsymbol{D} = \begin{bmatrix} \boldsymbol{P} \\ \boldsymbol{Q} \end{bmatrix}$. *It follows that* $z \in D^+$ *satisfies condition* 4.2 *if and only if*

$$\boldsymbol{P}^* z_a + \boldsymbol{Q}^* z_b = 0 , \tag{4.4}$$

and it follows that $y \in D$ *satisfies condition* 4.1 *if and only if*

$$\boldsymbol{y}_a^* = \boldsymbol{c}^* \boldsymbol{P}^* \quad \textit{and} \quad \boldsymbol{y}_b^* = -\boldsymbol{c}^* \boldsymbol{Q}^* \quad \textit{for some} \quad \boldsymbol{c} \in \mathscr{C}^n .$$

Proof. $z \in D^+$ satisfies 4.2 if and only if $\begin{bmatrix} z_a \\ z_b \end{bmatrix} = \begin{bmatrix} \boldsymbol{M}^* \\ -\boldsymbol{N}^* \end{bmatrix} \boldsymbol{c}$ for some $\boldsymbol{c} \in \mathscr{C}^n$. This holds if and only if $\begin{bmatrix} z_a \\ z_b \end{bmatrix}$ is in the column space of $\begin{bmatrix} \boldsymbol{M}^* \\ -\boldsymbol{N}^* \end{bmatrix}$ and this is equivalent to $\begin{bmatrix} z_a \\ z_b \end{bmatrix}$ being orthogonal to the orthogonal to the orthogonal complement of the column space of $\begin{bmatrix} \boldsymbol{M}^* \\ -\boldsymbol{N}^* \end{bmatrix}$. Eq. (4.4) is simply another way of stating that $\begin{bmatrix} z_a \\ z_b \end{bmatrix}$ is in the orthogonal complement of the column space of $\boldsymbol{D}$. The argument for the second assertion of the theorem is similar.

5. **Invertibility and Green's functions.** In this section we give a necessary and sufficient condition for the operator L_B, defined in §4, to be invertible and show how the inverse operator, when it exists, may be expressed as an integral operator of the Hilbert-Schmidt type.

THEOREM 5.1. *Let* $\boldsymbol{M}, \boldsymbol{N} \in \mathscr{M}^n$, *let* $\boldsymbol{B} = (\boldsymbol{M} \colon \boldsymbol{N})$, *and let* L_B *be as in* §4. *It follows that* L_B *is invertible if and only if the matrix* $\boldsymbol{M} + \boldsymbol{N}$ *is nonsingular.*

Proof. Since L_B is linear it is invertible if and only if the only solution to $L_B y = 0$ is the zero function. $L_B y = 0$ if and only if y satisfies the boundary condition 4.1 and y is a linear combination of the same sequence of solution $(\varphi_1, \cdots, \varphi_n)$ used to construct $\boldsymbol{y}_a$ and $\boldsymbol{y}_b$. Thus $L_B y = 0$ if and only if

$$\begin{aligned} &\boldsymbol{M} \lim_{t \to a} [\boldsymbol{K}(\varphi_1, \cdots, \varphi_n)(t)]^{-1} [\boldsymbol{K}(\varphi_1, \cdots, \varphi_n)(t)] \boldsymbol{c} \\ &+ \boldsymbol{N} \lim_{t \to b} [\boldsymbol{K}(\varphi_1, \cdots, \varphi_n)(t)]^{-1} [\boldsymbol{K}(\varphi_1, \cdots, \varphi_n)(t)] \boldsymbol{c} = 0 \\ &\text{or} \quad (\boldsymbol{M} + \boldsymbol{N}) \boldsymbol{c} = 0 \end{aligned} \tag{5.1}$$

where $\boldsymbol{c}$ is the vector in $\mathscr{C}^n$ such that

$$y = (\varphi_1, \cdots, \varphi_n) \boldsymbol{c} .$$

Since Eq. 5.1 is satisfied only for $\boldsymbol{c} = \boldsymbol{0}$ if and only if $\boldsymbol{M} + \boldsymbol{N}$ is nonsingular the theorem is proved.

THEOREM 5.2. *Let $M, N \in \mathscr{M}^n$, let $B = (M: N)$, let L_B be as in § 4, and suppose that L_B is invertible. If $f \in \mathscr{L}^2(w)$ then $y \in D_B$ and $L_B y = f$ if and only if $y(t) = \int_a^b G(t, s) f(s) w(s) ds$ for all $t \in (a, b)$ where*

$$G(t, s) = \begin{cases} [(\varphi_1, \cdots, \varphi_n)(t)](M + N)^{-1} M[(\theta_1, \cdots, \theta_n)(s)]^* \\ \quad \text{for } a < s < t < b \\ -[(\varphi_1, \cdots, \varphi_n)(t)](M + N)^{-1} N[(\theta_1, \cdots, \theta_n)(s)]^* \\ \quad \text{for } a < t < s < b \end{cases}$$

wherein $(\varphi_1, \cdots, \varphi_n)$ is a linearly independent sequence of solutions to $l(y) = 0$ on (a, b) and $(\theta_1, \cdots, \theta_n)$ is its adjoint.

Proof. $y \in D_B$ and $L_B y = f$ if and only if condition 4.1 holds and $l(y) = wf$ a.e. on (a, b). By Lemma 1.2 we see that this last differential equation holds if and only if

$$\begin{aligned} &[K(\varphi_1, \cdots, \varphi_n)(\tau)]^{-1} k(y)(\tau) \\ &= \Big\{[K(\varphi_1, \cdots, \varphi_n)(t)]^{-1} k(y)(t) \\ &\quad + \int_t^\tau [(\theta_1, \cdots, \theta_n)(s)]^* f(s) w(s) ds\Big\} \end{aligned}$$

whenever $t, \tau \in (a, b)$. Using the fact that each θ_k and f is in $\mathscr{L}^2(w)$ we may conclude that if $l(y) = wf$ a.e. on (a, b) then

$$\begin{aligned} y_a = &[K(\varphi_1, \cdots, \varphi_n)(t)]^{-1} k(y)(t) \\ &- \int_a^t [(\theta_1, \cdots, \theta_n)(s)]^* f(s) w(s) ds \end{aligned}$$

and

$$\begin{aligned} y_b = &[K(\varphi_1, \cdots, \varphi_n)(t)]^{-1} k(y)(t) \\ &+ \int_t^b [(\theta_1, \cdots, \theta_n)(s)]^* f(s) w(s) ds \end{aligned}$$

for all t in (a, b). Multiplying the first of these equations (on the left) by M and the second by N and adding we see that if $l(y) = f$ a.e. on (a, b) and 4.1 is satisfied then

$$\begin{aligned} &(M + N)[K(\varphi_1, \cdots, \varphi_n)(t)]^{-1} k(y)(t) \\ &= M \int_a^t [(\theta_1, \cdots, \theta_n)(s)]^* f(s) w(s) ds \\ &\quad - N \int_t^b [(\theta_1, \cdots, \theta_n)(s)]^* f(s) w(s) ds\,. \end{aligned} \tag{5.2}$$

Using the fact that $M + N$ is nonsingular (see Theorem 5.1), solving Eq. (5.2) for $k(y)(t)$, and examining the first components of the resultant equation we see that the integral equation indicated in the theorem is satisfied.

If the integral equation in the theorem is satisfied then differentiating we find that

$$
\begin{aligned}
y'(t) &= \sum_{k=1}^{n} \varphi_k(t)\overline{\theta_k(t)} f(t) w(t) \\
&\quad + \int_a^t [(\varphi_1, \cdots, \varphi_n)'(M + N)^{-1} M [(\theta_1, \cdots, \theta_n)]^* f(s) w(s) ds \qquad (5.3)\\
&\quad - \int_t^b [(\varphi_1, \cdots, \varphi_n)'(M + N)^{-1} N [(\theta_1, \cdots, \theta_n)]^* f(s) w(s) ds
\end{aligned}
$$

for all t in (a, b). Returning to Definition 1.1 we see that $\sum_{k=1}^{n} \varphi_k^{(j)} \bar{\theta}_k$ is the $(j + 1, n)$ entry of the $n \times n$ identity matrix. In case $n = 1$ Eq. (5.2) is immediate from the integral equation of the theorem, and in case $n \geqq 2$ the last observation and continued differentiation of Eq. (5.3) shows that Eq. (5.2) is satisfied. Taking the limits as $t \to a$ and as $t \to b$ in Eq. (5.2) we find that

$$
My_a + Ny_b = [- M(M + N)^{-1} N + N(M + N)^{-1} M] \mathscr{F} \qquad (5.4)
$$

where

$$
\mathscr{F} = \int_a^b [(\theta_1, \cdots, \theta_n)(s)]^* f(s) w(s) ds ,
$$

and adding and subtracting $M(M + N)^{-1} M$ in the term in brackets on the right side of Eq. (5.4) we see that condition 4.1 is satisfied.

Returning to Eq. (5.2), if we add and subtract

$$
N \int_a^t [(\theta_1, \cdots, \theta_n)(s)]^* f(s) w(s) ds
$$

on the right hand side we find that

$$
\begin{aligned}
k(y)(t) = K(\varphi_1, \cdots, \varphi_n)(t) \Big[&- (M + N)^{-1} N \mathscr{F} \\
&+ \int_a^t [(\theta_1, \cdots, \theta_n)(s)]^* f(s) w(s) ds \Big]
\end{aligned}
$$

for all t in (a, b). Letting t_0 be a point in (a, b) and adding and subtracting

$$
\int_{t_0}^a [(\theta_1, \cdots, \theta_n)(s)]^* f(s) w(s) ds
$$

in the term in brackets in the last equation we see that

$$
k(y)(t) = K(\varphi_1, \cdots, \varphi_n)(t) \Big[c + \int_{t_0}^t [(\varphi_1, \cdots, \varphi_n)(s)]^* f(s) w(s) ds \Big]
$$

for all $t \in (a, b)$ where c is a constant vector in $\mathscr{C}^n$. Thus by (a slight modification of) Lemma 1.2 y is a solution to $l(y) = wf$ a.e. on (a, b). Using Theorem 2.1 we may now conclude that $y \in D_B$ and $L_B y = f$.

References

1. E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1965.

2. R. H. Cole, *Theory of Ordinary Differential Equations*, Appleton-Century-Crofts, New York, 1968.

3. N. Dunford and J. T. Schwartz, *Linear Operators*, Part II, Interscience-Wiley, New York, and London, 1963.

4. H. E. Gollwitzer, *A note on a functional inequality*, Proc. Amer. Math. Soc., **23** (1969), 642–647.

5. K. S. Miller, *Linear Differential Equations in the Real Domain*, Norton, New York, 1963.

6. M. A. Naimark, *Linear Differential Operators*, Part II, Frederick Ungar, New York, 1968.

7. P. W. Walker, *Asympotics for a class of weighted eigenvalue problems*, Pacific J. Math., **40** (1972), 501–510.

Received August 18, 1972.

VIRGINIA POLYTECHNIC INSTITUTE
AND
STATE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

WHEN ARE WITT RINGS GROUP RINGS?

ROGER WARE

It has been shown that if C is a commutative connected semi-local ring with involution J then the Witt ring, $W(C, J)$, of hermitian forms over C is a factor ring of an integral group ring $Z[G]$, with G a group of exponent two. The purpose of this note is to characterize those pairs (C, J) whose Witt rings are actually isomorphic to integral group rings (Theorem 1).

I would like to express my thanks to Alex Rosenberg and Manfred Knebusch for several helpful suggestions.

This paper is in part motivated by the result of Elman and Lam which states that if F is a superpythagorean field [3, Th. 4.3, Def. 4.4] then the Witt ring, $W(F)$, of F is isomorphic to a group ring $Z[H]$, where H can be taken to be any subgroup of F^*/F^{*2} of index two, not containing the square class of -1 [3, Th. 5.13 (8)]. In Theorem 1 a different proof of the Elman-Lam result is given and it is shown that the converse is also true. In order to extend the notion of superpythagorean to semi-local rings, we employ the concept of signature as defined in [6].

In what follows C will always be a commutative connected ($=$ no idempotents other than 0 and 1) semi-local ring with involution J and A will be the fixed ring of J. We allow the possibility that J is the identity. The groups of units of C and A are denoted by C^* and A^* respectively, and $N: C^* \rightarrow A^*$ is the homomorphism given by $N(c) = cJ(c)$. We denote by $W(C, J)$ the Witt ring of hermitian spaces over C with respect to the involution J, as defined in [5]. The ring theoretic operations of $W(C, J)$ are induced by the orthogonal direct sum and tensor product of spaces respectively. For a in A^* we let $\langle a \rangle$ denote the class in $W(C, J)$ of the rank one hermitian space C with form $(c_1, c_2) \rightarrow c_1 J(c_2) a$ and $[a]$ the image of a in the group A^*/NC^*. Then $\langle a \rangle = \langle b \rangle$ in $W(C, J)$ if and only if $[a] = [b]$ in A^*/NC^* and $\langle a \rangle \langle b \rangle = \langle ab \rangle$. Hence the assignment $[a] \rightarrow \langle a \rangle$ induces a ring homomorphism $\psi: Z[A^*/NC^*] \rightarrow W(C, J)$. By [5, Th. 1.16], the mapping ψ is surjective.

A signature σ of (C, J) is a group homomorphism $\sigma: A^* \rightarrow \{\pm 1\}$ with the property that $\sigma(NC^*) = 1$ and if $\sigma: Z[A^*/NC^*] \rightarrow Z$ also denotes the induced ring homomorphism then $\sigma(\operatorname{Ker} \psi) = 0$. As remarked in [6], the signatures of (C, J) correspond bijectively with the ring homomorphisms from $W(C, J)$ to Z. By [5, Example 3.11] the latter set is in bijective correspondence with the set of non-maximal prime ideals of $W(C, J)$. If J is the identity and $C = A$ is a field then the signatures of C correspond to the (total) orderings on C (cf. [6, Remark 1.7 (ii)]).

Since the kernel of the natural map $\psi: \boldsymbol{Z}[A^*/NC^*] \to W(C, J)$ contains the element $[1] + [-1]$, [5, Cor. 1.17], it follows that any signature σ of (C, J) has the property that $\sigma(-1) = -1$. Suppose, in addition, (C, J) has the following property

(*) C has no maximal ideal M with $J(M) = M$ such that either $C/M = \boldsymbol{F}_2$ or $C/M = \boldsymbol{F}_4$ and $A/M \cap A = \boldsymbol{F}_2$ ($\boldsymbol{F}_n$ = finite field with n elements).

Then, by [6, Prop. 1.4], a homomorphism $\sigma: A^* \to \{\pm 1\}$ with $\sigma(NC^*) = 1$ and $\sigma(-1) = -1$ is a signature if and only if $\sigma(a) = 1$ implies $\sigma(N(c_1) + aN(c_2)) = 1$ for any c_1, c_2 in C with $N(c_1) + aN(c_2)$ in A^*.

The main result is the following.

THEOREM 1. *Assume C has property* (*) *and* -1 *is not in* NC^*. *Then the following statements are equivalent*:

(i) *For any a in A^* with $a \notin -NC^*$ we have* $(NC + aNC) \cap A^* = NC^* \cup aNC^*$.

(ii) *If $\sigma: A^* \to \{\pm 1\}$ is a homomorphism such that $\sigma(NC^*) = 1$ and $\sigma(-1) = -1$ then σ is a signature of (C, J).*

(iii) *If E is a finite subgroup of A^*/NC^* not containing the norm class $[-1]$ then there exists a signature σ of (C, J) such that $\sigma(E) = 1$.*

(iv) *If H is any subgroup of A^*/NC^* not containing* $[-1]$ *then there exists a signature σ such that $\sigma(H) = 1$.*

(v) *The kernel of the mapping $\psi: \boldsymbol{Z}[A^*/NC^*] \to W(C, J)$ is the ideal generated by* $[1] + [-1]$.

(vi) *$W(C, J) \cong \boldsymbol{Z}[H]$ where H is a subgroup of index two in A^*/NC^*. The group H can be taken to be any subgroup of index two not containing* $[-1]$.

(vii) *$W(C, J) \cong \boldsymbol{Z}[H]$ for some group H of exponent two.*

Proof. (i) $\Rightarrow$ (ii) As mentioned above it is enongh to show that if a is a unit of A with $\sigma(a) = 1$ and c_1, c_2 are elements of C such that $b = N(c_1) + aN(c_2)$ is also a unit then $\sigma(b) = 1$. Since $\sigma(a) = 1$ it follows that $a \notin -NC^*$. Hence by (i), b lies in $NC^* \cup aNC^*$ so that $\sigma(b) = 1$, as desired.

(ii) $\Rightarrow$ (iii) is clear.

(iii) $\Rightarrow$ (iv). Let Sign (C, J) denote the set of signatures of (C, J) and for a in A^*, let $V(a) = \{\sigma \text{ in Sign}(C, J) \mid \sigma(a) = 1\}$. The sets $V(a)$, a in A^*, can be taken as a subbase for a topology on Sign (C, J) which makes Sign (C, J) a compact Hausdorff space and each $V(a)$ a closed set [6, Lemma 2.3, Th. 2.18, Lemma 3.3 (i)]. Now let H be a subgroup of A^*/NC^* with $[-1] \notin H$ and choose $\{a_i\}_{i \in I} \subset A^*$ such that $H = \{[a_i]\}_{i \in I}$.

For any finite subset $I_0 \subset I$, the group H_0 generated by $\{[a_i]\}_{i \in I_0}$ is finite and $[-1] \notin H_0$. Hence, by (iii), there exists a signature σ such that $\sigma(H_0) = 1$, i.e., $\bigcap_{i \in I_0} V(a_i) \neq \emptyset$. Thus $\{V(a_i)\}_{i \in I}$ is a family of closed sets with the finite intersection property. Since Sign (C, J) is compact it follows that $\bigcap_{i \in I} V(a_i) \neq \emptyset$, i.e., there exists a signature σ with $\sigma(H) = 1$.

(iv) $\Rightarrow$ (v). Let $G = A^*/NC^*$ and let L be the ideal of $Z[G]$ generated by $[1] + [-1]$. Now, minimal prime ideals of $Z[G]$ correspond bijectively with group homomorphisms $G \rightarrow \{\pm 1\}$ and under this correspondence, prime ideals containing L correspond to homomorphisms sending $[-1]$ to -1 [5, Lemma 3.1]. By (iv), the latter set coincides with the set of signatures of (C, J). Thus if K is the kernel of the mapping ψ, then $L \subset K$ and if P is a minimal prime ideal of $Z[G]$ with $L \subset P$ then by definition of signature, we must also have $K \subset P$. Thus to prove (v) it is enough to show that L is the intersection of all such prime ideals. This is done by showing that $Z[G]/L \cong Z[H]$ where H is any subgroup of index two in G with $[-1] \notin H$. Statement (v) then follows because $Z[H]$ has no nonzero nilpotent elements. Note that this will also prove the implication (v) $\Rightarrow$ (vi).

Thus let H be a subgroup of index two in G with $[-1] \notin H$. Let $S = \{[1], -[-1]\}$ and $G' = H \times S$. Then $Z[G] = Z[G']$ and L is the ideal generated by all elements of the form $1 - s$, $s \in S$. Hence $Z[G]/L = Z[G']/L \cong Z[G'/S] \cong Z[H]$.

(v) $\Rightarrow$ (vi) is contained in the above argument.

(vi) $\Rightarrow$ (vii) is trivial.

(vii) $\Rightarrow$ (i). Let H be a group of exponent two and $f: W(C, J) \rightarrow Z[H]$ an isomorphism. Since for any a in A^*, $\langle a \rangle^2 = 1$ in $W(C, J)$ it follows that $(f(\langle a \rangle))^2 = 1$ in $Z[H]$ so by [5, Th. 3.23], $f(\langle a \rangle) = \pm h$ for some h in H. Now suppose a is a unit in A with $a \notin NC^*$ and c_1, c_2 are elements of C such that $b = N(c_1) + aN(c_2)$ is a unit in A. Then by [5, Th. 1.16 (iii) and Lemma 1.19] $(1 + \langle a \rangle)(1 - \langle b \rangle) = 0$ in $W(C, J)$. Hence

$$1 = -f(\langle a \rangle) + f(\langle b \rangle) + f(\langle ab \rangle)$$

in $Z[H]$. Thus either $f(\langle a \rangle) = -1$, or $f(\langle b \rangle) = 1$, or $f(\langle ab \rangle) = 1$. Since f is an isomorphism, $f(\langle a \rangle) = -1$ implies $\langle a \rangle = \langle -1 \rangle$ in $W(C, J)$ which implies $a \in -NC^*$, contrary to assumption. If $f(\langle b \rangle) = 1$ then $\langle b \rangle = 1$ in $W(C, J)$, i.e., $b = N(c_1) + aN(c_2) \in NC^*$, so we are done in this case. If $f(\langle ab \rangle) = 1$ then $f(\langle a \rangle) = f(\langle b \rangle)$ so $\langle a \rangle = \langle b \rangle$, i.e., $b \in aNC^*$, completing the proof.

REMARKS. (i) In [3], Elman and Lam studied formally real (= ordered) fields satisfying condition (iii) of Theorem 1. There, they proved a structure theorem, [3, Th. 5.13], for the Witt ring and algebraic k-groups of such fields which contains the statement that the Witt ring is an integral group ring. They also proved several equivalent conditions characterizing these fields which include the equivalence of (iii) and (iv) [3, Ths. 4.3, 4.7]. In fact, the foregoing proof of (iii) $\Rightarrow$ (iv) is the same as the proof S1 $\Rightarrow$ S2 in Theorem 4.3.

(ii) Diller and Dress [2] observed the equivalence of conditions (i) and (ii) when $J =$ Identity and $C = A$ is a field and showed that these are equivalent to the following:

For any a in A^* with $a \notin -A^{*2}$ the field $A(\sqrt{a})$ is pythagorean, i.e., sums of squares are squares [2, Satz 4].

Following Elman-Lam [3, Def. 4.4], we call (C, J) superpythagorean if (C, J) has property (*) and satisfies the conditions of Theorem 1.

COROLLARY 2. *If (C, J) is superpythagorean then every unit of A which is a sum of norms is itself a norm.*

Proof. This follows from condition (i) with $a = 1$. (See also [6, Prop. 3.13].)

COROLLARY 3. (cf. [3, Cor. 4.5]). *Assume A^*/NC^* is a finite group of order 2^n, $n \geqq 1$. Then (C, J) is superpythagorean if and only if C has exactly 2^{n-1} distinct signatures.*

Proof. Apply condition (ii) of the theorem together with the fact that there are exactly 2^{n-1} homomorphisms $A^*/NC^* \to \{\pm 1\}$ sending $[-1]$ to -1.

REMARK. In contrast to the Witt ring, the Witt-Grothendieck ring $K(C, J)$ of isometry classes of nondegenerate hermitian spaces over C is seldom an integral group ring. In fact, if $-1 \notin NC^*$, it is not difficult to show that the following statements are equivalent:

(a) $K(C, J)$ is the integral group ring of some group,

(b) $K(C, J) \cong \boldsymbol{Z}[A^*/NC^*]$,

(c) A^*/NC^* is cyclic of order two,

(d) $W(C, J) \cong \boldsymbol{Z}$,

(e) Ker ψ is *additively* generated by $[1] + [-1]$.

For the remainder of the paper we assume that J is the identity (so $C = A$ and $NC^* = A^{*2}$).

PROPOSITION 4. *Let A be a local ring with maximal ideal M and residue class field $k = A/M$. Assume $1 + M \subset A^{*2}$ (this happens, for example, if A is henselian* [1, Ex. 3, p. 126]). *Then*

(i) *A is superpythagorean if and only if k is a superpythagorean field.*

(ii) *If, in addition, A is a valuation ring with field of fractions F then A is superpythagorean if and only if F is a superpythagorean field.*

Proof. (i) By [4, Satz 7.1.1, N.B. 7.1.3] there is an isomorphism of Witt rings $W(A) \cong W(k)$ and hence A is superpythagorean if and only if k is.

(ii) Let A be a valuation ring with field of fractions F and assume A is superpythagorean. Since F is a field, in order to show a function $\sigma : F^* \to \{\pm 1\}$ with $\sigma(F^{*2}) = 1$ and $\sigma(-1) = -1$ is a signature it is enough to show that $\sigma(a) = 1$ implies $\sigma(1 + a) = 1$. Thus suppose a is an element of F^* with $\sigma(a) = 1$ and let $\bar{\sigma} = \sigma \mid A^*$. Then $\bar{\sigma}(A^{*2}) = 1$ and $\bar{\sigma}(-1) = -1$ so $\bar{\sigma}$ is a signature of A. Since A is a valuation ring of F, for any a in F, either a is a unit in A, or $a \in M$, or $a^{-1} \in M$. If a is a unit in A then $1 + a$ is also a unit (if $1 + a \in M$ then $1 + m = -a$ for some $m \in M$ and since $1 + M \subset A^{*2}$ this means $1 = \sigma(1 + m) = \sigma(-a) = -\sigma(a) = -1$, impossible). Since $\bar{\sigma}$ is a signature, $\sigma(1 + a) = \bar{\sigma}(1 + a) = 1$. If $a \in M$ then $1 + a \in A^{*2}$ so $\sigma(1 + a) = 1$ and if $a^{-1} \in M$ then $\sigma(1 + a^{-1}) = 1$ and
$\sigma(1 + a) = \sigma(a(a^{-1} + 1)) = \sigma(a)(a^{-1} + 1) = 1$, showing that F is superpythagorean.

Conversely, suppose F is superpythagorean. Since A is integrally closed in F, the inclusion $A^* \to F^*$ induces an inclusion $A^*/A^{*2} \to F^*/F^{*2}$. Since both are vector spaces over F_2 any homomorphims $\sigma : A^*/A^{*2} \to \{\pm 1\}$ extends to a homomorphism $\hat{\sigma} : F^*/F^{*2} \to \{\pm 1\}$. If $\hat{\sigma}$ is a signature of F then σ is a signature of A, completing the proof.

REMARK. The last part of the proof actually shows that if $A \subset B$ are rings with $A^* \cap B^{*2} = A^{*2}$ then A is superpythagorean if B is.

EXAMPLES. Assume k is a superpythagorean field. Then

(a) The ring of formal power series in n-variables, $k[[X_1, \cdots, X_n]]$, is superpythagorean. The ring of dual numbers over k, $k[\varepsilon]$, $\varepsilon^2 = 0$, is superpythagorean.

(b) If $n \geqq 2$ the quotient field $k((X_1, \cdots, X_n))$ of $k[[X_1, \cdots, X_n]]$ is not pythagorean (hence cannot be superpythagorean). However,

(c) [3, Cor. 4.6]. For any $n \geqq 1$ the field $k((X_1)) \cdots (X_n))$ of *iterated* Laurent series over k is superpythagorean.

Proof. (a) This is immediate from Proposition 4 (i).

(b) It is not difficult to check that $X^2 + Y^2$ cannot be a square in the field $k((X, Y))$.

(c) Here it is enough to show that $k((X))$ is superpythagorean. However, this follows from (a) and Proposition 4 (ii).

REFERENCES

1. N. Bourbaki, *Algèbre commutative,* Chaps. 3–4, Actualités Sci. Indust., no. 1293, Hermann, Paris, 1967.
2. J. Diller und A. Dress, *Zur Galoistheorie pythagoreischer Körper,* Arch. Math., **16** (1965), 148–152.
3. R. Elman and T. Y. Lam, *Quadratic forms over formally real fields and pythagorean fields,* Amer. J. Math., **94** (1972), 1155–1194.
4. M. Knebusch, *Grothendieck-und Wittringe von nichtausgearten symmetrischen Bilinear formen,* S.-B. Heidelberger Akad. Wiss. Math.—Natur. Kl. 1969, 93–157.
5. M. Knebusch, A. Rosenberg and R. Ware, *Structure of Witt rings and quotients of abelian group rings,* Amer. J. Math., **94** (1972), 119–155.
6. ———, *Signatures on semi-local rings,* J. Algebra, **26** (1973), 208–250.

Received July 5, 1972. Partially supported by NSF Grant GP-28915.

NORTHWESTERN UNIVERSITY
Current address: UNIVERSITY OF KANSAS

PACIFIC JOURNAL OF MATHEMATICS
Vol. 49, No. 1, 1973

PARACOMPACTIFICATIONS USING FILTER BASES

J. D. WINE

During the mid nineteen sixties, C. E. Aull presented a series of papers in which he distinguished several different types of paracompact subsets. Using these concepts, three classes of filter bases are introduced and their convergence and other properties are studied. A variety of characterizations of paracompactness based on the existence of certain types of these filter bases and z-filters are given. A paracompactification construction involving the addition of limit points for one of the classes of filter bases is presented in detail. Finally, properties of the paracompactification are explored with some attention given to its ring of continuous functions.

Generally, the notation of Gillman and Jerison [7], will be followed. To avoid confusion when considering a property which could be associated with any one of several sets under consideration the set symbol will be affixed before the property symbol.

Filter base classes. We begin with the definitions of the various paracompact subsets which were introduced by Aull.

DEFINITION 1. A subset M of a topological space X is *α-paracompact* if and only if given an X-open cover of M there is an X-open refinement which covers M and is locally finite at every point x in X (denote by X locally finite).

DEFINITION 2. A subset of X is β-paracompact if and only if it is paracompact as a subspace of X.

DEFINITION 3. A subset E of X is *σ-paracompact* if and only if every X-open cover of E has an X-open X σ-locally finite refinement which covers E.

DEFINITION 4. An *α-filter base* (respectively *β-filter base*, *σ-filter base*) is a family $\mathscr{F}$ of nonempty zero sets satisfying:

(1) If Z is in $\mathscr{F}$ then Z is not α-paracompact (respectively β-paracompact, σ-paracompact), and

(2) If U and V are in $\mathscr{F}$ then their intersection contains an element of $\mathscr{F}$.

In this section we wish to develop some properties of the three classes of filter bases defined above. For the most part, the proofs are identical for all three types except for the change in terminology. When such is the case we shall give the statement in all three forms, but only prove the result for the α-paracompact case, leaving the verifications of the others to the readers.

THEOREM 1. *An α-filter base (respectively β-filter base, σ-filter base) on a locally paracompact space cannot converge. (For properties of a locally paracompact space see* [13].)

Proof. By definition a space is locally paracompact if and only if every point has an α-paracompact neighborhood. No element of the α-filter base can be contained in the α-paracompact neighborhood of a point so the α-filter base can not converge.

THEOREM 2. *If $\mathscr{B}$ is an α-filter base (respectively β-filter base, σ-filter base), then there is a maximal α-filter base (respectively β-filter base, σ-filter base) which contains $\mathscr{B}$.*

Proof. This is a standard Zorn's lemma argument and is omitted.

DEFINITION 5. Let X be a space, then $\mathscr{A}(X)$ (respectively $\mathscr{B}(X)$, $\mathscr{S}(X)$) is the family of all maximal α-filter bases (respectively β-filter bases, σ-filter bases).

THEOREM 3. *If $\mathscr{F}$ is in $\mathscr{A}(X)$ (respectively $\mathscr{B}(X)$, $\mathscr{S}(X)$), then any zero set which contains an element of $\mathscr{F}$ is an element of $\mathscr{F}$.*

Proof. Let $\mathscr{F}$ be an element of $\mathscr{F}$ which is contained in the given zero set Z. Then Z cannot be α-paracompact because F is not. For any element B of $\mathscr{F}$, we must have $B \cap Z$ containing $B \cap F$ which contains an element of $\mathscr{F}$. Hence $\{z\} \cup \mathscr{F}$ is an α-filter base containing $\mathscr{F}$, which implies Z is an element of $\mathscr{F}$.

THEOREM 4. *If $\mathscr{F}$ is in $\mathscr{A}(X)$ (respectively $\mathscr{B}(X)$, $\mathscr{S}(X)$), then the intersection of two elements of $\mathscr{F}$ is an element of $\mathscr{F}$.*

Proof. The result follows immediately from Theorem 3 and the definition of an α-filter base.

As an immediate result of the last two theorems we have the theorem below.

THEOREM 5. *If $\mathscr{F}$ is in $\mathscr{A}(X)$ (respectively $\mathscr{B}(X)$, $\mathscr{S}(X)$), then $\mathscr{F}$ is a z-filter.*

THEOREM 6. *An α-filter base (respectively β-filter base, σ-filter* *base)* $\mathscr{F}$ *is maximal if and only if every zero set* Z*, such that* $Z \cap F$ *is not* α*-paracompact (respectively* β*-paracompact,* σ*-paracompact) for every* F *in* $\mathscr{F}$*, is an element of* $\mathscr{F}$*.*

Proof. Necessity. Suppose $\mathscr{F}$ is maximal and let Z be a zero set such that $Z \cap F$ is not α-paracompact for every F in $\mathscr{F}$. Then $\mathscr{U} = \{S : S = F \cap Z, F \text{ in } \mathscr{F}\} \cup \mathscr{F}$ is a family of non α-paracompact zero sets. Now let S and S^* be elements of $\mathscr{U}$. If both S and S^* are elements of $\mathscr{F}$, then $S \cap S^*$ is in $\mathscr{F}$ and hence $\mathscr{U}$. If either one or both of S and S^* is not in $\mathscr{F}$, then $S \cap S^*$ has the form $F \cap F^* \cap Z$ where F and F^* are in $\mathscr{F}$. In this case $E = F \cap F^*$ is in $\mathscr{F}$, and $E \cap Z$ is an element of $\mathscr{U}$. Therefore, $\mathscr{U}$ is an α-filter base, and since $X \cap Z$ is an element of $\mathscr{U}$, Z is in $\mathscr{U}$. Since $\mathscr{U}$ contains the maximal α-filter base $\mathscr{F}$ we must have $\mathscr{U} = \mathscr{F}$.

Sufficiency. Let $\mathscr{F}$ satisfy the hypothesis, and let $\mathscr{U}$ be an element of $\mathscr{A}(X)$ which contains $\mathscr{F}$. If $\mathscr{F}$ is not $\mathscr{U}$, then there is a Z in $\mathscr{U} - \mathscr{F}$. Since $\mathscr{U}$ is maximal and contains $\mathscr{F}$, we have $Z \cap F$ is not α-paracompact for every F in $\mathscr{F}$. Hence, Z is in $\mathscr{F}$, and $\mathscr{U}$ must equal $\mathscr{F}$.

We shall show later that maximal α-filter bases (respectively β-filter bases, σ-filter bases) on a Hausdorff completely regular space are very large in the sense that they are almost z-ultrafilters, being contained in only one z-ultrafilter.

Paracompact spaces and filter bases. We now present a variety of characterizations of paracompactness depending on the existence of α-filter bases, β-filter bases, σ-filter bases, and z-ultrafilters with α-paracompact elements. We continue our convention used in the last section regarding the three types of paracompact subsets and the statement and proof of theorems. For some of the results involving β-filter bases we will have need for the following result.

THEOREM 7. *If* $\mathscr{F}$ *is an* α*-filter base, let* $\mathscr{F}^0$ *be* $\{F^0 \text{ in } \mathscr{F} : F^0$ *is a neighborhood of some element of* $\mathscr{F}\}$*. Then* $\mathscr{F}$ *is a* β*-filter base.*

Proof. By Theorem 1 in [4], if some element of $\mathscr{F}^0$ were β-paracompact, the element of $\mathscr{F}$ for which it is a neighborhood would be α-paracompact.

THEOREM 8. *Let* X *be a completely regular space. Then* X *is paracompact if and only if there is no free* α*-filter base (respectively*

β-filter base, σ-filter base) on X. Also, X is paracompact if and only if there is no free maximal α-filter base (respectively β-filter base, σ-filter base) on X.

Proof. Necessity is obvious for all three cases.

Sufficiency. For the cases involving α-filter bases or σ-filter bases suppose X is not paracompact. Then there is an open cover $\mathscr{G}$ which has no locally finite (respectively σ-locally finite) refinement. Therefore, if $\mathscr{G}^*$ is a finite subfamily of $\mathscr{G}$, the union of elements of $\mathscr{G}^*$ is not X, and its complement is not α-paracompact (respectively σ-paracompact). Let $\mathscr{B} = \{F : F$ is the complement of the union of a finite subfamily of $\mathscr{G}\}$ and $\mathscr{F} = \{Z$ in $Z(X) : Z$ contains some F in $\mathscr{B}\}$. An easy calculation shows $\mathscr{F}$ to be an α-filter base (respectively σ-filter base). Since $\mathscr{G}$ is a cover of X, we may use complete regularity to show that $\mathscr{F}$ is free, and we are done.

For the case involving β-filter bases, we take the α-filter base $\mathscr{F}$ obtained above, and use Theorem 7 to obtain the β-filter base $\mathscr{F}^0$. Since $\mathscr{F}$ is free, we may use complete regularity to show that $\mathscr{F}^0$ is free.

The last assertion follows from Theorem 2.

COROLLARY 8A. *A completely regular space X is paracompact if and only if every free z-filter has an α-paracompact (respectively β-paracompact, σ-paracompact) element.*

Proof. The proof follows from Theorem 5.

COROLLARY 8B. *Let X be a regular space. Then X is paracompact if and only if there is no α-filter base (respectively β-filter base, σ-filter base) on X. If we drop the condition of regularity then we must drop the case for σ-filter bases in the conclusion.*

Proof. In the proof of the theorem, we needed complete regularity only to obtain $\mathscr{F}$ free, otherwise regularity sufficed, and regularity is needed only for the σ-filter base.

We note that for Corollary 8A results corresponding to Corollary 8B can be stated. If the space is locally paracompact and either regular or Hausdorff a slightly nicer result is possible. We will need the following results from [13].

THEOREM 9. *A Hausdorff locally paracompact space is regular.*

THEOREM 10. *Let E be an α-paracompact subset of a regular (respectively Hausdroff) locally paracompact space and let G be any open set containing E, then there is a closed α-paracompact neighborhood of E contained in G.*

THEOREM 11. *If X is a regular (respectively Hausdorff) locally paracompact space and A is an α-paracompact subset, then A is eompletely separated from any disjoint closed set.*

Proof. Let F be a closed set disjoint from A, then by Theorem 10 there is a closed α-paracompact neighborhood V of A which is contained in $X - F$. The subspace V is normal since it is regular and paracompact, hence there is a function g in $C(V)$ such that $0 \leqq g(x) \leqq 1$ and g is zero on A and one on V-int V. Define f taking X into the closed unit interval by $f \mid V = g$ and $f[X\text{-int } V] = 1$. Clearly f is continuous and completely separates A and F.

The Hausdorff case follows from Theorem 9.

THEOREM 12. *If X is a regular (respectively Hausdorff) locally paracompact space and $\mathscr{F}$ is in $\mathscr{A}(x)$, then given any α-paracompact set A there is an F in $\mathscr{F}$ which is disjoint from A.*

Proof. Let A and $\mathscr{F}$ be as in the hypothesis, and let V be a closed α-paracompact neighborhood of A. For each F in $\mathscr{F}$, the intersection of F and V is α-paracompact. Therefore, $E = F \cap (X\text{-int } V)$ is not α-paracompact. Since E is disjoint from A, there are disjoint zero set neighborhoods $Z(A)$ and $Z(E)$. By using Theorem 6, it is not difficult to show $Z(E)$ is an element of $\mathscr{F}$.

The Hausdorff case follows from Theorem 9.

COROLLARY 12A. *If X is a regular (respectively Hausdorff) locally paracompact space, then no maximal α-filter base has a cluster point.*

THEOREM 13. *A regular (respectively Hausdorff) locally paracompact space X is paracompact if and only if every free z-ultrafilter has an α-paracompact element.*

Proof. For the nontrivial part, suppose that every free z-ultrafilter has an α-paracompact element. Then no free z-ultrafilter can contain a maximal α-filter base by Theorem 12. Hence, any maximal α-filter base on X must be fixed. This is impossible because X is locally paracompact. Hence, X has no maximal α-filter base and is paracompact.

THEOREM 14. *If X is a regular (respectively Hausdorff) locally paracompact space, $\mathscr{A}(X)$ and the family of all free z-ultrafilters with no α-paracompact elements are identical.*

Proof. Every such z-ultrafilter must be an element of $\mathscr{A}(X)$. Now suppose that $\mathscr{F}$ is in $\mathscr{A}(X)$. Then there is a free z-ultrafilter $\mathscr{U}$ containing $\mathscr{F}$ and by Theorem 12 no element of $\mathscr{U}$ is α-paracompact. Hence $\mathscr{F}$ equals $\mathscr{U}$.

Paracompactifications using filter bases. In this section we take up the construction of paracompactifications obtained by adding limit points to the various classes of filter bases discussed above.

DEFINITION 6. Let K (respectively K_1, K_2) be an index set for the maximal α-filter bases (respectively β-filter bases, σ-filter bases) on X, then define $\mathscr{A}_F(X)$ (respectively $\mathscr{B}_F(X)$, $\mathscr{S}_F(X)$) to be $\{\mathscr{F}_k : k$ is in K (respectively K_1, K_2) and $\mathscr{F}_k$ is a free maximal α-filter base (respectively β-filter base, σ-filter base)$\}$. For each set G in X define $K(G)$ (respectively $K_1(G)$, $K_2(G)$) to be $\{k : k$ is in K (respectively K_1, K_2) and there is an F in $\mathscr{F}_k$ contained in $G\}$.

To establish the desired topologies on the extensions of X the following result is needed, its proof is obvious.

THEOREM 15. *If G and H are arbitrary sets, then $K(G) \cap K(H)$ equals $K(G \cap H)$ and similarly for K_1 and K_2.*

Using the previous result, it is a simple computation to obtain the following.

THEOREM 16. *Let $\pi X = \{y_k : k \in K, y_k$ is not in $X\} \cup X$ and $\mathscr{B} = \{H : H = G \cup \{y_k : k$ is in $K(G)\}$ where G is a cozero set$\}$. Then $\mathscr{B}$ is a base for a topology on πX. For K_1 and K_2 we may define $\pi_1 X$, $\mathscr{B}_1$, and $\pi_2 X$, $\mathscr{B}_2$ respectively with the analogous conclusion.*

From now on we will be using results from Gillman and Jerison [7], and hence will require all spaces to be Hausdorff and completely regular. Let W (respectively W_1, W_2) be the subspace of βX obtained by adding to X all points p in $\beta X - X$ for which the z-ultrafilter A^p on X contains a maximal α-filter base (respectively β-filter base, σ-filter base). Denote $\{A^p : p$ is in W (respectively W_1, W_2)$\}$ by $A(X)$ (respectively $B(X)$, $S(X)$).

THEOREM 17. *If $\mathscr{F}_k$ is in $\mathscr{A}(X)$ (respectively $\mathscr{B}(X)$, $\mathscr{S}(X)$) and A^p is an element of $A(X)$ (respectively $B(X)$, $S(X)$) such that $\mathscr{F}_k$ is contained in A^p, then $\mathscr{F}_k$ converges to p.*

Proof. Since the zero set neighborhoods of a point in a Hausdorff completely regular space are a base for its neighborhood system, it is sufficient to show that every zero set neighborhood of p contains an element of $\mathscr{F}_k$. Let $V(p)$ be a zero set neighborhood of p, and let $V^*(p)$ be a zero set neighborhood of p contained in the interior of $V(p)$.

Let $E = W\text{-int}_W V(p)$, since p is not in E there are disjoint zero set neighborhoods of E and p, call them $F(E)$ and $F(p)$. Let $U(p)$ be the intersection of $F(p)$ and $V^*(p)$. Then $U(p)$ is a zero set neighborhood of p, and $U(p)$ is contained in $\text{int}_W V(p)$. Let $Z(p) = V(p) \cap X$, $Z^*(p) = U(p) \cap X$, and $Z(E) = F(E) \cap X$; then $Z(p)$, $Z^*(p)$, and $Z(E)$ are zero sets in X. We also have $Z(E) \cap Z(p)$ empty and $Z(p) \cup Z(E) = X$. If p is in X, then it may be that $Z(p) = V(p)$, or $Z^*(p) = U(p)$, but all relations still hold.

Now suppose that $Z(p) \cap F$ is α-paracompact for some F in $\mathscr{F}_k$. We have that $Z^*(p) \cap F$ and $Z(E)$ are disjoint zero sets, and because $Z^*(p) \cap F$ must be in A^p, $Z(E)$ cannot be in $\mathscr{F}_k$. There is an F' in $\mathscr{F}_k$ such that $F' \cap Z(E)$ is α-paracompact, but then

$$F' \cap F = ((F' \cap F) \cap Z(E)) \cup ((F' \cap F) \cap Z(p))$$

which must be α-paracompact because components of the union are. This is impossible since $F' \cap F$ is in $\mathscr{F}_k$. Hence, $Z(E)$ must be in $\mathscr{F}_k$ or else $Z(p) \cap F$ is not α-paracompact, the former is impossible so the latter is the case. Therefore, $Z(p)$ is in $\mathscr{F}_k$.

The case for σ-filter bases has only the terminology changed in the above proof. The case for β-filter bases rests upon the fact that in a Hausdorff completely regular space the union of two closed β-paracompact sets is β-paracompact, so that the above proof holds with appropriate terminology changes.

COROLLARY 17A. *If X is a Hausdorff completely regular space, there is a one-to-one correspondence between the filter bases in $\mathscr{A}(X)$ (respectively $\mathscr{B}(X)$, $\mathscr{S}(X)$) and the z-ultrafilters containing them.*

COROLLARY 17B. *Every element in $\mathscr{A}(X)$ (respectively $\mathscr{B}(X)$, $\mathscr{S}(X)$) converges in W (respectively W_1, W_2).*

If $\mathscr{F}_k$ converges to p in W, denote $\mathscr{F}_k$ by $\mathscr{F}_{k(p)}$. The one-to-one correspondence given by Corollary 17A illustrates the previously mentioned fact that elements of $\mathscr{A}(X)$, $\mathscr{B}(X)$, and $\mathscr{S}(X)$ are nearly z-ultrafilters, since in general, a z-filter is contained in many z-ultrafilters.

THEOREM 18. *Let U be the relative topology of the subspace W of βX. The family $\mathscr{B}_w = \{H: H = G \cup \{p : k(p)$ is in $K(G)\}$ where G is* *a cozero set of $X\}$ is a base for U. Analogous statements may be made regarding W_1 and W_2.*

Proof. Let Z be in $Z(X)$, $G = X - Z$, and $H = W\text{-cl}_w Z$. If we show H has the desired form, since $\{\text{cl}_w Z : Z \text{ in } Z(X)\}$ is a base for the closed sets in W, we will have the topology generated by $\mathscr{B}$ contains $\mathscr{U}$.

Let q be an element of $W - X$. Suppose that q is in H, then since $\mathscr{F}_{k(q)}$ converges to q, and H is a neighborhood of q there is an F in $\mathscr{F}_{k(q)}$ such that F is contained in $H \cap X = G$. Therefore, $\{k(q) : q$ is in $H\}$ is contained in $K(G)$. Suppose q is not in H. Then q is in $\text{cl}_w Z$ and by the construction of βX given in [7], Z is an element of A^q. Hence, since F an element of $\mathscr{F}_{k(q)}$ implies that F is in A^q, we must have $F \cap Z$ nonempty and no F in $\mathscr{F}_{k(q)}$ is contained in $H \cap X = G$. Therefore, $K(G)$ is contained in $\{k(q) : q$ is in $H\}$ and they are equal. Hence, $H = G \cup \{q : q \text{ is in } H\} = G \cup \{q : k(q) \text{ is in } K(G)\}$.

To show that the topology generated by $\mathscr{B}$ is contained in $\mathscr{U}$, we show that every set in $\mathscr{B}$ is a member of $\mathscr{U}$. Let $H = G \cup \{p : k(p)$ is in $K(G)\}$ be an element of $\mathscr{B}$. If p is in H, then there exists an F in $\mathscr{F}_{k(p)}$ such that F is contained in G, so that $F \cap (X - g)$ is empty and the zero set $X - G$ cannot be an element of A^p. Hence p is not in $\text{cl}_w (X - G)$, and H is contained in $W\text{-cl}_w (X - G)$. If p is an element of $W - H$, no F in $\mathscr{F}_{k(p)}$ is in G. Hence every F in $\mathscr{F}_{k(p)}$ has nonvoid intersection with $X - G$. Therefore, $X - G$ is an element of A^p and p is not in $W\text{-cl}_w (X - G)$, so that $W\text{-cl}_w (X - G)$ is contained in H. Hence $H = W - \text{cl}_w (X - G)$ and H is an element of $\mathscr{U}$.

COROLLARY 18A. *The space πX (respectively $\pi_1 X$, $\pi_2 X$) is homeomorphic to a subspace of βX, namely W (respectively W_1, W_2). Hence πX (respectively $\pi_1 X$, $\pi_2 X$) is a completely regular Hausdorff space.*

Proof. This follows from Theorems 16, 17, and 18; identifying $\mathscr{B}$ and $\mathscr{B}_w$ in the obvious manner.

We now commence a series of lemmas which lead to the result that the extensions πX, $\pi_1 X$, and $\pi_2 X$ are paracompact. The first two are given simply for reference.

LEMMA 19. *If $\mathscr{F}$ is a free z-filter on a completely regular Hausdorff space, then $\mathscr{F}^0$ is also a free z-filter.*

LEMMA 20. *Let E be a dense subspace of a completely regular Hausdorff space Y. Then, if $\mathscr{F}$ is a free z-filter on Y with each* *element having nonvoid interior, the trace of* $\mathscr{F}$ *on* E *is a free* z*-filter.*

LEMMA 21. *If a* πX *closed (respectively* $\pi_1 X$ *closed,* $\pi_2 X$ *closed) set* F *is contained in* X*, then* F *is* X α*-paracompact (respectively* β*-paracompact,* X σ*-paracompact).*

Proof. Suppose that F is not X α-paracompact. Then there is an X open cover $\mathscr{G}$ which covers F and has no X locally finite refinement which covers F. Hence, $\mathscr{G}^* = \mathscr{G} \cup (X - F)$ covers X and has no locally finite refinement which covers X. Therefore, $\mathscr{F} = \{Z$ in $Z(X) : Z$ contains the complement of the union of a finite subfamily of $\mathscr{G}^*\}$ is a free α-filter base on X, and so F is contained in $\mathscr{F}_{k(q)}$, a member of $\mathscr{A}_F(X)$. Hence $\{Z$ in $Z(X) : F$ is contained in $Z\}$ is contained in $\mathscr{F}_{k(q)}$, which implies q is in $\mathrm{cl}_{\pi X} F = F$, and F is not contained in X.

The σ-paracompact case is analogous with the appropriate changes in terminology.

Now assume F is not β-paracompact, then F is not α-paracompact, and there is a free α-filter base $\mathscr{F}$ on X such that $\mathscr{F}$ contains $\{Z$ in $Z(X) : Z$ contains $F\}$. For some z-ultrafilter A^p and some $\mathscr{F}_{k(p)}$ in $\mathscr{A}_F(X)$ we have

$$\mathscr{F}^0 \subset \mathscr{F} \subset \mathscr{F}_{k(p)} \subset A^p .$$

Note that while A^p is an element of $A(X)$ we cannot assume A^p is in $B(X)$, however this is the case, as we now show. If Z_p is the intersection of a zero set neighborhood of p with X, Z_p contains an element Z of $\mathscr{F}_{k(p)}$ in $\mathrm{int}_X Z_p$. Hence Z_p is not β-paracompact. If F is an element of $\mathscr{F}^0$, let Z^* be an element of $\mathscr{F}$ which is contained in $\mathrm{int}_X F$, then

$$F \cap Z_p \supset \mathrm{int}_X (F \cap Z_p) = (\mathrm{int}_X F) \cap (\mathrm{int}_X Z_p) \supset Z^* \cap Z$$

and $F \cap Z_p$ is not β-paracompact. Since the trace $\mathscr{N}$ on the z-filter of zero set neighborhoods of p is a z-filter on X, and since p is a cluster point of $\mathscr{F}^0$, the family $\mathscr{N} \cup \mathscr{F}^0$ generates a β-filter base containing $\mathscr{F}^0$ and converging to p. Hence there is a maximal β-filter base $\mathscr{B}$ containing $\mathscr{F}^0$ and itself contained in A^p. Therefore A^p is in $B(X)$ and p would be in $\mathrm{cl}_{\pi_1 X} F$, which is a contradiction.

LEMMA 22. *Let* f *be an element of* $C(X)$ *which is bounded on the complement of an* α*-paracompact (respectively closed* β*-paracompact, closed* σ*-paracompact) set* A *by a real number* M*, then there is an extension* f^π *(respectively* f', f''*) of* f *in* $C(\pi X)$ *(respectively* $C(\pi_1 X)$*,* $C(\pi_2 X)$*) which is bounded on the remainder by* M*.*

Proof. Since any f in $C(X)$ equals $(f \vee 0) + (f \wedge 0)$, it is sufficient to show the result for a function $f \geqq 0$ satisfying the conditions in the hypothesis. Now define $g = (M + 1/2) \wedge f$, then g is in $C^*(X)$, and g restricted to $X - A$ is equal to f restricted to $X - A$. Since g is in $C^*(X)$, there is an extension g^π in $C^*(\pi X)$. Define f^π by f^π restricted to X is equal to f and $f^\pi(y_k) = g^\pi(y_k)$.

For each k, $f^\pi(y_k) \leqq M$, since if $g^\pi(y_k) = r > M$ then the set $S = g^{\pi\leftarrow}[(r - (r - m)/2, r + (r - m)/2)]$ is an element of the neighborhood system of y_k in πX. Hence S must contain a non X α-paracompact element Z of $\mathscr{F}_k$. But $S \cap X$ is contained in A so that Z would be X α-paracompact.

Now since $r = g^\pi(y_k) \leqq M$, let

$$V(r) = \left(r - \frac{M + 1/2 - r}{2}, r + \frac{M + 1/2 - r}{2}\right).$$

Then $g^{\pi\leftarrow}[V(r)]$ is a neighborhood of y_k in πX. The neighborhood $V(r)$ is contained in the interval $[0, M + 1/2]$, hence $f^{\pi\leftarrow}[V(r)] = g^{\pi\leftarrow}[V(r)]$. Any neighborhood of r contained in $V(r)$ is the preimage of a neighborhood of y_k under $f^{\pi\leftarrow}$. Hence if $U(r)$ is a neighborhood of r, then there exists an ε neighborhood $V_\varepsilon(r)$ such that $V_\varepsilon(r)$ is contained in $U(r) \cap V(r)$. The set $f^{\pi\leftarrow}[V_\varepsilon(r)]$ is in the neighborhood system of y_k in πX so that $f^{\pi\leftarrow}[U(r)]$, which contains $f^{\pi\leftarrow}[V_\varepsilon(r)]$, is a neighborhood of y_k. Therefore, f^π is continuous at y_k.

LEMMA 23. *Let F be an X α-paracompact (respectively σ-paracompact) set contained in a πX (respectively $\pi_2 X$) closed neighborhood V, which is contained in X. Then F is πX α-paracompact (respectively $\pi_2 X$ σ-paracompact).*

Proof. Let $\mathscr{G}$ be a πX open cover of F, and let $\mathscr{G}^* = \{G^* : G^*$ is equal to $G \cap \operatorname{int}_X V$, G an element of $\mathscr{G}\}$. Since V is X α-paracompact by Lemma 21, $K(\operatorname{int}_X V)$ is empty. Hence $\operatorname{int}_X V = \operatorname{int}_{\pi X} V$ and $\mathscr{G}^*$ in both πX and X open. We can now obtain an X locally finite is open refinement of $\mathscr{G}^*$ which has all of its elements contained in $\operatorname{int}_{\pi X} V$. Hence the refinement of $\mathscr{G}^*$ is πX open and πX locally finite, so F is πX α-paracompact.

LEMMA 24. *Let F be an X α-paracompact (respectively σ-paracompact) set with an X α-paracompact (respectively σ-paracompact) neighborhood. Then F is πX α-paracompact (respectively $\pi_2 X$ σ-paracompact).*

Proof. Let G be the X open set containing F with $\operatorname{cl}_X G$ being X α-paracompact. In $\operatorname{cl}_X G$, F and $\operatorname{cl}_X G - G$ are closed disjoint sets.

Since $\mathrm{cl}_X G$ is X α-paracompact and Hausdorff, it is a normal subspace of X. Hence there are disjoint zero set neighborhoods Z_1 and Z_2 in $Z(\mathrm{cl}_X G)$ such that F and $\mathrm{cl}_X G - G$ are contained in the interior with respect to $\mathrm{cl}_X G$ of Z_1 and Z_2 respectively. Let h be an element of $C(\mathrm{cl}_X G)$ such that $Z_1 = h^{\leftarrow}(1)$, $Z_2 = h^{\leftarrow}(0)$, and $0 \leqq h \leqq 1$.

Define f taking X into the closed unit interval by $f \mid \mathrm{cl}_X G = h$ and $f \mid (X - G) = 0$. The function f is in $C(X)$. If it is the case that $\mathrm{cl}_X G - G$ is empty, we may immediately define f in $C(X)$ by $f[G] = 1$ and $f[X - G] = 0$.

Let $Z_3 = \{x : f(x) \geqq 1/2\}$. Then we have $F \subset \mathrm{int}_G Z_1 \subset Z_1 \subset \mathrm{int}_G Z_3 \subset G$. The function f is bounded on the complement of Z_3 by $1/2$, and Z_3 is X α-paracompact. We apply Lemma 22 to get f^π in $C(\pi X)$ such that f^π is bounded on the remainder by $1/2$. Hence Z_1 is in $Z(\pi X)$ and is a πX closed neighborhood of F.

THEOREM 25. *πX (respectively $\pi_1 X$, $\pi_2 X$) is paracompact.*

Proof. If πX is not paracompact, there is a free maximal α-filter base $\mathscr{F}$ on πX Corollary 8C. By Theorem 5 $\mathscr{F}$ is a z-filter. Let $\mathscr{F}^*$ be the trace of the z-filter $\mathscr{F}^0$ on X. Then $\mathscr{F}^*$ is a free z-filter.

Suppose Z^* is an element of $\mathscr{F}^*$, we wish to show that Z^* is not X α-paracompact. There exists a Z in $\mathscr{F}^0$, such that $Z \cap X = Z^*$. Let Z' be an element of $\mathscr{F}$ such that Z' is contained in $\mathrm{int}_{\pi X} Z$. If Z' is contained in X then Z' is X α-paracompact by Lemma 21, and if we assume Z^* to be X α-paracompact then Z' is πX α-paracompact by Lemma 24. Therefore, $Z' \cap (\pi X - X)$ is not empty, but then $(\mathrm{int}_{\pi X} Z) \cap (\pi X - X)$ is not empty. For each y_k in $(\mathrm{int}_{\pi X} Z) \cap (\pi X - X)$, the $\mathrm{int}_{\pi X} Z$ is a neighborhood of y_k. Hence Z must contain an element F of $\mathscr{F}_k$. Since F is then contained in Z^* also, Z^* is not X α-paracompact.

Since $\mathscr{F}^*$ is an α-filter base on X, there is a maximal α-filter base $\mathscr{N}^*$ on X which contains $\mathscr{F}^*$, and since $\mathscr{F}^*$ is free so is $\mathscr{N}^*$. Hence $\mathscr{N}^*$ converges to some y_k in $\pi X - X$. Let $\mathscr{U}^*$ be the unique z-ultrafilter on X containing $\mathscr{N}^*$. Let $\overline{\mathscr{U}}$ be the z-filter on πX generated by $\{\mathrm{cl}_{\pi X} Z : Z \text{ is in } \mathscr{U}^*\}$. The set of cluster points of $\overline{\mathscr{U}}$ contains y_k, so there is a z-ultrafilter $\mathscr{U}$ on πX containing $\overline{\mathscr{U}}$ and converging to y_k.

Since $\mathscr{F}$ is a free z-filter on πX, it cannot be contained in the convergent z-ultrafilter $\mathscr{U}$. Therefore, there is an E in $\mathscr{F}$ and an E^* in $\mathscr{U}$ which are disjoint. Hence E and E^* have disjoint zero set neighborhoods V and V^*. The zero set V is in $\mathscr{F}^0$ which implies $V \cap X$ is in $\mathscr{F}^*$ and V contains $\mathrm{cl}_{\pi X}(V \cap X)$ so V is in $\overline{\mathscr{U}}$. Therefore, V and V^* are both elements of $\mathscr{U}$, which is absurd. We have a contradiction and the proof for the space πX is complete.

The proof for the space $\pi_2 X$ requires only the appropriate terminology changes in the above. For the space $\pi_1 X$, we may use the fact that β-paracompactness is a property which is dependent only on the set, so the application of Lemma 24 is not necessary. Other than this change the proof is analogous to that for πX.

COROLLARY 25A. *The three paracompactifications are T_4 spaces.*

The question as to the relation which exist among πX, $\pi_1 X$, and $\pi_2 X$ is answered in the following.

THEOREM 26. *The three paracompactifications constructed above are identical.*

Proof. Since all the paracompactifications are subsets of βX, we need only show that for points in the remainder, the families of z-ultrafilters are identical.

Let A^p be a z-ultrafilter converging in $\pi X - X$ with its associated maximal α-filter base $\mathscr{F}_{k(p)}$. Let V be a basic open set in πX containing p. Then there is an element F of $\mathscr{F}_{k(p)}$ contained in $V \cap X$. The set $Z = X - (V \cap X)$ is a zero set of X. Since Z and F are disjoint zero sets, they have disjoint zero set neighborhoods Z^0 and F^0. The set F^0 is an element of the β-filter base $\mathscr{F}^0_{k(p)}$, and F^0 is contained in $V \cap X$. Hence $\mathscr{F}^0_{k(p)}$ converges to p. We can then find a free maximal β-filter base converging to p and contained in A^p. Therefore, A^p is in $B(X)$.

Now suppose A^p is a z-ultrafilter on X converging to a point in $\pi_1 X - X$ with its associated maximal β-filter base $\mathscr{F}_{k(p)}$. Now $\mathscr{F}_{k(p)}$ is a free α-filter base which has a unique z-ultrafilter A^p containing it. Hence the free maximal α-filter base containing $\mathscr{F}_{k(p)}$ must be contained in A^p, and A^p is in $A(X)$. We have $A(X) = B(X)$ and πX is identical to $\pi_1 X$.

A similar argument established $\pi_1 X$ to be identical to $\pi_2 X$.

THEOREM 27. *A completely regular Hausdorff space X is locally paracompact if and only if X is open in πX.*

Proof. The proof of this is a straight forward argument and so is omitted.

THEOREM 28. *If X is a Hausdorff locally paracompact space, then F is X α-paracompact if and only if* $\mathrm{cl}_{\pi X} F = F$.

Proof. Necessity. If F is X α-paracompact, then for each y_k in

$\pi X - X$ we may apply Theorem 11 to get a zero set Z_k in $\mathscr{F}_k^0$ which is disjoint from F. The set given by $(\mathrm{int}_X Z_k) \cup \{Y_j : j \text{ is in } K(\mathrm{int}_X Z_k)\}$ is a neighborhood of y_k in πX and is disjoint from F.

Sufficiency. This follows from Lemma 21.

COROLLARY 28A. *If X is a locally paracompact Hausdorff space, and F is an X α-paracompact set, then F is πX α-paracompact.*

THEOREM 29. *Let X be Hausdorff and completely regular, let G be an open subset of X, and let $G^* = G \cup \{y_k : k \text{ is in } K(G)\}$. Then* $\mathrm{cl}_{\pi X}\, G = \mathrm{cl}_{\pi X}\, G^*$.

Proof. Clearly G is contained in G^*, so that $\mathrm{cl}_{\pi X}\, G$ is contained in $\mathrm{cl}_{\pi X}\, G^*$. If y_k is in G^*, then there is an F_k in $\mathscr{F}_k$ such that F_k is contained in G. Since $\mathscr{F}_k$ converges to y_k, every neighborhood of y_k has nonvoid intersection with F_k and hence G. Therefore, G^* is contained in $\mathrm{cl}_{\pi X}\, G$, and so $\mathrm{cl}_{\pi X}\, G^*$ equals $\mathrm{cl}_{\pi X}\, G$.

THEOREM 30. *Let νX be the Hewitt Realcompactification of X. Then νX is contained in πX. (This result and the next require all cardinals to be nonmeasurable.)*

Proof. We know νX is the smallest realcompactification of X. Since πX is paracompact, it is realcompact and $\nu X \cap \pi X$ is realcompact. Hence νX must be contained in this intersection.

THEOREM 31. *For a Hausdorff completely regular space X the following are equivalent.*

(1) *πX is identical to νX.*

(2) *$A(X)$ is identical to the family of real z-ultrafilters on X.*

(3) *For each f in $C(X)$ and each $\mathscr{F}_k$ in $\mathscr{A}_F(X)$ there is an element Z_f of $\mathscr{F}_k$ such that f is bounded on Z_f.*

Proof.

(1) $\Rightarrow$ (2) This is obvious.

(2) $\Rightarrow$ (3) If $A(X)$ is identical with the family of real z-ultrafilters on X, then every $\mathscr{F}_k$ in $\mathscr{A}_F(X)$ is contained in a unique z-ultrafilter with c.i.p. If some f in $C(X)$ is unbounded on every element of $\mathscr{F}_k$, then for every positive integer n, $Z_n = \{x : |f(x)| \geqq n\}$ has nonvoid intersection with every Z in $\mathscr{F}_k$. Hence $\mathscr{F}_k \cup \{Z_n : n = 1, 2, \cdots\}$ generates a z-filter which contains $\mathscr{F}_k$ and does not have c.i.p. This is a contradiction.

(3) $\Rightarrow$ (1) Let $\mathscr{F}_{k(q)}$ be an element of $\mathscr{A}_F(X)$. We show that

A^q is real. Let f^* be the Stone extension of f into the one-point-compactification of the real numbers (see § 7.5 of [7]). Then by Theorems 5.7 and 7.6 in [7] we have $f^*(q) \neq \infty$, since Z_f is an element of A^q for each f in $C(X)$. Hence A^q is real, πX is contained in νX.

We shall continue to use f^* to represent the Stone extension of a function f in $C(X)$ into the one-point-compactification of the real numbers.

THEOREM 32. *Let X be a Hausdorff completely regular space, and let f be in $C(X)$. Then there is a continuous extension f^π in $C(\pi X)$ of f if and only if the z-filter $f^\sharp(\mathscr{F}_k) = \{Z$, a zero set in the real numbers: $f^\leftarrow[Z]$ is in $\mathscr{F}_k\}$ converges, for every $\mathscr{F}_k$ in $\mathscr{A}_F(X)$.* (For properties of $f^\sharp$ see §§ 4.12 and 10.17 of [7].)

Proof. Necessity. Since f^π is continuous, the filter generated by the image of the neighborhood system of y_k in πX converges. Now $\mathscr{F}_k$ is a base for a filter $\mathscr{F}$ which contains the neighborhood system of y_k in πX. Therefore the filter generated by $\{S : S = f^\pi[F]$ where F is in $\mathscr{F}_k\}$ converges to the real number $r = f^\pi(y_k)$. Hence every zero set neighborhood of r is an element of this filter. Since every F in $\mathscr{F}_k$ is contained in X, the filter generated by $\{S : S = f^\pi[F], F$ is in $\mathscr{F}_k\}$ equals the filter generated by $\{S : S = f[F], F$ is in $\mathscr{F}_k\}$. For each zero set neighborhood $V(r)$ of r, $f^{\pi\leftarrow}[V(r)]$ is a zero set neighborhood of y_k, so $f^{\pi\leftarrow}[V(r)] \cap X$ is in $\mathscr{F}_k$, and equals $f^\leftarrow[V(r)]$ giving $f^\sharp(\mathscr{F}_k)$ convergent.

Sufficiency. To show f has an extension to πX it is sufficient to show that the limit of the filter generated by $S = \{f[X \cap V] : V$ is a neighborhood of y_k in $\pi X\}$ exists for every k in K. Let r be the limit of $f^\sharp(\mathscr{F}_k)$. Then if $V(r)$ is any zero set neighborhood of r, there is a Z in $f^\sharp(\mathscr{F}_k)$ which is contained in int $V(r)$. Hence $f^\leftarrow[Z]$ is in $\mathscr{F}_k$ so that $f^\leftarrow[V(r)]$ contains an element of $\mathscr{F}_k$, and $f(f^\leftarrow[V(r)]) = V(r)$ is an element of S, and we are through.

DEFINITION 7. For a Hausdorff completely regular space X, define $C^\pi(X) = \{f$ in $C(X) : f^*(p) \neq \infty$ for every p in $\pi X\}$.

We note without proof that C^π is identical to $\{f$ in $C(X) : \pi X$ is contained in $\nu_f X\}$ where $\nu_f X$ is the realcompact subspace of βX equal to $\{p$ in $\beta X : f^*(p) \neq \infty\}$. Hence $\pi X = \bigcap \{\nu_f X : f$ is in $C^\pi\}$. (See 8B2 and 8B3 in [7].)

THEOREM 33. *Let f be in $C(X)$. Then f has an extension f^π in $C(\pi X)$ if and only if f is in $C^\pi(X)$.*

Proof. Necessity. If f has an extension f^π in $C(\pi X)$, then f^* and f^π must agree on the dense subset X of πX. Both f^π and f^* may be considered as functions into the one-point-compactification of the real numbers. Hence both are defined for all of πX and must agree. Since $f^\pi(p) \neq \infty$ for all p in πX, neither is f^*, and f is in $C^\pi(X)$.

Sufficiency. If f is in $C^\pi(X)$, f^* is a real valued continuous extension of f to all of πX.

COROLLARY 33A. *The family $C^\pi(X)$ is a subring of $C(X)$ and is isomorphic to $C(\pi X)$.*

THEOREM 34. *Let f be an element of $C(X)$. Then $Z_f = \{p$ in $\beta X : f^*(p) \neq \infty\}$ is a zero set in βX.*

Proof. Define $g = |f| \vee 1$. Then $Z_f = Z_g$ and it will suffice to show the result for functions bounded away from zero and positive. Assume $f \geqq 1$ and define $h = 1/f$ then h is in $C^*(X)$ and has an extension h^β in $C(\beta X)$. The functions h^β and h^* are equal. Now $f^* \mid (\beta X - Z(h^\beta))$ and $1/h^\beta \mid (\beta X - Z(h^\beta))$ are both extensions of f and must be equal. Hence Z_f and $Z(h^\beta)$ are equal, for otherwise $h^\beta f^*$ would equal 1 for some point in Z_f or $Z(h^\beta)$.

THEOREM 35. *If $\beta X - \pi X$ is not empty, it has carinality greater than or equal to 2^c.*

Proof. Let f be in $C^\pi(X)$ and let Z_f be nonvoid. Then Z_f is contained in $\beta X - \pi X$. Since Z_f intersects the closure of πX but not $\pi X \cup X$ we apply Theorem 9.4 in [7] to get the result.

DEFINITION 8. A function f in $C(X)$ bounded on the complement of an X α-paracompact set is *essentially bounded.*

THEOREM 36. *A function f in $C(X)$ has an extension f^π in $C(\pi X)$ bounded on the remainder if and only if it is essentially bounded.*

Proof. Necessity. If f in $C(X)$ has the extension f^π in $C(\pi X)$ with bound M on the remainder, then for each y_k in the remainder the set $V_k(\varepsilon) = f^{\pi\leftarrow}[(f(y_k) - \varepsilon/2, f(y_k) + \varepsilon/2)]$ is an open neighborhood of y_k for $\varepsilon > 0$. The set $\pi X - \cup\{V_k(\varepsilon) : y_k$ is in $\pi X - X\}$ is an X α-paracompact set which has f bounded by $M + \varepsilon/2$ on its complement.

Sufficiency. This is simply Lemma 22.

THEOREM 37. *If $\pi X - X$ is pseudocompact, then the family of* *essentially bounded functions on X is $C^{\pi}(X)$.*

We wish to give some results now which present properties of α-filter bases in relation to various other properties.

THEOREM 38. *The maximal α-filter bases are prime z-filters.*

Proof. This is an easy computation and is omitted.

We recall that $\mathscr{F}_{k(q)}$ is the unique free α-filter base on X contained in the z-ultrafilter A^q.

THEOREM 39. *If $\mathscr{F}_k$ is in $\mathscr{A}(X)$, then $\mathscr{F}_{k(q)}$ contains $Z[0^q]$* (for properties of 0^q see 4I, 7.12 – 7.15, 7H *in* [7]).

Proof. Every prime ideal in $C(X)$ is contained in a maximal ideal M^p, and contains an 0^p for unique p. $Z^{\leftarrow}[\mathscr{F}_k]$ is a prime ideal. Hence $Z[Z^{\leftarrow}[\mathscr{F}_{k(q)}]]$ contains $Z[0^q]$.

COROLLARY 39A. *If X is a P-space, then $\mathscr{F}_{k(p)} = Z[0^p]$* (for properties of P-spaces see 4J, K, L, 5P, and 7L in [7]).

COROLLARY 39B. *If X is a P-space, then $cl_{\pi X}Z = Z$ for α-paracompact Z.*

THEOREM 40. *Let $\mathscr{F}_k$ be a free α-filter base such that the open cover consisting of the complements of members of $\mathscr{F}_k$ has no locally finite open refinement which covers X. Then the intersection of any subfamily $\mathscr{F}^*$ of $\mathscr{F}_k$, which has locally finite complements, is a member of $\mathscr{F}_k$.*

Proof. Let $\mathscr{F}_k$ be an element of $\mathscr{A}_F(X)$, and let $\mathscr{G}$ be $\{G : G = X - Z, Z \text{ in } \mathscr{F}_k\}$. Since $\mathscr{G}$ has no locally finite refinement which covers X, every locally finite subfamily $\mathscr{G}^*$ of $\mathscr{G}$ must have $\bigcap \{Z : Z = X - G, G \text{ in } \mathscr{G}^*\}$ to be nonvoid. Let Z^* be such an intersection. We show Z^* to be in F_k. If Z^* is not in F_k, then there is an F in $\mathscr{F}_k$ such that $F \cap Z^*$ is α-paracompact. Then $\mathscr{H} = \{G : G \text{ is in } \mathscr{G}^* \text{ or } G = X - F\}$ is an open locally finite refinement of $\mathscr{G}$. There is an open locally finite refinement $\mathscr{H}^*$ of $\mathscr{G}$ which covers $F \cap Z^*$. Then $\mathscr{H} \cup \mathscr{H}^*$ is a locally finite open refinement of $\mathscr{G}$ which covers X, and we have reached a contradiction.

THEOREM 41. *Let X be a space such that for every $\mathscr{F}_k$ in $\mathscr{A}_k(X)$, the family $\{G : G = X - Z, Z \text{ in } \mathscr{F}_k\}$ has no open locally finite refinement which covers X. Then for any paracompactification Y of X every*

$\mathscr{F}_k$ in $\mathscr{A}(X)$ must converge in Y.

Proof. Suppose that $\mathscr{F}_k$ in $\mathscr{A}_F(X)$ does not converge in Y. Let $\mathscr{F} = \{F : F = \mathrm{cl}_y Z, Z \text{ in } \mathscr{F}_k\}$. Then since $\mathscr{F}_k$ does not converge in Y, $\{G : G = y - F, F \text{ in } \mathscr{F}\}$ is a Y open cover of Y. Since Y is paracompact, there is a Y open locally finite refinement $\mathscr{G}^* = \{G_j^* : j \text{ is in } J\}$ which covers Y. Then $\{G_j : G_j = G_j^* \cap X, G_j^* \text{ in } \mathscr{G}^*\}$ is an X open locally finite refinement of $\{G : G = X - Z, Z \text{ in } \mathscr{F}_k\}$ which covers X.

THEOREM 42. *Let X be a space such that for every $\mathscr{F}_k$ in $\mathscr{A}_F(X)$, the family $\{G : G = X - Z, Z \text{ in } \mathscr{F}_k\}$ has no open locally finite refinement which covers X. Then if Y is any Hausdorff paracompactification of X, there exists a continuous function $f : \pi X \rightarrow Y$ that holds X pointwise fixed.*

Proof. Define the function $g : X \rightarrow Y$ by $g(x) = x$. Let $\bar{g}$ be the Stone extension of g taking βX into βY and define $X_0 = \bar{g}^{\leftarrow}[Y]$. Let $\bar{f} = \bar{g} \mid X_0$. Then $\bar{f}$ is continuous, onto Y, and holds X fixed. Since $\mathscr{F}_k$ in $\mathscr{A}_F(X)$ must converge in Y to some point y, $\mathscr{F}_k$ must converge to some element in $\bar{f}^{\leftarrow}(y)$. Therefore, πX is contained in X_0, and $f = \bar{f} \mid \pi X$ is the desired function.

COROLLARY 42A. *If X is a space such that for every $\mathscr{F}_k$ in $\mathscr{A}_F(X)$, the family $\{G : G = X - Z, Z \text{ in } \mathscr{F}_k\}$ has no locally finite open refinement which covers X, then πX is the smallest paracompactification contained in βX.*

THEOREM 43. *Let X be dense in a Hausdorff completely regular extension Y such that all $\mathscr{F}_k$ in $\mathscr{A}_F(X)$ converge in Y. Then Y contains a paracompactification of X. Further, if $Y - X$ consists only of limit points of the $\mathscr{F}_k$ in $\mathscr{A}_F(X)$ then Y is paracompact.*

Proof. Let $f : X \rightarrow Y$ be the identity map, let $\bar{f}$ be the Stone extension of f into βY, and let $X_0 = \bar{f}^{\leftarrow}[Y]$. If $f_0 = \bar{f} \mid X_0$, then f_0 is a closed continuous extension of f onto Y, and $f[X_0 - X]$ is contained in $Y - X$. (See §§ 10.13 and 10.15 in [7].)

We show that X_0 contains πX. Let p be an element in $\pi X - X$, let y_p be an element of Y such that A^p converges to y_p in Y, and let $y = f(p)$. Suppose that y is not identical to y_p. Then there are disjoint open neighborhoods $V(y)$ and $V(y_p)$ in Y. If $V(p) = \bar{f}^{\leftarrow}[V(y)]$, $G = \bar{f}^{\leftarrow}[V(y_p)]$, then we must have $V(p) \cap G$ empty. Since A^p converges to y_p there is a Z_p in A^p that is contained in $V(y_p)$, and hence G, since f is the identity. But $V(p)$ is a neighborhood of p in πX, so that there is an element Z_y of A^p contained in $V(p)$. Hence Z_y does not intersect Z_p which is impossible so $y_p = y$, and πX is contained in X_0.

The set $f_0[\pi X]$ contains X, and is β-paracompact in Y, since f_0 is a closed continuous function. Therefore, $f_0[\pi X]$ is a paracompactification of X contained in Y.

The second assertion follows, since if Y satisfies the given condition, $f_0[\pi X] = Y$, and $X_0 = \pi X$.

THEOREM 44. *Let X be a space such that for every $\mathscr{F}_k$ in $\mathscr{A}_F(X)$, the family $\{G : G = X - Z, Z \text{ in } \mathscr{F}_k\}$ has no open locally finite refinement which covers X, and let X be dense in a Hausdorff completely regular extension Y. Then Y contains a paracompactification of X if and only if $Y - X$ contains a limit point for every $\mathscr{F}_k$ in $\mathscr{A}_F(X)$.*

Proof. The proof follows from Theorems 41 and 43.

THEOREM 45. *If Y is a Hausdorff paracompactification of X such that some $\mathscr{F}_k$ in $\mathscr{A}_F(X)$ does not converge in Y, then there is an X locally finite partition of unity $\mathscr{L}$ contained in $Z_X^{\leftarrow}[\mathscr{F}_k]$.*

Proof. Suppose that $\mathscr{F}_k$ is in $\mathscr{A}_F(X)$ and does not converge in Y. Then $\{\mathrm{cl}_Y Z : Z \text{ is in } \mathscr{F}_k\}$ has void intersection and generates the free z-filter $\mathscr{F}^*$ on Y. Let $\mathscr{G} = \{G : G = Y - Z, Z \text{ in } \mathscr{F}^*\}$. Then $\mathscr{G}$ is an open cover of Y. Since Y is Hausdorff, there is a locally finite partition of unity subordinate to $\mathscr{G}$, call it $\mathscr{E}$. If f is in $\mathscr{F}$, then $Z_Y(f)$ contains some Z from $\mathscr{F}.^*$ Hence $f \mid X$ is in $Z_X^{\leftarrow}[\mathscr{F}_k]$, and there is the subfamily $\mathscr{L} = \{f \mid X : f \text{ is in } \mathscr{E}\}$ of $Z_X^{\leftarrow}[\mathscr{F}_k]$ which is an X locally finite partition of unity.

THEOREM 46. *Each $\mathscr{F}_k$ in $\mathscr{A}_F(X)$ has the property that the family $\{G : G = X - Z, Z \text{ is in } \mathscr{F}_k\}$ has no locally finite open refinement which covers X only if $Z^{\leftarrow}[\mathscr{F}_k]$ contains no locally finite partition of unity.*

Proof. A locally finite partition of unity which is contained in $Z^{\leftarrow}[\mathscr{F}_k]$ yields a locally finite open refinement of $\{G : G = X - Z, Z \text{ in } \mathscr{F}_k\}$ by cozero sets.

THEOREM 47. *A space X is nonparacompact if and only if for each free z-filter $\mathscr{F}$ there is a family $\mathscr{G}^*$ which consists of complements of elements of $\mathscr{F}$ and has no locally finite open refinement.*

Proof. Necessity. Let $\mathscr{F}$ be a free z-filter, and let $\mathscr{G} = \{G_j : j \text{ is in } J\}$ be an open cover of X which has no open locally finite refinement. For each x in X, let G_x be an element of $\mathscr{G}$ which contains x. Let Z be a zero set which contains x in its complement, and such that the complement of Z is contained in G_x. For each x in X let Z_x be that element of $\mathscr{F}$ which contains x in its complement. Then the family $\mathscr{G}^* = \{X - F_x : F_x = Z \cup Z_x, x \text{ in } X\}$ is the desired family of open sets.

Sufficiency. This is immediate.

J. Van der Slot [12] and H. Herrlich ([8], [9], [10]) have done work pertaining to extensions of spaces. If β is a property of topological spaces a *β-extension* of X is a space γX containing a dense homeomorphic image of X and having property β. A β-extension γX of X is *maximal* if for each continuous function f of X into a space Y having property β there is a continuous extension of f to all of γX. A space X is *β-regular* if it is homeomorphic to a subspace of a space which is the product of spaces each having property β. The following theorem is due to Van der Slot [12].

THEOREM 48. *Let β be a property possessed by Hausdorff spaces. Then a β-regular Hausdorff space has a maximal β-extension if and only if β is closed hereditary and productive.*

From the above theorem and the fact that paracompactness is not productive we have the following:

THEOREM 49. *A Hausdorff space X does not in general possess a maximal paracompactification.*

It would be of interest to obtain a characterization of those spaces X which have $A_F(X)$ such that if F is in $A_F(X)$ then $\{G : G = X - Z, Z \text{ in } F\}$ has no locally finite open refinements. It would also be interesting to know if it is true in general that πX is the smallest paracompactification of X contained in βX.

REFERENCES

1. R. A. Alo and H. L. Shapiro, *Paracompact subspaces*, Acta Math., **21** (1970).
2. C. E. Aull, *A certain class of topological spaces*, Prace Matematyczne, **11** (1967), 49-53.
3. ———, *Collectionwise normal subsets*, To appear in J. London Math. Soc.
4. ———, *Paracompact subsets*, Proceedings of the Second Prague Topological Symposium, (1966), 45-51.
5. ———, *Remarks on countable paracompactness*, Proc. Japan Acad., **54** (1968), 125-126.
6. J. A. Dieudonné, *Une généralization des espaces compacts*, J. Math. Pures Appl., **23** (1944), 65-76.
7. L. Gillman and M. Jerison, *Rings of Continuous Functions*, D. Van Nostrand.
8. H. Herrlich, *E-kompakte raume*, Math. Zeitschr, **96** (1967), 228-255.

9. ———, *Topologische Reflexionen and Coreflexion*, Lecture Notes in Mathematics, 78, Springer-Verlag, Berlin, 1968.

10. H. Herrlich and J. Van der Slot, *Properties which are closely related to compactness*, Indag, Math., **29** (1967), 524–530.

11. E. A. Michael, *Another note on paracompact spaces*, Proc. Amer. Math. Soc., **8** (1957), 822–823.

12. J. Van der Slot, *Some properties related to compactness*, Publication of Mathematisch Centrum, Amsterdam, 1968.

13. J. D. Wine, *Locally paracompact spaces*. To appear.

Received July 14, 1972 and in revised form April 5, 1973.

UNIVERSITY OF WISCONSIN-LA CROSSE

Pacific Journal of Mathematics

Vol. 49, No. 1 May, 1973