Pacific Journal of Mathematics

IN THIS ISSUE—

Vol. 36, No. 3 **BadMonth, 1971**

PACIFIC JOURNAL OF MATHEMATICS

Mathematical papers intended for publication in the *Pacific Journal of Mathematics* should be in typed form or offset-reproduced, (not dittoed), double spaced with large margins. Underline Greek letters in red, German in green, and script in blue. The first paragraph or two must be capable of being used separately as a synopsis of the entire paper. The editorial "we" must not be used in the synopsis, and items of the bibliography should not be cited there unless absolutely necessary, in which case they must be identified by author and Journal, rather than by item number. Manuscripts, in duplicate if possible, may be sent to any one of the four editors. Please classify according to the scheme of Math. Rev. Index to Vol. **39**. All other communications to the editors should be addressed to the managing editor, Richard Arens, University of California, Los Angeles, California, 90024.

50 reprints are provided free for each article; additional copies may be obtained at cost in multiples of 50.

The *Pacific Journal of Mathematics* is published monthly. Effective with Volume 16 the price per volume (3 numbers) is \$8.00; single issues, \$3.00. Special price for current issues to individual faculty members of supporting institutions and to individual members of the American Mathematical Society: \$4.00 per volume; single issues \$1.50. Back numbers are available.

Subscriptions, orders for back numbers, and changes of address should be sent to Pacific Journal of Mathematics, 103 Highland Boulevard, Berkeley, California, 94708.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS, A NON-PROFIT CORPORATION

Printed at Kokusai Bunken Insatsusha (International Academic Printing Co., Ltd.), 7-17, Fujimi 2-chome, Chiyoda-ku, Tokyo, Japan.

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

ON DOMINATED EXTENSIONS IN LINEAR SUBSPACES OF $\mathscr{C}_C(X)$

E. M. ALFSEN AND B. HIRSBERG

The main result is the following: Given a closed linear subspace A of $\mathscr{C}_C(X)$ where X is compact Hausdorff and A contains constants and separates points, and let F be a compact subset of the Choquet boundary $\partial_A X$ with the property that the restriction to F of every A-orthogonal boundary measure remains orthogonal. If $a_0 \in A|_F$ and $a_0 \leqq \Psi|_F$ for some strictly positive A-superharmonic function Ψ, then a_0 can be extended to a function $a \in A$ such that $a \leqq \Psi$ on all of X. It is shown how this result is related to various known dominated extension-and peak set-theorems for linear spaces and algebras. In particular, it is shown how it generalizes the Bishop-Rudin-Carleson Theorem.

The aim of this paper is to study extensions within a given linear subspace A of $\mathscr{C}_C(X)$ of functions defined on a compact subset of the Choquet boundary $\partial_A X$, in such a way that the extended function remains dominated by a given A-superharmonic function Ψ. (Precise definitions follow). Our main result is the possibility of such extensions for all functions in $A|_F$ provided F satisfies the crucial requirement that the restriction to F of every orthogonal boundary measure shall remain orthogonal (Theorem 4.5). Taking $\Psi \equiv 1$ in this theorem we obtain that F has the norm preserving extension property (Corollary 4.6). This was first stated by Björk [5] for a real linear subspace A of $\mathscr{C}_R(X)$ and for a metrizable X. A geometric proof of the latter result was given by Bai Andersen [3]. In fact, he derived it from a general property of split faces of compact convex sets, which he proved by a modification of an inductive construction devised by Pelczynski for the study of simultaneous extensions within $\mathscr{C}_R(X)$ [12]. Our treatment of the more general extension property proceeds along the same lines as Bai Andersen's work. It depends strongly upon the geometry of the state space of A, and Bai Andersen's construction is applied at an essential point in the proof. Note however, that this is no mere translation of real arguments. The presence of complex orthogonal measures seems to present a basically new situation. Applying arguments similar to those indicated above, we obtain a general peak set-and peak point criterion (Theorem 5.4 and Corollary 5.5) of which the latter has been proved for real spaces by Björk [6]. In § 6 (Theorem 6.1) it is shown how the Bishop-Rudin-Carleson Theorem follows from the general extension theorem mentioned above. In § 7 we assume that A is a sup-norm algebra over X and study the interrelationship between our conditions on F and a condition introduced by Gamelin and Glicksberg [9], [10]. Finally we should like point out that some related investigations have been carried out recently by Briem [7]. However, his methods are rather different. The geometry of the state space is not invoked, but instead he applies in an essential way a measurable selection theorem of Rao [14].

We want to thank Bai Andersen for many stimulating discussions of the problems of the present paper. Also we are indebted to A. M. Davie for the counterexample at the end of § 7.

1. Preliminaries and notation. In this note X shall denote a compact Hausdorff space and A a closed, linear subspace of $\mathscr{C}_C(X)$, which separates the points of X and contains the constant functions.

The *state space* of A, i.e.

$$S = \{p \in A^* \mid p(1) = \|p\| = 1\} ,$$

is convex and compact in the w^*-topology. Since A separates the points of X, we have a homeomorphic embedding Φ of X into S, defined by

$$\Phi(x)(a) = a(x), \quad \text{all} \quad a \in A .$$

Similary we have an embedding Ψ of A into the space $A_C(S)$ of all complex valued w^*-continuous affine functions on S; namely

$$\Psi(a)(p) = p(a), \quad \text{all} \quad p \in S .$$

By taking real parts of the functions $\Psi(a)$ we obtain the linear space of those real valued w^*-continuous affine functions on S, which can be extended to real valued w^*-continuous linear functionals on A^*, and this space $A_R(S, A^*)$ is dense in the space $A_R(S)$ of all real valued affine w^*-continuous functions on S, [1, Cor. I. 1.5].

We shall denote by $M(X)$, resp. $M(S)$, the Banach space of all complex Radon measures on X, resp. S; by $M^+(X)$ resp. $M^+(S)$ the cone of positive (real) measures, and by $M_1^+(X)$ resp. $M_1^+(S)$ the w^*-compact convex set of probability measures. The set of extreme points of S will be denoted by $\partial_e S$, and the *Choquet boundary* of X with respect to A is defined as the set

$$\partial_A X = \{x \in X \mid \Phi(x) \in \partial_e S\} .$$

From [13, p. 38] it follows that $\partial_e S \subset \Phi(X)$ so that Φ maps $\partial_A X$ homeomorphically onto $\partial_e S$.

A measure $\mu \in M(S)$ is said to be a *boundary measure on* S if the total variation $|\mu|$ is a maximal measure in Choquet's ordering of positive measures [1, ch. I, § 3], [13, p. 24]. A boundary measure is supported by $\overline{\partial_e S}$ [1, Prop. I. 4.6]. For a metrizable X (and S) a measure $\mu \in M(S)$ is a boundary measure if and only if $|\mu|(S \backslash \partial_e S) = 0$. We shall denote by $M(\partial_e S)$ the set of boundary measures on S (abuse of language). Observe that if $\mu \in M(\partial_e S)$, then the real and imaginary parts of μ are both boundary measures. The set of *boundary measures on* X is defined by

$$M(\partial_A X) = \{\mu \in M(X) \mid \Phi\mu \in M(\partial_e S)\} ,$$

where $\Phi\mu$ denotes the transport of the measure μ on X to a measure on S. For a metrizable X a measure μ on X belongs to $M(\partial_A X)$ if and only if $|\mu|(X \backslash \partial_A X) = 0$.

For every $\mu \in M_1^+(S)$ we shall use the symbol $r(\mu)$ to denote the *barycenter* of μ, i.e., the unique point in S such that $a(r(\mu)) = \mu(a)$ for all $a \in A_R(S)$. The Choquet-Bishop de Leeuw Theorem states that each point in S is the barycenter of a maximal (boundary) probability measure [1, Th. I. 4.8]. Accordingly we shall denote by $M_p^+(\partial_e S)$ the nonempty set of maximal (boundary) probability measures on S with barycenter $p \in S$. For $x \in X$ we define $M_x^+(\partial_A X)$ to be the set of all $\mu \in M_1^+(X)$ such that $\Phi\mu \in M_{\Phi(x)}^+(\partial_e S)$. Equivalently, $M_x^+(\partial_A X)$ consists of all $\mu \in M_1^+(\partial_A X)$ such that

$$a(x) = \int a d\mu \quad \text{all} \quad a \in A ,$$

i.e., μ *represents* x with respect to A. Also we denote by $M_x^+(X)$ the set of probability measures on all of X which represents x in this way. Similary we denote by $M_p^+(S)$ the set of probability measures on S with barycenter p. The *annihilator* of A in $M(X)$ is the set

$$A^\perp = \{\mu \in M(X) \mid \mu(a) = 0 \quad \text{all} \quad a \in A\} .$$

Finally we shall use the symbol $\mathscr{B}(X)$ to denote the class of all complex valued bounded Borel functions on X.

2. A dominated extension theorem.

We start by proving a general dominated extension theorem, which may be of some independant interest. In this connection we give the following:

DEFINITION 2.1. $\mathscr{A}$ is the class of all $f \in \mathscr{B}(X)$ such that

$$\mu(f) = 0 \quad \text{all} \quad \mu \in A^\perp . \tag{2.1}$$

Clearly $A \subset \mathscr{A}$.

THEOREM 2.2. *Let F be a closed subset of X for which $A|_F = \{a|_F \mid a \in A\}$ is closed in $\mathscr{C}_C(F)$; let $a_0 \in A|_F$ and let $\varphi\colon X \to \boldsymbol{R}^+ \cup \{\infty\}$ be* *a strictly positive l.s.c. function such that* $|a_0(x)| < \varphi(x)$ *for all* $x \in F$.

Now, if there exists a function $\bar{a}_0 \in \mathscr{A}$ *such that*

$$\bar{a}_0|_F = a_0, \quad |\bar{a}_0(x)| < \varphi(x) \quad \text{all} \quad x \in X \tag{2.2}$$

then there exists a function in A *with the same properties.*

Proof. Without lack of generality we can assume that φ is a bounded function with values in $\boldsymbol{R}^+$, and we assume for contradiction that

$$a_0 \notin G|_F = \{a|_F \mid a \in G\} , \tag{2.3}$$

where

$$G = \{a \in A \mid |a(x)| < \varphi(x)\} . \tag{2.4}$$

Since φ is l.s.c., G is an open subset of A. Since $A|_F$ is closed in $\mathscr{C}_C(F)$, we may apply the Open Mapping Theorem to the restriction map $R_F: A \to A|_F$. Hence $G|_F$ is an open subset of $A|_F$. Furthermore $G|_F$ is convex and circled. By the Hahn-Banach Theorem we can find a measure $\nu \in M(X)$ with $\operatorname{supp} \nu \subset F$ such that

$$\nu(a_0) \geqq 1 \geqq |\nu(b_0)| \quad \text{all} \quad b_0 \in G|_F . \tag{2.5}$$

Now we consider $\mathscr{C}_C(X)$ equipped with the norm

$$||f||_\varphi = \sup\left\{\frac{|f(x)|}{\varphi(x)} \,\middle|\, x \in X\right\} , \tag{2.6}$$

and observe that this norm is topologically equivalent with the customary, uniform norm. The dual of $(\mathscr{C}_C(X), ||-||_\varphi)$ is seen to be $M(X)$ equipped with the norm $||\mu||_\varphi = ||\varphi \cdot \mu||$, where $(\varphi \cdot \mu)(f) = \mu(\varphi f)$ for all $f \in \mathscr{C}_C(X)$.

It follows from (2.5) that the linear functional ξ on $(A, ||-||_\varphi)$ defined by

$$\xi(a) = \nu(R_F a) \quad \text{all} \quad a \in A , \tag{2.7}$$

is bounded with norm $||\xi||_\varphi \leqq 1$. Now we extend ξ with preservation of φ-norm to a bounded linear functional on $(\mathscr{C}_C(X), ||-||_\varphi)$. This gives a measure $\mu \in M(X)$, such that

$$\xi(a) = \mu(a) \quad \text{all} \quad a \in A, \quad ||\varphi \cdot \mu|| = ||\xi||_\varphi \leqq 1 . \tag{2.8}$$

It follows from (2.2) and (2.8) that

$$|\mu(\bar{a}_0)| = |(\varphi \cdot \mu)(\varphi^{-1}\bar{a}_0)| < 1 . \tag{2.9}$$

From (2.7) and (2.8) it follows that $\mu - \nu \in A^\perp$, and since $\bar{a}_0 \in \mathscr{A}$ we shall have

$$\left|\int_X \bar{a}_0 d\mu\right| = \left|\int_X \bar{a}_0 d\nu\right| = \int_F a_0 d\nu \geqq 1 . \tag{2.10}$$

This contradicts (2.9) and the proof is complete.

3. Applications of the geometry of the state space. We shall consider compact subsets F of $\partial_A X$ satisfying one or the other of the following two requirements:

$$\mu \in M(\partial_A X) \cap A^\perp \Rightarrow \mu|_F \in A^\perp \tag{A.1}$$

$$\mu \in M(\partial_A X) \cap A^\perp \Rightarrow \mu(F) = 0 . \tag{A.2}$$

We assume first (A.1). We also agree to write $S_F = \overline{co}(\Phi(F))$, and we observe that there is a canonical embedding Ψ_F of $A|_F$ into $A_C(S_F)$, defined by

$$\Psi_F(a_0)(p) = p(a), \quad \text{all} \quad p \in S_F \tag{3.1}$$

where $a \in A$; $a|_F = a_0$. In fact, it follows by the integral form of the Krein-Milman Theorem that p can be expressed as the barycenter of a probability measure on $\Phi(F)$, and hence that the particular choice of a is immaterial.

For every $a_0 \in A|_F$ we define

$$\bar{a}_0(x) = \int_F a_0 d\mu_x, \quad x \in X,\ \mu_x \in M_x^+(\partial_A X) , \tag{3.2}$$

and

$$\bar{\bar{a}}_0(p) = \int_{S_F} \Psi_F(a_0) d\mu_p, \quad p \in S,\ \mu_p \in M_p^+(\partial_e S) , \tag{3.3}$$

and we note that these definitions are legitimate by virtue of (A.1). We also note that $\mu_p(S_F) = \mu_p(\Phi(F))$ for all $p \in S$ and $\mu_p \in M_p^+(\partial_e S)$ [3, Lem. 1].

Clearly $\bar{a}_0$ is an extension of a_0 to a function defined on all of X; and if we think of Φ as an imbedding of X into S, then $\bar{\bar{a}}_0$ will in turn be an extension of $\bar{a}_0$ to a function defined on all S. More specifically, for every $\mu_x \in M_x^+(\partial_A X)$ the transported measure $\Phi\mu_x$ is in $M_{\Phi(x)}^+(\partial_e S)$ and so

$$\bar{\bar{a}}_0(\Phi(X)) = \int_{S_F} \Psi_F(a_0) d(\Phi\mu_x) = \int_F \Psi_F(a_0) \circ \Phi d\mu_x = \int_F a_0 d\mu_x ,$$

which entails

$$\bar{\bar{a}}_0 \circ \Phi = \bar{a}_0 . \tag{3.4}$$

LEMMA 3.1. *If F satisfies* (A.1) *and $a_0 \in A|_F$, then $\bar{a}_0 \in \mathscr{A}$.*

Proof. Let $\lambda = ||a_0||_F$ and define

$$a_1 = \operatorname{Re} \Psi_F(a_0) + \lambda, \quad a_2 = \operatorname{Im} \Psi_F(a_0) + \lambda .$$

Then $a_1, a_2 \in A_{\boldsymbol{R}}(S_F)^+$ and for any $p \in S$ and $\mu_p \in M_p^+(\partial_e S)$

$$\bar{\bar{a}}_0(p) = \int_{S_F} \Psi_F(a_0) d\mu_p = \int_{S_F} a_1 d\mu_p + i \int_{S_F} a_2 d\mu_p - \lambda\mu_p(S_F) - i\lambda\mu_p(S_F) .$$

At this point we shall appeal to the geometric theory of compact convex sets. We recall that a face Q of S is said to be *split* if the complementary face Q' (=the union of all faces disjoint from Q) is convex (hence a face) and every element of S can be expressed by a *unique* convex combination of an element of Q and an element of Q'. It is known ([1, Th. II. 6.12], [1, Th. II. 6.18], see also [2, Th. 3.5]) that for a closed face Q of S the following statements are equivalent:

(i) If a real measure $\mu \in M(\partial_e S)$ annihilates all continuous affine functions, then $\mu|_Q$ has the same property.

(ii) Q is a split face.

(iii) The u.s.c. concave upper envelope $\widehat{b\chi_Q}$ of the function $b\chi_Q$ which is equal to b on Q and 0 on $S \backslash Q$, is *affine* for every $b \in A_{\boldsymbol{R}}(Q)^+$.

It follows from the requirement (A.1) that S_F is a *split face* of S, and hence that

$$\bar{\bar{a}}_0(p) = \widehat{a_1\chi_{S_F}}(p) + i\widehat{a_2\chi_{S_F}}(p) - \lambda\hat{\chi}_{S_F}(p) - i\lambda\hat{\chi}_{S_F}(p) ,$$

where all the functions on the right hand side are u.s.c. and *affine*. In particular $\bar{\bar{a}}_0$ is a Borel function, and it follows from (3.4) that $\bar{a}_0$ is a Borel function as well. Since the barycentric calculus applies to real valued u.s.c. affine functions on S [1, Cor. I 1.4], we shall have:

$$\bar{\bar{a}}_0(p) = \int_S \bar{\bar{a}}_0 d\mu_p, \quad p \in S,\ \mu_p \in M_p^+(S) . \tag{3.5}$$

Let $\mu \in A^\perp$ be arbitrary and decompose

$$\mu = \sum_{i=1}^{4} \alpha_i \mu_i , \tag{3.6}$$

where $\alpha_1 \in \boldsymbol{R}^+$, $\alpha_2 \in -\boldsymbol{R}^+$, $\alpha_3 \in i\boldsymbol{R}^+$, $\alpha_4 \in (-i)\boldsymbol{R}^+$ and $\mu_i \in M_1^+(X)$ for $i = 1, 2, 3, 4$. Let $p_i \in S$ be the barycenter $\Phi\mu_i$ and let $\sigma_i \in M_{p_i}^+(\partial_e S)$ for $i = 1, 2, 3, 4$.

Since $\overline{\partial_e S} \subseteqq \Phi(X)$ we can transport σ_i back to X by the map Φ^{-1}, and it follows that the measures $\mu_i - \Phi^{-1}\sigma_i$ are (real) orthogonal measures for $i = 1, 2, 3, 4$.

Writing

$$\tau = \sum_{i=1}^{4} \alpha_i \Phi^{-1} \sigma_i$$

we obtain $\tau \in M(\partial_A X)$ and $\mu - \tau \in A^\perp$. In fact for every $a \in A$,

$$\int ad(\mu - \tau) = \int_S \Psi(a)d(\Phi(\mu - \tau)) = \sum_{i=1}^{4} \alpha_i \int_S \Psi(a)d(\Phi\mu_i - \sigma_i) = 0 .$$

Since $\mu \in A^\perp$, we shall also have $\tau \in A^\perp$ and then $\tau|_F \in A^\perp$ by virtue of (A.1). Hence by (3.3), (3.4), (3.5):

$$\begin{aligned}\int_X \bar{a}_0 d\mu &= \int_X \bar{\bar{a}} \circ \Phi d\mu = \int_S \bar{\bar{a}}_0 d(\Phi\mu) = \sum_{i=1}^{4} \alpha_i \int_S \bar{\bar{a}}_0 d(\Phi\mu_i) \\ &= \sum_{i=1}^{4} \alpha_i \bar{\bar{a}}_0(p_i) = \sum_{i=1}^{4} \alpha_i \int_{S_F} \Psi_F(a_0) d\sigma_i = \int_{S_F} \Psi_F(a_0) d(\Phi\tau) \\ &= \int_F a_0 d\tau = 0 .\end{aligned}$$

Hence $\bar{a}_0 \in \mathscr{A}$, and the proof is complete.

We next turn to the less restrictive requirement (A.2). It follows by a slight modification of the proof of [1, Th. II. 6.12], that the requirement (A.2) implies that S_F is a *parallel face* of S and hence that the function $\hat{\chi}_{S_F}$ is *affine* [15, Th. 12].

For every $x \in X$ we define

$$\bar{\chi}_F(x) = \int_F 1 d\mu_x, \quad \mu_x \in M_x^+(\partial_A X) \tag{3.7}$$

and we note that this definition is legitimate by virtue of (A.2). For $x \in X$ and $\mu_x \in M_x^+(\partial_A X)$ we shall have:

$$\hat{\chi}_{S_F}(\Phi(x)) = \int_{S_F} 1 d(\Phi\mu_x) = \int_F 1 d\mu_x = \bar{\chi}_F(x)$$

which entails

$$\hat{\chi}_{S_F} \circ \Phi = \bar{\chi}_F . \tag{3.8}$$

Applying (3.8) and proceeding as in the proof of Lemma 3.1, we can prove

LEMMA 3.2. *If F satisfies (A.2), then $\bar{\chi}_F \in \mathscr{A}$*

4. Extensions dominated by A-superharmonic functions.

We now proceed to the main theorem, but first we give some definitions.

DEFINITION 4.1. A function $\Psi: X \to \boldsymbol{R} \cup \{\infty\}$ is said to be *A-superharmonic* if it satisfies

(i) Ψ *l.s.c.*

(ii) $\Psi(x) \geqq \int_X \Psi d\mu_x$, all $x \in X$ and $\mu_x \in M_x^+(X)$.

DEFINITION 4.2. Let F be a compact subset of X. F has the *almost norm preserving extension property*, if for each $\varepsilon > 0$ and $a_0 \in A|_F$ there exists a function $a \in A$ such that

$$a|_F = a_0, \quad ||a||_X \leqq ||a_0||_F + \varepsilon . \tag{4.1}$$

If ε can be taken to be zero in (4.1), then F has the *norm preserving extension property*.

We shall need a criterion for the almost norm preserving extension property, which is due to Gamelin [9, p. 281] and Glicksberg [10, p. 420] (cf. also Curtis [8]). For the sake of completeness we present a short proof.

LEMMA 4.3. *A closed subset F of X has the almost norm preserving extension property if for each $\sigma \in A^\perp$:*

$$\inf_{\nu \in (A|_F)^\perp} ||\sigma|_F + \nu|| \leqq ||\sigma|_{X \setminus F}|| . \tag{4.2}$$

Proof. The almost norm preserving extension property is tantamaunt to the equality of the uniform norm on $A|_F$ and the *extension norm*:

$$||a_0||_{\text{ext.}} = \inf \{||a||_X \mid a \in A, \, a|_F = a_0\} .$$

In this norm $A|_F$ is isometrically isomorphic to the quotient space $A/F^\perp$ where $F^\perp = \{a \in A \mid a \equiv 0 \text{ on } F\}$; and we are to prove that the canonical imbedding $\rho \colon A/F^\perp \to A|_F$ is an isometry from the quotient norm to the uniform norm. By duality (i.e., by Hahn-Banach) we may as well prove that the transposed map ρ^* is an isometry. Representing the occuring functionals by measures, we can translate this statement into

$$\inf_{\sigma \in A^\perp} ||\mu + \sigma|| = \inf_{\nu \in (A|_F)^\perp} ||\mu + \nu||, \quad \text{all} \quad \mu \in M(F) . \tag{4.3}$$

To prove that (4.2) implies (4.5), we consider measures $\mu \in M(F)$, $\sigma \in A^\perp$ and an arbitrary $\varepsilon > 0$. Also we can choose $\nu_0 \in (A|_F)^\perp$ such that

$$||\sigma|_F - \nu_0|| \leqq \inf_{\nu \in (A|_F)^\perp} ||\sigma|_F - \nu|| + \varepsilon \leqq ||\sigma | X \setminus F|| + \varepsilon .$$

Then

$$\begin{aligned} ||\mu - \sigma|| = ||\mu - \sigma|_F|| + ||\sigma|_{X\setminus F}|| &\geqq ||\mu - \nu_0|| - ||\nu_0 - \sigma|_F|| \\ + ||\sigma|_{X\setminus F}|| \geqq ||\mu - \nu_0|| - \varepsilon &\geqq \inf_{\nu \in (A|_F)^\perp} ||\mu - \nu||| - \varepsilon , \end{aligned}$$

which completes the proof.

We remark for later purposes that for $\mu \in M(F)$:

$$(4.4) \qquad \sup\left\{\left|\int_F a_0 d\mu\right| \,\middle|\, \|a_0\|_F \leqq 1,\ a_0 \in A|_F\right\} = \inf_{\nu \in (A|F)^\perp} \|\mu - \nu\| .$$

PROPOSITION 4.4. *If F is a compact subset of $\partial_A X$ satisfying* (A.1), *then F has the almost norm preserving extension property.*

Proof. By Lemma 4.3 and the above remark (4.4), it suffices to prove that for every $\sigma \in A^\perp$:

$$\sup\left\{\left|\int_F a_0 d\sigma\right| \,\middle|\, \|a_0\|_F \leqq 1,\ a_0 \in A|_F\right\} \leqq \|\sigma|_{X\setminus F}\| .$$

Let $\sigma \in A^\perp$, and $a_0 \in A|_F$ with $\|a_0\|_F \leqq 1$. Applying Lemma 3.1 we obtain

$$0 = \sigma(\bar{a}_0) = \int_F a_0 d\sigma + \int_{F\setminus X} \bar{a}_0 d\sigma ,$$

such that

$$\left|\int_F a_0 d\sigma\right| = \left|\int_{X\setminus F} \bar{a}_0 d\sigma\right| \leqq \|\sigma|_{X\setminus F}\| ,$$

which completes the proof.

If F is a compact subset of $\partial_A X$ satisfying (A.1), then $A|_F$ is a closed subspace of $\mathscr{C}_C(F)$. In fact, $A|_F$ is isometrically isomorphic to $A/F^\perp$.

We are now able to state and prove the main theorem. The proof of this theorem is essentially based upon Theorem 2.1 and the technique developed by Bai Andersen [3].

THEOREM 4.5. *Let F be a compact subset of $\partial_A X$ satisfying* (A.1), *i.e.*

$$\mu \in M(\partial_A X) \cap A^\perp \Rightarrow \mu|_F \in A^\perp .$$

Let $a_0 \in A|_F$ and let ψ be a strictly positive A-superharmonic function on X such that $|a_0(x)| \leqq \psi(x)$ for all $x \in F$. Then there exists a function $a \in A$ such that

(i) $a|_F = a_0$,

(ii) $|a(x)| \leqq \psi(x)$ *all* $x \in X$.

Proof. Without loss of generality we may assume ψ to be bounded. Since F satisfies the requirement (A.1), $A|_F$ is closed and $\bar{a}_0 \in \mathscr{A}$.

Thus by Theorem 2.2 we can extend a_0 to a function $a_0' \in A$ such that $|a_0'(x)| < \varphi(x)$ for all $x \in X$, whenever φ is a bounded l.s.c. function on X such that $|\bar{a}_0(x)| < \varphi(x)$ for all $x \in X$.

Applying this to the function $\varphi_1 = 2\psi$, we can extend a_0 to a function $a_1 \in A$ such that $|a_1(x)| < 2\psi(x)$ for all $x \in X$.

Now define

$$\varphi_2 = 2\psi \wedge [2^2(\psi - 2^{-1}|a_1|)] .$$

The function φ_2 is strictly positive on all of X. For $x \in F$ we have $\varphi_2(x) = 2\psi(x)$, and hence for an arbitrary $x \in X$:

$$\begin{aligned}|\bar{a}_0(x)| &= \left|\int_F a_0 d\mu_x\right| \leqq \int_F |a_0| d\mu_x < \int_X 2^2(\psi - 2^{-1}|a_1|)d\mu_x \\ &= 2^2\left(\int_X \psi d\mu_x - 2^{-1}\int_X |a_1| d\mu_x\right) \leqq 2^2\left(\psi(x) - 2^{-1}\left|\int_X a_1 d\mu_x\right|\right) \\ &= 2^2(\psi(x) - 2^{-1}|a_1(x)|) .\end{aligned}$$

Hence $|\bar{a}_0(x)| < \varphi_2(x)$ all $x \in X$.

By Theorem 2.2 we can choose $a_2 \in A$ such that

$$|a_2| < \varphi_2, \quad a_2|_F = a_0 .$$

Assume for induction that extensions $a_1, \cdots, a_n \in A$ have been constructed such that

$$|a_p| < 2\psi \wedge \left[2^p\left(\psi - \sum_{r=1}^{p-1} 2^{-r}|a_r|\right)\right] = \varphi_p,\ p = 2, \cdots, n ,$$

and define

$$\varphi_{n+1} = 2\psi \wedge \left[2^{n+1}\left(\psi - \sum_{r-1}^{n} 2^{-r}|a_r|\right)\right] .$$

The function φ_{n+1} is strictly positive by induction hypothesis. For $x \in F$ we shall have

$$2^{n+1}\left(\psi(x) - \sum_{r=1}^{n} 2^{-r}|a_0(x)|\right) \geqq 2^{n+1}\left(\psi(x) - \sum_{r=1}^{n} 2^{-r}\psi(x)\right) = 2\psi(x)$$

such that $\varphi_{n+1}(x) = 2\psi(x)$. Hence for an arbitrary $x \in X$:

$$\begin{aligned}|\bar{a}_0(x)| &= \left|\int_F a_0 d\mu_x\right| \leqq \int_F \psi d\mu_x < \int_X 2^{n+1}\left(\psi - \sum_{r=1}^{n} 2^{-r}|a_r|\right)d\mu_x \\ &\leqq 2^{n+1}\left(\int_X \psi d\mu_x - \sum_{r=1}^{n} 2^{-r}\left|\int_X a_r d\mu_x\right|\right) \\ &\leqq 2^{n+1}\left(\psi(x) - \sum_{r=1}^{n} 2^{-r}|a_r(x)|\right) .\end{aligned}$$

Hence $|\bar{a}_0(x)| < \varphi_{n+1}(x)$ for all $x \in X$.

Again by Theorem 2.1 we can choose $a_{n+1} \in A$ such that

$$|a_{n+1}| < \varphi_{n+1}, \quad a_{n+1}|_F = a_0 .$$

Continuing in this way we obtain a sequence $\{a_n\}_{n=1}^{\infty} \subseteqq A$ such that for $n = 1, 2, \cdots$

(i) $a_n|_F = a_0$,

(ii) $\psi(x) - \sum_{r=1}^{n} 2^{-r}|a_r(x)| > 0$, all $x \in X$,

(iii) $||a_n|| \leqq 2 \sup_{x \in X} \psi(x)$.

By (iii) the sequence $\sum_{r=1}^{\infty} 2^{-r}a_r$ is uniformly convergent and $a = \sum_{r=1}^{\infty} 2^{-r}a_r \in A$. Clearly $a|_F = a_0$ and it follows from (ii) that $|a(x)| \leqq \psi(x)$ for all $x \in X$. This completes the proof.

Taking $\psi \equiv 1$ in Theorem 4.5 we obtain the following:

COROLLARY 4.6. *Let F be a compact subset of $\partial_A X$ satisfying* (A.1), *i.e.*

$$\mu \in M(\partial_A X) \cap A^{\perp} \Rightarrow \mu|_F \in A^{\perp} ,$$

then F has the norm preserving extension property.

REMARK. In the proof of Theorem 4.5 we have actually proved slightly more than was stated. The A-superharmonicity of the function ψ was used just once, namely in the verification that $|\bar{a}_0(x)| < \varphi_{n+1}(x)$ for $n = 1, 2, \cdots$ and all $x \in X$. However if x is a point of X such that

$$\mu \in M_x^{+}(\partial_A X) \Rightarrow \mu_x(F) = 0 ,$$

then by definition $\bar{a}_0(x) = 0$, and there is nothing to verify.

Hence, *Theorem* 4.5 *subsists if $\psi: X \rightarrow \boldsymbol{R}^{+} \cup \{\infty\}$ is allowed to be a l.s.c. function such that*

$$\psi(x) \geqq \int \psi d\mu_x ,$$

for all points $x \in X$ for which $\mu_x(F) \neq 0$ for some $\mu_x \in M_x^{+}(\partial_A X)$.

5. A peak set theorem.

In this section we shall deal with compact subsets F of $\partial_A X$ satisfying the requirement (A.2). For such an F we define the function $\bar{\chi}_F$ as in (3.7).

PROPOSITION 5.1. *If F is a compact subset of $\partial_A X$ satisfying* (A.2), *then the A-convex hull of F is equal to the set of all $x \in X$ such that $\bar{\chi}_F(x) = 1$.*

Proof. By definition, the A-convex hull of F is the set

$$F^{\wedge} = \{x \in X \mid |a(x)| \leqq ||a||_F, \quad \text{all} \quad a \in A\} . \tag{5.1}$$

We first assume that $\bar{\chi}_F(x) = 1$ i.e., $\mu_x(F) = 1$ for $\mu_x \in M_x^{+}(\partial_A X)$. Then we obtain for every $a \in A$,

$$|a(x)| = \left|\int_Y a d\mu_x\right| \leqq \int_F |a| d\mu_x \leqq \|a\|_F$$

such that $x \in F^\wedge$.

Next assume that $\bar{\chi}_F(x) < 1$. This implies that $\Phi(x) \notin S_F$. Hence we can separate $\Phi(x)$ and S_F by a w^*-continuous linear functional on A^* i.e., there exists a function $a \in A$ and an $\alpha \in \boldsymbol{R}$ such that

$$\operatorname{Re} \Psi(a)(\Phi(x)) > \alpha > \operatorname{Re} \Psi(a)(S_F) \geqq 0 ,$$

and hence again

$$\operatorname{Re} a(x) > \alpha > \operatorname{Re} a(F) \geqq 0 .$$

Now, for sufficiently large $\delta \in \boldsymbol{R}^+$, the function $a + \delta \in A$ satisfies

$$|a(x) + \delta| > \delta + \alpha > |a(y) + \delta| \quad \text{all} \quad y \in F .$$

In fact, it suffices to take

$$\delta > \frac{\gamma^2 + \beta^2 - \alpha\beta}{2(\alpha - \beta)} ,$$

where

$$\beta = \max\{\operatorname{Re} a(y) \mid y \in F\} < \alpha, \quad \gamma = \max\{|\operatorname{Im} a(y)| \mid y \in F\} .$$

Hence

$$\|a + \delta\|_F < |a(x) + \delta|$$

i.e., $x \in F^\wedge$, which completes the proof.

LEMMA 5.2. *Let F be a compact subset of $\partial_A X$ satisfying* (A.2), *for which $A|_F$ is closed in $\mathscr{C}_C(F)$. Let ψ be a strictly positive A-superharmonic function on X such that $1 \leqq \psi(x)$ for all $x \in F$.*

Then there exists a function $a \in A$ such that

$$a|_F = 1, \quad |a(x)| \leqq \psi(x) \quad \text{all} \quad x \in X \tag{5.2}$$

Proof. Since $\bar{\chi}_F$ is an element of $\mathscr{A}$ and $A|_F$ is assumed to be closed in $\mathscr{C}_C(F)$ we can use Theorem 2.2 with $a_0 \in A|_F$, $a_0 \equiv 1$. Now using the same technique as in the proof of Theorem 4.5 we obtain a function $a \in A$ satisfying (5.2).

LEMMA 5.3. *Let F be a compact subset of $\partial_A X$ satisfying* (A.2), *and let G be a compact subset of $X \backslash F^\wedge$. Then there exists an A-superharmonic function ψ on X such that:*

(i) $\psi(x) = 1$ *for all* $x \in F^\wedge$

(ii) $|\psi(x)| < 1$ *for all* $x \in G$

(iii) $0 < \psi(x) \leqq 1$ *for all* $x \in X$.

Proof. We write $S_G = \overline{co}(\Phi(G))$ and claim that $S_F \cap S_G = \emptyset$.

To prove this, we assume for contradiction that there exists a $p_0 \in S_F \cap S_G$, and we recall that $\hat{\chi}_{S_F}$ is u.s.c. and affine (since S_F is a parallel face) and that $\hat{\chi}_{S_F}$ is related to $\bar{\chi}_F$ by formula (3.8). Now we obtain

$$1 = \hat{\chi}_{S_F}(p_0) = \max_{p \in S_G} \hat{\chi}_{S_F}(p) = \max_{p \in \Phi(G)} \hat{\chi}_{S_F}(p) = \max_{p \in G} \bar{\chi}_F(p) .$$

By Proposition 5.1, this contradicts the hypothesis $G \cap F^\wedge = \emptyset$, and the claim is proved.

Now there exists a number δ such that

$$\max_{p \in S_G} \hat{\chi}_{S_F}(p) < \delta < 1 ,$$

and hence we can define two disjoint convex subsets of $A^* \times \boldsymbol{R}$ by the formulas:

$$F_0 = \{(p, \alpha) \mid p \in S, \alpha \in \boldsymbol{R}, 0 \leqq \alpha \leqq \hat{\chi}_{S_F}(p)\} \tag{5.3}$$

$$F_1 = \{(p, \alpha) \mid p \in S_G, \alpha \in \boldsymbol{R}, \delta \leqq \alpha\} . \tag{5.4}$$

The set F_0 is compact and the set F_1 is closed. Hence we can use Hahn-Banach separation to obtain a function $b \in A$ such that

$$\hat{\chi}_{S_F}(p) < \operatorname{Re} \psi(b)(p), \quad \text{all} \quad p \in G ,$$

and

$$\operatorname{Re} \psi(b)(p) < \delta < 1, \quad \text{all} \quad p \in S_G .$$

The function $\psi = \operatorname{Re}(b) \wedge 1$ is A-superharmonic and satisfies (i), (ii) and (iii).

THEOREM 5.4. *Let X be a metrizable compact Hausdorff space and let F be a compact subset of $\partial_A X$ which satisfies* (A.2) *i.e.*

$$\mu \in M(\partial_A X) \cap A^\perp \Rightarrow \mu(F) = 0 ,$$

and for which $A|_F$ is closed. Then there exists a function $a \in A$ such that

$$a|_{F^\wedge} = 1, \quad |a(x)| < 1 \quad \textit{all} \quad x \in X \backslash F^\wedge , \tag{5.5}$$

i.e. the A-convex hull of F is a peak set.

Proof. By metrizability $F^\wedge$ is a G_δ-set, and we can write $X \backslash F^\wedge = \bigcup_{u=1}^{\infty} K_n$, where K_n is closed.

Now we use Lemma 5.3 to obtain strictly positive A-superharmonic functions ψ_n on X such that

$$\psi_n(x) = 1 \quad \text{for all} \quad x \in F^\wedge, \quad \psi_n(x) < 1 \quad \text{for all} \quad x \in K_n, \quad n = 1, 2, \cdots$$

and $\psi_n(x) \leqq 1$ for all $x \in X$. It follows from Lemma 5.2 that there exist functions $a_n \in A$ such that $a_n|_F = 1$ and $|a_n(x)| \leqq \psi_n(x)$ for all $x \in X$. Now the function

$$a = \sum_{n=1}^{\infty} 2^{-n} a_n$$

satisfies (5.5) and the proof is complete.

REMARK. Actually the conclusion of Theorem 5.4 subsists under more general assumptions. The metrizability of X was only invoked to make $F^\wedge$ a G_δ-set. In particular we shall have the following:

COROLLARY 5.5. *Let $x \in \partial_A X$ be a G_δ-point satisfying* (A.2), *i.e.*

$$\mu \in M(\partial_A X) \cap A^\perp \Rightarrow \mu(\{x\}) = 0 \ ,$$

then x is a peak point for A.

Finally we remark that if X is a metrizable compact Hausdorff space and F is a compact subset of $\partial_A X$ satisfying the stronger condition (A.1) then *the A-convex hull of F is a peak set.*

6. Relations to the Bishop-Rudin-Carleson Theorem. In the present chapter we shall consider a compact subset F of X satisfying the requirement

(B) $$\mu \in A^\perp \Rightarrow \mu|_F = 0 \ .$$

Clearly (B) is more restrictive than (A.1), and *a fortiori* than (A.2). Note also that (B) implies $F \subset \partial_A X$ since $M_x^+(X) = \{\varepsilon_x\}$ for all $x \in F$.

If $x \notin F$ and $\mu_x \in M_x^+(X)$, then $\varepsilon_x - \mu_x \in A^\perp$. Now the requirement (B) implies $(\varepsilon_x - \mu_x)|_F = 0$, such that $\mu_x(F) = 0$. By the definition (3.2) we shall have $\bar{a}_0(x) = 0$. Hence

(6.1) $$\bar{a}_0 = a_0 \cdot \chi_F \ .$$

Transferring to the state space and making use of (3.8), we observe that the function $\hat{\chi}_{S_F}$ takes the value zero on $\Phi(X\backslash F)$. Geometrically, this means that *the canonical embedding Φ: $X \rightarrow S$ maps F into the (compact) split face $S_F = \overline{co}(\Phi(F))$, and $X\backslash F$ into the complementary (G_δ-) face S_F'* (*cf.* [2, Cor. 1.2]).

It follows from (6.1) that $\bar{\chi}_F = \chi_F$ and by Proposition 5.1 we obtain $F = \hat{F}$. Moreover, it follows from Proposition 4.4 that $A|_F$ is a closed subspace of $\mathscr{C}_C(F)$, and it follows from (B) that $(A|_F)^\perp = (0)$. Hence $A|_F = \mathscr{C}_C(F)$. Also it follows from the results of chapter 5 that if F is a G_δ, then it is a peak set.

In other words: *If F satisfies* (B) *then it is an interpolation set; and if in addition it is a G_δ, then it is a peak-interpolation set.*

Finally we note that we may apply Theorem 4.5 in the form stated in the Remark at the end of § 4, to obtain:

THEOREM 6.1. (*Bishop-Rudin-Carleson*) *Let F be a compact subset of X satisfying* (B), *i.e.*

$$\mu \in A^{\perp} \Rightarrow \mu|_F = 0 ;$$

let $f_0 \in \mathscr{C}_C(F)$, and let $\psi: X \to \boldsymbol{R}^+ \cup \{\infty\}$ be a strictly positive l.s.c. *function such that $|f_0(x)| \leqq \psi(x)$ for all $x \in F$. Then there exists an $a \in A$ such that $a|_F = f_0$ and $|a(x)| \leqq \psi(x)$ for all $x \in X$.*

REMARK. Theorem 6.1 is the most general form of the Bishop-Rudin-Carleson Theorem. Originally Bishop stated and proved this theorem for a continuous function ψ and strict inequality sign [4]. Appealing to the inductive construction of Pełczynski [12], Semadeni improved it to the form stated above [16]. (Cf. also Michael-Pełczynski [11, p. 569]).

7. The sup-norm algebra case. In this section we shall assume that A is a sup-norm algebra, and we shall consider two new requirements on a compact subset F of $\partial_A X$:

(G.1) $$\mu \in A^{\perp} \Rightarrow \mu|_F \in A^{\perp}$$

(G.2) $$\mu \in A^{\perp} \Rightarrow \mu|_{F^{\wedge}} \in A^{\perp} .$$

Clearly (B) implies (G.1) and (G.2), and each one of these implies (A.1). In fact, (G.2) implies (A.1) since $\mu|_{F^\wedge} = \mu|_F$ for every $\mu \in M(\partial_A X)$. This result in turn is elementary, but not entirely obvious, so we shall sketch a short proof: Note first that $F^\wedge = \Phi^{-1}(S_F)$, so that F can be thought of as the intersection of X with the ordinary closed convex hull of F in S. (This is standard for real function spaces, and the complex case is taken care of by the same argument as in the proof of Proposition 5.1.). Hence the problem is reduced to show the general implication:

$$\operatorname{Supp}(\nu) \subset \overline{co}(Q) \Rightarrow \operatorname{Supp}(\nu) \subset Q ,$$

where ν is a boundary measure and Q is a closed subset of S. By an elementary theorem ν is also a boundary measure on $\overline{co}(Q)$. (An explicite proof is given in [3, Lem. 1].) Hence ν is supported by the closure of the extreme points of $\overline{co}(Q)$. By Milman's Theorem $\operatorname{Supp}(\nu) \subset Q$, and the implication is proved.

In [9] and [10] Gamelin and Glicksberg have dealt with the requirement (G.1), and from their works we shall adopt the following:

DEFINITION 7.1. Let F be a compact subset of X and let $t > 0$. $A|_F$ is said to have the property E_t if the following conditions holds:

Given $f \in A|_F$ with $||f||_F < 1$ and a compact subset G of $X \backslash F$, there exists an extension $g \in A$ of f such that

$$||g||_X < \max\{1, t\}, \quad |g(x)| < t \quad \text{all} \quad x \in G .$$

The *extension constant* $e(A, F)$ *of* F *associated with* $A|_F$ is defined by the formula:

$$e(A, F) = \inf\{t \mid A|_F \ \text{ has property } \ E_t\} . \tag{7.1}$$

If $A|_F$ has property E_t for no t, then we define $e(A, F) = \infty$.

The connection between the extension constant and the requirement (G.1) is expressed in the following:

THEOREM 7.2. (*Gamelin-Glicksberg*). *Let F be a compact subset of X. Then the following conditions are equivalent*:

(i) $\mu \in A^\perp \Rightarrow \mu|_F \in A^\perp$

(ii) $e(A, F) = 0$

(iii) *F is an intersection of peak sets for A.*

Proof. See [9] and [10].

PROPOSITION 7.3. *Let A be a sup-norm algebra over X and let F be a compact subset of $\partial_A X$ satisfying the requirement* (A.1). *Also let G be a compact subset of $X \backslash F^\wedge$ and let $\varepsilon > 0$. Then there exists a function $a \in A$ such that*

(i) $a(x) = 1$ for all $x \in F^\wedge$

(ii) $|a(x)| < \varepsilon$ for all $x \in G$

(iii) $||a||_X = 1$.

Proof. Choose ψ as in Lemma 5.3 and let $a_0 \in A|_F$, $a_0 \equiv 1$. Using Theorem 4.5 we obtain a function $b \in A$ such that

$$b|_F = 1, \quad |b(x)| \leqq \psi(x) \quad \text{for all} \quad x \in X .$$

Cleary $b(x) = 1$ for all $x \in F^\wedge$ and $|b(x)| < 1$ for all $x \in G$. Now choose a natural number n such that $(||b||_G)^n < \varepsilon$ and define $a = b^n$. The proof is complete.

We are now able to clarify the connection between (A.1) and the extension constant of $F^\wedge$.

THEOREM 7.4. *Let A be a sup-norm algebra over X and let F be a compact subset of $\partial_A X$. Then $e(A, F^\wedge) = 0$ if and only if F satisfies* (A.1) *i.e.*

$$\mu \in M(\partial_A X) \cap A^\perp \Rightarrow \mu|_F \in A^\perp .$$

Proof. By virtue of Theorem 7.2 and the fact that $\mu|_{F^\wedge} = \mu|_F$ for every $\mu \in M(\partial_A X)$, if follows that $e(A, F^\wedge) = 0$ implies (A.1).

Now assume (A.1) and let $a_0 \in A|_{F^\wedge}$ with $\|a_0\|_{F^\wedge} = \|a_0\|_F < 1$. Let G be a compact subset of $X \backslash F^\wedge$ and let $\varepsilon > 0$. We choose $b \in A$ such that $\|b\|_X = \|a_0\|_F$ and $b|_F = a_0|_F$ according to Corollary (4.6), and we choose $h \in A$ according to Proposition (7.3) i.e.

$$h|_{F^\wedge} = 1, \quad |h(x)| < \varepsilon \quad \text{for all} \quad x \in G$$

and $\|h\|_X = 1$. Then we define $a = h \cdot b \in A$. Now, a is a norm preserving extension of a_0 and $|a(x)| < \varepsilon$ for all $x \in G$. Hence $A|_{F^\wedge}$ has property E_ε for all $\varepsilon > 0$, and so we have proved that $e(A, F^\wedge) = 0$.

Thus we see how the requirements (A.1), (G.1) and (G.2) are related for sup norm algebras. (A.1) and (G.2) are always equivalent for every compact subset F of $\partial_A X$, and if in addition F is A-convex, then they are equivalent to (G.1). This is not always the case even if A is an algebra and F satisfies (A.1), as can be seen from the following example

EXAMPLE 7.5. (*The "Tomato Can Algebra"*).

Let $X \subset \boldsymbol{R} \times \boldsymbol{C}$ be defined as $\{(t, z) \mid t \in [0, 1], |z| \leqq 1\}$; let A be the sup-norm algebra consisting all functions $f \in \mathscr{C}_C(X)$ such that $f(0, z)$ is analytic for $|z| < 1$; and let $F = \{(0, z) \mid |z| = 1\}$. Then F satisfies (A.1) *and $F^\wedge = \{(0, z) \mid |z| \leqq 1\}$.*

Proof. We first note that

$$\partial_A X = \{(t, z) \mid t \in \left]0, 1\right], |z| \leqq 1 \quad \text{or} \quad t = 0, |z| = 1\} .$$

Hence the Shilov boundary $\partial_S A = \overline{\partial_A X}$ is all of X, and it also follows that X is the maximal ideal space M_A of A.

If G is a compact subset of $X \backslash \{(0, z) \mid |z| \leqq 1\}$, then G is a peak interpolation set for A and $A|_G = \mathscr{C}_C(G)$. Hence if $\mu \in A^\perp$ then $\mu|_G = 0$. In other words $\operatorname{supp}(\mu) \subset \{0, z) \mid |z| \leqq 1\}$ for all $\mu \in A^\perp$.

Now assume $\mu \in M(\partial_A X) \cap A^\perp$. Then $\mu|_F = \mu \in A^\perp$. Hence F satisfies (A.1) but trivially $F^\wedge = \{(0, z) \mid |z| \leqq 1\}$; and the proof is complete.

This example shows also that *(A.1) and (G.1) need not be equivalent even if we consider A as a sup-norm algebra over the maximal ideal space or the Shilov boundary.*

Finally we remark that if X is a compact subset of $\boldsymbol{C}$ and

$A = R(X)|_{\partial X}$, i.e., if A is the uniform closure of the rational functions on X considered as a function algebra over the topological boundary ∂X, then the two conditions (A.1) and (G.1) are equivalent since $F = F^\wedge$ for every compact subset F of $\partial_A X$. In fact for a point $z_0 \in \partial X \backslash F$ we choose $f = (1/z - z_1) \in A$, where $z_1 \notin X$ and

$$|z_1 - z_0| = \tfrac{1}{2} \inf |z - z_0| ,$$

and obtain $|f(z_0)| = 2 \sup |f(z)|$.

References

1. E. M. Alfsen, *Compact convex sets and boundary integrals*, Ergebnisse der Matematik 57, Springer Verlag, Germany.

2. E. M. Alfsen and T. B. Andersen, *Split faces of compact convex sets*, to appear in Proc. London Math. Soc. **21** (1970), 415-442.

3. T. B. Andersen, *On dominated extensions of continuous affine functions on split faces*, (to appear).

4. Bishop, *A general Rudin-Carleson theorem*, Proc. Amer. Math. Soc., **13** (1962), 140-143.

5. J. E. Björk, Private communication.

6. ———, *Interpolation on closed linear subspaces of C(X)*, (to appear).

7. E. Brièm, *Restrictions of subspaces of C(X)*, Investigationes Math. **10** (1970), 288-297.

8. P. C. Curtis, *Topics in function spaces*, Lecture notes, Univ. of Århus 1969/70.

9. T. W. Gamelin, *Restrictions of subspaces of C(X)*, Trans. Amer. Math. Soc., **112** (1964), 278-286.

10. I. Glicksberg, *Measures orthogonal to algebras and sets of antisymmetry*, Trans. Amer. Math. Soc. **105** (1962), 415-435.

11. E. Michael and A. Pelczynski, *A linear extension theorem*, Illinois Journal of Mathematics, **11** (1967), 563-579.

12. A. Pelczynski, *Supplement to my paper "On simultaneous extensions of continuous functions*, Studia. Math. **25** (1964), 157-161.

13. R. R. Phelps, *Lectures on Choquet's theorem*, Van Nostrand, New York, 1966.

14. M. Rao, *Measurable selections of representing measures*, Århus Univ. Math. Inst. Preprint Ser. 1969/70 No. 24.

15. M. Rogalski, *Topologies faciales dans les convexes compacts; calcul functionel et décomposition spectrale dans le centre d'un espace A(X)*, Seminaire Choquet 1969-70 No. 4.

16. Z. Semadeni, *Simultaneous extensions and projections in spaces of continuous functions*, Lecture Notes, Univ. of Århus 1965.

Received June 18, 1970.

UNIVERSITY OF OSLO, NORWAY
AND
UNIVERSITY OF ÅRHUS, DENMARK

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

TOPOLOGIES FOR QUOTIENT FIELDS OF COMMUTATIVE INTEGRAL DOMAINS

JOBY MILO ANTHONY

In this paper topologies for the quotient field K of a commutative integral domain A are investigated. The topologies for K are defined so that convergence in K is stronger than convergence in A whenever A is a topological ring.

In particular, the Mikusinski field of operators is the quotient field of many commutative integral domains which are also topological rings. Each of these rings leads to a topological convergence notion in the Mikusinski field, which is stronger than the convergence notion introduced originally by Mikusinski. (The latter has recently been shown to be nontopological.)

In general, the algebraic and topological structures considered are not necessarily compatible; however, the question of compatibility is investigated. Necessary and sufficient conditions are given for the topology on A to be the restriction to A of the topology defined on K. In a theorem of S. Warner, necessary and sufficient conditions have been given for the neighborhood filter of zero in A to be a fundamental system of neighborhoods of zero for a topology on K. Moreover K, with this topology, is a topological field with A topologically embedded in K as an open set. For rings satisfying the conditions of this theorem, the topology for K which is defined in this paper is shown to reduce to that specified by Warner.

Let C_R^∞ denote the set of all infinitely differentiable, complex valued functions of a real variable with the support of each function contained in some right half-line. Endowed with the operations of addition and convolution, C_R^∞ becomes a commutative ring which has no divisors of zero. The quotient field of the ring C_R^∞ will be denoted by the symbol M. It is isomorphic to the field of Mikusiński operators [8]. If C_R^∞ is assigned the topology $\mathscr{T}^*$, in which a sequence $(\alpha_n \mid n \in Z^+)$ converges if the supports of the elements α_n are uniformly bounded on the left and the derivative sequences $(\alpha_n^{(k)} \mid n \in Z^+)$ converge uniformly on compact sets for all $k \in Z^+ \cup \{0\}$, then $(C_R^\infty, \mathscr{T}^*)$ is a topological ring.

Mikusiński has introduced a convergence concept for M which is equivalent to the following definition. If $(a_n \mid n \in Z^+)$ is a sequence in M, then $(a_n \mid n \in Z^+)$ converges if there exists a nonzero $p \in C_R^\infty$ such that $(pa_n \mid n \in Z^+)$ is a sequence in C_R^∞ which converges in the space $(C_R^\infty, \mathscr{T}^*)$ [6, pg. 144]. T. K. Boehme has shown that this convergence is nontopological in the sense that there is no topology for M in which sequential convergence is given by Mikusiński's definition [2]. E. F. Wagner has defined an analogous convergence concept for nets and filters and has shown that this leads to a limit space structure on M which is also nontopological [9].

It seems natural to ask how Mikusiński convergence can be modified so that it becomes topological. R. A. Struble has introduced such a modification [7], which has the property that the restriction of the resulting topology to the right-sided Schwartz distributions, which are embedded algebraically in M, is the topology which is ordinarily associated with them. The topology introduced by Struble is also defined by a convergence concept for sequences and appears to be unwieldy.

S. Warner has given necessary and sufficient conditions for a topological ring which has no zero divisors to be openly embeddable in a topological division ring [10, Theorem 5]. It is easy to see that the ring $(C_R^\infty, \mathscr{T}^*)$ does not satisfy these conditions. Consequently, there is no topology on M which makes M a topological field with C_R^∞ topologically embedded as an open set. Using some recent results of Boehme, we will prove an even stronger result concerning M; namely, there is no topology on M such that C_R^∞ is topologically embedded in M and multiplication in M is continuous. Essentially this means that M cannot be topologized in a "nice" way and efforts to "extend" the topology of C_R^∞ to M must be channelled in other directions.

In this paper we present a method for topologizing the quotient field of any commutative ring which has no zero divisors, using any topology which may be assigned to the ring. If the ring satisfies the conditions given by Warner in [10], then the topology which we will define has the property that the quotient field with this topology is a topological field with the ring topologically embedded as an open set. In general, however, the field topology will reflect only part of the algebraic and the topological structure of the ring and will not necessarily be compatible with the field structure. Although the ensuing development is applicable to very general algebraic and topological settings, it is strongly motivated by the unsatisfactory situation afforded by the Mikusiński operators. The field M will frequently be used as an example.

Throughout this paper, the symbol A will denote a commutative ring which has no zero divisors and K will denote the quotient field of A. We will use the symbol A^* to represent the set of nonzero elements of A. A topology on a set will be a collection of open sets and a neighborhood will be a set containing an open set. We will always assume that there is a topology associated with the set of elements of A and this topology will be denoted by $\mathscr{T}$. The topology $\mathscr{T}$ is not necessarily compatible with the algebraic structure of A.

Whenever we consider a ring of functions in which ring multiplication is the convolution of functions, multiplication will be represented by the symbol, $*$. For terminology concerning nets and filters, the reader should refer to [5] and [3].

The following development will be divided into two sections. In the first section we will deal with the definition and characterizations of a topology for the set of elements of the quotient field K. The second section will examine some specific properties of this topology relative to the algebraic and topological structures of A.

1. The definition and characterizations of the topology. Before defining a topology for the quotient field of an arbitrary commutative integral domain, let us examine the specific problem of extending the topology $\mathscr{T}^*$ of C_R^∞ to M.

LEMMA 1. *Let $\mathscr{T}'$ be any topology on M with the following properties.*

(i) $\mathscr{T}' | C_R^\infty > \mathscr{T}^*$. *(The restriction of $\mathscr{T}'$ to C_R^∞ is finer than $\mathscr{T}^*$.)*

(ii) *For each nonzero $p \in C_R^\infty$, the mapping $\xi_p: x \mapsto px$ of M into M is continuous.*

Then sequential convergence in $(M, \mathscr{T}')$ is stronger than Mikusiński convergence.

Proof. Let $(a_n | n \in Z^+)$ be a sequence in M and let $a \in M$ such that $(a_n | n \in Z^+) \xrightarrow{\mathscr{T}'} a$. (The net $(a_n | n \in Z^+)$ converges to a in the topology $\mathscr{T}'$.) A theorem of T. K. Boehme implies that any countable collection of elements in C_R^∞ has a common multiple in C_R^∞ [1]. This implies that there exists a nonzero element p in C_R^∞ such that $pa \in C_R^\infty$ and, for every $n \in Z^+$, $pa_n \in C_R^\infty$. Since multiplication by an element of C_R^∞ is continuous in $(M, \mathscr{T}')$, $(pa_n | n \in Z^+) \xrightarrow{\mathscr{T}'} pa$. But $\mathscr{T}' | C_R^\infty > \mathscr{T}^*$ and therefore $(pa_n | n \in Z^+) \xrightarrow{\mathscr{T}^*} pa$. This implies that $(a_n | n \in Z^+)$ Mikusiński-converges to a.

LEMMA 2. *Let $\mathscr{T}'$ be any topology on M with the following properties.*

(i) $\mathscr{T}^* > \mathscr{T}' | C_R^\infty$

(ii) *For each $a \in M$, the mapping $\xi_a: x \mapsto ax$ of M into M is continuous.*

Then Mikusiński convergence of sequences is stronger than sequential convergence in $(M, \mathscr{T}')$.

Proof. Let $(a_n | n \in Z^+)$ be a sequence in M and let $a \in M$ such that $(a_n \mid n \in Z^+)$ Mikusiński-converges to a. Then there exists a non-zero element p in C_R^∞ such that $(pa_n \mid n \in Z^+)$ is a sequence in C_R^∞, $pa \in C_R^\infty$ and $(pa_n \mid n \in Z^+) \xrightarrow{\mathscr{T}^*} pa$. Since $\mathscr{T}^* > \mathscr{T}' \mid C_R^\infty$, $(pa_n \mid n \in Z^+) \xrightarrow{\mathscr{T}'} pa$. But $\xi_{1/p}$ is continuous in $(M, \mathscr{T}')$ and therefore $(a_n \mid n \in Z^+) \xrightarrow{\mathscr{T}'} a$.

In [2] Boehme has shown that there is no topology on M which has the collection of Mikusiński-convergent sequences for its sequential convergence class. Combining this result with Lemma 1 and Lemma 2, we obtain the following theorem.

THEOREM 1. *There is no topology on M in which multiplication is continuous and for which $(C_R^\infty, \mathscr{T}^*)$ is topologically embedded in M.*

We will now examine the more general situation of an arbitrary commutative ring A with no zero divisors, and its associated quotient field K. For each $p \in A^*$, define a mapping ϕ_p from A into K by $\phi_p(\alpha) = \alpha/p$, $\alpha \in A$. Denote the image of A under the mapping ϕ_p by the symbol A_p. Let $\mathscr{T}_p$ be the finest topology on A_p which renders the mapping ϕ_p continuous. That is, $\mathscr{T}_p = \{0_p \subset A_p \mid 0_p = 0/p,\ 0 \in \mathscr{T}\}$. Since A has no zero divisors, $\alpha_1/p = \alpha_2/p$, $\alpha_1, \alpha_2 \in A$, if and only if $\alpha_1 = \alpha_2$. Consequently, ϕ_p is a bijection. Therefore $(A_p, \mathscr{T}_p)$ is homeomorphic to $(A, \mathscr{T})$. For each $\alpha \in A$, let $\mathscr{N}(\alpha)$ be the $\mathscr{T}$-neighborhood filter of α and if $a \in A_p$, let $\mathscr{N}_p(a)$ be the $\mathscr{T}_p$-neighborhood filter of a. We note that $K = \bigcup_{p \in A^*} A_p$. If $(a_\mu \mid \mu \in \mathrm{M})$ is a net in K, let $\mathrm{M}_p = \{\mu \in \mathrm{M} \mid a_\mu \in A_p\}$. Clearly if the net $(a_\mu \mid \mu \in \mathrm{M})$ is eventually in A_p, then $(a_\mu \mid \mu \in \mathrm{M}_p)$ is a subnet which is in A_p.

DEFINITION 1. Let $(a_\mu \mid \mu \in \mathrm{M})$ be a net in K and let $a \in K$. Then $(a_\mu \mid \mu \in \mathrm{M})$ is K-convergent to a, written $(a_\mu \mid \mu \in \mathrm{M}) \xrightarrow{K} a$, if the following condition is satisfied. Whenever $a \in A_p$, the net $(a_\mu \mid \mu \in \mathrm{M})$ is eventually in the space A_p and $(a_\mu \mid \mu \in \mathrm{M}_p) \xrightarrow{\mathscr{T}_p} a$.

The obvious generalization of Mikusiński convergence is the following. Let $(a_\mu \mid \mu \in \mathrm{M})$ be a net in K and let $a \in K$. Then $(a_\mu \mid \mu \in \mathrm{M})$ Mikusiński-converges to a if and only if for some $p \in A^*$, $a \in A_p$, $(a_\mu \mid \mu \in \mathrm{M})$ is eventually in A_p and $(a_\mu \mid \mu \in \mathrm{M}_p) \xrightarrow{\mathscr{T}_p} a$. Clearly K-convergence is stronger than Mikusiński convergence. We will now show that K-convergence is topological. This could be done directly by proving that the collection of K-convergent nets is the convergence class of a topology on K; however, it is slightly more interesting to give an analogous definition of K-convergence of filters, show that it is topological and then prove that it is equivalent to K-convergence of nets.

DEFINITION 2. If $\mathscr{F}$ is a filter on K and $a \in K$, then $\mathscr{F}$ is K-convergent to a, written $\mathscr{F}\tau_K a$, if and only if whenever $a \in A_p$, $\mathscr{F}$ is finer than the $\mathscr{T}_p$-neighborhood filter of a.

Clearly if $\mathscr{F}\tau_K a$ and $\mathscr{G} > \mathscr{F}$, then $\mathscr{G}\tau_K a$. Moreover, if $\mathscr{N}(a) = \bigcap_{\mathscr{F}\tau_K a} \mathscr{F}$, then $\mathscr{N}(a)\tau_K a$. Now for each $a \in K$, the collection of filters which K-converge to a is the collection of filters which are finer than $\mathscr{N}(a)$. Obviously $\mathscr{N}(a)$ is a candidate for the neighborhood filter of a in some topology. In [3, pg. 19, Proposition 2], sufficient conditions are given for a collection of filters on a set to uniquely determine a topology in which the specified filters are the neighborhood filters. The fact that these conditions are satisfied by the collection $\{\mathscr{N}(a) \mid a \in K\}$ constitutes the proof of Theorem 2; however, first we will prove the following lemma.

LEMMA 3. *For each* $a \in K$, $\mathscr{B}(a) = \{N_p(a) \mid a \in A_p$ *and* $N_p(a) \in \mathscr{N}_p(a)$ *for some* $p \in A^*\}$ *is a subbase for the filter* $\mathscr{N}(a)$.

Proof. Since every element of $\mathscr{B}(a)$ contains the point a, $\mathscr{B}(a)$ is a subbase for a filter on K. Let $\mathscr{B}'(a)$ be the collection of all finite intersections of elements of $\mathscr{B}(a)$ and let $\mathscr{B}''(a)$ be the filter generated by $\mathscr{B}'(a)$. Then $\mathscr{B}''(a)$ is the coarsest filter containing $\mathscr{B}(a)$. Now if $\mathscr{F}\tau_K a$, then $\mathscr{F}$ contains $\mathscr{B}(a)$ and so $\mathscr{B}''(a) < \mathscr{F}$. Therefore $\mathscr{B}''(a) < \mathscr{N}(a)$. On the other hand, if $a \in A_p$, then clearly $\mathscr{B}''(a) > \mathscr{N}_p(a)$ which implies that $\mathscr{B}''(a)\tau_K a$. Consequently $\mathscr{B}''(a) > \mathscr{N}(a)$.

THEOREM 2. *There is a unique topology on K with the property that a filter converges to a point if and only if it K-converges to that same point.*

Proof. Let a be a given element of K. Since $\mathscr{N}(a)$ is a filter, every subset of K which contains a set of $\mathscr{N}(a)$ is an element of $\mathscr{N}(a)$ and, moreover, $\mathscr{N}(a)$ has the finite intersection property. By Lemma 3 if $N(a) \in \mathscr{N}(a)$, then $a \in N(a)$ since every element of $\mathscr{B}(a)$ contains a. Also as a result of Lemma 3, there exists a finite intersection, $\bigcap_i N_{p_i}(a)$, of elements of $\mathscr{B}(a)$, which is contained in $N(a)$. Hence there exist open sets $O_{p_i} \in \mathscr{T}_{p_i}$ such that $O_{p_i}(a) \in \mathscr{N}_{p_i}(a)$ and $O_{p_i}(a) \subset N_{p_i}(a)$. Therefore $\bigcap_i O_{p_i}(a) \subset \bigcap_i N_{p_i}(a) \subset N(a)$. Moreover, $\bigcap_i O_{p_i}(a) \in \mathscr{N}(a)$. Let y be an arbitrary element of $\bigcap_i O_{p_i}(a)$. Since the sets $O_{p_i}(a)$ are open, they are elements of $\mathscr{B}(y)$. Consequently, $\bigcap_i O_{pi}(a) \in \mathscr{N}(y)$ and because $\bigcap_i O_{p_i}(a) \subset N(a)$, $N(a) \in \mathscr{N}(y)$.

It remains to be shown that K-convergence of nets and K-convergence of filters are equivalent. For a given net, its associated net filter is the collection of all sets which the net is "eventually in". In [5, pg. 83], it is shown that every filter is the net filter of some net. Therefore it is sufficient to prove the following theorem.

THEOREM 3. *Let $(a_\mu | \mu \in \mathrm{M})$ be a net in K and let $a \in K$. Then $(a_\mu | \mu \in \mathrm{M}) \xrightarrow{K} a$ if and only if its associated net filter, $\mathscr{F}(a_\mu | \mu \in \mathrm{M})$, K-converges to a.*

Proof. Suppose $(a_\mu | \mu \in \mathrm{M}) \xrightarrow{K} a$. Then if $a \in A_p$, $(a_\mu | \mu \in \mathrm{M})$ is eventually in every $\mathscr{T}_p$-neighborhood of a which implies that

$$\mathscr{F}(a_\mu | \mu \in \mathrm{M}) > \mathscr{N}_p(a) .$$

Therefore $\mathscr{F}(a_\mu | \mu \in \mathrm{M})\tau_K a$. Conversely, suppose $\mathscr{F}(a_\mu | \mu \in \mathrm{M})\tau_K a$. Then if $a \in A_p$, $\mathscr{F}(a_\mu | \mu \in \mathrm{M}) > \mathscr{N}_p(a)$ which implies that $(a_\mu | \mu \in \mathrm{M})$ is eventually in every $\mathscr{T}_p$-neighborhood of a. Therefore $(a_\mu | \mu \in \mathrm{M})$ is eventually in A_p, $(a_\mu | \mu \in \mathrm{M}_p) \xrightarrow{\mathscr{T}_p} a$, and hence $(a_\mu | \mu \in \mathrm{M}) \xrightarrow{K} a$.

Since K-convergence is topological, the topology determined by K-convergence will be denoted by $\mathscr{T}_K$. Moreover, since K-convergence of nets and K-convergence of filters are equivalent, we will use whichever notion of K-convergence is most appropriate to a particular situation.

The following theorem gives a simple characterization of the topology $\mathscr{T}_K$.

THEOREM 4. *$\mathscr{T}_K$ is the coarsest topology on K for which the collection $\mathscr{S} = \{O_p | O_p \in \mathscr{T}_p$ for some $p \in A^*\}$ is a collection of open sets.*

Proof. Let $O_p \in \mathscr{S}$ and suppose $(a_\mu | \mu \in \mathrm{M})$ is a net in K which K-converges to $a \in O_p$. Then $(a_\mu | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}_p} a$ which implies that $(a_\mu | \mu \in \mathrm{M})$ is eventually in O_p. Therefore $O_p \in \mathscr{T}_K$. Let $\mathscr{T}'$ be any topology on K with the property that $\mathscr{S} \subset \mathscr{T}'$. If $(a_\mu | \mu \in \mathrm{M})$ is a net in K which $\mathscr{T}'$-converges to a, then $(a_\mu | \mu \in \mathrm{M})$ is eventually in every $\mathscr{T}_p$-open-neighborhood of a. This implies that $(a_\mu | \mu \in \mathrm{M})$ is eventually in A_p, $(a_\mu | \mu \in \mathrm{M}_p) \xrightarrow{\mathscr{T}_p} a$, and hence $(a_\mu | \mu \in \mathrm{M}) \xrightarrow{K} a$. Therefore $\mathscr{T}_K$ is coarser than $\mathscr{T}'$.

Now we can make the following two observations. First, in view of Theorem 4, the topology $\mathscr{T}_K$ could have been defined as that topology on K which has the collection $\mathscr{S}$ as a subbase. From this point of view, Definition 1 and Definition 2 characterize convergence in this topology. Second, the algebraic characteristics of the ring A

and the field K are not essential in defining $\mathscr{T}_K$. In general, if $\{(\mathscr{B}_\ell, \mathscr{T}_\ell) \mid \ell \in \mathscr{L}\}$ is an arbitrary, indexed collection of topological spaces, then we may define a topology on $\bigcup_{\ell \in \mathscr{L}} B_\ell$ by taking the collection $\mathscr{S}_{\mathscr{L}} = \{O_\ell \mid O_\ell \in \mathscr{T}_\ell$ for some $\ell \in \mathscr{L}\}$ as a subbase. This topology is the coarsest one in which all the injection maps $i_\ell\colon B_\ell \to \bigcup_{\ell \in \mathscr{L}} B_\ell$ are open mappings. Convergence in this topology is characterized by definitions which are analogous to Definition 1 and Definition 2.

2. Properties of the topology. A few facts concerning the relationship between $(A, \mathscr{T})$ and $(K, \mathscr{T}_K)$ are immediate consequences of the characterizations of $\mathscr{T}_K$ which have already been given. For instance, it follows from Lemma 3 that if $(A, \mathscr{T})$ is a Hausdorff space, then $(K, \mathscr{T}_K)$ is a Hausdorff space, since distinct points of K are always elements of a common A_p-space and have disjoint neighborhoods in that space. There are several observations that can be made as a result of Theorem 4. Clearly if $(A, \mathscr{T})$ is a discrete space, then $(K, \mathscr{T}_K)$ is a discrete space. Also, it is obvious that for each $p \in A^*$, $\mathscr{T}_K \mid A_p$ is finer than $\mathscr{T}_p$. Another result of Theorem 4 is that if $\mathscr{T}^{(1)}$ and $\mathscr{T}^{(2)}$ are comparable topologies on A with $\mathscr{T}^{(1)}$ finer than $\mathscr{T}^{(2)}$, then the corresponding topologies of K-convergence, $\mathscr{T}_K^{(1)}$ and $\mathscr{T}_K^{(2)}$, have the same relationship. It is easy to construct examples to show that if $\mathscr{T}^{(1)}$ is strictly finer than $\mathscr{T}^{(2)}$, then $\mathscr{T}_K^{(1)}$ may be strictly finer than $\mathscr{T}_K^{(2)}$. Two major questions which remain to be answered are; "Under what conditions is $(A, \mathscr{T})$ topologically embedded in $(K, \mathscr{T}_K)$?", and "What is the relationship between the topology $\mathscr{T}_K$ and the algebraic structure of K?" It is the purpose of this section to examine these two questions.

For each $p \in A^*$, let ξ_p be the mapping of A into A defined by $\xi_p(x) = px$, $x \in A$.

LEMMA 4. *If for each $p \in A^*$ the mapping ξ_p is continuous, then $\mathscr{T}$ is finer than $\mathscr{T}_K \mid A$.*

Proof. Since the mappings ξ_p, $p \in A^*$, are continuous, if

$$(\alpha_\mu \mid \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha ,$$

then for every $p \in A^*$, $(\alpha_\mu p \mid \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha p$. From the way in which A is algebraically embedded in K, it follows that the elements of A are in every A_p-space. Specifically, if $p \in A^*$, then for each $\mu \in \mathrm{M}$, α_μ is identified with $\alpha_\mu p/p$ and α is identified with $\alpha p/p$. Therefore $(\alpha_\mu p \mid \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha p$ implies that $(\alpha_\mu \mid \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}_p} \alpha$. But this is true for every $p \in A^*$ and so we have $(\alpha_\mu \mid \mu \in \mathrm{M}) \xrightarrow{K} \alpha$.

In view of Lemma 4, we can make the following observation. If $(A, \mathscr{T})$ is a topological ring (recall that A is always a commutative ring which has no zero divisors), then K-convergence is a generalization of ring convergence in the sense that every $\mathscr{T}$-convergent net is K-convergent. There may, however, be nets in A which do not converge in $(A, \mathscr{T})$ but which are K-convergent. In fact, the following example shows that this is the case when $(A, \mathscr{T}) = (C_R^\infty, \mathscr{T}^*)$.

EXAMPLE 1. $\mathscr{T}$ may be strictly finer than $\mathscr{T}_K | A$.

If $(A, \mathscr{T}) = (C_R^\infty, \mathscr{T}^*)$ and $K = M$, we will denote the topology of K-convergence on M by $\mathscr{T}_K^*$. Then $\mathscr{T}^*$ is finer than $\mathscr{T}_K^* | C_R^\infty$ because $(C_R^\infty, \mathscr{T}^*)$ is a topological ring. Let $(\alpha_n | n \in Z^+)$ be a sequence in C_R^∞ with the following properties.

(i) For each $n \in Z^+$, the support of α_n is contained in $[0, 1/n]$.

(ii) For each $n \in Z^+$, $\max_t |\alpha_n(t)| = 1$.

Now if p is a nonzero element of C_R^∞, then $(\alpha_n * p | n \in Z^+) \xrightarrow{\mathscr{T}^*} 0$ which implies that $(\alpha_n | n \in Z^+) \xrightarrow{\mathscr{T}_p} 0$. Therefore $(\alpha_n | n \in Z^+) \xrightarrow{\mathscr{T}^*_K} 0$, but $(\alpha_n | n \in Z^+)$ does not converge in $\mathscr{T}^*$.

THEOREM 5. *$\mathscr{T} = \mathscr{T}_K | A$ if and only if $\mathscr{T}$ has the following property: If $(\alpha_\mu | \mu \in \mathrm{M})$ is a net in A and $\alpha \in A$, then $(\alpha_\mu p | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}}$ αp for every $p \in A^*$ if and only if $(\alpha_\mu | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha$.*

Proof. Suppose $\mathscr{T} = \mathscr{T}_K | A$. Let $(\alpha_\mu | \mu \in \mathrm{M})$ be a net in A with $\alpha \in A$ such that for each $p \in A^*$, $(\alpha_\mu p | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha p$. Now for each $p \in A^*$, $\alpha_\mu = (\alpha_\mu p / p)$, $\mu \in \mathrm{M}$, and $\alpha = (\alpha p / p)$. Therefore $(\alpha_\mu p | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha p$ implies that $(\alpha_\mu | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}_p} \alpha$. Consequently $(\alpha_\mu | \mu \in \mathrm{M}) \xrightarrow{K} \alpha$, and since $\mathscr{T} = \mathscr{T}_K | A$, $(\alpha_\mu | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha$. On the other hand, if $(\alpha_\mu | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha$ and $\mathscr{T} = \mathscr{T}_K | A$, then for every $p \in A^*$, $((\alpha_\mu p / p) | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}_p} (\alpha p / p)$ which implies that

$$(\alpha_\mu p | \mu \in \mu) \xrightarrow{\mathscr{T}} \alpha p \ .$$

Conversely, suppose $\mathscr{T}$ has the specified property. Then for each $p \in A^*$, the mapping ξ_p is continuous. By Lemma 4, $\mathscr{T} > \mathscr{T}_K | A$. Let $(\alpha_\mu | \mu \in \mathrm{M})$ be a net in A and let $\alpha \in A$ such that $(\alpha_\mu | \mu \in \mathrm{M}) \xrightarrow{K} \alpha$. Then for every $p \in A^*$, $(\alpha_\mu p | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha p$ and consequently

$$(\alpha_\mu | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha. \quad \text{Therefore } \mathscr{T}_k | A > \mathscr{T} \ .$$

COROLLARY a. *If A has an identity e and if for every $p \in A^*$ the mapping ξ_p is continuous, then $\mathscr{T} = \mathscr{T}_K | A$.*

Proof. Clearly the existence of an identity implies that the condition given in Theorem 5 is satisfied.

COROLLARY b. *Suppose that $(A, \mathscr{T})$ is a topological ring. If there exists $p' \in A^*$ such that $p'\mathscr{N}_{\mathscr{T}}(0) < \mathscr{N}_{\mathscr{T}}(0)$, then $\mathscr{T} = \mathscr{T}_K | A$.*

Proof. Since multiplication is continuous in A, we have $\mathscr{N}_{\mathscr{T}}(0) < p'\mathscr{N}_{\mathscr{T}}(0)$. Therefore the given condition is equivalent to the requirement that $p'\mathscr{N}_{\mathscr{T}}(0)$ be a base for $\mathscr{N}_{\mathscr{T}}(0)$. Because $(A, \mathscr{T})$ is a topological ring, if $(\alpha_\mu | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha$, then for every $p \in A^*$, $(\alpha_\mu p | (\mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha p$. Let $(\alpha_\mu | \mu \in \mathrm{M})$ be a net in A and let $\alpha \in A$ such that for every $p \in A^*$, $(\alpha_\mu p | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha p$. Since $p'\mathscr{N}_{\mathscr{T}}(0) < \mathscr{N}_{\mathscr{T}}(0)$, if $N_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$, it follows that $p'N_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$. Therefore $(\alpha_\mu p' - \alpha p' | \mu \in \mathrm{M})$ is eventually in $p'N_{\mathscr{T}}(0)$ which implies that $(\alpha_\mu - \alpha | \mu \in \mathrm{M})$ is eventually in $N_{\mathscr{T}}(0)$. Consequently $(\alpha_\mu | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha$.

COROLLARY c. *If $(A, \mathscr{T})$ is a compact, Hausdorff, topological ring, then $\mathscr{T} = \mathscr{T}_K | A$.*

Proof. Since $(A, \mathscr{T})$ is a topological ring, if $(\alpha_\mu | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha$, then for every $p \in A^*$, $(\alpha_\mu p | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha p$. Let $(\alpha_\mu | \mu \in \mathrm{M})$ be a net in A and let $\alpha \in A$ such that for every $p \in A^*$, $(\alpha_\mu p | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha p$. Let $(\beta_\lambda | \lambda \in \Lambda)$ be an arbitrary subnet of $(\alpha_\mu | \mu \in \mathrm{M})$. Since $(A, \mathscr{T})$ is compact, there exists a subnet $(\delta_\gamma | \gamma \in \Gamma)$ of $(\beta_\lambda | \lambda \in \Lambda)$ and $\delta \in A$ such that $(\delta_\gamma | \gamma \in \Gamma) \xrightarrow{\mathscr{T}} \delta$. If $p \in A^*$, then $(\delta_\gamma p | \gamma \in \Gamma) \xrightarrow{\mathscr{T}} \delta p$. But $(\delta_\gamma p | \gamma \in \Gamma)$ is a subnet of $(\alpha_\mu p | \mu \in \mathrm{M})$ which, by assumption, converges to αp. Therefore $(\delta_\gamma p | \gamma \in \Gamma) \xrightarrow{\mathscr{T}} \alpha p$ and since $(A, \mathscr{T})$ is Hausdorff, $\delta p = \alpha p$ which implies that $\delta = \alpha$. Now every subnet of $(\alpha_\mu | \mu \in \mathrm{M})$ has a subnet which converges to α. Consequently $(\alpha_\mu | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha$.

There are several important subsets of K which warrant special consideration, among which are A itself and the A_p-spaces. Another important subset of K is the intersection of all of the A_p-spaces. It is well known that the elements of K may be identified algebraically as either quotients (equivalence classes of ordered pairs of elements of A), or as partial homomorphisms of ideals of A into A whose domains are maximal in the sense that the partial homomorphisms cannot be extended to properly larger ideals [4]. In the latter situation, $\bigcap_{p \in A^*} A_p$

is identifiable as the set of those partial homomorphisms defined on all of A. This collection of mappings is denoted $\mathrm{Hom}_A(A, A)$. If $a \in A$, then a may be identified with the homomorphism $\xi_a: A \to A$ where $\xi_a(x) = ax$, $x \in A$. Hence $A \subset \mathrm{Hom}_A(A, A)$. Now $\mathrm{Hom}_A(A, A)$ may be considered as a collection of functions which map a common domain into a topological space. One way of topologizing such a function space is to use the so-called "weak" topology, the topology of pointwise convergence. Let $\mathscr{P}$ denote this topology. It is not difficult to see that $\mathscr{T}_K | \mathrm{Hom}_A(A, A) = \mathscr{P}$. An immediate corollary to Theorem 5 is that $(A, \mathscr{T})$ is topologically embedded in $(K, \mathscr{T}_K)$ if and only if it is topologically embedded in $(\mathrm{Hom}_A(A, A), \mathscr{P})$.

If $A = C_R^\infty$ and $K = M$, Struble has shown that $\mathrm{Hom}_A(A, A)$ is isomorphic to the collection of all right-sided Schwartz distributions [7]. The usual topology assigned to distribution is a "weak" topology. In this case it can be shown that these right-sided distributions are embedded both algebraically and topologically in the Mikusiński operator field.

In general, A and $\mathrm{Hom}_A(A, A)$ do not need to be either open or closed subsets of $(K, \mathscr{T}_K)$. In his paper on compact rings [10], Warner considers the problem of openly embedding a topological ring, which has no divisors of zero, in a division ring. The following theorem shows that if $(A, \mathscr{T})$ is a topological ring, then a weakened version of a condition used by Warner in [10, Theorem 5] is sufficient to guarantee that both $(A, \mathscr{T})$ and $(\mathrm{Hom}_A(A, A), \mathscr{P})$ are openly embedded in $(K, \mathscr{T}_K)$.

THEOREM 6. *Suppose that $(A, \mathscr{T})$ is a topological ring with the additional property that for each $N_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$, there exists $p \in A^*$ such that $pN_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$. Then*
(i) $\mathscr{T}_K | A = \mathscr{T}$
(ii) $A \in \mathscr{T}_K$
(iii) $\mathrm{Hom}_A(A, A) \in \mathscr{T}_K$.

Proof.
(i) Let $(\alpha_\mu | \mu \in \mathrm{M})$ be a net in A with $\alpha \in A$ such that

$$(\alpha_\mu p | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha p$$

for every $p \in A^*$. Then $(\alpha_\mu p - \alpha p | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} 0$ for every $p \in A^*$. Let $N_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$ and choose $p' \in A^*$ such that $p'N_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$. Then $(\alpha_\mu p' - \alpha p' | \mu \in \mathrm{M})$ is eventually in $p'N_{\mathscr{T}}(0)$ which implies that $(\alpha_\mu - \alpha | \mu \in \mathrm{M})$ is eventually in $N_{\mathscr{T}}(0)$. Therefore $(\alpha_\mu | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha$. Since $(A, \mathscr{T})$ is a topological ring, if $(\alpha_\mu | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha$, then for every $p \in A^*$, $(\alpha_\mu p | \mu \in \mathrm{M}) \xrightarrow{\mathscr{T}} \alpha p$. By Theorem 5, $\mathscr{T} = \mathscr{T}_K | A$.

(ii) Choose $p' \in A^*$ such that $p'A \in \mathscr{N}_{\mathscr{T}}(0)$. If $x \in p'A$, then there exists $y \in A$ such that $x = p'y$ and hence $x + p'A = p'y + p'A = p'A$. Therefore $p'A \in \mathscr{N}_{\mathscr{T}}(x)$. Now since $p'A$ is in the neighborhood filter of each of its points, it is an open subset of $(A, \mathscr{T})$. Consequently if $\alpha \in A$, then $p'A \in \mathscr{N}_{\mathscr{T}}(p'\alpha)$. But α has the representation $p'\alpha/p'$ in K, and since $p'A \in \mathscr{N}_{\mathscr{T}}(p'\alpha)$, by Lemma 3, $p'A/p' = A \in \mathscr{N}(\alpha)$. Thus A is in the $\mathscr{T}_K$-neighborhood filter of each of its points and hence is an open subset of $(K, \mathscr{T}_K)$.

(iii) We have shown in (ii) that it is possible to choose $p' \in A^*$ such that $p'A$ is an open subset of $(A, \mathscr{T})$. If $a \in \mathrm{Hom}_A(A, A)$, then there exists $\alpha \in A$ such that $a = \alpha/p'$. Now $\alpha + p'A \in \mathscr{N}_{\mathscr{T}}(\alpha)$ and by Lemma 3, $(\alpha + p'A)/p' \in \mathscr{N}(a)$. However, $(\alpha + p'A)/p' = \alpha/p' + p'A/p' = a + A$ which is a subset of $\mathrm{Hom}_A(A, A)$. Consequently $\mathrm{Hom}_A(A, A) \in \mathscr{N}(a)$. Now $\mathrm{Hom}_A(A, A)$ is in the $\mathscr{T}_K$-neighborhood filter of each of its points and hence is an open subset of $(K, \mathscr{T}_K)$.

The remainder of this paper will be devoted to an examination of the compatibility of the topology $\mathscr{T}_K$ with the algebraic structure of K.

DEFINITION 3. For each $a \in K$, let $D(a) = \{p \in A^* \mid a \in A_p\}$.

Note that $D(a) \cup \{0\}$ is an ideal in A. It is, in fact, the domain of a when a is identified as a partial homomorphism.

THEOREM 7. *Let a and b be elements of K such that $D(a + b) = D(a) \cap D(b)$. Then, if addition is continuous in $(A, \mathscr{T})$, the mapping $f: K \times K \to K$ defined by $f(x, y) = x + y$, $(x, y) \in K \times K$, is continuous at the point (a, b).*

Proof. Let $\mathscr{N}(a + b)$ be the $\mathscr{T}_K$-neighborhood filter of $a + b$ and let $N(a + b)$ be an arbitrary element of $\mathscr{N}(a + b)$. By Lemma 3, there exists a finite intersection, $\bigcap_i N_{p_i}(a + b)$, of $\mathscr{T}_{p_i}$-neighborhoods of $a + b$ contained in $N(a + b)$. Since $D(a + b) = D(a) \cap D(b)$, both a and b are elements of each A_{p_i}-space. Moreover, addition is continuous in each A_{p_i}-space and hence, for each i, there exists $N_{p_i}(a) \in \mathscr{N}_{p_i}(a)$ and $N_{p_i}(b) \in \mathscr{N}_{p_i}(b)$ such that $f(N_{p_i}(a) \times N_{p_i}(b)) \subset N_{p_i}(a + b)$. Therefore we have $f(\bigcap_i N_{p_i}(a) \times \bigcap_i N_{p_i}(b)) \subset \bigcap_i N_{p_i}(a + b) \subset N(a + b)$. If $\mathscr{N}(a, b)$ is the neighborhood filter of (a, b) in $K \times K$, then a base for $\mathscr{N}(a, b)$ is the collection of all sets of the form $N(a) \times N(b)$ where $N(a) \in \mathscr{N}(a)$ and $N(b) \in \mathscr{N}(b)$. Now by Lemma 3 we have $\bigcap_i N_{p_i}(a) \times \bigcap_i N_{p_i}(b) \in \mathscr{N}(a, b)$. Therefore f is continuous at (a, b).

COROLLARY. *If addition is continuous in $(A, \mathscr{T})$, then addition is continuous in* $(\mathrm{Hom}_A(A, A), \mathscr{P})$.

Proof. If a and b are elements of $\mathrm{Hom}_A(A, A)$, then $D(a) = D(b) = D(a + b) = A^*$.

THEOREM 8. *Let a and b be elements of K such that $D(ab) \subset D(a) \cdot D(b)$. ($D(a) \cdot D(b) = \{p \in A \mid p = qr,\ q \in D(a)$ and $r \in D(b)\}$). Then, if multiplication is continuous in $(A, \mathscr{T})$, the mapping $f\colon K \times K \to K$ defined by $f(x, y) = xy$, $(x, y) \in K \times K$, is continuous at the point (a, b).*

Proof. Let $\mathscr{N}(ab)$ be the $\mathscr{T}_K$-neighborhood filter of ab and let $N(ab)$ be an arbitrary element of $\mathscr{N}(ab)$. By Lemma 3, there exists a finite intersection, $\bigcap_i N_{p_i}(ab)$, of $\mathscr{T}_{p_i}$-neighborhoods of ab contained in $N(ab)$. Now, for each i, we have the following. Since $D(ab) \subset D(a) \cdot D(b)$, there exist $q_i \in A^*$ and $r_i \in A^*$ such that $a \in A_{q_i}$, $b \in A_{r_i}$, and $p_i = q_i r_i$. Therefore there exist ring elements α_i and β_i such that $a = \alpha_i/q_i$, $b = \beta_i/r_i$, and $ab = \alpha_i\beta_i/q_ir_i = \alpha_i\beta_i/p_i$. Moreover, there exists $N_{\mathscr{T}}(\alpha_i\beta_i) \in \mathscr{N}_{\mathscr{T}}(\alpha_i\beta_i)$ such that $N_{p_i}(ab) = N_{\mathscr{T}}(\alpha_i\beta_i)/p_i$. But multiplication is continuous in $(A, \mathscr{T})$, and therefore there exist $N_{\mathscr{T}}(\alpha_i) \in \mathscr{N}_{\mathscr{T}}(\alpha_i)$ and $N_{\mathscr{T}}(\beta_i) \in \mathscr{N}_{\mathscr{T}}(\beta_i)$ such that $N_{\mathscr{T}}(\alpha_i) \cdot N_{\mathscr{T}}(\beta_i) \subset N_{\mathscr{T}}(\alpha_i\beta_i)$. Let $N_{q_i}(a) = N_{\mathscr{T}}(\alpha_i)/q_i$ and $N_{r_i}(b) = N_{\mathscr{T}}(\beta_i)/r_i$. Then $N_{q_i}(a) \in \mathscr{N}_{q_i}(a)$ and $N_{r_i}(b) \in \mathscr{N}_{r_i}(b)$. Now we have

$$\begin{aligned} N_{q_i}(a) \cdot N_{r_i}(b) &= \frac{N_{\mathscr{T}}(\alpha_i)}{q_i} \cdot \frac{N_{\mathscr{T}}(\beta_i)}{r_i} = \frac{N_{\mathscr{T}}(\alpha_i) \cdot N_{\mathscr{T}}(\beta_i)}{q_ir_i} \\ &= \frac{N_{\mathscr{T}}(\alpha_i) \cdot N_{\mathscr{T}}(\beta_i)}{p_i} \subset \frac{N_{\mathscr{T}}(\alpha_i\beta_i)}{p_i} = N_{p_i}(ab)\ . \end{aligned}$$

That is, for each i, $f(N_{q_i}(a) \times N_{r_i}(b)) \subset N_{p_i}(ab)$. Therefore we have $f(\bigcap_i N_{q_i}(a) \times \bigcap_i N_{r_i}(b)) \subset \bigcap_i N_{p_i}(ab) \subset N(ab)$ and since $\bigcap_i N_{q_i}(a) \times \bigcap_i N_{r_i}(b) \in \mathscr{N}(a, b)$, f is continuous at the point (a, b).

COROLLARY. *If multiplication is continuous in $(A, \mathscr{T})$ and $A = A^2$, then multiplication is continuous in* $(\mathrm{Hom}_A(A, A), \mathscr{P})$.

Proof. If a and b are elements of $\mathrm{Hom}_A(A, A)$, then $D(a) = D(b) = D(ab) = A^*$. Since $A = A^2$ and A has no divisors of zero, $A^* = (A^*)^2$. Therefore $D(a) \cdot D(b) = (A^*)^2 = A^* = D(ab)$.

Theorems 7 and 8 give algebraic conditions which are sufficient for addition and multiplication to be locally continuous operations in $(K, \mathscr{T}_K)$. Since for every $a \in K$, $D(-a) = D(a)$, it is clear that if additive inversion is continuous in $(A, \mathscr{T})$, then it is also continuous in $(K, \mathscr{T}_K)$. Combining this fact with the corollaries to Theorems 7 and 8 yields the interesting result that if $(A, \mathscr{T})$ is a topological ring and $A = A^2$, then $(\mathrm{Hom}_A(A, A), \mathscr{P})$ is a topological ring.

The following examples demonstrate that multiplication and multiplicative inversion are not necessarily continuous operations in $(K, \mathscr{T}_K)$.

Example 2. Multiplication is not necessarily continuous in $(K, \mathscr{T}_K)$.

Let $(A, \mathscr{T}) = (C_R^\infty, \mathscr{T}^*)$ and $K = M$. Choose a nonzero element ϕ of C_R^∞ such that ϕ has compact support. For each $n \in Z^+$, let $m_n = \sup_t |\phi^{(n)}(t)|$. Consider the sequence $(f_n = s^n/nm_n | n \in Z^+)$, where s^n is the operator (homomorphism mapping C_R^∞ into itself) which maps a function in C_R^∞ to its nth derivative. If α is a nonzero element of C_R^∞, then for each $n \in Z^+, f_n$ has the representation $(\alpha^{(n)}/nm_n)/\alpha$. Choose a real number $\lambda > 1$ and let $\xi(t) = \phi(\lambda t)$. Then

$$\left(\frac{\xi^{(n)}/nm_n}{\xi}\,\middle|\, n \in Z^+\right) = \left(\frac{\lambda^n\phi^{(n)}(\lambda t)/nm_n}{\xi}\,\middle|\, n \in Z^+\right) \overset{\mathscr{T}_\xi}{\not\to} 0$$

$$\text{since}\left(\frac{\lambda^n\phi^{(n)}(\lambda t)}{nm_n}\,\middle|\, n \in Z^+\right) \overset{\mathscr{T}^*}{\not\to} 0 .$$

Hence $(f_n | n \in Z^+) \overset{K}{\not\to} 0$. If, however, ψ is any nonzero element of C_R^∞, then

$$\left(\frac{(\phi * \psi)^{(n)}}{nm_n}\,\middle|\, n \in Z^+\right) = \left(\frac{\phi^{(n)} * \psi}{nm_n}\,\middle|\, n \in Z^+\right) \xrightarrow{\mathscr{T}^*} 0 ,$$

and since $(C_R^\infty, \mathscr{T}^*)$ is a topological ring, by Lemma 4, it follows that

$$\left(\frac{(\phi * \psi)^{n}}{nm_n}\,\middle|\, n \in Z^+\right) \xrightarrow{K} 0 .$$

For each $n \in Z^+$, let $a_n = (\phi * \psi)^{(n)}/nm_n$ and let $b_n = (\phi * \psi)^{-1}$. Let $a = 0$ and $b = (\phi * \psi)^{-1}$. Now $(a_n | n \in Z^+) \xrightarrow{K} a$ and $(b_n | n \in Z^+) \xrightarrow{K} b$; however, $(a_n b_n | n \in Z^+) = (f_n | n \in Z^+) \overset{K}{\not\to} 0 = ab$. Therefore multiplication is not continuous on M.

Example 3. Multiplicative inversion is not necessarily continuous in $(K, \mathscr{T}_K)$.

Let $(A, \mathscr{T}) = (C_R^\infty, \mathscr{T}^*)$ and $K = M$. Consider the sequence $(1 - s/n | n \in Z^+)$. This is a sequence in M which clearly K-converges to the multiplicative identity; however, Mikusiński has shown that $((1 - s/n)^{-1} | n \in Z^+)$ does not converge according to his definition [6, pg. 147]. Therefore $((1 - s/n)^{-1} | n \in Z^+)$ does not K-converge. Consequently, multiplicative inversion is not continuous on M.

If addition is to be continuous in $(K, \mathscr{T}_K)$, then for each $a \in K$, the $\mathscr{T}_K$-neighborhood filter of a must be the translate to a of the

$\mathscr{T}_K$-neighborhood filter of zero. We will now discuss sufficient conditions on $(A, \mathscr{T})$ for $(K, \mathscr{T}_K)$ to have this property.

Suppose that $(A, \mathscr{T})$ is a topological ring. Then for each $p \in A^*$, the mapping $x \mapsto px$ is a continuous mapping of A into itself. Consequently, $p\mathscr{N}_{\mathscr{T}}(0)$ is a filter base for a filter which is finer than $\mathscr{N}_{\mathscr{T}}(0)$. In general, if p and q are distinct elements of A^*, then $p\mathscr{N}_{\mathscr{T}}(0)$ and $q\mathscr{N}_{\mathscr{T}}(0)$ are not equivalent filter bases; however, if for every pair (p, q) of elements of A^*, $p\mathscr{N}_{\mathscr{T}}(0)$ and $q\mathscr{N}_{\mathscr{T}}(0)$ are equivalent filter bases, then for each $p \in A^*$, $p\mathscr{N}_{\mathscr{T}}(0)$ and $p^2\mathscr{N}_{\mathscr{T}}(0)$ are equivalent filter bases. In this case, given $N_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$, there exists $N'_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$ such that $pN'_{\mathscr{T}}(0) \subset p^2N_{\mathscr{T}}(0)$ which implies that $N'_{\mathscr{T}}(0) \subset pN_{\mathscr{T}}(0)$. Therefore $pN_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$ and consequently, $p\mathscr{N}_{\mathscr{T}}(0)$ is a base for $\mathscr{N}_{\mathscr{T}}(0)$. Conversely, if for each $p \in A^*$, $p\mathscr{N}_{\mathscr{T}}(0)$ is a base for $\mathscr{N}_{\mathscr{T}}(0)$, then for every pair (p, q) of elements of A^*, $p\mathscr{N}_{\mathscr{T}}(0)$ and $q\mathscr{N}_{\mathscr{T}}(0)$ are equivalent filter bases.

LEMMA 5. *Let (R, T) be any topological ring. The following conditions on (R, T) are equivalent.*

(1) *Given an open neighborhood O of zero and a nonzero element p of R, then pO is an open set.*

(2) *Given a nonzero element p of R, then $p\mathscr{N}_T(0)$ is a base for $\mathscr{N}_T(0)$. ($\mathscr{N}_T(0)$ is the T-neighborhood filter of zero.)*

Proof.

(1) implies (2): Let p be a nonzero element of R and let $N_T(0) \in \mathscr{N}_T(0)$. Since (R, T) is a topological ring, the mapping $x \mapsto px$ is a continuous mapping of R into itself. Consequently, there exists an open neighborhood O of zero such that] $pO \subset N_T(0)$. By hypothesis, pO is an open neighborhood of zero. Therefore $p\mathscr{N}_T(0)$ is a base for $\mathscr{N}_T(0)$.

(2) implies (1): Let O be an open neighborhood of zero and let p be a nonzero element of R. Let $p\alpha$ be an arbitrary element of pO. Then O is a neighborhood of α. Consequently, there exists $O' \in \mathscr{N}_T(0)$ such that $O = \alpha + O'$. By hypothesis, $p\mathscr{N}_T(0)$ is a base for $\mathscr{N}_T(0)$. Therefore pO' is a neighborhood of zero. Now $pO = p\alpha + pO'$ and hence pO is an element of $\mathscr{N}_T(p\alpha)$. Therefore pO is in the neighborhood filter of each of its points which implies that pO is an open set.

THEOREM 9. *Suppose that $(A, \mathscr{T})$ is a topological ring. If for every $p \in A^*$, $p\mathscr{N}_{\mathscr{T}}(0)$ is a base for $\mathscr{N}_{\mathscr{T}}(0)$, then $\mathscr{K}(0) = \mathscr{N}_{\mathscr{T}}(0)$ is a base for $\mathscr{N}(0)$ and $\mathscr{N}(a) = a + \mathscr{N}(0)$ for every $a \in K$.*

Proof. By Lemma 3, for every $a \in K$, $\mathscr{B}(a) = \{N_p(a) \mid a \in A_p$ and

$N_p(a) \in \mathscr{N}_p(a)$ for some $p \in A^*\}$ is a subbase for $\mathscr{N}(a)$. If $N_p(0) \in \mathscr{B}(0)$, then there exists $N_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$ such that $N_p(0) = N_{\mathscr{T}}(0)/p$. Since $p\mathscr{N}_{\mathscr{T}}(0)$ is a base for $\mathscr{N}_{\mathscr{T}}(0)$, there exists $N'_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$ such that $pN'_{\mathscr{T}}(0) \subset N_{\mathscr{T}}(0)$. Therefore $N_p(0) = N_{\mathscr{T}}(0)/p \supset pN'_{\mathscr{T}}(0)/p = N'_{\mathscr{T}}(0)$. This implies that $\mathscr{B}(0) < \mathscr{K}(0)$. On the other hand, if $N_{\mathscr{T}}(0) \in \mathscr{K}(0)$ and $p \in A^*$, then $pN_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$. Now $N_{\mathscr{T}}(0) = pN_{\mathscr{T}}(0)/p$ which is an element of $\mathscr{B}(0)$. This implies that $\mathscr{K}(0) < \mathscr{B}(0)$. Therefore $\mathscr{K}(0)$ and $\mathscr{B}(0)$ are equivalent subbases. However, since $\mathscr{K}(0)$ is a filter on A, it is a filter base on K. Consequently, $\mathscr{K}(0)$ and $\mathscr{B}(0)$ are bases for the filter $\mathscr{N}(0)$. For each $a \in K$, let $\mathscr{K}(a) = a + \mathscr{K}(0)$. Clearly $\mathscr{K}(a)$ is a base for the filter $a + \mathscr{N}(0)$. If $a \in A_p$ and $N_p(a) \in \mathscr{N}_p(a)$, then there exists $\alpha \in A$ and $N_{\mathscr{T}}(\alpha) \in \mathscr{N}_{\mathscr{T}}(\alpha)$ such that $a = \alpha/p$ and $N_p(a) = N_{\mathscr{T}}(\alpha)/p$. Since $(A, \mathscr{T})$ is a topological ring, there exists $N_{\mathscr{T}}(0)$ such that $N_{\mathscr{T}}(\alpha) = \alpha + N_{\mathscr{T}}(0)$. Moreover, $p\mathscr{N}_{\mathscr{T}}(0)$ is a base for $\mathscr{N}_{\mathscr{T}}(0)$. Therefore there exists $N'_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$ such that $pN'_{\mathscr{T}}(0) \subset N_{\mathscr{T}}(0)$. Now we have

$$N_p(a) = \frac{N_{\mathscr{T}}(\alpha)}{p} = \frac{\alpha + N_{\mathscr{T}}(0)}{p} \subset \frac{\alpha + pN'_{\mathscr{T}}(0)}{p} = a + N'_{\mathscr{T}}(0) .$$

This implies that $\mathscr{B}(a) < \mathscr{K}(a)$. Conversely, if $a + N_{\mathscr{T}}(0) \in \mathscr{K}(a)$, choose $p \in A^*$ such that $a \in A_p$. Now let $\alpha \in A$ such that $a = \alpha/p$. Since $p\mathscr{N}_{\mathscr{T}}(0)$ is a base for $\mathscr{N}_{\mathscr{T}}(0)$, we have $pN_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(0)$. This implies that $\alpha + pN_{\mathscr{T}}(0) \in \mathscr{N}_{\mathscr{T}}(\alpha)$. Consider

$$a + N_{\mathscr{T}}(0) = \frac{\alpha}{p} + \frac{pN_{\mathscr{T}}(0)}{p} = \frac{\alpha + pN_{\mathscr{T}}(0)}{p} .$$

This is an element of $\mathscr{B}(a)$ and consequently $\mathscr{K}(a) < \mathscr{B}(a)$. Therefore $\mathscr{B}(a)$ is a filter base which is equivalent to $\mathscr{K}(a)$. Since $\mathscr{B}(a)$ is a base for $\mathscr{N}(a)$ and $\mathscr{K}(a)$ is a base for $a + \mathscr{N}(0)$, we have $\mathscr{N}(a) = a + \mathscr{N}(0)$.

What we have now demonstrated is that if $(A, \mathscr{T})$ is a topological ring which satisfies either of the conditions of Lemma 5, then $(K, \mathscr{T}_K)$ is homogeneous in the sense that the $\mathscr{T}_K$-neighborhood filter of any point is the translate to that point of the $\mathscr{T}_K$-neighborhood filter of zero. Moreover, the neighborhood filter of zero in $(A, \mathscr{T})$ is a base for the neighborhood filter of zero in $(K, \mathscr{T}_K)$. Also, since $(A, \mathscr{T})$ satisfies one of the conditions of Lemma 5, by Theorem 6 it follows that A is topologically embedded in $(K, \mathscr{T}_K)$ as an open set.

In [10, Theorem 5], Warner places the following conditions on a topological ring which has no divisors of zero.

(1) Given an open neighborhood O of zero and a nonzero ring element p, then pO and Op are open sets.

(2) The collection of ring elements which have an inverse relative

to the circle composition ($x \circ y = x + y - xy$) is an open set, and the mapping which sends an element of this open set to this inverse is continuous.

He concludes that these conditions are both necessary and sufficient for the ring to be algebraically embeddable in a division ring, where the neighborhood filter of zero in the original ring is a fundamental system of neighborhoods of zero for a topology on the division ring. Moreover, the specified topology on the division ring is compatible with the division ring structure and the original ring is topologically embedded as an open set. Therefore, by Lemma 5 and Theorem 9, we conclude that these conditions on $(A, \mathscr{T})$ are necessary and sufficient for $(K, \mathscr{T}_K)$ to be a topological field with A topologically embedded as an open set. In the process of proving this theorem of Warner's, condition (2) is used only to establish the continuity of multiplicative inversion in the division ring. Hence we conclude that $(K, \mathscr{T}_K)$ is a topological ring with A topologically embedded as an open set if and only if $(A, \mathscr{T})$ satisfies one of the conditions of Lemma 5.

Several questions concerning the topology $\mathscr{T}_K$ are suggested by this paper. For instance, what hypotheses are required for $(K, \mathscr{T}_K)$ to be a topological field without A necessarily being an open set? By Theorem 5, Corollary c, if $(A, \mathscr{T})$ is compact and Hausdorff, then it is topologically embedded in $(K, \mathscr{T}_K)$. What further hypotheses, if any, are needed to insure that $(K, \mathscr{T}_K)$ is at least a topological ring? There is also, of course, the observation that the concept of K-convergence provides a method for topologizing the Mikusiński field. In fact, the various algebraic models which generate the Mikusiński field lead to several topologies of K-convergence on it. What properties do they possess and how are they related?

References

1. T. K. Boehme, *On sequences of continuous functions and convolution*, Studia Math., **25** (1965), 333-335.
2. ———, *On Mikusiński operators*, To appear in Studia Math.
3. N. Bourbaki, *General Topology, Part* 1, Addison-Wesley Pub. Co., (1966), Reading, Mass.
4. G. D. Findlay, and J. Lambek, *A generalized ring of quotients I*, Canad. Math. Bull., **1** (1958), 77-85.
5. J. L. Kelley, *General Topology*, D. Van Nostrand Co., Inc., (1955), Princeton New Jersey.
6. J. Mikusiński, *Operational Calculus*, Pergamon Press, (1959), London and New York.
7. R. A. Struble, *An algebraic view of distributions and operators*, Unpublished manuscript.
8. ———, *On operators and distributions*, Canad. Math. Bull., **11** (1968), 61-63.

9. E. F. Wagner, *On the convergence structure of Mikusiński operators*, Studia. Math., **27** (1966), 39-48.
10. S. Warner, *Compact rings*, Math. Annalen, **145** (1962), 52-63.

Received February 15, 1970. This work was supported by the U. S. Army Research Office (Durham) and constitutes a portion of the author's Ph.D. thesis, written under the direction of Professor R. A. Struble at North Carolina State University, Raleigh.

NORTH CAROLINA STATE UNIVERSITY AT RALEIGH

FLORIDA TECHNOLOGICAL UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol, 36, No. 3, 1971

ORDERED CYCLE LENGTHS IN A RANDOM PERMUTATION

V. BALAKRISHNAN, G. SANKARANARAYANAN AND C. SUYAMBULINGOM

Let $x(t)$ denote the number of jumps occurring in the time interval $[0, t)$ and $v_k(t) = P\{x(t) = k\}$. The generating function of $v_k(t)$ is given by

$$\exp\{\lambda t[\phi(x) - 1]\},\ \phi(x) = \sum_{k=1}^{\infty} p_k x^k,\ \sum_{k=1}^{\infty} p_k = 1\ .$$

Lay off to the right of the origin successive intervals of length z^j/j^α, $j = 1, 2, \cdots$. Explicitly the end points are

$$t_1(z) = 0$$

$$t_j(z) = \sum_{k=1}^{j-1} z^k/k^\alpha,\ j = 2, 3, \cdots, \alpha > 0\ ,$$

and

$$t_\infty(z) = \sum_{k=1}^{\infty} z^k/k^\alpha\ .$$

Following Shepp and Lloyd L_r, the length of the rth longest cycle and S_r, the length of the rth shortest cycle have been defined for our choice of $x(t)$ and t_j, $j = 1, 2, \cdots$. This paper obtains the asymptotics for the mth moments of L_r and S_r suitably normalized by a new technique of generating functions. It is further shown that the results of Shepp and Lloyd are particular cases of these more general results.

Here we consider a problem involving a random permutation which is closely linked with the cycle structure of the permutation. Let S_n be the $n!$ permutation operators on n numbered places. Let $\alpha(\pi) = \{\alpha_1(\pi), \alpha_2(\pi), \cdots, \alpha_n(\pi)\}$ be the cycle class of $\pi \in S_n$. In this permutation π, there are $\alpha_1(\pi)$ cycles of length one, $\alpha_2(\pi)$ cycles of length two, etc. Usually the elements of S_n are assigned a probability $1/n!$ each. John Riordan has considered a model where he has assigned the probability

$$1.1 \qquad P\{\alpha_1 = a_1, \alpha_2 = a_2, \cdots, \alpha_n = a_n\} = \prod_{j=1}^{n} (1/j)^{a_j}/a_j! \text{ if } \sum_{j=1}^{n} ja_j = n,$$
$$= 0 \text{ otherwise}\ ,$$

for the cycle class $\alpha(\pi)$, the a's being nonnegative integers. Here α's would be independent if it were not for the condition $\sum ja_j = n$. Shepp and Lloyd has considered a sequence $\alpha = \{\alpha_1, \alpha_2, \cdots\}$ of mutually independent nonnegative integral valued random variables where for $j = 1, 2, \cdots$ the random variable α_j follows the Poisson distribution with mean z^j/j, $0 < z < 1$, z being same for all values of j. Accordingly

1.2 $$P_z\{\alpha_1 = a_1, \alpha_2 = a_2, \cdots\} = (1 - z)z^{\sum_{j=1}^{\infty} j a_j} \prod_{j=1}^{\infty} (1/j)^{a_j}/a_j!, \quad a_j > 0, j = 1, 2, \cdots .$$

From this it can be seen that the probability distribution of the random variable $\nu(\alpha) = \sum_{j=1}^{\infty} j\alpha_j$ is

1.3 $$P\{\nu(\alpha) = n\} = (1 - z)z^n, n = 0, 1, 2, \cdots .$$

Also

1.4 $$P_z\{\alpha_1 = a_1, \alpha_2 = a_2, \cdots | \nu(\alpha) = n\} = \prod_{j=1}^{n} (1/j)^{a_j}/a_j!, \sum_{j=1}^{\infty} j a_j = n$$
$$= 0 \text{ otherwise} .$$

Thus Shepp and Lloyd were able to recover 1.1 assumed in the model. In this paper, for the cycle class $\alpha(\pi)$ we have assigned the probability

1.5 $$P_z(\alpha_1 = a_1, \alpha_2 = a_2, \cdots, \alpha_n = a_n) = \mathrm{I}/\mathrm{II}, 0 < z < 1, \sum_{j=1}^{n} j a_j = n$$
$$= 0 \text{ otherwise} .$$

Here

1.6 $$I = \prod_{j=1}^{\infty} v_{a_j}(z^j/j^\alpha), \sum_{j=1}^{n} j a_j = n, a_{n+1} = a_{n+2} = \cdots = 0, \left(\sum_{j=1}^{\infty} j a_j = n\right)$$

where $v_{a_j}(z^j/j^\alpha)$ is the coefficient of x^{a_j} in $\exp\{\lambda(z^j/j^\alpha)[\phi(x) - 1]\}$,

1.7 $$\phi(x) = \sum_{k=1}^{\infty} p_k x^k \quad \text{and} \quad \sum_{k=1}^{\infty} p_k = 1 .$$

On detailed computation

1.8 $$v_{a_j}(z^j/j^\alpha) = e^{-\lambda z^j/j^\alpha} \sum_{n_1+2n_2+3n_3+\cdots=a_j} \frac{(p_1 z^j/j^\alpha)^{n_1}(p_2 z^j/j^\alpha)^{n_2}\cdots}{n_1!\, n_2!\cdots} .$$

In the special case when $\lambda = 1$, $p_1 = 1$, $p_2 = p_3 = \cdots = 0$ and $\alpha = 1$, $\exp\{\lambda(z^j/j^\alpha)[\phi(x) - 1]\}$ reduces to the generating function of the Poisson process with the time parameter equals to z^j/j, which has been considered by Shepp and Lloyd. Also the generating function of II which represents the distribution of $P\{\nu(\alpha) = n\}$, where for our choice of the sequence α_j's defined by 1.14

1.9 $$\nu(\alpha) = \sum_{j=1}^{\infty} j\alpha_j$$

is given by

1.10 $$\sum_{n=1}^{\infty} P\{\nu(\alpha) = n\}x^n = \prod_{j=1}^{\infty} \exp\{\lambda(z^j/j^\alpha)[\phi(x^j) - 1]\} .$$

On detailed computation we note that

1.11 $$P\{\nu(\alpha) = n\} = \exp\{-\lambda \sum_{j=1}^{\infty} z^j/j^\alpha\} \times \sum_{\substack{n_1+2n_2+3n_3+\cdots \\ +2(n_1'+2n_2'+\cdots) \\ +3(n_1''+2n_2''+\cdots) \\ +\cdots=n}} \left\{ \left[\frac{\left(\frac{p_1\lambda z}{1^\alpha}\right)^{n_1}\left(\frac{p_1\lambda z^2}{2^\alpha}\right)^{n_2}\cdots}{n_1!\, n_2!\,\cdots}\right] \times \left[\frac{\left(\frac{p_2\lambda z}{1^\alpha}\right)^{n_1'}\left(\frac{p_2\lambda z^2}{2^\alpha}\right)^{n_2'}\cdots}{n_1'!\, n_2'!\,\cdots}\right] \times \left[\frac{\left(\frac{p_3\lambda z}{1^\alpha}\right)^{n_1''}\left(\frac{p_3\lambda z^2}{2^\alpha}\right)^{n_2''}\cdots}{n_1''!\, n_2''!\,\cdots}\right] \times \cdots \right\} .$$

In particular when $\lambda = 1, \alpha = 1$ and $p_1 = 1, p_2 = p_3 = \cdots = 0$, the generating function of the distribution of 1.9 reduces to

1.12 $$\exp[-\sum (z^j/j) + \sum (x^j z^j/j)] = (1 - z)/(1 - zx) .$$

Hence

$$P\{\nu(\alpha) = n\} = (1 - z)z^n ,$$

which is in agreement with that considered by Shepp and Lloyd. In the special case mentioned above

1.13 $$\begin{aligned} \mathrm{I}/\mathrm{II} &= \prod_{j=1}^{n} (1/j)^{a_j}/a_j! \qquad \text{if} \qquad \sum_{j=1}^{n} j a_j = n, \\ &= 0 \qquad \text{otherwise} . \end{aligned}$$

This is also in agreement with the model discussed by Shepp and Lloyd.

If we take $\alpha = (\alpha_1, \alpha_2, \cdots)$ to be a sequence of mutually independent nonnegative integral valued random variables where for $j = 1, 2, \cdots$

1.14 $$P_z\{\alpha_j = a_j\} = v_{a_j}(z^j/j^\alpha),\ a_j = 0, 1, 2, \cdots ,$$

by using the Borel-Cantelli lemma, we can easily show that $\nu(\alpha) = \sum_{j=1}^{\infty} j\alpha_j$ is finite with probability one. Hence the joint distribution $(\alpha_1, \alpha_2, \alpha_3, \cdots, \nu(\alpha))$ can be written as

1.15 $$\begin{aligned} P_z\{\alpha_1 = a_1, \alpha_2 = a_2, \cdots, \nu(\alpha) = n\} &= \prod_{j=1}^{\infty} v_{a_j}(z^j/j^\alpha) \text{ if } \sum_{j=1}^{\infty} j a_j = n, \\ &= 0 \qquad \text{otherwise} . \end{aligned}$$

From this we can see that

1.16 $$P_z\{\alpha_1 = a_1, \alpha_2 = a_2, \cdots, \mid \nu(\alpha) = n\} = \mathrm{I}/\mathrm{II}\ ,$$

which we have assumed for the model.

Shepp and Lloyd have considered a Poisson process which takes place on $T = \{-\infty < t < +\infty\}$ at unit rate. That is, for any interval of length $I \subset T$, the probability that p jumps occur in I is

$$\exp[-|I|]\,|I|^p/p!,\ p = 0, 1, 2, \cdots$$

independently of any conditions on the process on $T - I$. They have taken the following end points for the time intervals

1.17 $$\begin{aligned} t_1(z) &= 0, \\ t_j(z) &= \sum_{k=1}^{j-1} z^k/k,\ j = 2, 3, \cdots\ , \\ t_\infty(z) &= \sum_{k=1}^{\infty} z^k/k \quad = \log(1-z)^{-1}\ , \end{aligned}$$

so that the jth interval is

$$t_j(z) < t < t_{j+1}(z),\ j = 1, 2, \cdots .$$

They define $\lambda_z(t)$; $-\infty < t < \infty$, to be a function whose value is 'j' on the jth interval, $j = 1, 2, \cdots$ and is zero if $t < 0$ or $t > t_\infty(z)$. Then for each $j = 1, 2, \cdots$ the interval $\{t; \lambda_z(t) = j\}$ has length z^j/j, the probability that a_j jumps of the Poisson process occur in this interval is

1.18 $$\exp(-z^j/j)\cdot(z^j/j)^{a_j}/a_j!,\ a_j = 0, 1, 2, \cdots$$

and that these various events for $j = 1, 2, \cdots$ are mutually independent. They have taken a sample function of the Poisson process, with jumps in the interval $[0, t_\infty(z))$, which are finite in number with probability one, occurring at times $\tau_1 \leqq \tau_2 \leqq \cdots \leqq \tau_\sigma$ (σ, random). They take the positive integers $\lambda_z(\tau_1) \leqq \lambda_z(\tau_2) \leqq \cdots \leqq \lambda_z(\tau_\sigma)$ as the lengths of the σ cycles of a permutation on $\nu = \sum_{s=1}^{\sigma} \lambda_z(\tau_s)$ places, and in this class S_ν, they choose a permutation at random with uniform distribution. For any given $r = 1, 2, \cdots$ let $S_r = S_r(\alpha)$ be the length of the rth shortest cycle in a permutation of the cycle class $\alpha \cdot S_r(\alpha) = 0$ if $\sum \alpha_j < r$. If the rth jump of the Poisson process occur at 't', then $S_r = \lambda_z(t)$ according to their model. Hence they have obtained the distribution of S_r. Similarly they have obtained the distribution of $L_r = L_r(\alpha)$, the length of the rth longest cycle. They have given asymptotics for the distribution and to all moments of the length of the rth longest and rth shortest cycles.

In this paper, instead of the Poisson process considered by Shepp and Lloyd, we consider a more general process which can have $k(k > 1)$

jumps at any moment. Let $x(t)$ denote the number of jumps in the interval $[0, t)$ and let

1.19 $$v_k(t) = P\{x(t) = k\} .$$

Let p_k be the probability of having k jumps at a chosen moment, if it is certain that jumps do occur generally at that moment. It has been shown in Khintchine that

1.20 $$F(t, x) = \sum_{k=0}^{\infty} v_k(t)x^k = \exp\{\lambda t[\phi(x) - 1]\} ,$$

where $\phi(x)$ is given by (1.7) and $\lambda > 0$. In our model, we take the end points of the time intervals to be

1.21 $$\begin{aligned} t_1(z) &= 0 \\ t_j(z) &= \sum_{k=1}^{j-1} z^k/k^\alpha, j = 2, 3, \cdots, \alpha > 0 , \end{aligned}$$

and

$$t_\infty(z) = \sum_{k=1}^{\infty} z^k/k^\alpha .$$

Here the probability that L_r, the length of the rth longest cycle is 'j' is given by

1.22 $$\begin{aligned} P_z\{L_r = j\} &= \frac{\lambda}{\sum_{k=1}^{r} p_k} \int_{t_j}^{t_{j+1}} \left\{\sum_{k=1}^{r} p_k v_{r-k}(t_\infty - t)\right\} dt, \\ &= \frac{\lambda}{P_r} \int_{t_j}^{t_{j+1}} \left\{\sum_{k=1}^{r} p_k v_{r-k}(t_\infty - t)\right\} dt , \end{aligned}$$

where

$$P_r = \sum_{k=1}^{r} p_k .$$

Also the probability that S_r, the length of the rth shortest cycle is 'j' is given by

$$P_z\{S_r = j\} = \frac{\lambda}{P_r} \int_{t_j}^{t_{j+1}} \left\{\sum_{k=1}^{r} p_k v_{r-k}(t)\right\} dt .$$

Here we use the technique of generating functions to estimate the asymptotics of $E\{L_r\}^m$ and $E\{S_r\}^m$ suitably normalized in a way different from that used by Shepp and Lloyd. While they have considered the case where the jumps occur according to Poisson law, we have considered a more general system of which Poisson process is a special case. By assuming the Poisson law for jumps they were able to recover the model based on the uniform distribution. By assuming a more general law for jumps we obtain a generalised probability model for the cycle class of which that derived on the basis of the uniform distribution is a special case. Thus we have in this paper discussed a generalization of the one given by Shepp and Lloyd with the help of the new technique.

2. A lemma. We now prove a lemma which we use extensively.

LEMMA. *Let*

2.1 $$A(z, x) = \sum_{r=1}^{\infty} a_r(z) x^r ,$$

and

2.2 $$A(x) = \sum_{r=1}^{\infty} a_r x^r ,$$

with $a_r(z) > 0$, *satisfying*

2.3 $$\sum_{r=1}^{\infty} a_r(z) = c,\ 0 < z < 1 ,$$

c, a constant. Then for

2.4 $$a_r(z) \longrightarrow a_r,\ z \longrightarrow 1^- ,$$

it is necessary and sufficient that for $0 < x < 1$

2.5 $$A(z, x) \longrightarrow A(x),\ z \longrightarrow 1^- .$$

Proof of the lemma. First let us suppose that (2.4) holds. Then for fixed x, $(0 < x < 1)$ and ε, we can choose a number n_0 such that $\{x^{n_0}/(1 - x)\} < \varepsilon$. Then,

2.6 $$|A(z, x) - A(x)| < \sum_{r=0}^{n_0} |a_r(z) - a_r| x^r + 2c\varepsilon .$$

Now each term in the right hand side tends to zero. Hence the necessary part. Now suppose that (2.5) holds. Since $\{a_r(z)\}$ is bounded it is always possible to find a converging subsequence. If (2.4) is not true then we can extract two subsequences converging to two different sequences $\{a_r^*\}$ and $\{a_r^{**}\}$ and the corresponding subsequences of $\{A(z, x)\}$ would converge to $A^*(x) = \sum a_r^* x^r$ and $A^{**}(x) = \sum a_r^{**} x^r$ which contradicts the assumption that (2.5) holds. Hence $\{a_r^*\} = \{a_r^{**}\} = \{a_r\}$. This proves the sufficiency part.

3. The rth longest cycle. The mth raw moment of the rth longest cycle is

3.1 $$E_z\{L_r\}^m = \lambda \sum_{j=1}^{\infty} \frac{j^m}{P_r} \int_{t_j}^{t_{j+1}} \sum_{k=1}^{r} p_k v_{r-k}(t_\infty - t)\, dt .$$

Hence

3.2
$$\begin{aligned}\sum_{r=1}^{\infty} P_r x^{r-1} E_z\{L_r\}^m &= \lambda \sum_{r=1}^{\infty} x^{r-1} \sum_{j=1}^{\infty} j^m \int_{t_j}^{t_{j+1}} \sum_{k=1}^{r} p_k v_{r-k}(t_\infty - t)dt, \\ &= \lambda \sum_{j=1}^{\infty} j^m \int_{t_j}^{t_{j+1}} \sum_{r=1}^{\infty} x^{r-1} \Big\{\sum_{k=1}^{r} v_{r-k}(t_\infty - t)p_k\Big\}dt, \\ &= \lambda \sum_{j=1}^{\infty} j^m \int_{t_j}^{t_{j+1}} e^{\lambda[\phi(x)-1](t_\infty - t)}\{\phi(x)/x\}dt .\end{aligned}$$

Let $F = F(\lambda)$ denotes the left hand side of (3.2) and $F' = F(\lambda s^{1-\alpha})$.

3.3
$$\begin{aligned}F' &= s^{1-\alpha}\lambda \sum_{j=1}^{\infty} j^m \int_{t_j}^{t_{j+1}} e^{\lambda s^{1-\alpha}[\phi(x)-1](t_\infty - t)}\{\phi(x)/x\}dt , \\ &= \sum_{r=1}^{\infty} P_r x^{r-1} E_z(L_r')^m ,\end{aligned}$$

where L_r' is the same as L_r with λ replaced by $\lambda s^{1-\alpha}$.
Let us now consider some analytical preliminaries regarding $t_j(z)$. With $z = e^{-s}$, $0 < s < \infty$. We have

3.4
$$t_\infty(e^{-s}) - t_j(e^{-s}) = \sum_{k=j}^{\infty} \{e^{-ks}/k^\alpha\} .$$

In the interval $\{y: ks < y < (k+1)s\}$, we have

$$\frac{e^{-ks}}{k^\alpha s^\alpha} > \frac{e^{-y}}{y^\alpha} > \frac{e^{-(k+1)s}}{(k+1)^\alpha s^\alpha} ,$$

and

3.5
$$\frac{e^{-ks}s^{1-\alpha}}{k^\alpha} > \int_{ks}^{(k+1)s} \frac{e^{-y}}{y^\alpha} dy > \frac{s^{1-\alpha}e^{-(k+1)s}}{(k+1)^\alpha} .$$

Summing with respect to k, we have,

3.6
$$s^{1-\alpha} \sum_{k=j}^{\infty} (e^{-ks}/k^\alpha) > \int_{js}^{\infty} (e^{-y}/y^\alpha)dy .$$

Let

3.7
$$E(\theta) = \int_{\theta}^{\infty} (e^{-y}/y^\alpha)dy .$$

Then from (3.6) $E(js) < s^{1-\alpha} \sum_{k=j}^{\infty} e^{-ks}/k^\alpha$. Also

$$\int_{(j-1)s}^{\infty} (e^{-y}/y^\alpha)dy > s^{1-\alpha} \sum_{k=j}^{\infty} \{e^{-ks}/k^\alpha\} .$$

Combining the two

3.8
$$E(js) < s^{1-\alpha} \sum_{k=j}^{\infty} \{e^{-ks}/k^\alpha\} < E\{(j-1)s\} .$$

Now consider the equation

3.9 $$s^{1-\alpha}\sum_{k=j}^{\infty}\{e^{-ks}/k^{\alpha}\} = E(X) .$$

If $X_j(s)$ is the root of the equation (3.9), we have

3.10 and (i) $(j-1)s < X_j(s) < js$
(ii) $X_j(s)$ is unique.

In (3.3) put $E(\theta) = s^{1-\alpha}(t_\infty - t)$ so that

$$s^{1-\alpha}\,dt = \{e^{-\theta}/\theta^{\alpha}\}d\theta .$$

Hence

3.11 $$F' = \lambda\sum_{j=1}^{\infty} j^m \int_{X_j(s)}^{X_{j+1}(s)} \{\phi(x)/x\}\frac{e^{\lambda[\phi(x)-1]E(\theta)-\theta}}{\theta^{\alpha}}d\theta .$$

Let

$$\mu_j = \int_{X_j(s)}^{X_{j+1}(s)} d\mu(\theta) ,$$

where

3.12 $$d\mu(\theta) = \{e^{\lambda[\phi(x)-1]E(\theta)-\theta}/\theta^{\alpha}\}d\theta .$$

But

3.13 $$(j-1)s < X_j(s) < js \quad\text{and}\quad js < X_{j+1}(s) < (j+1)s .$$

This implies that

$$X_j(s) < js < X_{j+1}(s) .$$

Thus

$$s^m F' = \frac{\lambda\phi(x)}{x}\sum_{j=1}^{\infty}(js)^m \int_{X_j(s)}^{X_{j+1}(s)} d\mu(\theta) .$$

Now

3.14 $$\frac{\lambda\phi(x)}{x}\sum_{j=1}^{\infty} X_j^m(s)\mu_j \leqq F's^m \leqq \frac{\lambda\phi(x)}{x}\sum_{j=1}^{\infty} X_{j+1}^m(s)\mu_j .$$

Consider

3.15 $$\int_{X_1(s)}^{\infty}\theta^m d\mu(\theta) = \sum_{j=1}^{\infty}\int_{X_j(s)}^{X_{j+1}(s)}\theta^m d\mu(\theta) .$$

We have

3.16 $$\sum_{j=1}^{\infty} X_j^m(s)\mu_j \leqq \int_{X_1(s)}^{\infty}\theta^m d\mu(\theta) \leqq \sum_{j=1}^{\infty} X_{j+1}^m(s)\mu_j .$$

i.e.,

$$I_1 \leqq I \leqq I_2 \quad \text{(say)} ,$$

where

$$I_1 = \sum_{j=1}^{\infty} X_j^m(s)\mu_j,\ I_2 = \sum_{j=1}^{\infty} X_{j+1}^m(s)\mu_j$$

and

$$I = \int_{X_1(s)}^{\infty} \theta^m d\mu(\theta) \ .$$

I_1 and I_2 are the Darboux sums for the Stieltjes integral based on the above meshes. Also $X_1(s) \to 0$ as $s \to 0^+$. Hence

3.17
$$\begin{aligned} s^m F' &\sim \{\phi(x)/x\}\lambda \int_0^\infty \theta^{m-\alpha} e^{\lambda[\phi(x)-1]E(\theta)-\theta} d\theta,\ s \to 0^+,\ m \geqq \alpha\ , \\ &\sim \lambda \int_0^\infty \theta^{m-\alpha} e^{-\theta} d\theta \sum_{r=1}^{\infty} x^{r-1}\Big\{\sum_{k=1}^{r} v_{r-k}[E(\theta)]p_k\Big\},\ s \to 0^+ \ . \end{aligned}$$

Now

3.18
$$\begin{aligned} s^m \sum_{r=1}^{\infty} P_r E_z(L'_r)^m &= \lambda s^{m+1-\alpha} \sum_{j=1}^{\infty} j^m \int_{t_j}^{t_j+1} dt = \lambda s^{m+1-\alpha} \sum_{j=1}^{\infty} j^m (t_{j+1} - t_j) \\ &= \lambda s^{m+1-\alpha} \sum_{j=1}^{\infty} j^m \{e^{-js}/j^\alpha\} = \lambda s^{m+1-\alpha} \sum_{j=1}^{\infty} \{e^{-js}/j^{\alpha-m}\} < \infty \ . \end{aligned}$$

Hence using the lemma

3.19
$$s^m P_r E_z(L'_r)^m \sim \lambda \int_0^\infty \Big[\sum_{k=1}^{r} v_{r-k}[E(\theta)]p_k\Big] e^{-\theta}\theta^{m-\alpha} d\theta,\ s \to 0^+ \ .$$

Since $s \sim (1 - z)$,

$$(1 - z)^m E_z(L'_r)^m \sim (\lambda/P_r) \int_0^\infty \Big[\sum_{k=1}^{r} v_{r-k}[E(\theta)]p_k\Big] e^{-\theta}\theta^{m-\alpha} d\theta,\ z \to 1^- \ .$$

Taking $\lambda = 1, \alpha = 1, p_1 = 1, p_2 = p_3 \cdots = 0$, we now have

3.20
$$\begin{aligned} (1 - z)^m E_z\{L_r\}^m &\sim \int_0^\infty v_{r-1}[E(\theta)] e^{-\theta}\theta^{m-1} d\theta,\ z \to 1^-, \\ &\sim \int_0^\infty e^{-E(\theta)-\theta}[E(\theta)]^{r-1}\{\theta^{m-1}/(r-1)!\} d\theta,\ z \to 1^- \ . \end{aligned}$$

This is in agreement with Shepp and Lloyd.

4. The rth shortest cycle. Let S_r be the length of the rth shortest cycle. Then

4.1
$$P\{S_r = j\} = (\lambda/P_r) \int_{t_j}^{t_{j+1}} \sum_{k=1}^{r} p_k v_{r-k}(t) dt \ .$$

Let

$$F_1 = F_1(\lambda) = \sum_{r=1}^{\infty} P_r x^{r-1} E_z\{S_r\}^m \ .$$

Then

4.2
$$F_1 = \lambda \sum_{r=1}^{\infty} x^{r-1} \sum_{j=1}^{\infty} j^m \int_{t_j}^{t_{j+1}} \sum_{k=1}^{r} p_k v_{r-k}(t)dt \,,$$
$$= \lambda \sum_{j=1}^{\infty} j^m \int_{t_j}^{t_{j+1}} e^{\lambda[\phi(x)-1]t}\{\phi(x)/x\}dt \,.$$

Also
$$F_1' = F_1(\lambda s^{1-\alpha}) = \sum P_r x^{r-1} E_z(S_r')^m \,,$$

where S_r' is the same as S_r with λ replaced by $\lambda s^{1-\alpha}$. Put $(t_\infty - t)s^{1-\alpha} = E(\theta)$ in F_1'.

4.3
$$F_1' = \lambda \sum_{j=1}^{\infty} j^m \int_{X_j(s)}^{X_{j+1}(s)} \{\phi(x)/(x\theta^\alpha)\} e^{\lambda[s^{1-\alpha}t_\infty - E(\theta)][\phi(x)-1]-\theta} d\theta \,.$$

Let
$$\mu_j = \int_{X_j(s)}^{X_{j+1}(s)} d\mu(\theta) \,,$$

where

4.4
$$d\mu(\theta) = \{\phi(x)/x\theta^\alpha\} e^{\lambda[s^{1-\alpha}t_\infty - E(\theta)][\phi(x)-1]-\theta} d\theta \,.$$

Hence

4.5
$$s^m F_1' = \lambda \sum_{j=1}^{\infty} (js)^m \int_{X_j(s)}^{X_{j+1}(s)} \{\phi(x)/x\theta^\alpha\} e^{\lambda[s^{1-\alpha}t_\infty - E(\theta)][\phi(x)-1]-\theta} d\theta \,.$$

Since $(j-1)s < X_j(s) < js < X_{j+1}(s) < (j+1)s$,

4.6
$$\lambda \sum_{j=1}^{\infty} X_j^m(s)\mu_j < F_1' s^m < \lambda \sum_{j=1}^{\infty} X_{j+1}^m(s)\mu_j \,.$$

Also
$$\sum_{j=1}^{\infty} X_j^m(s)\mu_j < \sum_{j=1}^{\infty} \int_{X_j(s)}^{X_{j+1}(s)} \theta^m d\mu(\theta) < \sum_{j=1}^{\infty} X_{j+1}^m(s)\mu_j \,.$$

That is

4.7
$$\sum_{j=1}^{\infty} X_j^m(s)\mu_j < \int_{X_1(s)}^{\infty} \theta^m d\mu(\theta) < \sum_{j=1}^{\infty} X_{j+1}^m(s)\mu_j \,.$$

Hence

4.8
$$s^m F_1' \sim \lambda \int_0^\infty \theta^{m-\alpha}\{\phi(x)/x\} e^{\lambda[s^{1-\alpha}t_\infty - E(\theta)][\phi(x)-1]-\theta} d\theta,\ s \to 0^+, \quad m \geqq \alpha \,.$$

Here also $s^m \sum_{r=1}^{\infty} P_r E_z(S_r')^m = s^{m+1-\alpha} \sum_{j=1}^{\infty} j^m (t_{j+1} - t_j) < \infty$ {by (3.18)}. Thus as in 3.17 by equating the coefficient of x^{r-1} on both sides we can obtain $\lim_{s\to 0} s^m P_r E_z(S_r')^m$.

Now let us consider the particular case of the above when $p_1 = 1$, $p_2 = p_3 = \cdots = 0$ $\lambda = 1$ and $\alpha = 1$. Here

4.9
$$s^m \sum_{r=1}^{\infty} x^{r-1} E_z(S_r)^m \sim \int_0^\infty \theta^{m-1} e^{(x-1)[\log(1-z)^{-1}]-(x-1)E(\theta)-\theta} d\theta,\ z \to 1^-,$$
$$\sim s \int_0^\infty \theta^{m-1} e^{-x[E(\theta)+\log s]+E(\theta)-\theta} d\theta,\ s \to 0^+ \,.$$

Hence

$$s^{m-1}\sum_{r=1}^{\infty} x^{r-1}E_z(S_r)^m \sim e^{x\log(s^{-1})}\int_0^{\infty} e^{-xE(\theta)+E(\theta)-\theta}\theta^{m-1}d\theta \ .$$

So

4.10 $$\frac{(1-z)^{m-1}}{(m-1)!}\sum_{r=1}^{\infty} x^{r-1}E_z(S_r)^m \sim \frac{1}{(m-1)!}\times \left[\int_0^{\infty} e^{E(\theta)-\theta}\theta^{m-1}\sum_{r=1}^{\infty}\frac{[-xE(\theta)]^{r-1}}{(r-1)!}d\theta\right]\times\left[\sum_{r=1}^{\infty}\frac{[x\log(1-z)^{-1}]^{r-1}}{(r-1)!}\right].$$

Equating coefficient of x^{r-1} on both sides of 4.10

$$\frac{(1-z)^{m-1}}{(m-1)}E_z(S_r)^m \sim \frac{1}{(m-1)!}\sum_{p=0}^{r-1}\left[\{[\log(1-z)^{-1}]^p/p!\}\times\left\{\int_0^{\infty}\frac{[-E(\theta)]^{r-1-p}\theta^{m-1}e^{E(\theta)-\theta}}{(r-1-p)!}d\theta\right\}\right],\ s\to 0^+$$

4.11 $$\sim \sum_{p=0}^{r-1}(1/p!)[\log(1-z)^{-1}]^pK(r-1-p,m),\ s\to 0^+\ ,$$

where

4.12 $$K(q,m)=\int_0^{\infty}\frac{\theta^{m-1}[-E(\theta)]^q e^{E(\theta)-\theta}}{(m-1)!\,q!}d\theta$$

which is in agreement with Shepp and Lloyd.

References

1. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol-1, Asia Pub. Co., 1969.
2. A. Y. Khintchine, *Mathematical methods in the theory of Queueing*, Griffin Statistical monographs, **7** (1960).
3. John Riordan, *An Introduction to Combinatorial Analysis*, Wiley, New York, 1958.
4. L. A. Shepp and S. P. Lloyd, *Ordered lengths in a random permutation*, Trans. Amer. Math. Soc. **121** (1966), 340-357.

Received October 23, 1969, and in revised form May 14, 1970.

ANNAMALAI UNIVERSITY
ANNAMALAINAGAR (P.O),
TAMIL NADU, INDIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

NONTANGENTIAL HOMOTOPY EQUIVALENCES

VICTOR A. BELFI

The purpose of this paper is to apply surgery techniques in a simple, geometric way to construct manifolds which are nontangentially homotopy equivalent to certain π-manifolds. Applying this construction to an H-manifold of the appropriate type yields an infinite collection of mutually nonhomeomorphic H-manifolds, all nontangentially homotopy equivalent to the given one.

The theorem proved is the following: If N^{4k} is a smooth, closed, orientable π-manifold and L^m is a smooth, closed, simply connected π-manifold, there is a countable collection of smooth, closed manifolds $\{M_i\}$ satisfying (1) no M_i is a π-manifold, (2) each M_i is homotopy equivalent but not homeomorphic to $N \times L$, (3) M_i is not homeomorphic to M_j if $i \neq j$.

1. Construction of the surgery problem. In [2] Milnor describes a $(2k-1)$-connected, bounded π-manifold of dimension $4k$ and Hirzebruch index 8 ($k \geqq 2$). This manifold, which we denote by Y^{4k}, is obtained by plumbing together 8 copies of the tangent disk bundle of S^{2k} according to a certain scheme. This implies that Y has the homotopy type of a bouquent of eight $2k$-spheres. The only other property of Y which we shall need is that ∂Y is a homotopy sphere. Let r be the order of ∂Y^{4k} in the group of homotopy spheres bP_{4k} [3] and take W^{4k} to be the r-fold connected sum along the boundary of Y^{4k}. By the choice of r, ∂W is diffeomorphic to S^{4k-1}. Attaching a $4k$-disk to W by a diffeomorphism along the boundary, we obtain a closed, smooth manifold $\hat{W}$, which is $(2k-1)$-connected and has index $8r$. By the Hirzebruch index theorem $\hat{W}$ is not a π-manifold, but is almost parallelizable.

Define $f: W^{4k} \rightarrow D^{4k}$ by the identity on the boundary, stretching a collar of ∂W over D^{4k}, and sending the remainder of W to a point. This gives a degree 1 map $f: (W, \partial W) \rightarrow (D^{4k}, \partial D^{4k})$ which is tangential since both W and D^{4k} are π-manifolds. f is already a homotopy equivalence on the boundary, so we have a surgery problem in the bounded case. The connectedness of W implies that f is already an isomorphism in homology below the middle dimension. However the kernel of f_* in dimension $2k$ is $\underbrace{Z \oplus \cdots \oplus Z}_{8r}$ and the index of the kernel is the index of W which is $8r$. Thus it is not possible to complete the surgery.

But if L^m is a closed, smooth, simply connected π-manifold, the surgery problem $f \times 1_L: W \times L \rightarrow D^{4k} \times L$ does have a solution. To see this note first that if m is odd, the problem is odd dimensional so there are no obstructions to modifying $W \times L$ and $f \times 1_L$ by surgery to obtain a homotopy equivalence. If $m \equiv 0 \pmod 4$, the problem has an index obstruction given by the product of the index obstruction of the map f and the index of the manifold L, i.e., $I(f \times 1_L) = I(f) \cdot I(L)$. This product vanishes since L is a π-manifold. The formula follows from the multiplicativity of the index of a manifold. If $m \equiv 2 \pmod 4$ the problem has a Kervaire invariant obstruction given by the mod 2 product of the Kervaire invariant obstruction of f and the Euler characteristic of L, the formula arising from Sullivan's characterization of the Kervaire invariant obstruction [8]. Since L is a π-manifold, $\chi(L) = 0$; so $K(f \times 1_L)$ vanishes as well.

Now we change the surgery problem discussed above into a problem for closed manifolds. Let N be a smooth, closed, π-manifold of dimension $4k$. Take a small disk D^{4k} in N and form the connected sum $N \# \hat{W}$ using this disk and the disk attached to W to make $\hat{W}$. Define $1_N \# f \colon N \# \hat{W} \to N$ by the identity on N-int D^{4k} and f on W. Although $(1_N \# f) \times 1_L$ is not tangential, it can be surgered to a homotopy equivalence. This is because it is already a homotopy equivalence except on $W \times L$, where it is tangential; so it suffices to do surgery on $W \times L$ leaving the boundary fixed to make $N \# \hat{W} \times L$ homotopy equivalent to $N \times L$. We have already seen that this can be done. Summing up the discussion we have

PROPOSITION 1. *Suppose N^{4k} is a closed, smooth, orientable π-manifold and L^m is a closed, smooth, simply connected π-manifold. Then there is a manifold M^{4k+m}, homotopy equivalent to $N \times L$ obtained by surgery on $(1_N \# f) \times 1_L$.*

Notice that if $W_i^{4k} = \underbrace{W^{4k} \# \cdots \# W^{4k}}_{i}$, and we define $f_i \colon W_i \to D^{4k}$ the same way as we defined f, the above considerations also apply to W_i. The only difference is that W_i has index $8ri$. We shall denote the solution to the surgery problem using W_i by M_i^{4k+m}.

We also remark here that M, as a solution to a given surgery problem, is unique up to PL homeomorphism, but not not always up to diffeomorphism. This follows from Novikov's results [5]. Since we shall be primarily concerned with the topological type of such solutions, we shall ignore this ambiguity.

2. Properties of the surgery solution.

PROPOSITION 2. *The manifold M^{4k+m} obtained by surgery on*

$$(1_N \# f) \times 1_L \colon N \# \hat{W} \times L \to N \times L$$

is not a π-manifold.

Proof. After surgery we have a homotopy equivalence $g \colon M \to N \times L$ and a cobordism Z between M and $N \# W \times L$ together with a map $F \colon Z \to N \times L$ whose restriction is g on M and $(1_N \# f) \times 1_L$ on $N \# \hat{W} \times L$. If $*$ is a point of L, $(1_N \# f) \times 1_L$ is transverse regular with respect to $N \times *$. Change g by a small homotopy to make it transverse regular with respect to $N \times *$. Finally leaving $(1_N \# f) \times 1_L$ and g fixed, make F transverse regular with respect to $N \times *$ to obtain the oriented cobordism $F^{-1}(N \times *)$ between $N \# \hat{W}$ and

$$S = g^{-1}(N \times *) .$$

Because $N \# \hat{W}$ and S are oriented cobordant, $I(S) = I(N \# \hat{W}) \neq 0$. We have the usual equivalence of tangent and normal bundles

$$\tau(M) | S \cong \tau(S) \oplus \nu(S \subset M) .$$

Since f is transverse regular with respect to $N \times *$ and

$$\nu(N \times * \subset N \times L)$$

is trivial, $\nu(S \subset M)$ is trivial. Thus if $\nu(M) | S$ were stably trivial, $\tau(S)$ would be stably trivial, contradicting $I(S) \neq 0$. Therefore $\tau(M) | S$ is not stably trivial and consequently $\tau(M)$ is not stably trivial.

PROPOSITION 3. *M is not homeomorphic to $N \times L$.*

Proof. Suppose $h \colon M \to N \times L$ is a homeomorphism. Denote by $p_j(M)$ the j^{th} Pontrjagin class of M (i.e., of $\tau(M)$) and by $p_j(M; \mathbf{Q})$ the j^{th} rational Pontrjagin class of M. In the proof of Proposition 2 it was shown that M^{4k+m} contains a closed submanifold S of dimension $4k$ and index $8r$. If $i \colon S \to M$ is inclusion, the Hirzebruch index theorem implies

$$\begin{aligned} 8r &= \langle L_k(p_1(S), \cdots, p_k(S)), [S] \rangle \\ &= \langle L_k(i^* p_1(M), \cdots, i^* p_k(M)), [S] \rangle \\ &= \langle L_k(p_1(M), \cdots, p_k(M)), i_*[S] \rangle . \end{aligned}$$

Now we may replace $p_j(M)$ by $p_j(M; \mathbf{Q})$ since any torsion evaluated on the orientation class is zero. By the topological invariance of rational Pontrjagin classes, $p_j(M; \mathbf{Q}) = h^*(p_j(N \times L); \mathbf{Q})$; but

$$p_j(N \times L; \mathbf{Q}) = 0$$

for every j because $N \times L$ is *a* π-manifold. Therefore $p_j(M; \mathbf{Q}) = 0$

for every j, a contradiction.

Observe that Propositions 2 and 3 are likewise valid for the manifolds M_i, each M_i containing a closed submanifold S_i of dimension $4k$ and index $8ri$.

Now we are in a position to prove the central theorem of this paper.

THEOREM 1. *Suppose N is a smooth, closed, orientable π-manifold of dimension $4k (k \geqq 2)$ and L is a smooth, closed simply connected π-manifold. Then there is a countable sequence of smooth, closed manifolds $\{M_i\}$ having the following properties*: (1) *no M_i is a π-manifold,* (2) *each M_i is homotopy equivalent but not homeomorphic to $N \times L$,* (3) *M_i is not homeomorphic to M_j if $i \neq j$.*

Proof. The M_i's are the surgery solutions already described. Propositions 2 and 3 establish (1) and (2). It remains to prove (3). We do this by expanding the idea of the proof of Proposition 3.

Suppose there exists a homeomorphism $h: M_j \rightarrow M_i$ and $i \neq j$, say $i > j$. (For the rest of this paragraph $t = i, j$.) Let $g_t: M_t \rightarrow N \times L$ be a homotopy equivalence which is transverse regular with respect to $N \times *$ so that $g_t^{-1}(N \times *) = S_t$ where $I(S_t) = 8rt$. (We may assume that g_t is still the identity on $(N - \operatorname{int} D^{4k}) \times L$ since no surgery is done there.) Then by the index theorem,

$$\langle L_k(p_1(M_t; \mathbf{Q}), \cdots, p_k(M_t; \mathbf{Q})), [S_t] \rangle = I(S_t) .$$

To simplify notation we omit explicit reference to the inclusion maps $S_t \subset M_t$ and abbreviate $L_k(p_1(X; \mathbf{Q}), \cdots, p_k(X; \mathbf{Q}))$ by $L_k(X)$. Let $\bar{g}_t$ be a homotopy inverse for g_t. The idea is then to show that $g_i h \bar{g}_j$ does not behave properly on rational homology. We shall be referring to the following diagram for the rest of the proof:

$$\begin{tikzcd}
S_j \subset M_j \arrow[rr, "h"] \arrow[dr, shift left, "g_j"] & & M_i \supset S_i \arrow[dl, shift right, "g_i"'] \\
& N \times L \arrow[ul, shift left, "\bar{g}_j"] \arrow[ur, shift right, "\bar{g}_i"'] &
\end{tikzcd}$$

By the transverse regularity of g_t, it follows that

$$g_{t*}[S_t] = [N \times *] = [N] \otimes 1 \in H_{4k}(N \times L; \mathbf{Q}) ,$$

so $g_{j*}\bar{g}_{i*}[S_i] = [S_j]$. Thus

$$I(S_j) = \langle L_k(M_j), \bar{g}_{j*}g_{i*}[S_i]\rangle = \langle L_k(M_i), h_*\bar{g}_{j*}g_{i*}[S_i]\rangle$$

by the topological invariance of rational Pontrjagin classes.

Define a bundle ξ over $N \times L$ by $\bar{g}_i^*(\tau(M_i))$. This means that $\tau(M_i) = g_i^*(\xi)$. Since g_i is the identity on $N - \operatorname{int} D^{4k} \times L$ and

$$\tau(M_i) \,|\, N - \operatorname{int} D^{4k} \times L$$

is trivial, it follows that $\xi | N - \operatorname{int} D^{4k} \times L$ is trivial. Now if

$$i \colon N - \operatorname{int} D^{4k} \times L \to N \times L$$

is inclusion, then if $x \otimes y \in H_*(N \times L; \mathbf{Q})$ and $\dim x < 4k$, $x \otimes y \in$ image i_*, say $x \otimes y = i_* z$. Thus $\langle L_k(\xi), x \otimes y\rangle = \langle L_k(i^*\xi), z\rangle = 0$ since $i^*\xi$ is trivial. This shows that if $\gamma_{4k} \in H_{4k}(N \times L; \mathbf{Q})$, then $\langle L_k(\xi), \gamma_{4k}\rangle$ is given by the product of the coefficient of $[N] \otimes 1$ in γ_{4k} and

$$\langle L_k(\xi), [N] \otimes 1\rangle .$$

Using the preceding observation, we can compute the coefficient of $[N] \otimes 1$ in $(g_i h \bar{g}_j)_*[N] \otimes 1$ as follows.

$$\begin{aligned}\langle L_k(\xi), (g_i h \bar{g}_j)_*[N] \otimes 1\rangle &= \langle L_k(M_i), h_*\bar{g}_{j*}[N] \otimes 1\rangle \\ &= \langle L_k(M_i), h_*\bar{g}_{j*}g_{i*}[S_i]\rangle \\ &= I(S_j) = (j/i)I(S_i) .\end{aligned}$$

But

$$I(S_i) = \langle L_k(M_i), [S_i]\rangle = \langle L_k(\xi), g_{i*}[S_i]\rangle = \langle L_k(\xi), [N] \otimes 1\rangle .$$

Hence this coefficient is j/i which is not an integer since $i > j$. This contradicts the fact that any induced map on rational homology must send integral classes to integral classes.

3. An extension of the results. It has been pointed out to me that the results of this paper can be extended in the following way:

If M^n is a simply connected smooth manifold where n is odd and $H^{4k}(M; \mathbf{Q}) \neq 0$ or some $4k < n$, the Pontrjagin character shows that $\widetilde{KO}(M)$ is infinite. (See, for example, Hsiang [2].) Thus the kernel of $\widetilde{KO}(M) \to J(M)$ is infinite. It can be shown that the result of doing surgery on the elements of the kernel is a collection of smooth manifolds homotopy equivalent to M containing an infinite subset $\{M_i\}$ of mutually non-homeomorphic manifolds. The condition on the rational cohomology of M is also necessary for the manifolds $\{M_i\}$ exist.

Although the theorem described above considerably extends the class of manifolds to which the principal result applies, its proof requires methods of a deeper sort and the geometric simplicity is lost.

4. Applications. By an *H-manifold* we mean a closed, orientable topological manifold having the structure of an H-space.

THEOREM 2. *Suppose N^{4k} and L^m are smooth H-manifolds, N and L are π-manifolds, and L is simply connected. Then there exists a sequence of mutually nonhomeomorphic smooth H-manifolds $\{M_i\}$ satisfying* (1) *no M_i is a π-maifold,* (2) *each M_i is homotopy equivalent, but not homeomorphic to $N \times L$.*

Proof. This is immediate from Theorem 1 since the product of 2 H-manifolds is an H-manifold and any manifold homotopy equivalent to an H-manifold is itself an H-manifold.

Examples of manifolds nontangentially homotopy equivalent to Lie groups were known before surgery techniques were introduced; however all these were nonsimply connected. An example due to Milnor of a manifold homotopy equivalent to $S^1 \times S^3 \times S^7$ with a nonzero Pontrjagin class is quoted by Browder and Spanier [1].

The recent results of a A. Zabrodsky [9] and J. Stasheff [7] have produced new homotopy types of H-manifolds (other than compact Lie groups) to which Theorem 2 applies. However if we restrict ourselves to simply connected, compact Lie groups, we can obtain a stronger conclusion.

THEOREM 3. *Suppose N^{4k} and L^m are simply connected compact Lie groups ($k \geqq 2$). Then there is a countable sequence of mutually nonhomeomorphic H-manifolds $\{M_i\}$ satisfying* (1) *no M_i is a π-manifold,* (2) *each M_i is homotopy equivalent to $N \times L$ but not homeomorphic to any Lie group.*

Proof. Since Lie groups are π-manifolds, Theorem 1 applies. H. Scheerer has proved [6] that homotopy equivalent, compact, simply connected Lie groups are isomorphic; so if M_i were homeomorphic to any Lie group, it would be homeomorphic to $N \times L$, contradicting Theorem 1.

The author is indebted to John W. Morgan for his invaluable suggestions in the course of this research and to the refree for pointing out related results.

REFERENCES

1. W. Browder and E. Spanier, *H-spaces and duality*, Pacific J. Math. **12** (1962), 411–414.

2. W.-C. Hsiang, *A note on free differentiable actions of S^1 and S^3 on homotopy spheres*, Ann. of Math. **83** (1966), 266–272.

3. M. A. Kervaire and J. W. Milnor, *Groups of homotopy spheres*: I, Ann. of Math. **77** (1963), 504-537.
4. J. W. Milnor, "Differential topology," *Lectures on Modern Mathematics*, Vol. II, T. L. Saaty, ed., John Wiley and Sons, Inc., New York, 1964.
5. S. P. Novikov, *Homotopically equivalent smooth manifolds*, I, Trans. Amer. Math. Soc. (2) **48** (1965), 271-396.
6. H. Scheerer, *Homotopieaquivalente kompacte Liesche Gruppen*, Topology **7** (1968), 227-232.
7. J. Stasheff, *Manifolds of the homotopy type of (non-Lie) groups*, Bull. Amer. Math. Soc. **75** (1969), 998-1000.
8. D. Sullivan, *Geometric Topology Seminar Notes* Mimeographed, Princeton University.
9. A. Zabrodsky, *Homotopy associativity and finite CW complexes*, Mimeographed Notes, University of Illinois, Chicago Circle, Ill., 1968.

Received October 3, 1969, and in revised form June 9, 1970. The author held a NASA graduate fellowship while this research was done for his Ph. D. thesis at Rice University in 1969.

TEXAS CHRISTIAN UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

COMPACT SEMIGROUPS WITH SQUARE ROOTS

JANE M. DAY

Suppose that S is a finite dimensional cancellative commutative clan with $E = \{0, 1\}$ and that H is the group of units of S. We show that if square roots exist in S/H, not necessarily uniquely, then there is a closed positive cone T in E^n for some n and a homomorphism $f: (T \cup \infty) \times H \rightarrow S$ which is onto and one-to-one on some neighborhood of the identity. $T \cup \infty$ denotes the one point compactification of T.

K. Keimel proved in (6), and Brown and Friedberg independently in (1), that if S/H is uniquely divisible, then it is isomorphic to $T \cup \infty$ for some closed positive cone T. Brown and Friedberg went on to show that if S is uniquely divisible, then S is isomorphic to the Rees quotient $((T \cup \infty) \times H)/(\infty \times H)$. What we do here is to weaken their hypothesis to assume just square roots in S/H and conclude that S is isomorphic to some quotient of such $(T \cup \infty) \times H$, which will be a Rees quotient if square roots are unique in $(S/H)\backslash 0$, but in general need not be Rees.[1] $f((T \cup \infty) \times 1)$ is a subclan of S and a local cross section at 1 for the orbits of the group action $H \times S \rightarrow S$ (which equal $\mathscr{H}$ classes here), but an example shows that it need not be a full cross section. Also, square roots exist (uniquely) in S if and only if they exist (uniquely) in S/H and H.

The proof consists essentially of showing that the ingenious constructions of (1) can still be done under the weaker hypothesis, in a sufficiently small neighborhood of H.

For basic information about semigroups, see (5), (8) or (9). The real intervals $(0, 1]$ and $[0, 1]$ are semigroups under usual real multiplication; as in (5), a *one parameter semigroup* is a homomorph of $(0, 1]$, and we also define here a *closed one parameter semigroup* to be a nonconstant homomorph of $[0, 1]$.

The Lemmas (I)-(III) are variations on standard themes so we omit proofs. (See (1), (3), (4), B-3 of (5), (6) and (7).) Throughout this paper let S be a clan with exactly two idempotents, a zero and an identity denoted by 0 and 1 respectively.

(I) *If R is a one parameter semigroup in S which is not contained in H and is not equal to 0, then $R \cup 0$ is a closed one parameter semigroup and an arc with endpoints 0 and 1. Let $\phi: (0, 1] \rightarrow R$ be the homomorphism that defines R; if $x = \phi(t) \in R$ and $k \geqq 0$, we write*

---

[1] Keimel has concurrently proved a further generalization, by a different method, assuming instead of cancellation that $x \times H \rightarrow xH$ is one-to-one for all x near H.

x^k for $\phi(t^k)$, and if $x \neq 0, 1$, each $y \in R\backslash 0$ equals x^k for unique k.

(II) *If H is normal and every element of S/H has a square root in S/H, then for each $x \in S$ there exists a closed one parameter semigroup in S intersecting Hx.*

(III) *Let T be a commutative uniquely divisible clan with group of units $H(T)$ and $E = \{0, 1\}$, and let V be a set containing a neighborhood of 1 in T such that $T\backslash V$ is an ideal. If S is commutative and ψ': $V \rightarrow S$ is a continuous function such that $\psi'(V\backslash H(T)) \cap H = \square$ and $\psi'(xy) = \psi'(x)\psi'(y)$ whenever $x, y, xy \in V$, then ψ' can be extended to a homomorphism ψ on all of T by defining $\psi(0) = 0$ and $\psi(x^n) = \psi'(x)^n$ for each $x \in V$ and positive integer n.*

The definition of independent family which follows agrees with the algebraic independence used in [1] when H is trivial and $W = S\backslash 0$, and that notion is due to Clark [2]. We include H in our definition so that we do not have to handle the case of S with trivial H separately first, and we define independence in neighborhoods of H rather than in S in order to apply the concept effectively to a clan with nonunique roots.

An *independent family* in S is a finite family $\{R_1, \cdots, R_n\}$ of closed one parameter semigroups in S such that there exists a neighborhood W of H with the property that for every partition of the set $\{1, \cdots, n\}$ into two nonnull disjoint sets A and B, this is true:

$$\underset{i \in A}{P}\{R_i\} \cap (\underset{i \in B}{P}\{R_i\})H \cap W \subset H .$$

We will also describe this situation by saying that $\{R_1, \cdots, R_n)$ is *independent in W*. We adopt the convention that if $X = \square$, then $P_{i \in X}\{x_i\} = 1$, for x_i's which are elements or subsets of S. S will be called *cancellative* if $x, y, z \in S$ and $xy = xz \neq 0$ implies $y = z$.

We will make frequent use of the following facts. $F(V)$ denotes boundary of V. Any neighborhood of H in compact S contains a neighborhood V of H such that $S\backslash V$ is an ideal (A-3.1, (5)), and if V is a set such that $S\backslash V$ is an ideal, then

$$0 \notin V,\ V = VH,\ F(V) = F(V)H,$$

$S\backslash V^*$ is an ideal if nonempty, and $xy \in V$ implies $x, y \in V$. If J is a closed ideal in compact S, shrinking J to a point gives a new compact semi-group denoted S/J and called the *Rees quotient* of S by J, and the natural map $S \rightarrow S/J$ is a homomorphism.

Part (i) of the lemma below is analogous to 1.4 of (1); part (ii) shows that the homomorphisms ϕ: $S\backslash 0 \rightarrow E^n$ and β: $S\backslash 0 \rightarrow H$ constructed in (1) can still be constructed here on a sufficiently small neighborhood of H. *Dim S* means inductive dimension of S.

LEMMA. *Let S be a cancellative commutative clan with $E = \{0, 1\}$ and let W be a closed neighborhood of* 1 *such that $S \backslash W$ is an ideal.*

(i) *If $\{R_1, \cdots, R_n\}$ is an independent family in W, and if $x_1 x_2 \cdots x_n h = x'_1 x_{2'} \cdots x'_n h' \in W$, where $x_i, x'_i \in R_i$ for each i and $h, h' \in H$, then $x_i = x'_i$ for each i and $h = h'$; consequently* $\dim S \geqq n$.

(ii) *Suppose* $\dim S \leqq N$ *or* $\dim S/H \leqq N$ *and that S/H has square roots. Then there exists a maximal independent family $\{R_1, \cdots, R_n\}$ of closed one parameter semigroups in S, and a closed neighborhood U of H may be chosen so that $S \backslash U$ is an ideal and if $x \in U$, x satisfies this condition.*

$(\nmid)$ There exists a unique partition (A, B) of $\{1, \cdots, n\}$ and unique elements $x_i \in R_i$ and $h \in H$ such that $i \in B$ whenever $x_i = 1$ and $x(P_{i \in A}\{x_i\}) = (P_{i \in B}\{x_i\})h \in W$.

Proof. (i) Since R_i is a closed one parameter semigroup and $x_i \neq 0$, we may factor x_i or x'_i for each i and then commute and cancel in the equality given to get $0 \neq P_{i \in A}\{r_i\} = (P_{i \in B}\{r_i\})h'h^{-1}$ for some partition (A, B) of $\{1, \cdots, n\}$. These points lie in W so by independence, $r_i = 1$, hence $x_i = x'_i$, for each i, and thus $h = h'$ also. There is a closed neighborhood V of 1 such that $V^n \subset W$, and then the multiplication function $(R_1 \cap V) \times \cdots \times (R_n \cap V) \rightarrow S$ is a homeomorphism so S contains an n-cell.

(ii) If $\dim S \leqq N$, then a maximal independent family exists by (i). If $\dim S/H \leqq N$ instead, S/H is cancellative since S is, so (i) can be applied to S/H to get a maximal independent family in S/H; a closed one parameter semigroup in S projects to a closed one parameter semigroup in S/H by (I), and it is easy to see that an independent family in S projects to one in S/H, so S can have no larger independent family than S/H does.

Now choose a maximal independent family $\{R_1, \cdots, R_n\}$ in S, and choose W smaller if necessary so that the R_i's are actually independent in a neighborhood of H containing W^2.

To prove the uniqueness assertion of $(\nmid)$, suppose that

$$x(\underset{i \in A}{P}\{x_i\}) = (\underset{i \in B}{P}\{x_i\})h \in W \quad \text{and} \quad x(\underset{i \in A'}{P}\{x'_i\}) = (\underset{i \in B'}{P}\{x'_i\})h' \in W ,$$

as described in $(\nmid)$. Then

$$(\underset{i \in A}{P}\{x_i\})(\underset{i \in B'}{P}\{x'_i\})h' = (\underset{i \in A'}{P}\{x'_i\})(\underset{i \in B}{P}\{x_i\})h \in W^2 ;$$

for each i, collect into one term the x_k's with $k = i$, on each side, and suppose there exists $j \in A \cap B'$; $j \in A$ implies that the factor on the left which is an element of R_j is not 1, and it has to equal one of the factors on the right by (i); therefore j has to be in A' or in B, because by independence an element of $(R_j \cap W^2)\backslash 1$ cannot arise from multiples of elements of R_i's for $i \neq j$. But $j \in B$ implies $j \notin A$ and $j \in A'$ implies $j \notin B'$, both contradictions. So $A \cap B'$ must be empty, similarly $A' \cap B$ is empty, hence $(A, B) = (A', B')$. Now apply (i).

Now let R be any closed one parameter semigroup in S.

$$\{R, R_1, \cdots, R_n\}$$

is not independent in any neighborhood of H (where R and R_i are each counted if $R = R_i$ for some i), so there is a particular partition (A_R, B_R) of $\{1, \cdots, n\}$ such that $T = RP \cap QH$ contains points arbitrarily near H in $S \backslash H$, where $P = P_{i \in A}\{R_i\}$ and $Q = P_{i \in B_R}\{R_i\}$. T is also a compact semigroup, so it contains a connected subsemigroup from 1 to 0 (B-4.9, (5)). $F(W)$ separates 0 and 1 in S, hence we may select $x_R \in R$ such that $x_R P \cap QH \cap F(W) \neq \square$. Every $x \geqq x_R$ in R satisfies $(\nmid)$ since the complement of an ideal in R is connected and $\{x \in R \mid xP \cap QH \subset S \backslash W\}$ is an ideal of R. It follows that every $x \geqq x_R$ in RH satisfies $(\nmid)$ also.

If we can find a closed neighborhood U of H such that $x_R \notin U$ for each closed one parameter semigroup R in S, then every $y \in U$ lies in some RH by (II), U may be chosen smaller so that $S \backslash U$ is an ideal, and then every $y \in U$ satisfies $(\nmid)$ by the preceding remark. Suppose no such U exists, so there is a net (x_R) of the x_R's clustering at some element of H; since there exist only a finite number of partitions of $\{1, \cdots, n\}$, we may suppose that for one particular partition (A, B) and for each x_R in the net, $(A_R, B_R) = (A, B)$. Then, since $F(W) = F(W)H$, any cluster point of (a_R) is an element of

$$\underset{i \in A}{P}\{R_i\} \cap (\underset{i \in B}{P}\{R_i\})H \cap F(W) ;$$

but this set is empty (by definition if $A = \square$, and if $A \neq \square$, by independence in W).

Euclidean n-space, denoted E^n, is a semigroup under vector addition with the origin as identity. If P^* is the set of nonnegative real numbers, N the set of negative real numbers, and juxtaposition denotes scalar multiplication, a *closed positive cone* in E^n is defined to be a closed subsemigroup T of E^n such that $P^* T \subset T$ and $NT \cap T = (0, \cdots, 0)$. The one point compactification $T \cup \infty$ of a nontrivial closed positive cone T is a continuum and becomes a clan with exactly two idempotents, a zero and an identity, when addition is extended by defining $z + \infty = \infty + z = \infty$ for each $z \in T \cup \infty$, and such clans are uniquely divisible (where the "nth root" of z would be $(1/n)z$ since the operation is addition).

THEOREM. *Suppose that S is a commutative cancellative clan with $E = \{0, 1\}$, such that every element of S/H has a square root in S/H.*

If $\dim S \leqq N$ *or* $\dim S/H \leqq N$, *then there is a closed positive cone* T *in* E^n *and an onto homomorphism* $f: (T \cup \infty) \times H \to S$ *which is a homeomorphism of some neighborhood of the identity onto a neighborhood of the identity in* S. f *maps* $(T \cup \infty) \times 1$ *to a subclan* T' *which is a local cross section at* 1 *for the natural projection homomorphism* $S \to S/H$.

Proof. Let W, U and $\{R_1, \cdots, R_n\}$ be as in (ii) of the Lemma and let $x_i \in R_i \cap F(U)$ for each i. These x_i's will remain fixed throughout the proof, and since $x_i \neq 0, 1$, by (I) each element of $R_i \backslash 0$ equals x_i^t for a unique nonnegative real number t. This together with (ii) of the Lemma implies that for each $x \in U$, there are a unique partition (A, B) of $\{1, \cdots, n\}$, unique real numbers $t_1, \cdots, t_n$, and unique $h \in H$ such that $x(P_{i \in A}\{x_i^{t_i}\}) = (P_{i \in B}\{x_i^{t_i}\})h \in W$ and $i \in B$ if $t_i = 0$; following the notation of (1), let $\varepsilon_i = 1$ if $i \in B$ and $\varepsilon_i = -1$ if $i \in A$, let $\phi(x) = (\varepsilon_1 t_1, \cdots, \varepsilon_n t_n)$, and let $\beta(x) = h$. Arguments just like those in (1) show that $\phi \times \beta$ is a homeomorphism, if one uses at judicious spots the facts that W is compact and that $S \backslash W$ is an ideal. Since S is commutative, ϕ and β are homomorphisms as far as they go.

Let $T = P^* \phi(U)$. We show next that $\phi(U)$ contains a neighborhood of the origin in T and that T is a closed positive cone in E^n. First, $T = P^* \phi(F(U))$ because each closed one parameter semigroup in S intersects $F(U)$, so T is closed in E^n because in general if A is closed in P^* and S is compact in E^n and does not contain the origin, then AB is closed. For this same reason, $[1, \infty)\phi(F(U))$ is closed, hence its complement in T is a neighborhood of the origin in T and also is a subset of $\phi(U)$ because $k\phi(x) = \phi(x^k)$ and $x \in U$ implies $x^k \in U$, for $k \in [0, 1)$. Since $\phi(U)$ contains a neighborhood of the origin in T and ϕ preserves multiplication on U, T is a subsemigroup of E^n. To see that $NT \cap T$ is the origin it suffices to prove that $(-1)\phi(U) \cap \phi(U)$ is, so suppose $x, x' \in U$ and $\phi(x) = (-1)\phi(x') = (t_1, \cdots, t_n)$. Then for some $h, h' \in H$, $x(P_{i \in A}\{x_i^{t_i}\}) = (P_{i \in B}\{x_i^{t_i}\})h \in W$ and $x'(P_{i \in B}\{x_i^{t_i}\}) = (P_{i \in A}\{x_i^{t_i}\})h' \in W$. Substituting from the first equation into the second and cancelling gives $x'xh^{-1} = h'$, hence $x, x' \in H$, hence $\phi(x)$ is the origin as required.

Now define $\psi: \phi(U) \to S$ by $\psi(z) = (\phi \times \beta)^{-1}(z, 1)$. ψ is a homeomorphism into and, if U is chosen small enough that ϕ is actually defined on U^2, ψ preserves multiplication on $\phi(U)$ also. T is uniquely divisible so by (III), ψ may be extended to a homomorphism of T into S. Now define $f: (T \cup \infty) \times H \to S$ by $f(z, h) = \psi(z)h$. f is a homomorphism because ψ is and S is commutative, and it is a homeomorphism of $\phi(U) \times H$ onto U because there it equals $(\phi \times \beta)^{-1}$. (We cannot use (III) to define f directly as an extension of $(\phi \times \beta)^{-1}$, because H need not be uniquely divisible.) Since the image of f is a subclan of S which contains a neighborhood of H and since S is divisible, f is onto. Therefore $T'H = S$ so $T' \to S/H$ is onto and the rest is clear.

In a semigroup with zero, a *nilpotent* is a nonzero element some finite power of which is zero.

COROLLARY. *Let everything be as in the theorem.*

(i) *If square roots are unique in* $(S/H)\backslash 0$ (but there could be nilpotents) *then* f *is one-to-one on the complement of* $f^{-1}(0)$, *hence* f *induces an isomorphism from the Rees quotient* $((T \cup \infty) \times H)/f^{-1}(0)$ *onto* S *and also* T' *is a full cross section for* $H \times S \to S$. *If square roots are unique in all of* S/H (so there are no nilpotents) *then* $f^{-1}(0) = \infty \times H$, *so* S *is isomorphic to* $((T \cup \infty) \times H)/(\infty \times H)$ (Theorem 2.2 of (1)).

(ii) *Square roots exist* (uniquely) *in* S *if and only if they exist* (uniquely) *in* H *and* S/H.

Proof. Let $p: S \to S/H$ be the natural map. If $f(t, h) = f(s, g) \neq 0$, then $f(t, 1)h = f(s, 1)g$ hence $pf(t, 1) = pf(s, 1)$. Uniqueness of roots in $(S/H)\backslash 0$ implies $pf(kt, 1) = pf(ks, 1)$ for all $k \geqq 1$ at least, and pf is one-to-one near the identity by the theorem, hence $kt = ks$ must be true for k sufficiently small. Therefore $t = s$ and cancelling $f(t, 1)$ now gives $h = g$ also. The rest is clear.

EXAMPLE 1. This was also discovered by D. Brown and M. Friedberg (and communicated orally to this author). It is a cancellative commutative clan S with $E = \{0, 1\}$ and trivial group of units, which has no nilpotents and is divisible but not uniquely divisible; in fact, any two distinct one parameter semigroups in S are independent near 1 and have no nondegenerate arc in common, but can intersect infinitely. Thus S is not a Rees quotient of any compactified cone. The author is indebted to Kermit Sigmon for the elegance of this description of the example.

Let T be the closed first quadrant of E^2, let D be the closed unit disc in the complex plane with usual complex multiplication, and define $g: T \cup \infty \to D$ by $g(x, y) = e^{-(x+y)+(x-y)\pi i}$ and $g(\infty) = 0$. g is a homomorphism by (III), so $S = g(T \cup \infty)$ is a clan, it has $E = \{0, 1\}$, is topologically a 2-cell, and is an egg-shaped subset of D with large end at 1 and small end at $-1/e$. S is commutative, cancellative and free of nilpotents since D is, has roots of all orders since $T \cup \infty$ does, and square roots are not unique since $\phi(1, 0) = \phi(0, 1)$ but $\phi(1/2, 0) \neq \phi(0, 1/2)$.

S can also be visualized without the aid of D: there is a congruence $\sim$ on $T \cup \infty$ such that S is isomorphic to $(T \cup \infty)/\sim$: it is the smallest congruence which identifies (0, 1) and (1, 0), and dividing by it has the effect geometrically of rolling up $T \cup \infty$ into a cone with pointed end at ∞.

EXAMPLE 2. This will show that the subclan T' of the theorem need not be a full cross section for H orbits, i.e., $\mathscr{H}$ classes. Let $T \cup \infty$ be as in the previous example, let G be the circle group with usual complex number notation, and let Q be the product semigroup $(T \cup \infty) \times G$. We will twist the $\mathscr{H}$ class of $(0, 1, 1)$ and then identity it with the $\mathscr{H}$ class of $(1, 0, 1)$. Formally, let $\sim$ be the smallest closed congruence on Q which identifies $(0, 1, 1)$ and $(1, 0, -1)$, let $S = Q/\sim$, and let $f\colon Q \to S$ be the natural projection. Thus if Δ is the diagonal of $Q \times Q$, $p = [(0, 1, 1), (1, 0, -1)]$, and $q = [(1, 0, -1), (0, 1, 1)]$, then $\sim$ is the smallest closed symmetric subsemigroup of $Q \times Q$ containing $p \cup \Delta$, and $pq \in \Delta$ so this equals $\Delta(\Gamma(p) \cup \Gamma(q) \cup \Delta)$. Clearly $[(0, 1, 1), (1, 0, 1)]$ is not in the semigroup generated by $p \cup q \cup \Delta$, and $\Gamma(p)$ and $\Gamma(q)$ have only one limit point, ∞, so this point is not in $\sim$, i.e., $f(0, 1, 1) \neq f(1, 0, 1)$. On the other hand, the $\mathscr{H}$ classes in S of these points are equal, because $H = f(0 \times 0 \times G)$ is the group of units of S and $f(0, 1, 1) = f(1, 0, 1)f(0, 0, -1)$.

f is a homeomorphism on $[0, 1) \times [0, 1) \times G$, which is a neighborhood of the identity, and we will show below that S is cancellative, so this is exactly the situation of the theorem. However, if T' denotes $f((T \cup \infty) \times 1)$, $T' \to S/H$ is not one-to-one.

Interestingly, there actually is a full cross section semigroup for the H orbits of this clan S; the problem in the above lies in the definition of f—that is, in the choice of the independent closed one parameter semigroups in S:

$$R_1 = f([0, \infty] \times 0 \times 1) \quad \text{and} \quad R_2 = f(0 \times [0, \infty] \times 1)$$

are independent but do not themselves intersect in some of the H orbits which they both go through. Rechoosing f so that R_2 actually does intersect R_1 at the levels where $Q \to S$ collapses two H orbits to one yields a subclan T'' of S which is isomorphic to S/H. In detail, define $g\colon Q \to Q$ by $g(x, y, e^{i\theta}) = (x, y, e^{i(\theta + \pi y)})$, let $f' = fg$, and let $T'' = f'((T \cup \infty) \times 1)$. To see that $T'' \to S/H$ is one-to-one, suppose

$$fg(x, y, 1) = fg(x', y', 1)fg(0, 0, e^{i\theta}) \neq 0 \ .$$

We will prove $e^{i\theta} = 1$. In $g(x, y, 1) = g(x', y', e^{i\theta})$ then we are done because g is one-to-one, so suppose $g(x, y, 1) \neq g(x', y', e^{i\theta})$. f identifies these points and not to 0 so for some n, $((g(x, y, 1), g(x', y', e^{i\theta})) \in \Delta p^n$. An arbitrary point of Δp^n is of the form $((s, n + t, e^{i\phi}), (n + s, t, e^{i(\phi + n\pi)}))$ for some s, t and ϕ, so we conclude $x' = x + n$, $y = y' + n$, $e^{i\pi y} = e^{i\phi}$, and $e^{i(\theta+\pi y')} = e^{i(\phi+n\pi)}$. These imply $e^{i(\theta+\pi y')} = e^{i\pi y'}$, so $e^{i\theta} = 1$ as asserted. From this it follows at once that $T'' \to S/H$ is one-to-one and in fact that S is isomorphic to $(T'' \times H)/(\infty \times H)$.

Now it is easy to show S cancellative, for it suffices to prove that T'' is, so suppose $fg(x, y, 1)fg(s, t, 1) = fg(x', y', 1)fg(s, t, 1)$. It follows that $x + s + n = x' + s$ and $y + t = y' + t + n$ for some n, hence $x + n = x'$ and $y = y' + n$. $fg(x, y, 1) = fg(x', y', 1)$ now is clear.

It seems at least possible that the technique used here for rechoosing f might work in general, so that there is always a full cross section semigroup for $S \to S/H$ when S is a homomorph of the direct product of H and a closed positive cone.

It also seems reasonable to conjecture that the theorem is still true with only H normal and S/H commutative, instead of S commutative. Under these weaker conditions ϕ and β still exist, but β need not be a homomorphism unless the R_i's commute with one another and with H; using Theorem VI of (5), it is possible to choose a maximal independent set in the centralizer of H, but the problem of choosing the R_j's to commute with one another also remains unsolved.

REFERENCES

1. D. R. Brown and M. Friedberg, *Representation theorems for uniquely divisible semigroups*, Duke Math. J., **35** (1968), 341-352.

2. C. E. Clark, *Homomorphisms of compact semigroups*, Dissertation, Louisiana State University, 1966.

3. A. H. Clifford, *Connected ordered topological semigroups with idempotent endpoints, I*, T. A.M.S., **91** (1958), 80-98.

4. W. M. Faucett, *Compact semigroups irreducibly connected between two idempotents*, P.A.M.S., **6** (1955), 741-747.

5. K. H. Hofmann and P. S. Mostert, *Elements of Compact Semigroups*, C. E. Merrill, Columbus, 1966.

6. K. Keimel, *Eine Exponentialfunktion für kompakte abelsche Halbgruppen*, Math. Zeit., **96** (1967), 7-25.

7. P. S. Mostert and A. L. Shields, *On the structure of semigroups on a compact manifold with boundary*, Ann. Math., **65** (1957), 117-143.

8. A. B. Paalman-de Miranda, *Topological semigroups*, Mathematisch Centrum, Amsterdam, 1964.

9. A. D. Wallace, *Project mob*, Department of Mathematics, University of Florida, 1965.

Received August 2, 1969, and in revised form June 23, 1970. The author is grateful to the American Association of University Woman for a postdoctoral fellowship which supported part of this research.

COLLEGE OF NOTRE DAME

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

QUASI REGULAR GROUPS OF FINITE COMMUTATIVE NILPOTENT ALGEBRAS

N. H. EGGERT

Let J be a finite commutative nilpotent algebra over a field F of characteristic p. J forms an abelian group under the "circle" operation, defined by $a \circ b = a + b + ab$. This group is called the quasi regular group of J.

Our main purpose is to investigate the relationship between the structure of J as an algebra, and the structure of its quasi regular group.

In particular, the structure of the quasi regular group is described in terms of certain subalgebras of J. These subalgebras are, for fixed j, the p^j powers of elements in J. They are denoted by $J^{(j)}$.

It is conjectured that the dimension of $J^{(j)}$ is greater than or equal to p times the dimension of $J^{(j+1)}$. If this is true, then Theorems 1.1 and 2.1 completely describe the possibilities for the quasi regular group of J. Paragraph 2 considers some special cases of the conjecture.

1. The quasi regular group of J. Let J be a finite commutative nilpotent algebra over a field F with p^u elements. Denote by $J^{(j)}$ the set of p^jth powers of elements in J, $j = 0, 1, \cdots$. The $J^{(j)}$ form a descending chain of subalgebras of J. If t is the minimum exponent such that $x^{p^t} = 0$ for all $x \in J$ then $J^{(t-1)} \neq (0)$ and $J^{(t)} = (0)$. The constant t will be called the height of J. Let the dimension of $J^{(j)}$ be r_j and set $s_h = r_{h-1} + r_{h+1} - 2r_h$, $h = 1, \cdots, t$.

We denote by $G(p, u; s_1, \cdots, s_t)$ the group which is the direct sum of us_h, $h = 1, \cdots, t$, copies of the cyclic group of order p^h.

THEOREM 1.1. *The quasi regular group of J is isomorphic to $G(p, u; s_1, \cdots, s_t)$.*

Proof. Since the pth power of $x \in J$ with respect to the operation "o" is x^p, the number of cyclic summands of order greater than p^h is the dimension of the quotient group $J^{(h)}/J^{(h+1)}$ over the integers modulo p, that is $u(r_h - r_{h+1})$ [1, page 27]. Hence the number of cyclic summands of order p^h in the quasi regular group J is $u(r_{h-1} + r_{h+1} - 2r_h)$, $h = 1, \cdots, t$.

2. The possibilities for the quasi regular group of J. Given certain p-groups, finite commutative nilpotent algebras can be constructed with these groups as their quasi regular groups.

THEOREM 2.1. *Let a_i be arbitrary nonnegative integers for $i = 1, \cdots, t$, $a_t \neq 0$. Then there exists a finite commutative nilpotent algebra J over a field F of order p^u where:*

(i) $r_t = 0$ *and* $r_{i-1} = pr_i + a_i$, $i = 1, \cdots, t$.

(ii) *the quasi regular group of J is $G(p, u; s_1, \cdots, s_t)$ where $s_h = r_{h-1} + r_{h+1} - 2r_h$.*

Proof. Let J_j be the Jacobson radical of $F[X]/(X^n)$, where $n = p^{j-1} + 1$. If $x = X + (X^n)$ then a basis for J_j over F is $\{x, x^2, \cdots, x^{n-1}\}$. Thus the dimension of $J_j^{(i)}$ is p^{j-i-1} for $i < j$. Let J be the direct sum of a_j copies of J_j for $j = 1, \cdots, t$. Then $r_i = \dim J^{(i)} = \sum_{j=i+1}^{t} a_j p^{j-i-1}$, $i < t$, $r_t = \dim J^{(t)} = 0$. A simple calculation gives $r_{i-1} - pr_i = a_i$. By using Theorem 1.1, the proof is complete.

The author conjectures that the converse of the above theorem is also true, that is:

(C) If J is a finite commutative nilpotent algebra over F then $\dim J^{(i-1)} - p \dim J^{(i)} = r_{i-1} - pr_i \geqq 0$.

This is immediate for algebras of height one, height two and $\dim J^{(1)} = 1$, and height two and $p = 2$. The following theorem establishes (C) for algebras of height two and $\dim J^{(1)} = 2$.

THEOREM 2.2. *Let J be a commutative nilpotent algebra over a perfect field F of characteristic p. Let x, y be elements of J and suppose x^p and y^p are linearly independent over F. Then the dimension of J is greater than or equal to $2p$.*

Proof. Suppose the theorem is false. That is, assume there is a finite commutative nilpotent algebra J over F and:

(i) $x, y \in J$ and x^p, y^p are independent over F,

(ii) $\dim J < 2p$.

We assume J is an algebra of least dimension over F which satisfies (i) and (ii). It then follows that:

(iii) J is generated by x and y, and

(iv) If I is an ideal of J and an algebra over F then $I = (0)$ or for some $a, b \in F$, $0 \neq ax^p + by^p \in I$.

If (iv) were false then J/I would satisfy (i) and (ii) and the dimension of J/I would be less than the dimension of J.

We may assume x^p is in the annihilator of J. This follows since, by (iv), there are elements a, b in F where $ax^p + by^p \neq 0$ is in the annilhilator. By replacing x by $x' = a'x + b'y$, where $a'^p = a$ and $b'^p = b$, conditions (i) through (iv) hold and x'^p is in the annihilator.

Let $\mathscr{C}$ be the cartesian product of the nonnegative integers with themselves less $(0, 0)$. Let the total ordering $\prec$ be defined in $\mathscr{C}$ by: $(s, t) \prec (i, j)$ if $s + t < i + j$ or $s + t = i + j$ and $s < i$.

LEMMA. *If $x^i y^j \neq 0$ then $i + j \leqq p$.*

Proof. Let $(n, m(0))$ be the maximum element in $\mathscr{C}$, with respect to $\prec$, such that $x^n y^{m(0)} \neq 0$. Suppose that $n + m(0) > p$.

Since x^p is in the annihilator of J, $n \leqq p$ and $m(0) > 0$, thus if $n > 0$ then $\mathscr{A} = \{(i, j) \in \mathscr{C} : i \leqq n, \text{ and } j \leqq m(0)\}$ has more than $2p$ elements. The monomials $x^i y^j$, $(i, j) \in \mathscr{A}$, are dependent, thus a nontrivial relation.

$$\Sigma a_{ij} x^i y^j = z = 0, (i, j) \in \mathscr{A}$$

exists. Let (s, t) be minimum such that $a_{st} \neq 0$. Consider

$$0 = z x^{n-s} y^{m(0)-t} .$$

For $(s, t) \prec (i, j)$ it follows that $(n, m(0)) \prec (i + n - s, j + m(0) - t)$. By the definition of $(n, m(0))$ we obtain $0 = a_{st} x^n y^{m(0)}$. This is a contradiction; thus $n = 0$.

Now define $m(i)$ to be the maximum integer such that $x^i y^{m(i)} \neq 0$, $i = 0, \cdots, p$. Since $x, \cdots, x^p, y, \cdots, y^p$ are dependent, let

$$z = \sum_{i=h}^{p} a_i x^i + \sum_{i=l}^{p} b_i y^i = 0 , \tag{1}$$

where $a_h \neq 0$ and $b_l \neq 0$. There is at least one nonzero a_j since $y, \cdots, y^p$ are independent. Likewise at least one b_i is nonzero. Thus considering $x^{p-h} z$ and $y^{m(0)-l} z$ we find $x^{p-h} y^l \neq 0$ and $x^h y^{m(0)-l} \neq 0$.

We will now show that, for $k = 0, \cdots, h$, if $i \geqq k$ and $x^i y^j \neq 0$ then $(i, j) \leqq (k, m(k))$. Suppose this has been shown for $0, \cdots, k - 1$. Since $(i + 1, m(i + 1)) \prec (i, m(i))$ for $i < k$, we see that $m(0) \geqq m(i) + 2i$. From $x^h y^{m(0)-l} \neq 0$ and $h < k - 1$ we have

$$(h, m(0) - l) \prec (k - 1, m(k - 1)) .$$

Therefore $h + m(0) - l < k - 1 + m(k - 1)$ and $l - h \geqq k$. Now let (u, v) be maximum such that $u \geqq k$ and $x^u y^v \neq 0$. Since $x^{p-h} y^l \neq 0$ and $p - h \geqq l - h \geqq k$ it follows that $u + v \geqq p - h + l \geqq p + k$. If $v = 0$ then $u = p$ and $k = 0$. Since for $k = 0$ our result is established, we consider $v > 0$. If $u > k$ then the set $\mathscr{A} = \{(i, j) \in \mathscr{C} : k \leqq i \leqq u, 0 \leqq j \leqq v\}$ contains $(u - k + 1)(v + 1) \geqq 2(u - k + v) \geqq 2p$ elements. Thus there is a nontrivial relation among the $x^i y^j$, $(i, j) \in \mathscr{A}$. As before, let (s, t) be minimum such that the coefficient, a_{st}, of $x^r y^t$ is nonzero. On multiplying the relation by $x^{u-s} y^{v-t}$ we obtain $0 = a_{st} x^u y^v$ which is contradictory. Therefore $u = k$ and $v = m(k)$. By the definition of (u, v), if $i \geqq u = k$ and $x^i y^j \neq 0$ then $(i, j) \prec (k, m(k))$.

We now have the inequality, $m(0) \geqq 2k + m(k)$, for $k = 0, \cdots, h$. Since $x^h y^{m(0)-l} \neq 0$, $m(h) \geqq m(0) - l$. That is $l \geqq 2h$.

Let $bh + c = p$ where $0 \leqq c < h$. Returning to equation (1) we obtain:

$0 \neq a_h^b x^p = x^c(\Sigma_i a_i x^i)^b = x^c(-\Sigma_i b_i y^i)^b = x^c y^{bl} Y$, where Y is a polynomial in y.

Hence $x^c y^{bl} \neq 0$. This implies $m(0) - 2c \geqq m(c) \geqq bl \geqq 2bh$. Therefore $m(0) \geqq 2p$ and $y, \cdots, y^{2p}$ are independent. This is a contradiction and the lemma is established.

Next we show that if $m + n = p$ and $n \neq p$ then $x^m y^n = c_n x^p$ where $c_n \in F$. Suppose this holds for the powers of y being $0, \cdots, n-1$. If $x^m y^n = 0$ then the result is established. Thus suppose $x^m y^n \neq 0$. There are $(m+1)(n+1) \geqq 2p$ monomials of the form x^p or $x^i y^j$, $i \leqq m, j \leqq n$. Thus there is a nontrivial relation

$$\sum a_{ij} x^i y^j + a x^p = 0 .$$

Let (s, t) be minimum such that the coefficient of $x^s y^t$ is nonzero. By multiplying the relation by $x^{m-s} y^{n-t}$ we obtain:

$$\begin{aligned} 0 &= \sum_{i+j=s+t} a_{ij} x^{i+m-s} y^{j+n-t} + a x^{p+m-s} y^{n-t} \\ &= \sum_{\substack{i+j=s+t \\ (i,j)\neq(s,t)}} c_{j+n-t} a_{ij} x^p + a' x^p + a_{st} x^m y^n . \end{aligned}$$

Since x^p is in the annihilator of J, $x^{p+m-s} y^{n-t}$ is x^p or 0. Therefore $x^m y^n = c_n x^p$.

Similarly we obtain: if $m + n = p$ and $m \neq p$, then $x^m y^n = b_m y^p$. Since x^p and y^p are independent, if $m + n = p$, $m \neq 0, p$ then $x^m y^n = 0$.

From equation (1) we may obtain, as before, $x^{p-h} y^l \neq 0$ and $x^h y^{p-l} \neq 0$ where $0 < h, l \leqq p$. Assuming, without loss of generality, $h \geqq l$ we have $h + (p - l) \geqq p$ and by the lemma we have equality, that is, $h = l$. Since $x^h y^{p-h} \neq 0$ we have, by the above paragraph, $h = l = p$. Equation (1) becomes $0 = a_p x^p + b_p y^p$ for nonzero a_p and b_p, a contradiction. This completes the proof of Theorem 2.2.

Reference

1. I. Kaplansky, *Infinite Abelian Groups*, Ann Arbor 1954.

Received January 26, 1968. This research was in part supported by the National Science Foundation, grant GP-1923.

Montana State University

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

SOME NUMBER THEORETIC RESULTS

(In memory of our good friend Leo Moser)

P. ERDÖS AND E. G. STRAUS

The paper first establishes the order of magnitude of maximal sets, S, of residues (mod p) so that the sums of different numbers of elements are distinct.

In the second part irrationalities of Lambert Series of the form $\sum f(n)/a_1 \cdots a_n$ are obtained where $f(n) = d(n)$, $\sigma(n)$ or $\varphi(n)$ and the a_i are integers, $a_i \geqq 2$, which satisfy suitable growth conditions.

This note consists of two rather separate topics. In §1 we generalize a topic from combinatorial number theory to get an order of magnitude for the number of elements in a maximal set of residues (mod p) such that sums of different numbers of elements from this set are distinct. We show that the correct order is $cp^{1/3}$ although we are unable to establish the correct value for the constant c.

Section 2 consists of irrationality results on series of the form $\Sigma f(n)/a_1a_2\cdots a_n$ where $f(n)$ is one of the number theoretic functions $d(n)$, $\sigma(n)$ or $\varphi(n)$ and a_n are integers $\geqq 2$. For $f(n) = d(n)$ it suffices that the a_n are monotonic while for $\sigma(n)$ and $\varphi(n)$ we needed additional conditions on their rates of growth.

1. Maximal sets in a cyclic group of prime order for which subsets of different orders have different sums. In an earlier paper [4] one of us has given a partial answer to the question:

What is the maximal number $n = f(x)$ of integers $a_1, \cdots, a_n$ so that $0 < a_1 < a_2 < \cdots < a_n \leqq x$ and so that

$$a_{i_1} + \cdots + a_{i_s} = a_{j_1} + \cdots + a_{j_t} \text{ for some } \begin{array}{l} 1 \leqq i_1 < \cdots < i_s \leqq n \\ 1 \leqq j_1 < \cdots < j_t \leqq n \end{array}$$

implies $s = t$? it is conjectured that the maximal set is obtained (loosely speaking) by taking the top $2\sqrt{x}$ integers of the interval $(1, x)$. We were indeed able to prove that $f(x) < c\sqrt{x}$ for suitable c (for example $4/\sqrt{3}$) by using the fact that a set of n positive integers has a minimal set of distinct sums of t-tuples $(1 \leqq t \leqq n)$ if it is in arithmetic progression.

It is natural to pose the analogous question for elements of cyclic groups of prime order, as was done at the Number Theory Symposium in Stony Brook [5]. Here again we may conjecture that a maximal set of residues (mod p) is attained by taking a set of consecutive residues, this time not at the upper end but near $p^{2/3}$.

Conjecture 1.1. Let $f(p)$ be the maximal cardinality of a set of residues mod p so that sums of different numbers of residues in this set are different, then $f(p) = (4p)^{1/3} + o(p^{1/3})$ where the maximum is attained, for example, by taking consecutive residues in an interval of length $(4p)^{1/3} + o(p^{1/3})$ containing the residue $[(p/2)^{1/3}]$.

It is easy to see that we can indeed get a set of about $(4p)^{1/3}$ residues by taking the residues in the interval $([(p/2)^{2/3} - (4p)^{1/3}], [(p/2)^{2/3}])$. Here sums of distinct numbers of elements are distinct integers, and since all sums are $< p$ it follows that they are distinct residues.

The observation which let to the upper bound in [4] is much less obvious (mod p):

Conjecture 1.2. A set $A = \{a_1, a_2, \cdots, a_k\}$ of residues (mod p) has a minimal number of distinct sums of subsets of t elements if A is in arithmetic progression.

Conjecture 1.2 would give us a simple upper bound for $f(p)$:

COROLLARY 1.3. *If Conjecture* 1.2 *holds then*

$$f(p) < (6p)^{1/3} + o(p^{1/3}) .$$

Proof. The sums of t elements from the set of residues

$$\{1, 2, \cdots, k - 1, k\}$$

fill the interval $(\binom{t+1}{2}, tk - \binom{t}{2})$ that is to say there are $tk - t^2 + O(t)$ such sums. Since for different t we get different sums we must have

$$p \geqq \sum_{t=1}^{k} (tk - t^2 + O(t)) = \frac{k^3}{6} + O(k^2)$$

$$\text{and hence } k < (6p)^{1/3} + o(p^{1/3}) .$$

Using methods employed by Erdös and Heilbronn [2] we can show that $f(p) = O(p^{1/3})$. We use the following lemma from [2].

LEMMA 1.4. *Let* $1 < m \leqq l < p/2$ *and let* $B = \{b_1, \cdots, b_l\}$, $A = \{a_1, \cdots, a_m\}$ *be sets of residues* (mod p). *Then there exists an* $a_i \in A$ *such that the number of solutions of* $a_i = b_j - b_k$; $b_j, b_k \in B$ *is less than* $l - m/6$.

We now can get a lower bound for the number of distinct sums of t elements from a set of residues.

LEMMA 1.5. *Let* $A = \{a_1, \cdots, a_k\}$ *be a set of residues* (mod p)

and let $A_t = \{a_{i_1} + \cdots + a_{i_t} | 1 \leqq i_1 < \cdots < i_t \leqq k\}$ *then for* $1 \leqq t \leqq k/4$ *we have*

$$|A_t| \geqq l + \frac{(t-1)m}{6} - \frac{t(t-1)}{6} \tag{1.6}$$

where

$$l = \left[\frac{k+1}{2}\right], \; m = \left[\frac{k}{2}\right].$$

Proof. We divide the set A into two disjoint sets

$$A = \{a_1, a_2, \cdots, a_l\}, \; B = \{b_1, b_2, \cdots, b_m\}$$

and prove the inequality (1.6) for the subset of A_t consisting of the sums

$$A_t^* = \{a_i + b_{2-\varepsilon_1} + b_{4-\varepsilon_2} + \cdots + b_{2t-2-\varepsilon_{t-1}} | \varepsilon_j = 0 \text{ or } 1\},$$

where the b_i are a suitable ordering of the elements of B.

The inequality holds for $t = 1$ since

$$A_t^* = \{a_i\} = A \text{ and } |A| = l.$$

Now assume that (1.6) holds for A_t^* with $t \leqq (m/2) - 1$. Then the set $A_t^* + b_{2t} \subset A^*_{t+1}$ and according to Lemma 1.3 there exists a $b_j \in \{b_{2t+1}, b_{2t+1}, \cdots, b_m\}$, say $b_j = b_{2t+1}$ so that the equation

$$b_{2t+1} - b_{2t} = a_i^* - a_j^*, \quad a_i^*, a_j^* \in A_t^*$$

has no more than $|A_t^*| - \frac{1}{6}(m - 2t)$ solutions. Hence the set

$$((b_{2t+1} - b_{2t}) + (A_t^* + b_{2t})) \cap (A_t^* + b_{2t})$$

contains no more than $A_t^* - \frac{1}{6}(m - 2t)$ elements and

$$\begin{aligned} |A_{t+1}^*| &= |(A_t^* + b_{t+1}) \cup (A_t^* + b_t)| \\ &\geqq |A_t^*| + \tfrac{1}{6}(m - 2t) \\ &\geqq l + \frac{(t-1)m}{6} - \frac{t(t-1)}{6} + \frac{1}{6}m - \frac{t}{3} \\ &= l + \frac{tm}{6} - \frac{(t+1)t}{6}. \end{aligned}$$

This completes the proof.

THEOREM 1.7. *The maximal number* $f(p)$ *of a set* A *of residues* (mod p) *so that sums of different numbers of distinct elements of* A *are distinct satisfies*

$$(4p)^{1/3} + o(p^{1/3}) < f(p) < (288p)^{1/3} + o(p^{1/3}). \tag{1.8}$$

Proof. According to Lemma 1.5 there are at least

$$k/2 + k(t-1)/12 - t^2/6 + O(t)$$

distinct sums of t elements (and hence, by symmetry, sums of $k-t$ elements) for $t < [k/4]$ out of a set A with k elements. Thus if A has the desired property we must have

$$p \geqq 2\sum_{t=1}^{k/4}(k/2 + k(t-1)/12 - t^2/6) + O(k^2)$$
$$= 2k^3\Big(\frac{1}{384} - \frac{1}{3}\frac{1}{384}\Big) + O(k^2) = k^3/288 + O(k^2)\ .$$

Thus

$$f(p) < (288\, p)^{1/3} + o(p^{1/3})\ .$$

The lower bound for $f(p)$ was established above.

2. On some irrational series.

One of us [1] proved that the series $\sum_{n=1}^{\infty} d(n)t^{-n}$ is irrational for every integer t, $|t| > 1$. In this section we generalize this result to series of the form

$$\xi = \sum_{n=1}^{\infty}\frac{d(n)}{a_1a_2\cdots a_n} \tag{2.1}$$

where the a_n are positive integers with $2 \leqq a_1 \leqq a_2 \leqq \cdots$. It is clear that we need some restriction, such as monotonicity, on the a_n since the choice $a_n = d(n)+1$ would lead to $\xi = 1$.

We divide the proof into two cases depending on the rate of increase of a_n. The first case is very similar to [1].

LEMMA 2.2. *The series (2.1) is irrational if there exists a $\delta > 0$ so that the inequality $a_n < (\log n)^{1-\delta}$ holds for infinitely many values of n.*

Proof. Let n be a large integer so that $a_n < (\log n)^{1-\delta}$. Then by the monotonicity of a_i there exists an interval I of length $n/\log n$ in $(1, n)$ so that for all integers $i \in I$ we have $a_i = t$ where t is a fixed integer, $t \leqq (\log n)^{1-\delta}$.

Now put $k = [(\log n)^{\delta/10}]$ and let $p_1, p_2, \cdots$ be the consecutive primes greater than $(\log n)^2$. Let

$$A = \Big(\prod_{1\leqq i\leqq k(k+1)/2} p_i\Big)^t$$

then

$$A < (2(\log n)^2)^{t\,k(k+1)/2} < e^{(\log n)^{1-\delta}(\log n)^{\delta/4}}$$
$$< e^{(\log n)^{1-\delta/2}}\ . \tag{2.3}$$

By the Chinese remainder theorem the congruences

$$
\begin{aligned}
x &\equiv p_1{}^{t-1} (\operatorname{mod} p_1{}^t) \\
x + 1 &\equiv (p_2 p_3)^{t-1} (\operatorname{mod} (p_2 p_3)^t) \\
&\vdots \\
x + k - 1 &\equiv (p_u p_{u+1} \cdots p_{u+k-1})^{t-1} (\operatorname{mod} (p_u p_{u+1} \cdots p_{u+k-1})^t)
\end{aligned} \tag{2.4}
$$

where $u = 1 + k(k-1)/2$, have solutions determined (mod A). The interval I contains at least $[n/(A \log n)]$ solutions of (2.4).

Now assume that $\xi = a/b$ and choose $x \in I$ to be a solution of (2.4) so that $(x, x+k) \subset I$. Then

$$
\begin{aligned}
b\, a_1 \cdots a_{x-1} \xi &= \text{integer} + b \sum_{l=0}^{k-1} \frac{d(x+l)}{t^{l+1}} \\
&\quad + b \sum_{s=0}^{\infty} \frac{d(x+k+s)}{t^k a_{x+k} \cdots a_{x+k+s}} .
\end{aligned} \tag{2.5}
$$

But (2.4) implies that $d(x+l) \equiv 0 \,(\operatorname{mod} t^{l+1})$ for $l = 0, 1, \cdots, k-1$. Thus (2.5) implies that

$$
b\, a_1 \cdots a_{x-1} \xi = \text{integer} + \frac{b}{t^k} \sum_{s=0}^{\infty} \frac{d(x+k+s)}{a_{x+k} \cdots a_{x+k+s}} . \tag{2.6}
$$

We now wish to show that for suitable choice of x the sum on the right side of (2.6) is less than 1 and hence $b\xi$ cannot be an integer. We first consider the sum

$$
\begin{aligned}
&\frac{b}{t^k} \sum_{s > 10 \log n} \frac{d(x+k+s)}{a_{x+k} \cdots a_{x+k+s}} \\
&< \frac{b}{t^k} \sum_{s > 10 \log n} \frac{x+k+s}{t^{s+1}} < b(x+k) \sum_{s > 10 \log n} \frac{s}{t^s} \\
&< \frac{2bn}{n^2} < \frac{1}{2} \text{ for large } n .
\end{aligned} \tag{2.7}
$$

Next we wish to show that it is possible to choose x so that

$$
d(x+k+s) < 2^{k/4} \text{ for } 0 \leqq s < 10 \log n . \tag{2.8}
$$

We first observe that

$$
(x+k+s, A) = 1 \text{ for all } 0 \leqq s < 10 \log n \tag{2.9}
$$

since otherwise

$$
x+k+s \equiv 0 \,(\operatorname{mod} p_j) \text{ for some } 1 \leqq j \leqq k(k+1)/2 \tag{2.10}
$$

and

$$
x+i \equiv 0 \,(\operatorname{mod} p_j) \text{ for some } 0 \leqq i < k . \tag{2.11}
$$

But

$$0 < k + s - i < 11 \log n < (\log n)^2 < p_j$$

so that (2.10) and (2.11) are incompatible.

Let $x = x_0, x_0 + A, \cdots, x_0 + zA$ be the solutions of (2.4) for which $(x, x + k) \subset I$. From (2.9) we get

$$\sum_{y=0}^{z} d(x_0 + k + s + yA) < 2 \sum_{l=1}^{\sqrt{n}} \left(\frac{n}{Al} + 1\right) < c\,\frac{n \log n}{A} . \tag{2.12}$$

Thus the number of y's for which $d(x_0 + k + s + yA) > 2^{k/4}$ is less than $c\,n \log n/(A.2^{k/4})$, and the number of y's so that for some $0 \leqq s < 10 \log n$ we have $d(x_0 + k + s + yA) > 2^{k/4}$ is less than

$$10c\, n \log^2 n/(A.2^{k/4}) < 1/2\, n/(A \log n) < z .$$

It is therefore possible to choose $x = x_0 + yA \in I$ so that (2.8) holds. For such a choice we get

$$\frac{b}{t^k} \sum_{s=0}^{10 \log n} \frac{d(x + k + s)}{a_{x+k} \cdots a_{x+k+s}} < \frac{b}{t^k}\, 2^{k/4} \sum_{s=0}^{\infty} \frac{1}{t^s} < b.2^{-3k/4} < \frac{1}{2} . \tag{2.13}$$

Combining (2.7) and (2.13) we see that ξ is irrational.

LEMMA 2.14. *If there exists a positive constant c so that $|a_n| > c(\log n)^{3/4}$ for all n then the series* (2.1) *is irrational.*

Note that in this lemma we need not assume the monotonicity of a_n (or even that they are positive, however for simplicity we give the proof for positive a_n only).

Proof. We use two results. The Dirichlet divisor theorem

$$\sum_{n=1}^{N} d(n) \sim N \log N \tag{2.15}$$

and the average order of $d(n)$, [3]

$$d(n) < (\log n)^{\log 2 + \varepsilon} \text{ for almost all } n . \tag{2.16}$$

From (2.15) we get the following.

LEMMA 2.17. *Given constants $b, c > 0$, then for almost all integers x*

(2.18) $$d(x+y) < b^{-1}(2c)^{-y}(\log x)^{3y/4};\ y = 3, 4, \cdots$$

Proof. If we choose x large enough so that $\log x > (2bce)^{4/3}$ then the right side of (2.18) is greater than e^y which exceeds $x+y$, and hence $d(x+y)$, whenever $y > 2\log x$. Thus, if (2.18) fails to hold for sufficiently large x then it must fail to hold for some y with $3 \leqq y \leqq 2\log x$.

Now if there are $c_1 N$ integers x below N so that (2.18) fails to hold then we have more than $c_2 N$ integers x with $\sqrt{N} \leqq x \leqq N - 2\log N$ and

(2.19) $$\begin{aligned} d(x+y) &> b^{-1}(2c)^{-y}(\log x)^{3y/4} \geqq b^{-1}(2c)^{-y}(\tfrac{1}{2}\log N)^{3y/4} \\ &\geqq b^{-1}(4c)^{-3}(\log N)^{9/4} = c_3(\log N)^{9/4} . \end{aligned}$$

Thus

$$\begin{aligned} \sum_{n=1}^{N} d(n) &\geqq c_2 N \cdot \frac{1}{2\log N} c_3 (\log N)^{9/4} \\ &= c_4 N(\log N)^{5/4} \end{aligned}$$

which contradicts (2.15) for large N.

Combining Lemma 2.17 with (2.16) we find that there exists an infinite set S of integers x so that

(2.21) $$d(x+1) < \frac{b^{-1}c}{2}(\log x)^{3/4},\ d(x+2) < \frac{b^{-1}c^2}{4}(\log x)^{3/4}$$

and (2.18) both hold.

Now assume that $\xi = a/b$ is a rational value of (2.1) and choose $n \in S$. Then

(2.22) $$a_1 \cdots a_n b\xi = \text{integer} + b\sum_{y=1}^{\infty} \frac{d(n+y)}{a_{n+1}\cdots a_{n+y}}$$

where

$$0 < \sum_{y=1}^{\infty} \frac{d(n+y)}{a_{n+1}\cdots a_{n+y}} < \sum_{y=1}^{\infty} \frac{(2c)^{-y}(\log n)^{3y/4}}{(c(\log n)^{3/4})^y} = 1 ,$$

in contradiction to the fact that the left side of (2.22) is an integer.

Summing up we have

THEOREM 2.23. *The series* (2.1) *is irrational whenever*

$$2 \leqq a_1 \leqq a_2 \leqq \cdots \leqq a_n \leqq \cdots .$$

With considerable additional effort one can weaken the monotonicity condition on the a_n to $a_m/a_n \geqq c > 0$ for all $m > n$.

We have not been able to prove the following

Conjecture 2.24. The series (2.1) is irrational whenever $a_n \to \infty$.

If we consider series of the form

$$\sum_{n=1}^{\infty} \frac{\varphi(n)}{a_1 \cdots a_n} \quad \text{or} \quad \sum_{n=1}^{\infty} \frac{\sigma(n)}{a_1 \cdots a_n} \tag{2.25}$$

then we cannot make conjectures analogous to 2.24 since the choice $a_n = \varphi(n) + 1$ or $\sigma(n) + 1$ would make these series converge to 1. It is reasonable to conjecture that the series (2.25) must be irrational if the a_n increase monotonically, however we can prove this only under more restrictive conditions.

THEOREM 2.26. *If $\{a_n\}$ is a monotonic sequence of integers with $a_n \geqq n^{11/12}$ for all large n then the series in* (2.25) *are irrational.*

For the proof we need the following simple lemmas.

LEMMA 2.27. *Let $\{a_n\}$ be a sequence of positive integers with $a_n \geqq 2$ and $\{b_n\}$ a sequence of positive integers so that $b_{n+1} = o(a_n a_{n+1})$. If*

$$\xi = \sum_{n=1}^{\infty} \frac{b_n}{a_1 \cdots a_n} \tag{2.28}$$

is rational then $a_n = O(b_n)$.

Proof. Assume $\xi = a/b$ and choose N so that for all $n > N$ we have $bb_n < a_{n-1}a_n/4$. If there existed an $n > N$ so that $a_n > 2bb_n$ then we would have

$$ba_1 \cdots a_{n-1}\xi = aa_1 \cdots a_{n-1} = \text{integer} + \sum_{k=0}^{\infty} \frac{bb_{n+k}}{a_n \cdots a_{n+k}}$$

but

$$0 < \sum_{k=0}^{\infty} \frac{bb_{n+k}}{a_n \cdots a_{n+k}} = \frac{bb_n}{a_n} + \sum_{k=1}^{\infty} \frac{bb_{n+k}}{a_{n+k-1} \cdots a_{n+k}} \cdot \frac{1}{a_n \cdots a_{n+k-2}}$$
$$< \tfrac{1}{2} + \tfrac{1}{4} \sum_{l=0}^{\infty} (\tfrac{1}{2})^l = 1\,,$$

a contradiction. Thus $a_n \leqq 2bb_n$ for all large n.

LEMMA 2.29. *If the series* (2.28) *is rational, say $\xi = a/b$, and $b_{n+1} = o(a_n a_{n+1})$, then there exists a sequence of positive integers $\{c_n\}$ so that for all large n we have*

$$bb_n = c_n a_n - c_{n+1}\,, \quad 0 < c_{n+1} < a_n, \text{ and } c_{n+1} = o(a_n)\,. \tag{2.30}$$

Conversely, if these conditions hold then the series (2.28) is rational.

Proof. Choose N so that for all $n > N$ we have $bb_n < a_n a_{n+1}/4$. Now for $n \geqq N$ choose c_n, c_{n+1} so that

$$bb_n = c_n a_n - c_{n+1}\,, \qquad c_n > 0, \quad 0 < c_{n+1} < a_n$$

and $c'_{n-1}\,, c'_{n+2}$

$$bb_{n+1} = c'_{n+1} a_{n+1} - c'_{n+2}\,, \qquad c'_{n+1} > 0, \quad 0 < c'_{n+2} < a_{n+1}\,.$$

Then

$$\begin{aligned}
ba_1 \cdots a_{n-1}\,\xi &= aa_1 \cdots a_{n-1} \\
&= \text{integer} + \frac{bb_n}{a_n} + \frac{bb_{n+1}}{a_n a_{n+1}} + \sum_{k=2}^{\infty} \frac{bb_{n+k}}{a_n \cdots a_{n+k}} \\
&= \text{integer} - \frac{c_{n+1}}{a_n} + \frac{c'_{n+1}}{a_n} - \frac{c'_{n+2}}{a_n a_{n+1}} \\
&\quad + \frac{1}{a_n} \sum_{k=2}^{\infty} \frac{bb_{n+k}}{a_{n+1} \cdots a_{n+k}} \qquad (2.31) \\
&= \text{integer} - \frac{c_{n+1}}{a_n} + \frac{c'_{n+1}}{a_n} - \frac{c'_{n+2}}{a_n a_{n+1}} + \frac{\theta}{a_n}\,, \qquad 0 < \theta < \tfrac{1}{2}\,.
\end{aligned}$$

Thus

$$\frac{1}{a_n}\left(- c_{n+1} + c'_{n+1} - \frac{c'_{n+2}}{a_{n+1}} + \theta\right) = \text{integer}$$

and since $0 < c_{n+1} < a_n$, $0 < c'_{n+1} \leqq [a_n/4] + 1$ $0 < c'_{n+2}/a_{n+1} < 1$, $0 < \theta < \frac{1}{2}$, this is possible only if $c_{n+1} = c'_{n+1}$.

Now choose N so large that $bb_{n+1} < \varepsilon a_n\, a_{n+1}$ for all $n > N$, then from (2.31) we have

$$\begin{aligned}
\text{integer} &= -\frac{c_{n+1}}{a_n} + \sum_{k=1}^{\infty} \frac{bb_{n+k}}{a_n a_{n+1} \cdots a_{n+k}} < -\frac{c_{n+1}}{a_n} + \varepsilon \sum_{k=1}^{\infty} \frac{1}{a_n \cdots a_{n+k-2}} \\
&\leqq -\frac{c_{n+1}}{a_n} + 2\varepsilon\,.
\end{aligned}$$

Thus $c_{n+1} < 2\varepsilon a_n$ for all $n > N$.

If condition (2.30) holds for all $n \geqq N$ then

$$\begin{aligned}
\sum_{n=N}^{\infty} \frac{bb_n}{a_1 \cdots a_n} &= \sum_{n=N}^{\infty} \frac{c_n a_n - c_{n+1}}{a_1 \cdots a_n} \\
&= \frac{c_N}{a_1 \cdots a_{N-1}} - \sum_{n=N}^{\infty} c_{n+1}\left(\frac{1}{a_1 \cdots a_n} - \frac{a_{n+1}}{a_1 \cdots a_{n+1}}\right) \\
&= \frac{c_N}{a_1 \cdots a_{N-1}}
\end{aligned}$$

is clearly rational.

Finally we need a fact from sieve theory. We are grateful to R. Miech for supplying the correct constants.

LEMMA 2.32. *Given an integer a and $\varepsilon > 0$ then for large y the number of integers m satisfying*

$$m \not\equiv 0,\ m \not\equiv a \pmod{p}$$

for all primes p, with $2 < p < y^{1/5}$ exceeds $y^{1-\varepsilon}$.

Proof of Theorem 2.26. Let $f(n)$ stand for either $\sigma(n)$ or $\varphi(n)$ and assume that

$$\sum_{n=1}^{\infty} \frac{f(n)}{a_1 \cdots a_n} = \frac{a}{b} .$$

Since $a_n > n^{11/12}$ for large n the hypothesis of Lemma 2.29 is satisfied and we get

$$bf(n) = c_n a_n - c_{n+1} \text{ for large } n . \tag{2.33}$$

Since $f(n) = o(n^{1+\varepsilon})$ for all $\varepsilon > 0$ we get

$$c_n < n^{1/12+\varepsilon} \text{ for large } n . \tag{2.34}$$

From Lemma 2.28 we get

$$a_n = O(f(n)) = O(n^{1+\varepsilon}) \tag{2.35}$$

and hence the number of integers $n \leqq x$ for which

$$\frac{a_{n+1}}{a_n} > 1 + x^{-1/2}$$

is $O(x^{3/4)})$, since otherwise we would have

$$a_x = \prod_{n<x} \frac{a_{n+1}}{a_n} > (1 + x^{-1/2})^{x^{3/4}} > x^2$$

for large x, in contradiction to (2.35). From now on we restrict our attention to integers n for which

$$\frac{a_{n+1}}{a_n} < 1 + n^{-1/2} . \tag{2.36}$$

For such integers we get from (2.33) and (2.35) that

$$\frac{f(n+1)}{f(n)} = \frac{c_{n+1}\,a_{n+1}}{c_n\,a_n}\Big(1 - \frac{c_{n+2}}{c_{n+1}\,a_{n+1}}\Big)\Big/\Big(1 - \frac{c_{n+1}}{c_n\,a_n}\Big)$$

$$(2.37) \qquad = \frac{c_{n+1}}{c_n}(1 + O(n^{-1/2}))\,(1 + O(n^{-3/4+\varepsilon}))$$

$$= \frac{c_{n+1}}{c_n} + O(n^{-1/2+\varepsilon})$$

Now consider a prime q, $\frac{1}{2}x^{1/11} \leqq q \leqq x^{1/11}$, then according to Lemma 2.32 there exist more than $y^{1-\varepsilon}$ integers $m \leqq y = x^{10/11}$ so that

$$(2.38) \qquad m \not\equiv 0,\ m \not\equiv -2q \pmod{p}$$

for all primes p with $2 < p < y^{1/5}$. We may even assume that m is odd. The number of integers $n = 2qm$ where m satisfies (2.38) exceeds $x^{10/11-\varepsilon} > x^{3/4}$ and hence we can pick such an n that satisfies (2.37) with $x/2 \leqq n \leqq x$.

Now

$$f(n) = f(2q)f(m)$$

where

$$\frac{f(2q)}{2q} = \begin{cases} \dfrac{3(q+1)}{2q} & \text{if } f = \sigma \\[2mm] \dfrac{q-1}{2q} & \text{if } f = \varphi \end{cases}$$

in either case

$$(2.39) \qquad f(2q) = A/q, \quad A \text{ an integer not divisible by } q.$$

Since m has at most 5 prime factors all exceeding $y^{1/5}$ we have

$$(1 - y^{-1/5})^5 < \frac{f(m)}{m} < (1 + y^{-1/5})^5$$

$$(2.40) \qquad f(m) = m(1 + O(y^{-1/5})) = m\,(1 + O(x^{-2/11}))\,.$$

By the same reasoning we get

$$(2.41) \qquad f(n+1) = n(1 + O(x^{-2/11}))\,.$$

Substituting (2.39), (2.40) and (2.41) in (2.37) we get

$$(2.42) \qquad \frac{f(n+1)}{f(n)} = \frac{A}{q}\,(1 + O(x^{-2/11})) = \frac{c_{n+1}}{c_n} + O(x^{-1/2+\varepsilon})\,.$$

But since $q > x^{1/12}$ and $c_n < x^{1/12}$ we get

$$(2.43) \qquad \frac{1}{qc_n} \leqq \left|\frac{A}{q} - \frac{c_{n+1}}{c_n}\right| < x^{-2/11+\varepsilon}\,.$$

Since $qc_n < x^{1/11+1/12} < x^{2/11-\varepsilon}$ this leads to a contradiction.

We could get similar irrationality results if the functions $\sigma(n)$ or $\varphi(n)$ are replaced by $\sigma_k(n) (k \geqq 1)$ or products of powers of $\sigma_k(n)$ and $\varphi(n)$. In each case we would need the assumption that the a_n are monotonic, increasing faster than a certain fractional power of the numerators.

From Lemma 2.29 it is clear that there is a set of power $2^{\aleph_0}$ of series (2.25) which are rational even if we restrict the integers c_n to the values 1 or 2 since for $c_n = 1$ we can choose $a_n = \sigma(n) - 1$ or $\sigma(n) - 2$ to get $c_{n+1} = 1$ or 2 respectively and for $c_n = 2$ we choose $a_n = [(\sigma(n)-1)/2]$ to get $c_{n+1} = 1$ if $\sigma(n)$ is odd and $c_{n+1} = 2$ if $\sigma(n)$ is even. For the series with numerators $\varphi(n)$ we would have to use $c_n = 1, 2$ or 3 since all $\varphi(n)$ are even for $n > 2$.

References

1. P. Erdös, *On arithmetical properties of Lambert series*, J. Indian Math. Soc., **12** (1948), 63-66.

2. P. Erdös and H. Heilbronn, *On the addition of residue classes mod p*, Acta Arithmetica **9** (1964), 149-159.

3. M. Kac, *Note on the distribution of values of the arithmetic function d(m)*, Bull. Amer. Math. Soc., **47** (1941), 815-817.

4. E. G. Straus, *On a problem in conbinatorial number theory*, J. Math. Sci., (Delhi), **1** (1966), 77-80.

5. ———, *Some problems concerning sum-free and average-free sets*, Lecture notes, Summer institute on Number Theory, Stony Brook, N.Y. 1969, Cll.

Received may 27, 1970. This work was supported under grant No. GP-13164

UNIVERSITY OF CALIFORNIA, LOS ANGELES

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

MONOTONE DECOMPOSITIONS OF IRREDUCIBLE HAUSDORFF CONTINUA

G. R. GORDH, JR.

It is shown that a number of important results concerning irreducible metric continua can be generalized to (nonmetric) irreducible continua. For example, if M is a (nonmetric) continuum which is irreducible between a pair of points and which contains no indecomposable subcontinuum with interior, then there exists a monotone continuous map of M onto a generalized arc, such that each point inverse has void interior. This result is applied to a study of hereditarily unicoherent, hereditarily decomposable continua. Certain properties of trees follow as corollaries. Also, trees are characterized as inverse limits of monotone inverse systems of dendrites.

In recent years there has been a growing interest in the study of (nonmetric) continua. It is well known (e.g., [6]) that some of the most useful and important properties of metric continua do not hold for (nonmetric) continua. It is the purpose of this paper to indicate that a substantial number of theorems concerning irreducible metric continua can be generalized to irreducible continua. These results are then applied to a study of certain hereditarily unicoherent continua.

In particular, § 2 contains generalizations of many of the results about irreducible metric continua appearing in Chapter 1 of [11]. These results are applied in § 3 to obtain generalizations of a number of theorems due to Miller [8] concerning hereditarily unicoherent continua. Section 4 contains several results about trees which follow as corollaries of theorems in § 3. Also, it is proved that every tree can be written as a monotone inverse limit of dendrites. In Chapter 2 of [11], Thomas discusses metric continua which are hereditarily of type A'. His definition is extended, in § 5, to (nonmetric) continua and several characterizations of such continua are obtained.

The reader is referred to [3], [5], and [14] for general results concerning continua (i.e., compact, connected Hausdorff spaces). It will be necessary to refer to results which are stated in the literature for metric continua; however, this will be done only when the proof for continua is essentially the same as that for metric continua.

The author is indebted to Professor F. Burton Jones for his advice and encouragement in the preparation of this paper.

2. Continua of type A. We observe that Theorem 1 and Theorem 7 of [11, Chapter 1] are true, as stated, for (non-metric) continua. To prove Theorem 1, apply [9, Theorem 47, page 16] to the proof as given in [11].

Let M be a continuum which is irreducible between a pair of points x and y. A decomposition $\mathscr{D}$ of M is said to be *admissible* in case each element of $\mathscr{D}$ is a nonvoid proper subcontinuum of M, and each element of $\mathscr{D}$ which does not contain x or y separates M. Notice that an admissible decomposition is not required by definition to be upper semi-continuous. However, we will show that an admissible decomposition must, in fact, be upper semi-continuous. Thus, for metric continua, our definition is equivalent to the definition in [11].

A *generalized arc* is a continuum A with precisely two non-separating points. It is well known that A can be totally ordered in such a way that the order topology and the original topology coincide. We will frequently denote A by $[a, b]$ where a and b are the non-separating points of A.

THEOREM 2.1. *Let M denote a continuum. Let $\mathscr{D} = \{D(x)\}$ be a decomposition of M such that* (1) *for each $x \in M$, $D(x)$ is a proper subcontinuum of M, and* (2) *there exist elements $D(a)$ and $D(b)$ of $\mathscr{D}$ such that every element $D(x)$ of $\mathscr{D}$ distinct from $D(a)$ and $D(b)$ separates $D(a)$ from $D(b)$. Then $\mathscr{D}$ is an upper semi-continuous decomposition, and $M/\mathscr{D}$ is a generalized arc.*

Proof. For each x in $M - [D(a) + D(b)]$, $M - D(x) = A_x + B_x$ where $a \in A_x$, $b \in B_x$, and A_x and B_x are connected. If x and y are in $M - [D(a) + D(b)]$ and $D(x) \neq D(y)$, then $D(y) \subset A_x$ if and only if $A_y \subset A_x$; also $D(y) \subset B_x$ if and only if $B_y \subset B_x$. Define $D(x) < (D(y)$ whenever $A_x \subset A_y$, and let $D(a) < D(z) < D(b)$ for all z in

$$M - [D(a) + D(b)] .$$

Then $<$ is a total order on $\mathscr{D}$. If $f: M \rightarrow \mathscr{D}$ denotes the natural map, then it is readily seen that f is continuous with respect to the order topology on $\mathscr{D}$. The conclusion of the theorem now follows.

COROLLARY 2.1. *Let M be a continuum which is irreducible from x to y. If $\mathscr{D}$ is an admissible decomposition for M, then $\mathscr{D}$ is upper semi-continuous and $M/\mathscr{D}$ is a generalized arc.*

A continuum M is of *type A* provided that it is irreducible between a pair of points and has an admissible decomposition; M is of *type A'* if it is of type A and has an admissible decomposition each of whose elements has void interior.

THEOREM 2.2. *Let M be a continuum irreducible from x to y. If M has an admissible decomposition, then it has one which is minimal* (with respect to partial order by refinement).

Proof. See the proof of [11, Theorem 3, page 8]. Notice that we are not required to prove the upper semi-continuity of the decomposition.

Suppose that M is a continuum irreducible between two points. If M is of type A, let Δ denote the collection of all admissible decompositions of M. For each $\mathscr{D} \in \Delta$, let $f: M \rightarrow M/\mathscr{D}$ denote the natural map. Thus f is a continuous monotone function from M onto a generalized arc. Observe that every monotone map from M onto a generalized arc is obtained in this manner.

THEOREM 2.3. *Let M be a continuum of type A, $\mathscr{D} \in \Delta$, and $f: M \rightarrow M/\mathscr{D}$. Suppose that K is a subcontinuum of M such that $f(K) = [r, s]$ where $[r, s]$ is a nondegenerate subinterval of $M/\mathscr{D}$. Then $f^{-1}(r) \cap K$ and $f^{-1}(s) \cap K$ are continua, and for $r < t < s$, $f^{-1}(t)$ is contained in and separates K. In particular $f|_K$ is a monotone map of K onto $[r, s]$; thus, if K is irreducible, K is of type A.*

Proof. Suppose that M is irreducible from x to y and $M/\mathscr{D} = [a, b]$. If $r < t < s$, then $f^{-1}(t) \subset K$; for if p is in $f^{-1}(t) - K$ then $f^{-1}([a, r]) + K + f^{-1}([s, b])$ is a proper subcontinuum containing x and y. Clearly $f^{-1}(t)$ separates K, since it separates M. To see that $f^{-1}(r) \cap K$ is connected, let $K' = \cap \{cl[f^{-1}((r, u))]; u \in (r, s)\}$. Then K' is a subcontinuum of $f^{-1}(r) \cap K$ which is easily seen to intersect each component of $f^{-1}(r) \cap K$. Thus $f^{-1}(r) \cap K$, as well as $f^{-1}(s) \cap K$, is connected.

THEOREM 2.4. *Let M be a continuum of type A; then Δ contains a unique minimal element.*

Proof. The proof of [11, Theorem 6, page 10] is valid, since we are not concerned with proving the upper semi-continuity of the decomposition.

COROLLARY 2.2. *Let M be a continuum of type A'. If $\mathscr{D} \in \Delta$ is such that each element of $\mathscr{D}$ has void interior, then $\mathscr{D}$ is the minimal element of Δ.*

Proof. Suppose that $\mathscr{D}' \in \Delta$ such that $\mathscr{D}' \leqq \mathscr{D}$. Let $D(a)$ and $D(b)$ denote the nonseparating elements of $\mathscr{D}$. Then $M - D(a)$ is connected, and since $D(a)^0 = \varnothing$, $[M - D(a)] + D'(a)$ is connected. Thus

$D(a) = D'(a)$ and $D(b) = D'(b)$. Given x in $M - [D(a) + D(b)]$, write $M - D(x) = A_x + B_x$ uniquely. Then $M = \bar{A}_x + \bar{B}_x$ and

$$\varnothing \neq \bar{A}_x \cap \bar{B}_x \subset D(x) .$$

Given z in $D(x)$, $D'(z)$ must separate $D(a)$ from $D(b)$; thus $\bar{A}_x \cap \bar{B}_x \subset D'(z)$. Consequently, $D'(x) = D(x)$ and $\mathscr{D}' = \mathscr{D}$.

The following useful result is a generalization of [11, Theorem 8, page 14].

THEOREM 2.5. *Let M be a continuum of type A, $\mathscr{D} \in \Delta$, and $f: M \rightarrow M/\mathscr{D} = [a, b]$. Then for $a \leqq r < s \leqq b$, $cl[f^{-1}((r, s))] = K$ is a subcontinuum of M which is irreducible from every point of*

$$K \cap f^{-1}(r) = K_r$$

to every point of $K \cap f^{-1}(s) = K_s$. Also K_r and K_s are subcontinua of K with void interior relative to K.

Proof. Since $K_r \subset K - f^{-1}((r, s))$, $K_r^0 = \varnothing$. By Theorem 2.3, K_r and K_s are subcontinua of K. That K is irreducible from K_r to K_s follows from the proof of [11, Theorem 8, page 14].

THEOREM 2.6. *Let M denote a continuum which is irreducible between two closed subsets H and K such that every subcontinuum of M with nonvoid interior is decomposable. Then the following hold.* (a) *There is a decomposition of M, $M = M_H + M_K$, where $H \subset M_H$, $K \subset M_K$ and $cl[M_H - M_K] \cap M_K$ is connected.* (b) *If U and V are open subsets of M such that $H \subset U \subset \bar{U} \subset V \subset M - K$ and both ∂U and ∂V are connected, then there is an open set W of M such that $\bar{U} \subset W \subset \bar{W} \subset V$ and ∂W is connected.*

Proof. The proof in [11, Theorem 9, page 14] is valid. Note that we have added the hypothesis that ∂U is connected in part (b).

THEOREM 2.7. *Let M be a continuum irreducible between a pair of points x and y. A necessary and sufficient condition that M be of type A' is that every subcontinuum of M with nonvoid interior be decomposable.*

Proof of sufficiency. Using the construction in [11, Theorem 10, page 15] we define a monotone function $f: M \rightarrow [0, 1]$. According to [4, Lemma 3, page 114] f is continuous. Thus $\{f^{-1}(t); t \in [0, 1]\}$ is an admissible decomposition for M. According to Theorem 2.2 there exists a minimal admissible decomposition for M, say $\mathscr{D}$. If some element $D \in \mathscr{D}$ has nonvoid interior then $cl(D^0)$ is of type A. Let $\mathscr{D}'$ denote an admissible decomposition for $cl(D^0)$. By combining $\mathscr{D}$ and $\mathscr{D}'$ in the natural way, one obtains an admissible decomposition which refines $\mathscr{D}$ properly. Thus no element of $\mathscr{D}$ has nonvoid interior, and M is of type A'.

Proof of necessity. See [11, Theorem 10, page 16].

By making the obvious necessary modifications, one can also generalize Theorems 17 through 22 of Chapter 1 of [11]. As in [11] we define $K(z) = \{y \in M;\ M$ is nonaposyndetic at z with respect to $y\}$ and $L(z) = \{y \in M;\ M$ is nonaposyndetic at y with respect to $z\}$. Observe that $L(z) = T(z)$ where T denotes the set function in [2]. The statements and proofs of Theorems 18 and 19 can be shortened by observing that $K(z) = L(z)$ for any point z of an irreducible continuum [2, Theorem 2, page 116]. Since Theorem 19 provides a concise topological characterization for continua of type A', we include its statement as Theorem 2.8.

THEOREM 2.8. *Let M denote a continuum irreducible from x to y. Then M is of type A' if and only if $K(z)^0 = \varnothing$ for each z in M.*

3. Hereditarily unicoherent, hereditarily decomposable continua. In [8] Miller proves that every irreducible, hereditarily decomposable metric continuum is of type A (this is a corollary of our Theorem 2.7). By applying this result she obtains a number of conditions which imply that a hereditarily decomposable metric continuum is hereditarily unicoherent, and she also shows that hereditarily unicoherent, hereditarily decomposable metric continua have certain properties analogous to properties of acyclic continuous curves (i.e., dendrites). In this section we will apply Theorem 2.7 to show that most (but not all) of Miller's results can be generalized to (nonmetric) continua.

It is easy to see that a continuum M is hereditarily unicoherent if and only if for each pair of distinct points x and y of M there exists exactly one subcontinuum of M which is irreducible from x to y.

By a *generalized simple closed curve* we mean a continuum which is separated by the omission of any two of its points. A point p is said to *cut* the continuum M in case there exist points x and y in M such that each subcontinuum of M containing x and y also contains p. Such a point, p, is said to cut x from y in M, or to cut between x and y in M.

The theorems that follow extend and generalize (to nonmetric continua) Theorems 2.4 through 2.9 of [8].

THEOREM 3.1. *Let M be a continuum of type A, and $\mathscr{D} \in \Delta$. If* *each element of $\mathscr{D}$ is unicoherent then M is unicoherent.*

Proof. Let $f: M \to M/\mathscr{D} = [a, b]$ denote the natural map. Suppose that H and K are proper subcontinua of M such that $M = H + K$. If $f(H) = [a, c]$ and $f(K) = [c, b]$ then $H \cap K \subset f^{-1}(c)$. Now

$$[H \cap f^{-1}(c)] + [K \cap f^{-1}(c)] = f^{-1}(c) .$$

Since $H \cap f^{-1}(c)$ and $K \cap f^{-1}(c)$ are continua (Theorem 2.3), and $f^{-1}(c)$ is unicoherent, $H \cap K = [H \cap f^{-1}(c)] \cap [K \cap f^{-1}(c)]$ is connected. The other cases are handled in a similar manner, although they do not depend on the unicoherence of the elements of $\mathscr{D}$.

THEOREM 3.2. *Let M be a continuum of type A, and $\mathscr{D} \in \Delta$. If $f: M \to M/\mathscr{D} = [a, b]$ is an open map, then M is unicoherent.*

Proof. Let H and K be proper subcontinua of M such that $M = H + K$. If $f(H) = [a, c]$ and $f(K) = [c, b]$ then

$$H \cap f^{-1}(c) = f^{-1}(c) = K \cap f^{-1}(c)$$

since f is open. Thus $H \cap K = f^{-1}(c)$ which is connected. The other cases are handled as in Theorem 3.1.

THEOREM 3.3. *If M is a hereditarily decomposable continuum which is not unicoherent, then M contains a continuum N which is a generalized simple closed curve with respect to the elements of a monotone upper semi-continuous decomposition $\mathscr{D}$. Furthermore, if D_1 and D_2 are in $\mathscr{D}$ then $N - (D_1 + D_2) = U + V$ where U and V are disjoint connected open sets such that* (1) *$N = \bar{U} + \bar{V}$,* (2) *$\bar{U}$ and $\bar{V}$ are irreducible from D_1 to D_2, and* (3) *any subcontinuum of $D_1 + D_2 + U$ which intersects D_1 and D_2 contains $\bar{U}$.*

Proof. Apply Theorem 2.7 to the proof of [8, Theorem 2.6, page 187].

THEOREM 3.4. *Let M be a hereditarily decomposable continuum. M is hereditarily unicoherent if and only if M contains no subcontinuum N which is a generalized simple closed curve with respect to the elements of a monotone upper semi-continuous decomposition.*

Proof. If M is not hereditarily unicoherent, apply Theorem 3.3. Conversely, suppose that $f: N \to C$, where N is a subcontinuum of M, f is monotone and onto, and C is a generalized simple closed curve. Write $C = A + B$ where A and B are generalized arcs. Then $f^{-1}(A) \cap f^{-1}(B)$ is not connected.

THEOREM 3.5. *Let M be a hereditarily decomposable continuum. Suppose that there exists a cardinal number $k \leqq c$ such that given k points of M there exists one of them which cuts between some pair of them. Then M is hereditarily unicoherent.*

Proof. Suppose M is not hereditarily unicoherent. According to Theorem 3.4 there exists a subcontinuum N of M, a generalized simple closed curve C, and a monotone map f from N onto C. Choosing k distinct points of C it is clear that no one cuts between any pair of them. The theorem follows.

THEOREM 3.6. *If M is a hereditarily decomposable continuum every subcontinuum of which is irreducible about a closed proper subset having only countably many components, then M is hereditarily unicoherent.*

Proof. Apply [5, Theorem 6, page 173] to the proof of [8, Theorem 2.9].

Theorem 3.6 does not remain true if "countably many components" is replaced by "c components". A simple modification of Example 2 [11, page 12] produces a metric continuum which is irreducible about a closed set with uncountably many components and is not unicoherent.

In order to obtain generalizations of theorems in [8, Section 3, page 190] we prove a generalization of a theorem due to R. L. Moore [10].

THEOREM 3.7. *Let M denote a hereditarily unicoherent continuum, and suppose that each indecomposable subcontinuum of M is irreducible. If H is an irreducible subcontinuum of M then H is contained in a maximal irreducible subcontinuum.*

Proof. Throughout this proof $\langle x, y\rangle$ denotes the unique irreducible continuum from x to y.

Suppose that H is irreducible from a to b. Let $\{H_\alpha\}$ be a maximal monotonic collection of continua such that $H \subset H_\alpha$ for each α, and $H_\alpha = \langle a, h_\alpha\rangle$ for some h_α in M. Let $K = cl(\bigcup_\alpha H_\alpha)$. We will prove that the continuum K is irreducible from a to some point k. Assume not. Observe that if A is a proper subcontinuum of K which contains a, then $K - A$ is connected. There are two cases to consider.

Case 1. Suppose that $cl(K - A)$ is indecomposable for some subcontinuum A of K which contains a. Let $T = cl(K - A)$. Then $T \cap$

A is a proper subcontinuum of T; hence $T \cap A$ is contained in a composant C of T. Since T is irreducible, it contains at least two composants. Choose $k \in T - C$. Then $\langle a, k \rangle = K$. To see this, suppose that $\langle a, k \rangle \neq K$. Then $\langle a, k \rangle \cap T$ is a continuum which intersects two composants of T; thus $T \subset \langle a, k \rangle$. Choose $h_\alpha \in A - \langle a, k \rangle$, $h_\beta \in K - A$. Then $H_\alpha \not\subset H_\beta$ and $H_\beta \not\subset H_\alpha$, which is a contradiction.

Case 2. $cl(K - A)$ is decomposable for each subcontinuum A of K containing a. If

$$cl(K - A) = E + F$$

is any decomposition of $cl(K - A)$, then $A \cap F = \emptyset$ or $A \cap E = \emptyset$. Using this fact it is easy to verify that there exists an H_β such that $A \subset H_\beta^\circ$. In particular, given an H_α, there exists an H_β such that $H_\alpha \subset H_\beta^\circ$. Choosing k in $\bigcap_\alpha cl(K - H_\alpha)$ it follows that $\langle a, k \rangle = K$.

In either case, K is "maximally irreducible" from a to some point k. If $\langle x, y \rangle$ contains $K = \langle a, k \rangle$ properly, then $\langle x, y \rangle = \langle x, k \rangle$ or $\langle x, y \rangle = \langle y, k \rangle$. For suppose not and let $x \notin K$. Then $k \notin \langle a, x \rangle$; hence $y \notin \langle a, x \rangle$. Since $\langle x, k \rangle$ is properly contained in $\langle x, y \rangle$, $y \notin \langle x, k \rangle$. But $K \subset \langle a, x \rangle + \langle x, k \rangle$; thus $y \notin K$. Now $\langle x, y \rangle \subset \langle a, x \rangle + \langle a, y \rangle$ which misses k. This is a contradiction.

Let L be a continuum containing K which is "maximally irreducible" from k to some point. Then L is a maximal irreducible subcontinuum containing H. For if $L \subset \langle z, y \rangle$ then $K \subset \langle x, y \rangle$. According to the argument above we can assume that $\langle x, y \rangle = \langle x, k \rangle$. It follows immediately that $\langle x, y \rangle = L$.

COROLLARY 3.1. *Let M denote a hereditarily unicoherent, hereditarily decomposable continuum. If H is an irreducible subcontinuum of M, then H is contained in a maximal irreducible subcontinuum.*

COROLLARY 3.2 (Moore). *Let M denote a hereditarily unicoherent metric continuum. If H is an irreducible subcontinuum of M, then H is contained in a maximal irreducible subcontinuum.*

Proof. Every indecomposable metric continuum is irreducible.

As in [8], we define a point p to be a *terminal point* of the continuum M in case every irreducible subcontinuum of M which contains p is irreducible from p to some point. By making use of Theorem 2.7 and Corollary 3.1 we obtain the following generalizations of theorems in [8, § 3, page 190].

THEOREM 3.8. *Every point of a hereditarily unicoherent continuum*

M is either a terminal point or a cut point of M.

THEOREM 3.9. *A continuum which is hereditarily unicoherent and hereditarily decomposable has at least two terminal points.*

THEOREM 3.10. *A continuum which is hereditarily unicoherent and hereditarily decomposable is irreducible about the set of all its terminal points.*

THEOREM 3.11. *If the continuum M is hereditarily decomposable and K is a subset of M consisting of some of the terminal points of M, then $M - K$ is connected.*

In § 4 we will see that Theorem 3.7 of [8] does not generalize to nonmetric continua.

4. Some properties of trees. A continuum M is said to be a *tree* [12] if and only if given two distinct points p and q of M, there exists a third point which separates p from q. The point p of a tree M is said to be an *end point* of M if and only if p is a nonseparating point of every generalized arc containing p. It is known [12] that a continuum M is a tree if and only if M is locally connected and hereditarily unicoherent. If M is a metric continuum then M is a tree if and only if M is a dendrite [13, (1.1), page 88]. In Theorem 4.1 we show that a number of familiar properties of dendrites are also shared by trees.

THEOREM 4.1. *Let M denote a tree. Then* (1) *M is connected by generalized arcs,* (2) *each point of M is a separating point or an end point,* (3) *each generalized arc in M is contained in a maximal generalized arc,* (4) *M has at least two end points,* (5) *M is irreducible about the set of all its end points,* (6) *if K is a subset of the end points of M, then $M - K$ is connected.*

Proof. Let A be a subcontinuum of M irreducible from p to q. Since M is hereditarily unicoherent, each point of $A - (p + q)$ cuts p from q in M; thus, since M is locally connected, each point of $A - (p + q)$ actually separates p from q in M. Consequently, A is a generalized arc. Since M is hereditarily decomposable, properties (2) through (6) follow from Theorems 3.7 through 3.11.

For a metric continuum M the following properties are equivalent [13, (1.1), page 88]: (a) M is a tree, (b) M is locally connected and contains no (generalized) simple closed curve, (c) every subcontinuum of M contains uncountably many separating points of M.

For (nonmetric) continua we have seen that condition (a) implies conditions (b) and (c). However, neither of these implications can be reversed. Mardešić has shown [6] that there exists a locally connected continuum which contains no proper locally connected subcontinuum. This example clearly satisfies condition (b), but is not a tree. The following example satisfies condition (c) but not (a); and also shows that [8, Theorem 3.7, page 193] does not generalize to (nonmetric) continua.

EXAMPLE. Let C denote a circle, and let $M = C \times [0, 1]$. We define a basis $\mathscr{B}$ for the topology on M as follows: V is in $\mathscr{B}$ if and only if (1) $V = p \times (r, s)$, (2) $V = p \times (r, 1]$, or (3)

$$V = (U \times [0, 1]) - \bigcup_{i=1}^{n} \{p_i \times [q_i, 1]\} ,$$

where U is open in the usual topology for C, p_i is in U, and $0 < q_i < 1$. If $\mathscr{T}$ denotes the topology generated by $\mathscr{B}$ then $(M, \mathscr{T})$ is seen to be a (compact Hausdorff) continuum with the desired properties.

Finally, we give a characterization of trees in terms of inverse limits. For a discussion of inverse limits systems, see [1].

THEOREM 4.2. *The continuum M is a tree if and only if M is homeomorphic to the inverse limit of a monotone inverse system $(D_\alpha, \pi_{\alpha\beta}, \Lambda)$ where each D_α is a (metric) dendrite.*

Proof. According to [12] we must show that M is locally connected and hereditarily unicoherent. M is locally connected by [1, Theorem 4.3, page 241]. A simple application of [1, page 235, 2.9] shows that M is hereditarily unicoherent. On the other hand, since M is locally connected, M can be written as the inverse limit of a monotone inverse system $(D_\alpha, \pi_{\alpha\beta}, \Lambda)$ where each D_α is a locally connected metric continuum [7]. According to [1], $\pi_\alpha: M \rightarrow D_\alpha$ is monotone. It follows easily that D_α is a tree, hence a dendrite.

5. Continua hereditarily of type A'. As in Chapter 2 of [11], we define a continuum M to be *hereditarily of type A'* if and only if every nondegenerate subcontinuum of M is of type A'. If M is a hereditarily decomposable metric continuum then M is hereditarily of type A' if and only if M is snake-like [11, Theorem 13, page 50]. In this section we obtain several topological characterizations of (nonmetric) continua which are hereditarily of type A'.

THEOREM 5.1. *If the continuum M is hereditarily of type A', then M is hereditarily unicoherent and atriodic.*

Proof. The proof of [11, Theorem 6, page 41] is valid for (non-metric) continua.

LEMMA 5.1. *If the continuum M is hereditarily unicoherent and atriodic, then given three points of M, one cuts between the other two.*

THEOREM 5.2. *The continuum M is hereditarily of type A' if and only if M is hereditarily unicoherent, hereditarily decomposable, and atriodic.*

Proof. Suppose that M is hereditarily unicoherent, hereditarily decomposable, and atriodic. According to Theorem 2.7 it suffices to show that every subcontinuum N of M is irreducible. Let A be a maximal irreducible subcontinuum of N (Theorem 3.7) which is irreducible from p to q. If there exists a point r in $N - A$ then, since A is maximal irreducible, it follows that none of p, q, and r cuts between the other two. This contradicts Lemma 5.1; hence $N = A$. The converse follows from Theorem 5.1.

THEOREM 5.3. *Let M denote a hereditarily decomposable continuum. Then M is hereditarily of type A' if and only if given any three points of M one cuts between the other two.*

Proof. If M is hereditarity of type A' apply Theorem 5.1 and Lemma 5.1. If given any three points one cuts between the other two then M is hereditarily unicoherent (Theorem 3.5). Clearly M contains no triods. Thus, by Theorem 5.2, M is hereditarily of type A'.

REFERENCES

1. C. E. Capel, *Inverse limit spaces*, Duke Math. J., **21** (1954), 233-246.
2. H. S. Davis, D. P. Stadtlander and P. M. Swingle, *Properties of the set functions T^n*, Portugaliae Mathematica, **21** (1962), 113-133.
3. J. G. Hocking and G. S. Young, *Topology*, Addison-Wesley, Reading, Mass., 1961.
4. J. K. Kelley, *General Topology*, Van Nostrand, Princeton, 1955.
5. K. Kuratowski, *Topology II*, PWN-Academic Press, Warsaw-New York, 1968.
6. S. Mardešić, *A locally connected continuum which contains no proper locally connected subcontinuum*, Glasnik Matematicki, **2** (22) (1967), 167-178.
7. ———, *Locally connected, ordered and chainable continua*, Rad Jugoslav. Akad. Znan. Umjetn., **319** (1960), 147-166.
8. H. C. Miller, *On unicoherent continua*, Trans. Amer. Math. Soc., **69** (1950), 179-194.

9. R. L. Moore, *Foundation of point set theory*, Amer. Math. Soc. Colloquium Publications 13, Revised Edition, Providence, 1962.
10. ———, *Concerning compact continua which contain no continuum which separates the plane*, Proc. Nat. Acad. Sci., U.S.A., **29** (1934), 41-45.
11. E. S. Thomas, Jr., *Monotone decompositions of irreducible continua*, Rozprawy Matematyczne 50, Warszawa, 1966.
12. L. E. Ward, Jr., *Mobs, trees and fixed points*, Proc. Amer. Math. Soc., **8** (1957), 798-804.
13. G. T. Whyburn, *Analytic Topology*, Amer. Math. Soc. Colloquium Publications 28, Providence, 1942.
14. R. L. Wilder, *Topology of Manifolds*, Amer. Math. Soc. Colloquium Publications 32, Providence, 1949.

Received September 10, 1970. This work was supported by a National Science Foundation Traineeship.

UNIVERSITY OF CALIFORNIA, RIVERSIDE

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

THE MATRIX EQUATION $AXB = X$

D. J. HARTFIEL

This paper considers the solutions of the matrix equation $AXB = X$ where we specify A and B to be n-square and doubly stochastic. Solutions are found explicitly and do not depend on either the Jordan or Rational canonical forms. We further find all doubly stochastic solutions of this equation, by noting that $J_n = (1/n)$, the n-square doubly stochastic matrix in which each entry is $1/n$, is always a solution and that the doubly stochastic solutions form a compact convex set. We solve the equation by characterizing the vertices of this convex set.

Matrices considered in this paper are real matrice unless otherwise stated. Most of the definitions and notation may be found in [5], although some will be presented below.

If $A_1, A_2, \cdots, A_s$ are square matrices, by $\sum_{k=1}^{\cdot s} A_k$ we mean the direct sum of the A_k's. If $s = 2$ we may write $A_1 \oplus A_2$ for this direct sum. We say that a square matrix A is reducible if there exists a permutation matrix P so that $PAP^t = \begin{pmatrix} X & O \\ Y & Z \end{pmatrix}$ where X and Z are square and P^t denotes the transpose of P. If A is not reducible, then it is said to be irreducible. A square matrix $A = (a_{ij})$ is doubly stochastic if $a_{ij} \geqq 0$ and $\sum_k a_{ik} = \sum_k a_{kj} = 1$ for all i, j. It readily follows that if A is doubly stochastic, then there exists a permutation matrix P such that $PAP^t = \sum_{k=1}^{\cdot s} A_k$ where each A_k is doubly stochastic and irreducible.

The following two celebrated theorems in matrix theory are used in the paper.

BIRKHOFF'S THEOREM. *The set of all n-square doubly stochastic matrices, Ω_n, forms a convex polyhedron with the permutation matrices as vertices* [5, p. 97].

PERRON-FROBENIUS THEOREM. *Let A be an n-square nonnegative irreducible matrix. Then:*

(i) *A has a real positive characteristic root r which is simple. If λ is any characteristic root of A, then $|\lambda| \leqq r$.*

(ii) *If A has h characteristic roots of modulus*

$$r{:}\ \lambda_0 = r, \lambda_1, \cdots, \lambda_{h-1}$$

then these are h distinct roots of $\lambda^h - r^h = 0$, h is called the index of imprimitivity of A. If $h = 1$ the matrix is called primitive.

(iii) *If* $\lambda_0, \lambda_1, \cdots, \lambda_{n-1}$ *are all the characteristic roots of* A, *and* $\theta = e^{i(2\pi/n)}$ *then* $\lambda_0\theta, \cdots, \lambda_{n-1}\theta$ *are* $\lambda_0, \cdots, \lambda_{n-1}$ *in some order.*

(iv) *If* $h > 1$, *then there exists a permutation matrix* P *such that*

$$PAP^t = \begin{pmatrix} 0 & A_{12} & 0 & \cdots & 0 & 0 \\ 0 & 0 & A_{23} & \cdots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & A_{h-1,h} \\ A_{h,1} & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} .$$

where the zero blocks down the main diagonal are square [5, p. 125].

If A is a nonnegative matrix and

$$a_{i_1j_1}, a_{i_1j_2}, a_{i_2j_2}, \cdots \cdot a_{i_{m-1}j_m}, a_{i_mj_m} = a_{i_1j_1}$$

are all positive elements in A, then A is said to have a loop of length m. If $A = (a_{ij})$ is such that all a_{ij} are equal, then we say that A is flat. If A is partitioned into block matrices A_{ij}, i.e., $A = (A_{ij})$, and each A_{ij} is flat, then a block loop is defined similarly.

1. Preliminary results. First we note that if P and Q are permutation matrices then $AXB = X$ if and only if

$$PAP^tPXQQ^tBQ = PXQ .$$

Since A and B can each be put into a direct sum of irreducible matrices by simultaneous row and column permutations we may assume by the Perron-Frobenius Theorem that

$$A = \sum_{\alpha=1}^{s}{}^{\cdot} A_\alpha , \qquad B = \sum_{\beta=1}^{r}{}^{\cdot} B_\beta$$

$$A_\alpha = \begin{pmatrix} 0 & A_1^\alpha & 0 & \cdots & 0 \\ 0 & 0 & A_2^\alpha & \cdots & 0 \\ & \cdots & \cdots & \cdots & \\ 0 & 0 & 0 & \cdots & A_{s_\alpha-1}^\alpha \\ A_{s_\alpha}^\alpha & 0 & 0 & \cdots & 0 \end{pmatrix}, \qquad B_\beta = \begin{pmatrix} 0 & B_1^\beta & 0 & \cdots & 0 \\ 0 & 0 & B_2^\beta & \cdots & 0 \\ & \cdots & \cdots & \cdots & \\ 0 & 0 & 0 & \cdots & B_{r_\beta-1}^\beta \\ B_{r_\beta}^\beta & 0 & 0 & \cdots & 0 \end{pmatrix}$$

where A_α is irreducible with index of imprimitivity s_α; B_β is irreducible with index of imprimitivity r_β. Further the 0 blocks down the main diagonal on A_α and B_β are all square.

Note that the dimension of each A_k^α $(k = 1, 2, \cdots, s_\alpha)$ is the same for each fixed α. For fixed β the dimensions of the B_k^β $(k = 1, 2, \cdots, r_\beta)$ are also equal. Hence

$$A_{\alpha}^{s_\alpha} = \begin{pmatrix} C_1 & 0 & 0 & \cdots & 0 \\ 0 & C_2 & 0 & \cdots & 0 \\ & & \cdots\cdots & & \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & C_{s_\alpha} \end{pmatrix}, \qquad B_{\beta}^{r_\beta} = \begin{pmatrix} D_1 & 0 & 0 & \cdots & 0 \\ 0 & D_2 & 0 & \cdots & 0 \\ & & \cdots\cdots & & \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & D_{r_\beta} \end{pmatrix}$$

where each C_k $(k = 1, 2, \cdots, s_\alpha)$, D_k $(k = 1, 2, \cdots, r_\beta)$ is a primitive doubly stochastic matrix. Now let p be a sufficiently large integer so that A^p and B^p are direct sums of primitive matrices.

LEMMA 1.1. *If T is a linear operator on a convex set S whose vertices are X_i $(i = 1, 2, \cdots, m)$, then $T(S)$ is a convex set whose vertices are in $\{T(X_i) \mid i = 1, \cdots, m\}$.*

THEOREM 1.2. *The set of doubly stochastic solutions of the matrix equation $A^pXB^p = X$ (p previously defined) is the convex hull of*

$$\left\{\lim_{k\to\infty} (A^p)^k P_l (B^p)^k \mid P_l \textit{ is a permutation matrix, } l = 1, 2, \cdots, n!\right\}.$$

Proof. If V is an $m \times m$ primitive doubly stochastic matrix, then $\lim_{k\to\infty} V^k = J_m$, the flat $m \times m$ doubly stochastic matrix.

$$\lim_{k\to\infty} (A^p)^k \quad \text{and} \quad \lim_{k\to\infty} (B^p)^k$$

exist, their limits being direct sums of flat doubly stochastic matrices. Let $L(X) = \lim_{k\to\infty} (A^p)^k X (B^p)^k$. This is a linear operator defined on the set of $n \times n$ matrices.

By Lemma 1.1, $L(\Omega_n)$ is the convex hull of $\{L(P_l) \mid P_l$ is a permutation matrix$\}$ i.e., of $\{\lim_{k\to\infty} (A^p)^k P_l (B^p)^k \mid P_l$ is a permutation matrix$\}$.

Now if $A^pXB^p = X$, $X \in \Omega_n$, then $L(X) = X$ and by Birkhoff's Theorem, X is in the convex hull of the $\{L(P_l) \mid P_l$ is a permutation matrix$\}$. Furthermore, if X is in the convex hull of the $\{L(P_l) \mid P_l$ is a permutation matrix$\}$ i.e., $X = \Sigma\lambda_l L(P_l)$ where $\lambda_l \geqq 0$ and $\Sigma\lambda_l = 1$, then

$$X = \Sigma\lambda_l L(P_l) = \Sigma\lambda_l \lim_{k\to\infty} (A^p)^k P_l (B^p)^k$$
$$= A^p\left[\Sigma\lambda_l \lim_{k\to\infty} (A^p)^{k-1} P_l (B^p)^{k-1}\right]B^p = A^pXB^p,$$

and X is a solution of the matrix equation.

THEOREM 1.3. *$Y \in \Omega_n$ is a solution of $AXB = X$ if and only if $Y = \sum_{k=0}^{p-1} A^k W B^k / p$ where $W \in \Omega_n$ is a solution of $A^pXB^p = X$.*

Proof. If $Y = \sum_{k=0}^{p-1} A^k W B^k / p$, W a solution of $A^pXB^p = X$, then $AYB = Y$.

Further if Y is solution of $AXB = X$ then Y is a solution of $A^pXB^p = X$ and so $Y = \sum_{k=0}^{p-1} A^k YB^k/p$.

Let $M(Z) = \sum_{k=0}^{p-1} A^k ZB^k/p$. Then M is a linear operator defined on the set of $n \times n$ matrices.

COROLLARY 1.4. *The vertices of the set of doubly stochastic solutions of* $AXB = X$ *is a subset of* $\{M[L(P_l)] \mid P_l$ *is a permutation matrix*$\}$.

Proof. The proof follows from Lemma 1.1, Theorem 1.2, and Theorem 1.3.

COROLLARY 1.5. *If one of* A *or* B *is primitive, then the only doubly stochastic solution of the equation* $AXB = X$ *is* J_n.

Proof. Either $\lim_{k\to\infty} (A^p)^k$ or $\lim_{k\to\infty} (B^p)^k$ is J_n. Thus if X is doubly stochastic, then $L(X) = J_n$.

2. The operator L. Our primary aim here is to investigate the structure of the convex set $L(\Omega_n)$: in particular its vertices.

From § 1 we know for P_l a permutation matrix

$$L(P_l) = \lim_{k\to\infty} (A^p)^k P_l (B^p)^k = \Big(\sum_\gamma{}^{\cdot} J_\gamma^A\Big) P_l \Big(\sum_\sigma{}^{\cdot} J_\sigma^B\Big)$$

where J_τ^A and J_σ^B are flat doubly stochastic matrices whose dimensions correspond to the dimension of the primitive matrices in the direct sums A^p and B^p respectively.

Suppose $a_\gamma \times a_\gamma$ is the dimension of J_γ^A and $b_\sigma \times b_\sigma$ is the dimension of J_σ^B. Set $(\sum^{\cdot}_\gamma J_\gamma^A) P_l (\sum^{\cdot}_\sigma J_\sigma^B) = V_l$. Partition V_l into blocks $V_{\gamma\sigma}$ of dimension $a_\gamma \times b_\sigma$.

LEMMA 2.1. *If* $X \in L(\Omega_n)$ *is partitioned into block matrices* $X_{\gamma\sigma}$ *of dimension* $a_\gamma \times b_\sigma$, *then each* $X_{\gamma\sigma}$ *is flat.*

THEOREM 2.2. *If* $X \in L(\Omega_n)$ *is partitioned into block matrices* $X_{\gamma\sigma}$ *of dimension* $a_\gamma \times b_\sigma$, *then* X *is a vertex of* $L(\Omega_n)$ *if and only if* X *does not have a block loop.*

Proof. Suppose X has a block loop

$$X_{\gamma_1\sigma_1}, X_{\gamma_1\sigma_2}, X_{\gamma_2\sigma_2}, \cdots, X_{\gamma_m\sigma_m} = X_{\gamma_1\sigma_1} .$$

Add $\varepsilon > 0$ to each element in the $\gamma_1\sigma_1$ block. Subtract $(b_{\sigma_1}/b_{\sigma_2})\varepsilon$ from each element in the $\gamma_1\sigma_2$ block. All the row sums of the matrix are now one. Now add $(a_{\gamma_1}b_{\sigma_1}/a_{\gamma_2}b_{\sigma_2})\varepsilon$ to each element in the $\gamma_2\sigma_2$ block.

All the column sums of the matrix are now one. Now subtract

$$\frac{b_{\sigma_2}a_{\gamma_1}b_{\sigma_1}}{b_{\sigma_3}a_{\gamma_2}b_{\sigma_2}}\varepsilon$$

from each element in the $\gamma_2\sigma_3$ block. All the row sums of the matrix are now one. Continuing in this manner we see that in the $\gamma_m\sigma_m$ block we add $(a_{\gamma_{m-1}}b_{\sigma_{m-1}}\cdots b_{\sigma_1}/a_{\gamma_m}b_{\sigma_m}\cdots b_{\sigma_2})\varepsilon = \varepsilon$. This is exactly what is in the $\gamma_m\sigma_m$ or $\gamma_1\sigma_1$ block. Now all rows and columns sum to one. Call this generated matrix X'. Now considering the same block loop we generate X'' by replacing ε by $-\varepsilon$ in X'. Again all rows and columns sum to one. Now $X = \frac{1}{2}(X' + X'')$, and since X' and $X'' \in L(\Omega_n)$ for ε sufficiently small, X is an interior point.

On the other hand if $X \in L(\Omega_n)$ and interior to it, there are X' and X'' in $L(\Omega_n)$ so that $X = \frac{1}{2}(X' + X'')$. We may pick X' and X'' in $L(\Omega_n)$ so that they have zero blocks in the block position if and only if X does. Now if $X' \neq X''$ then there is a $\gamma_1\sigma_1$ block so that $X'_{\gamma_1\sigma_1} < X''_{\gamma_1\sigma_1}$ where $X'_{\gamma_1\sigma_1}$ is a block in X', $X''_{\gamma_1\sigma_1}$ is a block in X'' and the relation is elementwise. Hence there is a $X'_{\gamma_1\sigma_2} > X''_{\gamma_1\sigma_2}$ and so on. This generates a block loop in X.

COROLLARY 2.3. *X is a vertex of the convex set of doubly stochastic matrices if and only if X does not have a loop.*

Proof. Consider the matrix equation $IXI = X$ and apply the Theorem 2.2.

We are now in a position to find the vertices of $L(\Omega_n)$. Partition each permutation matrix P_l into blocks $P^l_{\gamma\sigma}$ of dimension $a_\gamma \times b_\sigma$. Let $n_{\gamma\sigma}$ be the number of ones in the $\gamma\sigma$ block of P_l. Then

$$\Big(\sum_\gamma{}^{\cdot} J^A_\gamma\Big)P_l\Big(\sum_\sigma{}^{\cdot} J^B_\sigma\Big) = V_l$$

and $V_{\gamma\sigma}$ has all its elements equal to $n_{\gamma\sigma}/a_\gamma b_\sigma$. We may now use Theorem 2.2 on this finite set to establish exact vertices.

EXAMPLE.

$$\begin{pmatrix} \begin{pmatrix} \frac12 & \frac12 \\ \frac12 & \frac12 \end{pmatrix} & 0 \\ 0 & \begin{pmatrix} \frac12 & \frac12 \\ \frac12 & \frac12 \end{pmatrix} \end{pmatrix} P_l \begin{pmatrix} \begin{pmatrix} \frac12 & \frac12 \\ \frac12 & \frac12 \end{pmatrix} & 0 \\ 0 & \begin{pmatrix} \frac12 & \frac12 \\ \frac12 & \frac12 \end{pmatrix} \end{pmatrix}.$$

Partitioning the matrices P_l we have

$$(1) \qquad L\left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{cc|cc} \frac{1}{2} & \frac{1}{2} & 0 & \\ \frac{1}{2} & \frac{1}{2} & & \\ \hline 0 & & \frac{1}{2} & \frac{1}{2} \\ & & \frac{1}{2} & \frac{1}{2} \end{array}\right), \quad \text{a vertex} ,$$

$$(2) \qquad L\left(\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right) = \left(\begin{array}{cc|cc} 0 & & \frac{1}{2} & \frac{1}{2} \\ & & \frac{1}{2} & \frac{1}{2} \\ \hline \frac{1}{2} & \frac{1}{2} & 0 & \\ \frac{1}{2} & \frac{1}{2} & & \end{array}\right), \quad \text{a vertex} .$$

All vertices are of the form $L(P_l)$ for some permutation P_l. However, $L(P_l)$ is not always a vertex for every l. For example,

$$(3) \qquad L\left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{cc|cc} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \hline \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array}\right) \quad \text{an interior point} .$$

We can further note by Theorem 2.2 that 1 and 2 are the only vertices of $L(\Omega_n)$.

3. General solutions of $A^pXB^p = X$. We already know from Theorem 1.2 that for each $W \in \Omega_n$, $L(W)$ is a solution of $A^pXB^p = X$. Actually we have shown that if W is any $n \times n$ matrix then $L(W)$ is a solution of $A^pXB^p = X$. Further if W is a solution of the equation then $L(W) = W$. i.e., $(\sum_\gamma^{\cdot} J_\gamma^A) W (\sum_\sigma^{\cdot} J_\sigma^B) = W$. Partition W into blocks $W_{\gamma\sigma}$ as in § 2. Now $J_\gamma^A W_{\gamma\sigma} J_\sigma^B = W_{\gamma\sigma}$ implies that $W_{\gamma\sigma}$ is flat. Also if each $W_{\gamma\sigma}$ of W is flat, then W is a solution. Hence we know all solutions of the matrix equation $A^pXB^p = X$.

4. Orbits in matrices. Let $C = (c_{ij})$ be a $p \times q$ matrix. Suppose we pick some $c_{i_1j_1}$. Then by the orbit of $c_{i_1j_1}$ we mean the set of positions $(i_1 - k, j_1 + k)$ $[k = 0, 1, \cdots]$ where the row index is modulo p and the column index is modulo q.

EXAMPLE.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \qquad \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$$

The numbers in the positions of the orbit of

(1) 5 are 5, 3, 7
(2) 2 are 2, 9, 4
(3) 1 are 1, 8, 6
(4) a are a, e, c, d, b, f.

Consider the group $Z/p \oplus Z/q$ where Z is the additive group of integers. Note that $K = \{(-k \bmod p, k \bmod q) \mid k \in Z\}$ is a subgroup of $(Z/p \oplus Z/q)$. Hence we can consider orbits as cosets in $(Z/p \oplus Z/q)/K$ by looking at indices. We now see:

1. The number of elements in each orbit is the same.
2. If two orbits intersect, they are the same.
3. If one orbit contains a row index k times then all orbits contain that row index k times. The same property holds for columns.
4. Each row index and column index appear at least once in each orbit.
5. If p and q are relatively prime, then there is only one distinct orbit.

Finally we note that since orbits are defined by indices, we may consider block orbits in partitioned matrices.

5. **The operator M.** Our aim here is to investigate the structure of the convex set $M[L(\Omega_n)]$: in particular to find its vertices. Let $X \in L(\Omega_n)$. Partition X into blocks $X_{\gamma\sigma}$ of dimension $a_\gamma \times b_\sigma$, then

$$M(X) = \frac{1}{p}\sum_{k=0}^{p-1} A^k X B^k = \frac{1}{p}\sum_{k=0}^{p-1} \sum_\alpha{}^{\cdot} \begin{pmatrix} 0 & A_1^\alpha & 0 & 0 & \cdots & 0 \\ 0 & 0 & A_2^\alpha & 0 & \cdots & 0 \\ & & \cdots\cdots & & & \\ A_{s_\alpha}^\alpha & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}^k X \sum_\beta{}^{\cdot} \begin{pmatrix} 0 & B_1^\beta & 0 & 0 & \cdots & 0 \\ 0 & 0 & B_2^\beta & 0 & \cdots & 0 \\ & & \cdots\cdots & & & \\ B_{r_\beta}^\beta & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}^k$$

and since the blocks $X_{\gamma\sigma}$ of X are flat we may write

$$M(X) = \frac{1}{p}\sum_{k=0}^{p-1} \sum_\alpha{}^{\cdot} \begin{pmatrix} 0 & J_1^\alpha & 0 & 0 & \cdots & 0 \\ 0 & 0 & J_2^\alpha & 0 & \cdots & 0 \\ & & \cdots\cdots & & & \\ J_{s_\alpha}^\alpha & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}^k X \sum_\beta{}^{\cdot} \begin{pmatrix} 0 & J_1^\beta & 0 & 0 & \cdots & 0 \\ 0 & 0 & J_1^\beta & 0 & \cdots & 0 \\ & & \cdots\cdots & & & \\ J_{r_\beta}^\beta & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}^k$$

where $J_k^\alpha\,(k = 1, 2, \cdots, s_\alpha)$ and $J_k^\beta\,(k = 1, 2, \cdots, r_\beta)$ are flat doubly stochastic matrices whose dimensions are the same as those of A_k^α and B_k^β, respectively. Suppose the irreducible blocks A_α of $\sum^{\cdot}_\alpha A_\alpha$ have dimension $p_\alpha \times p_\alpha$ and the irreducible blocks B_β of $\sum^{\cdot}_\beta B_\beta$ have the dimension $q_\beta \times q_\beta$. Partition X into blocks $X'_{\alpha\beta}$ of dimension $p_\alpha \times q_\beta$. We call these blocks the major blocks of X. Now since X is already partitioned into blocks of dimension $a_\gamma \times b_\sigma$, we see that the major blocks are partitioned into the $X_{\gamma\sigma}$ blocks in the first partitioning. We call each block in the original partition a minor block. Note that inside each major block, all minor blocks are of the same dimension.

Now suppose $X'_{\alpha\beta}$ is a major block of X. Then we see the sequence

$$X'_{\alpha\beta},\begin{pmatrix}0 & J_1^\alpha & 0 & \cdots & 0\\ 0 & 0 & J_2^\alpha & \cdots & 0\\ \cdots & & & & \\ J_{s_\alpha}^\alpha & 0 & 0 & \cdots & 0\end{pmatrix}X'_{\alpha\beta}\begin{pmatrix}0 & J_1^\beta & 0 & \cdots & 0\\ 0 & 0 & J_2^\beta & \cdots & 0\\ \cdots & & & & \\ J_{r_\beta}^\beta & 0 & 0 & \cdots & 0\end{pmatrix},\cdots,$$
$$\begin{pmatrix}0 & J_1^\alpha & 0 & \cdots & 0\\ 0 & 0 & J_2^\alpha & \cdots & 0\\ \cdots & & & & \\ J_{s_\alpha}^\alpha & 0 & 0 & \cdots & 0\end{pmatrix}^{p-1}X'_{\alpha\beta}\begin{pmatrix}0 & J_1^\beta & 0 & \cdots & 0\\ 0 & 0 & J_2^\beta & \cdots & 0\\ \cdots & & & & \\ J_{r_\beta}^\beta & 0 & 0 & \cdots & 0\end{pmatrix}^{p-1}$$

is such that each minor block in $X'_{\alpha\beta}$ moves through its orbit in $X'_{\alpha\beta}$ at least once.

By the definition of M and the remarks made above we see that $M(X)$, $X \in L(\Omega_n)$, is found as follows. Let X be partitioned into major and minor blocks. Consider the orbit of the minor blocks in each major block. Sum the blocks in each orbit with sufficiently many copies in order that there are p blocks. Then divide the sum by p and replace each block in the orbit by this block. From this we see that $X \in M[L(\Omega_n)]$ if and only if

1. $X \in L(\Omega_n)$.
2. If $X_{\gamma_1\sigma_1}$ and $X_{\gamma_2\sigma_2}$ are in the same major block and in the same orbit in the major block, then they are equal.

We now find necessary and sufficient conditions for X to be a vertex of $M[L(\Omega_n)]$.

DEFINITION. If $X_{\alpha_1\beta_1}, X_{\alpha_1\beta_2}, \cdots, X_{\alpha_m\beta_m} = X_{\alpha_1\beta_1}$ are major blocks of X, $X \in M[L(\Omega_n)]$ and each $X_{\alpha_k\beta_k}$ $(k = 1, 2, \cdots, m)$, $X_{\alpha_k\beta_{k+1}}$ $(k = 1, 2, \cdots, m-1)$ has exactly one positive minor block orbit, then

$$X_{\alpha_1\beta_1}, X_{\alpha_1\beta_2}, \cdots, X_{\alpha_m\beta_m}$$

is an orbital block loop in X.

THEOREM 5.1. *$X \in M[L(\Omega_n)]$ is a vertex if and only if*

1. *there do not exist two different positive minor block orbits in any major block of X, and*
2. *there does not exist an orbital block loop in X.*

Proof. First suppose $X \in M[L(\Omega_n)]$ and X has two positive block orbits in a major block $X_{\alpha\beta}$ of X. Then we add $\varepsilon > 0$ to each element in each block of one of these orbits and subtract ε from each element of each block in the other orbit. Call this matrix X'. To generate the matrix X'' replace ε by $-\varepsilon$ in X'. Now for ε sufficiently small, X' and $X'' \in M[L(\Omega_n)]$. Since $X = \frac{1}{2}(X' + X'')$, X is interior and therefore if X is a vertex it must satisfy 1.

Now suppose $X \in M[L(\Omega_n)]$ satisfies 1 but not 2. This means X

has an orbital block loop, say $X_{\alpha_1\beta_1}, X_{\alpha_1\beta_2}, \cdots, X_{\alpha_m\beta_m} = X_{\alpha_1\beta_1}$. Each of these major blocks has a positive orbit by definition. Flatten each major block; i.e., if $X_{\alpha\beta}$ is a block in the orbital block loop and has s different orbits, divide the element c in the positive orbit by s and replace all elements in the major block by c/s. If we call this matrix X' then $X' \in M[L(\Omega_n)]$. We may now use the scheme of Theorem 2.2 to alternately add and subtract $\varepsilon > 0$ from this major block loop, thereby generating X_1' and $X_2' \in M[L(\Omega_n)]$ and $X' = \frac{1}{2}(X_1' + X_2')$. Now absorb the flat major blocks back into the original orbits, i.e., if $X_{\alpha\beta}$ is a major block in the orbital block loop with s different orbits then replace each element c in each block of the original positive orbit by sc. Put zero blocks in all other orbits in this major block. Doing this to X', X_1', and X_2' we generate X, X_1, and X_2, respectively. Note $X_1, X_2 \in M[L(\Omega_n)]$. Further $X = \frac{1}{2}(X_1 + X_2)$. Hence X is interior.

Finally suppose X satisfies 1 and 2. Suppose that there exist $X_1, X_2 \in M[L(\Omega_n)]$ so that $X = \frac{1}{2}(X_1 + X_2)$. We may suppose X_1 and X_2 have the same zero pattern as X. If $X_1 \neq X_2$ and X_1, X_2 satisfy 1 we can see by an argument similar to Theorem 2.2, that X has an orbital block loop. This contradicts X having property 2. Hence we see that X is a vertex.

Using this theorem and the remarks preceeding this theorem we see that we have characterized the vertices of $M[L(\Omega_n)]$.

EXAMPLE.

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} X \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

There are three orbits for X given in the following diagram.

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 1 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

They are the positions occupied by 1, 2 and 3 respectively. Consider the vertices of $L(\Omega_n)$. Using 1 of Theorem 5.1 we see

(a)
$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

has a one in each orbit; hence

$$M\left[\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}\right] = \begin{pmatrix}1/3 & 1/3 & 1/3\\ 1/3 & 1/3 & 1/3\\ 1/3 & 1/3 & 1/3\end{pmatrix}$$

which is interior.

(b)
$$\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}$$

has 3 ones in the same orbit, hence

$$M\left[\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}\right] = \begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}$$

which is a vertex. The other vertices are

$$\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{pmatrix} \quad\text{and}\quad \begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}.$$

6. **General solutions of $AXB = X$.** Partition X into the major and minor blocks. Since $AXB = X$ would imply $A^pXB^p = X$ we see that each minor block of X must be flat. If we add the further condition that minor blocks on the same orbit are all equal then we see from § 5 that X is a solution and all solutions are of this form.

7. **General remarks.** It is interesting to note that in order to obtain solutions of $AXB = X$ it is only necessary to know the block form of A and B, i.e., if A_1 is doubly stochastic and has the same block form as A and B_1 is doubly stochastic and has the same block form as B then $AXB = X$ if and only if $A_1XB_1 = X$.

From § 4, property 5, we see that if A and B are irreducible, where the index of imprimitivity of A and the index of imprimitivity of B are relatively prime, then J_n is the only doubly stochastic solution. The only general solution is flat. This follows since there is only one orbit in X. Each block in the orbit is flat and all blocks in the orbit are equal.

Finally we point out that our result can be extended to a more general setting by considering the following result due to Sinkhorn (7):

THEOREM. *Let D be the set of all $n \times n$ matrices with row and column sums equal to* 1, *M_{n-1} the set of $(n-1) \times (n-1)$ matrices.*

Let $R = 1 \oplus M_{n-1}$. *Then there is a nonsingular matrix* P *so that* $PDP^{-1} = R$.

From this we know that if A_1 and A_2 are $(n-1) \times (n-1)$ matrices then there are nonsingular matrices P and Q so that $P^{-1}(1 \oplus A_2)P$ and $Q(1 \oplus A_1)Q^{-1}$ have row and column sums equal to 1. If $P^{-1}(1 \oplus A_2)P$ and $Q(1 \oplus A_1)Q^{-1}$ are nonnegative and real and hence doubly stochastic, then since

$$A_1XA_2 = X$$

if and only if

$$(1 \oplus A_1)(1 \oplus X)(1 \oplus A_2) = 1 \oplus X$$

if and only if

$$Q(1 \oplus A_1)Q^{-1}Q(1 \oplus X)PP^{-1}(1 \oplus A_2)P = Q(1 \oplus X)P ,$$

we can also find the solutions to the matrix equation

$$A_1XA_2 = X .$$

References

1. R. A. Brualdi, *Convex sets of nonnegative matrices*, to appear.
2. H. O. Foulkes, *Rational solutions of the matrix equation $XA = BX$*, Proceedings of the London Mathematical Society (2), **50** (1948), 196–209.
3. F. R. Gantmacher, *The Theory of Matrices*, Chelsea Publishing Co., New York, 1960.
4. W. B. Jurkat and H. J. Ryser, *Term ranks and permanents of nonnegative matrices*, J. of Algebra, **5**, #3, March 1967.
5. Marvin Marcus and Henryk Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Boston, 1964.
6. W. V. Parker, *The Matrix Equation $AX = XB$*, Duke Math. J. **17** (1950), 43–51.
7. Richard Sinkhorn, *On the Factor Space of the Complex Doubly Stochastic Matrices*. Abstract 62T-243, Notices American Mathematical Society, 9, (1962), 334–335.

Received February 17, 1970.

TEXAS A&M UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

EXPANSIVE AUTOMORPHISMS OF BANACH SPACES, II

JAMES H. HEDLUND

An automorphism of a complex Banach space is shown to be uniformly expansive if and only if its approximate point spectrum is disjoint from the unit circle.

The problem of giving a spectral characterization of the property that an operator be uniformly expansive was investigated in [2], but the theorem stated above was obtained only for automorphisms of a Hilbert space. The proof given in this note is both more general and more transparent than the special version. We also note some topological properties of the various classes of expansive operators in the space of all invertible operators.

1. Uniformly expansive automorphisms.

If T is an automorphism (a bounded, invertible, linear operator) on a complex Banach space X denote its spectrum by $\Lambda(T)$, its compression spectrum by $\Gamma(T)$, its approximate point spectrum by $\Pi(T)$, and its point spectrum by $\Pi_0(T)$. Denote the unit circle $\{\lambda: |\lambda| = 1\}$ in the complex plane by C. The automorphism T is said to be

(1) expansive if for each $x \in X$ with $||x|| = 1$ there exists some non-zero integer n with $||T^n x|| \geqq 2$;

(2) uniformly expansive if there exists some positive integer n such that if $x \in X$ with $||x|| = 1$ then either $||T^n x|| \geqq 2$ or $||T^{-n}x|| \geqq 2$;

(3) hyperbolic if there exists a splitting $X = X_s \oplus X_u$, $T = T_s \oplus T_u$, where X_s and X_u are closed T-invariant linear subspaces of X, $T_s = T|X_s$ is a proper contraction, and $T_u = T|X_u$ is a proper dilation.

A discussion of these classes of automorphisms may be found in [2].

It is well-known [2, Lemma 1] that an automorphism T is hyperbolic if and only if $\Lambda(T) \cap C = \varnothing$. The principal result weakens both conditions.

THEOREM 1. *Let T be an automorphism of a complex Banach space X. Then T is uniformly expansive if and only if $\Pi(T) \cap C = \phi$.*

The proof requires the Banach space version of an interesting numerical lemma.

LEMMA 1. *Given any complex numbers $c_1, \cdots, c_s$ there exists $\lambda \in C$ such that $\sum_{j=1}^{s} \lambda^j c_j \geq 0$.*

Proof. [2, Lemma 2]

LEMMA 2. *Given any complex numbers $c_{-r}, \cdots, c_s$ with $c_0 \neq 0$ there exists $\lambda \in C$ such that $|\sum_{j=-r}^{s} \lambda^j c_j| \geq |c_0|$.*

Proof. We may assume that $c_0 > 0$: otherwise set $d_j = (\bar{c}_0/|c_0|)\, c_j$ and proceed. Let $f(\lambda) = \sum_{j=1}^{s} \lambda^j c_j$, $g(\lambda) = \sum_{j=-r}^{-1} \lambda^j c_j$, and $h(\lambda) = \sum_{j=1}^{r} \lambda^j \bar{c}_{-j}$. Since $\lambda^{-j} = (\bar{\lambda})^j$ for $\lambda \in C$ it follows that $\operatorname{Re} g(\lambda) = \operatorname{Re} h(\lambda)$, and therefore $\operatorname{Re}[f(\lambda) + g(\lambda)] = \operatorname{Re}[f(\lambda) + h(\lambda)]$. Now $f(\lambda) + h(\lambda)$ is a polynomial in λ; by Lemma 1 there exists $\lambda \in C$ such that $f(\lambda) + h(\lambda) \geq 0$. Thus $f(\lambda) + h(\lambda) + c_0 \geq c_0$, and

$$\left| \sum_{j=-r}^{s} \lambda^j c_j \right| \geq \operatorname{Re}\left(\sum_{j=-r}^{s} \lambda^j c_j \right) = \operatorname{Re}[f(\lambda) + h(\lambda) + c_0] \geq c_0 .$$

LEMMA 3. *Given any vectors $x_{-r}, \cdots, x_s$ in a Banach space X with $x_0 \neq 0$ there exists $\lambda \in C$ such that*

$$\left\| \sum_{j=-r}^{s} \lambda^j x_j \right\| \geqq \| x_0 \| .$$

Proof. By the Hahn-Banach Theorem choose $x^* \in X^*$ with $\| x^* \| = 1$ and $x^*(x_0) = \| x_0 \|$. It suffices to find $\lambda \in C$ with

$$\left| x^*\left(\sum_{j=-r}^{s} \lambda^j x_j \right) \right| \geqq | x^*(x_0) | .$$

Set $c_j = x^*(x_j)$ and apply Lemma 2: there exists $\lambda \in C$ such that

$$\left| x^*\left(\sum_{j=-r}^{s} \lambda^j x_j \right) \right| = \left| \sum_{j=-r}^{s} \lambda^j c_j \right| \geqq | c_0 | = | x^*(x_0) | .$$

Proof of Theorem 1. Necessity is proved in [2, Theorem 1]. To prove sufficiency, suppose that T is not uniformly expansive. Then for each positive integer n there exists $x_n \in X$ with $\| x_n \| = 1$ and $\max \{\| T^n x_n \|, \| T^{-n} x_n \|\} < 2$. For infinitely many n we produce a vector $y_n \in X$ and a number $\lambda_n \in C$ such that $\| (T - \lambda_n^{-1})\, y_n \| / \| y_n \| \to 0$. This will suffice. In fact, if $\mu \in C$ is a limit point of $\{\lambda_n^{-1}\}$ choose a subsequence $\{\lambda_m^{-1}\}$ of $\{\lambda_n^{-1}\}$ with $\lambda_m^{-1} \to \mu$. Then

$$\| (T - \mu)\, y_m \| / \| y_m \| \leqq \| (T - \lambda_m^{-1})\, y_m \| / \| y_m \| + | \lambda_m^{-1} - \mu | .$$

The right-hand side approaches 0 as $m \to \infty$, so that $\mu \in \Pi(T)$.

To construct y_n we must consider two cases. Define

$$\phi(n) = \max_{k=-n,0} \sup_{\lambda \in C} \left\| \sum_{j=k}^{k+n-1} \lambda^j T^j x_n \right\| .$$

Case 1. $\phi(n)$ is unbounded. Fix n, choose k where the maximum in the definition of ϕ is attained, and let λ_n be the $\lambda \in C$ where the supremum is attained. Define

$$y_n = \sum_{j=k}^{k+n-1} \lambda_n^j T^j x_n$$

so that $\| y_n \| = \phi(n)$. Now

$$(T - \lambda_n^{-1})\, y_n = \lambda_n^{n-1}\, T^n x_n - \lambda_n^{-1}\, x_n \text{ if } k = 0 \;,$$

and

$$(T - \lambda_n^{-1})\, y_n = \lambda_n^{-1} x_n - \lambda_n^{-n-1} T^{-n} x_n \text{ if } k = -n \;.$$

In either event,

$$\| (T - \lambda_n^{-1})\, y_n \| \leqq 3. \quad \text{Thus } \| (T - \lambda_n^{-1})\, y_n \| / \| y_n \| \leqq 3/\phi(n) \;.$$

Since $\phi(n)$ is unbounded, $3/\phi(n_j) \to 0$ for some subsequence $n_j \to \infty$.

Case 2. $\phi(n)$ is bounded. Assume that $\phi(n) \leqq A$ for all n and define

$$y_n = \sum_{j=-n}^{-1} (n + 1 + j)\, \lambda_n^j T^j x_n + \sum_{j=0}^{n-1} (n - j)\, \lambda_n^j T^j x_n \;,$$

where we choose $\lambda_n \in C$ by Lemma 3 to insure that $\| y_n \| \geqq n$, the norm of the term with index 0.

$$\begin{aligned} \| (T - \lambda_n^{-1})\, y_n \| &= \left\| -\sum_{j=-n}^{-1} \lambda_n^{j-1}\, T^j x_n + \sum_{j=1}^{n} \lambda_n^{j-1}\, T^j x_n \right\| \\ &\leqq \left\| \sum_{j=-n}^{-1} \lambda_n^j\, T^j x_n \right\| + \left\| T\Big(\sum_{j=0}^{n-1} \lambda_n^j T^j x_n \Big) \right\| \\ &\leqq A(1 + \| T \|) \;. \end{aligned}$$

Hence

$$\| (T - \lambda_n^{-1})\, y_n \| / \| y_n \| \leqq A(1 + \| T \|)/n \to 0 \;.$$

Note that the hypothesis that T is not uniformly expansive is not used in Case 2. But it is easy to see directly (by Lemma 3) that T is not uniformly expansive if $\phi(n)$ is bounded. Note also that it follows immediately from Theorem 1 that a hyperbolic automorphism is uniformly expansive.

2. Density. Denote the class of all hyperbolic automorphisms of a fixed Banach space X by $\mathscr{H}$, of uniformly expansive by $\mathscr{U}\mathscr{E}$, of expansive by $\mathscr{E}$, of all automorphisms by $\mathscr{I}$, and of all bounded linear operators by $\mathscr{B}$. If dim $X < \infty$ then $\mathscr{H} = \mathscr{U}\mathscr{E} = \mathscr{E}$ and is precisely the class of all automorphisms whose spectrum is disjoint from C. In general the situation is much different.

THEOREM 2. *Let X be separable infinite dimensional Hilbert space. Then:*

(1) $\mathscr{H} \subset \mathscr{U}\mathscr{E} \subset \mathscr{E} \subset \mathscr{I} \subset \mathscr{B}$;

(2) *$\mathscr{H}$ and $\mathscr{U}\mathscr{E}$ are open (in $\mathscr{B}$, in the uniform operator topology) but $\mathscr{E}$ is not;*

(3) *no class is dense in the next larger.*

The tools necessary for the proof are two results on semicontinuity of pieces of the spectrum due to Halmos and Lumer.

THEOREM A. [4, Theorem 2] *$\Pi(T)$ and $\Lambda(T)$ are upper semicontinuous: to every $T \in \mathscr{B}$ and every open set G containing $\Pi(T)$ [respectively, $\Lambda(T)$] there corresponds a positive number ε such that $\Pi(S) \subset G$ [$\Lambda(S) \subset G$] whenever $\|S - T\| < \varepsilon$.*

THEOREM B. [4, Theorem 3] *$\Lambda(T)\backslash\Pi(T)$ is lower semicontinuous: to every $T \in \mathscr{B}$ and every compact set K contained in $\Lambda(T)\backslash\Pi(T)$ there corresponds a positive number ε such that $K \subset \Lambda(S)\backslash\Pi(S)$ whenever $\| S - T \| < \varepsilon$.*

Proof of Theorem 2. (2) If $T \in \mathscr{H}$ then $\Lambda(T) \cap C = \varnothing$. By semicontinuity, $\Lambda(S) \cap C = \varnothing$ for S sufficiently near T. Since $\mathscr{I}$ is open, $S \in \mathscr{H}$. The proof for $\mathscr{U}\mathscr{E}$ is identical. To see that $\mathscr{E}$ is not open fix an orthonormal base $\{e_n\}_1^\infty$ and let T be the diagonal operator $Te_n = n/(n+1)\, e_n$. T is expansive [2, Example 2]. Given $\varepsilon > 0$ let $Se_n = Te_n$ for $|1 - n/(n+1)| \geqq \varepsilon$ and $Se_n = e_n$ otherwise. Then $\| S - T \| < \varepsilon$ but S is not expansive since $1 \in \Pi_0(S)$.

(3) $\mathscr{I}$ is not dense in $\mathscr{B}$: [3, Problem 109].

$\mathscr{E}$ is not dense in $\mathscr{I}$: let $\{e_n\}_{-\infty}^{\infty}$ be an orthonormal base and let T be the backward bilaterial weighted shift defined by $Te_n = 2e_{n-1}$ for $n \geqq 1$, $Te_n = 1/2\, e_{n-1}$ for $n \leqq 0$. Then [2, Example 4]

$$\Pi_0(T) = \{1/2 < |\lambda| < 2\}$$

so that $T \notin \mathscr{E}$. Now $\Lambda(T^*)\backslash\Pi(T^*) = \{1/2 < |\lambda| < 2\}$; by Theorem B if $\| S^* - T^* \|$ is small then $C \subset \Lambda(S^*)\backslash\Pi(S^*) \subset \Gamma(S^*)$. Hence $C \subset \Pi_0(S)$ and $S \notin \mathscr{E}$.

$\mathscr{H}$ is not dense in $\mathscr{U}\mathscr{E}$: in fact $\mathscr{U}\mathscr{E}\backslash\mathscr{H}$ is open. If $T \in \mathscr{U}\mathscr{E}\backslash\mathscr{H}$ then $\Pi(T) \cap C = \varnothing$ but $\Lambda(T) \cap C \neq \varnothing$. So there exists a compact set $K \subset C \cap [\Lambda(T)\backslash\Pi(T)]$. By Theorem B, $K \subset \Lambda(S)$ for $\| S - T \|$ small, so that $S \notin \mathscr{H}$.

$\mathscr{U}\mathscr{E}$ is not dense in $\mathscr{E}$: let X be represented as H^2 (of the unit circle) and let T be the multiplication operator $Tf(e^{it}) = (e^{it} + 3/2) f(e^{it})$. Let $A_r = \{|\lambda - 3/2| \leqq r\}$. Either direct calculation or appeal to the spectral properties of Toeplitz operators ([1], for instance) shows that $\Lambda(T) = A_1$, $\Pi_0(T) = \varnothing$, $\Pi(T) = \text{bdy } A_1$, and $\Gamma(T) = \text{int } A_1$. By Theorems A and B there exists $\varepsilon > 0$ such that if $\| S - T \| < \varepsilon$ then $A_{3/4} \subset \Gamma(S)$ and $\Lambda(S) \subset A_{3/2}$. Now the arc $\alpha(t) = e^{it}$, $0 \leqq t \leqq \pi/2$, on the unit circle has $\alpha(0) \in A_{3/4}$ and $\alpha(\pi/2) \notin A_{3/2}$. Thus $\alpha(t) \in \text{bdy } \Lambda(S)$ for some t; hence $\Pi(S) \cap C \neq \varnothing$ and $S \notin \mathscr{U}\mathscr{E}$. To verify that T is expansive let $a \in [0, \pi]$ with $| e^{ia} + 3/2 | = 1$. Fix $f \in H^2$ with $\| f \|_2 = 1$. Then either

$$1/2\pi \int_{-a}^{a} | f(e^{it}) |^2 \, dt \geqq 1/2 \text{ or } 1/2\pi \int_{a}^{2\pi - a} | f(e^{it})|^2 \, dt \geqq 1/2 .$$

If the former holds choose $-a < b < c < a$ with

$$1/2\pi \int_{b}^{c} | f(e^{it}) |^2 \, dt \geqq 1/4 ,$$

let $K = \min \{| e^{ib} + 3/2 |, | e^{ic} + 3/2 |\} > 1$, and choose an integer n with $K^n \geqq 4$. If $m \geqq n$

$$\begin{aligned} \| T^m f \|_2^2 &= 1/2\pi \int_0^{2\pi} | e^{it} + 3/2 |^{2m} | f(e^{it})|^2 \, dt \\ &\geqq 1/2\pi \int_b^c | e^{it} + 3/2|^{2m} | f(e^{it})|^2 \, dt \\ &\geqq K^{2m} 1/2\pi \int_b^c | f(e^{it}) |^2 \, dt \\ &\geqq 4 . \end{aligned}$$

If the other alternative holds then $\| T^{-m} f \|_2 \geqq 2$ for large m. Hence T is expansive.

REFERENCES

1. P. L. Duren, *On the spectrum of a Toeplitz operator*, Pacific J. Math., **14** (1964), 21-29.
2. M. Eisenberg and J. H. Hedlund, *Expansive automorphisms of Banach spaces*, Pacific J. Math., **34** (1970), 647-656.
3. P. R. Halmos, *A Hilbert Space Problem Book*, Van Nostrand, Princeton, 1967.
4. P. R. Halmos and G. Lumer, *Square roots of operators*, II, Proc. Amer. Math. Soc., **5** (1954), 589-595.

Received August 3, 1970.

UNIVERSITY OF MASSACHUSETTS, AMHERST

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

THE p-PARTS OF CHARACTER DEGREES IN p-SOLVABLE GROUPS

I. M. ISAACS

Let G be a finite group and $Irr(G)$ the set of irreducible complex characters of G. Fix a prime integer p and let $e(G)$ be the largest integer such that $p^{e(G)}$ divides $\chi(1)$ for some $\chi \in Irr(G)$. The purpose of this paper is to obtain information about the structure of G, and in particular about a Sylow p-subgroup of G, from a knowledge of $e(G)$. If G is solvable, we obtain the bound $2e(G)+1$ for the derived length of an S_p subgroup of G. We also obtain some information about the normal structure of G in terms of $e(G)$.

When $e(G) = 0$, our result is equivalent to the theorem of N. Ito which asserts that G has a normal abelian Sylow p-subgroup. Actually, Ito's result, [7], holds for p-solvable groups. This may readily be proved by induction on the group order, as follows. The hypothesis $e(G) = 0$ is inherited by factor groups and by normal subgroups and it follows easily that a minimal counterexample, G, has a normal p-complement, H. Now let $\chi \in Irr(G)$. It follows from Clifford's theorem that $t | \chi(1)$, where t is the index in G of the inertia group of an irreducible constituent of the restriction χ_H. Since t is a power of p, we have $t = 1$, and every irreducible constituent of χ_H is invariant in G. It follows by Frobenius reciprocity that every irreducible character of H is invariant in G. Now Lemma 2.1 of [4] applies to yield the result.

Although it might be conjectured that our present bounds hold for all p-solvable groups when $e(G) > 0$, the proofs given here fail even when $e(G) = 1$. However in this case, we do obtain a result which is valid for p-solvable G with $p > 3$, and shows that $\mathcal{O}_p(G)$ is either abelian or else is a Sylow subgroup of G.

1. The following lemma is well known and will be used repeatedly. Since its proof is quite short, we present it here.

LEMMA 1.1. *Let $N \lhd G$ and $\chi \in Irr(G)$. Suppose θ is an irreducible constituent of χ_N. Let $T = \mathscr{I}_G(\theta)$, the inertia group of θ. Then there exists a unique irreducible constituent ψ of χ_T such that θ is a constituent of ψ_N. Furthermore $\chi = \psi^G$ and $[\chi_N, \theta] = [\psi_N, \theta]$.*

Proof. Choose any irreducible constituent ψ of χ_T such that θ is a constituent of ψ_N. By Clifford's theorem, $\chi_N = a \sum_{i=1}^{t} \theta_i$ where $\theta_1 = \theta$

and $\chi(1) = at\theta(1)$. We have $t = |G: T|$ and $\psi_N = a_0 \theta$, $a_0 \leqq a$. Now χ is a constituent of ψ^G and so

$$\chi(1) \leqq \psi^G(1) = t\psi(1) = ta_0\theta(1) \leqq ta\theta(1) = \chi(1) .$$

We have equality throughout, so that $\chi(1) = \psi^G(1)$ and $a = a_0$. Thus $\chi = \psi^G$ and $[\chi_N, \theta] = a = a_0 = [\psi_N, \theta]$. The uniqueness of ψ also follows from $a = a_0$.

If $e(G) = e$ and $N \lhd G$, let $\theta \in Irr(N)$ and $T = \mathscr{I}_G(\theta)$. Suppose that $|G: T|_p = p^r$, where n_p denotes the p-part of the integer n. Let ψ be any irreducible constituent of θ^T, and let χ be an irreducible constituent of ψ^G. Then by Frobenius reciprocity and Lemma 1.1, it follows that $\chi = \psi^G$ and hence $\psi(1)_p \leqq p^{e-r}$. It does not follow, however, that $e(T) \leqq e - r$. We wish to prove our results by induction in a manner similar to this and hence we define a quantity which "inducts" properly.

DEFINITION 1.2. Let $N \lhd G$ and $\theta \in Irr(N)$. Suppose θ is invariant in G. Then $e(G, N, \theta) = e$ is the largest integer such that $p^e | (\chi(1)/\theta(1))$ for some irreducible constituent χ of θ^G.

Note that $e(G, 1, 1) = e(G)$ and that if $N \subseteqq H \lhd G$, then

$$e(H, N, \theta) \leqq e(G, N, \theta) .$$

The following is immediate.

COROLLARY 1.3. *Suppose $e(G, N, \theta) = e$ and $N \subseteqq M \lhd G$. Let ψ be an irreducible constituent of θ^M and let $p^f = (\psi(1)/\theta(1))_p$. Set $T = \mathscr{I}_G(\psi)$ and $p^r = |G: T|_p$. Then $e(T, M, \psi) \leqq e - f - r$.*

It would suffice for our purposes to show that if $N \lhd G$, G/N is solvable and $e(G, N, \theta) = e$ for some $\theta \in Irr(N)$, then the derived length of an S_p subgroup of G/N is bounded by a function of e. We in fact will prove this for certain special characters θ and also for certain groups G/N. In order to prove results like these, it is necessary to be able to produce irreducible characters of degrees divisible by "large" powers of p. This is done using the following result of Gallagher ([1], Theorem 2).

PROPOSITION 1.4. *Let $N \lhd G$ and suppose $\chi \in Irr(G)$ and*

$$\chi_N = \theta \in Irr(N) .$$

Then the irreducible constituents of θ^G are uniquely of the form $\beta\chi$ where $\beta \in Irr(G/N)$ is viewed as a character of G. For all such β,

$\beta\chi$ is irreducible.

LEMMA 1.5. *Let $N \lhd G$, $N \subseteqq H \lhd G$ with G/H a p-group. Let $\theta \in Irr(N)$ be invariant in G. If $e(G, N, \theta) = e(H, N, \theta)$, then G/H is abelian. If $e(G, N, \theta) > e(H, N, \theta)$, then there exists $L \lhd G$ with $H \subseteqq L$, G/L abelian and $e(L, N, \theta) < e(G, N, \theta)$.*

Proof. Let $K \lhd G$, $K \supseteqq H$ be minimal such that

$$e(K, N, \theta) = e(G, N, \theta) = e .$$

Let ψ be an irreducible constituent of θ^K with $p^e | (\psi(1)/\theta(1))$. Let χ be any irreducible constituent of ψ^G. Then $p^{e+1} \nmid (\chi(1)/\theta(1))$ and therefore $p \nmid (\chi(1)/\psi(1))$. Since G/K is a p-group, $\chi(1)/\psi(1)$ is a power of p and thus $\chi(1) = \psi(1)$ and $\chi_K = \psi \in Irr(K)$. Let β be an arbitrary irreducible character of G/K. By Proposition 1.4, $\beta\chi$ is an irreducible constituent of ψ^G and we may apply the above reasoning to $\beta\chi$ in place of χ. Hence $(\beta\chi)(1) = \psi(1) = \chi(1)$ and $\beta(1) = 1$. Thus G/K is abelian. If $e(G, N, \theta) = e(H, N, \theta)$ then $H = K$ and the first statement is proved.

Otherwise $K > H$ and we may choose $L \lhd G$ with $H \subseteqq L < K$ and $|K: L| = p$. By the choice of K, $e(L, N, \theta) < e$ and hence ψ_L is reducible. Therefore $\chi_L = \psi_L$ is a sum of p distinct irreducible characters, conjugate in K. Let φ be one of these characters and put $T = \mathscr{I}_G(\varphi)$ so $|G: T| = p$. Thus $T \lhd G$ and $G' \subseteqq T$. We also have $G' \subseteqq K$ and $K \cap T = L$ so that G/L is abelian and the result follows.

LEMMA 1.6. *Let $N \lhd G$ and suppose that G/N is p-solvable with p'-length $\leqq 1$. Suppose $\theta \in Irr(N)$ and is invariant in G with*

$$e(G, N, \theta) = e .$$

Then the derived length of an S_p subgroup of G/N is $\leqq e + 2$. If G/N is a p-group, d.l.$(G/N) \leqq e + 1$.

Proof. Let $K/N = \mathscr{O}^p(G/N)$, the minimum normal subgroup with factor group a p-group. By hypothesis, K/N has the normal S_p subgroup P/N. Suppose $e(K, N, \theta) < e$. Then by Lemma 1.5, there exists $L \lhd G$, $K \subseteqq L$ with G/L abelian and $e(L, N, \theta) < e$. Both statements now follow by induction on $|G: N|$. Suppose then $e(K, N, \theta) = e$. Then G/K is abelian by Lemma 1.5. If $K = N$, then d.l.$(G/N) \leqq e + 1$ is trivial. Suppose, then, $K > N$. Then $P < K$ and $e(P, N, \theta) \leqq e$ so by induction, d.l.$(P/N) \leqq e + 1$. Since G/K is abelian, the derived length of an S_p subgroup of G/N is $\leqq e+2$. However, since $K>N$, G/N is not a p-group and the proof is complete.

2. Suppose $N \lhd G$ and $\theta \in Irr(N)$ and is invariant in G. It will occasionally be necessary in what follows to be able to extend θ to an irreducible character of G. This is, of course, not always possible. We discuss some sufficient conditions below.

Given any character χ of a finite group G, we define the determinant $\det \chi = \lambda$ to be the linear character of G given by

$$\lambda(g) = \det \mathfrak{X}(g) ,$$

where $\mathfrak{X}$ is any representation affording χ. Let $o(\chi)$ denote the order of λ as an element of the group of linear characters of G. Clearly $o(\chi) = o(\lambda) = |G\colon \ker \lambda|$. Gallagher [1] has shown that if $\theta \in Irr(N)$, $N \lhd G$, θ is invariant G and $(\theta(1), |G\colon N|) = 1$, then θ is extendible to G if and only if $\det \theta$ is extendible to G. Furthermore, Gallagher proved that if $\lambda = \det \theta$ and μ is an extension of λ, then there is a unique extension χ of θ with $\det \chi = \mu$. Since θ is invariant in G, so is λ and it follows that $\ker \lambda \lhd G$ and $N/\ker \lambda \subseteqq \mathbf{Z}(G/\ker \lambda)$. If $(o(\theta), |G\colon N|) = 1$, then $N/\ker \lambda$ is a direct factor of $G/\ker \lambda$ and hence there is a unique extension μ of λ to G with $o(\mu) = o(\lambda)$. Summarizing these results, we obtain the following.

PROPOSITION 2.1. *Let $N \lhd G$ and let $\theta \in Irr(N)$ with θ invariant in G. Suppose $o(\theta)$ and $\theta(1)$ are both relatively prime to $|G\colon N|$. Then there exists a unique extension, $\hat{\theta}$, of θ to G with $o(\hat{\theta}) = o(\theta)$.*

DEFINITION 2.2. Let $\chi \in Irr(G)$. Then χ is a *p-character* of G if $\chi(1)$ and $o(\chi)$ are powers of p.

LEMMA 2.3. *Let $N \lhd G$ and suppose $\theta \in Irr(N)$ is a p-character which is invariant in G. Suppose G/N has a normal p-complement K/N and that $\mathcal{O}_p(G/N) = 1$. Then* $\text{d.l.}(G/K) \leqq e(G, N, \theta) = e$.

Proof. Use induction on $|G\colon N|$. Suppose $e>0$. If $e(K, N, \theta) = e$, then by Lemma 1.5, G/K is abelian and we are done. Otherwise, $e(L, N, \theta) < e$ for some $L \lhd G$ with $K \subseteqq L$ and G/L abelian. By induction, $\text{d.l.}(L/K) \leqq e-1$ and the result follows. The only remaining case is where $e=0$. Here we must show that $K=G$.

Since θ is a p-character of N, there is an extension $\hat{\theta}$ of θ to K. Let χ be any irreducible constituent of $\hat{\theta}^G$. Then $\chi(1)/\theta(1)$ is a power of p and θ is a constituent of χ_N so $\chi(1) = \theta(1)$ since $e(G, N, \theta) = 0$. Thus if β is any irreducible character of G/N, $\beta\chi \in Irr(G)$ and since θ is a constituent of $(\beta\chi)_N$, it follows that $p \nmid \beta(1)$. Hence $e(G/N) = 0$

and therefore G/N has a normal S_p subgroup. Since $\mathcal{O}_p(G/N) = 1$, $p \nmid |G: N|$ and thus $K = G$ and the proof is complete.

The following lemma will be used to prove that a given character is a p-character.

LEMMA 2.4. *Let $N \lhd G$ and suppose that G/N has no proper normal subgroup of p'-index. Let $\chi \in Irr(G)$ and suppose θ is an irreducible constituent of χ_N and $o(\theta)$ is a power of p. Then* $o(\chi)$ is *a power of p.*

Proof. Let $\lambda = \det \chi$, and let $K = \{g \in G \mid \lambda(g)^{p^e} = 1$ for some $e \geqq 0\}$. It suffices to show that $K = G$. Clearly, $K \lhd G$ is a subgroup, and $p \nmid |G: K|$. The result will follow if we show $N \subseteq K$. Now $\chi_N = a\Sigma\theta_i$ by Clifford's theorem, where the θ_i are all conjugate to θ. Let $\mu_i = \det \theta_i$ so that $\lambda_N = (\Pi\mu_i)^a$. Each μ_i has order equal to $o(\theta)$ which is a power of p. Therefore, for suitable e, and for $x \in N$, we have $\mu_i(x)$ is a p^e-th root of 1. It follows that $N \subseteq K$ and the proof is complete.

3. We define functions u, v as follows.

DEFINITION 3.1. Let u, v be functions from the set of nonnegative integers into the same set with ∞ adjoined, where $u(e) =$ maximum derived length of an S_p subgroup of G/N where G is a finite group, $N \lhd G$, G/N is solvable and there exists a p-character, θ, of N, invariant in G and such that $e(G, N, \theta) \leqq e$. Set $u(e) = \infty$ if there is no maximum. Define $v(e)$ similarly, except that only those situations are considered where $\mathcal{O}_p(G/N) = 1$.

LEMMA 3.2. *Let P be a p-group and suppose that $P_0 \subseteq P$ with $|P: P_0| = p^r$. Then* $\text{d.l.}(P) \leqq r + \text{d.l.}(P_0)$.

Proof. Use induction on r. The result is trivial if $r = 0$. Otherwise $P_0 < P$ and hence $P_0P' < P$ since $P' \subseteq \Phi(P)$, the Frattini subgroup of P. By induction, $\text{d.l.}(P_0P') \leqq (r - 1) + \text{d.l.}(P_0)$. However, $P_0P' \lhd P$ and P/P_0P' is abelian. The result follows.

LEMMA 3.3. *Let $N \subseteq H$ be normal subgroups of L. Assume $(|H: N|, |L: H|) = 1$. Let $\theta \in Irr(N)$ and suppose θ is extendible to H. If θ is invariant in L, then some extension of θ to H is also invariant in L.*

Proof. Let $\mathcal{S}$ be the set of extensions of θ to H, and let U be the group of linear characters of H/N. Then U acts on the set $\mathcal{S}$

by multiplication and by Proposition 1.4, this action is transitive. Set $A = L/H$. We have $U \cong H/H'N$ and thus $(|A|, |U|) = 1$. Clearly, A acts on $\mathscr{S}$ and on the group U and if $\chi \in \mathscr{S}$, $\lambda \in U$, then $(\chi\lambda)^a = \chi^a\lambda^a$ for all $a \in A$. Therefore Glauberman's Lemma (Theorem 4 of [2]) applies and hence A fixes some $\chi \in \mathscr{S}$. Thus χ is invariant in L.

Before going on to our main result, we digress briefly to give an application of some of the lemmas we have already accumulated.

COROLLARY 3.4. *Let $N \lhd G$ with G/N p-solvable. Suppose θ is a p-character of N which is invariant in G and that $e(G, N, \theta) = 0$. Then θ is extendible to G and G/N has a normal abelian S_p subgroup.*

Proof. If θ is extendible to G, then it follows from Proposition 1.4 that $e(G/N) = 0$ and hence G/N has a normal abelian S_p subgroup. We prove extendibility by induction on $|G: N|$. Let $M/N = \mathscr{O}^p(G/N)$. If $M < G$, then θ is extendible to $\psi \in Irr(M)$. Let χ be any irreducible constituent of ψ^G. Since G/M is a p-group, it follows that $\chi(1)/\psi(1)$ is a power of p. Since $e(G, N, \theta) = 0$, $\chi(1) = \psi(1)$ and the result follows.

Suppose then $M = G$ and let $V/N = \mathscr{O}^{p'}(G/N)$. Then $V < G$ and θ is extendible to V. Let $W/N = (V/N)'$. Then V/W is a p-group. Now if $x \in G$, then ψ^x is an extension of θ so $\psi^x = \lambda\psi$ for some linear character λ of G/N (Proposition 1.4). Then $\lambda_W = 1$ and $\psi^x_W = \psi_W$. Hence ψ_W is invariant in G and by Lemma 3.3 we may assume that ψ is invariant in G. By Lemma 2.4, ψ is a p-character of V and thus is extendible to G. The proof is complete.

THEOREM 3.5. *The functions u and v are finite valued, $v(0) = 0$, $u(0) = 1$ and*

$$v(e) \leqq \max_{0<f\leqq e} (f + u(e - f)) \text{ for } e > 0 \text{ and}$$
$$u(e) \leqq 1 + \max_{0<f\leqq e} (f + u(e - f)) \text{ for } e > 0 .$$

Proof. If u ever takes on the value ∞, choose $e \geqq 0$ minimal with $u(e) = \infty$. Otherwise pick e arbitrarily. Choose a group G, $N \lhd G$, θ a p-character of N, invariant in G with $e(G, N, \theta) \leqq e$. Let P/N be an S_p subgroup of G/N. If $e > 0$, write $b = \max\{f + u(e - f) \mid 0 < f \leqq e\}$. If $e = 0$, set $b = 0$. We claim that (a) if $\mathscr{O}_p(G/N) = 1$, then d.l.$(P/N) \leqq b$ and in any case (b) d.l.$(P/N) \leqq b + 1$. The proof will be complete when these claims are established. In particular, the inequality involving $v(e)$ will follow when (a) is proved. Note that when $e = 0$, the result follows from Corollary 3.4, however this case also follows from the general argument and we do not appeal to the previous result. We shall prove (a) and (b) by induction on $|G: N|$, for the fixed value of e chosen above.

Case 1. $\mathcal{O}_{p'}(G/N) > 1$. Let K/N be a minimal normal p'-subgroup of G/N so that K/N is an elementary abelian q-group for some $q \neq p$. Let $\hat{\theta}$ be the unique extension of θ to K with $o(\hat{\theta}) = o(\theta)$. Because of the uniqueness, $\hat{\theta}$ is invariant in G and by definition, $\hat{\theta}$ is a p-character. Clearly $e(G, K, \hat{\theta}) \leqq e$ and thus d.l.$(PK/K) \leqq b + 1$ by induction. Since $PK/K \cong P/N$, (b) follows in this case. If $\mathcal{O}_p(G/K) = 1$, then d.l.$(PK/K) \leqq b$ and (a) follows.

Assume that $\mathcal{O}_p(G/N) = 1$ but that $\mathcal{O}_p(G/K) = H/K > 1$. Let ψ be an irreducible constituent of θ^H with $(\psi(1)/\theta(1))_p = p^f$ as large as possible. Let φ be an irreducible constituent of ψ_K which is a constituent of θ^K. Since K/N is abelian, it follows from Proposition 1.4 that $\varphi = \hat{\theta}\lambda$ for a linear character λ of K/N. Thus $\varphi(1) = \theta(1)$ is a power of p. Since H/K is a p-group, $\psi(1)/\varphi(1)$ is a power of p and hence $\psi(1)$ is a power of p. We claim that ψ is a p-character of H. This will follow from Lemma 2.4 when we establish that H/N has no nontrivial p'-factor group.

Now $H'N \cap K \lhd G$ and by the minimality of K, we have either $H'N \cap K = N$ or $H'N \cap K = K$. In the first situation, $K/N \subseteqq \mathbf{Z}(H/N)$ and it follows that $\mathcal{O}_p(H/N) \cong H/K > 1$, a contradiction. Thus $H'N \cap K = K$. Since any p'-factor group of H/N is abelian, this shows that only the trivial one exists.

Let $T = \mathcal{I}_G(\psi)$ and set $p^r = |G: T|_p$. By Corollary 1.3,

$$e(T, H, \psi) \leqq e - f - r\,.$$

Let P_0/H be an S_p subgroup of T/H and assume that $P_0 \subseteqq PK$ since PK/H is an S_p subgroup of G/H. Now d.l.$(P_0/H) \leqq u(e - f - r)$ and $|PK: P_0| = p^r$ so that d.l.$(PK/H) \leqq r + u(e - f - r)$ by Lemma 3.2. We have $e(H, N, \theta) = f$ and $\mathcal{O}_p(H/N) = 1$ and hence $0 <$ d.l.$(H/K) \leqq f \leqq e$ by Lemma 2.3. It follows that d.l.$(PK/K) \leqq r + f + u(e - f - r) \leqq b$ and the proof of Case 1 is complete. In particular, since only Case 1 can occur when $\mathcal{O}_p(G/N) = 1$, we have shown that $v(e) \leqq b$.

Case 2. $\mathcal{O}_{p'}(G/N) = 1$. Let $H/N = \mathcal{O}_p(G/N) > 1$ and let

$$f = e(H, N, \theta)\,.$$

Since H/N is a p-group, we can pick $\psi \in Irr(H)$ with $\psi_N = p^f\theta$. Also ψ is a p-character by Lemma 2.4. Let $T = \mathcal{I}_G(\psi)$ and $p^r = |G: T|_p$. Reasoning exactly as before, we get

$$\text{d.l.}(P/N) \leqq r + u(e - f - r) + \text{d.l.}(H/N)\,.$$

By Lemma 1.6, d.l.$(H/N) \leqq f + 1$, and thus

$$\text{d.l.}(P/N) \leqq 1 + f + r + u(e - f - r)\,.$$

If $f + r > 0$, then $e > 0$ and we obtain d.l.$(P/N) \leqq b + 1$ and we are done in this case.

Assume then that $f = 0 = r$ for all irreducible constituents ψ of θ^H. From $f = 0$, it follows that θ is extendible to H and by Lemma 3.3, we may choose an extension ψ which is invariant in L where $L/H = \mathcal{O}_{p'}(G/H)$. Now let $T = \mathcal{I}_G(\psi)$. Since $r = 0$, we may assume $P \subseteqq T$. Also $L \subseteqq T$. We claim that $U/H = \mathcal{O}_p(T/H) = 1$. We have $[L, U] \subseteqq H$ and hence by Lemma 1.2.3 of [3], it follows that $U \subseteqq H$. Therefore, d.l.$(P/H) \leqq v(e)$ since $e(T, H, \psi) \leqq e$. By Lemma 1.6, d.l.$(H/N) \leqq 1$ and thus d.l.$(P/N) \leqq 1 + v(e)$. Since we have already shown that $v(e) \leqq b$, the result follows.

COROLLARY 3.6. *$v(e) \leqq 2e$, $u(e) \leqq 2e + 1$ for all $e \geqq 0$.*

Proof. Use induction on e. The Corollary is immediate if $e = 0$. For $e > 0$ we have $v(e) \leqq \max\{f + u(e - f) \mid 0 < f \leqq e\} \leqq \max\{f + 2(e - f) + 1\}$. This maximum occurs when $f = 1$ and yields $v(e) \leqq 2e$. Similarly $u(e) \leqq 2e + 1$.

4. Some improvement on the bounds of Theorem 3.5 can be obtained, especially for $e < p - 1$. We shall use Theorem B of Hall and Higman [3] and also the following result of Passman (Corollary 2.4 (i) of [8]).

PROPOSITION 4.1. *Let P be a p-group which acts faithfully on a solvable p'-group A. Suppose that every element of A lies in an orbit of size $\leqq p^e < p^p$ under the action of P. Then some element of A lies in a regular orbit and hence $|P| \leqq p^e$.*

LEMMA 4.2. *Let $N \subseteqq H$ be normal subgroups of L. Suppose H/N is solvable and that $(|L{:}\, H|, |H{:}\, N|) = 1$. Let $\theta \in Irr(N)$ and suppose $\mathcal{I}_L(\theta)$ covers L over H. Then some irreducible constituent ψ of θ^H is invariant in L.*

Proof. We use induction on $|H{:}\, N|$. The result is trivial if $N = H$. Let $M \lhd L$, $M < H$ be maximal such. By the Schur-Zassenhaus Theorem, applied to the group $\mathcal{I}_L(\theta)/N$ which has the normal Hall subgroup, $\mathcal{I}_H(\theta)/N$, we can find a subgroup $S \subseteqq L$ with $S \cap H = N$, $SH = L$ and $S \subseteqq \mathcal{I}_L(\theta)$. By induction applied to the situation $N \lhd M \lhd SM$, there exists an irreducible constituent φ of θ^M which is invariant under S. Since H/M is an elementary abelian chief factor of L, Proposition 3, Part 2 of [5] applies and we conclude that there are only three cases to consider. They are (a) $\varphi^H = \psi$ is irreducible, (b) $\varphi^H = a\psi$ where ψ is irreducible or (c) φ is extendible to H. In either of cases (a) or (b), ψ is invariant under S and since $L = HS$, we are done. In the remaining case, φ is invariant in L and the result follows from Lemma 3.3.

We state below the special case of Theorem B of Hall and Higman which will be needed in what follows.

PROPOSITION 4.3. *Let G be a p-solvable group which acts faithfully and irreducibly on an elementary abelian p-group U. Suppose $|U| < p^{p-1}$. Then $p \nmid |G|$.*

THEOREM 4.4. *Let $e < p - 1$. Then $u(e) \leqq e + 2$ and $v(e) \leqq e$. If $e(G, N, \theta) < p - 1$, where θ is p-character and G/N is solvable, then $G/N = \mathcal{O}_{pp'pp'}(G/N)$.*

Proof. The first statement follows from the second by Lemmas 1.6 and 2.3 since in calculating $u(e)$ and $v(e)$, it is sufficient to consider only cases where $G/N = \mathcal{O}^{p'}(G/N)$. We proceed to prove the second statement.

Let $N \lhd G$, θ an invariant p-character of N and

$$e(G, N, \theta) = e < p - 1 \,.$$

It suffices to show that $\mathcal{O}^{p'pp'p}(G/N) = 1$ and this is done by induction on $|G\colon N|$. If $\mathcal{O}^{p'}(G/N) = L/N$ and $L < G$, then since

$$e(L, N, \theta) \leqq e(G, N, \theta) \,,$$

the result follows by induction. Thus we may assume that

$$\mathcal{O}^{p'}(G/N) = G/N$$

and similarly, $\mathcal{O}^{pp'pp'p}(G/N) = 1$. Let $H/N = \mathcal{O}^{pp'pp'}(G/N)$ and $U/N = \mathcal{O}^{pp'p}(G/N)$ so that U/N has the normal S_p subgroup H/N. We may assume $U > N$. Let $V/N = \mathcal{O}^{pp'}(G/N)$ so that V/U is a p-group. Suppose $U \subseteqq Y < V$ with $Y \lhd G$ and $|V\colon Y| < p^{p-1}$. Let Y be a maximal such subgroup. Then V/Y is an elementary abelian p-group which is an irreducible G/V module. Let $C/V = \mathbf{C}_{G/V}(V/Y)$ so V/Y is a faithful irreducible G/C module. By Proposition 4.3, G/C is a p'-group and since $G/N = \mathcal{O}^{p'}(G/N)$, we have $C = G$. It follows that V/Y is a direct factor of M/Y where $M/N = \mathcal{O}^{p}(G/N)$. Since $V/Y > 1$, this contradicts $\mathcal{O}^{p}(M/N) = M/N$ and therefore no such Y exists.

Now let $U_0/H = (U/H)'$ and let $Y/U = \mathbf{C}_{V/U}(U/U_0)$. Then $Y \lhd G$. Now $U/U_0 \subseteqq \mathbf{Z}(Y/U_0)$ and U/U_0 is a nontrivial $S_{p'}$ subgroup of Y/U_0 since $U > H$ and U/H is a solvable p'-group. It follows that $\mathcal{O}^{p'}(Y/U_0) < Y/U_0$. Since $\mathcal{O}^{p'}(V/N) = V/N$, it must be that $Y < V$.

We will have the desired contradiction when we show $|V: Y| \leqq p^e < p^{p-1}$.

By Lemma 4.2, there exists an irreducible constituent ψ of θ^H such that ψ is invariant in U. Since H/N is a p-group, it follows from Lemma 2.4 that ψ is a p-character of H and hence there exists a unique extension $\hat{\psi}$ of ψ to U with $o(\psi) = o(\hat{\psi})$. It follows from the uniqueness that $\mathscr{I}_G(\psi) = \mathscr{I}_G(\hat{\psi})$. Now let λ be any linear character of U/H. Then $\lambda\hat{\psi} \in Irr(U)$. Let $T = \mathscr{I}_G(\lambda\hat{\psi})$ and put $|G: T| = p^r$. By Corollary 1.3, $e(T, U, \lambda\hat{\psi}) \leqq e - r$ and thus $r \leqq e$. Let $x \in T$. We have

$$\lambda\hat{\psi} = (\lambda\hat{\psi})^x = \lambda^x\hat{\psi}^x .$$

Restricting this to H, we obtain $\psi = \psi^x$ since $\lambda_H = 1$ and $\hat{\psi}_H = \psi$. Thus $x \in \mathscr{I}(\psi) = \mathscr{I}(\hat{\psi})$. Therefore $\lambda\hat{\psi} = \lambda^x\hat{\psi}$ and it follows from Proposition 1.4 that $\lambda = \lambda^x$. Thus $T \subseteqq \mathscr{I}_G(\lambda)$. Since $|G: T|_p = p^r$ and V/U is a normal p-subgroup of G/U, it follows that $|V: T \cap V| \leqq p^r$. Thus $|V: \mathscr{I}_V(\lambda)| \leqq p^r \leqq p^e < p^p$. Therefore, in the action of the p-group V/U on the group of linear characters of U/H, all orbits have size $\leqq p^e$. The kernel of this action is Y/U and thus by proposition 4.1, $|V/Y| \leqq p^e$ which yields the desired contradiction and the proof is complete.

COROLLARY 4.5. *If* $e \geqq p - 1$, $u(e) \leqq 2e - p + 4$ *and*

$$v(e) \leqq 2e - p + 3 .$$

Proof.

$$\begin{aligned} u(p-1) &\leqq \max_{0<f\leqq p-1} \{u(p-1-f) + f\} + 1 \\ &\leqq \max_{0<f\leqq p-1} \{p - 1 - f + 2 + f\} + 1 = p + 2 \end{aligned}$$

and similarly $v(p-1) \leqq p + 1$. Thus the desired inequalities hold when $e = p - 1$. For $e > p - 1$, apply induction.

5. In this section we consider the case $e = 1$ in more detail. From Theorem 4.4 we have $u(1) \leqq 3$ and $v(1) \leqq 1$ when $p \geqq 3$. For $p = 2$, Corollary 3.6 yields $u(1) \leqq 3$ and $v(1) \leqq 2$. An example (see 6.1) shows that $u(1) = 3$ for $p = 3$.

THEOREM 5.1. *For all primes,* $v(1) = 1$.

Proof. That $v(1) \geqq 1$ is clear. Let $e(G, N, \theta) = 1$ with G/N solvable and θ an invariant p-character. Suppose $\mathscr{O}_p(G/N) = 1$. We must show that an S_p subgroup, P/N, of G/N is abelian. Let K/N be a minimal normal subgroup of G/N so that K/N is an elementary abelian q-group for some prime $q \neq p$. Let $\hat{\theta}$ be the unique extension of θ to K with $o(\hat{\theta}) = o(\theta)$. Then $\hat{\theta}$ is an invariant p-character of K. If $\mathcal{O}_p(G/K) = 1$, then the result follows by induction on $|G: N|$. Assume then that $H/K = \mathcal{O}_p(G/K) > 1$. Let λ be any linear character of H/K. Then $\mathcal{I}_G(\lambda) = \mathcal{I}_G(\hat{\theta}\lambda)$ and thus $p^2 \nmid |G: \mathcal{I}_G(\lambda)|$. It follows that λ lies in an orbit of size 1 or p under the action of H/K on the group of linear characters of K/N. Since $\mathcal{O}_p(G/N) = 1$, $\mathbf{C}_{H/K}(K/N) = 1$ and thus H/K acts faithfully on the linear characters of K/N. By Proposition 4.1, $|H/K| = p$.

Now choose λ as above in an orbit of size p. Then

$$(\lambda\hat{\theta})^H = \psi \in Irr(H)$$

and ψ is a p-character of H by Lemma 2.4 (using the minimality of K). Let $T = \mathcal{I}_G(\psi)$ and $T_0 = \mathcal{I}_G(\lambda\hat{\theta})$ so that $HT_0 \subseteqq T$ and $T_0 \cap H = K$. By the usual argument, $p^2 \nmid |G: T_0|$ and hence $p \nmid |G: HT_0|$ and we may assume that $P \subseteqq HT_0$. Then $PK/K = (H/K)(P_0/K)$ where $P_0 = PK \cap T_0$. Now $e(T, H, \psi) = 0$ by Corollary 1.3 and since $u(0) = 1$, we have PK/H is abelian. But $PK/H \cong P_0/K$ and $H/K \subseteqq \mathbf{Z}(PK/K)$ and thus $PK/K \cong P/N$ is abelian. The proof is complete.

We now prove a result which is valid for p-solvable groups with $p > 3$. It will enable us to conclude for solvable groups that $u(1) \leqq 2$ with respect to these primes.

THEOREM 5.2. *Let $N \lhd G$ with G/N p-solvable and $p > 3$. Suppose θ is a p-character of N which is invariant in G and that $e(G, N, \theta) = 1$. Let $P/N = \mathcal{O}_p(G/N)$ and suppose that P/N is not abelian. Then P/N is an S_p subgroup of G/N.*

Proof. Use induction on $|G: N|$ and assume that $P/N \notin Syl_p(G/N)$. Then P/N is a Sylow subgroup of every proper normal subgroup of G/N which contains it. It follows that $\mathcal{O}^{p'}(G/N) = G/N$. Also $M/P = \mathcal{O}^p(G/P) < G/P$ and $|G: M| = p$. By Lemma 4.2, there exists an irreducible constituent η of θ^P which is invariant in M. Now η is a p-character of P by Lemma 2.4 and thus there exists a unique extension $\hat{\eta}$ of η to M with $o(\eta) = o(\hat{\eta})$. We have either $\eta(1) = \theta(1)$ or $\eta(1) = p\theta(1)$. In the latter case, it is clear that $\hat{\eta}$ must be invariant in G and hence it is extendible to $\chi \in Irr(G)$. Now G/P does not have a normal S_p subgroup and thus has some irreducible character β of degree divisible by p. Since χ_P is irreducible, $\beta\chi \in Irr(G)$ and this contradicts $e(G, N, \theta) = 1$. Therefore we must have $\eta(1) = \theta(1)$.

We claim now that $e(G/N) = 1$. Let $\varphi \in Irr(M/N)$ with $p | \varphi(1)$. It suffices to show that $p^2 \nmid \varphi(1)$ and that φ is invariant in G. Now $\hat{\eta}\varphi \in Irr(M)$ and $p^2\theta(1) \nmid (\hat{\eta}\varphi)(1)$. Thus $p^2 \nmid \varphi(1)$. Also $(\hat{\eta}\varphi)^G$ is not irreducible so that $\hat{\eta}\varphi$ is invariant in G. Now let $x \in G$. Then $\hat{\eta}\varphi = (\hat{\eta}\varphi)^x = \hat{\eta}^x\varphi^x$. Since $\hat{\eta}^x$ and $\hat{\eta}$ are both extensions of θ to M, there exists a linear character λ of M/N with $\hat{\eta}^x = \lambda\hat{\eta}$. Substituting in the above, we obtain $\hat{\eta}\varphi = \lambda\hat{\eta}\varphi^x$. Since $\hat{\eta}$ is an extension of θ and φ and $\lambda\varphi^x$ are irreducible characters of H/N, it follows by Proposition 1.4 that $\varphi = \lambda\varphi^x$. Applying this to the complex conjugate character $\bar{\varphi}$, we obtain $\bar{\varphi} = \lambda\bar{\varphi}^x$, and thus $\varphi = \bar{\lambda}\varphi^x$. This yields $\lambda\varphi^x = \bar{\lambda}\varphi^x$ and $\lambda^2\varphi^x = \varphi^x$.

Now $o(\eta) = o(\hat{\eta}^x)$. We have $\det(\hat{\eta}^x) = \det(\lambda\hat{\eta}) = \lambda^f \det(\hat{\eta})$ where $f = \hat{\eta}(1)$ is a power of p. It follows that $o(\lambda)$ is a power of p, and since $p > 2$, λ is a power of λ^2. Since $\varphi^x = \lambda^2\varphi^x$, we obtain $\varphi^x = \lambda\varphi^x = \varphi$. Since $x \in G$ was arbitrary, φ is invariant in G and we have thus shown that $e(G/N) = 1$.

We may now assume without loss that $N = 1$. In the notation of [6], P has $r.x.1$ and by Theorem C of that paper, either P has an abelian subgroup of index p or else $|P\colon \mathbf{Z}(P)| = p^3$. It follows that either P has a characteristic abelian subgroup of index p or $|P\colon \mathbf{Z}(P)| \leqq p^3$. We claim that there exists $A \lhd G, A \subseteqq P$ with $|P\colon A| = p$ and A abelian. If this is not the case then $|P\colon\mathbf{Z}(P)| \leqq p^3$. Let S be an $S_{p'}$ subgroup of M. Then $U = [P, S] \lhd G$ since for $g \in G$, $S^g = S^x$ for some $x \in P$. We claim that $U \subseteqq \mathbf{Z}(P)$. Otherwise $V = U\mathbf{Z}(P) > \mathbf{Z}(P)$ and we choose $Y \lhd G$, maximal such that $\mathbf{Z}(P) \subseteqq Y < V$. Let $C/Y = \mathbf{C}_{G/Y}(V/Y)$. Then V/Y is a faithful irreducible G/C module. Now $|V/Y| \leqq p^3$ and $p \geqq 5$ and hence it follows from Proposition 4.3 that $p \nmid |G/C|$. Since $\mathcal{O}^{p'}(G) = G$, it follows that $G = C$ and thus $[V, G] \subseteqq Y$. In particular $[U, S] \subseteqq Y \cap U < U$. Since $U = [P, S] = [P, S, S]$, this is a contradiction and thus $U \subseteqq \mathbf{Z}(P)$. It follows that $\mathbf{Z}(P) \supseteqq P \cap \mathcal{O}^p(G)$.

Since P is not abelian, $P/\mathbf{Z}(P)$ is not cyclic and thus $G/\mathcal{O}^p(G)$ is not cyclic. It follows that there exists a subgroup $M_0 \lhd G$, with $M \neq M_0$ and $|G\colon M_0| = p$. Now $\mathcal{O}_p(M_0) = M_0 \cap P$ is not an S_p subgroup of M_0. By induction, $M_0 \cap P$ is abelian. Since $|P\colon M_0 \cap P| = p$ and $M_0 \cap P \lhd G$, the claim is established and A exists.

Suppose λ is a linear character of A which is not invariant in P. Let $T = \mathcal{I}_G(\lambda)$. Then, $P \cap T = A$ and hence $p \mid |G\colon T|$. By Corollary 1.3, it follows that $e(T, A, \lambda) = 0$ and $p^2 \nmid |G\colon T|$. Since λ is obviously a p-character, it follows from Corollary 3.4 that T/A has a normal S_p subgroup, of order exactly p. Let U be the group of linear characters of A. Then G/A acts on U and we let $Z = \mathbf{C}_U(P/A)$. The above argument shows that if $u \in U - Z$, then $\mathbf{C}_{G/A}(u)$ has a normal S_p subgroup of order p.

Let $P/A = \langle x \rangle$ and let $W = [U, x]$. Then the map $f\colon u \to [u, x]$ defines a homomorphism from U onto W and $\ker f = Z$. Set $Y = G/A$ and $P_0 = P/A \lhd Y$. Now $\mathbf{C}_Y(P_0)$ has index dividing $p - 1$. However

$\mathcal{O}^{p'}(Y) = Y$ and it follows that $P_0 \subseteq \mathbf{Z}(Y)$. Therefore, for $y \in Y$ and $u \in U$, we have $f(u^y) = f(u)^y$ and f is a homomorphism of Y-modules. Also, from $P_0 \subseteq \mathbf{Z}(Y)$, it follows that Y has a normal p-complement and thus so does every subgroup.

Since P is not abelian, $A \not\subseteq \mathbf{Z}(P)$ and it follows that x acts nontrivially on U. Therefore $W > 1$ and hence $V = W \cap Z > 1$. Now choose $w \in V$, $w \neq 1$. Let K be the normal p-complement of $\mathbf{C}_Y(w)$. Then K fixes the inverse image of w under f, which is a coset of Z. It follows (by Theorem 1 of [2] for instance), that K fixes some element $u \in U$ with $f(u) = w$. In particular, $u \notin Z$ so $\mathbf{C}_Y(u)$ has the normal S_p subgroup, P_1, of order p. Now K is a full p-complement for $\mathbf{C}_Y(u)$ since $\mathbf{C}_Y(u) \subseteq \mathbf{C}_Y(w)$. Hence $\mathbf{C}_Y(u) = K \times P_1$ and

$$\mathbf{C}_Y(w) = K \times P_1 \times P_0 .$$

Now, $\mathbf{C}_Y(V) \subseteq \mathbf{C}_Y(w)$ and thus has a normal S_p subgroup. Since $\mathcal{O}_p(Y) = P_0$, P_0 is a full S_p subgroup of $\mathbf{C}_Y(V)$.

Now suppose $v \in V$ with $P_1 \not\subseteq \mathbf{C}(v)$. Let P_2 be the subgroup of order p in $\mathbf{C}_Y(vu)$. Then $P_2 \neq P_1$ and $P_2 \neq P_0$. Furthermore, since $f(vu) = w$, $P_2 \subseteq \mathbf{C}_Y(vu) \subseteq \mathbf{C}_Y(w)$ and thus $P_2 \subseteq P_0P_1$. We may therefore choose $y \in P_1$ with $xy \in P_2$. Then $uv = (uv)^{xy} = u^{xy}v^y = u^xv^y$. However, $w = f(u) = u^{-1}u^x$ and $u^x = uw$. Hence $uv = uwv^y$ and $[y, v] = v^{-y}v = w$. Since $wy = yw$, it follows that $1 = [y^p, v] = w^p$ and w has order p. Since $w \in V$ was arbitrary, V is elementary abelian. Also from $[y, v] = w$, it follows that $[P_1, v] = \langle w \rangle$. Since $v \in V$ was arbitrary, not centralized by P_1, it follows that $[P_1, V] = \langle w \rangle$. Therefore $\mathbf{C}_V(P_1)$ has codimension 1 in V. Now choose $w^* \in \mathbf{C}_V(P_1)$ with $w^* \neq 1$. Repeating the above reasoning with w^* in place of w, we conclude that $[P_1^*, V] = \langle w^* \rangle$, where $P_1^* \times P_0$ is a normal S_p subgroup of $C_Y(w^*)$. By the choice of w^*, $P_1 \subseteq \mathbf{C}_Y(w^*)$ and thus $P_1 \subseteq P_1^* \times P_0$. Since $[P_0, V] = 1$, $\langle w \rangle = [P_1, V] \subseteq [P_1^*, V] = \langle w^* \rangle$. It follows that $\mathbf{C}_V(P_1) = \langle w \rangle$ and hence $|V| = p^2$. Given any basis $\{v, w\}$ for V, the above argument shows that there exists $y \in Y$ with $[y, v] = w$ and thus Y acts irreducibly on V. Since $p > 3$, Proposition 4.3 applies and $p \nmid |Y : \mathbf{C}_Y(V)|$. It follows that Y centralizes V which is a contradiction. The proof is complete.

COROLLARY 5.3 *If $p > 3$, then $u(1) = 2$.*

Proof. It suffices to show $u(1) \leqq 2$. Let $e(G, N, \theta) = 1$ with θ a p-character and G/N solvable. If $\mathcal{O}_p(G/N) = H/N$ is an S_p subgroup of G/N, then by Lemma 1.6, d.l.$(H/N) \leqq 2$ and nothing remains to be shown. Otherwise H/N is abelian. Choose an irreducible constituent ψ of θ^H which is invariant in U, where $U/H = \mathcal{O}_{p'}(G/N)$.

(Lemma 4.2). Let $T = \mathscr{I}_G(\psi)$. Then $\mathscr{O}_p(T/H) = 1$ and $e(T, H, \psi) \leqq 1$. If $e(T, H, \psi) = 0$ then since $v(0) = 0$, $p \nmid |T:H|$ and $p^2 \nmid |G:T|$ by Corollary 1.3. Thus $p^2 \nmid |G: H|$ and the result follows. If $e(T, H, \psi) = 1$ then $p \nmid |G: T|$ and the result follows from $v(1) = 1$.

6. The assumption $p > 3$ was used twice in the proof of Theorem 5.2. In this section we give examples to show that both uses were essential.

EXAMPLE 6.1. Let P be the group of matrices of the form

$$\begin{bmatrix} 1 & x & y \\ 0 & 1 & x^3 \\ 0 & 0 & 1 \end{bmatrix} = M(x, y)$$

where $x, y \in GF(27)$. Then $|P| = 3^6$ and

$$P' = \mathbf{Z}(P) = \{M(0, y) | y \in GF(27)\} .$$

Let $\lambda \in GF(27)$ have order 13. Then the map $M(x, y) \rightarrow M(x\lambda, y\lambda^4)$ is an automorphism of P of order 13. Denote this automorphism by σ_λ and let M be the split extension $P \langle\sigma_\lambda\rangle$. Now $GF(27)$ has an automorphism τ of order 3 and we let τ act on M in the natural manner, with $(\sigma_\lambda)^\tau = \sigma_{\lambda^\tau}$. Let $G = M \langle\tau\rangle$. We claim that $e(G) = 1$, but $\mathscr{O}_3(G) = P$ is not abelian.

It suffices to check that every irreducible character of P is stabilized by some element of order 3 in G/P. Now τ fixes the two linear characters of P whose kernel is $[P, \tau]$. It is not hard to show that $P \langle\tau\rangle/[P', \tau]$ has center of index 3^3 so all of its irreducible nonlinear characters have degree 3. It follows that τ fixes all six nonlinear irreducible characters of P with kernel containing $[P', \tau]$. Since σ acts transitively on hyperplanes of P/P' and of P', it follows that every irreducible character of P is conjugate in M to a character fixed by τ and this proves the claim. Note that G contains no normal abelian subgroup A of index 3 in P. Also, d.l.$(P \langle\tau\rangle) = 3$.

EXAMPLE 6.2. Let $A = \langle x_1, x_2, y_1, y_2\rangle$ be elementary abelian of order 3^4. Let $Y = \langle\sigma\rangle \times S$ where σ has order 3 and $S \cong SL(2, 3)$. Let Y act on A so that S acts in its natural manner on $\langle x_1, x_2\rangle$ and on $\langle y_1, y_2\rangle$ with $x_1 \rightarrow y_1$ and $x_2 \rightarrow y_2$ defining an S-isomorphism. Let $x_i^\sigma = x_i y_i$ and $y_i^\sigma = y_i$. Let G be the split extension AY. Now $\mathscr{O}_3(G) = A \langle\sigma\rangle$ is not abelian.

To show that $e(G) = 1$, it suffices to show that every linear character of A is fixed by some element of Y of order 3. Let U be the group of linear characters of A and let $V \subseteqq U$ be those whose kernels contain $\langle y_1, y_2\rangle$. The unique element of order 2 of Y fixes no nonidentity element of U and hence for $1 \neq u \in U$, $C_Y(u)$ is a 3-group. Now the 3-subgroups of Y, either contain σ or else have order 3. Since $C_U(\sigma) = V$, it follows that if $u \in U - V$, then $|C_Y(u)| \leqq 3$.

Each subgroup of order 3 of Y must centralize a subgroup of order at least 9 in U since U is elementary abelian of order 3^4. Since $C_Y(V) = \langle\sigma\rangle$, it follows that each of the 12 subgroups of Y of order 3, different from $\langle\sigma\rangle$, centralize at least six elements of U-V. Since these sets are disjoint, this accounts for all 72 elements of $U - V$ and the result follows.

In example 6.2, even though the normal abelian subgroup A does exist, the conclusion of Theorem 5.2 does not hold. Therefore, the second assumption that $p > 3$ was essential. Note that Example 6.1 shows that $u(3) = 3$.

REFERENCES

1. P. X. Gallagher, *Group characters and normal Hall subgroups*, Nagoya Math. J., 21 (1962), 223-230.

2. G. Glauberman, *Fixed points in groups with operator groups*, Math. Zeit., **84** (1964), 120-125.

3. P. Hall and G. Higman, *On the p-length of p-solvable groups...*, Proc. London Math. Soc., (3) **6** (1956) 1-42.

4. I. M. Isaacs, *Finite groups with small character degrees and large prime divisors*, Pacific J. Math., **23** (1967), 273-280.

5. ———, *Fixed points and characters in groups with non-coprime operator groups*. Canad. J. Math., **20** (1968), 1315-1320.

6. I. M. Isaacs and D. S. Passman, *A characterization of groups in terms of the degrees of their characters*, Pacific J. Math., **15** (1965), 877-903.

7. N. Itô, *Some studies on group characters*, Nagoya Math. J., **2** (1951), 17-28.

8. D. S. Passman, *Groups with normal solvable Hall p′-subgroups*, Trans. Amer. Math. Soc., **123** (1966), 99-111.

Received October 6, 1969.

UNIVERSITY OF WISCONSIN

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

RINGS OF QUOTIENTS OF Φ-ALGEBRAS

D. G. JOHNSON

Let $\mathscr{X}$ be a completely regular (Hausdorff) space. Fine, Gillman, and Lambek have studied the (generalized) rings of quotients of $C(\mathscr{X}) = C(\mathscr{X}; \mathbf{R})$, with particular emphasis on the maximal ring of quotients, $Q(\mathscr{X})$. In this note, we start with a characterization of $Q(\mathscr{X})$ that differs only slightly from one of theirs. This characterization is easily altered to fit more general circumstances, and so serves to obtain some results on non-maximal rings of quotients of $C(\mathscr{X})$, and to generalize these results to the class of Φ-algebras.

We consider only commutative rings with unit. Let A be one such, and recall that the (unitary) over-ring B of A is called a *rational extension* or *ring of quotients* of A if it satisfies the following condition: given $b \in B$, for every $0 \neq b' \in B$ there is $a \in A$ with $ba \in A$ and $b'a \neq 0$. A ring without proper rational extensions is said to be *rationally complete*. For the rings to be considered here (all are semi-prime), the condition above can be replaced by the simpler condition: for $0 \neq b \in B$, there exists $a \in A$ such that $0 \neq ba \in A$ ([1], p. 5). Accordingly, we make the following

DEFINITION. If B is an over-ring of A and $0 \neq b \in B$, say that b is *rational over* A if there is $a \in A$ with $0 \neq ba \in A$.

Let $m\beta\mathscr{X}$ denote the minimal projective extension of $\beta\mathscr{X}$ and $\tau: m\beta\mathscr{X} \to \beta\mathscr{X}$ the minimal perfect map ([2]). In [1], it is shown that $Q(\mathscr{X})$ is a dense, point-separating subalgebra of $D(m\beta\mathscr{X})$, the set of all continuous maps from $m\beta\mathscr{X}$ into the two-point compactification of the real line which are real-valued on a dense subset of $m\beta\mathscr{X}$ (see, also, [3]). Since $Q(\mathscr{X})$ contains every ring of quotients of $C(\mathscr{X})$, this leads to

PROPOSITION 1. *If B is any ring of quotients of $C(\mathscr{X})$, then there exist a compact (Hausdorff) space $\mathscr{Y}$ and minimal perfect maps α and γ such that B is a point-separating subalgebra of $D(\mathscr{Y})$ and the following diagram commutes*:

$$\begin{tikzcd} m\beta\mathscr{X} \arrow[r, "\alpha"] \arrow[dr, "\tau"'] & \mathscr{Y} \arrow[d, "\gamma"] \\ & \beta\mathscr{X} \end{tikzcd}$$

$\mathscr{Y}$ is the obvious identification space, and the proof consists of a routine argument to show that the quotient map α is closed, whence $\mathscr{Y}$ is Hausdorff. Since $C(\mathscr{X}) \subseteqq B$, the existence of γ follows immediately. (Note that, although $D(m\beta\mathscr{X})$ is an algebra, $D(\mathscr{Y})$ for other spaces $\mathscr{Y}$ is, in general, only a partial algebra.)

For our purposes, it is convenient to view $C(\mathscr{X})$ as a subalgebra of $D(\beta\mathscr{X})$. *This allows us to decree that all spaces are compact (Hausdorff).*

Let us say that any space $\mathscr{Y}$ that is situated in a commutative diagram of the form

$$\begin{tikzcd} m\mathscr{X} \arrow[r, "\alpha"] \arrow[dr, "\tau"'] & \mathscr{Y} \arrow[d, "\gamma"] \\ & \mathscr{X}\,, \end{tikzcd}$$

where all maps are minimal perfect, is *near to* $\mathscr{X}$. (Of course, the existence of γ automatically guarantees the existence of α.) Note that we have already adopted the convention of identifying $f \in D(\mathscr{X})$ with its image $f \circ \gamma$ in $D(\mathscr{Y})$ whenever convenient. With this convention, if A is a subalgebra of $D(\mathscr{Y})$ and $f \in D(\mathscr{Y})$ then we may consider f as an element of an over-ring of A—$D(m\mathscr{X})$—, even if there is no subalgebra of $D(\mathscr{Y})$ containing both A and f.

Now let A be a Φ-algebra that is closed under bounded inversion; i.e., an archimedean lattice ordered algebra with a multiplicative identity that is a weak order unit, in which $1/a \in A$ whenever $1 \leqq a \in A$. Let $\mathscr{X} = \mathscr{M}(A)$, the space of maximal ideals of A with the hull-kernel topology. It is shown in [4] that A is (isomorphic with) a point-separating subalgebra of $D(\mathscr{X})$. If $\mathscr{Y}$ is any space that is near to $\mathscr{X}$, let $A_{\mathscr{Y}} = \{f \in D(\mathscr{Y})$: for each nonempty open set $\mathscr{U}$ in $\mathscr{Y}$, there are a nonempty open set $\mathscr{V} \subseteqq \mathscr{U}$ and $g \in A$ such that $f|_{\mathscr{V}} = g|_{\mathscr{V}}\}$. Note that $A_{\mathscr{Y}}$ is always a lattice. However, it need not be an algebra:

EXAMPLE. Let $\mathscr{X} = \mathscr{Y}$, the one-point compactification of the countable discrete space, and let $A = C(\mathscr{X})$. Then $A_{\mathscr{Y}} = D(\mathscr{Y})$, which is not an algebra.

REMARK. One readily shows that the open sets $\mathscr{V}$ appearing in the definition of $A_{\mathscr{Y}}$ can always be shown to have the form $\gamma^{\leftarrow}[\mathscr{V}_1]$, where $\mathscr{V}_1$ is open in $\mathscr{X}$. It follows that

$$A_{\mathscr{Y}} = \{f \in D(\mathscr{Y}) \colon f \circ \alpha \in A_{m\mathscr{X}}\}\,.$$

PROPOSITION 2. (i) *Every element of* $A_{\mathscr{Y}}$ *is rational over* A^*

(and, hence, over A).

(ii) *$A_{\mathscr{Y}}$ contains every rational extension of A and A^* in $D(\mathscr{Y})$.*

Proof. (i) Let $0 \neq f \in A_{\mathscr{Y}}$, and let $\mathscr{U}$ be a nonempty open set contained in $\operatorname{coz} f$. Since $f \in A_{\mathscr{Y}}$, there exist a nonempty open set $\mathscr{V} = \gamma^{\leftarrow}[\mathscr{V}_1] \subseteq U$, where $\mathscr{V}_1$ is open in $\mathscr{X}$, and $h \in A^*$ such that $f|_{\mathscr{V}} = h|_{\mathscr{V}}$. Choose $0 \neq g \in A^*$ with $\overline{\operatorname{coz} g} \subseteq \mathscr{V}_1$. Then $0 \neq fg = hg \in A^*$.

(ii) Let $f \in D(\mathscr{Y}) \backslash A_{\mathscr{Y}}$. Then, there is a nonempty open set $\mathscr{U}$ such that f agrees with no member of A on any nonempty open subset of $\mathscr{U}$. Choose $g \in A^*$ with $\phi \neq \overline{\operatorname{coz} g} \subseteq \mathscr{U}$.

There is no $h \in A$ with $hg \neq 0$ while $fh \in A$. For, such h would agree with a unit h_1 of A on some nonempty open subset $\mathscr{V}$ of $\mathscr{U}$ (since A is closed under bounded inversion), whence

$$f|_{\mathscr{V}} = (h/h_1)f|_{\mathscr{V}} ,$$

while $(1/h_1)hf \in A$, a contradiction. Thus, f is contained in no rational extension of A.

Although $A_{\mathscr{Y}}$ may contain many different rational extensions of A, it is not true that it is the union of such extensions, as is seen in the example preceding Proposition 2. However, in those spaces $\mathscr{Y}$ for which $A_{\mathscr{Y}}$ is an algebra, $A_{\mathscr{Y}}$ is a Φ-algebra and is the largest ring of quotients of A that "lives on" $\mathscr{Y}$. In particular, this happens when $D(\mathscr{Y})$ is an algebra (e.g., when $\mathscr{Y}$ is basically disconnected or an F-space). Hence, $A_{m\mathscr{X}}$ is a Φ-algebra, since $m\mathscr{X}$ is extremally disconnected, and we obtain the following generalizations of results in [1].

THEOREM 1. *$A_{m\mathscr{X}}$ is rationally complete; thus, $A_{m\mathscr{X}} = \mathscr{Q}(A)$, the maximal ring of quotients of A.*

THEOREM 2. *$A_{m\mathscr{X}}$ is uniformly dense in $D(m\mathscr{X})$.*

THEOREM 3 ([1]). *$D(m\mathscr{X})$ is rationally complete.*

The proofs of Theorems 1 and 3 are virtually identical, and are related to one found on p. 30 of [1]; we prove 1. To do so, we will employ the following characterization of rational completeness (see [1], p. 7).

The commutative ring B is rationally complete if and only if it satisfies: for any dense ideal I of B, every element of $\operatorname{Hom}_B(I, B)$ *is a multiplication by an element of B. (In the present setting, an ideal I of $A_{m\mathscr{X}}$ is dense if and only if $\bigcup\{\operatorname{coz} f : f \in I\}$ is dense in $m\mathscr{X}$.)*

Proof of Theorem 1. Let I be a dense ideal in A, and let

$\phi \in \mathrm{Hom}_{A_{m\mathscr{X}}}(I, A_{m\mathscr{X}})$. By Zorn's lemma, choose a family $\{\mathscr{U}_\kappa: \kappa \in K\}$ of open sets in $m\mathscr{X}$ satisfying:

(i) $\mathscr{U} = \bigcup \mathscr{U}_\kappa$ is dense in $m\mathscr{X}$;

(ii) the $\mathscr{U}_\kappa$ are pairwise disjoint;

(iii) for each κ, there is $f_\kappa \in I$ such that f_κ is bounded away from zero on $\mathscr{U}_\kappa$ and both f_κ and $\phi(f_\kappa)$ agree with members of A on $\mathscr{U}_\kappa$.

Let $f \in D(m\mathscr{X})$ satisfy

$$f\Big|_{\mathscr{U}_\kappa} = \frac{\phi(f_\kappa)}{f_\kappa}\Big|_{\mathscr{U}_\kappa}$$

for each $\kappa \in K$. This is possible, since $m\mathscr{X}$ is extremally disconnected, so $m\mathscr{X} = \beta\mathscr{U}$.

If $g \in I$ and $x \in \mathscr{U}_\kappa$, then

$$f(x)g(x) = \frac{\phi(f_\kappa)(x)}{f_\kappa(x)}g(x) = \frac{g\phi(f_\kappa)}{f_\kappa}\Big|_{\mathscr{U}_\kappa}(x) = \frac{f_\kappa\phi(g)}{f_\kappa}\Big|_{\mathscr{U}_\kappa}(x) = \phi(g)(x) .$$

It follows that ϕ is multiplication by f. Clearly, $f \in A_{m\mathscr{X}}$, and the proof is complete.

Proof of Theorem 2. Let $f \in D(m\mathscr{X})$, $\varepsilon > 0$. By Zorn's lemma, choose a family $\{\mathscr{U}_\kappa: \kappa \in K\}$ of open sets in $m\mathscr{X}$ which satisfies:

(i) $\mathscr{U} = \bigcup \mathscr{U}_\kappa$ is dense in $m\mathscr{X}$;

(ii) the $\mathscr{U}_\kappa$ are pairwise disjoint;

(iii) for $x, y \in \mathscr{U}_\kappa$, $|f(x) - f(y)| < \varepsilon$ (in particular, f is real-valued on $\mathscr{U}_\kappa$).

For each $\kappa \in K$, choose $x_\kappa \in \mathscr{U}_\kappa$, and define $g: \mathscr{U} \to \mathbf{R}$ by

$$g(y) = f(x_\kappa) \quad \text{if} \quad y \in \mathscr{U}_\kappa .$$

Since $m\mathscr{X} = \beta\mathscr{U}$, g can be extended to $\hat{g} \in D(m\mathscr{X})$. Clearly, $\hat{g} \in A_{m\mathscr{X}}$, and

$$|f - \hat{g}| \leqq \varepsilon .$$

Now the analogue of Proposition 1 for Φ-algebras is routinely obtained.

In case $\mathscr{Y} = m\mathscr{X}$ and $A = C(\mathscr{X})$ one readily translates the definition of $A_{\mathscr{Y}}$ (using the fact that $m\mathscr{X}$ is extremally disconnected, and hence that every dense subspace is C^*-embedded) as follows:

$$A_{m\mathscr{X}} = \varinjlim \{C(\mathscr{S}): \mathscr{S} \text{ is a dense open subset of } \mathscr{X}\} .$$

Thus, the Fine-Gillman-Lambek result that this direct limit is $Q(\mathscr{X})$ follows from Theorem 1.

It is easily seen that any Φ-algebra A is a rational extension of its bounded subring A^*, and hence that $(A^*)_{\mathscr{Y}} = A_{\mathscr{Y}}$ for any space $\mathscr{Y}$ near to $\mathscr{M}(A)$. Thus, if A is closed under uniform convergence, then $\mathscr{Q}(A) = \mathscr{Q}(A^*) = Q(\mathscr{M}(A))$, since $A^* = C(\mathscr{M}(A))$. In the general case, this may fail to hold. (So, more generally, $A_{\mathscr{Y}} \neq C(\mathscr{M}(A))_{\mathscr{Y}}$ even when $A \subsetneqq C(\mathscr{M}(A))$.)

EXAMPLE. Let $A = Q(\mathbf{R})$. Then (see [1], p. 34),

$$A = \mathscr{Q}(A^*) \neq D(m\mathbf{R}) = D(M(A^*)) = Q(M(A^*)) .$$

For any Φ-algebra A and any space $\mathscr{Y}$ near to $\mathscr{X} = \mathscr{M}(A)$, every subalgebra of $A_{\mathscr{Y}}$ that contains A is a ring of quotients of A. Of interest are those that separate points of $\mathscr{Y}$; prime candidates are the maximal subalgebras of $A_{\mathscr{Y}}$ containing A, which are easily seen to exist.

The results that follow are obtained using ideas and methods employed by Nanzetta in [6] (see his 2.1, 2.3, 4.1). Conversion of his arguments to the present setting is largely an exercise in careful bookkeeping, and the details are omitted.

THEOREM 4. *If B is a maximal subalgebra of $A_{\mathscr{Y}}$, then B is a lattice* (hence, a Φ-algebra).

We will use the term "maximal subalgebra of $A_{\mathscr{Y}}$" to denote only those that contain A.

DEFINITION. Let B be a subalgebra of $D(\mathscr{Y})$. A function $f \in D(\mathscr{Y})$ is said to be *locally in* B if each point of $\mathscr{Y}$ has a neighborhood on which f coincides with some member of B. The subalgebra B is said to be *local* (in $D(\mathscr{Y})$) if each member of $D(\mathscr{Y})$ that is locally in B is a member of B.

THEOREM 5. *Every maximal subalgebra of $A_{\mathscr{Y}}$ is local.*

As in [6], this fact yields the following result.

THEOREM 6. *Let B be a maximal subalgebra of $A_{\mathscr{Y}}$, and let $\mathscr{S}$ be a stationary set of B. If $|\mathscr{S}| > 1$, then*

(i) *$\mathscr{S}$ is closed;*

(ii) *$\mathscr{S}$ is nowhere dense;*

(iii) *$\mathscr{S}$ is connected.*

COROLLARY. *If $\mathscr{Y}$ is totally disconnected, then every maximal subalgebra of $A_{\mathscr{Y}}$ separates points of $\mathscr{Y}$.* (Note that this may occur even when $A_{\mathscr{Y}}$ is not an algebra: see the example preceding Proposition 2.)

It is not known whether every space $\mathscr{Y}$ near to $\mathscr{X}$ supports (i.e., is the structure space of) a ring of quotients of $C(\mathscr{X})$. Apparently, an answer to this question awaits a more systematic description of the collection of spaces near to $\mathscr{X}$.

Note that $(A_{\mathscr{Y}})^*$, the set of bounded elements of $A_{\mathscr{Y}}$, is always a Φ-algebra. Hence, it is always a ring of quotients of A^*—the largest bounded ring of quotients of A^* in $D(\mathscr{Y})$. As mentioned above, it is not known whether $(A_{\mathscr{Y}})^*$ always separates points of $\mathscr{Y}$; it clearly does so if and only if $A_{\mathscr{Y}}$ does. However, the example that follows shows that $A_{\mathscr{Y}}$ may separate points in $\mathscr{Y}$ even though $\mathscr{Y}$ supports no ring of quotients of A.

EXAMPLE. Let $\mathscr{S} = \{(x, \sin(1/x)); x \in (0, 1]\}$, let $\mathscr{X}$ denote the one-point compactification of $\mathscr{S}$, and let $\mathscr{Y} = \mathscr{S} \cup (\{0\} \times [-1, 1])$. Let A denote the Φ-algebra of all functions $f \in D(\mathscr{X})$ that satisfy the following condition:

There is a real number x_0, $0 < x_0 < 1$, and a real polynomial p such that

$$f\left(x, \sin\frac{1}{x}\right) = p\left(\frac{1}{x}\right) \quad \text{for} \quad 0 < x < x_0$$

(cf. [4], 3.6). Then $(A_{\mathscr{Y}})^* = C(\mathscr{Y})$, whereas no subalgebra of $D(\mathscr{Y})$ containing A separates points in $\mathscr{Y}$ ([6], Theorem 4.6).

In passing, it should be noted that the development here has proceeded independently of [1]. The only results from that work that have been employed in an essential way came from Chapter 1 of [1], which consists of standard facts about rings of quotients of commutative rings (see, e.g., [5]). Thus, one can rapidly and efficiently reach the high points of the theory developed in [1] along the lines suggested by this note.

REFERENCES

1. N. J. Fine, L. Gillman, and J. Lambek, *Rings of Quotients of Rings of Functions*, McGill Univ. Press, Montreal 1966.
2. A. M. Gleason, *Projective topological spaces*, Ill. J. Math., **2** (1958), 482-489.
3. A. W. Hager, *Isomorphism with a $C(Y)$ of the maximal ring of quotients of $C(X)$*, Fund. Math., **66** (1969), 7-13.
4. M. Henriksen and D. G. Johnson, *On the structure of a class of archimedean lattice-ordered algebras*, Fund. Math., **50** (1961), 73-94.
5. J. Lambek, *Lectures on Rings and Modules*, Blaisdell, Waltham, Mass. 1966.

6. P. Nanzetta, *Maximal lattice-ordered algebras of continuous functions*, Fund. Math., **63** (1968), 53-75.

Received April 13, 1970. Sincere thanks are due to A. W. Hager whose critical comments greatly improved a hastily prepared manuscript.

NEW MEXICO STATE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

TRANSLATION PLANES CONSTRUCTED FROM SEMIFIELD PLANES

NORMAN L. JOHNSON

Let π be an affine plane of order q^2 that is coordinatized by a "derivable" semifield $\mathscr{S} = (\mathscr{S}, +, \cdot)$. If $(\mathscr{S}, +)$ is a right vector space over $F = GF(q)$ then a plane π' may be constructed from π using Ostrom's method of "derivation."

The purpose of this article is to examine the planes π' and their coordinate structures $(\mathscr{S}, +, *)$. It is shown, in particular, that $(\mathscr{S}, +, *)$ is a (right) quasifield which is neither a nearfield nor a semifield. Furthermore, it is shown that π' is always of Lenz-Barlotti class IVa. 1.

The automorphism groups of semifields of square order are also briefly investigated.

1. The Construction of Quasifields from Derivable Semifields.

We will assume that the reader is familiar with the concept of "derivation." For background material the reader is referred to [2], [4], [6], and [7].

DEFINITION 1.1. A semifield $\mathscr{S} = (\mathscr{S}, +, \cdot)$ of order q^2, $q = p^r$, p a prime, will be said to be *derivable* if and only if $(\mathscr{S}, +)$ is a vector space over $GF(q) = F$ where $F \subseteqq \mathscr{S}$ and $x \cdot \alpha = x\alpha$ (or $\alpha \cdot x = \alpha x$) is scalar product.

If a semifield $\mathscr{S}$ is derivable then either $\mathscr{S}$ or dual $\mathscr{S}$ (i.e., right multiplication becomes left multiplication, and conversely) is a right vector space over $GF(q)$ and hence either the affine plane π coordinatized by $\mathscr{S}$ or an affine restriction of the dual of the projective extension of π is derivable (see sections 3 and 4, [7]).

A projective plane is a semifield plane if and only if it can be coordinatized by a semifield or if and only if the plane is $(P, \underline{l})$-transitive $\forall$ points $P \in \underline{l}$, and $(\underline{Q}, l)$-transitive $\forall$ lines $l \in \underline{Q}$ and $\underline{Q} \in \underline{l}$.

If $\underline{Q}, \underline{l}$ are chosen to be (∞) and l_∞, respectively, then the coordinate structure obtained is a semifield. In dualizing the semifield plane π we shall let $(\infty) \leftrightarrow l_\infty$ and then delete l_∞ to obtain an affine plane coordinatized by a semifield dual to a semifield which coordinatizes π.

DEFINITION 1.2. Let $\mathscr{S} = (\mathscr{S}, +, \cdot)$ be a derivable semifield. $\mathscr{S}$ is *subcommutative* if and only if $a\alpha = \alpha a$ for all $a \in \mathscr{S}$ and for all $\alpha \in GF(q)$.

DEFINITION 1.3. A semifield $\mathscr{S}$ of order q^2 containing $GF(q)$ is a weak nucleus semifield (wn-semifield) if and only if $(ab)c = a(bc)$ whenever any two of a, b, c are in $GF(q)$.

Note that a wn-semifield of order q^2 is derivable and a derivable subcommutative semifield is a wn-semifield.

Let $\mathcal{S}$ be a derivable semifield which is a *right* 2-dimensional vector space over $GF(q)$. Let $\{1, t\}$, $t \in \mathcal{S} - GF(q)$ be a basis for $\mathcal{S}$ over $GF(q)$.

Then let $\beta(t\alpha) = th(\beta, \alpha) + k(\beta, \alpha)$ and $(t\alpha)(t\beta) = tf(\alpha, \beta) + g(\alpha, \beta)$ for $\alpha, \beta \in GF(q)$ where h, k, f, g are bilinear functions: $GF(q) \times GF(q) \rightarrow GF(q)$ which introduce no zero divisors into the multiplication.

Then multiplication in the semifield is given by:

$$(t\alpha + \delta)(t\beta + \gamma) = t(f(\alpha, \beta) + h(\delta, \beta) + \alpha\gamma) + (g(\alpha, \beta) + k(\delta, \beta) + \delta\gamma) .$$

Thus, if $\mathcal{S}$ is any derivable semifield then either the multiplication of $\mathcal{S}$ or dual $\mathcal{S}$ is of the above form.

THEOREM 1.4. *Let $\mathcal{S} = (\mathcal{S}, +, \cdot)$ be a derivable semifield which is a right vector space of dimension 2 over $F = GF(q)$, $q = p^r$, p a prime. Let the multiplication in $\mathcal{S}$ be given by*:

$$(t\alpha + \delta) \cdot (t\beta + \gamma) = t(f(\alpha, \beta) + h(\delta, \beta) + \alpha\gamma) + (g(\alpha, \beta) + k(\delta, \beta) + \delta\gamma) \ \forall\, \alpha, \beta, \delta, \gamma \in F$$

where f, h, g, k are bilinear functions: $F \times F \rightarrow F$.

Define a system $\mathcal{S}^* = (\mathcal{S}, +, *)$ when the $*$-multiplication is given by

$$t * \alpha = t\alpha,\ (t\alpha + \beta) * \gamma = t(\alpha\gamma) + \beta\gamma \quad \text{and if } \delta \neq 0$$

$(t\alpha + \beta) * (t\delta + \gamma) = t\rho + \chi$ where

(1) $h(\delta, \mu_1) = 1$,

(2) $k(\delta, \mu_1) + \delta\mu_2 = \gamma$,

(3) $f(\alpha, \mu_1) + h(\rho, \mu_1) + \alpha\mu_2 = \beta$,

(4) $g(\alpha, \mu_1) + k(\rho, \mu_1) + \rho\mu_2 = \chi$

$\forall \alpha, \beta, \delta \neq 0, \gamma \in F$ where μ_1, μ_2 and thus $\rho, \chi \in F$ are determined from the above equations.

Then $\mathcal{S}^* = (\mathcal{S}, +, *)$ is a (right) quasifield.

Proof. The affine plane π coordinatized by $\mathcal{S}$ is derivable (see [2], [6], [7]). Ostrom [6] has shown that the plane π' derived from π is a translation plane and may be coordinatized by a system

$$(\mathcal{S}, +, {}^*) \ni t\alpha = t * \alpha,\ (t\alpha + \beta) * (t\delta + \gamma) = t\rho + \chi$$

if and only if $(t\alpha + \rho)(t\mu_1 + \mu_2) = t\beta + \chi$ where $\delta(t\mu_1 + \mu_2) = t + \gamma$ for $\delta \neq 0$, and $(t\alpha + \beta) * \gamma = (t\alpha + \beta)\gamma$ for all $\alpha, \beta, \delta, \gamma \in GF(q)$. Our equations are obtained by merely equating vector components.

We shall now specialize (1.4) to the case where $\mathscr{S}$ is a *wn*-semifield.

Knuth [4] has shown that if $\mathscr{S}$ is a *wn*-semifield then a basis $\{1, t\}$ can be chosen so that $\alpha t = t\alpha^\sigma \forall \alpha \in GF(q)$ where σ is some automorphism of $GF(q)$. In this case, $h(\delta, \beta) = \delta^\sigma \beta$ and $k(\delta, \beta) = 0$ for all $\delta, \beta \in GF(q)$.

Thus $h(\delta, \mu_1) = \delta^\sigma \mu_1 = 1$ implies $\mu_1 = \delta^{-\sigma}$ and $k(\delta, \mu_1) + \delta\mu_2 = \gamma$ implies that $\mu_2 = \delta^{-1}\gamma$ for $\delta \neq 0$. Thus $f(\alpha, \mu_1) + h(\rho, \mu_1) + \alpha\mu_2 = \beta$ implies that $f(\alpha, \delta^{-\sigma}) + \rho^\sigma \delta^{-\sigma} = \alpha\delta^{-1}\gamma = \beta$. Hence

$$\rho = ((\beta - f(\alpha, \delta^{-\sigma}) - \alpha\delta^{-1}\gamma)\delta^\sigma)^{\sigma^{-1}} = (\beta - f(\alpha, \delta^{-\sigma}) - \alpha\delta^{-1}\gamma)^{\sigma^{-1}}\delta .$$

Also, $g(\alpha, \mu_1) + k(\rho, \mu_1) + \rho\mu_2 = \chi$ implies that $g(\alpha, \delta^{-\sigma}) + \rho\delta^{-1}\gamma = \chi$.

Thus, we have the following theorem.

THEOREM 1.5. *If $\mathscr{S} = (\mathscr{S}, +, \cdot)$ is a weak nucleus semifield of order $q^2 \ni$ multiplication in $\mathscr{S}$ is given by*

$$(t\alpha + \delta)(t\beta + \gamma) = t(f(\alpha, \beta) + \delta^\sigma \beta + \alpha\gamma) + (g(\alpha, \beta) + \delta\gamma) .$$

Define a system $\mathscr{S}^* = (\mathscr{S}, +, *)$ by defining a $*$-multiplication as follows:

$$\begin{aligned} t * \alpha = t\alpha, (t\alpha + \delta) * (t\beta + \gamma) &= t(\delta - f(\alpha, \beta^{-\sigma}) - \alpha\beta^{-1}\gamma)^{\sigma^{-1}}\beta \\ &+ g(\alpha, \beta^{-\sigma}) + (\delta - f(\alpha, \beta^{-\sigma}) - \alpha\beta^{-1}\gamma)^{\sigma^{-1}}\gamma \end{aligned}$$

for $\delta \neq 0$ and σ an automorphism of $GF(q)$, and

$$(t\alpha + \delta) * \gamma = (t\alpha + \delta)\gamma \forall \alpha, \beta, \delta, \gamma \in GF(q) .$$

Then $\mathscr{S}^*$ is a (right) quasifield.

REMARKS 1.6. Under the assumptions of (1.5)

(i) $\alpha * a = a * \alpha^{\sigma^{-1}} \forall \alpha \in GF(q)$ and $\forall a \in \mathscr{S} - GF(q)$,

(ii) $(a * b) * c = a * (b * c)$ whenever any two of a, b, c are in $GF(q)$.

Proof. The proof of (1.6) is routine and is left to the reader.

2. Automorphisms of derivable semifields which fix $GF(q)$ elementwise. The semifields of order 16 have been tabulated, [3], and are all isotopic (Sec. 3, [4]) to one of two weak nucleus semifields, each of which admits a group of automorphisms of order 3 which fixes $GF(q)$ elementwise (see [4]). The multiplications for the two systems are given by $(t\alpha)(t\delta) = t\alpha^2\delta^2 + \alpha^2\delta$, $\beta t = t\beta^2 \forall \alpha, \delta, \beta \in GF(4)$ and $(t\alpha)(t\delta) = \omega\alpha^2\delta$, $\beta t = t\beta^2$ where ω is a primitive root of $GF(4)$.

The semifields of order 16 are exceptions among derivable semifields of order q^2 in that no derivable semifield of order q^2, $q > 4$ can admit such automorphism groups.

THEOREM 2.1 *Let* $(\mathscr{S}, +, \cdot)$ *be a derivable proper semifield of order* q^2. *Then* $\mathscr{S}$ *is of order* 16 *if and only if a derivable isotopic image of* $\mathscr{S}$ *admits a group of automorphisms of order* $q - 1$ *which fixes* $GF(q)$ *elementwise.*

Proof. Suppose the indicated automorphisms τ_ρ that the form $t^\tau\rho = t\rho \forall \rho \in GF(q) - \{0\}$. (Note: This would be true by (2.2) if $\mathscr{S}$ is a wn-semifield and $\sigma \neq 1$, but we are not necessarily assuming this property.) If $\mathscr{S}$ is a left vector space over $GF(q)$, consider dual $\mathscr{S}$. Let $\{1, t\}$ be a basis for $\mathscr{S}$ or dual $\mathscr{S}$.

$((t\alpha)(t\beta))^{\tau_\rho} = (tf(\alpha, \beta) + g(\alpha, \beta))^{\tau_\rho}$ where f, g are bilinear functions: $GF(q) \times GF(q) \to GF(q)$. Thus,

$$(t(\rho\alpha))(t(\rho\beta)) = t(\rho f(\alpha, \beta)) + g(\alpha, \beta)$$

which implies that $\rho f(\alpha, \beta) = f(\rho\alpha, \rho\beta)$ and $g(\alpha, \beta) = g(\rho\alpha, \rho\beta)$. Since we have $q - 1$ automorphisms τ_ρ these previous equations are true for all $\alpha, \beta, \rho \in GF(q) - \{0\}$. If characteristic $F \neq 2$ then $g(2\rho, 2\rho) = g(2, 2)$. But g is bilinear so $g(2, 2) = 4g(1, 1)$. Also $g(\alpha, \alpha) = g(1, 1)$ so that $4g(1, 1) = g(1, 1)$. Moreover $g(1, 1) \neq 0$ since $t^2 = tf(1, 1) + g(1, 1)$ and multiplication of nonzero elements is a loop.

Hence $4 \equiv 1$ so that characteristic $F = 3$.

Since $g(\rho\alpha, \rho\beta) = g(\alpha, \beta) \forall \alpha, \beta, \rho \in GF(q) - \{0\}$ then

$$g(1, (\alpha + \gamma)^{-1}) = g(\alpha + \gamma, 1) = g(\alpha, 1) + g(\gamma, 1)$$

for $\alpha + \gamma \neq 0$.

Thus, $g(1, (\alpha + \gamma)^{-1}) - (g(\alpha, 1) + g(\gamma, 1)) = 0$, which implies that $g(1, (\alpha + \gamma)^{-1}) + 2(g(\alpha, 1) + g(\gamma, 1)) = 0$.

Clearly, $2g(\beta, 1) = g(2\beta, 1) \forall \beta \in GF(q)$, and $g(2\beta, 1) = g(1, 2\beta^{-1})$, so

$$\begin{aligned} g(1, (\alpha + \gamma)^{-1}) &+ g(2\alpha, 1) + g(2\gamma, 1) \\ &= g(1, (\alpha + \gamma)^{-1}) + g(1, 2\alpha^{-1}) + g(1, 2\gamma^{-1}) \\ &= g(1, (\alpha + \gamma)^{-1} + 2\alpha^{-1} + 2\gamma^{-1}) \\ &= g(1, (\alpha + \gamma)^{-1} - (\alpha^{-1} + \gamma^{-1})) . \end{aligned}$$

If $(\alpha + \gamma)^{-1} \neq \alpha^{-1} + \gamma^{-1}$, then

$$t(t((\alpha + \gamma)^{-1} - (\alpha^{-1} + \gamma^{-1})) = tf(1, (\alpha + \gamma)^{-1} - (\alpha^{-1} + \gamma^{-1}))$$

which cannot be the case. Hence $(\alpha+\gamma)^{-1}=\alpha^{-1}+\gamma^{-1}$. It is easy to see that in this situation $GF(q)=GF(3)$.

But then $\mathscr{S}$ would be a field ([4], p. 208) contrary to our assumption.

Hence, characteristic $F=2$. Then, using the bilinearity of g we may argue as before (except that $-1=+1$) to obtain $(\alpha+\gamma)^{-1}=\alpha^{-1}+\gamma^{-1}$ from which it follows that $GF(q)=GF(4)$.

To complete the proof of (2.1) we must show that the automorphisms τ_ρ have the form $t^\tau\rho=t\rho$.

Let π be the affine plane coordinatized by $\mathscr{S}$ and let π_0 be the subplane of π coordinatized by $GF(q)$.

The automorphism group of $\mathscr{S}$ induces a collineation group of π which fixes π_0 pointwise. In the derived plane there is a collineation group of order $q-1$ fixing the line $\{(x, y)\mid x=0\}$ pointwise. (The validity of this last statement may be seen by choosing coordinates for the derived plane so that π_0 in π is the point set $\{(x, y)\mid x=0\}$ in the derived plane. See e.g. [6], Theorem 10.)

Thus, the derived plane π' admits a $(P, x=0)$-homology group of order $q-1$ (see [2], remarks following (2.6)). Moreover, this group must fix the set points of π_0' on the line at infinity of the derived plane where π_0' is the line $x=0$ in π (see [6], Theorem 7). Hence, $P=(\alpha)$ where $\alpha\in GF(q)$. If $\alpha\neq 0$ we can rechoose t in $\mathscr{S}$ so that P is represented by (0).

Now $\{(t\delta+\alpha\delta, t\beta+\alpha\beta)\}$ in π is the same as $\{(t\delta+\beta, t\alpha\delta+\alpha\beta)\}$ in π' ([6], Theorem 10). If we let $t=t+\alpha$ then $\{(t\delta, t\beta)\}$ is $\{(t\delta+\beta, 0)\}$ in π'. Hence, we have relabeled $\{(x, y)\mid y=x\alpha\}$ in π' by $\{(x, y)\mid y=0\}$. Thus, $P=(\alpha)$ is relabeled by (0).

Now a group of $((0), x=0)$-collineations which fix π_0' induce automorphisms of the form $\tau_\rho\ni(t\alpha+\beta)\tau_\rho=t(\rho\alpha)+\beta$ in $\mathscr{S}$ (see [2], (2.10), and the proof of (3.10)).

Hence (2.1) is proved.

PROPOSITION 2.2. *Let $(\mathscr{S}, +, \cdot)$ be a wn-semifield of order q^2 with multiplication defined by $(t\alpha)(t\beta)=tf(\alpha,\beta)+g(\alpha,\beta)$, $\delta t=t\delta^\sigma$, σ an automorphism of $GF(q)$, $\forall\alpha,\beta,\delta\in GF(q)$. If $\sigma\neq 1$, and if τ is any automorphism of $(\mathscr{S}, +, \cdot)$ fixing $GF(q)$ elementwise then $(t\alpha+\beta)^\tau=t(\rho\alpha)+\beta$ for some $\rho\in GF(q)$.*

Proof. $(\alpha t)^\tau=\alpha^\tau t^\tau=\alpha t^\tau$. Let $t^\tau=t\rho+\theta$ for some $\rho,\theta\in GF(q)$. Then $\alpha t^\tau=t\alpha^\sigma\rho+\alpha\theta$ and $(\alpha t)^\tau=(t\alpha^\sigma)^\tau$, $t^\tau\alpha^\sigma=t\rho\alpha^\sigma+\theta\alpha^\sigma$. Hence, $\alpha\theta=\theta\alpha^\sigma$ which implies $\theta=0$.

THEOREM 2.3. *If a derivable semifield $\mathscr{S}=(\mathscr{S}, +, \cdot)$ of order $q^2, q>2$ admits a nontrivial automorphism group $\mathscr{G}$ which fixes*

$GF(q) = F$ elementwise and $|\mathscr{G}|\,|q$ then $\mathscr{G}$ is an elementary abelian 2-group whose order is strictly less than q.

Proof. Without loss of generality, suppose that $(\mathscr{S}, +)$ is a *right* vector space over F. Then it follows directly from [5], Theorem 1, that if $\tau \in \mathscr{G}$ and $\{1, t\}$ is a basis for $(\mathscr{S}, +)$ over F then $t^\tau = t + \gamma$ for some $\gamma \in F$.

Let $\delta(t\beta) = th(\delta, \beta) + k(\delta, \beta)$,

$$(t\alpha)(t\beta) = tf(\alpha, \beta) + g(\alpha, \beta) \forall \alpha, \beta, \delta \in GF(q)$$

where f, g, h, k are bilinear functions: $GF(q) \times GF(q) \rightarrow GF(q)$.

Then, $(t\alpha)(t\beta)^\tau = (tf(\alpha, \beta) + g(\alpha, \beta))^\tau$ if and only if

$$\begin{aligned}(t\alpha)(t\beta) &+ t(h(\gamma\alpha, \beta) + \alpha\gamma\beta) \\ &+ k(\gamma\alpha, \beta) + \gamma^2\alpha\beta = (t\alpha)(t\beta) + \gamma f(\alpha, \beta) .\end{aligned}$$

Equating vector components:

(1) $h(\gamma\alpha, \beta) = -\alpha\gamma\beta \forall \alpha, \beta$ and

(2) $k(\gamma\alpha, \beta) + \gamma^2\alpha\beta = \gamma f(\alpha, \beta)$.

If $\alpha = \gamma^{-1}$ in (1), then $h(1, \beta) = -\beta$. But, $h(1, \beta) = \beta$. $\therefore F$ is of characteristic 2. Thus, $\mathscr{G}$ is an elementary abelian 2-group.

Now assume $|\mathscr{G}| = q$. Then, by (2), $k(1, \beta) + \gamma\beta = \gamma f(\gamma^{-1}, \beta) = \gamma\beta$ so that $f(\gamma^{-1}, \beta) = \beta$ for all $\gamma \in F$. But

$$f(\mathscr{N}^{-1} + \gamma^{-1}, \beta) = f(\mathscr{N}^{-1}, \beta) + f(\gamma^{-1}, \beta) = 0$$

since f is bilinear and F is of characteristic 2.

Hence, (2.3) is proved.

COROLLARY 2.4. *If $\mathscr{S} = (\mathscr{S}, +, \cdot)$ is a wn-semifield of order q^2 which admits a nontrivial automorphism group $\mathscr{G}$ such that $|\mathscr{G}|\,|q$ then $|\mathscr{G}| = 2$.*

Proof. By (2.3)(2), $k(\gamma\alpha, \beta) + \gamma^2\alpha\beta = \gamma f(\alpha, \beta)$.

We may choose $t \in \mathscr{S} - F \ni k(\gamma\alpha, \beta) \equiv 0 \forall \alpha, \beta, \gamma \in F$ so $\gamma^2\alpha\beta = \gamma f(\alpha, \beta) \Rightarrow \gamma\alpha\beta = f(\alpha, \beta)$. Clearly $|\mathscr{S}| = 2$ for otherwise it would follow that $\gamma\alpha\beta = \mathscr{N}\alpha\beta$ for $\gamma \neq \mathscr{N} \forall \alpha, \beta \in F$.

COROLLARY 2.5. *If $\mathscr{S} = (\mathscr{S}, +, \cdot)$ is a wn-semifield which admits a group $\mathscr{G}$ of (2.4) then there is a $t \in \mathscr{S} - F$ such that*

$$(t\alpha + \delta)(t\beta + \gamma) = t(\alpha\beta f + \delta\beta + \alpha\gamma) + (g(\alpha, \beta) + \delta\gamma)$$

where g is a bilinear function $F \times F \rightarrow F$ and f is a nonzero constant in F.

Proof. $\exists t \in \mathscr{S} - F \ni \alpha t = t\alpha^{\sigma} \forall \alpha \in F$, σ an automorphism of F. By (2.2), $\sigma = 1$. By (2.4), $|\mathscr{S}| = 2$ and if $\tau \in \mathscr{S} \ni t^{\tau} = t + ff(\alpha, \beta) = \alpha\beta f$.

COROLLARY 2.6. *Let* $(\mathscr{S}, +, \cdot)$ *satisfy the hypothesis of* (2.3) *and* $(\mathscr{S}, +, *)$ *the quasifield of* (1.4). *Consider the following distributive law*:

$$c * (\alpha + b) = c * \alpha + c * b$$

for all $c, b \in \mathscr{S}$ *and for some* $\alpha \in F$.

Then

(i) if char $F \neq 2$ this distributive law cannot hold for any nonzero $\alpha \in F$,

(ii) if char $F = 2$ and $(\mathscr{S}, +, \cdot)$ is a *wn*-semifield then the distributive law holds for at most a single nonzero element of F,

(iii) if char $F = 2$ this distributive law cannot hold for all $\alpha \in F$.

Thus, in particular, $(\mathscr{S}, +, *)$ is not a semifield.

Proof. The given distributive law induces a $((\infty), x = 0, \pi_0)$-collineation in the affine plane coordinatized by $(\mathscr{S}, +, *)$ and hence ([2], see the proof of (3.10)) an automorphism group in $(\mathscr{S}, +, \cdot)$ as in (2.3).

We have seen that $(\mathscr{S}, +, *)$, if $\mathscr{S}$ is a *wn*-semifield, admits some associative properties ((1.6) (ii)). In general, however, we note that $(\mathscr{S}, +*)$ cannot be associative.

THEOREM 2.7. *If* $\mathscr{S} = (\mathscr{S}, +, \cdot)$ *is a derivable semifield* $\ni (\mathscr{S}, +)$ *is a right vector space over* $GF(q)$ *then* $(\mathscr{S}, +, *)$ *is neither associative nor distributive.*

Proof. The affine plane coordinatizing $(\mathscr{S}, +, \cdot)$ is $((\infty), x = 0, \pi_0)$-transitive ([2], [6]) and thus $(\mathscr{S}, +, *)$ admits a group of automorphisms of order q which fix $GF(q)$ elementwise. But regular nearfields clearly cannot admit such automorphisms. The irregular nearfields all have order p^2 where p is a prime. If $\mathscr{S}$ has order p^2 then $\mathscr{S}$ is a field ([4]) in which case $(\mathscr{S}, +, *)$ is a quasifield which coordinatizes a Hall plane.

3. The Knuth multiplication.

Let $(\mathscr{S}, +) = (GF(q^2), +)$. Let $t \in \mathscr{S} - GF(q)$ and define $\alpha t = t\alpha^{\sigma}$ where σ is an automorphism of $GF(q)$. The functions $f(\alpha, \beta) = \alpha^{\mathscr{N}}\beta^{\chi}f$, $g(\alpha, \beta) = \alpha^{\rho}\beta^{\delta}g$ where $\mathscr{N}, \chi, \rho, \delta$ are automorphisms of $GF(q)$, $\alpha, \beta \in GF(q)$, f, g constants in $GF(q)$ are bilinear functions: $GF(q) \times GF(q) \to GF(q)$.

$$\alpha t = t\alpha^{\sigma},\ (t\alpha)(t\beta) = t\alpha^{\mathscr{N}}\beta^{\chi}f + \alpha^{\rho}\beta^{\delta}g$$

will define multiplication of a semifield $\mathscr{S} = (\mathscr{S}, +, \cdot)$ provided no zero divisors are introduced by the choices of $\sigma, \mathscr{N}, \chi, \rho, \delta, f$ and g. If no zero divisors occur, we shall say that the semifield so defined is a Knuth Semifield.

THEOREM 3.1. *(Knuth [4]). Let*

$$\mathscr{S} = (\mathscr{S}, +, \cdot) \ni (\mathscr{S}, +) = GF(q^2)$$

and

$$(t\alpha + \delta)(t\beta + \gamma) = t[\alpha^{\mathscr{N}}\beta^{\chi}f + \alpha\gamma + \delta^{\sigma}\beta] + [\alpha^{\rho}\beta^{\delta}g + \delta\gamma]\forall\alpha, \beta, \delta, \gamma \in GF(q)$$

where $\mathscr{N}, \chi, \sigma, \rho, \delta$ are automorphisms of $GF(q)$ and f, g elements of $GF(q)$.

(a) *If $f = 0$ and g is a nonsquare in $GF(q)$ then the above multiplication defines a Knuth Semifield for an arbitrary choice of automorphisms σ, ρ, δ.*

That is, $\alpha t = t\alpha^{\sigma}$, $(t\alpha)(t\beta) = \alpha^{\rho}\beta^{\sigma}g$ for arbitrary automorphisms ρ, δ of $GF(q)$ and g a nonsquare in $GF(q)$ define a semifield.

(b) *If $f \neq 0$ and σ, f, g are chosen so that $y^{\sigma+1} + fy - g = 0$ has no solutions in $GF(q)$ and $(\mathscr{N}, \chi, \rho, \delta) = (\sigma, \sigma^{-1}, \sigma, \sigma^{-2})$, $(\sigma, 1, \sigma, 1)$, $(1, \sigma^{-1}, \sigma^{-1}, \sigma^{-2})$ or $(1, 1, \sigma^{-1}, 1)$ then the above multiplication defines a Knuth Semifield. That is, each of the following multiplications define a class of semifields:*

I. $\alpha t = t\alpha^{\sigma}$, $(t\alpha)(t\beta) = t\alpha^{\sigma}\beta^{\sigma^{-1}}f + \alpha^{\sigma}\beta^{\sigma^{-2}}g$

II. $\alpha t = t\alpha^{\sigma}$, $(t\alpha)(t\beta) = t\alpha^{\sigma}\beta f + \alpha^{\sigma}\beta g$

III. $\alpha t = t\alpha^{\sigma}$, $(t\alpha)(t\beta) = t\alpha\beta^{\sigma^{-1}}f + \alpha^{\sigma^{-1}}\beta^{\sigma^{-2}}g$

IV. $\alpha t = t\alpha^{\sigma}$, $(t\alpha)(t\beta) = t\alpha\beta f + \alpha^{\sigma^{-1}}\beta g$.

Furthermore, Knuth [4] has characterized types II, III and IV in terms of the nuclei.

DEFINITION 3.2. Let $(Q, +, \cdot)$ be a ternary system. Let

$$\{x \in Q \mid (ab)x = a(bx)\forall a, b \in Q\} = \mathscr{N}_{\mathscr{R}Q},$$
$$\{x \in Q \mid (ax)b = a(xb)\forall a, b \in Q\} = \mathscr{N}_{\mathscr{M}Q},$$
$$\{x \in Q \mid (xa)b = x(ab)\forall a, b \in Q\} = \mathscr{N}_{\mathscr{L}Q},$$

$\mathscr{N}_{\mathscr{R}Q}, \mathscr{N}_{\mathscr{M}Q}, \mathscr{N}_{\mathscr{L}Q}$ will be called the *right*, *middle*, and *left nucleus* of Q, respectively.

THEOREM 3.3. *(Knuth [4]). Let $(\mathscr{S}, +, \cdot)$ be a Knuth Semifield of order q^2. Then $GF(q) = \mathscr{N}_{\mathscr{R}\mathscr{S}} = \mathscr{N}_{\mathscr{M}\mathscr{S}}$ if and only if $\mathscr{S}$ is of type II. $GF(q) = \mathscr{N}_{\mathscr{L}\mathscr{S}} = \mathscr{N}_{\mathscr{M}\mathscr{S}}$ if and only if $\mathscr{S}$ is of type III, and $GF(q) = \mathscr{N}_{\mathscr{L}\mathscr{S}} = \mathscr{N}_{\mathscr{R}\mathscr{S}}$ if and only if $\mathscr{S}$ of type IV.*

By applying (1.4) to (3.1), we obtain the following result:

THEOREM 3.4. *Each of the following multiplications* $*$ *(with field addition) defines a (right) quasifield which is neither a semifield or nearfield. If* $\beta \neq 0$,

(1) $$(t\alpha + \delta) * (t\beta + \gamma) = t(\delta - \alpha\beta^{-1}\gamma)^{\sigma^{-1}}\beta + (\delta - \alpha\beta^{-1}\gamma)^{\sigma^{-1}}\gamma + \alpha^{\mathscr{N}}\beta^{-\sigma\chi}g,\ g \text{ a nonsequare in } F$$

(2) $$(t\alpha + \delta) * (t\beta + \gamma) = t(\delta - \alpha^{\sigma}\beta^{-1}f - \alpha\beta^{-1}\gamma)^{\sigma^{-1}}\beta + (\delta - \alpha^{\sigma}\beta^{-1}f - \alpha\beta^{-1}\gamma)^{\sigma^{-1}}\gamma + \alpha^{\sigma}\beta^{-\sigma^{-1}}g,\ \sigma \neq 1,\ f \neq 0$$

(3) $$(t\alpha + \delta) * (t\beta + \gamma) = t(\delta - \alpha^{\sigma}\beta^{-\sigma}f - \alpha\beta^{-1}\gamma)^{\sigma^{-1}}\beta + (\delta - \alpha^{\sigma}\beta^{-\sigma}f - \alpha\beta^{-1}\gamma)^{\sigma^{-1}}\gamma + \alpha^{\sigma}\beta^{-\sigma}g,\ \sigma \neq 1,\ f \neq 0$$

(4) $$(t\alpha + \delta) * (t\beta + \gamma) = t(\delta - \alpha\beta^{-1}f - \alpha\beta^{-1}\gamma)^{\sigma^{-1}}\beta + (\delta - \alpha\beta^{-1}f - \alpha\beta^{-1}\gamma)^{\sigma^{-1}}\gamma + \alpha^{\sigma^{-1}}\beta^{\sigma^{-1}}g,\ \sigma \neq 1,\ f \neq 0$$

(5) $$(t\alpha + \delta) * (t\beta + \gamma) = t(\delta - \alpha\beta^{-\sigma}f - \alpha\beta^{-1}\gamma)^{\sigma^{-1}}\beta + (\delta - \alpha\beta^{-\sigma}f - \alpha\beta^{-1}\gamma)^{\sigma^{-1}}\gamma + \alpha^{\sigma^{-1}}\beta^{-\sigma}g,\ \sigma \neq 1,\ f \neq 0.$$

Also, $(t\alpha + \delta) * \gamma = t(\alpha\gamma) + \delta\gamma$ *where* σ *is an automorphism of* F *and in cases* (2) *through* (5) $y^{\sigma+1} + fy - g \neq 0 \forall y \in GF(q)$ *and* $\mathscr{N}, \chi$ *automorphisms of* F *in case* (1).

Proof. See (1.4), (2.7) and (3.1).

4. The planes coordinatized by the $(\mathscr{S}, +, *)$ **quasifields.** A plane Σ is of Lenz-Barlotti Class IV.a.2 or IV.a.3 if Σ can be coordinatized by a (right) nearfield, and of Class V.1 if Σ can be coordinatized by a semifield. Σ is of Class IV.a.1 if Σ is coordinatized by (right) quasifield but no coordinate system for Σ is a (right) nearfield or semifield.

The planes coordinatized by the $(\mathscr{S}, +, *)$ quasifields are therefore of L-B Classes IV.a.1, a.2, a.3, or V.

THEOREM 4.1. *Let* $\mathscr{S} = (\mathscr{S}, +, \cdot)$ *be a derivable semifield* $\ni (\mathscr{S}, +)$ *is a right vector space over* $GF(q)$. *Let* π *be the semifield plane coordinatized by* $\mathscr{S}$. π *is derivable, so let* π' *be the plane derived from* π. *Then* π' *is of Lenz-Barlotti Class* IV.a.1.

Proof. We must show that π' cannot be of type IV.a.2, a.3, or V.1.

Suppose π' is of type V.1, then π' is $((m), l)$-transitive for all lines l incident with (m) where $mI l_\infty$. By (2.7), $(m) \neq (\infty)$ since $\mathscr{S}^* = (\mathscr{S}, +, *)$ is not a semifield. Clearly (m) is fixed by the full collineation group of π' (otherwise π' is Desarguesian and every coordinatizing structure is a field). Recall (see proof of (2.7)), $\mathscr{S}^*$ admits an automorphism group of order q fixing F pointwise such that $t \to t + \alpha$ for all $\alpha \in F$ (see (2.3) and (2.7)). Hence, $m \in F$ if π' is $((m), l)$-transitive.

We consider two cases:

$$(1) \quad (m) = (0), \qquad (2) \quad m \neq (0) .$$

Case (1). If $(m) = (0)$, consider changing coordinates as follows in $\mathscr{S}$ (in π):

$$(tx_1 + x_2, ty_1 + y_2) \xrightarrow[\text{change } \sigma]{\text{coordinate}} (tx_2 + x_1, ty_2 + y_1) \forall x_1, x_2, y_1, y_2 \in F .$$

$\mathscr{S}\sigma$ is a derivable semifield (see [2], the proofs of (3.6) and (3.7)).

The coordinate change appears as $(x, y) \to (y, x)$ in π' (see [2], (3.7)) and thus induces a Hall coordinate system $\mathscr{S}_{\mathscr{R}}^* \ni \pi'$ is $((\infty), x = 0)$-transitive. $\therefore$ $\mathscr{S}_{\mathscr{R}}^*$ is a (derivable) semifield. However, $\mathscr{S}_{\mathscr{R}}^*$ is constructed from $\mathscr{S}_{\mathscr{R}} = \mathscr{S}^{\sigma}$ in the same manner that $\mathscr{S}^*$ is constructed from $\mathscr{S}$. $\therefore$ we have a contradiction by (2.7).

(2) $(m) \neq (0)$.

Choose $\bar{t} = t + m$ (recall $m \in F$) in $(\mathscr{S}, +, \cdot)$. Then in π'

$$(y = xm) = \{(x, y) \mid x = t\alpha + \beta, y = t(\alpha m) + (\beta m)\}$$

is the same as $\{(t\alpha + \alpha m, t\beta + \beta m) = (\bar{t}\alpha, \bar{t}\beta)\} \equiv y = 0$ in π'. Hence, by case (1) we have a contradiction.

Assume that π' is of type IV.a.2 or a.3. Then π' is $((P), (Q))$-transitive for some pair of points $(P), (Q), P \neq Q$.

Moreover, every collineation of π' must fix $\{(P), (Q)\}$. Therefore, since $\mathscr{S}^*$ admits an automorphism group of order q it must be that $P, Q \in F$ or $P, Q = \infty$.

Now if we can change coordinates so that $\mathscr{S}_{\mathscr{R}}^*$ is a nearfield and $\mathscr{S}_{\mathscr{R}}^*$ admits an automorphism group of order q, then we have a contradiction since the order of an automorphism group of a nearfield of order $q^2 (q = p^r, r > 1)$ is never this large.

Let $(P) = (\alpha)$ and $(Q) = (\beta)$, $\alpha, \beta \in F$ or $\alpha, \beta = \infty$.

Case (1). $(\alpha) = (\infty)$. Since $\mathscr{S}^*$ is not a nearfield (see (2.7)), $(\beta) \neq (0)$. We can rechoose t in $\mathscr{S}$ (in π) so that $y = x\beta$ is $y = 0$ in π' (i.e., if $\bar{t} = t + \beta$) and (∞) in π' is left fixed. $\therefore$ $\mathscr{S}^*$ with the basis $\{1, \bar{t}\}$ is a nearfield and admits q automorphisms.

Case (2). $(\alpha) \neq (\infty), (\beta) \neq (\infty), (\alpha) = (0)$. We can move (0) to (∞) by the $(x, y) \to (y, x)$ coordinate change of $\mathscr{S}^*$ of the previous argument. Therefore, π' is $((\infty), (\gamma))$-transitive for $(\gamma) \neq (0)$. Then, we may rechoose t in $\mathscr{S}_{\mathscr{R}}$ so that (γ) is (0) in $\mathscr{S}_{\mathscr{R}}^*$ (or in π'). Since $\mathscr{S}_{\mathscr{R}}$ is a (derivable) semifield, $\mathscr{S}_{\mathscr{R}}^*$ admits an automorphism group of order q which is a contradiction.

Case (3). $(\alpha), (\beta) \neq (\infty)$ or (0). First rechoose t in $\mathscr{S}$ so that (α) is (0), then repeat Case 2.

REMARKS. If $(\mathscr{S}, +, \cdot)$ is a derivable subcommutative semifield then a "derivable chain" (see [1]) can be constructed based on the affine plane coordinatized by $(\mathscr{S}, +, \cdot)$.

$(\mathscr{S}, +, \cdot)$ actually need not be finite to construct $(\mathscr{S}, +, *)$. That is, Ostrom's "derivation process" extends for infinite translation planes. We shall explore this in a later paper.

REFERENCES

1. N. L. Johnson, *Derivable chains of planes*, Bol. Un. Mat. Ital. N. 2, (1970) 167-184.
2. ———, *Derivable semi-translation planes*, Pacific J. Math., **34** (1970).
3. E. Kleinfeld, *Techniques for enumerating Veblen-Wedderburn systems*, J. Assoc. Comput. Math., **7** (1960), 330-337.
4. D. E. Knuth, *Finite semifields and projective planes*, J. Algebra, **2** (1965), 182-217.
5. D. L. Morgan and T. G. Ostrom, *Coordinate systems of some semi-translation planes*, Trans. Amer. Math. Soc., **111** (1964), 19-32.
6. T. G. Ostrom, *Semi-translation planes*, Trans. Amer. Math. Soc., **111** (1964), 1-18.
7. ———, *Vector spaces and construction of finite projective planes*, Arch. Math., **19** (1968), 1-25.

Received June 3, 1970.

UNIVERSITY OF IOWA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

QUASI-PROJECTIVE AND QUASI-INJECTIVE MODULES

ANNE KOEHLER

This paper contains results which are needed to prove a decomposition theorem for quasi-projective modules over left perfect rings.

An R-module M is called quasi-projective if and only if for every R-module A, every R-epimorphism $q: M \to A$, and every R-homomorphism $f: M \to A$, there is an $f' \in \mathrm{End}_R(M)$ such that the diagram

$$\begin{tikzcd} & M \arrow[dl, dashed, "f'"'] \arrow[d, "f"] & \\ M \arrow[r, "q"] & A \arrow[r] & 0 \end{tikzcd}$$

commutes, that is, $q \circ f' = f$. An R-module M is called quasi-injective if and only if for every R-module A, every R-monomorphism $j: A \to M$, and R-homomorphism $f: A \to M$, there is an $f' \in \mathrm{End}_R(M)$ such that the diagram

$$\begin{tikzcd} 0 \arrow[r] & A \arrow[r, "j"] \arrow[d, "f"'] & M \arrow[dl, dashed, "f'"] \\ & M & \end{tikzcd}$$

commutes.

The first section of this paper contains results which are needed to prove a decomposition theorem for quasi-projective modules over left perfect rings (Theorem 1.10). This decomposition is a characterization for quasi-projective modules over left perfect rings. A ring is left perfect if a projective cover (the dual concept of injective envelope) exists for every left R-module [4, p. 467]. It is known, for example, that left Artinian rings are left perfect [4, p. 467]. Some of the propositions are stated for semiperfect rings which are rings such that every finitely generated module has a projective cover [4, p. 471].

In the second section the decomposition for quasi-projective modules is used to obtain a decomposition for quasi-injective modules over a special class of rings. For these rings this decomposition characterizes quasi-injective modules. This decomposition theorem (Theorem 2.5) is specialized to the cases where the ring is quasi-Frobenius and where it is a finite dimensional algebra.

It will be assumed that all rings have an identity and that the modules are unital. Modules will be left R-modules, and homomorphisms will be R-homomorphisms unless otherwise stated. When S is the centralizer of ${}_RM$ in the sense of Jacobson [8], the notation will be abused and be written $S = \mathrm{End}_R(M)$. Actually, S operates on the right is anti-isomorphic to $\mathrm{End}_R(M)$. The radical will mean the Jacobson radical and be denoted by N. A direct sum of card(I) copies of M will be written M^I unless card$(I) = n < \infty$, and then M^n will be used in place of M^I. Also $\sum_{i=1}^k \oplus M_i^{g(i)}$ is a direct sum where $M_i^{g(i)}$ is $g(i)$ copies of M, and $g(i)$ can be any cardinal number. If $g(i) = 0$, then $M_i^{g(i)} = 0$.

I wish to thank Professor Azumaya who suggested that I investigate quasi-projective modules.

1. Quasi-projective modules. The goal of this section is to prove Theorem 1.10 which is a characterization of quasi-projective modules over left perfect rings. The first proposition to be presented was proved by Wu and Jans.

PROPOSITION 1.1. *Let R be a semi-perfect ring. Then M is a finitely generated, indecomposable, quasi-projective module if and only if $M = Re/Je$ where e is an indecomposable idempotent, and J is an ideal of R* [12, Thm. 3.1].

PROPOSITION 1.2. *Let R be a semi-perfect ring. If $Re/Je \neq 0$ where e is an indecomposable idempotent and J is an ideal, then $Je = J'e$ where J' is an ideal contained in the radical N.*

Proof. The module Ne is small in Re [4, p. 473]. Since Re is indecomposable and the projective cover of Re/Ne, Re/Ne is indecomposable. It is known that R/N is completely reducible if R is semi-perfect [4, Thm. 2.1]. Thus Re/Ne is simple, and Ne is maximal in Re. Now $Je \subseteqq Ne$ because Ne is both maximal and small in Re. Let $J' = J \cap N$.

PROPOSITION 1.3. *If M is quasi-projective and has a projective cover $P \xrightarrow{\pi} M \longrightarrow 0$ and if $P = \Sigma \oplus P_a$ ($a \in I$, and indexing set), then $M = \sum \oplus M_a$ and $P_a \xrightarrow{\pi_a} M_a \longrightarrow 0$ is the projective cover of M_a where $\pi_a = \pi \mid P_a$.*

Proof. The proof for the finite case [12, Prop. 2.4] will work here also.

PROPOSITION 1.4. *Let* $P_a \xrightarrow{\pi_a} M_a \longrightarrow 0$ *be the projective cover of* M_a *where* $a \in I$, *an indexing set. If* $f(Ker\ \pi_a) \subseteqq Ker\ \pi_b$ *for every* $a, b \in I$ and $f \in Hom_R(P_a, P_b)$, *then* $\Sigma \oplus M_a$ *is quasi-projective.*

Proof. It is sufficient to show that $\Sigma \oplus Ker\ \pi_a$ is an $\mathrm{End}_R(\Sigma \oplus P_a)$-module [12, Prop. 1.1]. Let q_c be the projection of $\Sigma \oplus P_a$ onto P_c and $f \in \mathrm{End}_R(\Sigma \oplus P_a)$. We will be done if we show $f(Ker\,\pi_b) \subseteqq \Sigma \oplus \mathrm{Ker}\,\pi_a$. Let $x \in \mathrm{Ker}\,\pi_b$. Since $q_a \circ (f|P_b) \in \mathrm{Hom}_R(P_b, P_a)$, $f(x) = (q_{a_1} \circ f)(x) + \cdots + (q_{a_n} \circ f)(x) \in \mathrm{Ker}\,\pi_a + \cdots + \mathrm{Ker}\,\pi_{a_n} \subseteqq \Sigma \oplus \mathrm{Ker}\,\pi_a$.

REMARK. If I is finite or R is left perfect, then the converse is true, that is, if $\Sigma \oplus M_a$ is quasi-projective, then $f(\mathrm{Ker}\,\pi_a) \subseteqq \mathrm{Ker}\,\pi_b$ for every $a, b \in I$ and $f \in \mathrm{Hom}_R(P_a, P_b)$.

COROLLARY 1.5. If M is quasi-projective and has a projective cover, then M^I is quasi-projective.

PROPOSITION 1.6. *If* M_1 *and* M_2 *are quasi-projective and have projective covers* P_1 *and* P_2 *which are isomorphic and* $M_1 \oplus M_2$ *is quasi-projective, then* $M_1 \cong M_2$.

Proof. The proof of the dual theorem for quasi-injective modules [7, Prop. 2.4] can be dualized.

Bass has shown [4, p. 473] that if R is a left perfect ring and P is a projective module, then $P = \Sigma \oplus Re_i$ where Re_i/Ne_i is simple and e_i is an idempotent in R. This result will be stated in a different form in the next proposition.

PROPOSITION 1.7. *Let* R *be left perfect. Then* P *is projective if and only if* $P = \sum_{i=1}^{k} \oplus (Re_i)^{g(i)}$ *where* Re_i *is the projective cover of a simple module,* e_i *is an indecomposable idempotent,* k *is the number of non-isomorphic simple modules, and* $Re_i \not\cong Re_j$ *if* $i \neq j$.

Proof. Ne_i is small in Re_i [4, p. 473]. Hence Re_i is the projective cover of Re_i/Ne_i and is indecomposable by Proposition 1.3. Since R/N is left Artinian [4, p. 467], and the simple R-modules and the simple R/N-modules are the same, there are only a finite number of nonisomorphic simple modules. Also, simple modules are isomorphic if and only if their projective covers are isomorphic.

PROPOSITION 1.8. *Let* R *be semi-perfect and* M *be a finitely generated, quasi-projective module. Then* M *is indecomposable* (*non-zero*) *if and only if* $End_R(M)$ *is a local ring.*

Proof. (i) If M is not indecomposable, then $\mathrm{End}_R(M)$ has a nonzero idempotent e which is different from the identity. Since neither e nor $1-e$ is a unit, $\mathrm{End}_R(M)$ is not a local ring.

(ii) If M is indecomposable, then $M = Re/Je$ where J' is an ideal of R and e is an indecomposable idempotent. Thus $M = R^*e^*$ where $R^* = R/J$. R^* is semi-perfect [4, Lemma 2.2]. Since M is indecomposable as an R^*-module, e^* is an indecomposable idempotent. In addition $\mathrm{End}_R(M) = \mathrm{End}_{R^*}(R^*e^*) = e^*R^*e^*$. Finally, $e^*R^*e^*$ is a local ring because R^* is semi-perfect and e^* is indecomposable [10, p. 76].

LEMMA 1.9. *Let R be semi-perfect and $1 = e_1 + \cdots + e_n$ where $e_1, \cdots e_n$ are orthogonal, indecomposable idempotents. If*

$$Re_1/J_1e_1 \oplus Re_2/J_2e_2 \oplus \cdots \oplus Re_m/J_me_m$$

is quasi-projective where J_i, $i = 1, \cdots, m$, is an ideal, then there is an ideal J such that $Je_i = J_ie_i$ for $i = 1, \cdots, m$.

Proof. The projective cover of $\sum_{1=i}^{m} \oplus Re_i/J_ie_i$ is

$$\sum_{i=1}^{m} \oplus Re_i \xrightarrow{\pi} \sum_{i=1}^{m} \oplus Re_i/J_ie_i \longrightarrow 0$$

where $\mathrm{Ker}\,\pi = \sum_{i=1}^{m} \oplus J_ie_i$. Since $\mathrm{End}_R(\sum_{i=1}^{m} \oplus Re_i) = \sum_{j=1}^{m} \oplus_{i=1}^{m} \oplus e_iRe_j$, it follows that $J_ie_i \cdot e_iRe_j \subseteqq J_je_j$ for $i, j = 1, \cdots, m$ [12. Prop. 2.2]. Let

$$J = \sum_{i=1}^{m} \oplus J_ie_i \oplus \left(\sum_{i=1}^{m} \oplus J_ie_i\right) R \left(1 - \sum_{i=1}^{m} e_i\right).$$

Then J is an ideal because $R = \sum_{j=1}^{n} \oplus \sum_{i=1}^{n} \oplus e_iRe_j$. Also, $Je_i = J_ie_i$ for $i = 1 \cdots, m$.

REMARKS. 1. The proof for Lemma 1.9 remains valid if any subcollection of $e_1, \cdots e_n$ is used rather than the first m of them.

2. The result that for a semi-perfect ring $1 = e_1 + \cdots + e_n$ where $e_1, \cdots, e_n$ are orthogonal indecomposable idempotents can be found in [10].

THEOREM 1.10. *Let R be left perfect. Then M is a quasi-projective module if and only if*

$$M = \sum_{i=1}^{k} \oplus (Re_i/Je_i)^{g(i)}$$

where J is an ideal, $e_1, \cdots, e_k$ are indecomposable idempotents, the number of nonisomorphic simple R-modules is k, and $Re_1, \cdots, Re_k$ are the corresponding nonisomorphic projective covers. In addition *the decomposition is unique up to automorphism.*

Proof. (i) Let M be quasi-projective. If $M = 0$, then we can choose $J = R$ and be done. If $M \neq 0$, let $P \to M \to 0$ be the projective cover of M. By Proposition 1.7 $P = \sum_{i=1}^{k} \oplus (Re_i)^{g(i)}$ where $Re_1, \cdots, Re_k$ are the nonisomorphic indecomposable projective covers of all the simple modules. By Proposition 1.3 $M = \sum_{i=1}^{k} \oplus \sum_{a \in I_i} \oplus M_{ai}$ where $\operatorname{card}(I_i) = g(i)$. Proposition 1.6 shows that $M_{ai} \cong M_{bi}$ for every $a, b \in I_i$. From Proposition 1.1 $M_{ai} = Re_i/J_ie_i$ with J_i an ideal and e_i an indecompotent. As a result of Lemma 1.9 and the remark following it, there is an ideal J such that $Je_i = J_ie_i$ for $i = 1, \cdots, k$.

(ii) Conversely, if $M = \sum_{i=1}^{k} \oplus (Re_i/Je_i)^{g(i)}$ with the same notation as in the statement of the theorem and $J \neq R$, then Propositions 1.2 and 1.4 show that M is quasi-projective. If $J = R$, then $M = 0$ and is, of course, quasi-projective.

(iii) Uniqueness. Using Proposition 1.8 and a generalized Krull-Remark-Schmidt theorem which was proved by Azumaya [1, Thm. 1], we have the following result: if $\sum_{a \in A} \oplus M_a$ and $\sum_{b \in B} \oplus M'_b$ are two decompositions of quasi-projective module into indecomposable, modules, then there is a 1 to 1, onto mapping $f\colon A \to B$ such that $M_a \cong M'_{f(a)}$.

REMARKS. 1. Theorem 1.10 is true for semi-perfect rings if M is finitely generated.

2. If M is nonzero in Theorem 1.10, then J can be chosen is the radical of the ring.

2. Quasi-injective modules.

In the first section a decomposition theorem for quasi-projective modules was obtained. The motivation for attempting to prove this proposition came from a paper by Harada on quasi-injective modules [7]. Now Theorem 1.10 will be used to obtain a characterization for quasi-injective modules over left Artinian rings which have a finitely generated, lower distinguished (contains an isomorphic copy of every simple module), and injective module. This class of rings includes quasi-Frobenius rings and finitely generated algebras over commutative Artinian rings [2].

PROPOSITION 2.1. Let R *be left Artinian. Then* R *has a finitely generated, lower distinguished, injective module if and only if the injective envelope of every simple module is finitely generated.*

Proof. (Given by G. Azumaya). Assume Q is finitely generated, lower distinguished, and injective. Let $Q_1, \cdots, Q_k$ be the nonisomorphic injective envelopes of all the simple modules. Then

$Q = \sum_{i=1}^{k} \oplus Q_i^{h(i)}$ where $0 \leqq h(i) < \infty$ [2, Thm. 11, p. 268]. Since Q is lower distinguished, $h(i) \neq 0$ for each $i = 1, \cdots, k$. It follows that each Q_i is finitely generated. The converse is clear.

PROPOSITION 2.2. *If R is left Artinian and has a finitely generated, lower distinguished, injective module, then every indecomposable quasi-injective module is finitely generated.*

Proof. Let M be indecomposable and quasi-injective, and let Q be its injective envelope. Q is indecomposable [7, Proposition 2.3], so it is the injective envelope of a simple module [2, Thm. 1, p. 268]. Hence, Q is finitely generated by Proposition 2.1. Since R is left Noetherian, M is finitely generated.

REMARK. If R is perfect, then every indecomposable quasi-projective module is finitely generated by Proposition 1.7.

The following proposition was proved by Azumaya for the class of rings in the last two propositions and will be stated without giving his proof.

PROPOSITION 2.3. (Duality Theorem). *Let R be a left Artinian ring which has a finitely generated, injective, and lower distinguished module Q, and let $S = End_R(Q)$. Then for any finitely generated left R-module X, $X^* = \mathrm{Hom}_R(X, Q)$ is a finitely generated right S-module and $(X^*)^* = \mathrm{Hom}_S(X^*, Q) = {}_RX$. The same is true for finitely generated S-modules* [2, Thm. 8, p. 262].

PROPOSITION 2.4. *If R is left Noetherian and M is quasi-injective, then M^I is quasi-injective.*

Proof. Let Q be the injective, envelope of M. Since R is left Noetherian, Q^I is the injective envelope of M^I. With this result and a theorem of Johnson and Wong [9, Thm. 1.1], a procedure which is similar to the one found in the proof of Proposition 1.4 can be used to see that M^I is quasi-injective.

THEOREM 2.5. *Let R be left Artinian and have a finitely generated, lower distinguished, and injective module Q. Then M is quasi-injective if and only if*

$$M = \sum_{i=1}^{k} \oplus (\mathrm{Hom}(e_iS/e_iJ, Q))^{g(i)}$$

where $S = \mathrm{End}_R(Q)$, e_i is an indecomposable idempotent in S for $i = 1, \cdots, k$, J is an ideal of S, the number of nonisomorphic simple

R-modules is k, and for $i \neq j$ $e_iS \ncong e_jS$. This decomposition is unique up to automorphism.

Proof. If $M = 0$, we can choose $J = S$. Thus we will assume that M is a nonzero quasi-injective module. It is known that if R is left Artinian, then it is left Noetherian and has only a finite number of simple R-modules. Harada has shown that for left Noetherian rings $M = \sum \oplus M_a$ where the M_a's are indecomposable quasi-injective modules and that this decomposition is unique up to automorphism [7, Prop. 2.5]. If Q_a is the injective envelope of M_a, then it is the injective envelope of a simple module (see proof of Prop. 2.2). By the dual theorem of Proposition 1.6 and the result that nonisomorphic simple modules have nonisomorphic injective envelopes, $M = \sum_{i=1}^{k} \oplus M_i^{g(i)}$ and $M_i \ncong M_j$ for $i \neq j$.

As a result of Proposition 2.2, M_i is finitely generated. By the Duality Theorem $\mathrm{Hom}_R(M_i, Q)$ is a finitely generated, indecomposable, quasi-projective, right S-module. Also, S is right Artinian [2, Thm. 6, p. 259]. Hence, $\mathrm{Hom}_R(M_i, Q) = e_iS/e_iJ_i$ where e_i is an indecomposable idempotent in S, and J_i is an ideal of S. Since $\sum_{g(i)\neq 0} \oplus M_i$ is a direct summand of M it is quasi-injetive. It follows that $\mathrm{Hom}\,(\sum_{g(i)\neq 0} \oplus M_i, Q) = \sum_{g(i)\neq 0} \oplus \mathrm{Hom}(M_i, Q) = \sum_{g(i)\neq 0} \oplus e_iS/e_iJ_i$ and is quasi-projective. For $i \neq j$ $M_i \ncong M_j$, so $e_iS \ncong e_jS$. By Lemma 1.9 and a remark following it we can choose $J_i = J$ for $g(i) \neq 0$. In addition $M_i = \mathrm{Hom}\,(\mathrm{Hom}\,(M_i, Q), Q) = \mathrm{Hom}_S\,(e_iS/e_iJ, Q)$.

(ii) Suppose $M = \sum_{i=1}^{k} \oplus (\mathrm{Hom}_S\,(e_iS/e_iJ, Q)^{g(i)}$ with the same notation as in the statement of the theorem. Let $M' = \sum_{g(i)\neq 0} \oplus \mathrm{Hom}_S(e_iS/e_iJ, Q)$. Then $\mathrm{Hom}_R\,(M', Q) = \sum_{g(i)\neq 0} \oplus e_iS/e_iJ$ which is quasi-projective by Theorem 1.10. Thus M' is quasi-injective. Let $m = \max\{g(i)\}_{i=1,\ldots,k}$ and $M'' = (M')^m$. Proposition 2.4 gives us that M'' is quasi-injective. Therefore the direct summand M is quasi-injective.

COROLLARY 2.6. *Let R be quasi-Frobenius. Then M is quasi-injective if and only if*

$$M = \sum_{i=1}^{k} \oplus (\mathrm{Hom}_R\,(e_iR/e_iJ, R))^{g(i)}.$$

Proof. R being quasi-Frobenius implies R is left Artinian, self injective, lower distinguished, and finitely generated [2, Thm. 6, p. 259]. Also, $R = \mathrm{End}_R(R)$.

COROLLARY 2.7. *Let R be a finitely generated algebra over a commutative Artinian ring K. Then M is quasi-injective if and only if*

$$M = \sum_{i=1}^{k} \oplus (\mathrm{Hom}_K(e_iR/e_iJ, F))^{g(i)}$$

where F is the K-injective envelope of $K/\mathrm{rad}\, K$.

Proof. R has a finitely generated, lower distinguished, injective module Q such that $R = \mathrm{End}_R(Q)$ [2, Prop. 19, p. 273]. The functors $\mathrm{Hom}_K(\cdot, F)$ and $\mathrm{Hom}_R(\cdot, Q)$ are naturally equivalent for finitely generated R-modules [2, Thm. 20, 275].

COROLLARY 2.8. *Let R be a finite dimensional algebra over a field K. Then M is quasi-injective if and only if*

$$M = \sum_{i=1}^{k} \oplus (\mathrm{Hom}_K (e_iR/e_iJ, K))^{g(i)}.$$

Proof. $K = F$ in Corollary 2.7.

REFERENCES

1. G. Azumaya, *Corrections and supplementaries to my paper concerning Krull-Remak-Schmidt's theorem*, Nagoya Math. J. **1** (1950), 117-124.
2. ———, *A duality theory for injective modules*, Amer. J. Math. **81** (1959), 249-278.
3. ———, *Completely faithful modules and self-injective rings*, Nagoya Math. J. **27** (1966), 697-708.
4. H. Bass, *Finitistic dimension and a homological generalization of semi-primary rings*, Trans. Amer. Math. Soc. **95** (1960), 466-488.
5. ———, *The Morita theorems*, mimeographed notes.
6. S. Eilenberg, *Homological dimension and syzygies*, Ann. of Math. **64** (1956), 328-336.
7. M. Harada, *Note on quasi-injective modules*, Osaka J. Math. **2** (1965). 351-356.
8. N. Jacobson, *Structure of rings*, Amer. Math. Soc., Providence, 1964.
9. R. E. Johnson and E. T. Wong, *Quasi-injective modules and irreducible rings*, J. London Math Soc., **36** (1961), 260-268.
10. J. Lambek, *Lectures on Rings and Modules*, Blaisdell Publishing Co., Toronto, 1966.
11. E. Matlis, *Injective modules over Noetherian rings*, Pacific J. Math., **8** (1958), 511-528.
12. L. E. T. Wu and J. P. Jans, *On quasi-projectives*, Ill. J. Math. **11** (1967), 439-447.

Received August 19, 1968. This paper is part of the author's thesis at Indiana University and was partially supported by NSF Grant No. GP5799 and an NSF Traineeship.

INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47401
AND
MIAMI UNIVERSITY
OXFORD, OHIO 45056

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

COMPLETIONS OF DEDEKIND PRIME RINGS AS SECOND ENDOMORPHISM RINGS

JAMES KUZMANOVICH

The purpose of this paper is to show that if M is a maximal two-sided ideal of a Dedekind prime ring R and P is any maximal right ideal containing M, then the M-adic completion $\bar{R}$ of R can be realized as the second endomorphism ring of $E=E(R/P)$, the R-injective hull of R/P; that is, as end $({}_K E)$ where $K=\text{end}\,(E_R)$. The ring K turns out to be a complete, local, principal ideal domain.

This paper was motivated by a result of Matlis [6] which says that if P is a prime ideal of a commutative Noetherian ring R, then the P-adic completion of the localization of R at P can be realized as the ring of endomorphisms of $E=E(R/P)$, the R-injective hull of R/P.

Since $\bar{R}$ is a full matrix ring over a complete local domain L [4], we are able to approach the problem by considering first the case that R is a complete local domain, then by means of the Morita theorems we pass to the case $\bar{R}=R$, and finally pass to the general case.

1. Introduction. A prime ring R is called a Dedekind prime ring if it is Noetherian, hereditary, and a maximal order in its classical quotient ring Q (see [3]). A ring R is called local if the nonunits of R form an ideal.

If R is a Dedekind prime ring with a nonzero prime ideal M, then M is a maximal two-sided ideal and $\cap M^n = 0$ (see Robson [7]). Let $\bar{R} = \bar{R}_M$ be the completion of R at M in the sense of Goldie [3]. In this situation combining results of Goldie ([3], Theorem 4.5) and Gwynne and Robson ([4], Theorem 2.3) yields the following theorem.

THEOREM 1.1. *Let R be a Dedekind prime ring with a maximal ideal M. Then* (i) *$\bar{R}$ has a unique maximal two-sided ideal $\bar{M}$, $\bar{M}$ is the Jacobson radical of $\bar{R}$, and $R \cap \bar{M}^p = M^p$.*

(ii) *$\bar{R}$ is a full $k \times k$ matrix ring over a domain L which has a unique maximal ideal N, and $L/N = F$ where F is a division ring. Also $R/M^p \simeq \bar{R}/\bar{M}^p$* (each coset of $\bar{M}^p$ has a representative in R).

(iii) *$\bar{R}$ is a prime principal ideal ring and L is a complete, local, principal ideal domain. The only one-sided ideals of L are the powers of N.*

For the rest of this section let R, M, $\bar{R}$, $\bar{M}$, L, and N be as in Theorem 1.1. Let x be the generator of N; then $N = xL = Lx$ and $N^k = x^k L = Lx^k$.

2. **The Ring L.** This section will be concerned with the construction of the L-injective hull of $(L/N)_L$ and with showing that Theorem 4.4 holds for L.

LEMMA 2.1. *L/N^k can be embedded in L/N^{k+1} as a right L-module via the map h_k: $L/N^k \to L/N^{k+1}$ defined by $h_k([u + N^k]) = [xu + N^{k+1}]$.*

Proof. h_k is clearly additive and right L-linear. Suppose $h_k([u + N^k]) = [0 + N^{k+1}]$. From the definition of h_k it follows that $xu \in N^{k+1}$ so that $xu = x^{k+1}u$, for some u' in L and $u = x^k u' \in N^k$. Hence $[u + N^k] = [0 + N^k]$ and n_k is a monomorphism. A similar argument shows that h_k is well-defined.

The maps $\{h_k\}$ and the right L-modules $\{(L/N^k)_L\}$ give rise to a directed system. Let E_L be the direct limit of this system. Then E_L can be considered as an ascending union of a family of submodules, $\{(S_j)_L\}$, which is totally ordered by inclusion and where each $(S_j)_L$ is isomorphic to $(L/N^i)_L$.

LEMMA 2.2. *Consider $(L/N^{p+t+1})_L$. Take $a \in N^p/N^{p+t+1}$ and $d \in N^p \backslash N^{p+1}$. The equation $yd = a$ has a solution in $(L/N^{p+t+1})_L$.*

Proof. $a \in N^p/N^{p+t+1}$ so that $a = [x^p v + N^{p+t+1}]$. $d \in N^p \backslash N^{p+1}$ so that $d = x^p u$ where u is a unit in L. In L, $x^p vu^{-1} = wx^p$ since

$$\begin{aligned} N^p = x^p L = Lx^p. \quad \text{Let } y = [w + N^{p+t+1}].\ yd \\ = [w + N^{p+t+1}]d = [wd + N^{p+t+1}] = [wx^p u + N^{p+t+1}] \\ = [x^p vu^{-1}u + N^{p+t+1}] = [x^p v N^{p+t+1}] = a. \end{aligned}$$

PROPOSITION 2.3. *E_L is isomorphic to the L-injective hull of the simple right L-module $(L/N)_L$.*

Proof. E_L contains a copy of $(L/N)_L$, namely S_1. Thus it is enough to show that E is an essential injective extension of S_1. S_1 is essential in E for if $a \in E$, $a \in S_k$ for some integer k. Let t be the first such integer: then $a \in S_t \backslash S_{t-1}$, a is a generator for S_t, and $aL = S_t$. Thus $aL \cap S_1 = S_1$ and S_1 is essential. Since L is a principal ideal domain, it is a hereditary two-sided order in its quotient division ring. In order to prove E_L is injective it is sufficient by a result of Levy ([5], Theorem 3.4) to show that it is L-divisible. Take $a \in E$ and $0 \neq d \in L$. $a \in S_t$ for some t and $d \in N^p \backslash N^{p+1}$ for some p. $yd = a$ has a solution in S_{p+t+1}, and hence in E, by Lemma 2.2. E is thus an essential injective extension of S_1 and hence is its injective hull.

Let $K = \text{end}_L(E)$ and let K act on E by left multiplication; E then becomes a left K-module. Let $H = \text{end}_K(E)$; in similar manner E then becomes a right H-module. $Ed = E$ (since E is L-divisible) for all nonzero d in L; thus E is a faithful right L-module. Hence L may be considered as a unital subring of H.

LEMMA 2.4. *The S_k's are the only proper L-submodules of E_L.*

Proof. Suppose M_L is a submodule of E with generating set $\{m_i\}$. Since $E = \cup S_k$, each m_i is in some S_k. Let k_i be the first k for which $m_i \in S_k$. Then $m_i \in S_{k_i} \backslash S_{k_i-1}$ and $m_i L = S_{k_i}$. $M = \Sigma m_i L = \Sigma S_{k_i}$ so that if $\{k_i\}$ is bounded, $M = S_{k_t}$ where $k_t = \max\{k_i\}$, and if $\{k_i\}$ is not bounded, then $M = E_L$.

LEMMA 2.5. *If $a \in S_n$ and if $b \in S_{n-1}$, then there is a $q \in K$ such that $q(b) = a$.*

Proof. Assume that t is the first integer for which $b \in S_{n+t}$. Then $\text{ann}_L(b) = N^{n+t}$ which is contained in N^n which in turn is contained in $\text{ann}_L(a)$. Thus the map $\bar{q}\colon bL \to aL$ defined by $\bar{q}(bd) = ad$ is well defined. E_L is L-injective so that $\bar{q}$ can be extended to an endomorphism q of E. $q \in K$.

PROPOSITION 2.6. *Each S_n is a cyclic left K-submodule of ${}_KE$, the composition length of ${}_K(S_n)$ is n, and the S_n's are the only proper K-submodules of E.*

Proof. If $q \in K, q(S_n)$ is an L-submodule of E of composition length less than or equal to n and hence must be contained in S_n by Lemma 2.4; hence each S_n is a left K-submodule. Each ${}_K(S_n)$ is cyclic via Lemma 2.5; in fact, any L generator of S_n will be a K generator of S_n. This implies that ${}_K(S_1)$ is simple and inductively that the composition length of ${}_K(S_n)$ is n. The proof of Lemma 2.4 shows that these are the only K-submodules of E.

LEMMA 2.7. *Let H_i be the annihilator of S_i in H. Then H_i is a two-sided ideal of H, H_{i+1} is properly contained in H_i, and $\cap H_i = 0$.*

Proof. H_i is clearly a right ideal of H. If $h \in H$, then $(S_i)h$ is a K-submodule of E of composition length less than or equal to i. By Proposition 2.6 it must be that $(S_i)h \subset S_i$ so that each S_i is H-invariant. As a result H_i is a left ideal and hence an ideal. The inclusions are proper, for $H_i \cap L = N^i$ and $N_i \neq N^{i+1}$. Since $E = \cup S_i$, anything in $\cap H_i$ would annihilate all of E and hence be zero.

PROPOSITION 2.8. $H = L$. *That is, L is the second endomorphism ring of E_L.*

Proof. Take $f \in H$. By Proposition 2.6 there is a nonzero $y \in S_1$ such that $S_1 = Ky = yL$. Hence there is a $p_1 \in L$ such that $yf = yp_1$. Also, if $z \in S_1$, $z = ky$ for some $k \in K$ and

$$z(f - p_1) = (ky)(f - p_1) = k0 = 0 . \quad \text{Hence } f - p_1 \in \operatorname{ann}_H(S_1) = H_1 .$$

Inductively suppose that there is a $p_i \in L$ such that $f - p_i \in H_i$. Now take $y \in S_{i+1} \backslash S_i$. $y(f - p_i) \in S_{i+1}$ so that there is a $d \in L$ such that $y(f - p_i) = yd$. If $z \in S_{i+1}$, $z = ky$ for some $k \in K$. Then $z(f - p_i) = (ky)(f - p_i) = k(y(f - p_i)) = k(yd) = (ky)d = zd$ and hence $f - p_i - d$ is in H_{i+1}. Let $p_{i+1} = p_i + d$; then $f - p_{i+1} \in H_{i+1}$.

The sequence $\{p_i\}$ is Cauchy in L, for $p_n - p_m = (p_n - f) + (f - p_m)$ an element of $H_n + H_m$; but $H_n + H_m = H_n$ if $n \leqq$ m. Thus $p_n = p_m$ is in $H_n \cap L = N^n$. L is complete; therefore $\{p_i\}$ converges to some element p of L. It only remains to be shown that $p = f$. Take $z \in E$; $z \in S_n$ for some n. $\{p_i\}$ converges to p so that there is a positive integer M such that $p_m - p \in N^n$ for all m greater than M. Take m greater than $M + n$. $zf = zp_m = zp$. z was arbitrary; therefore $f = p$.

3. The Ring K.

In this section it will be shown that K is a complete, local, principal ideal domain.

LEMMA 3.1. *Let L, E, and K be as in §2. Let J denote the Jacobson radical of K and let $A_n = \operatorname{ann}_K(S_n)$. Then*

(i) *K is a local domain.*

(ii) *$J = A_1$, $J^n \subset A_n \cap A_n = 0$, and $\cap J^n = 0$.*

(iii) *K is complete in the topology induced by the A_n's.*

Proof. (i) K is local since it is the endomorphism ring of an indecomposable injective module. To prove that K is a domain it is sufficient to show that every nonzero endomorphism of E_L is an epimorphism. Let $0 \neq k \in K$. If $k(E) \neq E$, $k(E) = S_n$ for some n by Lemma 2.4. $\operatorname{Ann}_L(S_n) = N^n$; take $0 \neq b \in N^n$. Since E is L-divisible, $Eb = E$. As a result $S_n = k(E) = k(Eb) = k(E)b = S_n b = 0$ contradicting the fact that $k \neq 0$.

(ii) The radical of K, J, is the set of all endomorphisms of E_L whose kernel is essential (see [2], page 44). Since $(S_1)_L$ is the unique minimal submodule of E, $\ker(k)$ is essential if and only if $k(S_1) = 0$;

therefore $J = A_1$ and $JS_1 = 0$. Inductively suppose that $J^{n-1}S_{n-1} = 0$. $JS_n \subset S_{n-1}$ since it is contained in the radical of $K(S_n)$, S_{n-1}. Hence $J^n s_n = J^{n-1}(Js_n)$ which is contained in $J^{n-1}S_{n-1}$ which is zero, hence $J^n \subset A_n \cap A_n = 0$ since anything in $\cap A_n$ would annihilate all of the S_n's and hence all of E. $\cap J^n = 0$ since $J^n \subset A_n$.

(iii) Let $\{f_i\}$ be a Cauchy sequence in K with respect to the topology induced by the decreasing family $\{A_n\}$. Let $x \in E$. $x \in S_p$ for some p. Since $\{f_i\}$ is Cauchy, there is an integer M such that $f_n - f_m \in A_p$ for n, m greater than M. Define $f(x) = f_{M+1}(x)$. It is clear that $f \in K$ and that $f_i \to f$ by the nature of the construction.

Pick $j \in J\backslash A_2$. There is such a j, for if $y_2 \in S_2\backslash S_1$ and if $0 \neq y_1 \in S_1$, then there is a $j \in K$ such that $j(y_2) = y_1$ by Lemma 2.5. $j \in J\backslash A_2$. In fact if $s \in S_{n+1}\backslash S_n$, then $j^n s$ is a nonzero element of S_1. The proof is by induction. If $s \in S_2\backslash S_1$, then $s = y_2 u$ for u a unit in L. Hence $js = jy_2 u = y_1 u \neq 0$. Inductively suppose that $j^{n-1}s$ is nonzero for all s in $S_n\backslash S_{n-1}$ and take $s \in S_{n+1}\backslash S_n$. $js \in S_n$ by an argument in the previous proof. The claim is that $js \notin S_{n-1}$. If it were, then $j^{n-1}s = 0$ which contradicts the induction hypothesis since $sd \in S_n\backslash S_{n-1}$ for some d in L. Hence $js \notin S_{n-1}$ so again by the induction hypothesis $j^n s = j^{n-1}(js) \neq 0$.

LEMMA 3.2. *Let K, J, j, E, and L be as above.*
(i) $J = jK$.
(ii) $J = Kj$.
(iii) $J^n = j^n K = Kj^n$.

Proof. (i) Let $x \in$ J. Let $y_2 \in S_2\backslash S_1$. $x(y_2) = y \in S_1$ since $x \in J$. Let $j(y_2) = y_1$; y_1 is a nonzero element of S_1 since $j \in J\backslash A_2$. Then there is an element d in L such that $y = y_1 d = j(y_2)d = j(y_2 d)$. By Lemma 2.5 there exists $k_1 \in K$ such that $k_1(y_2) = y_2 d$. If $s \in S_2$, then $s = y_2 c$ for some c in L. $x(s) = x(y_2 c) = X(y_2)c = uc = (jk_1(y_?))c = jk_1(y_2 c) = jk_1(s)$. This says that $x - jk_1 \in A_2$.

Inductively suppose that there exist $k_1, \cdots, k_{n-1}$ such that

$$z = x - (jk_1 + j^2k_2 + \cdots + j^{n-1}k_{n-1}) \in \mathrm{A}_n. \quad \text{If}$$
$$y_{n+1} \in S_{n+1}\backslash S_n, \text{ then } j^n(y_{n+1}) = y_1$$

a nonzero element of S_1 by the above choice of j. Also $z(y_{n+1}) \in S_1$ since $z \in A_n$. Hence by the argument above there is a $k_n \in K$ such that $z - j^n k_n \in A_{n+1}$. The sequence $\{jk_1 + \cdots + j^n k_n\}$ converges to x in the A_n topology by the nature of the construction. Also, since $J^n \subset A_n$ the sequence $\{k_1 + \cdots + j^{n-1}k_n\}$ is Cauchy and hence by the completeness of K converges to some element k of K. Also by the construction $jk = x$. Since x was arbitrary in J, $J = jk$.

(ii) is proven by an argument similar to that of (i).

(iii) $J = jK = Kj$ by (i) and (ii). Inductively suppose that $J^n = j^nK = Kj^n$. Then $J^{n+1} = J^nJ = (j^nK)(jK) = j^n(Kj)K = j^n(jK)K = j^{n+1}K$. Similarly $J^{n+1} = Kj^{n+1}$.

PROPOSITION 3.3. *K as above.*

(i) $J^n = A_n$ for all n.

(ii) J^n are the only one-sided ideals of K.

(iii) K is a complete principal ideal domain.

Proof. (i) $J = A_1$ by Lemma 3.1. Inductively suppose that $A_n = J^n$. $J^{n+1} \subset A_{n+1} \subset A_n = J^n$. $J^n/J^{n+1} = j^nK/j^{n+1}K \simeq K/jK = K/J$ which is simple. Therefore either $A_{n+1} = J^{n+1}$ or $A_{n+1} = J^n$. But by the induction hypothesis $j^n \notin A_{n+1}$ so that $A_{n+1} = J^{n+1}$.

(ii) It is sufficient to show that given $x \in K$, $xK = K$ or that $xK = J^p$ for some p. Take $x \in K$ and suppose that $xK \neq K$, then x is not a unit and hence $x \in J^{p+1}$ for some p. By Lemma 3.1 $x = j^pk$, and k must be a unit; for otherwise $k = jk_1$ for some k_1 in K and $x = j^pjk_1 \in J^{p+1}$. As a result $xK = j^pkK = j^pK = J^p$. Similarly $Kx = J^p$.

(iii) K is a principal ideal domain by Lemma 3.2 and (ii). K is complete by (i) and Lemma 3.1.

4. **The Ring R.** Let R, M, $\bar{R}$, and L be as in Theorem 1.1. Then $\bar{R}$ is the full $k \times k$ matrix ring over L. Let e_{ij}, $i, j = 1, 2, \cdots, n$ be a complete set of matrix units for $\bar{R}$. Let M_L be a right L-module and let $M^* = M_1 \oplus \cdots \oplus M_n$, a direct sum of n copies of M. Let f_1 be the identity map on M_1, and let f_i, $i = 2, \cdots, n$ be an isomorphism from M_1 to M_i. Then M^* can be made into an $\bar{R}$-module by defining $f_i(m)e_{ij} = f_j(m)$ and $f_i(m)e_{kj} = 0$ if $i \neq k$. "*" is a category isomorphism from the category of right L-modules to the category of right $\bar{R}$-modules. There is also a category isomorphism e_{11} from the category of right $\bar{R}$-modules to the category of right L-modules defined by $(A_{\bar{R}})e_{11} = Ae_{11}$. M and M^*e_{11} are isomorphic for any right L-module M (see [1], or [5] page 137).

PROPOSITION 4.1. *$\bar{R}$ is the second endomorphism ring of the $\bar{R}$-injective hull of the simple right $\bar{R}$-module.*

Proof. Let E be the L-injective hull of the simple right L-module as in §2. Then E^* is the $\bar{R}$-injective hull of *a* simple right $\bar{R}$-module since * is a category isomorphism. $\bar{R}/\bar{M}$ is simple Artinian and $\bar{M}$ is the Jacobson radical of $\bar{R}$ so there is only one isomorphism class of simple right-$\bar{R}$-modules. Let $K = \mathrm{end}_{\bar{R}}(E^*)$ and take $q \in K$.

$q(E^*e_{ii}) = q(E^*{}_{ii}e_{ii}) = q(E^*e_{ii})e_{ii}$; thus each E^*e_{ii} is K-invariant and ${}_KE^* = {}_KF^*e_{11} \oplus {}_KEe_{22} \oplus \cdots \oplus {}_KE^*e_{kk}$. Each e_{ij} is a K-isomorphism so that E^* is decomposed as a direct sum of k mutually isomorphic K-modules. Thus each K-endomorphism of E^* can be given by multiplication by a matrix of homomorphisms. The remainder of the proof shows that the entries in this matrix are of the desired forms. Each $q \in K$ restricted to $E^*{}_{ii}$ is an L-endomorphism of $E^*{}_{ii}$. Each L-endomorphism of E^*e_{ii} can be extended in one and only one way to an $\bar{R}$-endomorphism of E^*; namely, if $\bar{q}$ is an L-endomorphism of E^*e_{ii}, then its unique extension q is defined by $q(z) = \sum_{j=1}^{k}\bar{q}(ze_{ji})e_{ij}$ for $z \in E^*$. Hence $K \simeq \mathrm{end}_L(E^*e_{ii})$ via the restriction map. By proposition 2.8 each element of $\mathrm{end}_K(E^*e_{ii})$ can be given by right multiplication by an element of $e_{ii}\bar{R}e_{ii}$. If $h: E^*e_{ii} \to E^*e_{jj}$ is a K-homomorphism, then $h\bar{e}_{ji}$ is a K-endomorphism of E^*e_{ji} where $\bar{e}_{ji}$ denotes right multiplication by e_{ji}. Hence $h\bar{e}_{ji} = \bar{e}_{ii}r\bar{e}_{ii}$ for some $r \in \bar{R}$. If $z \in E^*{}_{ii}$, then $(z)h = zhe_{jj} = zhe_{ji}e_{ij} = ze_{ii}re_{ii}e_{ij} = ze_{ii}re_{ij}$ so that h is given by right multiplication by an element of $e_{ii}\bar{R}e_{ij}$. As a result every K-endomorphism of E^* is given by right multiplication by an element of $\bar{R}$.

R can be considered as a subring of $\bar{R}$; as a result every $\bar{R}$-module is automatically an R-module. Also, if $\bar{M}$ is the maximal two-sided ideal of $\bar{R}$, then $\bar{M}^p \cap R = M^p$ and every coset of $\bar{R}/\bar{M}^p$ has a representative in R (Theorem 1.1).

LEMMA 4.2. *E^* as in the proof of Proposition* 4.1, *then $(E^*)_R$ is the ascending union of $\bar{R}$-modules $0 \subset B_1 \subset B_2 \subset \cdots$ where the composition length of B_n is n. These are the only $\bar{R}$-submodules of E^*. Furthermore, the B_i's are the only R-submodules of E^* and every R-endomorphism of E^* is an $\bar{R}$-endomorphism. That is, the structure of E^* as an R-module is identical to its structure as an $\bar{R}$-module.*

Proof. The first part follows since it was true of E and $*$ is a category isomorphism. Let $B_i = S_i{}^*$. A category isomorphism preserves the submodule lattice. Note that the composition length of $(B_n)_{\bar{R}}$ is n; since $\bar{M}$ is the radical of $\bar{R}$, $B_n\bar{M}^n = 0$. In order to prove that the B_n's are the only R-submodules of E^* it is sufficient to show that $aR = a\bar{R}$ for all $a \in E^*$. Take $a \in E^*$. Clearly $aR \subset a\bar{R}$. Take $\bar{r} \in \bar{R}$. $a \in B_n$ for some n so that $a\bar{M}^n = 0$. By theorem 1.1 there is an m in $\bar{M}^n$ so that $\bar{r} + m = r \in R$, then $a\bar{r} = a\bar{r} + 0 = a\bar{r} + am = a(\bar{r} + m)ar$. Thus $a\bar{R} \subset aR$ and $a\bar{R} = aR$.

Let q be an R-endomorphism of E^* and take $a \in E^*$ and $\bar{r} \in \bar{R}$. It must be shown that $q(a\bar{r}) = q(a)\bar{r}$. Since $a \in E^*$, $a \in B_n$ for some n. The B_n's are the only R-submodules of E^* and the composition length of B_n is n, so that $q(B_n) \subset B_n$ and $q(a) \in B_n$. As above there is an $m \in \bar{M}^n$ such that $\bar{r} + m = r \in R$. $B_n \bar{M}^n = 0$. Then

$$\begin{aligned} q(a\bar{r}) &= q(a\bar{r} + 0) = q(a\bar{r} + am) = q(a(\bar{r} + m)) = q(ar) \\ &= q(a)r = q(a)(\bar{r} + m) = q(a)\bar{r} + q(a)m \\ &= q(a)\bar{r} + 0 = q(a)\bar{r} \,. \end{aligned}$$

Thus q is an $\bar{R}$-endomorphism.

LEMMA 4.3 *E^* is the R-injective hull of $(B_1)_R$.*

Proof. By Lemma 4.2 $(B_1)_R$ is an essential submodule of ${E^*}_R$. E^* is an injective $\bar{R}$-module since * is a category isomorphism; in particular E^* is a divisible $\bar{R}$-module so that E^* is a divisible R-module. R is a hereditary two-sided order so that E^* is an injective R-module by [5], Theorem 3.4.

THEOREM 4.4. *Let R be a Dedekind prime ring with a maximal two-sided ideal M, and let P be a maximal right ideal of R containing M. Then the R-endomorphism ring of the R-injective hull of R/P is a complete principal ideal domain.*

Proof. Let R, $\bar{R}$, L, E_L, and E^* be as above. Then by Lemma 4.3 E^* is the injective hull of a simple right R-module which is annihilated by M. $(B_1)_R \simeq R\backslash P$ since both are simple modules over the simple Artinian ring R/M; thus $E^* \simeq E(R/P)$. By Lemma 4.2 $\mathrm{end}_R(E^*) = \mathrm{end}_{\bar{R}}(E^*)$ which is isomorphic to $\mathrm{end}_L(E)$ since* is a category isomorphism. Hence the result follows by Proposition 3.3.

THEOREM 4.5. (Main Theorem) *Let R be a Dedekind prime ring with a nonzero prime ideal M, and let P be a maximal right ideal containing M with $E(R/P)$ the R-injective hull of R/P. Then $\bar{R}$, the completion of R at M, is isomorphic to the second endomorphism ring of $E(R/P)$.*

Proof. Consider E^*; as above $E^* \simeq E(R/P)$. By Lemma 4.2 the R and $\bar{R}$ structures of E^* are identical. Thus $\bar{R}$ is second endomorphism ring of $E(R/P)$ by Proposition 4.1.

REFERENCES

1. P. M. Cohn, *Morita Equivalence and Duality*, Queen Mary College, London.
2. C. Faith, *Lectures on Injective Modules and Quotient Rings*, Springer-Verlag, Berlin, 1967.
3. A. W. Goldie, *Localization in non-commutative Noetherian rings*, J. Algebra, **5** (1967), 89-105.

4. W. D. Gwynne and J. C. Robson, *Completions of Non-Commutative Dedekind Prime Rings*, Unpublished.
5. L. S. Levy, *Torsion-free and divisible modules over non-integral Domains*, Canadian J. Math., **15** (1963), 132-151.
6. E. Matlis, *Injective modules over Noetherian rings*, Pacific J. Math., **8** (1958), 511-528.
7. J. C. Robson, *Non-commutative Dedekind rings*, J. Algebra, **9** (1968), 249-265.

Received April 8, 1970.

UNIVERSITY OF WISCONSIN

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

ON GENERALIZED TRANSLATED QUASI-CESÀRO SUMMABILITY

B. KWEE

Let $\alpha > 0, \beta > -1$. The (C_t, α, β) transformation of the sequence $\{s_k\}$ is defined by

$$t_n = \frac{\Gamma(\beta+n+2)\Gamma(\alpha+\beta+1)}{\Gamma(n+1)\Gamma(\beta+1)\Gamma(\alpha)}\sum_{k=0}^{\infty}\frac{\Gamma(\alpha+k)\Gamma(k+n+1)}{\Gamma(k+1)\Gamma(\alpha+\beta+n+k+2)}s_k ,$$

and the (C_t, α, β) transformation of the function $s(x)$ is defined by

$$g(y) = \frac{\Gamma(\alpha+\beta+1)}{\Gamma(\alpha)\Gamma(\beta+1)}y^{\beta+1}\int_0^\infty \frac{x^{\alpha-1}s(x)}{(x+y)^{\alpha+\beta+1}}dx .$$

Some properties of the above two transformations are given in this paper and the relation between the summability methods defined by these transformations is discussed.

1. For any sequence $\{\mu_n\}$ the Hausdorff summability (H, μ_n) is defined by the transformation

$$t_n = \sum_{k=0}^{n}\binom{n}{k}(\Delta^{n-k}\mu_k)s_k ,$$

where

$$\Delta^0_{\mu_k} = \mu_k ,$$
$$\Delta\mu_k = \mu_k - \mu_{k+1} ,$$
$$\Delta^m_{\mu_k} = \Delta\Delta^{m-1}\mu_k .$$

Transposing the matrix of the (H, μ_n), transformation we get the matrix of the quasi-Hausdorff transformation

$$t_n = \sum_{k=n}^{\infty}\binom{k}{n}(\Delta^{k-n}\mu_n)s_k ,$$

which will be denoted by (H^*, μ_n). Ramanujan [8] introduced the (S, μ_n) summability, which is defined by the transformation

$$t_n = \sum_{k=0}^{\infty}\binom{k+n}{n}(\Delta^k\mu_n)s_k .$$

Thus the elements of row n of the matrix of the (S, μ_n) transformation are those of the corresponding row of the (H^*, μ_n) transformation moved n places to the left.

It is known [8] that if (H, μ_n) is regular and if $\mu_n \to 0$ as $n \to \infty$, then (S, μ_{n+1}) is regular; conversely, if (S, μ_{n+1}) is regular, then (H, μ_n)

can be made regular by a suitable choice of μ_0.

When

$$\mu_n = \frac{1}{\binom{n+\alpha}{n}},$$

(H, μ_n) reduces to the Cesàro summability (C, α). Borwein [3] introduced the generalized Cesàro summability (C, α, β) which is (H, μ_n) with

$$\mu_n = \frac{\binom{n+\beta}{n}}{\binom{n+\alpha+\beta}{n}}. \tag{1}$$

The aim of this paper is to discuss properties of the (S, μ_{n+1}) summability with μ_n given by (1) for $\alpha > 0, \beta > -1$ and of the analogous functional transformation. We shall denote this summability by (C_t, α, β). The case in which $\beta = 0$ has been considered by Kuttner [6] and a summability method similar to (C_t, α, β) has been discussed by me [7].

A straightforward calculation shows that the (C_t, α, β) transformation is given by

$$\begin{aligned} t_n = t(n, \alpha, \beta) &= \frac{(\beta+1)(\beta+2)\cdots(\beta+n+1)}{n!} \\ &\quad \times \sum_{k=0}^{\infty} \frac{\alpha(\alpha+1)\cdots(\alpha+k-1)(k+1)(k+2)\cdots(k+n)}{(\alpha+\beta+1)(\alpha+\beta+2)\cdots(\alpha+\beta+n+1+k)} s_k \\ &= \frac{\Gamma(\beta+n+2)\Gamma(\alpha+\beta+1)}{\Gamma(n+1)\Gamma(\beta+1)\Gamma(\alpha)} \sum_{k=0}^{\infty} \frac{\Gamma(\alpha+k)\Gamma(k+n+1)}{\Gamma(k+1)\Gamma(\alpha+\beta+n+k+2)} s_k . \end{aligned} \tag{2}$$

It is clear that, if (2) converges for one value of n, then it converges for all n. Further, a necessary and sufficient condition for this to happen is that

$$\sum_{k=1}^{\infty} \frac{s_k}{k^{\beta+2}} \tag{3}$$

should converge.

Let $s(x)$ be any function L-integrable in any finite interval of $x \geqq 0$ and bounded in some right-hand neighbourhood of the origin. Let $\alpha > 0, \beta > -1$, and let

$$g(y) = g(y, \alpha, \beta) = \frac{\Gamma(\alpha+\beta+1)}{\Gamma(\alpha)\Gamma(\beta+1)} y^{\beta+1} \int_0^{\infty} \frac{x^{\alpha-1} s(x)}{(x+y)^{\alpha+\beta+1}} dx . \tag{4}$$

If $g(y)$ exists for $y > 0$ and if

$$\lim_{y\to\infty} g(y) = s\ ,$$

we say that $s(x)$ is summable (C_t, α, β) to s.

It is clear that a necessary and sufficient condition for the convergence of (4) is that

$$\int_1^\infty \frac{s(x)}{x^{\beta+2}}dx \tag{5}$$

should converge.

2. The relationship between sequence-to-sequence and function-to-functions transformations. Given any sequence $\{s_n\}$, let the function $f(x)$ be defined by

$$f(x) = s_n \quad (n \leqq x < n + 1;\ n = 0, 1, 2, \cdots)\ .$$

Then the (C_t, α, β) summability of $\{s_n\}$ is equivalent to the (C_t, α, β) summability of $f(x)$ for $\alpha > 0$, $\beta = 0$ (see [6] Theorem 4). However, the proof breaks down when $\beta > 0$. We can prove that they are equivalent for $-1 < \beta \leqq 0$ as follows. Write

$$a(n, k) = \frac{\Gamma(\alpha + k)\Gamma(k + n + 1)}{\Gamma(k + 1)\Gamma(\alpha + \beta + n + k + 2)}$$

$$b(y, k) = \int_k^{k+1} \frac{x^{\alpha-1}}{(x + y)^{\alpha+\beta+1}}dx\ .$$

As in [6], we may suppose that $s_0 = 0$. Then the result would follow if, corresponding to equation (11) of [6], we proved that, if (3) converges, then uniformly for $0 \leqq \theta < 1$,

$$\sum_{k=1}^{\infty} [a(n, k) - b(n + \theta, k)]s_k = o\left(\frac{1}{n^{\beta+1}}\right)\ . \tag{6}$$

Choose an integer Q such that $Q \geqq \beta + 3$. From equations analogous to those of the last line and line 6 from bottom of p. 709 of [6], we find that

$$a(n, k) - b(n + \theta, k) = \Sigma p(\theta)\frac{k^{\alpha-q}}{(n + k)^{\alpha+\beta+r}} + O\left(\frac{k^{\alpha-Q}}{(k + n)^{\alpha+\beta+1}}\right)\ , \tag{7}$$

where $p(\theta)$ is a polynomial in θ (which may be different for each term in the sum), and the sum is taken over those integers q, r which are such that

$$q \geqq 1,\ r \geqq 1,\quad q, r \text{ not both } 1,\quad q + r \leqq Q\ .$$

Since the convergence of (3) implies that

$$s_k = o(k^{\beta+2})\ ,$$

and since $\alpha > 0$, $Q \geqq \beta + 3$, we see that the contribution to the expression on the left of (6) of the "0" term in (7) is

$$o\left(\frac{1}{n^{\beta+1}}\right).$$

Hence the result would follow if (corresponding to Lemma 2 of [6]) we could prove that the convergence of (3) implied that, for relevant q, r,

$$\sum_{k=1}^{\infty} \frac{k^{\alpha-q}}{(k+n)^{\alpha+\beta+r}} s_k = o\left(\frac{1}{n^{\beta+1}}\right). \tag{8}$$

Now write

$$v_k = \sum_{m=k}^{\infty} \frac{s_m}{m^{\beta+2}}$$

so that $v_k \to 0$ (and this is all we know). The sum on the left of (8) is

$$\begin{aligned} &\sum_{k=1}^{\infty} \frac{k^{\alpha+\beta+2-q}}{(k+n)^{\alpha+\beta+r}}(v_k - v_{k+1}) \\ &= \frac{v_1}{(n+1)^{\alpha+\beta+r}} + \sum_{k=2}^{\infty} v_k \left\{ \frac{k^{\alpha+\beta+2-q}}{(k+n)^{\alpha+\beta+r}} - \frac{(k-1)^{\alpha+\beta+2-q}}{(k-1+n)^{\alpha+\beta+r}} \right\}. \end{aligned} \tag{9}$$

The first term on the right of (9) is $o(1/n^{\beta+1})$ (since $r \geqq 1, \alpha > 0$). The expression in curly brackets in the second term is

$$O\left(\frac{k^{\alpha+\beta+1-q}}{(k+n)^{\alpha+\beta+r}}\right)$$

(and this result is best possible). This gives the required result when $\beta \leqq 0$; but if $\beta > 0$, all that we can deduce in the "worst" cases (which are $q = 1, r = 2$ or $q = 2, r = 1$) is that the sum (9) is $o(1/n)$.

Of course, the fact that the proof breaks down does not imply that the theorem itself is false. My guess is that the theorem probably is false for $p > 0$; but I have not actually got a counter example.

3. Theorems. The following two theorems with $\beta = 0$ are Theorem 1′ and Theorem 2′ given by Kuttner [6]. The proof of Theorem 1 is similar to that of Theorem 1′ in [6], and Theorem 2 follows from Lemma 1 and Lemma 2 of this paper.

THEOREM 1. *Let $\alpha > 0, \beta > -1$ and $r \geqq 0$ and let $s(x)$ be summable (C, r)*[1] *to s and* (4) *converge. Then $s(x)$ is summable (C_t, α, β) to s.*

THEOREM 2. *Let $\alpha > \alpha' > 0, \beta > -1$, and let $s(x)$ be summable (C_t, α, β) to s. Then $s(x)$ is summable (C_t, α', β) to s.*

[1] For definition of the (C, r) summability of $s(v)$, see [7].

In § 5, we shall prove

THEOREM 3. *Let* $\alpha > 0$, $\beta > \beta' > -1$. *Suppose that* $s(x)$ *is summable* (C_t, α, β) *to* s *and the integral*

$$\int_1^\infty \frac{s(x)}{x^{\beta'+2}} dx$$

converges. Then $s(x)$ *is summable* (C_t, α, β') *to* s.

The sequence $\{s_n\}$ is said to be summable A_λ to s if

$$f_\lambda(x) = (1-x)^{\lambda+1} \sum_{n=0}^\infty \binom{n+\lambda}{n} s_n x^n$$

converges for all x in the interval $0 \leqq x < 1$ and tend to a finite limit s as $x \to 1-$. The A_0 method is the ordinary Abel method.

It is known (see [1] and [2]) that $A_\mu \supset A_\lambda$ for $\lambda > \mu > -1$. For other properties of this summability method, see [1] and [6]. We shall prove

THEOREM 4. *Let* $\lambda > -1$, $\beta > -1$. *Suppose that the sequence* $\{s_n\}$ *is summable* A_λ *to* s *and that* (3) *converges. Then the sequence is summable* $(C_t, \lambda+1, \beta)$ *to* s.

4. Lemmas.

LEMMA 1. *Let* $\alpha > \alpha' > 0$, $\beta > -1$. *Suppose that* (5) *converges. Then*

$$y^{\alpha-1} g(y, \alpha', \beta) = \frac{\Gamma(\alpha)}{\Gamma(\alpha')\Gamma(\alpha-\alpha')} \int_0^y t^{\alpha'-1}(y-t)^{\alpha-\alpha'-1} g(t, \alpha, \beta) dt .$$

The proof of this lemma is similar to that of Lemma 4 in [6].

LEMMA 2. *Let*

$$t(x) = \int_0^\infty c(x, y) s(y) dy .$$

Then in order that

$$s(y) \to s \qquad (y \to \infty)$$

should imply

$$t(x) \to s \qquad (x \to \infty)$$

for every bounded $s(y)$, *it is sufficient that*

$$\int_0^\infty |c(x, y)| dy < H ,$$

where H is independent of x, that

$$\int_0^Y |c(x, y)|dy \to 0$$

when $x \to \infty$, *for every finite* Y, *and that*

$$\int_0^\infty c(x, y)dy \to 1$$

when $x \to \infty$.

This Theorem 6 in [4].

5. Proof of Theorem 3. Let

$$\phi(x) = \int_x^\infty \frac{s(u)}{u^{\beta+2}}du$$

for $x > 0$. Then $\phi(x)$ is continuous in $(0, \infty)$, and $\phi(x) \to 0$ as $x \to \infty$; hence $\phi(x)$ is bounded in (B, ∞) for any $B > 0$, say

$$|\phi(x)| \leqq M$$

for $x \geqq B$, where M may depend on B if B is small, but may be taken as an absolute constant for large B. It follows that

$$\begin{aligned} \left|\int_B^\infty \frac{x^{\alpha-1}s(x)}{(x+t)^{\alpha+\beta+1}}dx\right| &= \left|\int_B^\infty \left(\frac{x}{x+t}\right)^{\alpha+\beta+1} d\phi(x)\right| \\ &= \left|\left(\frac{B}{B+t}\right)^{\alpha+\beta+1}\phi(B)\right. \\ &\quad \left. + (\alpha+\beta+1)t\int_B^\infty \frac{1}{(x+t)^2}\left(\frac{x}{x+t}\right)^{\alpha+\beta}\phi(x)dx\right| \\ &\leqq |\phi(B)| + (\alpha+\beta+1)tM\int_B^\infty \frac{dx}{(x+t)^2} \\ &\leqq (\alpha+\beta+2)M\,. \end{aligned} \tag{10}$$

Since $s(x)$ is bounded in some right-hand neighbourhood of the origin, there exists $B_0 > 0$ such that

$$|s(x)| \leqq K$$

for $0 < x < B_0$. By partial integration, we obtain

$$\left|t^{\beta+1}\int_0^{B_0} \frac{x^{\alpha-1}s(x)}{(x+t)^{\alpha+\beta+1}}dx\right| \leqq \frac{K(\alpha+2\beta+2)}{\alpha(\beta+1)}\,. \tag{11}$$

By combining (10) and (11) it follows that $g(t, \alpha, \beta)$ is bounded in any finite interval $(0, T)$. Since it tends to s as $t \to \infty$, $g(t, \alpha, \beta)$ is bounded in $(0, \infty)$. Thus, for $y > 0$, the integral

$$I = \frac{\Gamma(\alpha)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+1)}\int_y^\infty t^{-\beta-1}(t-y)^{\beta-\beta'-1}g(t, \alpha, \beta)dt$$

converges. In view of the definition of $g(t, \alpha, \beta)$ it follows that

$$I = \lim_{A\to\infty} I(A) \tag{12}$$

where

$$I(A) = \int_y^A (t-y)^{\beta-\beta'-1}dt\int_0^\infty \frac{x^{\alpha-1}s(x)}{(x+t)^{\alpha+\beta+1}}dx\ .$$

It follows from (10) by dominated convergence that, for fixed A,

$$\int_y^A (t-y)^{\beta-\beta'-1}dt\int_B^\infty \frac{x^{\alpha-1}s(x)}{(x+t)^{\alpha+\beta+1}}dx \to 0$$

as $B \to \infty$. Hence, by Fubini's theorem

$$I(A) = \int_0^\infty x^{\alpha-1}s(x)dx\int_y^A \frac{(t-y)^{\beta-\beta'-1}}{(x+t)^{\alpha+\beta+1}}dt\ . \tag{13}$$

We will now show that, for fixed y,

$$\int_0^\infty x^{\alpha-1}s(x)dx\int_A^\infty \frac{(t-y)^{\beta-\beta'-1}}{(x+t)^{\alpha+\beta+1}}dt \to 0 \tag{14}$$

as $A \to \infty$. It is clear that for large A the inner integral in (14) is $O(A^{-\alpha-\beta'-1})$ uniformly in $0 \leqq x \leqq 1$, so that the contribution to (14) of the range $0 < x < 1$ tends to 0 as $A \to \infty$. Now write

$$\psi(x) = \int_x^\infty \frac{s(u)}{u^{\beta'+2}}du;$$

thus we are given that $\psi(x)$ exists and that it tends to 0 as $x \to \infty$. The contribution to (14) of $x > 1$ may now be written

$$-\int_1^\infty x^{\alpha+\beta'+1}d\psi(x)\int_A^\infty \frac{(t-y)^{\beta-\beta'-1}}{(x+t)^{\alpha+\beta+1}}dt\ . \tag{15}$$

It is easily seen that, for fixed y, A and large x, the inner integral in (15) is $O(x^{-\alpha-\beta'-1})$; thus, integrating by parts, (15) becomes

$$\psi(1)\int_A^\infty \frac{(t-y)^{\beta-\beta'-1}}{(1+t)^{\alpha+\beta+1}}dt \tag{16}$$
$$+\int_1^\infty x^{\alpha+\beta'}\psi(x)dx\int_A^\infty \frac{(t-y)^{\beta-\beta'-1}[(\alpha+\beta'+1)t-(\beta-\beta')x]}{(x+t)^{\alpha+\beta+2}}dt\ .$$

Now for fixed y and large A, uniformly in $0 \leqq x \leqq A$, the inner integral in (16) is

$$O\left\{\int_A^\infty t^{-\alpha-\beta'-2}dt\right\} = O(A^{-\alpha-\beta'-1}) .$$

Hence

$$\int_1^A x^{\alpha+\beta'}\psi(x)dx\int_A^\infty \frac{(t-y)^{\beta-\beta'-1}[(\alpha+\beta'+1)t-(\beta-\beta')x]}{(x+t)^{\alpha+\beta+2}}dt$$
$$= \left(\int_1^{A/\log A} + \int_{A/\log A}^A\right)x^{\alpha+\beta'}\psi(x)O(A^{-\alpha-\beta'-1})dx$$
$$= O\left(A^{-\alpha-\beta'-1}\int_1^{A/\log A} x^{\alpha+\beta'}dx\right)$$
$$+ O\left(A^{-\alpha-\beta'-1}\sup_{x\geqq(A/\log A)}|\psi(x)|\int_{A/\log A}^A x^{\alpha+\beta'}dx\right) = O(1) .$$

Nothing that for fixed y and large t

$$(t-y)^{\beta-\beta'-1} = t^{\beta-\beta'-1} + O(t^{\beta-\beta'-2}) ,$$

and also that

$$\int_0^\infty \frac{t^{\beta-\beta'-1}[(\alpha+\beta'+1)t-(\beta-\beta')x]}{(x+t)^{\alpha+\beta+2}}dt = 0 ,$$

we see that, for large A uniformly in $x \geqq A$, the inner integral in (16) is

$$-\int_0^A \frac{t^{\beta-\beta'-1}[(\alpha+\beta'+1)t-(\beta-\beta')x]}{(x+t)^{\alpha+\beta+2}}dt$$
$$+ O\left\{\int_A^\infty \frac{t^{\beta-\beta'-2}|(\alpha+\beta'+1)t-(\beta-\beta')x|}{(x+t)^{\alpha+\beta+2}}dt\right\}$$
$$= O\left\{x^{-\alpha-\beta-1}\int_0^A t^{\beta-\beta'-1}dt\right\} + O\left\{x^{-\alpha-\beta-1}\int_A^x t^{\beta-\beta'-2}dt\right\} + O\left\{\int_x^\infty t^{-\alpha-\beta'-3}dt\right\}$$
$$= O(x^{-\alpha-\beta-1}A^{\beta-\beta'}) + O(x^{-\alpha-\beta'-2})$$

(except that, in the case $\beta-\beta'=1$, we must insert an extra term $O(x^{-\alpha-\beta-1}\log x)$). It is now clear that the expression (16) tends to 0 as $A\to\infty$, and this completes the proof of (14). We deduce from (12), (13) and (14) that

$$I = \int_0^\infty x^{\alpha-1}s(x)dx\int_y^\infty \frac{(t-y)^{\beta-\beta'-1}}{(x+t)^{\alpha+\beta+1}}dt$$
$$= \frac{\Gamma(\beta-\beta')\Gamma(\alpha+\beta'+1)}{\Gamma(\alpha+\beta+1)}\int_0^\infty \frac{x^{\alpha-1}s(x)}{(x+y)^{\alpha+\beta'+1}}dx$$
$$= \frac{\Gamma(\beta-\beta')\Gamma(\alpha)\Gamma(\beta'+1)}{\Gamma(\alpha+\beta+1)}y^{-\beta'-1}g(y,\alpha,\beta') .$$

Thus, in view of the definition of I, we have

$$g(y, \alpha, \beta') = \frac{\Gamma(\beta + 1)}{\Gamma(\beta - \beta')\Gamma(\beta' + 1)} y^{\beta'+1} \int_y^\infty t^{-\beta-1}(t - y)^{\beta-\beta'-1} g(t, \alpha, \beta) dt .$$

The kernel of this last transformation can easily be verified to satisfy the conditions of Lemma 2, and the theorem now follows.

6. Proof of Theorem 4. It follows from the convergence of (3) that for $\beta > -1$, $s_\nu = o(\nu^{\beta+2})$. We can easily prove that the function $t^{n+k}(1 - t)^{\lambda+\beta'+1}$ has a maximum when

$$t = \frac{k + n}{k + n + \lambda + \beta' + 1} .$$

For large $k + n$, this maximum is $O((k + n)^{-\lambda-\beta'-1})$. Hence, if $\beta' > \beta + 2$, we have, the inversion in the order of integration and summation being justified by absolute convergence,

$$\begin{aligned}
&\frac{\Gamma(\beta' + n + 2)}{\Gamma(n + 1)\Gamma(\beta' + 1)} \int_0^1 t^n (1 - t)^{\lambda+\beta'+1} \left\{ \sum_{k=0}^\infty \binom{\lambda + k}{k} s_k t^k \right\} dt \\
(17) \quad &= \frac{\Gamma(\beta' + n + 2)}{\Gamma(n + 1)\Gamma(\beta' + 1)} \sum_{k=0}^\infty \binom{\lambda + k}{k} s_k \int_0^1 t^{k+n}(1 - t)^{\lambda+\beta'+1} dt \\
&= \frac{\Gamma(\beta'+n+2)\Gamma(\lambda+\beta'+2)}{\Gamma(n+1)\Gamma(\beta'+1)\Gamma(\lambda+1)} \sum_{k=0}^\infty \frac{\Gamma(\lambda+k+1)\Gamma(k+n+1)}{\Gamma(k+1)\Gamma(\lambda+\beta'+n+k+3)} s_k \\
&= t(n, \lambda + 1, \beta') .
\end{aligned}$$

By analytic continuation, (17) holds for $\beta' \geqq \beta$. Hence

$$\begin{aligned}
t(n, \lambda + 1, \beta) &= \frac{\Gamma(\beta + n + 2)}{\Gamma(n + 1)\Gamma(\beta + 1)} \int_0^1 t^n (1 - t)^\beta f_\lambda(t) dt \\
&= \frac{\Gamma(\beta + n + 2)}{\Gamma(n + 1)\Gamma(\beta + 1)} \int_0^\infty (1 - e^{-y})^n e^{-(\beta+1)y} f_\lambda(1 - e^{-y}) dy .
\end{aligned}$$

By Lemma 2 the result with follow if

$$\text{(i)} \qquad \frac{\Gamma(\beta + n + 2)}{\Gamma(n + 1)\Gamma(\beta + 1)} \int_0^\infty (1 - e^{-y})^n e^{-(\beta+1)y} dy < H$$

where H is independent of n,

$$\text{(ii)} \qquad \frac{\Gamma(\beta + n + 2)}{\Gamma(n + 1)\Gamma(\beta + 1)} \int_0^Y (1 - e^{-y})^n e^{-(\beta+1)y} dy \to 0$$

when $n \to \infty$, for every finite Y, and

$$\text{(iii)} \qquad \frac{\Gamma(\beta + n + 2)}{\Gamma(n + 1)\Gamma(\beta + 1)} \int_0^\infty (1 - e^{-y})^n e^{-(\beta+1)y} dy \to 1 ,$$

when $n \to \infty$. Since

$$\int_0^\infty (1 - e^{-y})^n e^{-(\beta+1)y} dy = \frac{\Gamma(n+1)\Gamma(\beta+1)}{\Gamma(\beta+n+2)} ,$$

(i) and (iii) are satisfied. We have $\Gamma(n + \beta + 2) \sim n^{\beta+1}\Gamma(n + 1)$, and the integral in (ii) is, by changing the variable,

$$\int_0^{1-e^{-Y}} t^n(1 - t)^\beta dt .$$

Hence (ii) is satisfied.

REFERENCES

1. D. Borwein, *On a scale of Abel-type summability methods*, Proc. Camb. Phil. Soc., **53** (1957), 318-322.
2. ———, *On methods of summability based on power series*, Proc. Royal Soc. Edin, **64** (1957), 342-349
3. ———, *Theorems on some methods of summability*, Quart. J. Math., (Oxford), (2) **9** (1958), 310-316.
4. G. H. Hardy, *Divergent series*.
5. A. Jakimovski, *Some relations between the methods of summability of Abel, Borel,, Cesàro, Hölder and Hausdorff*, J. Anal. Math., **3** (1954), 346-381.
6. B. Kuttner, *On translated quasi-Cesàro summability*, Proc. Camb. Phil. Soc., **62** (1966), 705-712.
7. B. Kwee, *On Perron's method of summation*, ibid., **63** (1967), 1033-1040.
8. M. S. Ramanujan, *On Hausdorff and quasi-Hausdorff methods of summability*, Quart, J. Math. (Oxford), (2) **8** (1957), 197-213.

Received December 1, 1969.

UNIVERSITY OF MALAYA,
KUALA LUMPUR, MALAYSIA.

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

DIFFERENTIAL SIMPLICITY AND COMPLETE INTEGRAL CLOSURE

YVES LEQUAIN

Let R be an integral domain containing the rational numbers, and let R' denote the complete integral closure of R. It is shown that if R is differentiably simple, then R need not be equal to R', even when R is Noetherian, and then the relationship between R and R' is studied.

Let $\mathscr{D}$ be any set of derivations of R. Seidenberg has shown that the conductor $C = \{x \in R \mid xR' \subset R\}$ is a $\mathscr{D}$-ideal of R, so that when R is $\mathscr{D}$-simple and $C \neq 0$, then $R = R'$. We investigate here the situation when $C = 0$.

The first observation that one must make is that it is no longer true that $R = R'$ when R is differentiably simple, even when R is Noetherian. We show this in Example 2.2 where we construct a 1-dimensional local domain containing the rational numbers which is differentiably simple but not integrally closed. This counterexamples a conjecture of Posner [4, p. 1421] and also answers affirmatively a question of Vasconcelos [6, p. 230].

Thus, it is not a redundant task to study the relationship between a differentiably simple ring R and its complete integral closure. An important tool in this study is the technique of § 3 which associates to any prime ideal P of R containing no D-ideal a rank-1, discrete valuation ring centered on P; by means of this, we show in Theorem 3.2 that over such a prime ideal P of R there lies a unique prime ideal of R'. When R is a Noetherian $\mathscr{D}$-simple ring with $\{P_\alpha\}_{\alpha \in \Lambda}$ as set of minimal prime ideals, Theorem 3.3 asserts that $R' = \bigcap_{\alpha \in \Lambda} \{R_\alpha \mid R_\alpha$ is the valuation ring associated with the minimal prime ideal $P_\alpha\}$; Corollary 3.5 asserts that R' is the largest $\mathscr{D}$-simple overring of R having a prime ideal lying over every minimal prime ideal of R.

1. Preliminaries. Our notation and terminology adhere to that of Zariski-Samuel [7] and [8]. Throughout the paper we use R to denote a commutative ring with 1, K to denote the total quotient ring of R, and A to denote an ideal of R; A is proper if $A \neq R$. A derivation D of R is a map of R into R such that

$$D(a + b) = D(a) + D(b) \quad \text{and} \quad D(ab) = aD(b) + bD(a)$$

for all $a, b \in R$.

Such a derivation can be uniquely extended to K, and we shall also denote the extended derivation by D. D is said to be regular on a subring S of K if $D(S) \subset S$. If $\mathscr{D}$ is a family of derivations of R, A is called a $\mathscr{D}$-ideal if $D(A) \subset A$ for every $D \in \mathscr{D}$; when $\mathscr{D} = \{D\}$, we merely say D-ideal. If R has no $\mathscr{D}$-ideal different from (0) and (1), R is said to be $\mathscr{D}$-simple. We use $D^{(0)}(x)$ to denote x, and for $n \geqq 1$ $D^{(n)}(x)$ to denote $D(D^{(n-1)}(x))$, i.e. the n^{th} derivative of x; by induction one proves Leibnitz's rule:

$$D^{(n)}(ab) = \sum_{i=0}^{n} C_n^i D^{(n-i)}(a) D^{(i)}(b) .$$

We assume henceforth that $\mathscr{D}$ is a family of derivations of R and that $D \in \mathscr{D}$. Let $\varphi: R \rightarrow S$ be a homomorphism onto; then

$$D'(\varphi(r)) = \varphi(D(r))$$

defines a derivation D' on S if and only if the kernel I of φ is a D-ideal. Suppose that I is a $\mathscr{D}$-ideal, and write $\mathscr{D}'$ to denote the set of derivations of S thus induced by $\mathscr{D}$; if A is a $\mathscr{D}$-ideal of R, then $\varphi(A)$ is a $\mathscr{D}'$-ideal of S, and conversely if B is a $\mathscr{D}'$-ideal of S, then $\varphi^{-1}(B)$ is a $\mathscr{D}$-ideal of R containing I. Thus, in particular, if A is a maximal proper $\mathscr{D}$-ideal of R, then R/A is $\mathscr{D}'$-simple.

LEMMA 1.1. *Let D be a derivation of R, M a multiplicative system of R, and $h: R \rightarrow R_M$ the canonical homomorphism. Then, we can define a derivation on R_M, which we also call D, by*

$$D(h(r)(h(m))^{-1}) = [h(m)h(D(r)) - h(r)h(D(m))](h(m^2))^{-1} .$$

Furthermore, if A is a D-ideal of R, then $h(A)R_M$ is a D-ideal of R_M, and if B is a D-ideal of R_M, then $h^{-1}(B)$ is a D-ideal of R.

Proof. $\ker h = \{x \in R \mid xm = 0 \text{ for some } m \in M\}$ is a D-ideal of R since $0 = D(xm) = xD(m) + mD(x) = xmD(m) + m^2D(x) = m^2D(x)$. Hence D induces a derivation on $R/\ker h$, a derivation which can be then extended to R_M. The remainder of the lemma is straightforward.

LEMMA 1.2. *Let $\mathscr{D}$ be a family of derivations of R, and suppose that R contains the rational numbers. Then, the radical of a $\mathscr{D}$-ideal of R is a $\mathscr{D}$-ideal.*

Proof. See [2, Lemma 1.8, p. 12].

COROLLARY 1.3. *If P is a minimal prime divisor of a $\mathscr{D}$-ideal*

A, and P does not contain an integer $\neq 0$, then P is a $\mathscr{D}$-ideal.

Proof. Localize at P and apply 1.1 and 1.2.

THEOREM 1.4. *Let A be a maximal proper $\mathscr{D}$-ideal of R, then*

(i) *A is primary.*

(ii) *If R/A has characteristic $p \neq 0$, then $\sqrt{A}$ is a maximal ideal.*

(iii) *If R/A has characteristic 0, then A is prime.*

Proof. (i) Suppose $x, y \in R, x \notin A$ and $xy \in A$; then, $\bigcup_{n=0}^{\infty} (A: y^n) \supset A: y > A$. But $\bigcup_{n=0}^{\infty} (A: y^n)$ is a $\mathscr{D}$-ideal; hence, by the maximality of A, $\bigcup_{n=0}^{\infty} (A: y^n) = R$ and there exists n such that $y^n \in A$.

(ii) Let P be a maximal ideal of R containing A. Consider the ideal $B = (A, \{x^p \mid x \in P\}) \subset P$; since R/A has characteristic p, B is a $\mathscr{D}$-ideal; hence, by the maximality of A, $B = A$ and $P = \sqrt{A}$.

(iii) Since R/A has characteristic 0, A contains no integer other than 0, hence the prime ideal $P = \sqrt{A}$ contains no integer either, and by 1.3 P is a $\mathscr{D}$-ideal. Then, by the maximality of A, $P = A$.

COROLLARY 1.5. *Let R be of characteristic 0. Then R is $\mathscr{D}$-simple if R contains the rational numbers and has no prime $\mathscr{D}$-ideal different from* (0) *and* (1). *If R is $\mathscr{D}$-simple, then R is a domain.*

One should note that a $\mathscr{D}$-simple ring R always contains a field, namely $F = \{x \in R \mid D(x) = 0 \text{ for all } D \in \mathscr{D}\}$; moreover, if the characteristic of R is $p \neq 0$, 1.4 shows that R is a primary ring and hence is equal to its total quotient ring; so this case will not be of interest in our further considerations, and throughout the remainder of this section we shall be dealing with a $\mathscr{D}$-simple ring of characteristic 0, which is then a domain containing the rational numbers.

DEFINITION 1.6. Let R be a domain with quotient field K. An element $x \in K$ is said to be quasi-integral over R if there exists an element $d \in R$, $d \neq 0$, such that $dx^n \in R$ for all $n \geqq 1$. The set R' of all elements of K that are quasi-integral over R is a ring, called the complete integral closure of R. R is said to be completely integrally closed if $R = R'$. Note that if R is Noetherian, the concepts of integral dependence and quasi-integral dependence over R for elements of K become the same.

LEMMA 1.7. *Let R be a domain with quotient field K, S a ring* *such that $R \subset S \subset K$, and $\mathscr{D}$ a family of derivations of R regular on S. Then S is $\mathscr{D}$-simple if R is $\mathscr{D}$-simple.*

Proof. If B is any $\mathscr{D}$-ideal of S, then $B \cap R$ is a $\mathscr{D}$-ideal of R, and if B is different from (0) then $B \cap R$ is also different from (0) since $S \subset K$.

THEOREM 1.8. *Let R be a domain of characteristic 0 and R' its complete integral closure. Then R' is $\mathscr{D}$-simple if R is $\mathscr{D}$-simple.*

Proof. By [5, p. 168], any $D \in \mathscr{D}$ is regular on R', hence the theorem follows from 1.7.

2. Example of a 1-dimensional local ring which is D-simple but not integrally closed. First, in this section, we modify an idea of Akizuki in [1] to construct some 1-dimensional local ring R of arbitrary characteristic such that the integral closure $\bar{R}$ is not a finite R-module.

THEOREM 2.1. *Let k be a field of arbitrary characteristic, Y an indeterminate over k, $\pi = a_1 Y + a_2 Y^3 + \cdots + a_r Y^{2^r-1} + \cdots$ an element of $k[[Y]]$ which is transcendental over $k[Y]$*[1]. *Set*

$$\theta_1 = \pi Y^{-1}, \theta_r = (\theta_{r-1} - a_{r-1}) Y^{-2^{r-1}}$$

for $r \geqq 2$ (alternatively $\theta_r = a_r + a_{r+1} Y^{2^r} + \cdots + a_s Y^{2^s - 2^r} + \cdots$); for $r \geqq 1$, set

$$t_r = (\theta_r - a_r)^2 \quad \text{and} \quad \pi_r = \pi - (a_1 Y + \cdots + a_r Y^{2^r-1}).$$

Set also $T = k[Y, \pi, t_1, t_2, \cdots, t_r \cdots]$ and $P = (Y, \pi)T$. Note that $T \subset k[[Y]]$ and that $P \subset Yk[[Y]]$. Then,

(i) *For $r > 1$, $t_{r-1} = Y^{2^r}(a_r^2 + t_r) + 2a_r Y \pi_r$ and P is a maximal ideal of T.*

(ii) *For $r \geqq 1$, $\pi_r^2 = Y^{2^{r+1}-2}$tr and $k(Y, \pi)$ is the quotient field of T.*

(iii) *The ring $R = T_P$ is a 1-dimensional local domain.*

(iv) *The integral closure $\bar{R}$ of R is not a finite R-module.*

Proof. (i) For $r > 1$, we have

$$t_{r-1} = (\theta_{r-1} - a_{r-1})^2 = (Y^{2^{r-1}} \theta_r)^2 = Y^{2^r}(a_r^2 + t_r) + 2a_r Y^{2^r}(\theta_r - a_r).$$

But

$$Y^{2^r}(\theta_r - a_r) = Y[\pi - (a_1 Y + \cdots + a_r Y^{2^r-1})] = Y\pi_r,$$

[1] Such an element exists; take for example $\pi = a_1 Y + a_2 Y^3 + \cdots + a_r Y^{2^{r!}-1} + \cdots$ with $a_r \neq 0$ for every $r \geqq 1$.

hence $t_{r-1} = Y^{2^r}(a_r^2 + t_r) + 2a_r Y\pi_r$. Since furthermore $P \subset Yk[[Y]]$, $1 \notin P$, and P is a maximal ideal of T.

(ii)

$$\begin{aligned}\pi_r &= \pi - (a_1 Y + \cdots + a_r Y^{2^r-1}) \\ &= Y^{2^r-1}(a_{r+1}Y^{2^r} + \cdots + a_{r+\ell}Y^{2^{r+\ell}-2^r} + \cdots) \\ &= Y^{2^r-1}(\theta_r - a_r)\ ;\end{aligned}$$

thus $\pi_r^2 = Y^{2^{r+1}-2}t_r$ and $k(Y, \pi)$ is the quotient field of T.

(iii) Let us show that Y belongs to every nonzero prime ideal of R. Since $k(Y, \pi)$ is the quotient field of R it suffices to show that $R[Y^{-1}] = k(Y, \pi)$. Let $\beta \in k[Y, \pi]$; then $\beta = \sum_{i=0}^n s_i\pi^i$ with $s_i \in k[Y]$. For any integer $r \geqq 1$, set $f_r = \sum_{i=0}^n s_i(a_1 Y + \cdots + a_r Y^{2^r-1})^i$; then

$$f_{r+1} = \sum_{i=0}^n s_i(a_1 Y + \cdots + a_r Y^{2^r-1} + a_{r+1}Y^{2^{r+1}-1})^i = f_r + Y^{2^{r+1}-1}h_{r+1}$$

with $h_{r+1} \in k[Y]$, and since $2^{r+1} - 1 > r$, we have $f_r = b_0 + b_1 Y + \cdots + b_r Y^r + Y^{r+1}g_r$ and

$$f_{r+1} = b_0 + b_1 Y + \cdots + b_r Y^r + b_{r+1}Y^{r+1} + Y^{r+2}g_{r+1}$$

with $b_0, \cdots, b_r, b_{r+1} \in k$ and $g_r, g_{r+1} \in k\,[Y]$. Now, since

$$\pi = \pi_r + (a_1 Y + \cdots + a_r Y^{2^r-1}), \qquad \beta = \sum_{i=0}^n s_i\pi^i = \pi_r\delta_r + f_r$$

with $\delta_r \in T$. Hence, there exists $b_0, b_1, \cdots, b_r, \cdots \in k$, $\delta_1, \cdots, \delta_r, \cdots \in T$ and $g_1, \cdots, g_r, \cdots \in k[Y]$ such that

$$(*) \qquad \beta = \sum_{j=0}^r b_j Y^j + \pi_r\delta_r + Y^{r+1}g_r\ .$$

Note that $\pi_r \in P$ and therefore that π_r is a nonunit in R.

If $b_0 \neq 0$, with $r = 1$, the relation (*) gives that $\beta = b_0 + (b_1 Y + \pi_1\delta_1 + Y^2 g_1)$ is a unit in R and thus that $\beta^{-1} \in R \subset R[Y^{-1}]$.

If $b_0 = b_1 = \cdots = b_{r-1} = 0$ and $b_r \neq 0$, the relation (*) gives $\beta = Y^r(b_r + Yg_r) + \pi_r\delta_r$ where $w_r = b_r + Yg_r$ is a unit in R; then

$$\beta(Y^r w_r - \pi_r\delta_r) = Y^{2r}w_r^2 - \pi_r^2\delta_r^2 = Y^{2r}(w_r^2 - Y^{2^{r+1}-2r-2}t_r\delta_r^2)$$

where $w_r^2 - Y^{2^{r+1}-2r-2}t_r\delta_r^2$ is a unit in R, so that $\beta^{-1} \in R[Y^{-1}]$.

If $b_r = 0$ for every $r \geqq 0$, then by the relation (*) we have

$$\beta \in \bigcap_{r=1}^\infty (\pi_r, Y^{r+1})T \subset \bigcap_{r=1}^\infty Y^{r+1}k[[Y]] = (0)\ .$$

Thus, if $\beta \in k[Y, \pi]$, either $\beta^{-1} \in R[Y^{-1}]$ or $\beta = 0$. If $\eta \in k(Y, \pi)$, then $\eta = \nu\lambda^{-1}$ with $\nu, \lambda \in k[Y, \pi], \lambda \neq 0$, so that $\eta \in R[Y^{-1}]$; hence $R[Y^{-1}] = k(Y, \pi)$.

Now,

$$\pi^2 = (Y\theta_1)^2 = [a_1 Y + (\theta_1 - a_1) Y]^2 = (t_1 - a_1^2) Y^2 + 2a_1 Y\pi$$

so that $Y^{-1} \in R[\pi^{-1}]$, $k(Y, \pi) = R[Y^{-1}] \subset R[\pi^{-1}]$, and π belongs also to every nonzero prime ideal of R. Thus $PR = (Y, \pi)R$, which is the unique maximal ideal of R and which is contained in every nonzero prime ideal of R, is the only nonzero prime ideal of R. As furthermore PR is finitely generated, R is a 1-dimensional local ring.

(iv) First, let us show that $\theta_1 = \pi Y^{-1} \notin T$. Suppose that $\theta_1 \in T = k[Y, \pi, t_1, \cdots, t_r, \cdots]$; then $\theta_1 = f(\pi, t_1, \cdots, t_\ell)$ where f is a polynomial in $\ell + 1$ indeterminates over $k[Y]$. For $r < \ell$, by (i), t_r can be expressed as a linear combination of 1, t_ℓ and π with coefficients in $k[Y]$, hence $\theta_1 = f(\pi, t_1, \cdots, t_\ell) = F(\pi, t_\ell) = F(Y\theta_1, (\theta_\ell - a_\ell)^2)$ where F is a polynomial in two indeterminates over $k[Y]$. Furthermore, by definition $\theta_{r-1} = Y^{2^{r-1}}\theta_r + a_{r-1}$, hence $\theta_1 = Y^{2^\ell - 2}\theta_\ell + \beta_\ell$ with $\beta_\ell \in k[Y]$ and we have

$$(^{**}) \qquad Y^{2^\ell - 2}\theta_\ell = G(Y^{2^\ell - 1}\theta_\ell, (\theta_\ell - a_\ell)^2)$$

where G is a polynomial in two indeterminates over $k[Y]$; but π being transcendental over $k[Y]$, θ_ℓ is transcendental over $k[Y]$ also, and the relation $(^{**})$ has to be an identity, which is absurd. Thus, $\theta_1 \notin T$.

Now, let R^* be the completion of R with the (PR)-adic topology; $\{\pi_r\}_{r \geq 0}$ is a Cauchy sequence in R. Suppose that $\pi_r \in P^2R$ for some $r \geq 1$; since P^2 is a primary ideal of T, we have $\pi_r \in P^2R \cap T = P^2 \subset YT$, and $\pi = \pi_r + (a_1 Y + \cdots + a_r Y^{2^r - 1}) \in YT$ which is absurd since $\theta_1 \notin T$. Thus, for every $r \geq 0$, $\pi_r \notin P^2R$ and $\beta = \lim_r \pi_r$ is $\neq 0$. However, we also have $\beta^2 = \lim_r \pi_r^2 = \lim_r Y^{2^{r+1} - 2} t_r = 0$; hence R^* has a nonzero nilpotent element and $\bar{R}$ is not a finite R-module [1, p. 330].

EXAMPLE 2.2. Let Q be the rational numbers, $(X_1, \cdots, X_r, \cdots)$ a set of indeterminates over Q and $k = Q(X_1, \cdots, X_r, \cdots)$. Let

$$\pi = b_1 X_1 Y + \cdots + b_r X_r Y^{2^r - 1} + \cdots$$

be transcendental over $k[Y]$ with $b_i \in Q - \{0\}$ for every $i \geq 1$[2]. Construct the rings $T = k[Y, \pi, t_1, \cdots, t_r, \cdots]$ and $R = T_P$ as in 2.1. On the quotient field $k(Y, \pi) = Q(X_1, \cdots, X_r, \cdots; Y, \pi)$ define a derivation D by

$$\begin{aligned} D(q) &= 0 \quad \text{for every} \quad q \in Q \\ D(Y) &= 1 \\ D(\pi) &= 3b_2 X_2 Y^2 + b_1 X_1 \\ D(X_1) &= 0 \end{aligned}$$

[2] There exists such a π since k is countable.

$$D(X_2) = -7b_3b_2^{-1}X_3Y^3$$
$$\vdots$$
$$D(X_i) = -(2^{i+1}-1)b_{i+1}b_i^{-1}X_{i+1}Y^{2^{i+1}-2^{i}-1}$$
$$\vdots$$

Then,

(i) D is regular on R

(ii) R is a 1-dimensional local D-simple ring which is not integrally closed.

Proof. (i) Since $R = T_P$, it suffices to show that $D(T) \subset R$. By definition of D we already have $D(k) \subset R$, $D(Y) \in R$ and $D(\pi) \in R$; hence it remains to show that $D(t_r) \in R$ for every $r \geqq 1$. Differentiating $\pi_r^2 = Y^{2^{r+1}-2}t_r$, we get $2\pi_r D(\pi_r) = Y^{2^{r+1}-2}D(t_r) + (2^{r+1}-2)Y^{2^{r+1}-3}t_r$; but $t_r \in YR$ by 2.1, hence $D(t_r) \in R$ if and only if $\pi_r D(\pi_r) \in Y^{2^{r+1}-2}R$. Let us show that in fact we have $D(\pi_r) \in Y^{2^{r+1}-2}R$. From $\pi_1 = \pi - b_1X_1Y$ we get $D(\pi_1) = D(\pi) - b_1X_1 = 3b_2X_2Y^2$; by induction, if we suppose that $D(\pi_{r-1}) = (2^r-1)b_rX_rY^{2^r-2}$ and if we differentiate the relation $\pi_r = \pi_{r-1} - b_rX_rY^{2^r-1}$, we get $D(\pi_r) = (2^{r+1}-1)b_{r+1}X_{r+1}Y^{2^{r+1}-2} \in Y^{2^{r+1}-2}R$. Hence D is regular on R.

(ii) The only prime ideal of R which is not (0) or (1) is $PR = (Y, \pi)R$; it is not a D-ideal since $D(Y) = 1$; thus by 1.5, R is D-simple. Furthermore by 2.1. R is a 1-dimensional local, not integrally closed, domain.

3. On the complete integral closure of a $\mathscr{D}$-simple ring. We have seen in the preliminaries that a $\mathscr{D}$-simple ring of characteristic $p \neq 0$ is equal to it total quotient ring. In this section we are concerned with rings of characteristic 0. Henceforth, R will denote a ring containing the integers.

THEOREM 3.1. *Let R be a ring, D a derivation on R, P a prime ideal of R containing no D-ideal other than* (0). *Define $v: R\backslash\{0\} \to$ {nonnegative integers} by $v(x) = n$ if $D^{(i)}(x) \in P$ for $i = 0, \cdots, n-1$ and $D^{(n)}(x) \notin P$. Then,*

(i) *R is domain.*

(ii) *v is rank-1-discrete valuation whose valuation ring R_v contains R and whose maximal ideal M_v lies over P.*

(iii) *D is regular on R_v and R_v is D-simple.*

Proof. (i) If n is any integer, $D(n) = 0$ and nR is a D-ideal of R; hence 0 is the only integer contained in P. Now, (0) is a D-ideal, hence by 1.3 any minimal prime divisor Q of (0) is a D-ideal also; then, by the hypothesis made on P, we have $(0) = Q$ and R is a domain.

(ii) Let x and y be two nonzero elements of R, and let $v(x) = n$, $v(y) = m$, $n \leqq m$. For every i such that $0 \leqq i \leqq n-1$, both $D^{(i)}(x)$ and $D^{(i)}(y)$ belong to P, hence $D^{(i)}(x+y) \in P$ and

$$v(x+y) \geqq n = \inf\{v(x), v(y)\} .$$

Let k be such that $0 \leqq k \leqq n+m-1$. For $0 \leqq i \leqq \inf\{k, n-1\}$ we have $D^{(i)}(x) \in P$, hence also $C_k^i D^{(i)}(x) D^{(k-i)}(y) \in P$; for $n \leqq k$ and $n \leqq i \leqq k$ we have $0 \leqq k-i \leqq k-n \leqq m-1$, hence $D^{(k-i)}(y) \in P$ and $C_k^i D^{(i)}(x) D^{(k-i)}(y) \in P$; thus

$$D^{(k)}(xy) = \sum_{i=0}^{k} C_k^i D^{(i)}(x) D^{(k-i)}(y) \in P .$$

Now,

$$D^{(n+m)}(xy) = \sum_{i=0}^{n+m} C_{n+m}^i D^{(i)}(x) D^{(n+m-i)}(y); \sum_{i=0}^{n-1} C_{n+m}^i D^{(i)}(x) D^{(n+m-i)}(y)$$
$$+ \sum_{i=n+1}^{n+m} C_{n+m}^i D^{(i)}(x) D^{(n+m-1)}(y) \in P$$

whereas $C_{n+m}^n D^{(n)}(x) D^{(m)}(y) \notin P$ since C_{n+m}^n, $D^{(n)}(x)$, $D^{(m)}(y) \notin P$; thus

$$D^{(n+m)}(xy) \notin P , \qquad v(xy) = n+m = v(x) + v(y)$$

and v is a valuation, rank-1-discrete since its value group is the group of integers. Furthermore, we obviously have $R \subset R_v$ and $M_v \cap R = P$.

(iii) Let ab^{-1} be any element of R_v with $a, b \in R$, $b \neq 0$, $v(a) \geqq v(b)$; then $D(ab^{-1}) = [bD(a) - aD(b)]b^{-2}$. If $v(a) > v(b)$, then $v(D(a)) = v(a) - 1 \geqq v(b)$ and $v(D(b)) \geqq v(b) - 1$ so that

$$v(bD(a) - aD(b)) \geqq \inf\{v(b) + v(D(a)), v(a) + v(D(b))\} \geqq 2v(b)$$

and $D(ab^{-1}) \in R_v$. If $v(a) = v(b) = 0$, then $v(bD(a) - aD(b)) \geqq 0 = 2v(b)$ and $D(ab^{-1}) \in R_v$. If $v(a) = v(b) = n > 0$, then $v(bD(a)) = v(aD(b)) = 2n-1$, so that $D^{(k)}(bD(a) - aD(b)) \in P$ for every $k \leqq 2n-2$; furthermore we have

$$D^{(2n-1)}(bD(a)) = \sum_{i=0}^{2n-1} C_{2n-1}^i D^{(i)}(b) D^{(2n-i)}(a) = \alpha_1 + C_{2n-1}^n D^{(n)}(b) D^{(n)}(a)$$

with $\alpha_1 \in P$, and similarly $D^{(2n-1)}(aD(b)) = \alpha_2 + C_{2n-1}^n D^{(n)}(a) D^{(n)}(b)$ with $\alpha_2 \in P$, so that $D^{(2n-1)}(bD(a) - aD(b)) = \alpha_1 - \alpha_2 \in P$; hence, $v(bD(a) - aD(b)) \geqq 2n$ and $D(ab^{-1}) \in R_v$. Thus D is regular on R_v. Moreover, R_v is D-simple since if $A \neq (0)$ were a D-ideal of R_v, then $A \cap R \neq (0)$ would be a D-ideal of R contained in P, which would be absurd.

THEOREM 3.2. *Let R be a domain with quotient field K, S a ring such that $R \subset S \subset K$ and D a derivation of R regular on S.*

Let P be a prime ideal of R such that R_P is D-simple. Then,

(i) *There is at most one prime ideal Q of S lying over P, Q being a minimal prime ideal when P is.*

(ii) *If S is the complete integral closure R' of R there is exactly one prime ideal P' of R' lying over P.*

Proof. (i) Let Q be a prime ideal of S such that $Q \cap R = P$. Being regular on S, D is also regular on S_Q, and S_Q is D-simple since $S_Q \supset R_P$. Define v: $R\backslash\{0\} \to \{\text{nonnegative integers}\}$ by $v(x) = n$ if

$$D^{(0)}(x), \cdots, D^{(n-1)}(x) \in P \quad \text{and} \quad D^{(n)}(x) \notin P ,$$

and w: $S\backslash\{0\} \to \{\text{nonnegative integers}\}$ by

$$w(y) = m \text{ if } D^{(0)}(y), \cdots, D^{(m-1)}(y) \in Q$$

and $D^{(m)}(y) \notin Q$. By 3.1, v and w extend to valuations of K; furthermore, for $x \in R$ we have $D^{(k)}(x) \in P$ if and only if $D^{(k)}(x) \in Q$ since $Q \cap R = P$; hence $v = w$, and $Q = M_V \cap S$ where M_V is the maximal ideal of the valuation ring R_v of v.

If P is a minimal prime ideal of R, suppose that Q' is a prime ideal of S such that $0 < Q' \subset Q$. We have $0 < Q' \cap R \subset Q \cap R = P$ and $Q' \cap R = P$ by the minimality of P; then $Q' = Q$ since Q is the only prime ideal of S lying over P.

(ii) By [5, p. 168] every derivation of R is regular on R'. Being a rank-1 valuation ring, R_v is completely integrally closed and contains R'. Then, $P' = M_V \cap R'$ is a prime ideal of R' lying over P; of course, by (i), P' is unique.

THEOREM 3.3. *Let R be a Noetherian $\mathscr{D}$-simple ring and $\bar{R}$ its integral closure. Let $\{P_\alpha\}_{\alpha \in A}$ be the set all the minimal prime ideals of R. Then,*

(i) *For every $\alpha \in A$, there exists $D \in \mathscr{D}$ such that R_{P_α} is D-simple, and there exists a unique prime ideal $\bar{P}_\alpha$ of $\bar{R}$ lying over P_α.*

(ii) *$\{\bar{P}_\alpha\}_{\alpha \in A}$ is the set of all the minimal prime ideals of $\bar{R}$.*

(iii) *Let $D \in \mathscr{D}$ such that $D(P_\alpha) \not\subset P_\alpha$, w_α the valuation associated by 3.1, and R_α its valuation ring. Then $R_\alpha = \bar{R}_{\bar{P}_\alpha}$ (hence, any two derivations D and D' such that $D(P_\alpha) \not\subset P_\alpha$ and $D'(P_\alpha) \not\subset P_\alpha$ give rise to the same valuation w_α).*

(iv) $\bar{R} = \bigcap_{\alpha \in A} R_\alpha$.

Proof. (i) Being $\mathscr{D}$-simple, R is a domain containing the rational numbers, and for any $\alpha \in A$, there exists $D \in \mathscr{D}$ such that $D(P_\alpha) \not\subset P_\alpha$, and by 1.3, R_{P_α} is D-simple. Then, by 3.2, there exists a unique prime ideal $\bar{P}_\alpha$ of $\bar{R}$ lying over P_α.

(ii) That every P_α is a minimal prime ideal of $\bar{R}$ is given by 3.2. Now, let $\bar{P}$ be a minimal prime ideal of $\bar{R}$, and let $P = \bar{P} \cap R$; let M be a minimal prime ideal of R contained in P; by [3, (10.8), p. 30] there exists a prime ideal $\bar{M}$ of $\bar{R}$ lying over M; since $\bar{P}$ is the only prime ideal of $\bar{R}$ lying over P, we have $\bar{M} \subset \bar{P}$ by [3, (10.9), p. 30], hence $\bar{M} = \bar{P}$, and $P = \bar{P} \cap R = M$ is a minimal prime ideal of R.

(iii) Since R is Noetherian, $\bar{R}$ is a Krull ring [3, (33.10), p. 118], and $\bar{R}_{\bar{P}_\alpha}$ is a rank-1-discrete valuation ring. As furthermore $\bar{R}_{\bar{P}_\alpha} \subset R_\alpha$ we get $\bar{R}_{\bar{P}_\alpha} = R_\alpha$.

(iv) $\bar{R}$ is a Krull ring and $\{\bar{P}_\alpha\}_{\alpha \in A}$ is the set of all the minimal prime ideals of $\bar{R}$; thus $\bar{R} = \bigcap_{\alpha \in A} \bar{R}_{\bar{P}_\alpha} = \bigcap_{\alpha \in A} R_\alpha$.

COROLLARY 3.4. *Let R be a Noetherian $\mathscr{D}$-simple ring with quotient field K. Let S be a ring such that $R \subset S \subset K$ and such that every $D \in \mathscr{D}$ is regular on S. Then, the following statements are equivalent*:

(i) *For every minimal prime ideal P of R there exists a (unique) prime ideal Q of S lying over P.*

(ii) *S is integral over R.*

(iii) *For every prime ideal M of R there exists a (unique) prime ideal N of S lying over M.*

Proof. That (ii) $\Rightarrow$ (iii) is a consequence of [3, (10.7), p. 30] and 3.2; that (iii) $\Rightarrow$ (i) is obvious. Now, let $\{P_\alpha\}_{\alpha \in A}$ be the set of the minimal prime ideals of R, $\{w_\alpha\}_{\alpha \in A}$ the associated valuations and $\{R_\alpha\}_{\alpha \in A}$ the valuation rings of the w_α's. For any $\alpha \in A$, let $D \in \mathscr{D}$ be such that $D(P_\alpha) \not\subset P_\alpha$, and let Q_α be a prime ideal of S lying over P_α; S_{Q_α} is D-simple, the valuation associated to Q_α is equal to w_α and $S \subset R_\alpha$. Hence, $S \subset \bar{R} = \bigcap_{\alpha \in A} R_\alpha$.

COROLLARY 3.5. *Let R be a Noetherian $\mathscr{D}$-simple ring with quotient field K, and $\bar{R}$ its integral closure. Then,*

(i) *$\bar{R}$ is the largest $\mathscr{D}$-simple overring of R in K having a prime ideal lying over every prime ideal of R.*

(ii) *$\bar{R}$ is the largest $\mathscr{D}$-simple overring of R in K having a prime ideal lying over every minimal prime ideal of R.*

Proof. Apply 3.4.

The author wishes to acknowledge the many helpful discussions on the topics of this paper he had with Professor Ohm.

REFERENCES

1. Y. Akizuki, *Einige Bemerkungen uber primare Integritatsbereiche mit Teilerkettensatz*, Proc. Physico-Math Soc. Japan **17** (1935), 327-336.

2. I. Kaplansky, *An Introduction to Differential Algebra*, Actualités Scientifiques et industrielles, Herman, 1957.

3. M. Nagata, *Local Rings*, John Wiley and Sons, New York, 1962.

4. E. C. Posner, *Integral closure of rings of solutions of linear differential equations*, Pacific J. Math. **12** (1962), 1417-1422.

5. A. Seidenberg, *Derivations and integral closure*, Pacific J. Math. **16** (1966), 167-173.

6. W. V. Vasconcelos, *Derivations of commutative Noetherian rings*, Math. Zeit **112** (1969), 229-233.

7. O. Zariski and P. Samuel, *Commutative Algebra*, V. 1, Van Nostrand Co., Princeton, 1958.

8. ———, *Commutative Algebra*, V. 2, Van Nostrand Co., Princeton, 1960.

Received February 9, 1970. This paper is part of the author's doctoral thesis written under the direction of Professor Jack Ohm at Louisiana State University.

LOUISIANA STATE UNIVERSITY
BATON ROUGE, LOUISIANA
AND
INSTITUTE DE MATEMATICA PURA AND APLICADA
RUA LUIS DE CAMÕES 68
RIO DE JANEIRO 58, GB, BRAZIL

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

ON NONNEGATIVE MATRICES

M. LEWIN

The following characterisation of totally indecomposable nonnegative n-square matrices is introduced: A nonnegative n-square matrix is totally indecomposable if and only if it diminishes the number of zeros of every n-dimensional nonnegative vector which is neither positive nor zero. From this characterisation it follows quite easily that:

I. The class of totally indecomposable nonnegative n-square matrices is closed with respect to matrix multiplication.

II. The $(n-1)$-st power of a matrix of that class is positive.

A very short proof of two equivalent versions of the König-Frobenius duality theorem on $(0, 1)$-matrices is supplied at the end.

A matrix is called *nonnegative* or *positive* according as all its elements are nonnegative or positive respectively. An n-square matrix A is said to be *decomposable* if there exists a permutation matrix P such that $PAP^T = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix}$, where B and D are square matrices; otherwise it is *indecomposable*. A is said to be *partly decomposable* if there exist permutation matrices P, Q such that

$$PAQ = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix}, \text{ where } B \text{ and } D \text{ are square}$$

matrices; otherwise it is *totally indecomposable*.

Whereas the notion of indecomposable matrices first appeared in 1912 in a paper by Frobenius [2] dealing with the spectral properties of nonnegative matrices, totally indecomposable matrices were introduced fairly recently apparently by Marcus and Minc [10]. Their properties have been studied in several papers on inequalities for the permanent function.

In [11] Minc gives the following characterisation of totally indecomposable matrices:

A nonnegative n-square matrix A, $n \geqq 2$, is totally indecomposable if and only if every $(n-1)$-square submatrix of A has a positive permanent.

A well-known theorem states:

THEOREM 1. *If A is an indecomposable nonnegative n-square matrix then*

$$(A + I)^{n-1} > 0 \ [3], [9] .$$

An indecomposable matrix is primitive if its characteristic value of maximum modulus is unique.

Wielandt [15] states (without proof) that for primitive n-square matrices we have

$$A^{n^2-2n+2} > 0 .$$[1]

By using solely the properties of total indecomposability we establish a different characterisation for totally indecomposable matrices from the one given by Minc. Using part of the characterisation we show that if A is a totally indecomposable nonnegative n-square matrix then $A^{n-1} > 0$. This result is best possible as for every n there exist totally indecomposable n-square matrices A for which $A^{n-2} \not> 0$. Theorem 1 then follows as a corollary of the latter result.

We should like to point out that Theorem 2 is by no means essential for the proof of Theorem 3. Two independent proofs of Theorem 3 are given in § 4. It seems justified however to present Theorem 2 on its own merit.

We conclude with a very short proof of two equivalent versions of König's theorem on matrices.

2. Preliminaries. $|S|$ denotes the number of elements of a given set S. Let M_n be the set of all nonnegative n-square matrices, let D_n be the subset of M_n of indecomposable matrices and let T_n be the subset of D_n of totally indecomposable matrices. Let $A \in M_n$ and let p and q be nonempty subsets of $N = \{1, \cdots, n\}$. Then $A[p \mid q]$, $A(p \mid q)$ is the $|p| \times |q|$ submatrix of A consisting precisely of those elements a_{ij} of A for which $i \in p$ and $j \in q$, $i \notin p$ and $j \notin q$ respectively. $A[p \mid q)$ and $A(p \mid q]$ are defined accordingly. We can now formulate equivalent definitions for matrices in D_n and T_n:

D. 1. *$A \in D_n$ if $A[p \mid N - p] \neq 0$ for every nonempty $p \subset N$.*

D. 2. *$A \in T_n$ if $A[p \mid q] \neq 0$ for any nonempty subsets p and q of N such that $|p| + |q| = n$.*

Let us now establish some connections between indecomposable and totally indecomposable matrices.

LEMMA 1. *If $A \in (D_n - T_n)$ then A has a zero on its main diagonal.*[2]

Proof. Since $A \notin T_n$ there exists a zero-submatrix $A[p \mid q]$ with $|p| + |q| = n$; but since $A \in D_n$, $p \cap q \neq \varnothing$, which means that A has

[1] A proof is supplied in [5].

[2] Lemma 1 is part of Lemma 2.3 in [1] but the shortness of our proof seems to justify its presentation.

a zero on its main diagonal.

COROLLARY 1. *If* $A \in D_n$ *then* $A + I \in T_n$.

Proof obvious.

3. **The main results.** Let $A = (a_{ij}) \in M_n$ and let v denote an n-dimensional vector with $a_i(v)$ its ith entry.

Define: $J_k = \{j: a_{kj} = 0\}$, $I_k = \{i: a_{ik} = 0\}$,

$$I_0(v) = \{i: a_i(v) = 0\}\ , \quad I_+(v) = \{i: a_i(v) > 0\}\ .$$

Let R_n denote the space of n-tuples of real numbers.

Let X_n be the set of all nonnegative vectors in R_n which are neither positive nor zero. We then have the following

THEOREM 2. *A nonnegative n-square matrix A is totally indecomposable if and only if* $|I_0(Ax)| < |I_0(x)|$ *for every* $x \in X_n$.

Proof. Let $A \in T_n$ and $x \in X_n$. A necessary and sufficient condition for $a_{i_0}(Ax) = 0$ for some i_0 is

$$(1) \qquad I_+(x) \subseteq J_{i_0}\ .$$

If $I_0(Ax) = \varnothing$, then there is nothing to prove, so we may assume

$$(2) \qquad I_0(Ax) \neq \varnothing\ .$$

$x \in X_n$ implies

$$(3) \qquad I_+(x) \neq \varnothing\ .$$

(1), (2) and (3) imply that $A[I_0(Ax) \mid I_+(x)]$ is a zero-submatrix of A. Since $A \in T_n$ by assumption, we have (by D. 2.)

$$|I_0(Ax)| + |I_+(x)| < n = |I_0(x)| + |I_+(x)|$$

and hence $|I_0(Ax)| < |I_0(x)|$ which proves the first part of the theorem. (It is not generally true however that $I_0(Ax) \subseteq I_0(x)$ as it may happen that $a_i(x) > 0$ and $a_i(Ax) = 0$, a situation which differs somewhat from that in the similar case for indecomposable matrices (5.2.2 in [9])).

Let now $A \notin T_n$. Then A contains a zero-submatrix $A[I \mid J]$ such that $I, J \neq \varnothing$ and $|I| + |J| = n$. Choose now $x \in R_n$ such that

$$(4) \qquad I_+(x) = J\ .$$

Then clearly $x \in X_n$. We have $I_0(x) = N - I_+(x) = N - J$, and hence $|I_0(x)| = |I|$. For $i \in I$ we have $J_i \supseteq J$, and hence by (4) $I_+(x) \subseteq J_i$, so that for $i \in I$ according to (1) $a_i(Ax) = 0$ and hence $I_0(Ax) \supseteq I$. Then $|I_0(Ax)| \geqq |I| = |I_0(x)|$. This completes the proof.

X_n in Theorem 2 may of course be replaced by its subset Y_n consisting of the $2^n - 2$ zero-one vectors.

Theorem 2 admits of two simple corollaries which we present as Theorems 3 and 4.

THEOREM 3. *If A is a totally indecomposable nonnegative n-square matrix then*

$$A^{n-1} > 0 .$$

Proof. If for some j_0 we had $|I_{j_0}| \geqq n - 1$ then A would be partly decomposable and hence $|I_{j_0}| \leqq n - 2$ for $j \in N$ and the rest follows.

Theorem 1 follows from Theorem 3 as an immediate consequence of Corollary 1. For $A = I + P$ where P is the n-square permutation matrix with ones in the superdiagonal, so that $a_{ij} = 1$ if $i = j$ or $i = j - 1$, $a_{n1} = 1$ and $a_{ij} = 0$ otherwise, it is easy to show that $A^{n-2} \not> 0$, which shows that our result is best possible.

THEOREM 4. *The product of any finite number of totally indecomposable nonnegative n-square matrices is totally indecomposable.*

Proof. It is clearly sufficient to prove the statement for two matrices. Let therefore $A, B \in T_n$. Choose an arbitrary element x of X_n. We then have

$$|I_0(ABx)| \leqq |I_0(Bx)| < |I_0(x)| \tag{5}$$

by Theorem 2. Since x was arbitrary, (5) applies to all elements of X_n. Again by Theorem 2 it follows that AB is totally indecomposable, which proves the theorem.

4. Independent proofs of Theorem 3. A lemma of Gantmacher [3] states that if $A \in D_n$ and $x \in X_n$, then $I_0[(A + I)x] \subset I_0(x)$.

The following proof of Theorem 3 assuming the lemma has been suggested by London[3]: Let $A \in T_n$. Using the fact that a matrix in T_n possesses a positive diagonal d, put

$$A_1 = \frac{1}{\alpha} P^T(A - \alpha P) = \frac{1}{\alpha} \quad P^T A - I \text{ where } \quad 0 < \alpha < \min a_{ij}(a_{ij} \in d)$$

[3] D. London, oral communication.

and $P = (p_{ij})$ is an n-square permutation matrix such that $p_{ij} = 1$ if and only if $a_{ij} \in d$. Then $A \in T_n$ implies $A_1 \in T_n$.

We have $A = \alpha P(A_1 + I)$; since $A_1 \in D_n$ we obtain

$$I_0(Ax) = I_0[P(A_1 + I)x] = I_0[(A_1 + I)x] \subset I_0(x) ,$$

for $x \in X_n$. Then $I_0(A^{n-1}x) = \varnothing$, and $A^{n-1} > 0$.

Another proof has been kindly suggested by the referee of this paper: We show that if A is totally indecomposable, then if $x \in X_n$, then

$$| I_0(Ax) | < | I_0(x) | .$$

The theorem then follows immediately.

Suppose $| I_0(Ay) | \geqq | I_0(y) |$ for some $y \in X_n$.

Put $| I_0(y) | = s$. There are permutation matrices P and Q such that

$$PAy = \begin{bmatrix} 0 \\ u \end{bmatrix} \quad \text{and} \quad Q^T y = \begin{bmatrix} 0 \\ v \end{bmatrix}$$

where u is an $(n-s)$-dimensional nonnegative victor and v is an $(n-s)$-dimensional positive vector: The 0's represent s zero components in each case.

We now write $PAQ = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$ where A_1 is $s \times s$, A_2 is $s \times (n-s)$, A_3 is $(n-s) \times s$ and A_4 is $(n-s) \times (n-s)$. Then $\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} 0 \\ V \end{bmatrix} = \begin{bmatrix} 0 \\ u \end{bmatrix}$ and so $A_2 V = 0$. Thus $A_2 = 0$ and hence $A \notin T_n$, a contradiction.

5. **König's Theorem.** Let A be an $m \times n$ matrix. A *covering* of A is a set of *lines* (rows or columns) containing all the positive elements of A. A covering of A is a *minimal covering* of A if there does not exist a covering of A consisting of fewer lines. Let $M(A)$ denote the number of lines in a minimal covering of A. A basis of A is a positive subdiagonal of A of maximal length. $m(A)$ denotes the length of a basis of A. The jth column of A is *essential* to A if $M(A(\varnothing J)) < M(A)$.

We now give the two versions of König's Theorem and their proofs:

K. T. 1. *If A is an $m \times n$ matrix, then $m(A) = M(A)$.*

K. T. 2. *If A is an n-square matrix, then A has k zeros on every diagonal $(k > 0)$ if and only if A contains an $s \times t$ zero-submatrix with $s + t = n + k$.*

This is a generalized version of a theorem of Frobenius. The following theorem appears in [8] (we reproduce it here in a hypothetical form).

E. T.: *If A is an $m \times n$ matrix and* K. T. I. *holds for A, then there exists a minimal covering of A* (called *essential covering*) *containing precisely the essential columns of A* (and may be some rows).

Proof of K. T. 1. $m(A) \leqq M(A)$ holds trivially. The theorem is clearly true for $1 \times n$ matrices for all n. Assume that the theorem is true for all $\mu \times n$ matrices, $\mu < m$ and all n. Let A be an $m \times n$ matrix. Consider $A' = A(\{m\} \mid N]$. A' is an $(m-1) \times n$ matrix so that K. T. 1, holds for A' and hence E. T. holds for A'. Let Q be the essential covering of A'.

Case 1. Q is a covering of A. Then $m(A) \geqq m(A') = M(A') \geqq M(A)$.

Case 2. Q is not a covering of A. Then there exists $j_0 \in N$ for which $a_{mj_0} > 0$ which is not covered by Q and hence the j_0th column is not essential to A'. Then clearly there exists a basis b' of A' without elements in the j_0th column. Then $b = b' \cup \{a_{mj_0}\}$ is a subdiagonal of A and hence $M(A) \leqq M(A') + 1 = m(A') + 1 \leqq m(A)$. This proves K. T. 1.

Proof of K. T. 2. *Necessity.* If A has k zeros on every diagonal then $m(A) \leqq n - k$. By K. T. 1, $M(A) \leqq n - k$. Apply a minimal covering to A. Then there remains an $s \times t$ zero-matrix of A which is not covered, with $s + t \geqq 2n - M(A) \geqq n + k$.

Sufficiency. Let A contain an $s \times t$ zero-submatrix with $s + t = n + k$. Then there are positive elements on at most $2n - (n + k) = n - k$ lines, meaning that there are at least k zero-rows, which proves the sufficiency.

References

1. R. A. Brualdi, S. V. Parter and H. Schneider, *The diagonal equivalence of a nonnegative matrix to a stochastic matrix*, J. Math. Anal. Appl. **16** (1966) 31-50.
2. G. Frobenius, *Über Matrizen aus nichtnegativenE lementen*, Sitzb. d. Preuss. Akad. d. Wiss, (1912), 456-477.
3. F. R. Gantmacher, *The Theory of Matrices*, vol. 2 Chelsea, New York (1959).
4. D. J. Hartfiel, *A simplified form for nearly reducible and nearly decomposable matrices*, (To appear in the Proc. Amer. Math. Soc.).
5. J. C. Holladay and R. S. Varga, *On powers of nonnegative matrices*, Proc. Amer. Math. Soc., **9** (1958), 631-634.
6. D. König, *Theorie der endlichen und unendlichen Graphen*, New York, Chelsea (1950).
7. R. Sinkhorn and P. Knopp, *Concerning nonnegative matrices and doubly stochastic matrices*, Pacific J. Math., **21** (1967), 343-348.

8. M. Lewin, *Essential coverings of matrices*, Proc. Camb. Phil. Soc., **67** (1970), 263-267.
9. M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Boston (1964).
10. M. Marcus and H. Minc, *Disjoint pairs of sets and incidence matrices*, Illinois J. Math., **7** (1963), 137-147.
11. H. Minc, *On lower bounds for permanents of* (0, 1)*-matrices*. Proc. Amer. Math. Soc. **22** (1969), 233-237.
12. H. Minc. *Nearly decomposable matrices*, (To appear).
13. H. J. Ryser, *Combinatorial Mathematics*, The Carus Mathematical Monographs (1963).
14. R. Sinkhorn, *Concerning a conjecture of Marshall Hall*. Proc. Amer. Math. Soc., **21** (1969), 197-201.
15. H. Wielandt, *Unzerlegbare nicht negative Matrizen*, Math. Z., **52** (1950), 642-648.

Received January 27, 1970.

TECHNION, ISRAEL INSTITUTE OF TECHNOLOGY
HAIFA, ISRAEL.

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

SPECIALITY OF QUADRATIC JORDAN ALGEBRAS

KEVIN MCCRIMMON

In this paper we extend to quadratic Jordan algebras certain results due to P. M. Cohn giving conditions under which a Jordan algebra is special, the most important of these being the Shirshov-Cohn Theorem that a Jordan algebra with two generators and no extreme radical is always special. We also prove that the free algebra on two generators x, y modulo polynomial relations $p(x) = 0, q(y) = 0$ is special, and by taking a particular $p(x)$ we show that most of the properties of the Peirce decomposition of a Jordan algebra relative to a supplementary family of orthogonal idempotents follow immediately from the analogous properties of Peirce decompositions in associative algebras.

Throughout we will work with algebras over an arbitrary (commutative, associative) ring of scalars Φ. A (unital) *quadratic Jordan algebra* is defined axiomatically in terms of a product $U_x y$ linear in y and quadratic in x [4, p. 1072]. We can introduce a quadratic Jordan structure $\mathfrak{A}^+$ in any unital associative algebra $\mathfrak{A}$ by taking

$$U_x y = xyx .$$

Any (Jordan) subalgebra of such an algebra $\mathfrak{A}^+$ is called a *special Jordan algebra*. A *specialization* of a quadratic Jordan algebra $\mathfrak{J}$ is a homomorphism of $\mathfrak{J}$ into an algebra of the form $\mathfrak{A}^+$.

With any quadratic Jordan algebra $\mathfrak{J}$ we can associate its *special universal envelope*, consisting of a unital associative algebra $su(\mathfrak{J})$ and a (universal) specialization $\sigma_u\colon \mathfrak{J} \to su(\mathfrak{J})^+$ such that any specialization $\sigma\colon \mathfrak{J} \to \mathfrak{A}^+$ factors uniquely through an associative homomorphism $su(\sigma)\colon su(\mathfrak{J}) \to \mathfrak{A}$,

$$\begin{tikzcd} \mathfrak{J} \arrow[rr, "\sigma"] \arrow[dr, "\sigma_u"'] & & \mathfrak{A}^+ \\ & su(\mathfrak{J}) \arrow[ur, "su(\sigma)"'] & \end{tikzcd} \tag{1}$$

$su(\mathfrak{J})$ carries a unique involution, the *main involution* π, such that the elements of $\mathfrak{J}^{\sigma_u}$ are symmetric: $x^{\sigma_u \pi} = x^{\sigma_u}$. This association is functorial—if $\varphi\colon \mathfrak{J} \to \tilde{\mathfrak{J}}$ is a homomorphism of quadratic Jordan algebras there is induced an associative homomorphism $su(\varphi)$ making

$$\begin{tikzcd} \mathfrak{J} \arrow[r, "\varphi"] \arrow[d, "\sigma_u"'] & \tilde{\mathfrak{J}} \arrow[d, "\tilde{\sigma}_u"] \\ su(\mathfrak{J}) \arrow[r, "su(\varphi)"'] & su(\tilde{\mathfrak{J}}) \end{tikzcd} \tag{2}$$

commutative. An algebra $\mathfrak{J}$ is special if and only if it is imbedded in $su(\mathfrak{J})$ via σ_u.

For any set X we have a *free quadratic Jordan algebra* $FJ(X)$, a *free special Jordan algebra* $FS(X)$, and a *free associative algebra* $F(X)$ on the set X (over the ring Φ). We have $FS(X)$ imbedded in $F(X)$ as the (Jordan) subalgebra of $F(X)^+$ generated by X, and $F(X)$ with this inclusion map serves as special universal envelope for $FS(X)$. When X consists of just two elements $X = \{x, y\}$ we know $FJ(x, y) = FS(x, y)$ by *Shirshov's Theorem*. For all these see [3].

1. Cohn's theorem and criterion. We consider a set $X = \{x_i\}_{i \in I}$ where the indices are linearly ordered. The free associative algebra $F(X)$ carries a *reversal involution*, whose action on a typical monomial is

$$(x_{i_1} \cdots x_{i_n})^* = x_{i_n} \cdots x_{i_1} .$$

The subspace $\mathfrak{H}(F(X), *)$ of $*$-symmetric elements is a Jordan subalgebra of $F(X)^+$ containing X, hence containing $FS(X)$. Cohn's Theorem measures how far $FS(X)$ is from being all of $\mathfrak{H}(F(X), *)$.

COHN'S THEOREM [1, p. 257; 2, ex. 2 p. 9]. *$\mathfrak{H}(F(X), *)$ is the Jordan subalgebra of $F(X)^+$ generated by 1, X, and all the n-tads*

$$\{x_{i_1} \cdots x_{i_n}\} = x_{i_1} \cdots x_{i_n} + x_{i_n} \cdots x_{i_1}$$

where $n \geqq 4$ and $i_1 < i_2 < \cdots < i_n$.

Proof. Clearly $\mathfrak{H} = \mathfrak{H}(F(X), *)$ contains X and all n-tads. Conversely, to show the subalgebra $\mathfrak{K}$ generated by such elements is all of $\mathfrak{H}$ we must show $\mathfrak{K}$ contains all $\{x_{i_1} \cdots x_{i_n}\} = x_{i_1} \cdots x_{i_n} + x_{i_n} \cdots x_{i_1}$ and all $x_{i_1} \cdots x_{i_n} y x_{i_n} \cdots x_{i_1}$ (where y is either 1 or one of the x_i) since these clearly span $\mathfrak{H}$. Now the $x_{i_1} \cdots x_{i_n} y x_{i_n} \cdots x_{i_1} = U_{x_{i_1}} \cdots U_{x_{i_n}} y$ are generated by X alone, so we need only generate the $\{x_{i_1} \cdots x_{i_n}\}$. We do this by induction on n. The result is trivial for $n = 2, 3$ since $\{x_{i_1} x_{i_2}\} = x_{i_1} \circ x_{i_2}$, $\{x_{i_1} x_{i_2} x_{i_3}\} = U_{x_{i_1}, x_{i_3}} x_{i_2}$ where $x \circ y$ and $U_{x,z} y$ are the linearizations of $x^2 (= U_x 1)$ and $U_x y$. We assume $n \geqq 4$ and that all $\{x_{i_1} \cdots x_{i_m}\}$ for $m < n$ are in $\mathfrak{K}$.

Our first task is to show

$$\{x_{i_{\pi(1)}} \cdots x_{i_{\pi(n)}}\} \equiv \pm \{x_{i_1} \cdots x_{i_n}\} \pmod{\mathfrak{K}} \tag{3}$$

for any permutation π. It suffices to do this for the generators $(12 \cdots n)$ and $(1n)$ of the symmetric group S_n. For the transposition $(1n)$ we have

$$\{x_{i_1} \cdots x_{i_n}\} + \{x_{i_n} x_{i_2} \cdots x_{i_{n-1}} x_{i_1}\} = U_{x_{i_1}, x_{i_n}} \{x_{i_2} \cdots x_{i_{n-1}}\} \equiv 0$$

by our induction hypothesis, and for the cycle $(12 \cdots n)$

$$\{x_{i_1} \cdots x_{i_n}\} + \{x_{i_2} \cdots x_{i_n} x_{i_1}\} = x_{i_1} \circ \{x_{i_2} \cdots x_{i_n}\} \equiv 0 \,.$$

If all the indices are distinct then (3) shows that $\{x_{i_1} \cdots x_{i_n}\}$ is congruent to $\pm$ an n-tad, which belongs to $\mathfrak{K}$ by hypothesis, so $\{x_{i_1} \cdots x_{i_n}\}$ also belongs to $\mathfrak{K}$. If two indices coincide, (3) shows $\{x_{i_1} \cdots x \cdots x \cdots x_{i_n}\} \equiv \pm\{x x_{i_1} \cdots x_{i_n} x\} = U_x\{x_{i_1} \cdots x_{i_n}\} \equiv 0$ by induction. In either case, $\{x_{i_1} \cdots x_{i_n}\} \in \mathfrak{K}$.

Since there are no n-tads for $n \geqq 4$ if there are only three variables, we have the following useful corollary.

COROLLARY. *For $m \leqq 3$, the subalgebra of $F(x_1, \cdots, x_m)^+$ generated by $x_1, \cdots, x_m$ is all of $\mathfrak{H}(F(x_1, \cdots, x_m), {}^*)$.*

The next result gives a criterion for when a homomorphic image of a special Jordan algebra is again special.

COHN'S CRITERION [1, p. 255; 2, p. 10]. *If $\mathfrak{J}$ is a special Jordan algebra and $\mathfrak{K}$ an ideal in $\mathfrak{J}$ then $\mathfrak{J}/\mathfrak{K}$ is special if and only if $\mathfrak{J} \cap \bar{\mathfrak{K}} = \mathfrak{K}$ where $\bar{\mathfrak{K}}$ is the ideal in $su(\mathfrak{J})$ generated by $\mathfrak{K}$.*

Proof. A standard functorial argument shows that the algebra $su(\mathfrak{J}/\mathfrak{K}) = su(\mathfrak{J})/\bar{\mathfrak{K}}$ and the specialization of $\mathfrak{J}/\mathfrak{K}$ induced from $\mathfrak{J} \to su(\mathfrak{J}) \to su(\mathfrak{J})/\bar{\mathfrak{K}}$ by passage to the quotient serve as special universal envelope for $\mathfrak{J}/\mathfrak{K}$ (i.e., satisfy the universal property (1)). The kernel of this specialization is $\mathfrak{J} \cap \bar{\mathfrak{K}}/\mathfrak{K}$, so the specialization is injective (i.e., $\mathfrak{J}/\mathfrak{K}$ is special) if and only if $\mathfrak{J} \cap \bar{\mathfrak{K}} = \mathfrak{K}$.

In particular, for $\mathfrak{J} = FS(X)$ and $su(\mathfrak{J}) = F(X)$ we obtain

COROLLARY. *$FS(X)/\mathfrak{K}$ is special if and only if $\bar{\mathfrak{K}} \cap FS(X) = \mathfrak{K}$ where $\bar{\mathfrak{K}}$ is the associative ideal in $F(X)$ generated by the Jordan ideal $\mathfrak{K}$ in $FS(X)$.*

2. Shirshov-Cohn theorem. The *extreme radical* of a unital quadratic Jordan algebra $\mathfrak{J}$ is the set of elements z such that $U_z = U_{z,x} = 0$ for all x in $\mathfrak{J}$; this always forms an ideal. Since $2z = z \circ 1 = 0$ for such elements, the extreme radical is always zero when $\frac{1}{2} \in \Phi$.

PROPOSITION [1, p. 260]. *If $\mathfrak{K}$ is an ideal in $FS(x, y, z)$ having a set of generators $\{k\}$ such that all tetrads $\{xyzk\}$ belong to $\mathfrak{K}$, and if $FS(x, y, z)/\mathfrak{K}$ has zero extreme radical, then $FS(x, y, z)/\mathfrak{K}$ is special.*

Proof. By the Corollary to Cohn's Criterion $FS(x, y, z)/\mathfrak{K}$ will be special if $\bar{\mathfrak{K}} \cap FS(x, y, z) \subset \mathfrak{K}$. To prove that any $p(x, y, z)$ in $\bar{\mathfrak{K}} \cap$

$FS(x, y, z)$ belongs to $\mathfrak{K}$ it will suffice to show it is in the extreme radical modulo $\mathfrak{K}$,

(i) $U_p r = prp \in \mathfrak{K}$

(ii) $U_{p,q} r = prq + qrp \in \mathfrak{K} \qquad (q, r \in FS(x, y, z))$

since we are assuming $FS(x, y, z)/\mathfrak{K}$ has no extreme radical.

It will be enough to prove the stronger results

(i)′ $prp^* \in \mathfrak{K}$

(ii)′ $p + p^* \in \mathfrak{K} \qquad (p \in \bar{\bar{\mathfrak{K}}}, r \in FS(x, y, z))$

since $p = p^*$ if $p \in \bar{\bar{\mathfrak{K}}} \cap FS(x, y, z)$ and then $prq \in \bar{\bar{\mathfrak{K}}}$ has $prq + (prq)^* = prq + qrp$.

We tackle (ii)′ first. The proof is the standard one [2, p. 11]. It suffices to consider $p = skt$ for s, t monomials in x, y, z and k a generator of $\mathfrak{K}$, since such elements span $\bar{\bar{\mathfrak{K}}}$. As $swt + t^*ws^*$ is a symmetric element of the free algebra $F(x, y, z, w)$, by Cohn's Theorem it is a sum of Jordan products of x, y, z, w and the tetred $\{xyzw\}$ where each term in the sum has a factor w or $\{xyzw\}$. But then (applying the homomorphism $F(x, y, z, w) \to F(x, y, z)$ sending $x \to x$, $y \to y$, $z \to z$, $w \to k$) we see $p + p^* = skt + t^*ks^*$ is a sum of Jordan products of x, y, z, k and the tetrad $\{xyzk\}$ where each term has a factor $k \in \mathfrak{K}$ or $\{xyzk\} \in \mathfrak{K}$ (by our hypothesis), so $p + p^*$ falls in the ideal $\mathfrak{K}$.

Since (i)′ is not linear in p we must first consider a general $p = \Sigma p_i = \Sigma s_i k_i t_i$. Here $prp^* = \Sigma_i p_i r p_i^* + \Sigma_{i<j}(p_i r p_j^* + p_j r p_i^*)$. By (ii)′ the latter sum is in $\mathfrak{K}$ since the $p_i r p_j^*$ belong to $\bar{\bar{\mathfrak{K}}}$ if p_i does, so once again we need only consider an individual p_i: to consider prp^* for $p = skt$. Now $prp^* = sktrt^*ks^* = skhks^*$ for

$$h = trt^* \in \mathfrak{H}(F(x, y, z), *) = FS(x, y, z)$$

by the Corollary to Cohn's Theorem. But since $\mathfrak{K}$ is an ideal in $FS(x, y, z)$ this yields $k' = khk = U_k h \in \mathfrak{K}$, and if $s = s_1 \cdots s_m$ where each s_i is an x, y, or z then $sk's^* = U_{s_1} \cdots U_{s_m} k' \in \mathfrak{K}$. Thus $prp^* \in \mathfrak{K}$ in all cases, finishing (i)′ and the Proposition.

Shirshov-Cohn Theorem [1, p. 261; 2, p. 48]. *Any unital quadratic Jordan algebra on two generators without extreme radical is special.*

Proof. By universal properties, any quadratic Jordan algebra $\mathfrak{J}$ on two generators is a homomorphic image of the free quadratic Jordan algebra $FJ(x, y)$ on two generators, hence (by Shirshov's Theorem) of $FS(x, y)$: $\mathfrak{J} \cong FS(x, y)/\mathfrak{K}$ for some ideal $\mathfrak{K}$. We now apply the Proposition; we can forget about tetrads, since we are not concerned with the variable z.

More precisely, let $\{k\}$ be a set of generators for $\mathfrak{K}$, let $\mathfrak{Z}$ be the ideal in $FS(x, y, z)$ generated by z, and let $\mathfrak{L}$ be the ideal generated by z together with the k's. Then $FS(x, y) \cong FS(x, y, z)/\mathfrak{Z}$ and

$$FS(x, y)/\mathfrak{K} \cong (FS(x, y, z)/\mathfrak{Z})/(\mathfrak{L}/\mathfrak{Z}) \cong FS(x, y, z)/\mathfrak{L} .$$

Each $\{xyzk(x, y)\}$ or $\{xyzz\}$ belongs to $\mathfrak{L}$—the latter is $\{xyz^2\} = U_{x,z^2}y$ and the former is a sum of Jordan products of x, y, z each term of which has a factor z, so in fact the tetrads belong to $\mathfrak{Z} \subset \mathfrak{L}$. Since $FS(x, y, z)/\mathfrak{L} \cong \mathfrak{J}$ has no extreme radical, we apply the Proposition to conclude $\mathfrak{J}$ is special.

Note that if $\frac{1}{2} \in \Phi$ then the extreme radical is automatically zero, so in that case we obtain the usual Shirshov-Cohn Theorem that *any* Jordan algebra on two generators is special. A standard example [2, ex. 3 p. 12] shows that this stronger form does not hold in general: if $\mathfrak{K}$ is the ideal spanned by $x^2, x^4, x^5, x^6 \cdots$ in the free algebra

$$FJ(x) = FS(x) = F(x)$$

on a single generator over a field Φ of characteristic 2 then the coset $\bar{x}$ in $FS(x)/\mathfrak{K}$ has $\bar{x}^2 = 0$ but $\bar{x}^3 \neq 0$ so $FS(x)/\mathfrak{K}$ cannot be special. (Of course, $\bar{x}^3$ is in the extreme radical).

An algebra $\mathfrak{J}$ is *power-associative* if each subalgebra $\Phi[z]$ generated by a single element forms an associative algebra under the natural structure induced from $\mathfrak{J}$ [5, p. 293], and *strictly power-associative* if it remains power-associative under all scalar extensions. Power-associativity amounts to the condition that a polynomial relation $p(z) = 0$ implies $zp(z) = 0$. In the previous example it was the failure of this condition which led to trouble. However, the following example shows that imposing power-associativity is not by itself enough to guarantee speciality; the condition is necessary but not sufficient.

EXAMPLE. If $\mathfrak{K}$ is the ideal in $FJ(x, y)$ over a field Φ of characteristic 2 generated by U_xy and all monomials of degree $\geqq 6$, then $\mathfrak{J} = FJ(x, y)/\mathfrak{K}$ is a strictly power-associative algebra generated by two elements which is not special.

Proof. $\mathfrak{J} = FJ(x, y)/\mathfrak{K} = FS(x, y)/\mathfrak{K}$ is not special by Cohn's Criterion since $\bar{\mathfrak{K}} \cap FS(x, y) > \mathfrak{K}$; indeed, $U_xU_yx = xyxyx = xy(U_xy)$ belongs to $\bar{\mathfrak{K}}$ and to $FS(x, y)$, yet not to $\mathfrak{K}$. To see this, recall that the ideal generated by U_xy is spanned by all $M_1 \cdots M_n(U_xy)$ and $M_1 \cdots M_n(U_{U(x)y})m$ for $m \in FS(x, y)$ and $M_i = U_x, U_y, U_{x,y}, V_x, V_y$, or I. The part of the homogeneous ideal $\mathfrak{K}$ of x-degree 3 and y-degree 2 is spanned by $U_{x,y}(U_xy)$, $V_xV_y(U_xy)$, $V_yV_x(U_xy)$, i.e., by

$$x^2yxy + yxyx^2, 2xyxyx + x^2yxy + yxyx^2, yx^2yx + xyx^2y + x^2yxy + yxyx^2 ,$$

hence by $x^2yxy + yxyx^2$ and $yx^2yx + xyx^2y$ in characteristic 2, so that $xyxyx$ is not in $\mathfrak{K}$.

We will show $\mathfrak{J}$ is power-associative; since any extension $\mathfrak{J}_\Omega$ has the same form over Ω that $\mathfrak{J}$ does over Φ, the same argument will apply to all $\mathfrak{J}_\Omega$, and consequently $\mathfrak{J}$ will be strictly power-associative. We must show that if $p(z) \in \mathfrak{K}$ for some polynomial p then also $zp(z) \in \mathfrak{K}$.

First we get rid of the constant terms. Let $z = \alpha_0 1 + w$ where w contains the homogeneous parts of z of degree $\geqq 1$. Then the degree zero part of $p(z) \in \mathfrak{K}$ is $p(\alpha_0)$, and since $\mathfrak{K}$ is homogeneous and contains only terms of degree $\geqq 3$ we have $p(\alpha_0) = 0$. Thus if $q(\lambda) = p(\lambda + \alpha_0)$ we have $q(0) = p(\alpha_0) = 0$, so q has zero constant term, and

$$p(z) = q(z - \alpha_0 1) = q(w) .$$

Therefore

$$zp(z) = \alpha_0 p(z) + wp(z) = \alpha_0 p(z) + wq(w) ,$$

and it will be enough if $wq(w)$ lies in $\mathfrak{K}$.

This shows we may assume (after replacing p, z by q, w) that $p(\lambda)$ and z have no constant term:

$$p(\lambda) = \gamma_1 \lambda + \cdots + \gamma_n \lambda^n \quad z = z_1 + \cdots + z_m$$

for z_i homogeneous of degree i. We next get rid of the degree one term $z_1 = \alpha x + \beta y$. If $\gamma_1 = \cdots = \gamma_{r-1} = 0$ but $\gamma_r \neq 0$ then the degree r term of $p(z) \in \mathfrak{K}$ is $\gamma_r z_1^r$, so by the homogeneity of $\mathfrak{K}$

$$z_1^r = (\alpha x + \beta y)^r = \alpha^r x^r + \beta^r y^r + \cdots$$

lies in $\mathfrak{K}$. Since all elements of $\mathfrak{K}$ have x-degree $\geqq 2$ and y-degree $\geqq 1$ we see $\alpha^r = \beta^r = 0$. Thus $\alpha = \beta = 0$ and $z_1 = 0$ as desired.

We are reduced to considering $z = z_2 + z_3 + z_4 + z_5$ (modulo terms of degree $\geqq 6$); in this case z^k for $k \geqq 3$ consists entirely of terms of degree $\geqq 6$, so $p(z) \equiv \gamma_1 z + \gamma_2 z^2$ and $zp(z) \equiv \gamma_1 z^2 \bmod \mathfrak{K}$. If $\gamma_1 = 0$ trivially $zp(z) \in \mathfrak{K}$, while if $\gamma_1 \neq 0$ then $\gamma_1 z + \gamma_2 z^2 \equiv \gamma_1 z_2 + \gamma_1 z_3 + (\gamma_1 z_4 + \gamma_2 z_2^2) + (\gamma_1 z_5 + \gamma_2 z_2 \circ z_3) \in \mathfrak{K}$ implies $z_2, z_3 \in \mathfrak{K}$ by homogeneity, so $\gamma_1 z^2 \equiv \gamma_1(z_2^2 + z_2 \circ z_3) \in \mathfrak{K}$. In all cases $zp(z)$ belongs to $\mathfrak{K}$, and $\mathfrak{J}$ is power-associative.

We can improve slightly on the theorem. In dealing with associative algebras $\mathfrak{A}$ with involution * in situations where $\frac{1}{2} \notin \Phi$ it is sometimes more convenient to work with certain "ample" subalgebras of $\mathfrak{H}(\mathfrak{A}, *)$ rather than just with $\mathfrak{H}(\mathfrak{A}, *)$ itself. A subspace $\mathfrak{K}$ of $\mathfrak{H}(\mathfrak{A}, *)$ is *ample* if $\mathfrak{K}$ contains 1 and all aka^* for $a \in \mathfrak{A}$ and $k \in \mathfrak{K}$. (In particular, $\mathfrak{K}$ contains all norms aa^* and traces $a + a^*$, so if $\frac{1}{2} \in \Phi$ then $\mathfrak{K} = \mathfrak{H}$). We will say a Jordan algebra is *reflexive* if $\mathfrak{J}^{\sigma_u}$ is an ample subspace of $\mathfrak{H}(su(\mathfrak{J}), \pi)$ (and *strongly reflexive* if $\mathfrak{J}^{\sigma_u} = \mathfrak{H}(su(\mathfrak{J}), \pi)$).

By the Corollary to Cohn's Theorem $\mathfrak{J} = FJ(x_1, \cdots, x_m)$ is strongly reflexive for $m \leqq 3$, but its homomorphic images may not be. However, they do inherit reflexivity:

THEOREM [2, p. 77] *If $\mathfrak{J}$ is reflexive so is any homomorphic image.*

Proof. Let $\varphi\colon \mathfrak{J} \to \tilde{\mathfrak{J}}$ be an epimorphism. To see that $\tilde{\mathfrak{J}}^{\tilde{\sigma}_u}$ is ample in $\mathfrak{H}(su(\tilde{\mathfrak{J}}), \tilde{\pi})$ we use (2) to see that (setting $\psi = su(\varphi)$) any $\tilde{a}\tilde{x}\tilde{a}^{\tilde{\pi}}$ for $\tilde{a} = \psi(a) \in su(\tilde{\mathfrak{J}}) = \psi(su(\mathfrak{J}))$, $\tilde{x} = \psi(x) \in \tilde{\mathfrak{J}}^{\tilde{\sigma}_u} = \varphi(\mathfrak{J})^{\tilde{\sigma}_u} = \psi(\mathfrak{J}^{\sigma_u})$ has the form $\psi(a)\psi(x)\psi(a)^{\pi} = \psi(axa^{\pi}) \in \psi(\mathfrak{J}^{\sigma_u}) = \tilde{\mathfrak{J}}^{\tilde{\sigma}_u}$ and hence belongs to $\tilde{\mathfrak{J}}^{\tilde{\sigma}_u}$.

COROLLARY. *Any quadratic Jordan algebra with three or fewer generators is reflexive.*

Since any algebra $\mathfrak{J}$ which is both special and reflexive has $\mathfrak{J} \cong \mathfrak{J}^{\sigma_u}$ ample in $\mathfrak{H}(su(\mathfrak{J}), \pi)$ we have the improved result

SHIRSHOV-COHN THEOREM [2, p. 77]. *Any quadratic Jordan algebra on two generators without extreme radical is isomorphic to an ample subalgebra of $\mathfrak{H}(\mathfrak{A}, {}^*)$ for some associative algebra $\mathfrak{A}$ with involution.*

Again, if $\frac{1}{2} \in \Phi$ the only ample subspace of $\mathfrak{H}(\mathfrak{A}, {}^*)$ is $\mathfrak{H}(\mathfrak{A}, {}^*)$ itself.

3. An example. In this section we consider the free special algebra $FS(x, y, z)$ on three generators, together with three relations $p(x) = 0$, $q(y) = 0$, $r(z) = 0$ where $p(\lambda)$, $q(\lambda)$, $r(\lambda)$ are monic polynomials of degree n, m, l respectively. (We allow any of these to be zero, in which case we take the degree to be ∞).

By singling out powers of x, y, z greater than or equal to n, m, l we can write any monomial in $F(x, y, z)$ uniquely as a word

$$w = a_1 w_1 a_2 w_2 \cdots w_k a_{k+1}$$

where (i) each w_α is an x^i, y^j, or z^k for $i \geqq n, j \geqq m, k \geqq l$; (ii) each a_α is a monomial containing only powers x^i, y^j, z^k for $i < n, j < m, k < l$; (iii) there is no coalescing between the w_α's and the a_α's in the sense that if $w_\alpha = x^i$ then a_α cannot end nor $a_{\alpha+1}$ begin with a factor x (similarly if w_α is y^j or z^k). Since p, q, r are monic it is easy to see (writing $i \geqq n$ as $i = \varepsilon + ne$, $j \geqq m$ as $j = \eta + mf$, $k \geqq l$ as $k = \gamma + lg$ for $0 \leqq \varepsilon < m$, $0 \leqq \eta < n$, $0 \leqq \gamma < l$ and $e, f, g \geqq 1$) that $F(x, y, z)$ has a basis consisting of the

$$m = a_1 m_1 a_2 m_2 \cdots m_k a_{k+1} \tag{4}$$

where the a_α satisfy (ii) and (iii) and the m_α are either $x^\varepsilon p(x)^e$, $y^\eta q(y)^f$, or $z^\gamma r(z)^g$. We say m_α has *weight* $\omega(m_\alpha) = e, f,$ or g and m has *weight* $\omega(m) = \Sigma\omega(m_\alpha)$.

THEOREM. *If $\mathfrak{K}$ is the (Jordan) ideal in $FS(x, y, z)$ generated by the elements $p(x)$, $xp(x)$, $q(y)$, $yq(y)$, $r(z)$, $zr(z)$ for some monic $p(\lambda)$, $q(\lambda)$, $r(\lambda)$ then $FS(x, y, z)/\mathfrak{K}$ is special.*

Proof. By the Corollary to Cohn's Criterion it suffices to show $\bar{\mathfrak{K}} \cap FS(x, y, z) \subset \mathfrak{K}$. So suppose $f(x, y, z) \in \bar{\mathfrak{K}}$ is symmetric. It is easy to see that the elements m (as in (4)) of weight $\geqq 1$ form a basis for $\bar{\mathfrak{K}}$ (they are all contained in $\bar{\mathfrak{K}}$, and they span an associative ideal containing p, xp, q, yq, r, zr which are the Jordan generators for $\mathfrak{K}$ and associative generators of $\bar{\mathfrak{K}}$). Since the reverse m^* of an element m again has the form (4), $f(x, y, z)$ is a linear combination of elements $m + m^*$ and of symmetric elements $m = m^*$.

Consider the homomorphism of the free algebra $F(x, y, z, p, q, r)$ on 6 free generators onto $F(x, y, z)$ sending $x \to x$, $y \to y$, $z \to z$, $p \to p(x)$, $q \to q(y)$, $r \to r(z)$. Each $m + m^*$ has a pre-image of the form $n + n^*$ where if m is as in (4) then $n = a_1 n_1 a_2 n_2 \cdots n_k a_{k+1}$ for a_α as before and n_α either $x^\varepsilon p^e$, $y^\eta q^f$, or $z^\gamma r^g$; such $n + n^*$ is symmetric in $F(x, y, z, p, q, r)$, hence by Cohn's Theorem a Jordan product of x, y, z, p, q, r and n-tads $\{x_{i_1} \cdots x_{i_n}\}$ for $4 \leqq n \leqq 6$, where we order the variables $x < p < y < q < z < r$. Applying the homomorphism, $m + m^*$ is a sum of Jordan products of x, y, z, $p(x)$, $q(y)$, $r(z)$ and n-tads. But all the n-tads reduce to Jordan products of x, y, z, $p(x)$, $q(y)$, $r(z)$ together with $xp(x)$, $yq(y)$, $zr(z)$—for example, the 6-tad

$$\{x\, p(x)\, y\, q(y)\, z\, r(z)\} = \{xp(x)\, yq(y)\, zr(z)\}\ .$$

Thus $m + m^*$ is a sum of Jordan products at least one factor of which is a $p(x)$, $q(y)$, $r(z)$ or $xp(x)$, $yq(y)$, $zr(z)$ (since m is of weight $\geqq 1$ and so has at least one factor $p(x)$, $q(y)$, or $r(z)$). This means that $m + m^*$ falls in the Jordan ideal $\mathfrak{K}$.

A similar but more involved argument works for the symmetric $m = m^*$. Consider the homomorphism of the free algebra on 9 generators $F(x, y, z, p, q, r, p', q', r')$ to $F(x, y, z)$ sending $x \to x$, $y \to y$, $z \to z$, $p \to p(x)$, $q \to q(y)$, $r \to r(z)$, $p' \to xp(x)$, $q' \to yq(y)$, $r' \to zr(z)$. We claim $m = m^*$ has a pre-image $n = n^*$ which is symmetric in $F(x, y, z, p, q, r, p', q', r')$. (Once we have this we argue as before; we have to worry about n-tads for $4 \leqq n \leqq 9$ now, where we order the variables $x < p < p' < y < q < q' < z < r < r'$, but again all n-tads reduce to ordinary Jordan products in $FS(x, y, z)$ since $xpp' \to xp(x)^2x$, $xp \to xp(x)$, $pp' \to p(x)xp(x)$ etc.—for example, the 7-tad $\{x\, y\, q\, q'\, z\, r\, r'\}$ reduces to $\{x\, y\, q(y)\, yq(y)\, z\, r(z)\, zr(z)\} = \{x\, yq(y)^2 y\, zr(z)^2 z\}$—and thus again $m = m^*$ falls in $\mathfrak{R}$). If $m = a_1 m_1 a_2 \cdots m_k a_{k+1} = m^* = a_{k+1}^* m_k \cdots a_2^* m_1 a_1^*$ we have $a_1 = a_{k+1}^*, a_2 = a_k^*, \cdots, a_{k+1} = a_1^*$ and $m_1 = m_k, m_2 = m_{k-1}, \cdots$ by uniqueness of the representation (4). Therefore $n = a_1 n_1 a_2 \cdots n_k a_{k+1}$ will be a symmetric pre-image of m if the n_α are symmetric pre-images of m_α. So consider $m_\alpha = x^\varepsilon p(x)^e$. Now $x^\varepsilon p^e$ is not symmetric when x, p are free variables, so we must find an alternate representation. If $\varepsilon = 2\varepsilon'$ is even then $x^\varepsilon p(x)^e = x^{\varepsilon'} p(x)^e x^{\varepsilon'}$ has the symmetric pre-image $x^{\varepsilon'} p^e x^{\varepsilon'}$, similarly if $e = 2e'$ is even then $x^\varepsilon p(x)^e = p(x)^{e'} x^\varepsilon p(x)^{e'}$ has pre-image $p^{e'} x^\varepsilon p'$, while if $\varepsilon = 2\varepsilon' + 1$ and $e = 2e' + 1$ are both odd $x^\varepsilon p(x)^e = x^{\varepsilon'} p(x)^{e'} (xp(x)) p(x)^{\varepsilon'} x^{\varepsilon'}$ has symmetric pre-image $x^{\varepsilon'} p^{e'} p' p^{\varepsilon'} x^{\varepsilon'}$ (here we need the extra free variables p', q', r'). We also note that since m is of weight $\geqq 1$, n contains at least one factor p, q, r or p', q', r'. As we said above, this is enough to allow us to complete the proof that $m = m^*$ falls in $\mathfrak{R}$.

Since $FJ(x, y) = FS(x, y)$ by Shirshov's Theorem, specializing $z \to 0$ gives

COROLLARY. *If $p(\lambda), q(\lambda)$ are monic polynomials then $FJ(x, y)/\mathfrak{R}$ is special for $\mathfrak{R}$ the ideal generated by $p(x), xp(x), q(y), yq(y)$.*

It is essential (in the general case where $\frac{1}{2} \in \Phi$) that we take $xp(x)$ and $yq(y)$ along with $p(x)$ and $q(y)$. Indeed, in our pathological one-generator example we divided out by x^2 but not x^3, and it was this x^3 that came back to haunt us. However, the Example of § 2 shows that the condition $p(z) \in \mathfrak{R} \Rightarrow zp(z) \in \mathfrak{R}$ is not by itself enough to guarantee speciality.

It is also essential that the relations involve only one variable at a time. The situation becomes much more complex when the variables are intermixed. For example, if $\mathfrak{R}$ in $FS(x, y, z)$ is generated by $x^2 - y^2$ then $FS(x, y, z)/\mathfrak{R}$ is not special, but it $\mathfrak{R}$ is generated by $U_x y - x$, $U_x y^2 - 1$ then $F/\mathfrak{R}$ is special. Thus speciality depends very much on the particular relations chosen.

4. Applications to Peirce decompositions. We define the *free Jordan algebra on X with n (supplementary, orthogonal) idempotents* $FJ(X; e_1, \cdots, e_n)$ to be the quotient $FJ(X \cup Y)/\mathfrak{R}$ where $Y = \{y_1, \cdots, y_n\}$ is disjoint from X and $\mathfrak{R}$ is the ideal generated by $1 - \Sigma y_i$, $y_i^2 - y_i$, $U_{y_i} y_j$, $y_i \circ y_j (i \neq j)$. The cosets $e_i = y_i + \mathfrak{R}$ are supplementary orthogonal idempotents in $FJ(X; e_1, \cdots, e_n) = FJ(X \cup Y)/\mathfrak{R}$, and one has the universal property that any map $X \to \mathfrak{J}$ of X into a Jordan algebra $\mathfrak{J}$ with n supplementary orthogonal idempotents $f_1, \cdots, f_n$ extends uniquely to a homomorphism $FJ(X; e_1, \cdots, e_n) \to \mathfrak{J}$ sending $e_i \to f_i$.

Consider the following properties of the Peirce decomposition of an arbitrary Jordan algebra $\mathfrak{J}$ relative to a supplementary family of orthogonal idempotents $e_1, \cdots, e_n$ [2, p. 120–1; 4, p. 1074–5].

(PD 0) $E_{ii} = U_{e_i}$ and $E_{ij} = U_{e_i, e_j} = E_{ji}$ form a supplementary family of orthogonal projections on $\mathfrak{J}$, so $\mathfrak{J} = \bigoplus \mathfrak{J}_{ij}$ for $\mathfrak{J}_{ij} = E_{ij}(\mathfrak{J}) = \mathfrak{J}_{ji}$,

and for elements x_{pq} of the Peirce spaces $\mathfrak{J}_{pq}$ and distinct indices i, j, k, l,

(PD 1) $x_{ii}^2 \in \mathfrak{J}_{ii}$, so $\mathfrak{J}_{ii}^2 \subset \mathfrak{J}_{ii}$

(PD 2) $x_{ij}^2 \in \mathfrak{J}_{ii} + \mathfrak{J}_{jj}$, so $\mathfrak{J}_{ij}^2 \subset \mathfrak{J}_{ii} + \mathfrak{J}_{jj}$

(PD 3) $x_{ii} \circ y_{ij} \in \mathfrak{J}_{ij}$, so $\mathfrak{J}_{ii} \circ \mathfrak{J}_{ij} \subset \mathfrak{J}_{ij}$

(PD 4) $x_{ij} \circ y_{jk} \in \mathfrak{J}_{ik}$, so $\mathfrak{J}_{ij} \circ \mathfrak{J}_{jk} \subset \mathfrak{J}_{ik}$

(PD 5) $x_{pq} \circ y_{rs} = 0$, so $\mathfrak{J}_{pq} \circ \mathfrak{J}_{rs} = 0$ if $\{p, q\} \cap \{r, s\} = \varnothing$

(PD 6) $U_{x_{ii}} y_{ii} \in \mathfrak{J}_{ii}$, so $U_{\mathfrak{J}_{ii}} \mathfrak{J}_{ii} \subset \mathfrak{J}_{ii}$

(PD 7) $U_{x_{ij}} y_{ii} \in \mathfrak{J}_{jj}$, so $U_{\mathfrak{J}_{ij}} \mathfrak{J}_{ii} \subset \mathfrak{J}_{jj}$

(PD 8) $U_{x_{ij}} y_{ij} = x_{ij} \circ U_{e_i}(x_{ij} \circ y_{ij}) - y_{ij} \circ U_{e_j}(x_{ij}^2)$, so $U_{\mathfrak{J}_{ij}} \mathfrak{J}_{ij} \subset \mathfrak{J}_{ij}$

(PD 9) $U_{x_{pq}} y_{rs} = 0$, so $U_{\mathfrak{J}_{pq}} \mathfrak{J}_{rs} = 0$ if $\{r, s\} \not\subset \{p, q\}$

(PD 10) $\{x_{ii} y_{ij} z_{jj}\} = (x_{ii} \circ y_{ij}) \circ z_{jj} = x_{ii} \circ (y_{ij} \circ z_{jj})$, so $\{\mathfrak{J}_{ii} \mathfrak{J}_{ij} \mathfrak{J}_{jj}\} \subset \mathfrak{J}_{ij}$

(PD 11) $\{x_{ii} y_{ij} z_{jk}\} = (x_{ii} \circ y_{ij}) \circ z_{jk} = x_{ii} \circ (y_{ij} \circ z_{jk})$, so $\{\mathfrak{J}_{ii} \mathfrak{J}_{ij} \mathfrak{J}_{jk}\} \subset \mathfrak{J}_{ik}$

(PD 12) $\{x_{ij} y_{jj} z_{jk}\} = (x_{ij} \circ y_{jj}) \circ z_{jk} = x_{ij} \circ (y_{jj} \circ z_{jk})$, so $\{\mathfrak{J}_{ij} \mathfrak{J}_{jj} \mathfrak{J}_{jk}\} \subset \mathfrak{J}_{ik}$

(PD 13) $\{x_{ij} y_{jk} z_{kl}\} = (x_{ij} \circ y_{jk}) \circ z_{kl} = x_{ij} \circ (y_{jk} \circ z_{kl})$, so $\{\mathfrak{J}_{ij} \mathfrak{J}_{jk} \mathfrak{J}_{kl}\} \subset \mathfrak{J}_{il}$

(PD 14) $\{x_{ij} y_{jk} z_{ki}\} = U_{e_i}\{(x_{ij} \circ y_{jk}) \circ z_{ki}\} = U_{e_i}\{x_{ij} \circ (y_{jk} \circ z_{ki})\}$, so $\{\mathfrak{J}_{ij} \mathfrak{J}_{jk} \mathfrak{J}_{ki}\} \subset \mathfrak{J}_{ii}$

(PD 17) $\{x_{ii} y_{ii} z_{ij}\} = x_{ii} \circ (y_{ii} \circ z_{ij})$, so $\{\mathfrak{J}_{ii} \mathfrak{J}_{ii} \mathfrak{J}_{ij}\} \subset \mathfrak{J}_{ij}$

(PD 18) $\{x_{ij} y_{ji} z_{ik}\} = x_{ij} \circ (y_{ji} \circ z_{ik})$, so $\{\mathfrak{J}_{ij} \mathfrak{J}_{ji} \mathfrak{J}_{ik}\} \subset \mathfrak{J}_{ik}$

(PD 19) $\{x_{pq} y_{rs} z_{tv}\} = 0$, so $\{\mathfrak{J}_{pq} \mathfrak{J}_{rs} \mathfrak{J}_{tv}\} = 0$ unless the indices may be linked

(PD 20) $U_{x_{ij}} e_i = U_{e_j} x_{ij}^2$

(PD 21) $e_i \circ y_{ij} = y_{ij}$, $x_{ii}^2 \circ y_{ij} = x_{ii} \circ (x_{ii} \circ y_{ij})$, $U_{x_{ii}} z_{ii} \circ y_{ij} = x_{ii} \circ (z_{ii} \circ (x_{ii} \circ y_{ij}))$ so that $V_{e_i} = I$, $V x_{ii}^2 = V_{x_{ii}}^2$, $V_{U(x_{ii}) z_{ii}} = V_{x_{ii}} V_{z_{ii}} V_{x_{ii}}$ on $\mathfrak{J}_{ij}$.

It is an easy matter to verify these for special Jordan algebras, since if $\mathfrak{A} = \Sigma_{i,j} \mathfrak{A}_{ij}$ is the Peirce decomposition of the associative algebra $\mathfrak{A}$ then $\mathfrak{J} = \Sigma_{i \leq j} \mathfrak{J}_{ij}$ for $\mathfrak{J}_{ij} = \mathfrak{A}_{ij} + \mathfrak{A}_{ji}$ is the Peirce decomposition of the Jordan algebra $\mathfrak{J} = \mathfrak{A}^+$.

We claim that if these relations hold in $\tilde{\mathfrak{J}} = FJ(\tilde{x}; \tilde{e}_1, \cdots, \tilde{e}_n)$ (taking $X = \{\tilde{x}\}$ to consist of one element) they hold in any $\mathfrak{J}$. (This is why there are two "missing" relations

(PD 15) $\{x_{ij} y_{jj} z_{ji}\} = U_{e_i}\{(x_{ij} \circ y_{jj}) \circ z_{ji}\} = U_{e_i}\{x_{ij} \circ (y_{jj} \circ z_{ji})\}$, so $\{\mathfrak{J}_{ij} \mathfrak{J}_{jj} \mathfrak{J}_{ji}\} \subset \mathfrak{J}_{ii}$

(PD 16) $\{x_{ii} y_{ij} z_{ij}\} = U_{e_i}\{(x_{ii} \circ y_{ij}) \circ z_{ji}\}$ so $\{\mathfrak{J}_{ii} \mathfrak{J}_{ij} \mathfrak{J}_{ji}\} \subset \mathfrak{J}_{ii}$;

these do not seem to follow from $\tilde{\mathfrak{J}}$, and must be verified directly).

The reason for this is that for any collection of elements x_{ij} from *distinct* Peirce spaces $\mathfrak{J}_{ij}$ there is an element $x = \Sigma x_{ij}$ having the x_{ij} as its Peirce ij-compoments; there is a homomorphism $\tilde{\mathfrak{J}} \to \mathfrak{J}$ sending $\tilde{x} \to x$ and $\tilde{e}_i \to e_i$, so the Peirce components $\tilde{x}_{ij}$ of $\tilde{x}$ map into the Peirce components x_{ij} of x. Hence any relation holding among the $\tilde{x}_{ij}$ will also hold for the x_{ij}. That is, any relation involving elements from *distinct* Peirce spaces will hold in $\mathfrak{J}$ if it holds in $\tilde{\mathfrak{J}}$. This immediately applies to (PD 1–5), (PD 7), (PD 9–14), (PD 19–20), and the first two parts of (PD 21). The same argument works for (PD 0): if $\tilde{I} = \Sigma \tilde{E}_{ij}$, $\tilde{E}_{ij}^2 = \tilde{E}_{ij}$, $\tilde{E}_{pq}\tilde{E}_{rs} = 0$ on $\tilde{x}$ then $I = \Sigma E_{ij}$, $E_{ij}^2 = E_{ij}$, $E_{pq}E_{rs} = 0$ an any x, so the E_{ij} are supplementary orthogonal idempotents).

The remaining formulas can be derived from the previous ones by various stratagems. For (PD 17–18) we use the relation

$$\{abb\} = a \circ b^2 \quad \{abc\} + \{acb\} = a \circ (b \circ c)$$

valid in any Jordan algebra. In (PD 18) $\{x_{ij}y_{ji}z_{ik}\} = x_{ij} \circ (y_{ji} \circ z_{ik}) - \{x_{ij}z_{ik}y_{ji}\} = x_{ij} \circ (y_{ji} \circ z_{ik})$ since $U_{\mathfrak{J}_{ij}}\mathfrak{J}_{ik} = 0$ by (PD 9), and similarly in (PD 17) since $U_{\mathfrak{J}_{ii}}\mathfrak{J}_{ij} = 0$. (This argument also shows either one of (PD 15), (PD 16) implies the other).

For (PD 6), (PD 8), and the last part of (PD 21) we use

$$\partial_y\{x^3\}|_x = U_x y + U_{x,y}x = U_x y + \{xxy\} = U_x y + x^2 \circ y \,.$$

Now the relations

(PD 6)′ $U_{x_{ii}}x_{ii} \in \mathfrak{J}_{ii}$

(PD 8)′ $U_{x_{ij}}x_{ij} = x_{ij} \circ U_{e_i}(x_{ij}^2)$

(PD 21)′ $V_{U(x_{ii})x_{ii}} = V_{x_{ii}}^3$ on $\mathfrak{J}_{ij}$

will be inherited from $\tilde{\mathfrak{J}}$, and this remains true over any scalar extension Ω of Φ, so we can linearize to get

$$\begin{aligned}
&U_{x_{ii}}y_{ii} + x_{ii}^2 \circ y_{ii} \in \mathfrak{J}_{ii}\\
&U_{x_{ij}}y_{ij} + x_{ij}^2 \circ y_{ij} = y_{ij} \circ U_{e_i}(x_{ij}^2) + x_{ij} \circ U_{e_i}(x_{ij} \circ y_{ij})\\
&V_{U(x_{ii})z_{ii}} + V_{x_{ii}^2 \circ z_{ii}} = V_{x_{ii}}V_{z_{ii}}V_{x_{ii}} + V_{x_{ii}}^2 V_{z_{ii}} + V_{z_{ii}}V_{x_{ii}}^2 \,.
\end{aligned}$$

The first of these implies (PD 6) via (PD 1), the second implies (PD 8) via (PD 2), and the third implies (PD 21) since we already know $V_{x_{ii}^2} = V_{x_{ii}}^2$ and so $V_{x_{ii} \circ y_{ii}} = V_{x_{ii}}V_{y_{ii}} + V_{y_{ii}}V_{x_{ii}}$.

Thus the task of verifying Peirce relations for an arbitrary Jordan algebra $\mathfrak{J}$ reduces to verifying them for the free Jordan algebra $\tilde{\mathfrak{J}}$ on one generator with idempotents. The whole point of this reduction is that $\tilde{\mathfrak{J}}$ is *special,* and we already remarked that the relations were easily verified in any special algebra.

THEOREM. *The free Jordan algebra $FJ(x; e_1, \cdots, e_n)$ on one generator with n supplementary orthogonal idempotents is special.*

To show $FJ(x; e_1, \cdots, e_n) = FJ_\Phi(x; e_1, \cdots, e_n)$ is special it will be enough if it is imbedded in a special algebra $FJ_\Phi(x; e_1, \cdots, e_n)_\Omega = FJ_\Omega(x; e_1, \cdots, e_n)$. We choose Ω as follows. Consider the polynomial ring $\Phi[\lambda_1, \cdots, \lambda_n]$. The element $\mu = \prod_{i<j} (\lambda_i - \lambda_j)$ is homogeneous in the λ's and the coefficient of $\lambda_1^{n-1}\lambda_2^{n-2} \cdots \lambda_{n-1}^1$ in μ is 1, so μ is not a zero divisor in $\Phi[\lambda_1, \cdots, \lambda_n]$. This guarantees Φ is imbedded in $\Omega = \Phi[\lambda_1, \cdots, \lambda_n][1/\mu]$; the important thing about Ω is that each $\lambda_i - \lambda_j$ is invertible in Ω. Since μ is not a zero-divisor in

$$FJ_\Phi(X; e_1, \cdots, e_n) \otimes \Phi[\lambda_1, \cdots, \lambda_n] ,$$

$FJ_\Phi(X; e_1, \cdots, e_n)$ is imbedded in $FJ_\Phi(X; e_1, \cdots, e_n)_\Omega = FJ_\Omega(X; e_1, \cdots, e_n)$.

PROPOSITION. *For any X, $FJ_\Omega(X; e_1, \cdots, e_n) \cong FJ_\Omega(X, y)/\mathfrak{R}$ where $\mathfrak{R}$ is the ideal generated by $p(y) = \prod (y - \lambda_i 1)$ and $yp(y)$.*

Proof. Consider the polynomials $p(\lambda) = \prod (\lambda - \lambda_i)$ and $p_i(\lambda) = \prod_{j\neq i} (\lambda - \lambda_j)/\prod_{j\neq i} (\lambda_i - \lambda_j)$ in Ω. We have $p_i(\lambda_i) = 1$, $p_i(\lambda_j) = 0$ if $j \neq i$. Therefore $1 - \sum p_i(\lambda)$ is of degree $\leqq n - 1$ yet has n roots $\lambda_1, \cdots, \lambda_n$, so it must be identically zero, and similarly for $\lambda = \sum \lambda_i p_i(\lambda)$:

$$\sum p_i(\lambda) = 1, \sum \lambda_i p_i(\lambda) = \lambda .$$

(We always assume $n > 1$ since for $n = 1$ $FJ(X; e_1) = FJ(X; 1) = FJ(X)$ has only the trivial idempotent $e_1 = 1$). Also

$$U_{p_i(\lambda)} p_j(\lambda) = p_i(\lambda)^2 p_j(\lambda),\ p_i(\lambda) \circ p_j(\lambda) = 2p_i(\lambda)p_j(\lambda),$$
$$p_i(\lambda)^2 - p_i(\lambda) = p_i(\lambda)^2 - \sum p_i(\lambda)p_j(\lambda) = \sum_{j\neq i} p_i(\lambda)p_j(\lambda)$$

are all divisible by $p(\lambda)$ and belong to the (Jordan) ideal generated by $p(\lambda)$ and $\lambda p(\lambda)$.

These conditions imply that the elements $\tilde{e}_i = p_i(y)$ in $FJ_\Omega(X, y)$ satisfy $\sum \tilde{e}_i = 1$, $\sum \lambda_i \tilde{e}_i = y$, $U_{\tilde{e}_i} \tilde{e}_j \in \mathfrak{R}$, $\tilde{e}_i \circ \tilde{e}_j \in \mathfrak{R}$, $\tilde{e}_i^2 - \tilde{e}_i \in \mathfrak{R}$, so the cosets $e_i = \tilde{e}_i + \mathfrak{R}$ in $FJ_\Omega(X, y)/\mathfrak{R}$ form a supplementary family of orthogonal idempotents. (Note $p_i(y)$ is defined since we are allowed to divide by $\lambda_i - \lambda_j$ in Ω). We show $FJ_\Omega(X, y)/\mathfrak{R}$ is isomorphic to $FJ_\Omega(X; e_1, \cdots, e_n)$ by showing it has the universal property of the latter. Given any map φ of X into a Jordan algebra $\mathfrak{J}$ with idempotents $f_1, \cdots, f_n$ we have a homomorphism $FJ_\Omega(X, y) \to \mathfrak{J}$ sending $x \to \varphi(x)$, $y \to \sum \lambda_j f_j$. Then $\tilde{e}_i = p_i(y)$ is mapped into

$$p_i(\sum \lambda_j f_j) = \sum p_i(\lambda_j) f_j = f_i ,$$

$p(y)$ into $p(\sum \lambda_j f_j) = \sum p(\lambda_j) f_j = 0$, and $yp(y)$ into $\sum \lambda_j p(\lambda_j) f_j = 0$. Since $p(y)$ and $yp(y)$ generate $\mathfrak{R}$ we have an induced homomorphism

$$FJ_\Omega(X, y)/\mathfrak{R} \longrightarrow \mathfrak{J}$$

sending $e_i \to f_i$. The uniqueness follows since $FJ_\Omega(X, y)/\mathfrak{K}$ is generated over Ω by X and the e_i (because $\sum \lambda_i e_i = y$).

Applying the Proposition when $X = \{x\}$, we have

$$FJ_\Omega(x; e_1, \cdots, e_n) \cong FJ_\Omega(x, y)/\mathfrak{K}$$

where $\mathfrak{K}$ is generated by $p(y)$ and $yp(y)$. By the Corollary to the Theorem of the previous Section (with $q(\lambda) = 0$), $FJ_\Omega(x, y)/\mathfrak{K}$ is special. Therefore $FJ(x; e_1, \cdots, e_n) \subset FJ_\Omega(x; e_1, \cdots, e_n)$ is special too, completing the proof of the theorem.

The algebra $FJ(x, y; e_1, \cdots, e_n)$ on two generators is no longer special, since it has the exceptional algebra $\mathfrak{H}(\mathfrak{C}_3)$ as a homomorphic image ($\mathfrak{C}$ a Cayley algebra); indeed, the exceptional algebra can be generated by two elements x, y and the idempotents e_1, e_2, e_3 [2, ex. 1 p. 51].

REFERENCES

1. P. M. Cohn, *On homomorphic images of special Jordan algebras*, Canadian J. Math., **6** (1954), 253-264.

2. N. Jacobson, *Structure and Representations of Jordan Algebras*, Amer. Math. Soc. Colloq. Publ., vol. **39**, Providence, 1968.

3. R. Leward and K. McCrimmon, *Macdonald's theorem for quadratic Jordan algebras*, Pacific J. Math., **35** (1970), 681-707.

4. K. McCrimmon, *A general theory of Jordan rings*, Proc. Nat. Acad. Sci. U.S.A. **56** (1966), 1072-1079.

5. ———, *The Freudenthal-Springer-Tits constructions revisited*, Trans. Amer. Math. Soc., **148** (1970), 293-314.

Received June 30, 1970.

UNIVERSITY OF VIRGINIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

SINGULAR PERTURBATIONS OF DIFFERENTIAL EQUATIONS IN ABSTRACT SPACES

HUSSAIN S. NUR

In a recent paper, Kisynski studied the solutions of the abstract Cauchy problem $\varepsilon x^{\cdot\cdot}(t) + x^{\cdot}(t) + Ax(t) = 0$, $x(0) = x_0$ and $x^{\cdot}(0) = x_1$ where $0 \leqq t \leqq T$, $\varepsilon > 0$ is small parameter and A is a nonnegative self-adjoint operator in a Hilbert space H. With the aid of the functional calculus of the operator A, he has showed that as $\varepsilon \to 0$ the solution of this problem converges to the solution of the unperturbed Cauchy problem $x^{\cdot}(t) + Ax(t) = 0$, $x(0) = x_0$. Smoller has proved the same result for equation of higher order.

The purpose of this paper is to study the solution of a similar problem and allowing the operator A to depend on t.

To be precise, we shall show that if the initial data is taken from a suitable dense subset of H, then the solution of the Cauchy problem:

$$\varepsilon x^{\cdot\cdot}(t) + x^{\cdot}(t) + A(t)x(t) = 0, x(0) = x_0, x^{\cdot}(0) = x_1 \tag{1.1}$$

converges to the solution of the unperturbed Cauchy problem

$$x^{\cdot}(t) + A(t)x(t) = 0, x(0) = x_0 \tag{1.2}$$

as $\varepsilon \to 0$ where $0 \leqq t \leqq T$, $\varepsilon > 0$ is a small parameter, $A(t)$ is a continuous semi-group of nonnegative self-adjoint operators in H with infinitesimal generator A.

2. **The problem (1.1) when $H = R_1$.** Before considering (1.1) in the general case, it is necessary to consider (1.1) in the case when $H = R_1$ (i.e., the real line). Thus we consider the Cauchy problem:

$$\varepsilon u^{\cdot\cdot}(t) + u^{\cdot}(t) + e^{\mu t}u(t) = 0 \; . \quad u(0) = x_0,\ u^{\cdot}(0) = x_1 \tag{2.1}$$

when $t \geqq 0$, $\mu \geqq 0$. $\varepsilon > 0$.

According to theorem (1) in [2], equation (2.1) has two linearly independent solutions:

$$u_1 = \sum_0^{m-1} u_{1j}(t)\varepsilon^j + \varepsilon^m E_0 \; , \qquad u_1^{\cdot} = \sum_0^{m-1} u_{1j}^{\cdot}(t)\varepsilon^j + \varepsilon^{m-1}E_1$$

$$u_2 = \sum_0^{m-1} u_{2j}(t)\varepsilon^j e^{-t/\varepsilon} + \varepsilon^m E_0 \; , \quad u_2^{\cdot} = \sum_0^{m-1} (d/dt)[u_{2j}(t)e^{-t\varepsilon}]\varepsilon^j + \varepsilon^{m-1}E_1$$

where $u_{ij}(t)$ $(i = 1, 2)$ are C^∞ functions on $[0, T]$ and $u_{i0}(t)$ $(i = 1, 2)$ does not vanish at any point of $[0, T]$ and E_0, E_1 are functions of ε and others, but bounded for small $\varepsilon \geqq 0$.

Hence the general solution of equation (2.1) is $u = c_1 u_1 + c_2 u_2$. Solving for c_1 and c_2 by using the initial condition we obtain $u = x_0 s_{00} + x_1 s_{01}$ and $u^\cdot = x_0 s_{10} + x_1 s_{11}$ where

$$
\begin{aligned}
s_{00} &= H^{-1}(\varepsilon)[u_2^\cdot(0)u_1(t) - u_1^\cdot(0)u_2(t)] \\
s_{01} &= H^{-1}(\varepsilon)[u_1(0)u_2(t) - u_2(0)u_1(t)] \\
s_{10} &= s_{00}^\cdot = \frac{d}{dt}s_{00} \\
s_{11} &= s_{01}^\cdot = \frac{d}{dt}s_{01}
\end{aligned}
\tag{2.3}
$$

and

$$H(\varepsilon) = u_1(0)u_2^\cdot(0) - u_2(0)u_1^\cdot(0)$$

How taking the limit as $\varepsilon \to 0$, we find that

$$
\begin{aligned}
s_{00}(t, \varepsilon, \mu) &\longrightarrow x_0 u_{10}(t) \\
s_{01}(t, \varepsilon, \mu) &\longrightarrow 0 \,.
\end{aligned}
\tag{2.4}
$$

Consequently, $u(t, \varepsilon) \to x_0 u_{10}(t)$. From equation 15 in [2] we find that $u_{10}(t)$ is the solution of the equation

$$u^\cdot + e^{\mu t}u = 0 \tag{2.5}$$

and this is what we wished to show.

3. Estimates for the Functions $s_{ij}(t, \varepsilon, \mu)$. In this section we would like to find estimates for the functions $s_{ij}(t, \varepsilon, \mu)$ $(i, j = 0, 1)$. We may do so by solving for $u_{ij}(t)$ $(i = 1, 2; j = 0, 1, \cdots, m - 1)$ from equation 15 in [2]. Since this would be rather tedious we will take the simpler approach of estimating $u_i(t, \varepsilon, \mu)$ and $u_i^\cdot(t, \varepsilon, \mu)$ $(i = 1, 2)$. Multiplying (2.1) by $u^\cdot$ and integrating between 0 and t we obtain:

$$\frac{\varepsilon u^{\cdot 2}}{2} + \int_0^t u^{\cdot 2} + \frac{u^2}{2}e^{\mu t} - \frac{1}{2}\mu\int_0^t u^2 e^{\mu t} = c \,.$$

Consequently

$$u^2 \leqq 2\,|\,c\,| + \mu\int_0^t u^2 e^{\mu t}dt \,.$$

Now using Bellman's lemma, we obtain

$$u^2 \leqq 2/c/e^{e^{\mu t}} \tag{3.1}$$

For estimating $u^{\cdot}(t)$, we multiply equation (2.1) by $e^{-\mu t}u^{\cdot}$, integrating between 0 and t and using Bellman's lemma we obtain:

$$u^{\cdot 2}(t) \leqq 2\varepsilon^{-1}/c/e^{2\mu t} \tag{3.2}$$

In [2] page 323 we proved that for all small $\varepsilon \geqq 0$ $H(\varepsilon) \neq 0$, therefore we see that (2.3), (3.1), and (3.2) yield,

$$|s_{00}| \leqq K(\varepsilon) \exp\left(\frac{e^{\mu t}}{2}\right) \tag{3.3}$$

$K(\varepsilon)$ is a bounded function in ε, and

$$|s_{01}| \leqq \bar{K}(\varepsilon) \exp\left(e^{e^{\mu t}/2}\right) \tag{3.4}$$

$\bar{K}(\varepsilon)$ is a bounded function in ε.

To obtain an estimate for s_{ij} $(i, j = 1, 2)$ we write equation (2.1) in amatrix form as:

$$U^{\cdot} = AU$$

when

$$A = \begin{pmatrix} 0 & 1 \\ -\bar{\varepsilon}^1 \exp(\mu t) & -\bar{\varepsilon}^1 \end{pmatrix}.$$

Hence

$$U = \exp\left[\int A(s)ds\right] = \begin{pmatrix} s_{00} & s_{01} \\ s_{10} & s_{11} \end{pmatrix}$$

and from the equation

$$\begin{aligned}(d/dt)\begin{pmatrix} s_{00} & s_{01} \\ s_{10} & s_{11} \end{pmatrix} &= \begin{pmatrix} s_{00} & s_{01} \\ s_{10} & s_{11} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -\bar{\varepsilon}^1 \exp(\mu t) & -\bar{\varepsilon}^1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -\bar{\varepsilon}^1 \exp(\mu t) & -\bar{\varepsilon}^1 \end{pmatrix}\begin{pmatrix} s_{00} & s_{01} \\ s_{10} & s_{11} \end{pmatrix}\end{aligned} \tag{3.5}$$

we obtain

$$s_{10} = -s_{01}\varepsilon^{-1} \exp(\mu t) \tag{3.6}$$

$$s_{11} = s_{00} - \varepsilon^{-1}s_{01}. \tag{3.7}$$

4. The problem (1.1) in abstract Hilbert space. We shall now consider the problem (1.1) in any Hilbert space H with norm $\|\cdot\|$.

Since $\{A(t)\}$ is a semi-group of a nonnegative selfadjoint operator in H, with infinitesimal generator A, there is a resolution of the identity E_μ such that $A(t)$ has the spectral representation:

$$A(t) = \int_0^\infty e^{\mu t} dE_\mu \,.$$

We shall next use the functional calculus of the operator $A(t)$. For fixed $\varepsilon > 0$, $t \geqq 0$, we define the operator S_{ij} on H by

$$S_{ij}(t, \varepsilon) = \int_0^\infty s_{ij}(t, \varepsilon, \mu) dE_\mu \qquad (i, j = 0, 1) \tag{4.1}$$

where the $s_{ij}(t, \varepsilon, \mu)$ are defined by (2.3). If we let D denote the dense domain of the operator $e^{A^2(t)}$ for all t, then our estimates (3.2) through (3.7) imply that D is contained in the domain of $S_{ij}(t, \varepsilon)$ for every $i, j = 0, 1$.

For x_0 and x_1 in D, we write

$$x_\varepsilon(t) = S_{00}(t, \varepsilon)x_0 + S_{01}(t, \varepsilon)x_1 \tag{4.2}$$

and we see that $x_\varepsilon(t)$ is in the domain of $A(t)$ for every $\varepsilon > 0$. We now state the main theorem.

THEOREM. *Let $x_\varepsilon(t)$ be defined as in* (4.2) *when x_0, x_1 are in D. Then $x_\varepsilon(t)$ is the unique solution of the Cauchy problem* (1.1) *and $x_\varepsilon(t)$ converges to the solution of* (1.2) *as $\varepsilon \to 0$.*

To prove this theorem we first prove the following lemmas:

LEMMA 1. *For $x \in D$, $(d/dt)S_{ij}(t, \varepsilon)x$ exists and*

$$(d/dt)S_{ij}(t, \varepsilon)x = \int_0^\infty (d/dt)s_{ij}(t, \varepsilon, \mu)dE_\mu x \qquad (i, j = 0, 1) \,. \tag{4.3}$$

Proof. We shall prove the lemma for $i = j = 0$. Since the proofs for the other cases are similar, they will be omitted. For $x \in D$ and $t \geqq 0$ fixed, we have:

$$\begin{aligned} &\left\| \frac{S_{00}(t + \Delta t, \varepsilon) - S_{00}(t)}{\Delta t} \times - S_{10}(t, \varepsilon)x \right\|^2 \\ &= \int_0^\infty \left[\frac{s_{00}(t + \Delta t, \varepsilon, \mu) - s_{00}(t, \varepsilon, \mu)}{\Delta t} - s_{10}(t, \varepsilon, \mu) \right]^2 d \,||\, E_\mu x \,||^2 \\ &= \int_0^\infty [s_{10}(t', \varepsilon, \mu) - s_{10}(t. \varepsilon. \mu)]^2 d \,||\, E_\mu x \,||^2 \,, \end{aligned}$$

where $t \leqq t' \leqq t + \Delta t$, using the theorem of the mean and (2.3).

Now there is a T such that $t + \Delta t \leqq T$ for all Δt sufficiently small, so that if we use (3.3) through (3.7) we see that

$$\begin{aligned}|s_{10}(t', \varepsilon, \mu) - s_{10}(t, \varepsilon, \mu)| &\leqq |s_{10}(t', \varepsilon, \mu)| + |s_{10}(t, \varepsilon, \mu)| \\ &\leqq \varepsilon^{-1} e^{\mu T} K(\varepsilon) e^{(1/2) e^{\mu T}} \leqq N(\varepsilon, T) e^{e^{\mu T}}\end{aligned}$$

where $N(\varepsilon, T)$ is a constant depending on T and ε only. Therefore the function $|s_{10}(t', \varepsilon, \mu) - s_{10}(t, \varepsilon, \mu)|^2$ is summable with respect to the measure $d\,||E_\mu x||^2$ if Δt is sufficiently small. Furthermore,

$$\lim_{\Delta t \to 0} [s_{10}(t', \varepsilon, \mu) - s_{11}(t, \varepsilon, \mu)]^2 = 0 .$$

So that the Lebseque dominated convergence theorem yields:

$$\lim_{\Delta t \to 0} \int_0^\infty [s_{10}(t', \varepsilon, \mu) - s_{10}(t, \varepsilon, \mu)]^2 d\,||E_\mu x||^2 = 0 .$$

This completes the proof of the lemma.

LEMMA 2. *For $x \in D$ and $t \geqq 0$, we have*

$$\lim_{\varepsilon \to 0} \left\| S_{00}(t, \varepsilon)x - \exp\left(-\int A(s)ds\right)x \right\| = 0 \tag{4.4}$$

$$\lim_{\varepsilon \to 0} ||S_{01}(t, \varepsilon)x|| = 0 . \tag{4.5}$$

Proof.

$$\begin{aligned}&\left\| S_{00}(t, \varepsilon)x - \exp\left(-\int A(s)ds\right) \times \right\|^2 \\ &= \int_0^\infty \left| \left(s_{00}(t, \varepsilon, \mu) - \exp\left(-\int^t e^{\mu s} ds\right)\right) \right|^2 d\,||E_\mu x||^2 .\end{aligned}$$

From (3.3) we see that $\left[s_{00}(t, \varepsilon, \mu) - \exp\left(-\int^t e^{\mu s} ds\right)\right]^2$ is summable with respect to the measure $d\,||E_\mu x||^2$ and, as we have seen in (2.4) and (2.5), the integrand converges pointwise to zero. We apply the Lebesgue dominated convergence theorem to conclude that the integral likewise converges to zero as $\varepsilon \to 0$. This proves (4.4). Relation (4.5) follows from (2.4) and (2.5) likewise.

LEMMA 3. *Let B be a bounded operator in H. If $x^\cdot(t) + Bx(t) = 0$, $0 \leqq t \leqq 0$, and $x(0) = 0$, then $x(t) \equiv 0$.*

The proof of the above lemma is in [3] and therefore will be omitted.

The proof of the theorem. That $x_\varepsilon(t)$ defined by (4.2) is a solution of (1.1) follows at once from Lemma 1 by direct verification. The uniqueness of $x_\varepsilon(t)$ follows from Lemma 3 just as in [1]. Finally, since $\exp\left(-\int^t A(s)ds\right)x_0$ is the solution of (1.2) Lemma 2 shows that.

$$\lim_{\varepsilon\to 0}\left\| x_\varepsilon(t) - \exp\left(-\int^t A(s)ds\right)x_0 \right\| = 0 .$$

This completes the proof of the theorem.

References

1. J. Kisynski, *Sur les Equations Hyperboliques avec Petit Parametre*, Colloq. Math., **10** (1963), 331–343.

2. Hussain S. Nur, *Singular perturbation of linear partial differential equations*, J. Differential Equations **6** (1969), 1–12.

3. J. A. Smoller, *Singular perturbation and a theorem of Kisynski*, J. Math. Anal. Appl., **12**, No. 1, (1965).

Received December 2, 1969.

Fresno State College

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

A NON-COMPACT KREIN-MILMAN THEOREM

D. K. OATES

This paper describes a class of closed bounded convex sets which are the closed convex hulls of their extreme points. It includes all compact ones and those with the positive binary intersection property.

Let K be a closed bounded convex subset of a Hausdorff locally convex linear topological space F. Denote by EK the extreme points of K, by $\mathrm{co}\, EK$ their convex hull and let $\overline{\mathrm{co}}\, EK$ be its closure. We are interested in showing when

$$K = \overline{\mathrm{co}}\, EK\ .$$

The principal known results are the following:

THEOREM 1.1. *If either*
(a) *K is compact;*
or (b) *K has the positive binary intersection property;*
then $K = \overline{\mathrm{co}}\, EK$.

Case (a) is the Krein-Milman Theorem [3, p. 131]. Case (b) was proved by Nachbin in [6], and he poses in [5, p. 346] the problem of obtaining a theorem of which both (a) and (b) are specializations. This is answered by Theorem 4.2. For the whole of this paper, S is a Stonean (extremally disconnected compact Hausdorff) space.[1]

A simplified version of Theorem 4.2 reads as follows:

THEOREM 1.2. *Let X be a normed linear space. Then any norm-closed ball in the space $\mathfrak{B}(X, C(S))$ of continuous linear operators from X to $C(S)$ is the closure of the convex hull of its extreme points in the strong neighborhood topology.*

The result concerning the unit ball of a dual Banach space in its weak*-topology and that concerning the unit ball in $C(S)$ in its norm topology are special cases of Theorem 1.2.

A *sublinear function* P from a vector space X to a partially ordered space V satisfies

$$P(x + y) \leqq P(x) + P(y)$$

and

[1] Theorem 2.3 and its proof are valid when S is zero-dimensional.

$$P(tx) = tP(x)$$

for all x, y in X and $t \geqq 0$.

A linear operator T from X to V is *dominated by* P if $Tx \leqq Px$ for all x in X. The set of all linear operators from X to V dominated by P will be written $L(P)$.

2. Let P be a sublinear function into $C(S)$, where S is Stonean. We obtain a compact approximation to $L(P)$ by considering a finite partition $\mathscr{U} = \{U_1, \cdots, U_M\}$ of S into disjoint open-and-closed sets. Let $C(S_{\mathscr{U}})$ denote the set of all function in $C(S)$ whose restrictions $f \mid U_k$ are constant. The constant values will be written as $f(U_k)$.

LEMMA 2.1. *Let P be a sublinear function from a vector space X to $C(S_{\mathscr{U}})$ and let $L(P_{\mathscr{U}})$ be the set of all linear operators from X to $C(S_{\mathscr{U}})$ dominated by P. Then*

$$EL(P_{\mathscr{U}}) \subseteqq EL(P) .$$

Proof. Suppose $T \in EL(P_{\mathscr{U}})$. For $k = 1, \cdots, M$ let t_k be chosen arbitrarily in U_k. If $G, H \in L(P)$ and $T = 1/2(G+H)$ define $G', H' \in L(P_{\mathscr{U}})$ by

$$G'x = \sum_{k=1}^{M} (Gx)(t_k)\chi_k \qquad H'x = \sum_{k=1}^{M} Hx(t_k)\chi_k$$

where χ_k is the characteristic function of U_k. Since $1/2(G' + H') = T$ and $T \in EL(P_{\mathscr{U}})$, we have $G' = H' = T$. Hence, for each $x \in X$ and $k = 1, \cdots, M$,

$$G'x(U_k) = H'x(U_k) = Tx(U_k)$$

so that

$$Gx(t_k) = Hx(t_k) = Tx(t_k) .$$

Since t_k was chosen arbitrarily in U_k, $G = H = T$. Hence $T \in EL(P)$.

DEFINITION 2.2. Let X and E be linear topological spaces and let $\mathfrak{B}(X, E)$ be the space of all continuous linear operators from X to E. The *strong neighborhood topology* for $\mathfrak{B}(X, E)$ is the topology with a base given by sets of the form

$$N(T;\ x_1, \cdots, x_n;\ U) = \{S \in \mathfrak{B}(X, E) :\ (T-S)x_i \in U, i = 1, \cdots, n\}$$

where $x_1, \cdots, x_n \in X$ and U is a neighborhood of 0 in E.

If E is normed, then we write

$N(T;\ x_1, \cdots, x_n;\ \varepsilon)$ for $N(T;\ x_1, \cdots, x_n;\ U)$ when U is the open ε-ball about 0.

THEOREM 2.3. *Let $\mathscr{W}$ be a finite partition of S into nonempty open-and-closed subsets. Let P be a sublinear function from a linear space X into $C(S_{\mathscr{W}})$. Then $L(P) = \overline{\text{co}}\ EL(P)$, with the closure in the strong neighborhood topology of $\mathfrak{B}(X, C(S))$.*

Proof. Let $\mathscr{U}$ be any finite partition of S into nonempty open-and-closed sets. From Lemma 2.1, $\overline{\text{co}}\ EL(P) \supseteqq \overline{\text{co}}\ EL(P_{\mathscr{U}})$. Now $L(P_{\mathscr{U}})$ can be linearly identified with a certain compact convex subset of a finite product $X^* \times \cdots \times X^*$, where X^* is the algebraic dual of X with the topology $w(X^*, X)$. Hence, from the Krein-Milman Theorem, $\overline{\text{co}}\ EL(P_{\mathscr{U}}) = L(P_{\mathscr{U}})$.

Let $T \in L(P)$ and let $N(T;\ x_1, \cdots, x_n;\ \varepsilon)$ be a strong neighborhood of T. The functions $\{Tx_i\colon\ i = 1, \cdots, n\}$ are continuous so for each fixed i there is a finite covering

$$\mathscr{V}^{(i)} = \{V_1^i \cdots,\ V_{N_i}^i\}$$

of S by open sets such that

$$\operatorname{Var}(Tx_i,\ V_k^i) < \varepsilon$$

for all k.

Since S is zero-dimensional, there is a finite partition

$$\mathscr{U} = \{U_1, \cdots, U_M\}$$

of S into nonempty open-and-closed sets that simultaneously refines $\mathscr{V}^{(1)}, \cdots, \mathscr{V}^{(n)}$. By taking a further refinement if necessary, $\mathscr{U}$ may also be assumed to be a refinement of $\mathscr{W}$ and then the functions $P(x)$ are constant on each of the sets U_k.

For each $k = 1, \cdots, M$ define a sublinear functional q_k on X by $q_k(x) = \sup\{Tx(t)\colon\ t \in U_k\}$. From the Hahn-Banach Theorem, there exists a linear functional ϕ_k on X dominated by q_k. Define $T_1\colon X \to C(S_{\mathscr{U}})$ by

$$T_1 x = \sum_{k=1}^{M} \phi_k(x)\ \chi_{U_k}\ .$$

Then $T_1 \in L(P_{\mathscr{U}})$ and, for $i = 1, \cdots, n$,

$$\|(T_1 - T)x_i\| \leqq \sup_{R} \operatorname{Var}(Tx_i,\ U_k) < \varepsilon\ .$$

DEDUCTION OF THEOREM 1.2. With X and S as in the statement of the theorem, let $\mathfrak{B}_1$ be the closed unit ball in $\mathfrak{B}(X, C(S))$.

The set $\mathfrak{B}_1$ is $L(P)$, where P is the sublinear function $P(x) = \|x\| e$, e being the unit function in $C(S)$. By Theorem 2.3 $\mathfrak{B}_1 = \text{co } E\mathfrak{B}_1$ and the result for any closed ball then follows by a scalar multiplication and translation.

3. Nachbin's problem. Let K be a closed bounded convex subset of a linear topological space E. Recall that K has the *positive binary intersection property* if every pairwise-intersecting subfamily of

$$\{x + \lambda K \colon x \in E, \ \lambda \geqq 0\}$$

has nonempty intersection.

If K is bounded and has the above property, it may be shown to be centrally symmetric with a unique centre c, and to have the *binary intersection property* where the restriction $\lambda \geqq 0$ is removed. This is proved in [6].

Results in [4] and [2] then show that the set $K_0 = K - c$ generates a subspace of E which is a hyperconvex normed space and isomorphic to $C(S)$, with S Stonean.

THEOREM 3.1. *Let E be a locally convex Hausdorff linear topological space containing a closed bounded convex subset K with the positive binary intersection property. Let p be a continuous sublinear functional on a locally convex Hausdorff linear topological space X.*

If L is the set of linear maps $T \colon X \to E$ such that for all x in X

$$Tx \in \frac{1}{2}[p(x) - p(-x)]\, e + \frac{1}{2}[p(x) + p(-x)]\, K_0$$

where e is any extreme point of K_0, then $L = \overline{\text{co}}\, L$, with the closure taken in $\mathfrak{B}(X, E)$ with the strong neighborhood topology.

Proof. Because p is continuous the set $L(P)$ is closed in the space $\mathfrak{B}(X, E)$ in the strong neighborhood topology. Since K is centrally symmetric, K_0 has the binary intersection property and is linearly isomorphic to the unit ball in a space $C(S)$ with S Stonean. The isomorphism may be chosen as in [4] so that e is mapped onto the unit function of $C(S)$. Using e to denote also this unit function, we may define a sublinear function $P(x) = p(x)\, e$ from X to $C(S)$, which is the situation of Theorem 3.1. with $\mathscr{W} = \{S\}$.

Given $T \in L(P)$, $x_1, \cdots, x_n \in X$ and $\varepsilon > 0$ there exists $A \in \text{co } EL(P)$ with

$$(T - A)\,x_i \in \varepsilon K_0 \quad (i = 1, \cdots, n)\,.$$

Given a neighborhood U of 0 in E, there exists $r > 0$ with $K_0 \subseteqq rU$, since K is bounded. So choosing $\varepsilon = r^{-1}$ there exists $A \in \mathrm{co}\, EL(P)$ with

$$(T - A)\,x_i \in r^{-1} K_0 \subseteqq U \quad (i = 1, \cdots, n)\,,$$

which completes the proof.

DEDUCTION OF THEOREM 1.1. (a) Let p_K be the sublinear functional defined on F^* by

$$p_K(f) = \sup \{f(k)\colon\ k \in K\}\,.$$

Then, from the bipolar theorem,

$$L = \{g \in F^{**}\colon\ g(f) \leqq p_K(f) \quad \text{for all } f \in F^*\}$$

is identical with the canonical image $\hat{K}$ of K under the evaluation map. Now apply Theorem 3.1 with $E = \mathbf{R}$, $K = [-1, 1]$, $e = 1$ and $X = F^*$, taken with the topology of uniform convergence on compact subsets of F. This shows that $\hat{K}$ is the closure of $\mathrm{co}\, E\hat{K}$ in the topology $w(F^{**}, F^*)$, which is equivalent to K being the $w(F, F^*)$ and hence the strong closure of $\mathrm{co}\, EK$ in F.

(b) Apply Theorem 3.1 with $X = \mathbf{R}$ and $E = F$. Then, under the natural isomorphism of $\mathfrak{B}(X, E)$ and E, K_0 corresponds to L, which satisfies $L = \overline{\mathrm{co}}\, EL$. Since E is a linear topological space we have

$$K = \overline{\mathrm{co}}\, EK\,.$$

REFERENCES

1. N. Aronszajn and P. Panitchpakdi, *Extensions of unifomly continuous transformations and hyperconvex metric spaces*, Pacific J. Math. **6** (1956), 405–439.
2. J. L. Kelley, *Banach spaces with the extension property*, Trans. Math. Soc. **72** (1952), 323–326.
3. J. L. Kelley, I. Namioka, et al., *Linear Topological Spaces*, van Nostrand, 1963.
4. L. Nachbin, *A theorem of the Hahn-Banach type for linear transformations*, Trans. Amer. Math. Soc. **68** (1950), 28–46.
5. ———, *Some problems in extending and lifting continuous linear transformations*, Proc. Int. Symposium on Linear Spaces, Jerusalem 1960, 340–350.
6. ———, *Sur l'abondance des points extrémaux d'un ensemble convexe borné et fermé*, Anais Acad. Brasileira Ciên. **34** (1962), 445–448.

Received August 11, 1969.

UNIVERSITY OF EXETER
U. K.

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

OPERATORS THAT COMMUTE WITH A UNILATERAL SHIFT ON AN INVARIANT SUBSPACE

LAVON B. PAGE

A co-isometry on a Hilbert space $\mathscr{H}$ is a bounded operator having an isometric adjoint. If V is a co-isometry on $\mathscr{H}$ and $\mathscr{M}$ is an invariant subspace for V, then every bounded operator on $\mathscr{M}$ that commutes with V on $\mathscr{M}$ can be extended to an operator on $\mathscr{H}$ that commutes with V, and the extension can be made without increasing the norm of the operator. This paper is concerned with unilateral shifts. The questions asked are these: (1) Do shifts enjoy the above property shared by co-isometries and self-adjoint operators? (The answer to this question is "rarely".) (2) Why not? (3) If S is a shift, $\mathscr{M}$ is an invariant subspace for S, S_0 is the restriction of S to $\mathscr{M}$, and T is a bounded operator on $\mathscr{M}$ satisfying $TS_0 = S_0T$, how tame do T and $\mathscr{M}$ have to be in order that T can be extended (without increasing the norm) to an operator in the commutant of S? Extension is possible in a large number of cases.

The result mentioned above for co-isometries is due to Sz.-Nagy and Foias [8]. (An excellent exposition on the problem is found in [3]; see Theorem 4 in particular.) For self-adjoint operators the statement is trivial for the simple reason that every invariant subspace is then reducing and any commuting operator on a subspace can be extended by simply requiring it to be zero on the orthogonal complement of the subspace.

Recall that a unilateral shift S is an isometry having the property that $\bigcap_{n=0}^{\infty} S^n\mathscr{H} = \{0\}$. The Hilbert space dimension of the subspace $(S\mathscr{H})^\perp$ is called the multiplicity of S. Within the class of partial isometries on $\mathscr{H}$ the unilateral shifts are in a sense as far removed as possible from the co-isometries and the self-adjoint partial isometries. For shifts have no self-adjoint part, and far from being co-isometric if S is a shift S^{*n} goes strongly to zero. (These and other simple properties of shifts may be deduced from problem 118 and the surrounding material in Halmos [5].)

II. We begin with a complex Hilbert space $\mathscr{H}$ (not necessarily separable) and a unilateral shift S on $\mathscr{H}$. It is well known that shifts decompose the underlying Hilbert space in the following way:

$$\mathscr{H} = \bigoplus\sum_{n=0}^{\infty} S^n\mathscr{C} \quad \text{where} \quad \mathscr{C} = (S\mathscr{H})^\perp ,$$

(See for example Halmos [5], problem 118).

We also fix an invariant subspace $\mathscr{M}$ of S. By S_0 we denote the restriction of S to $\mathscr{M}$, $S_0 = S|\mathscr{M}$. The commutant of S is the algebra of bounded operators on $\mathscr{H}$ which commute with S and is denoted by $\mathscr{A}_S$.

The invariant subspaces of S are known to the following extent. Every invariant subspace of S is the range of a partial isometry in $\mathscr{A}_S$ whose initial space reduces S. (This well known result appears in many forms. The particular form cited here appears in [7], see proof of Theorem 1.) Particularly when a function space model is used these operators are often referred to as inner functions or rigid functions.

Finally we will fix a bounded operator T on $\mathscr{M}$ which commutes with S_0. As indicated earlier the problem being considered is that of extending T to an operator on $\mathscr{H}$ lying in $\mathscr{A}_S$ and having norm equal to $\|T\|$.

THEOREM 2.1. *If S is the simple shift, i.e., if* $\dim \mathscr{C} = 1$, *then T has an extension in $\mathscr{A}_S$ whose norm is equal to $\|T\|$.*

Proof. This theorem will follow from a later result. (See Remark 2.4 below.) The simple shift can be represented as the usual shift on the Hardy space H^2 of complex valued functions on the unit circle (Helson [6], chapter 1). It is instructive to sketch a proof in this setting where $\mathscr{M} = BH^2$ with B an inner function in H^2. Also $B \in \mathscr{M}$, and $T: B \to Bg$ for some g in H^2. The fact that $TS_0 = S_0T$ allows one to argue that $T: Bf \to Bfg$ for all $f \in H^\infty$, and finally using standard techniques one shows that $g \in H^\infty$, that T is multiplication by g on $\mathscr{M}$, and hence that T has an obvious extension to an operator on H^2 which commutes with S. The extension does not increase the norm.

EXAMPLE 2.2. *T does not necessarily have a bounded extension which commutes with S if S is a shift of multiplicity two, i.e., if* $\dim \mathscr{C} = 2$.

Proof. Here we let $\mathscr{H} = H^2 \oplus H^2$. Vectors in $\mathscr{H}$ will be written as ordered pairs (f, g). Let χ be the identity function on the unit circle, $\chi(e^{it}) = e^{it}$, and then the shift S of multiplicity two on $\mathscr{H}$ is $S: (f, g) \to (\chi f, \chi g)$.

Let $\mathscr{M}$ be the subspace of $\mathscr{H}$ consisting of all vectors of the form $(f, \chi g)$ where $f, g \in H^2$. Clearly $S\mathscr{M} \subseteq \mathscr{M}$. Define T on $\mathscr{M}$ by $T: (f, g) \to (\bar{\chi} g, 0)$, the bar denoting complex conjugate. It is trivial to verify that T is a bounded operator mapping $\mathscr{M}$ into $\mathscr{M}$, and that $TS = ST$ on $\mathscr{M}$. But it is equally easy to see that T can have no extension in $\mathscr{A}_S$. For if T' is an extension of T to $\mathscr{H}$, then we must have $T'S: (0, 1) \rightarrow (1, 0)$, whereas everything in the range of ST' must be orthogonal to $(1, 0)$.

It becomes apparent in the discussion which follows that the reason we obtain different answers in the case of the simple shift as opposed to nonsimple shifts is that the simple shift is the only shift having an abelian commutant. Recall that $\mathscr{M} = B\mathscr{H}$ where B is a partial isometry in $\mathscr{A}_S$ and $B^*\mathscr{H}$ reduces S. Let A_T be the operator on $\mathscr{H}$ defined by

$$A_T = B^*TB .$$

Since BB^* is the orthogonal projection onto $\mathscr{M}$ we have

$$BB^*TBS = TBS = STB = SBB^*TB = BSB^*TB ,$$

or $BA_TS = BSA_T$. Now the range of A_T is contained in the range of B^* which is a reducing subspace for S. Since B is isometric on the range of B^* we can infer from the last equation that $A_TS = SA_T$. Thus A_T satisfies the three conditions

(i) $A_T \in A_S$

(ii) $TB = BA_T$

(iii) $||A_T|| \leqq ||T||$.

Clearly an operator A in $\mathscr{A}_S$ is an extension of T if and only if $AB = TB$. Thus it follows that T has an extension in $\mathscr{A}_S$ if and only if there exists an operator $A \in \mathscr{A}_S$ such that $AB = BA_T$, i.e., the problem is now one of solving the operator equation $AB = BA_T$ for $A \in \mathscr{A}_S$. (B and A_T are already in $\mathscr{A}_S$.)

A *hyperinvariant* subspace for S is a subspace which is invariant under every operator which commutes with S.

PROPOSITION 2.3. *If $\mathscr{M}$ is a hyperinvariant subspace of S, then T has an extension in $\mathscr{A}_S$ whose norm is $||T||$.*

Proof. The fact that $\mathscr{M}$ is hyperinvariant guarantees that B can be chosen so as to have the additional property that B commutes with every operator in $\mathscr{A}_S$. (Douglas and Pearcy [2], Theorem 5). Thus $A_TB = BA_T$, and T possesses the desired extension by the remarks above.

REMARK 2.4. Since every invariant subspace for the simple shift is hyperinvariant, the above proposition contains Theorem 2.1.

There is a relationship between T having an extension in $\mathscr{A}_S$ and a factorization of a familiar type. From the definition of A_T it is clear that range $A_T^* \subseteq$ range B^*. Thus by a standard factorization result (Douglas [1]) there exists a bounded operator D on $\mathscr{H}$ such that $A_T = DB$.

PROPOSITION 2.5. *If $A_T = DB$ where $D \in \mathscr{A}_S$, then T has an extension in $\mathscr{A}_S$.*

Proof. Suppose $D \in \mathscr{A}_S$ and $A_T = DB$. Then $BA_T = BDB$. Setting $A = BD$ it follows from the remarks made preceeding Proposition 2.3 that T has an extension in $\mathscr{A}_S$.

III. In order that an operator A on $\mathscr{H}$ commute with the shift S it is necessary that every subspace $S^n\mathscr{H}(n \geqq 0)$ be invariant under A. The proposition below is a slight generalization of this statement. For $n \geqq 0$, let $P_n = I - S^nS^{*n}$, the orthogonal projection onto the orthogonal complement of $S^n\mathscr{H}$.

PROPOSITION 3.1. *If $A \in \mathscr{A}_S$, then there is a constant α such that $||P_nAf|| \leqq \alpha||P_nf||$ for every $n \geqq 0$ and every $f \in \mathscr{H}$. In fact α can be chosen to be $||A||$.*

Proof. If $n \geqq 0$ and $f \in \mathscr{H}$ write $f = S^ng + h$ where $g = S^{*n}f$ and $h = P_nf$. Then since $S^{*n}S^n = I$ and $P_nA^* = P_nA^*P_n$, $||P_nAf|| = ||P_nAh|| \leqq ||A|| \, ||h|| = ||A|| \, ||P_nf||$.

With T defined initially on $\mathscr{M}$ Proposition 3.1 indicates that it is fruitless to look for an extension of T in $\mathscr{A}_S$ unless T initially satisfies a similar condition on $\mathscr{M}$. Henceforth we assume that there exists a constant α such that

$$(*) \qquad ||P_nTf|| \leqq \alpha||P_nf||$$

for all $f \in \mathscr{M}$ and $n \geqq 0$.

It is easy now to see that in Example 2.2 T could have no extension in $\mathscr{A}_S$ because condition (*) is not satisfied. If in that example we take $f = (0, \chi)$, then $||P_nf|| = 0$ but $||P_nTf|| = 1$ when $n = 1$.

Whether condition (*) is sufficient to guarantee that T has an extension in A_S we have been unable to determine (see Remark 3.6). We have been able to show, Example 3.5 below, that such an extension cannot always be made without increasing the norm.

The next theorem indicates the existence of a certain subspace

$\mathscr{W}$ between $\mathscr{M}$ and $\mathscr{H}$ and also invariant under S to which T, if T satisfies condition (*), can always be extended without increasing the norm and so as to commute with S. Two corollaries indicate that frequently $\mathscr{W}$ is all of $\mathscr{H}$.

If $f \in \mathscr{H}$, let $\rho(f, \mathscr{M}) = \inf\{\|f - g\|: g \in \mathscr{M}\}$.

THEOREM 3.2. *Let* $\mathscr{W}$ *be the set of all* $f \in \mathscr{H}$ *such that*

$$\rho(S^n f, \mathscr{M}) \to 0$$

as $n \to \infty$. *Then* $\mathscr{W}$ *is a (closed) subspace of* $\mathscr{H}$ *which is invariant under* S, *and if* T *satisfies condition* (*) *on* $\mathscr{M}$ *then* T *has an extension to an operator* T' *on* $\mathscr{W}$ *satisfying* $T'S = ST'$ *on* $\mathscr{W}$ *and* $\|T'\| = \|T\|$.

Proof. It is easy to verify that $\mathscr{W}$ is a linear manifold and that $S\mathscr{W} \subseteq \mathscr{W}$. To see that $\mathscr{W}$ is closed, suppose that f is in the closure of $\mathscr{W}$. Then for $g \in \mathscr{W}$,

$$\rho(S^n f, \mathscr{M}) \leqq \|S^n f - S^n g\| + \rho(S^n g, \mathscr{M}) .$$

By choosing g sufficiently near to f and n sufficiently large, the two terms on the right can be made as small as desired.

We next describe the manner in which T extends to $\mathscr{W}$. Suppose f is in $\mathscr{W}$. Let $\{g_n\}$ be a sequence in $\mathscr{M}$ such that $\lim \|S^n f - g_n\| = 0$, and set $h_n = S^n f - g_n$. Now if $m \geqq n$,

$$\begin{aligned} &\|S^{*n}Tg_n - S^{*m}Tg_m\| = \|S^{*m}TS^{m-n}g_n - S^{*m}Tg_m\| \\ &\leqq \|T\|\,\|S^{m-n}g_n - g_m\| = \|T\|\,\|S^{m-n}h_n - h_m\| \\ &\leqq \|T\|(\|h_n\| + \|h_m\|) , \end{aligned}$$

and the last expression goes to zero as $n, m \to \infty$. Thus we have shown that the sequence $\{S^{*n}Tg_n\}$ is a Cauchy sequence. To extend T to $\mathscr{W}$, if $f \in \mathscr{W}$ we select a sequence $\{g_n\}$ in $\mathscr{M}$ such that

$$\|S^n f - g_n\| \to 0$$

as $n \to \infty$ and set $T'f = \lim S^{*n}Tg_n$. In light of the earlier remarks in this paragraph it is easy to see that the way in which $T'f$ is defined here is independent of the sequence $\{g_n\}$ chosen and coincides with the original operator T in case $f \in \mathscr{M}$. It is also clear that the extension does not increase the norm.

To see that $T'\mathscr{W} \subseteq \mathscr{W}$, we assume $f \in \mathscr{W}$. Let $\{g_n\}$ be a sequence in $\mathscr{M}$ such that $\|S^n f - g_n\| \to 0$. Now making use of the fact that T satisfies condition (*) we have $\|P_n T g_n\| \leqq \alpha \|P_n g_n\|$, and the right-hand side here goes to zero. Furthermore,

$$\rho(S^n T'f, \mathscr{M}) \leqq \| S^n T'f - S^n S^{*n} T g_n \| + \rho(S^n S^{*n} T g_n, \mathscr{M}) .$$

The first term on the right goes to zero by the definition of $T'f$, and the second goes to zero because $Tg_n \in \mathscr{M}$ and $\| P_n T g_n \| \to 0$. Thus $T'f \in \mathscr{W}$.

Finally we show that $T'S = ST'$ on $\mathscr{W}$. If $f \in \mathscr{W}$, let $\{g_n\}$ be a sequence in $\mathscr{M}$ such that $\| S^n f - g_n \| \to 0$. Then

$$\begin{aligned} &\| T'Sf - ST'f \| \leqq \limsup \| S^{*n} T S g_n - S S^{*n} T g_n \| \\ &= \limsup \| S^{*(n-1)} T g_n - S S^{*n} T g_n \| \\ &\leqq \limsup \| S^{(n-1)} S^{*(n-1)} T g_n - T g_n \| + \limsup \| T g_n - S^n S^{*n} T g_n \| \\ &\leqq \alpha \limsup \| P_{n-1} g_n \| + \alpha \limsup \| P_n g_n \| = 0 . \end{aligned}$$

Frequently the subspace $\mathscr{W}$ of Theorem 3.2 will be all of $\mathscr{H}$. The two corollaries below give examples of this occurrence.

COROLLARY 3.3. *If* $\dim \mathscr{M}^\perp < \infty$, *and if* T *satisfies* (*), *then* T *has an extension in* $\mathscr{A}_S$ *whose norm is* $\| T \|$.

Proof. Let $\mathscr{W}$ be the subspace of Theorem 3.2. Assume that x is an eigenvector for the operator on $\mathscr{W}^\perp$ obtained by compressing S to $\mathscr{W}^\perp$, the operator $(I - P)S | \mathscr{W}^\perp$ where P is the orthogonal projection of $\mathscr{H}$ onto $\mathscr{W}$. Let λ be the corresponding eigenvalue, so $|\lambda| \leqq 1$ and $Sx = y + \lambda x$ where $y = PSx$.

Then $S^2 x = Sy + \lambda Sx = (Sy + \lambda y) + \lambda^2 x$. In general

$$S^n x = y_n + \lambda^n x$$

where $y_n \in \mathscr{W}$. Now if $|\lambda| = 1$ then $\| Sx \|^2 = \| y \|^2 + \| x \|^2$, implying that $y = 0$ since S is a contraction. But this would imply that λ is an eigenvalue of S, and since S is a shift S has no eigenvalues.

Thus $|\lambda| < 1$, and $\lambda^n x \to 0$ as $n \to \infty$, implying that $x \in \mathscr{W}$. This too is a contradiction and we have shown that in fact $(I - P)S | \mathscr{W}^\perp$ can have no eigenvalues and hence since $\mathscr{W}^\perp$ is finite demensional we must have $\dim \mathscr{W}^\perp = 0$. The proof is now complete in light of Theorem 3.2.

There is a special type of invariant subspace for nonsimple shifts which is encountered frequently in the literature. Such subspaces are the ones which, in the Hardy space model (Helson [6], chapter 6), correspond to operator valued analytic functions on the unit disk assuming unitary values on the boundary. For a general invariant subspace the corresponding rigid function (see Halmos, [4]) can be required only to assume partially isometric values.

There is an equivalent abstract formulation of the condition that an invariant subspace correspond to a unitary valued function. First of all it is evident that the minimal unitary extension of a unilateral shift is a bilateral shift of the same multiplicity. If we continue to let S and $\mathscr{H}$ denote respectively a unilateral shift and the space on which it acts and now let U and $\mathscr{K}$ denote respectively the minimal unitary extension of S and the space $\mathscr{K}$ on which U acts, then for each subspace $\mathscr{M}$ of $\mathscr{H}$ invariant under S it is clear that $\mathscr{M}$ is invariant under U as well. It can be shown without great difficulty that in the Hardy space model $\mathscr{M}$ corresponds to a unitary function if and only if the smallest reducing subspace for U containing $\mathscr{M}$ is $\mathscr{K}$ itself.

COROLLARY 3.4. *If the smallest reducing subspace for U which contains $\mathscr{M}$ is $\mathscr{K}$ (where U and $\mathscr{K}$ are as in the preceeding paragraph) then every operator T on $\mathscr{M}$ satisfying (*) has an extension in $\mathscr{A}_S$ whose norm is $\|T\|$.*

Proof. Recall that $\mathscr{M} = B\mathscr{H}$ where B is a partial isometry in $\mathscr{A}_S$. From the folklore of the field we know that B has a unique extension to an operator on $\mathscr{K}$, call it B', which commutes with U. (This also can be deduced from the lifting theorem of Sz-Nagy and Foias, Theorem 4 of [3].) Now the range of B' reduces U and contains $\mathscr{M}$. Hence by assumption $B'\mathscr{K} = \mathscr{K}$.

Let $f \in \mathscr{H}$. Since the subspaces $U^{*n}\mathscr{H}$, $n \geqq 0$, span $\mathscr{K}$, for each $\varepsilon > 0$ there is an integer $n \geqq 0$ and a $g \in U^{*n}\mathscr{H}$ such that

$$\|B'g - f\| < \varepsilon \, .$$

We have $U^nB'g = B'U^ng \in B\mathscr{H} = \mathscr{M}$, and $\|S^nf - U^nB'g\| < \varepsilon$. Thus we have shown that $\mathscr{W} = \mathscr{H}$ in Theorem 3.2 and therefore that T has the desired extension.

Our final task will be to show that in general condition (*) on T and $\mathscr{M}$ is not sufficient to guarantee an extension in $\mathscr{A}_S$ with norm equal to $\|T\|$. Because the condition is sufficient in the rather inclusive instances already considered, it is not surprising that some care must be exercised in constructing the following example.

EXAMPLE 3.5. We take S to be a shift of multiplicity 7 on $\mathscr{H}$. Let $\{e_i\}_{i=1}^{7}$ be an orthonormal basis for $(S\mathscr{H})^{\perp}$. We take the subspace $\mathscr{M}$ of $\mathscr{H}$ to be the smallest invariant subspace for S containing the following vectors:

$$u_1 = e_1 + Se_2,\ u_2 = e_3 + Se_4,\ u_3 = e_5 + Se_6,\ u_4 = e_5 + Se_7 \, .$$

The operator T is defined on a dense linear manifold in $\mathscr{M}$ by requiring that

$$Tu_1 = u_3,\ Tu_2 = u_4, \quad \text{and} \quad Tu_3 = Tu_4 = 0$$

and by requiring that $TS = ST$. (The linear manifold referred to is the linear span of the vectors $P(S)u_k$, $k = 1, 2, 3, 4$, where $P(S)$ is a polynomial in S.)

Some elementary calculations show that T is in fact bounded on this linear manifold, and that moreover $\|T\| \leqq \sqrt{3}/\sqrt{2}$. Furthermore it can be shown that T on $\mathscr{M}$ satisfies condition (*) where the constant α can be taken to be $\sqrt{2}$.

Finally one shows that any extension of T to $\mathscr{H}$ which is to commute with S on $\mathscr{H}$ must map $e_1 + e_3$ to $2e_5$, and must hence have norm not less than $\sqrt{2}$. Thus T cannot be extended to an operator which commutes with S on $\mathscr{H}$ without increasing the norm.

REMARK 3.6. It is peculiar in the above example that we could show only that any extension of T to an operator in $\mathscr{A}_S$ must have norm not less than α where α is the constant in (*). This leads naturally to the following conjecture.

CONJECTURE. If T on $\mathscr{M}$ satisfies (*) then T has an extension in $\mathscr{A}_S$ having norm less than or equal to α.

REFERENCES

1. R. G. Douglas, *On factorization, majorization, and range inclusion of operators on Hilbert Space*, Proc. Amer. Math. Soc., **17** (1966), 413-415.

2. R. G. Douglas, and C. Pearcy, *On a topology for invariant subspaces*, Journal of Functional Analysis, **2** (1968), 323-341.

3. R. G. Douglas, P. S. Muhly, and C. Pearcy, *Lifting commuting operators*, Michigan Math. Journal, **15** (1968), 385-396.

4. P. R. Halmos, *Shifts on Hilbert Spaces*, J. Reine Angew, Math., **208** (1961), 102-112.

5. ———, *A Hilbert Space problem book*, Van Nostrand, Princeton, 1967,

6. H. Helson, *Lectures on invariant subspaces*, Academic Press, New York, 1964.

7. M. Rosenblum, *Vectorial Toeplitz operators and the Fejer-Riesz theorem*, J. Math. Anal. Appl., **23** (1968), 139-147.

8. B. Sz.-Nagy, and C. Foias, *Dilations des commutants d'operateurs*, Comptes Rendus, **266** (1968), 493-495.

Received April 6, 1970. This research was supported in part by a grant from the North Carolina Engineering Foundation.

NORTH CAROLINA STATE UNIVERSITY AT RALEIGH

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

PROPERTIES OF FIXED POINT SETS ON DENDRITES

HELGA SCHIRMER

Every nonempty closed subset of a dendrite can be the fixed point set of a self-map, but in general it cannot be the fixed point set of a map with special properties. Necessary conditions found here for the fixed point sets of homeomorphisms and monotone surjections of dendrites are mainly concerned with the order of the possible fixed points, and extend earlier results by G. E. Schweigert and L. E. Ward, Jr.

1. Introduction. It was proved in [3, 4] that every closed, nonempty subset of the n-ball B^n can be the fixed point set of a self-map of B^n, but that not all such subsets can be the fixed point set of a homeomorphism of B^n. We investigate in this paper related questions for dendrites. The first result (Theorem 3.1) shows that again every closed nonempty subset can be the fixed point set of a self-map of a dendrite.

It is already known that not every closed nonempty subset A of a dendrite D can be the fixed point set of a homeomorphism of D, or even of a monotone surjection of D. Results for homeomorphisms by G. E. Schweigert [5] and generalizations for monotone maps by L. E. Ward, Jr. [7] show that A cannot consist of one end point of D:

THEOREM 1.1. (*Schweigert and Ward*). *Let f: $D \rightarrow D$ be a monotone surjection of a dendrite D which leaves one end point e of D fixed. Then there exists at least one fixed point distinct from e.*

We extend this theorem in several ways. In § 4 we prove more details about the order (see [8, p. 48]) of the possible fixed points if the fixed point set consists of only finitely many points. The theorem by Schweigert and Ward states that the fixed point set of a monotone surjection cannot consist of one end point, i.e., of one point of order one. We show in Theorem 4.1 that it also cannot consist of two points of order two, and in the case of a homeomorphism it cannot consist of three points of order three. But it can consist of n points of order n for all $n > 3$. We further strengthen Theorem 1.1 by proving a restriction on the fixed point different from e: if f is a homeomorphism, then it can be chosen of an order $\neq 2$ (Theorem 4.5). This is no longer true for monotone surjections.

The work by Schweigert and Ward is concerned with fixed point sets containing one end point. In § 5 we investigate fixed point sets which contain almost all of the end points, and show that they must contain also all points of a sufficiently high order (Theorem 5.1). In particular we can conclude that if a monotone surjection leaves all but one of the end points fixed, then it leaves in fact all points of order $\neq 2$ fixed (Corollary 5.5).

In § 4 we saw that a distinction exists between fixed point sets of homeomorphisms and of monotone surjections. In the final paragraph (§ 6) we show that such a distinction no longer holds for finite dendrites, i.e., that a subset of a finite dendrite can be the fixed point set of a homeomorphism if and only if it can be the fixed point set of a monotone surjection (Theorem 6.1). The same is true for open maps of finite dendrites, but nothing is known so far about fixed point sets of open maps of arbitrary dendrites.

Ward actually proved Theorem 1.1 not only for dendrites, but more generally for trees, i.e., he did not assume that the space has a metric. It is likely that most or all of the results of this paper can be extended to trees. The metric of the dendrite is used crucially in the proof of Theorem 3.1, and it is also used implicitly in the parts of the paper concerning the order of a point as this concept was developed in [8] for the metric case.

2. **Dendrites.** A dendrite D is a metric continuum (i.e., compact connected Hausdorff space) in which every pair of distinct points is separated by a third. We use the partial order structure of dendrites which was developed by Ward [6, 7]. Take an arbitrary point $r \in D$ as root, and define a partial order $\leqq$ on D by $x \leqq y$ if $x = r$, x separates r and y, or $x = y$. Then $r \leqq x$ for every $x \in D$. Define

$$L(a) = \{y \in D \mid y \leqq a\} ,$$
$$M(a) = \{y \in D \mid a \leqq y\} .$$

The sets $L(a)$ and $M(a)$ are closed in D. Let $[a, b] = M(a) \cap L(b)$; it is a nonempty closed chain (i.e., it is linearly ordered) if $a < b$. Let (a, b) be the interior of $[a, b]$. A point m is called a maximum of a subset A of D, written max A, if $m \nless x$ for each $x \in A$. It is shown in [6, Theorem 1] that every nonempty closed subset of D has a maximum.

We also need in the following some results about dendrites, in particular about the order of points and about arcwise connectedness, which can be found in [8]. Frequently we use the next lemma which characterizes the order $o(a)$ of a point a [8, p. 48] in the case where it is finite.

LEMMA 2.1. *Let a be a point of a dendrite D. If either the order $o(a)$ or the number of the components of $D\backslash\{a\}$ is finite, then these two numbers are equal* [8, p. 88].

$a \in D$ is called an end point if $o(a) = 1$, a cut point if $o(a) \geqq 2$, and a branch point if $o(a) \geqq 3$.

LEMMA 2.2. *Every maximum of D is an end point, and every end point is either a maximum or a root.*

Proof. Let m be a maximum of D. If m is not an end point, then it is a cut point [8, p. 88], and therefore m separates D into two disjoint separated sets A and B [8, p. 42]. Choose A and B so that the root r is in A, and take any $y \in B$. Then m separates r and y, i.e., $m < y$. But this is impossible if m is a maximum. Hence m is an end point. Let now e be an end point with $e \neq r$. As e is not a cut point, the set $D\backslash\{e\}$ is connected, and e cannot separate any two points of $D\backslash\{e\}$. So $e < x$ is not possible for any $x \in D$, and hence e is a maximum of D.

It follows from [6, Theorem 5] that $M(x)$ is connected for all $x \in D$, and therefore $M(x)$ is a subdendrite with root x [8, p. 89]. The space D, and hence $M(x)$, are not only connected, but they are also arcwise connected, and the arc between any two of their points is unique [8, p. 89]. We write arc ab for the unique arc from a to b if $a, b \in D$.

LEMMA 2.3. *If $b_1, b_2 \in D$ and $m =$* max $[L(b_1) \cap L(b_2)]$, *then arc $b_1b_2 = [mb_1] \cup [mb_2]$.*

Proof. The sets $[mb_i] = M(m) \cap L(b_i)$, where $i = 1, 2$, are connected chains and hence arcs [7, Theorem 1; 6, Theorems 4 and 6; 8, p. 36]. As $[mb_1]$ and $[mb_2]$ have exactly one point in common, $[mb_1] \cup [mb_2]$ is an arc, and hence it is the unique arc b_1b_2.

An immediate consequence of Lemma 2.3 is

LEMMA 2.4. *If the connected subset A of D contains the points b_1 and b_2, then it also contains* max $[L(b_1) \cap L(b_2)]$.

We finally state a lemma concerning homeomorphisms and monotone maps (i.e., maps where $f^{-1}(y)$ is connected for all points of the range of f) which is crucial in most of the following work. Its proof can be found in [6, Lemma 13 and p. 156].

LEMMA 2.5. *If f: $D \twoheadrightarrow D$ is a monotone surjective self-map of a dendrite D, then it is isotone* (i.e., $x < y$ *implies* $f(x) \leqq f(y)$). *If f: $D \twoheadrightarrow D$ is a homeomorphism, then it is strictly isotone* (*i.e.*, $x < y$ *implies* $f(x) < f(y)$).

From now on all monotone surjections are assumed to be continuous.

3. Fixed point sets of arbitrary maps on dendrites. We show in this paragraph that any closed nonempty subset can be the fixed point set of a self-map of a dendrite.

THEOREM 3.1. *Let A be an arbitrary closed nonempty subset of a dendrite D. Then there exists a map f: $D \to D$ with A as its fixed point set.*

Proof. Give D the convex metric d (see [1, 2]). As D is acyclic and complete, it follows that for every $x, y \in D$ the point

$$z = tx + (1 - t)y \quad (0 \leqq t \leqq 1)$$

is a unique point of D. As D is compact, it is bounded, hence the diameter diam (D) is finite. Select a point $a \in A$, and define

$$f(x) = \frac{d(x, A)}{\operatorname{diam}(D)} a + \left[1 - \frac{d(x, A)}{\operatorname{diam}(D)}\right] x \text{ for every } x \in D .$$

Then f is the desired map.

Note that the result is not true any longer if we ask in addition that f is surjective. It is e.g., not possible to construct a map from the unit interval onto itself such that its fixed point set consists of one end point of the interval.

4. Nonexistence of some finite fixed point sets. Theorem 1.1 by Schweigert and Ward shows that the fixed point set of a monotone surjection on a dendrite cannot consist of one point of order one. We investigate in this paragraph the existence of fixed point sets on dendrites consisting of n points of order n, for arbitrary positive integers n. The main result is stated in the following theorem.

THEOREM 4.1. *Let f: $D \twoheadrightarrow D$ be a surjective self-map of a dendrite.*

(i) *If f is monotone, then the fixed point set of f cannot consist of n points of order n for $n = 1$ or $n = 2$.*

(ii) *If f is a homeomorphism, then the fixed point set of f cannot consist of n points of order n for $n = 1$, $n = 2$ or $n = 3$.*

The proof of Theorem 4.1 is lengthy and will be accomplished in several parts. The next lemma is used in the proof of part (i) of Theorem 4.1 and in the proof of Theorem 4.5 below.

LEMMA 4.2. *If a is a point of order two in D and different from the root, then it is a point of order one in the subdendrite $M(a)$.*

Proof. As $o(a) = 2$ in D, we can assume that $D\backslash\{a\} = K_1 \cup K_2$, where K_1 and K_2 are the two components of $D\backslash\{a\}$ and the root $r \in K_1$. As K_1 is arcwise connected, we have

$$\begin{aligned} K_1 &= \{x \mid a \notin \text{arc } rx\} \\ &= \{x \mid a \notin [rx]\} = D\backslash M(a) \,. \end{aligned}$$

Hence $K_2 = M(a)\backslash\{a\}$, so that $M(a)\backslash\{a\}$ is connected and $o(a) = 1$ in $M(a)$.

Proof of part (i) *of Theorem* 4.1. Because of Theorem 1.1 we only have to prove the nonexistence of a fixed point set consisting of two points of order two.

Let $f: D \twoheadrightarrow D$ be a monotone surjection which has two fixed points of order two. Take one as root r, and let a be the other fixed point. As f is isotone (Lemma 2.5), we have $f(M(a)) \subseteqq M(a)$. The restriction $f \mid M(a)$: $M(a) \rightarrow M(a)$ is monotone, as for any $y \in f(M(a))$ the counterimage $f^{-1}(y)$ is connected in D and hence (see [8, p. 88]) $f^{-1}(y) \cap M(a)$ is connected in $M(a)$. If $f \mid M(a)$: $M(a) \rightarrow M(a)$ is onto, then it follows from Theorem 1.1 and Lemma 4.2 that f has a second fixed point on $M(a)$, and part (i) of Theorem 4.1 is proved.

Assume now that $f(M(a)) \neq M(a)$, and choose $q \in M(a)\backslash f(M(a))$. As f is surjective, there exists $p \in D\backslash M(a)$ with $f(p) = q$, and because f is isotone, we have $f([ra]) = [ra]$, so that in fact $p \in D\backslash\{M(a) \cup [ra]\}$. Let $m = \max\,[L(p) \cap L(q)]$. Then $r \leqq m < a$ and hence $r \leqq f(m) \leqq a$. But in fact $f(m) = a$: as $f([mp]) = [f(m)q]$ and $a \in [f(m)q]$, there exists an $x \in [mp]$ with $f(x) = a$. But we also have $f(a) = a$, so that by Lemma 2.4 the connected set $f^{-1}(a)$ must contain

$$\max\,[L(x) \cap L(a)] = m, \text{ i.e., } f(m) = a \,.$$

So we see that if f has no other fixed points but r and a, then there exists $m \in \text{arc } ra \backslash \{a\}$ with $f(m) = a$. If we take a instead of r as root, then an analogous argument shows: if f has no other fixed points but r and a, then there exists $n \in \text{arc } ra \backslash \{r\}$ with $f(n) = r$. But as $f(\text{arc } ra) = \text{arc } ra$, the existence of m and n implies the existence of a fixed point on arc ra different from r and a. Hence f must have a fixed point different from r and a, and part (i) of Theorem 4.1 follows.

We now set out to prove part (ii) of Theorem 4.1. This is done with the help of the next two lemmas. The first is stated in much more generality than is needed here for the sake of its use in the proof of Theorem 5.1 below. We say that $f: D \rightarrow D$ permutes the set of n points $\{b_i \mid i = 1, 2, \cdots, n\}$ of D if it transforms the set $\{b_i\}$ bijectively onto itself; the identity transformation of the b_i is included as a possibility.

LEMMA 4.3. *Assume that the monotone surjection $f: D \twoheadrightarrow D$ leaves the root of D fixed and that it permutes the set of points*

$$\{b_i \mid i = 1, 2, \cdots, n\} ,$$

where $n \geqq 2$. Then

$$m = \max\left[\bigcap_{i=1}^{n} L(b_i)\right]$$

is a fixed point of f.

Proof. Let r be the root of D. As $r \leqq m \leqq b_i$, the fact that f is isotone (see Lemma 2.5) implies

$$r \leqq f(m) \leqq f(b_i) = b_k\, (i, k = 1, 2, \cdots, n)$$

and hence $f(m) \leqq m$. But $f([rb_i]) = [rb_k]$, so that there exists for $i = 1, 2, \cdots, n$ an x_i with $r \leqq m \leqq x_i \leqq b_i$ and $f(x_i) = m$. Therefore the connected set $f^{-1}(m)$ contains all x_i, and as

$$\max\left[\bigcap_{i=1}^{n} L(x_i)\right] = \max\left\{L\left[\max \bigcap_{i=1}^{n-1} L(x_i)\right] \cap L(x_n)\right\} ,$$

it follows by induction from Lemma 2.4 that $m \in f^{-1}(m)$. Thus $m = f(m)$ is a fixed point of f.

LEMMA 4.4. *Let $f: D \twoheadrightarrow D$ be a homeomorphism which leaves the root r of D and a point a fixed. Then f maps $M(a)$ homeomorphically onto itself.*

Proof. We see from Lemma 2.5 that $f(M(a)) \subsetneqq M(a)$, so that $f|M(a)$ is an injection. As f is a homeomorphism, its inverse f^{-1} is a homeomorphism too, hence $f^{-1}(M(a)) \subsetneqq M(a)$ or $M(a) \subsetneqq f(M(a))$. Therefore $f(M(a)) = M(a)$, and $f|M(a)$ is a homeomorphism of $M(a)$.

Proof of part (ii) *of Theorem* 4.1. As a homeomorphism is a monotone map, it only remains to show that the fixed point set of a homeomorphism cannot consist of three points of order three.

Let a, b, and c be three distinct fixed points of order three of the homeomorphism $f\colon D \twoheadrightarrow D$. Take a as root. Then

$$m = \max[L(b) \cap L(c)]$$

is a fixed point according to Lemma 4.3. So part (ii) of Theorem 4.1 is proved if m is different from a, b, and c.

Assume now that $m = a$, i.e., that a separates b and c. Define

$$M_b(a) = \{x \mid a \text{ separates } b \text{ and } x\}\ ,$$
$$M_c(a) = \{x \mid a \text{ separates } c \text{ and } x\}\ ,$$

(i.e., $M_b(a)$ is the set $M(a)$ if b is taken as root, and $M_c(a)$ is the set $M(a)$ if c is taken as root). Hence $M_b(a)$ and $M_c(a)$, and therefore [8, p. 88] the set $Q = M_b(a) \cap M_c(a)$, are continua. By definition $a \in Q$, but $b \notin Q$ and $c \notin Q$. It follows from Lemma 4.4 (with b resp. c as root) that f induces a homeomorphism of $M_b(a)$ and of $M_c(a)$, and hence of Q. Therefore in this case part (ii) of Theorem 4.1 follows from Theorem 1.1 if we can show that a is of order one in Q.

If a is not of order one in Q, then $Q\backslash\{a\}$ is not connected. Hence we can select two points $p, q \in Q\backslash\{a\}$ so that $a \in \text{arc } pq$ and therefore $a = \max[L(p) \cap L(q)]$. As $q \in M_b(a)$, we see that a separates b and q. So we have $a \in \text{arc } bq$ and hence $a = \max[L(b) \cap L(q)]$. Similarly $a = \max[L(x) \cap L(y)]$ if $x = b$ or $x = c$, and $y = p$ or $y = q$. This shows that the subdendrite $D' = L(b) \cup L(c) \cup L(p) \cup L(q)$ consists of the four arcs $[ab]$, $[ac]$, $[ap]$, and $[aq]$, and that the order of a in D' is four. As a is of order three in D this is impossible. So a must be of order one in Q, and Theorem 4.1 (ii) holds if $m = a$.

If $m = b$ then b separates a and c. Therefore the same argument, but with b and a interchanged, proves Theorem 4.1 (ii) in this case. If $m = c$ we proceed analogously. This concludes the proof of part (ii) of Theorem 4.1, and hence of Theorem 4.1.

REMARKS. (i) One might ask whether part (ii) of Theorem 4.1 can be extended to monotone maps. That this is not the case is shown by the following example of a monotone surjection of a dendrite which has a fixed point set consisting of three points of order three.

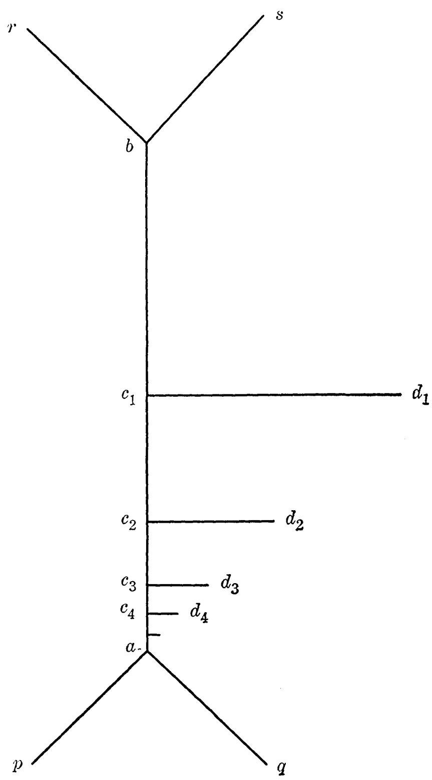

FIGURE 1

Let D be the dendrite illustrated in Figure 1. It is constructed by attaching to the finite dendrite with vertices a, b, p, q, r, and s countably many line segments $[c_i d_i]$, $i = 1, 2, 3, \cdots$, so that c_1 is the mid point of $[ab]$, that c_{i+1} is the mid point of $[ac_i]$ for $i = 1, 2, 3, \cdots$, and that the length of $[c_i d_i]$ equals the length of $[ac_i]$. Define $f\colon D \twoheadrightarrow D$ first on the vertices of D as follows:

$$\begin{aligned}
&f(x) = x \text{ if } x = a, b, \text{ or } c_1 ,\\
&f(p) = q,\ f(q) = p,\ f(r) = s,\ f(s) = r ,\\
&f(d_1) = c_1 ,\\
&\left.\begin{aligned} f(c_{i+1}) &= c_i \\ f(d_{i+1}) &= d_i \end{aligned}\right\} \text{ for } i = 1, 2, 3, \cdots .
\end{aligned}$$

Now extend f linearly over $D \backslash [c_1 b]$, and define it over $[c_1 b]$ as a monotone map of $[c_1b]$ onto itself with the only fixed points c_1 and b. Then $f\colon D \twoheadrightarrow D$ is monotone and has the fixed points a, b, and c_1, each of order three.

(ii) Tempted by one's habit of mathematical induction one might also ask whether it is possible to prove the nonexistence of a fixed point set under a homeomorphism consisting of n points of order n if $n > 3$. That this cannot be done is shown by the next example, in which for any positive integer $n > 3$ a dendrite D_n and a homeomorphism f_n of D_n are constructed so that the fixed point set of f_n consists of exactly n points of order n.

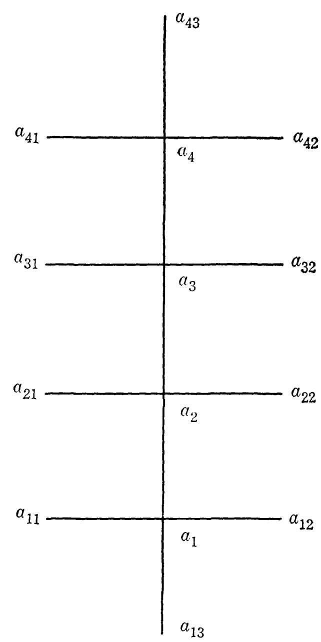

FIGURE 2

Take a chain of n vertices $a_1, a_2, \cdots, a_n$ (see Figure 2 for the case $n = 4$). To both a_1 and a_n attach $n - 1$ segments with end points $a_{ij}(i = 1 \text{ or } n; j = 1, 2, \cdots, n-1)$; to each of $a_2, a_3, \cdots, a_{n-1}$ attach $n - 2$ segments with end points $a_{ij}(i = 2, 3, \cdots, n - 1; j = 1, 2, \cdots, n - 2)$. Then $o(a_i) = n$ for $i = 1, 2, \cdots, n$. Define f_n on the vertices of D_n by

$$f_n(a_i) = a_i \,, \qquad i = 1, 2, \cdots, n \,,$$
$$f_n(a_{ij}) = a_{ik} \,, \qquad i = 1, 2, \cdots, n \,,$$

where $j \neq k$ and $f_n(a_{ij}) \neq f(a_{ij'})$ if $j \neq j'$. Extend f_n as a homeomorphism with no further fixed points over all edges of D_n. Then the fixed point set of f_n is $\{a_1, a_2, \cdots, a_n\}$.

We conclude this paragraph by strengthening Schweigert's theorem in a different direction.

THEOREM 4.5. *If a homeomorphism of a dendrite leaves one end point e fixed, then there exists at least one fixed point distinct from e and of order* $\neq 2$.

Proof. Take the fixed end point as root r of the dendrite D, and let A be the fixed point set of the homeomorphism f of D. It follows from Theorem 1.1 that $A \backslash \{r\} \neq \varnothing$, and hence that $a \neq r$ if a is a maximum of A. Assume that $o(a) = 2$. Then Lemmas 4.2 and 4.4 show that f induces a homeomorphism of $M(a)$ which leaves the end point a of $M(a)$ fixed. Therefore Theorem 1.1 implies the existence of a fixed point of f on $M(a)$ different from a, in contradiction to $a = \max A$. So it is necessary that $o(a) \neq 2$.

REMARK. Theorem 4.5 cannot be extended to monotone maps, for we can construct a monotone surjection of a dendrite such that its fixed point set consists of one point of order one and of one point of order two. For this purpose, let D' be the subdendrite obtained from the dendrite D in Figure 1 by deleting the end points p, q, r, and s, and the four edges which have these points as end points. Let t be the mid point of c_1 and b. Define f on the vertices of D' as follows:

$$\begin{gathered} f(a) = a,\ f(t) = t\ , \\ f(b) = t,\ f(c_1) = t\ , \\ f(d_1) = b\ , \\ \left.\begin{aligned} f(c_{i+1}) &= c_i \\ f(d_{i+1}) &= d_i \end{aligned}\right\} \text{ for } i = 1, 2, 3, \cdots\ , \end{gathered}$$

and extend it linearly over the edges of D'. Then f is a monotone map $D' \twoheadrightarrow D'$ with fixed points a of order one and t of order two.

5. Fixed point sets of monotone maps which leave almost all end points fixed. Theorem 1.1 by Schweigert and Ward considers the case where a map is known to leave one end point fixed. The main result of this paragraph, Theorem 5.1, considers a case which is, in a sense, the other extreme: a map is known to leave almost all, or all, of the end points fixed.

THEOREM 5.1. *Let* $f\colon D \twoheadrightarrow D$ *be a monotone surjection of a dendrite D and assume that at most n of the end points of D do not belong to the fixed point set of f. Then every point of order n (where*

$n \geqq 3$) *belongs to the fixed point set of* f.

We will prove Theorem 5.1 with the help of the following two lemmas.

LEMMA 5.2. *If* $a \in D$, *then every component of* $D \backslash \{a\}$ *contains either the root or an end point.*

Proof. Let K be a component of $D \backslash \{a\}$, and let r be the root of D. We can assume that $r \notin K$. Take an arbitrary $x \in K$. Then $x \neq r$ and $x \neq a$. For any $y \in M(x) \backslash \{x\}$ we have $a \notin [xy]$, as $x < a \leqq y$ would imply $[rx] \subset D \backslash \{a\}$ and hence x would be contained in the same component of $D \backslash \{a\}$ as r. Therefore $M(x) \subset K$. But $M(x)$ is closed in D and hence has a maximum. According to Lemma 2.2 this is an end point, as a maximum of $M(x)$ is clearly a maximum of D.

LEMMA 5.3. *Let* f: $D \twoheadrightarrow D$ *be a monotone surjection which leaves the root of* D *fixed. Then the counterimage of any maximum of* D *contains a maximum.*

Proof. Let m be a maximum. As f is surjective, there exists an $x \in D$ with $f(x) = m$. As f is isotone (Lemma 2.5), we have $f(M(x)) = m$, and as $M(x)$ contains a maximum, so does $f^{-1}(m)$.

Proof of Theorem 5.1. Let $a \in D$ be a point of order n (with $n \geqq 3$). If D has only n end points, then it is a finite dendrite of the form $\bigcup_{i=1}^{n} [ae_i]$, where $[ae_i]$ are arcs. Hence a is fixed. (This can easily be seen directly; it also follows from the arguments used in the proof of Theorem 6.1 below.)

If D has more than n end points, then at least one of them belongs to the fixed point set of f; take it as the root r of D. According to Lemma 5.2 we can select in each of the components of $D \backslash \{a\}$ which do not contain r an end point, thus obtaining at least $n-1$. Choose them as fixed points if possible, and then select from this set exactly $n-1$ end points, again including as many fixed points as possible. We now continue with the proof by investigating three possible cases.

Case 1. At least two of the selected end points, say e' and e'', are fixed points. Consider $m = \max [L(e') \cap L(e'')]$, which is fixed in consequence of Lemma 4.3. If $a \neq m$, then arc $e'e'' \subset D \backslash \{a\}$ by Lemma 2.3, in contradiction to the selection of e' and e'' in different components of $D \backslash \{a\}$. Hence $a = m$ is fixed.

Case 2. Only one of the selected $n-1$ end points, say e, is a fixed point. Let $e_1, e_2, \cdots, e_{n-2}$ (where $n-2 \geqq 1$) be the other selected end points; they are not fixed. If one or two more nonfixed end points of D exist, call them e_{n-1} and e_n. Otherwise put e_n or both of e_{n-1} and e_n equal to e. Define

$$m = \max\,[L(e) \cap \bigcap_{i=1}^{n} \mathrm{L}(e_i)]\ .$$

It follows from Lemmas 2.2 and 5.3 that the set $\{e\} \cup \{e_i \mid i = 1, 2, \cdots, n\}$ is permuted by f. Hence Lemma 4.3 shows that m is a fixed point. As e and $e_i (i = 1, 2, \cdots, n-2)$ are in different components of $D \backslash \{a\}$, we see that $a = \max\,[L(e) \cap L(e_i)]$ and therefore

$$a = \max\,[L(e) \cap \bigcap_{i=1}^{n-2} L(e_i)]\ .$$

This implies $m \leqq a$.

If $m = a$, then a is fixed. If $m < a$, then at least one of e_{n-1} or e_n is $\ngtr a$. Without loss of generality we can assume that $e_{n-1} \ngtr a$ and that $m = \max\,[L(e) \cap L(e_{n-1})]$. As f is isotone, we have $f([me]) = [me]$, hence there exists an x with $m < x < e$ and $f(x) = a$. If $f(e_{n-1}) > a$, then there must also exist a point y with $m < y < e_{n-1}$ and $f(y) = a$. As $f^{-1}(a)$ is connected, it must contain

$$\max\,[L(x) \cap L(y)] = \max\,[L(e) \cap L(e_{n-1})] = m$$

by Lemma 2.4. But $f(m) = m < a$. Therefore $f(e_{n-1}) \ngtr a$, and thus $f(e_{n-1}) = e_n$ and $e_n \ngtr a$.

Assume now that $f(e_n) > a$. Then there exists a point z with $m < z < e_n$ and $f(z) = a$, and $f^{-1}(a)$ contains

$$k = \max\,[L(x) \cap L(z)] = \max\,[L(e) \cap L(e_n)]\ .$$

As $e_n \ngtr a$, we have $m \leqq k < a$. But as $a \in [me]$, we see that $m \leqq f(a) \leqq e$, and as $a \in [me_i]$ for some e_i with $i \leqq n-2$ and $f(e_i) = e_{n-1}$, we see that $m \leqq f(a) \leqq e_{n-1}$. As $m = \max\,[L(e) \cap L(e_{n-1})]$, this implies $f(a) = m$, and therefore $f([ma]) = m$ in contradiction to $f(k) = a$. So it is necessary that $f(e_n) \ngtr a$, i.e., $f(e_n) = e_{n-1}$. We can now apply Lemma 4.3 to the set $\{e\} \cup \{e_i \mid i = 1, 2, \cdots, n-2\}$ to see that a is fixed.

Case 3. None of the selected end points is a fixed point. Denote these end points by $e_1, e_2, \cdots, e_{n-1}$ (where $n-1 \geqq 2$). If one other nonfixed end point exists, call it e_n, otherwise put $e_n = e_1$. As in case 2 we see that

$$m = \max\left[\bigcap_{i=1}^{n} L(e_i)\right]$$

is a fixed point, that

$$a = \max\left[\bigcap_{i=1}^{n-1} L(e_i)\right],$$

and hence that $m \leqq a$. If $m = a$, then a is fixed. If $m < a$, then $e_n \not> a$. Choose $e_i (i \leqq n-1)$ such that $f(e_i) \neq e_n$, i.e., $f(e_i) > a$. Then $a \in f([me_i])$, hence there exists an x with $m < x < e_i$ and $f(x) = a$. But also $a \in f([me_n])$, therefore there exists y with $m < y < e_n$ and $f(y) = a$. Thus the connected set $f^{-1}(a)$ must contain

$$\max\,[L(x) \cap L(y)] = \max\,[L(e_i) \cap L(e_n)] = m$$

by Lemma 2.4, so that $f(m) = a$. But $f(m) = m$, so $m < a$ is impossible.

This completes the proof of Theorem 5.1.

Putting $n = 3$ in Theorem 5.1 we obtain the following special case.

COROLLARY 5.4. *Let $f: D \twoheadrightarrow D$ be a monotone surjection of a dendrite D which leaves at most three of the end points not fixed. Then f leaves all branch points fixed.*

We conclude this paragraph by formulating one further consequence of Theorem 5.1, which is a complement to the theorem by Schweigert and Ward.

COROLLARY 5.5. *Let $f: D \twoheadrightarrow D$ be a monotone surjection of a dendrite D which leaves at most one of the end points not fixed. Then f leaves all end points and all branch points fixed.*

Proof. Assume that we know that f leaves all end points fixed with the possible exception of one end point, say e. Take any of the fixed end points as root r. It follows from Lemmas 2.2 and 5.3 that $f^{-1}(e)$ contains an end point which must of necessity be e. So $f(e) = e$ is fixed. That all branch points of D are fixed follows now from Corollary 5.4.

REMARK. It is not possible to strengthen Theorem 5.1 to include the points with order $n - 1$. To see this, consider the finite subdendrite D'' of Figure 1 with vertices a, b, p, q, r, and s, define f on the vertices of D'' by $f(a) = b$, $f(b) = a$, $f(p) = r$, $f(q) = s$, $f(r) = p$, $f(s) = q$, and extend it linearly over the five edges of D''. Then f leaves none of the four end points of D'' fixed. Take $n = 4$ in Theorem 5.1, and check that the two branch points a and b of order $n - 1 = 3$ are not fixed.

6. Fixed point sets of monotone maps on finite dendrites. In § 4 we found it necessary to distinguish between fixed point sets of monotone maps and fixed point sets of homeomorphisms on dendrites. We will show now that this distinction is superfluous in the case of finite dendrites, i.e., dendrites with finitely many vertices.

THEOREM 6.1. *A subset of a finite dendrite D can be the fixed point set of a homeomorphism of D if and only if it can be the fixed point set of a monotone surjection of D.*

Proof. It is only necessary to show that a subset $A \subset D$ which is the fixed point set of a monotone surjection $f\colon D \twoheadrightarrow D$ can be the fixed point set of a homeomorphism of D.

A is nonempty; select a root r of D with $r \in A$. Take the branch points and end points of D, as well as r if not yet included, as the set V of vertices of a simplicial complex K which is a triangulation of D. We first show that $f \mid V$ determines a simplicial map $\varphi\colon K \twoheadrightarrow K$ (i.e., $f \mid V$ is a function of the vertices of K onto themselves such that adjoining vertices are mapped onto adjoining vertices).

As D is finite, Lemmas 2.2 and 5.3 imply that the image under f of an end point is an end point. Similarly it follows that the image of a branch point is a branch point if we can show that the counterimage of every branch point contains a branch point. Assume by way of contradiction that $b \in D$ is a branch point such that $o(x) = 2$ for all $x \in f^{-1}(b)$. As $f^{-1}(b)$ is closed and connected, it must be of the form $[mn]$, where $m \leqq n$ and $[mn]$ is contained in an edge of D. As f is isotone (see Lemma 2.5), we have $f(M(n)\backslash\{n\}) \subseteqq M(b)\backslash\{b\}$. As $o(n) = 2$, the set $M(n)\backslash\{n\}$ is connected, hence $f(M(n)\backslash\{n\})$ is connected. But $o(b) > 2$, therefore $M(b)\backslash\{b\}$ is not connected, and thus $f(M(n)\backslash\{n\}) \neq M(b)\backslash\{b\}$. Choose $y \in M(b)\backslash\{b\}$ such that

$$f^{-1}(y) \cap [M(n)\backslash\{n\}] = \varnothing .$$

As f is surjective, there exists $x \in D$ with $f(x) = y$, and we see that then $x \not> n$ and even $x \not> m$. As f is isotone, we have $f([rx]) = [ry]$, therefore there exists $x' \in D$ with $r < x' < x$ and $f(x') = b$. This implies $m \leqq x' \leqq n$ and hence $m < x$ in contradiction to $x \not> m$. So it follows that $f^{-1}(b)$ must contain a branch point.

We complete the argument that $f|V$ determines a simplicial map by showing that f maps adjoining vertices of K onto adjoining vertices. Let a and b be adjoining vertices; we can assume that $a < b$. Then $f(a) < f(b)$. If there exists a vertex between $f(a)$ and $f(b)$, then there exists a vertex c which is its counterimage, i.e., $f(a) < f(c) < f(b)$. As f is isotone and as there is no vertex between a and b, this implies $c \notin L(b) \cup M(b)$. Let $m = \max [L(b) \cap L(c)]$, then $o(m) > 2$. As $m > a$ would imply $m \geqq b$ and $c \in M(b)$, it follows that $m \leqq a$. Now $f([ab]) = [f(a) f(b)]$, hence there exists an x with $a < x < b$ and $f(x) = f(c)$. The set $f^{-1}f(c)$ is connected and therefore contains $\max [L(x) \cap L(c)] = \max [L(b) \cap L(c)] = m$, so that $f(m) = f(c)$. As f determines a bijective transformation of V, we must have $m = c$. But this would imply $c \leqq a$ in contradiction to $c \notin L(b)$. So the vertex $f(c)$ cannot exist, and $f|V$ determines a simplicial map φ: $K \twoheadrightarrow K$.

As the image of an edge $[ab]$ under a monotone map f must be the edge $[f(a) f(b)]$, it is now easy to check that the fixed point set A of f must be of the following form:

(1) $a \in A$ for every $a \in V$ with $\varphi(a) = a$;

(2) $A \cap (ab)$ is an arbitrary (possibly empty) closed set for every edge $[ab]$ of D with $\varphi(a) = a$ and $\varphi(b) = b$;

(3) $A \cap (ab) = \varnothing$ for all other edges.

But we can construct a homeomorphism of D with the same images of the vertices as f and with this set A as fixed point set. Therefore Theorem 6.1 holds.

Using a theorem by Whyburn [8, p. 182 Theorem 1.1] we can extend Theorem 6.1 to open maps if D is finite and not an interval, for a study of the proof of Whyburn's theorem shows that in this case $f|V$ again determines a simplicial map. Hence we have

THEOREM 6.2. *If the finite dendrite D is not an interval, then a subset of D can be the fixed point set of a homeomorphism of D if and only if it can be the fixed point set of an open surjection of D.*

The case where D is an interval has to be excluded, as e.g., the subset $\{1/3, 2/3\}$ of the unit interval $[01]$ can be the fixed point set of an open surjection but not of a homeomorphism. It would be interesting to know whether any or all of the results of § 4 and § 5 generalize to open maps. The method of proof will have to be different, though, as an open map of a dendrite need not be isotone.

References

1. R. H. Bing, *Partitioning continuous curves*, Bull. Amer. Math. Soc. **58** (1952), 536–556.
2. R. L. Plunkett, *A fixed point theorem for continuous multi-valued transformations*, Bull. Amer. Math. Soc. **7** (1956), 160–163.
3. H. Robbin, *Some complements to Brouwer's fixed point theorem*, Israel J. Math. **5** (1967), 225–226.
4. H. Schirmer, *On fixed point sets of homeomorphisms of the n-ball*, Israel J. Math. **7** (1969), 46–50.
5. G. E. Schweigert, *Fixed elements and periodic types for homeomorphisms on s.l.c. continua*, Amer. J. Math. **66** (1944), 229–244.
6. L. E. Ward, Jr., *Partially ordered topological spaces*, Proc. Amer. Math. Soc. **5** (1954), 144–161.
7. ———, *A note on dendrites and trees*, Proc. Amer. Math. Soc. **5** (1954), 992–994.
8. G. T. Whyburn, *Analytic Topology*, Providence, R.I., 1942.

Received July 9, 1970. This research was partially supported by the National Research Council of Canada (Grant A7579).

Carleton University,
Ottawa, Canada

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

ON THE NUMBER OF NON-ALMOST ISOMORPHIC MODELS OF T IN A POWER

SAHARON SHELAH

Let T be a first order theory. Two models are almost isomorphic if they are elementarily equivalent in the language $L_{\infty,\omega}$. We investigate the number of non almost-isomorphic models of T of power λ as a function of λ, $I(T, \lambda)$. We prove $\mu > \lambda \geqq |T|$, $I(T, \lambda) \leqq \lambda$ implies $I(T, \mu) \leqq I(T, \lambda)$. In fact, we generalize the downward Skolem-Lowenheim theorem for infinitary languages. Th. (1, 4, 5).

Let L be a set of predicates with finite number of places and sufficient number of variables. (We assume there are no function symbols in L for simplicity only.) $|L|$ will denote the number of predicates in L plus $\aleph_0$. Models will be denoted by M, N. The set of elements of M will be $|M|$, the cardinality of a set A by $|A|$ and so the cardinality of M by $\|M\|$. Unless specified otherwise, every model is an L-model. Cardinals will be denoted by $\lambda, \mu, \kappa, \chi$ ordinals i, j, α, β. T will denote a theory, i.e., set of sentences. We define $\mu^{(\lambda)} = \sum_{\kappa<\lambda} \mu^\kappa$. For cardinals λ, μ we define the language $L(\lambda, \mu)$ i.e., a set of formulas. This set is defined as the well known first-order language where we adjoin to its operations conjunction and disjunction on a set of $<\lambda$ formulas (i.e., $\bigwedge_{i \in I} \phi_i$, where $|I| < \lambda$) and existential or universal quantifications over a sequence of $<\mu$ variables. $L^*(\lambda, \mu)$ will be defined as $L(\lambda, \mu)$ where in addition we permit quantification of the form

$$[\forall \bar{x}^1)(\exists \bar{y}^1) \cdots (\forall \bar{x}^n)(\exists \bar{y}^n) \cdots]_{n<\omega}$$

if

$$|\{x_0^1, x_1^1, \cdots, y_0^1, y_1^1, \cdots, x_0^n \cdots\}| < \mu .$$

$RL^*(\lambda, \mu)$ will denote the subset of $L^*(\lambda, \mu)$ consisting of the formulas Φ of $L^*(\lambda, \mu)$ such that for every subformula ϕ of Φ, if $\phi = [(\forall \bar{x}^1)(\exists \bar{y}^1) \cdots]\psi$, then $\models \phi \leftrightarrow \neg[(\exists \bar{x}^1)(\forall \bar{y}^1) \cdots] \neg \psi$. Clearly $RL^*(\lambda, \mu) \supset L(\lambda, \mu)$. K will denote any of those languages. Satisfaction (i.e., if $\phi = \phi(\bar{x})$, and $\bar{a}$ is a sequence from $|M|$, then $M \models \phi[\bar{a}]$) is defined naturally. (See Hanf [2] and Henkin [3].) The only nontotally trivial case is

$$\psi(\bar{z}) = [(\forall \bar{x}^0)(\exists \bar{y}^0)(\forall \bar{x}^1)(\exists \bar{y}^1) \cdots]\phi(\bar{z}, \bar{x}^0, \bar{x}^1, \cdots, \bar{y}^0, \bar{y}^1 \cdots) .$$

$M \models \psi[\bar{a}]$ if and only if there are functions $f_i^n(\bar{x}^0, \cdots, \bar{x}^n)$ such that for every sequence $\bar{a}^0, \bar{a}^1, \cdots$ from M, $M \models \phi[\bar{a}, \bar{a}^0, \bar{a}^1, \cdots, \bar{b}^0, \bar{b}^1, \cdots]$ where $\bar{b}^n = \langle \cdots, f_i^n(\bar{a}^0, \bar{a}^1, \cdots, \bar{a}^n), \cdots \rangle$. For a sentence ψ, $\models \psi$ if for every M, $M \models \psi$. (Such languages were first defined in Henkin [3].)

If Γ is a set of formulas (for example one of the languages defined above), M is a Γ elementary submodel of N, if the set of elements of M, $|M|$ is included in the set of elements of N, $|N|$, and for every formula $\phi(\bar{x})$, $\phi(\bar{x}) \in \Gamma$, and sequence $\bar{a}$ from $|M|$, $M \models \phi[\bar{a}]$ if and only if $N \models \phi[\bar{a}]$, M, N are Γ-elementarily equivalent if for every sentence $\phi \in \Gamma$, $M \models \phi$ if and only if $N \models \phi$.

THEOREM 1. *Let $\lambda > \mu$, λ regular and T be a theory in $RL^*(\lambda, \mu)$ [i.e., $T \subset RL^*(\lambda, \mu)$] and Γ be the set of subformulas of the formulas in T. Then for every model M we can add $<\lambda + |T|^+$ functions of $<\mu$ places such that: If $A \subset M$, and A is closed under those functions, then there exists a Γ-elementary submodel N of M, $|N| = A$. So if $\kappa \geqq \lambda + |T|$ (or $\kappa \geqq$ the number of those functions) and $\kappa^{(\mu)} = \kappa$, and T has a model of power $\geqq \kappa$, then T has a model of power κ.*

Proof. This theorem is proved in [9], and is a straight-forward generalization of a theorem of Hanf in [2].

DEFINITION 1.

$$L(\infty, \mu) = \bigcup_{\lambda} L(\lambda, \mu), \; L^*(\infty, \mu) = \bigcup_{\lambda} L^*(\lambda, \mu),$$
$$RL^*(\infty, \mu) = \bigcup_{\lambda} RL^*(\lambda, \mu) \; .$$

DEFINITION 2. (1) M and N are μ-almost isomorphic, $M \sim_{\mu} N$ if M, N are $L(\infty, \mu)$-elementarily equivalent. We say M and N are almost isomorphic if $M \sim_{\aleph_0} N$, and we write $M \sim N$.

(2) $I(T, \lambda, \mu)$, is the number of non-μ-almost-isomorphic models of T of power λ. We assume always λ is $\geqq$ then $|T|$.

See footnote 1.

THEOREM 2. *If T is a theory in $RL^*(\lambda, \mu)$, $\mu = \aleph_0$ or $\mu = \mu_1^+$, $\kappa \geq \chi = \chi^{(\mu)} + \lambda + |T|$ and $I(T, \chi, \mu) \leqq \chi$ then $I(T, \kappa, \mu) \leqq I(T, \chi, \mu)$.*

The proof is broken into a series of lemmas.

REMARKS. (1) It is not hard to show that if $T \subset L(\lambda, \aleph_0)$, $I(T, \chi, \aleph_0) \leqq \chi$, then for every $\kappa_1, \kappa_2 \geqq \beth_{(2^{\lambda+\chi})^+}$, $I(T, \kappa_1, \aleph_0) = I(T, \kappa_2, \aleph_0)$. (See Makkai [7] and Eklof [15].)

[1] The results here appear in the notices [10] Th. 5 [11] Th. 3. The lemma has other uses: see [12] Th. 2.5 and Remark (4): in [11] their consequences are better formulated. In Th. 2 we can replace $T \subset RT^*(\lambda, \mu)$ by $T \subset RL^*(\lambda^+, \mu)$ and similarly in other cases.

(2) Let $\lambda = \mu = \aleph_0$ and suppose $|T| \leqq \kappa_0$. Then as the class of such theories is a set, there is a number $\kappa = HAI_{\kappa_0}$ (Hanf number of almost isomorphism) such that: for all T, $|T| \leqq \kappa_0$, $I(T, \kappa, \aleph_0) \leqq \kappa$ if and only if there is a χ, $I(T, \chi, \aleph_0) \leqq \chi$, and κ is the first such cardinality. (The existence of such numbers for a wide class of cases was proved by Hanf in [2].)

Question 1. What is HAI_{κ_0}? (Clearly if $\lambda \to (\kappa_0^+)^{<\omega}_{2^{\kappa_0}}$ then $HAI_{\kappa_0} < \lambda$).

(3) It is known that $M \sim N$, $\aleph_0 = ||M|| = ||N||$ implies that M, N are isomorphic (see Scott [8]).

(4) Ehrenfeucht in [1] defined a model to be rigid if it has no nontrivial automorphisms and tried to investigate what can be the class of cardinals in which a certain theory has a rigid model. He gives some examples, but does not prove any theorem of the form: If T has a rigid model of one power, then it has a rigid model in another power.

DEFINITION. M is μ-rigid if there do not exist two different sequences of length $<\mu$, $\bar{a}$, $\bar{b}$, such that $(M, \bar{a}) \sim_\mu (M, \bar{b})$. ($(M, \bar{a})$ is the model M when we adjoin the a's as individual constants.) See footnote 2. Clearly

THEOREM. *If $\mu < \lambda$, and M is μ-rigid, then it is λ-rigid and also rigid. By a proof similar to that of Theorem 2, we can prove*:

THEOREM. *If a first-order theory T has a μ-rigid model of power λ, $|T| + \aleph_0 \leqq \kappa = \kappa^{(\mu)} \leqq \lambda$, $\mu = \mu_1^+$ or $\mu = \aleph_0$, then T has a μ-rigid model of power κ.*

Proof of Theorem 2.

DEFINITION 3. (1) Let L_1 be L where we adjoin to it one two-place predicate E and variables $y, y_0, y_1, \cdots$ we assume $E, y, y_0 \cdots \neq L$. We shall write xEy instead $E(x, y)$.

(2) If $R \in L$ then R^M will denote the relation of M corresponding to R.

(3) Let $\{M_i: i \in I\}$ be a set of L-models and we define their sum $N = \bigoplus_{i \in I} M_i$, (or $\bigoplus\{M_i: i \in I\}$). For simplicity we assume that the sets $|M_i|$ are pairwise disjoint. N will be an L_1-model $|N| = \bigcup_{i \in I} |M_i|$, $R^N = \bigcup_{i \in I} R^{M_i}$ for every $R \in L$, and $E^N = \{\langle a, b\rangle: (\exists i)[a, b \in |M_i|]\}$.

(4) For every formula ϕ of any language, we define by induction

[2] Barwise [14] suggests a similar definition and argues its naturality.

$\bar{\phi}$: if ϕ is atomic $\bar{\phi} = \phi$; $\overline{\neg\phi} = \neg\bar{\phi}$, $\overline{\phi \vee \psi} = \bar{\phi} \vee \bar{\psi}$, (likewise for the other connectives), $\overline{\exists(\exists\bar{x})\phi} = (\exists\bar{x})[\bar{\phi} \wedge \bigwedge_i x_i E y]$, (where $\bar{x} = \langle \cdots x_i \cdots \rangle$) $\overline{(\forall\bar{x})\phi} = (\forall\,\bar{x})[\bigwedge_i x_i E y \rightarrow \bar{\phi}]$, and

$$\overline{[(\forall\bar{x}^1)(\exists\bar{y}^1)\cdots]\phi} = [(\forall\bar{x}^1)(\exists\bar{y}^1)\cdots](\bigwedge_{i,n} x_i^n E y \rightarrow \bar{\phi} \wedge \bigwedge_{i,n} y_i^n E y)$$

if the language contains such formulas. Clearly for any language K, $\phi \in K \Rightarrow \bar{\phi} \in K$. Also, if ϕ is a sentence $(\forall y)\bar{\phi}$ is a sentence.

(5) Define

$$\bar{T} = \{(\forall y)\bar{\phi}\colon \phi \in T\} \cup \{(\forall x)xEx,\ (\forall x_0 x_1 x_2)(x_0 E x_1 \wedge x_0 E x_2 \rightarrow x_1 E x_2)\}\ .$$

LEMMA 3. *Each M_i is an L-model of T if and only if $\bigoplus_{i\in I} M_i$ is an L_1-model of $\bar{T}$.*

Proof. Immediate

DEFINITION 4.

$$\begin{aligned}\psi_\alpha^n = \psi_\alpha^n(\bar{x}^0, \bar{x}^1, \cdots, \bar{x}^n, \bar{y}^0, \cdots, \bar{y}^n) = \bigwedge \{&R(x_{j_1}^{i_1}, \cdots, x_{j_k}^{i_k} \cdots)\\ &\leftrightarrow R(y_{j_1}^{i_1}, \cdots, y_{j_k}^{i_k} \cdots)\colon i_1, \cdots, i_k \cdots \in \{0, \cdots, n\},\\ &R \in L, j_1, \cdots, j_k \cdots < \alpha\}\end{aligned}$$

where

$$\bar{x}^n = \langle \cdots x_i^n \cdots \rangle_{i<\alpha},\ \bar{y}^n = \langle \cdots y_i^n \cdots \rangle_{i<\alpha}\ .$$

Also let

$$\begin{aligned}\Phi_\alpha^m = [\bigwedge_{\substack{i<\alpha\\ 2n<m}} x_i^{2n} E x \wedge \bigwedge_{\substack{i<\alpha\\ 2n+1<m}} y_i^{2n+1} E y] \rightarrow [\bigwedge_{\substack{i<\alpha\\ 2n+1<m}} x_i^{2n+1} E x \wedge \bigwedge_{\substack{i<\alpha\\ 2n<m}} y_i^{2n} E y\\ \wedge \bigwedge_{n<m} \psi_\alpha^n(\bar{x}^0, \cdots, \bar{x}^n, \bar{y}^0, \cdots, \bar{y}^n)]\ :\end{aligned}$$

$$\phi_\alpha^\omega = \bigwedge_{m<\omega} \Phi_\alpha^m = \phi_\alpha^\omega(x, y, \bar{x}^0, \bar{y}^0, \bar{x}^1, \bar{y}^1, \cdots)\ .$$

For even n

$$\phi_\alpha^n = \phi_\alpha^n(x, y, \bar{x}^0, \bar{y}^0, \cdots, \bar{x}^{n-1}, \bar{y}^{n-1}) = [(\forall\bar{x}^n)(\exists\bar{y}^n)(\forall\bar{y}^{n+1})(\exists\bar{y}^{n+1})\cdots]\phi_\alpha^\omega\ .$$

For odd n

$$\phi_\alpha^n(x, y, \bar{x}^0, \bar{y}^0, \cdots, \bar{x}^{n-1}, \bar{y}^{n-1}) = [(\forall\bar{y}^n)(\exists\bar{x}^n)(\forall\bar{x}^{n+1})(\exists\bar{y}^{n+1})(\forall\bar{y}^{n+2})\cdots]\phi_\alpha^\omega\ .$$

LEMMA 4. *If*

$$a \in |M|,\ b \in |N|,\ M, N \in \{M_i\colon i \in I\},\ M^* = \bigoplus_{i\in I} M_i\ ,$$

and $\mu = \kappa^+$ or $\mu = \aleph_0$, and κ is finite, then $M \sim_\mu N$ if and only if $M^* \models \phi_\kappa^0[a, b]$.

REMARK. Keisler in [5] used sentences similar to ϕ^n_α. These sentences can be seen as asserting something about an appropriate game (between a player choosing $\bar{x}^0, y^1, x^2, \cdots$ and a player choosing $\bar{y}^0$, $\bar{x}^1, \cdots$). In this presentation a similar theorem appears in Karp [4].

Added in proof. See also Benda [13].

Proof.

Part A- Suppose $M \sim_\mu N$.

For every two sequences $\bar{a}, \bar{b}$ of elements of M, either there is a formula $\phi_{\bar{a},\bar{b}}(\bar{x})$ of $L(\infty, \mu)$ such that $M \models \phi_{\bar{a},\bar{b}}[\bar{a}]$, $M \models \neg\, \phi_{\bar{a},\bar{b}}[\bar{b}]$, or there is no such ϕ and in this case, we let $\phi_{\bar{a},\bar{b}}(\bar{x}) = (x_0 = x_0)$.

Let $\phi_{\bar{a}}(\bar{x}) = \bigwedge_{\bar{b}} \phi_{\bar{a},\bar{b}}(\bar{x}) \in L(\infty, \mu)$. Let $\overline{\phi_{\bar{a}}(\bar{x})} = \phi'_{\bar{a}}(y, \bar{x})$. Let $\alpha < \mu$. We define the functions

$$f^{2n}_i(\bar{x}^0, \bar{y}^0, \bar{y}^1, \bar{x}^1, \bar{x}^2, \cdots, \bar{y}^{2n-1}, \bar{x}^{2n-1}, \bar{x}^{2n}),$$
$$f^{2n+1}_i(\bar{x}^0, \bar{y}^0, \bar{y}^1, \bar{x}^1, \bar{x}^2, \cdots, \bar{x}^{2n}, \bar{y}^{2n}, \bar{y}^{2n+1})$$

for $i < \alpha$ such that: If $\bar{a}^0, \bar{b}^0, \bar{a}^1, \bar{b}^1 \cdots$ are sequences of length α, $\bar{a}^{2n}$ a sequence of elements of M, and $\bar{b}^{2n+1}$ a sequence of elements of N, and for every n

$$\bar{b}^{2n} = \langle \cdots f^{2n}_i(\bar{a}^0, \bar{b}^0, \cdots, \bar{a}^{2n}) \cdots \rangle_{i<\alpha}$$
$$\bar{a}^{2n+1} = \langle \cdots f^{2n+1}_i(\bar{a}^0, \cdots, \bar{b}^{2n+1}) \cdots \rangle_{i<\alpha}$$

then $M^* \models \phi^\omega_\alpha[a, b, \bar{a}^0, \bar{b}^0, \cdots]$.

Suppose we have defined f^n_i for $n < 2m$, and let us define f^{2m}_i for $i < \alpha$. (f^{2m+1}_i are defined similarly.)

If for some $n < 2m$, $i < \alpha$ $b^n_i \notin |N|$, or for some $i < \alpha$, $n \leqq 2m$ $a^n_i \notin |M|$, then $f^{2m}_i(\bar{a}^0, \cdots, a^{2m})$ is defined as an arbitrary element of M^*. Also if there exists a formula $\psi(\bar{z}^1, \cdots, \bar{z}^n) \in L(\infty, \mu)$ such that

$$M \models \psi[\bar{a}^0, \bar{a}^1, \cdots, \bar{a}^{2m-1}] N \models \neg\, \psi[\bar{b}^0, \cdots, \bar{b}^{2m-1}],$$

we define $f^{2m}_i(\bar{a}^0 f^0 \cdots \bar{a}^{2m})$ arbitrarily.

So assume none of the previous cases occur. Define $\bar{a}[n] = \bar{a}^0 \frown \bar{a}^1 \frown \cdots \frown \bar{a}^n$ (the concatenation of $\bar{a}_1, \cdots, \bar{a}^n$) and $\bar{b}[n] = \bar{b}^0 \frown \cdots \frown \bar{b}^n$. Clearly

$$M \models (\forall \bar{x})(\phi_{\bar{a}[2m-1]}(\bar{x}) \rightarrow (\exists \bar{z})\phi_{\bar{a}[2m]}(\bar{x}, \bar{z})).$$

As $M \sim_\mu N$, N also satisfies the above sentence; so there exists $\bar{b}^{2m}$ such that for every $\phi \in L(\infty, \mu)$, $M \models \phi[\bar{a}^0, \cdots, \bar{a}^{2m}]$ if and only if $N \models \phi[\bar{b}^0, \cdots, \bar{\phi}^{2m}]$. Let $f^{2m}_i(\bar{a}^0, \bar{b}^0, \cdots, \bar{a}^{2m}) = \bar{b}^{2m}_i$.

Clearly this shows that $M^* \models \phi^0_\alpha[a, b]$ for every $\alpha < \mu$, and in particular for κ.

Part B. We now assume that $M^* \models \phi^0_1[a, b]$, and $\mu = \aleph_0$. The proof in the case $\mu = \kappa^+$ or $1 < \kappa < \aleph_0$ is similar. For simplicity, we shall not distinguish between $\bar{a} = \langle a_0 \rangle$ and a_0.

Two sequences, $\bar{a}$ from M and $\bar{b}$ from N, of length n, $n < \omega$, will be called equivalent if $M^* \models \phi^n_1[a, b, \bar{a}, \bar{b}]$. If $n = 2m$, clearly for every $b^{n+1} \in |N|$ there exists $a^{n+1} \in |M|$ such that $\bar{a} \frown \langle a^{n+1} \rangle$ and $\bar{b} \frown \langle b^{n+1} \rangle$ are equivalent, and similarly for $n = 2m + 1$.

Let $\phi(\bar{x}) \in L(\infty, \mu)$, $\bar{x}$ a finite sequence of variables. We shall prove that if $\bar{a}$, $\bar{b}$ are equivalent then $M \models \phi[\bar{a}]$ if and only if $N \models \phi[\bar{b}]$. As subformulas of formulas with $< \aleph_0$ free variables have $< \aleph_0$ free variables we can prove it by induction. For atomic formulas it follows from the definition of ϕ^n_1. For $\neg\phi$, $\phi \vee \psi$, it is immediate, and so also for the other connectives. For quantification it follows by the fact mentioned above after the definition of equivalent sequences.

So we have proved that if $\bar{a}$, $\bar{b}$ are equivalent sequences, $\phi(\bar{x}) \in L(\infty, \mu)$, then $M \models \phi[\bar{a}]$ if and only if $N \models \phi[\bar{b}]$. Since the sequences of length zero from M and N are equivalent (by our hypotheses $M^* \models \phi^0_1(a, b)$), we get our conclusion that $M \sim N$. This proves Lemma 4.

LEMMA 5. $\phi^0_\alpha(x, y) \in RL^*(\infty, \mu)$. See footnote 3.

Proof. It is easily seen that the only thing we have to prove is:

$$\models [(\forall \bar{x}^0)(\exists \bar{y}^0)(\forall y^1)(\exists x^1) \cdots] \bigwedge_{n<\omega} \phi^n_\alpha \leftrightarrow \neg [(\exists \bar{x}^0)(\forall \bar{y}^0)(\exists \bar{y}^1)(\forall x^1) \cdots] \bigvee_{n<\omega} \neg \phi^n_\alpha .$$

For simplicity, let $\alpha = 1$.

It is not hard to see that if $M \models [(\forall x^0)(\exists y^0) \cdots] \bigwedge_{n<\omega} \phi^n_1$, then $M \models \neg[(\exists x^0)(\forall y^0) \cdots] \bigvee_{n<\omega} \neg \phi^n_1$. (See, for example, Keisler [6].)

So suppose $M \models \neg [(\exists \bar{x}^0)(\forall y^0) \cdots] \bigvee_{n<\omega} \neg \phi^n_1$. It is not hard to see that for every $n < \omega$, and formula ϕ

$$\models \neg [(\forall z_1)(\exists z_2)(\forall z_3) \cdots]\phi \leftrightarrow (\exists z_1) \neg [(\exists z_2)(\forall z_3) \cdots]\phi$$
$$\models (\exists z_1) \neg [(\exists z_2)(\forall z_3) \cdots]\phi \leftrightarrow (\exists z_1)(\forall z_2) \neg [(\forall z_3) \cdots]\phi , \qquad \text{etc.}$$

Now let us define functions $g_n(x^0, y^0, y^1, \cdots, x^i \cdots y^j \cdots)_{i,j<n}$. Let

$$\theta_n(x, y, x^0, y^0, x^1, y^1, \cdots, x^n, y^n) = \neg [\forall x^n)(\exists y^n)(\forall y^{n+1})(\exists x^{n+1}) \cdots] \bigvee_{n<\omega} \neg \phi^n_1 .$$

[3] This lemma is, in fact, a translation of a well known theorem from game theory.

(This is for even n, the definition for odd n is clear.) The functions will be such that if $a^0, \cdots, a^n \in |M|, b^0, \cdots, b^n \in |N|$, and for every $2m \leqq nb^{2m} = g_{2m}(a^0, b^0, \cdots)$, and for every $2m + 1 \leqq na^{2m+1} = g_{2m+1}(a^0, b^0, \cdots)$; then $M^* \models \theta_n[a, b, a^0, b^0 \cdots]$. The definition is self-evident. Let $a^0 \cdots a^n \cdots \in |M|, b^0 \cdots b^n \cdots \in |N|$ be such that for every $2mb^{2m} = g_{2m}(a^0, b^0 \cdots)$ and for every $2m + 1\ a^{2m+1} = g_{2m+1}(a^0, b^0 \cdots)$ and let $n < \omega$. As $M^* \models \theta_{n+1}[a, b, a^0, b^0 \cdots a^n, b^n]$, clearly $M^* \models \phi_1^n(a, b, a^0, b^0 \cdots a^n, b^n)$.

So $M^* \models \bigwedge_{n<\omega} \phi_1^n(a, b, a^0, b^0, \cdots, a^n b^n)$, and hence $M^* \models \phi_1^\omega[a, b, a^0, b^0 \cdots]$. So $M^* \models \phi_1^0[a, b]$ (as this is true for every $a^0, b^1, a^2, b^3 \cdots$) and this is the desired conclusion.

LEMMA 6. *Let* $\mu = \kappa^+$ or $\mu = \aleph_0$, $\kappa = 1$, *T a theory in $RL^*(\lambda, \mu)$, $\chi = \chi^{(\mu)} + \lambda + |T|$, and $I(T, \chi, \mu) \leqq \chi$. Then for every model N of T of power $>\chi$, there exists a model M of T of power χ such that $M \sim_\mu N$.*

REMARK. This clearly proves Theorem 2.

Proof. Let $\{M_i : i \in I\}$ be a maximal set of non-μ-almost-isomorphic models of T of power χ, and let N be a model of T of power $>\chi$ such that for no $i \in I$, $N \sim_\mu M_i$.

Let $M^* = \bigoplus (\{N\}\{M_i : i \in I\})$. Clearly M^* is a model of $T_1 = \bar{T} \cup \{(\forall x, y)[\neg xEy \rightarrow \neg \phi_\kappa^0(x, y)]\}$. Let $a \in |N|$, and $A = \{a\} \cup \bigcup \{|M_i| : i \in I\}$. Clearly, $|A| = \chi$.

Let Γ be the set of subformulas of formulas $\in T_1$. By Theorem 1, it follows that M^* has a Γ-elementary submodel $N^*, |N^*| \supset A, \chi = ||N^*|| =$ (the power of N^*), such that every equivalence class (of E) in N^* has exactly χ elements. Clearly, $N^* = \bigoplus (\{N_1\} \cup \{M_i : i \in I\})$, and for every i, N_1, M_i are models of T, and they are non-μ-almost-isomorphic. So N_1 contradicts the definition of $\{M_i : i \in I\}$, thus proving Lemma 6.

This ends the proof of Theorem 2.

REFERENCES

TM will denote the Proc. 1963 Berkeley Symposium on Theory of models, North Holland Publ. Co., 1965.

1. A. Ehrenfeucht, *Elementary Theories with Models Without Automorphisms*, TM, pp. 70–76.
2. W. Hanf, *Doctoral Dissertation*, University of California, 1962.
3. L. Henkin, *Some Remarks on Infinitely Long Formulas, Infinitistic Methods*, Warsaw, (1961), 167–183.

4. C. Karp, *Finite Quantifier Equivalence*, TM, pp. 407-412.
5. H. J. Keisler, *Some applications of infinitely long formulas*, J. Symbolic Logic **30** (1965), 339-349.
6. ———, *Formulas With Linearly Ordered Quantifiers*, Lecture Notes in Math. 72. The Syntex and Semantics of Infinitary Languages, (1968), 96-131.
7. M. Makkai, Notices of A.M.S., vol. **16** (1964), p. 322.
8. D. Scott, *Logic with denumerable long formulas and finite strings of quantifiers*, TM, pp. 329-341.
9. S. Shelah, Master's thesis written under the guidence of Professor H. Gaifman. The Hebrew Univ. Jerusalem 1967.
10. *On the number of the non-isomorphic models of a theory in a cardinality*, Notices of Amer. Math. Soc., **17** (1970), 576.
11. ———, *Some unconnected results in model theory*, Notices of the Amer. Math. Soc., **18** (1971), 576. April.
12. ———, *A combinatorial problem, stability and order for models and theories in infinitary languages*, to appear (Pacific J. Math.)
13. M. Benda, *Reduced products and nonstandard logics*, J. Symbolic Logic, **34** (1969), 424-436.
14. J. Barwise, *Back and forth thru infinitary logic*, in a forthcoming book edited by Morley.
15. P. C. Eklof, *On the existence of $L_{\infty,k}$-indiscernibles*, Proc. Amer. Math. Soc., **25** (1970), 798-800.

Received August 19, 1969.

THE HEBREW UNIVERSITY
AND
THE UNIVERSITY OF CALIFORNIA, LOS ANGELES

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

MINIMAL FIRST COUNTABLE HAUSDORFF SPACES

R. M. STEPHENSON, JR.

If $\mathscr{P}$ is a property of topologies, a $\mathscr{P}$-space $(X,\mathscr{T})$ is called a $\mathscr{P}$-minimal space if there exists no $\mathscr{P}$-topology on X properly contained in $\mathscr{T}$. Throughout the following, $\mathscr{H}$ = first countable and Hausdorff and $\mathscr{C}$ = first countable and completely Hausdorff (a space X is called completely Hausdorff if the continuous real valued functions defined on X separate the points of X).

In this paper we give examples of $\mathscr{H}$-minimal $\mathscr{C}$-spaces that are (i) not regular and (ii) regular but neither completely regular nor countably compact.

Two other results obtained are the following. (a) Every locally pseudocompact zero-dimensional $\mathscr{H}$-space can be embedded densely in a pseudocompact zero-dimensional $\mathscr{H}$-space. (b) Let $\mathscr{P} = \mathscr{C}$, completely regular $\mathscr{H}$, or zero-dimensional $\mathscr{H}$, and suppose that X is a $\mathscr{P}$-space such that for every $\mathscr{P}$-space Y and continuous mapping $f\colon X \to Y$, f is closed. Then X is countably compact.

N will denote the set of natural numbers, and $C(X, Y)$ will denote the family of continuous mappings of X into Y. For definitions, see [4].

1. An embedding theorem and some examples.

Recall that a space $(X, \mathscr{T})$ is said to be *semiregular* if $\{\mathring{\bar{T}} \mid T \in \mathscr{T}\}$ is a base for $\mathscr{T}$. If $(X, \mathscr{T})$ has a property $\mathscr{P}$, then $(X, \mathscr{T})$ is said to be $\mathscr{P}$-*closed* provided that it is a closed subset of every $\mathscr{P}$-space in which it can be embedded.

For many properties $\mathscr{P}$, it is known that $\mathscr{P}$-minimal and $\mathscr{P}$-closed spaces are closely connected. For the case $\mathscr{P} = \mathscr{H}$, the following two results, established in [11], will be used below. An $\mathscr{H}$-space X is $\mathscr{H}$-closed if and only if every countable open filter base on X has nonempty adherence. An $\mathscr{H}$-space is $\mathscr{H}$-minimal if and only if it is semiregular and $\mathscr{H}$-closed.

We shall now describe constructions which can be used to densely embed certain $\mathscr{C}$-spaces in $\mathscr{H}$-minimal ($\mathscr{H}$-closed) $\mathscr{C}$-spaces. As special cases, we shall obtain examples with the properties mentioned in the introduction. First some terminology is needed.

A space X is said to be *locally pseudocompact* (W. W. Comfort) if every point of X has a pseudocompact neighborhood.

A filter base $\mathscr{F}$ is said to be *pseudocompact* if for every $F \in \mathscr{F}$ and $G \in \mathscr{F}$, $F - G$ is pseudocompact. $\mathscr{F}$ is called *zero-dimensional* if the sets belonging to it are open- and-closed.

Notation. (B. Banaschewski). Let $\mathscr{M}$ be a family of open filter bases on a space X. Let $\{p(\mathscr{F}) | \mathscr{F} \in \mathscr{M}\}$ be a new set of distinct points, and let $X(\mathscr{M})$ be the space whose points are the elements of $X \cup \{p(\mathscr{F}) | \mathscr{F} \in \mathscr{M}\}$ and whose topology has as a base sets of the form $V^* = V \cup \{p(\mathscr{F}) | V$ contains some member of $\mathscr{F}\}$, where V is any open subset of X.

THEOREM 1.1. *Let X be an $\mathscr{H}$-space containing a point a such that X-$\{a\}$ is a zero-dimensional locally pseudocompact space. Let $\mathscr{N} = \{\mathscr{F} | \mathscr{F}$ is a free, countable, pseudocompact, zero-dimensional filter base on $X\}$, and denote by $\mathscr{M}$ a maximal subset of $\mathscr{N}$ such that whenever $\mathscr{F}, \mathscr{G} \in \mathscr{M}$ with $\mathscr{F} \neq \mathscr{G}$, then there exist disjoint sets $F \in \mathscr{F}$ and $G \in \mathscr{G}$.*

Then the space $X(\mathscr{M})$ is an $\mathscr{H}$-closed $\mathscr{C}$-space in which X is embedded as a dense subset, and $X(\mathscr{M})$ is $\mathscr{H}$-minimal if and only if X is semiregular.

Proof. $X(\mathscr{M})$ is clearly an $\mathscr{H}$-space. Furthermore, it follows from the hypothesis that each point of $X(\mathscr{M}) - \{a\}$ has a fundamental system of feebly compact open neighborhoods. Thus the characteristic functions of open-and-closed subsets of $X(\mathscr{M})$ separate the points of $X(\mathscr{M})$ and $X(\mathscr{M})$ is a $\mathscr{C}$-space.

Suppose that $\mathscr{F}$ is a countable open filter base on $X(\mathscr{M})$ and no point of X is an adherent point of $\mathscr{F}$. A slight modification of the proof of Lemma 2.17 in [11] shows that there exists a free, countable, pseudocompact, zero-dimensional filter base $\mathscr{G}$ on X which is stronger than the filter base $\mathscr{F} | X$. By the maximality of $\mathscr{M}$, there exists $\mathscr{K} \in \mathscr{M}$ with $G \cap H$ nonempty for all $G \in \mathscr{G}$ and $H \in \mathscr{K}$. Thus $p(\mathscr{K})$ is an adherent point of $\mathscr{F}$.

To check semiregularity, it suffices to observe that if

$$a \in V = \mathrm{Int}_X Cl_X V, \text{ then } V^* = \mathrm{Int}_{X(\mathscr{M})} Cl_{X(\mathscr{M})} V^*.$$

THEOREM 1.2. *Let X and a be as in Theorem* 1.1, *and suppose that $\{V_n | n \in N\}$ is a fundamental system of open neighborhoods for a such that $V_1 = X$ and each $V_n \supset Cl_X V_{n+1}$. Let $\mathscr{M}$ be a maximal family of free, countable, pseudocompact, zero-dimensional filter bases on X such that* (a) *whenever $\mathscr{F}, \mathscr{G} \in \mathscr{M}$ with $\mathscr{F} \neq \mathscr{G}$, then there* *exist disjoint sets* $F \in \mathscr{F}$ *and* $G \in \mathscr{G}$, *and* (b) *for every* $\mathscr{F} \in \mathscr{M}$ *there exists* $n \in N$ *such that* $\cup \mathscr{F} \subset V_n - V_{n+1}$.

Then $X(\mathscr{M})$ *is a regular* $\mathscr{C}$*-space that is* $\mathscr{H}$*-minimal and contains* X *as a dense subspace. If each* V_n *is closed in* X, *then* $X(\mathscr{M})$ *is zero-dimensional.*

Proof. Since $\{p(\mathscr{F}) | \mathscr{F} \in \mathscr{M}\} - \{a\}$ is a closed discrete subset of $X(\mathscr{M}) - \{a\}$, it follows from (b) that $Cl_{X(\mathscr{M})} V^*_{n+1} = V^*_{n+1} \cup Cl_X V_{n+1}$. Thus $X(\mathscr{M})$ is regular, and if each V_n is closed in X, then $X(\mathscr{M})$ is zero-dimensional.

The proof that $X(\mathscr{M})$ is feebly compact is similar to the corresponding proof given for Theorem 1.1-one just notes that for some n, $\mathscr{F} | (Cl_X V_n - Cl_X V_{n+1})$ is a filter base, and so $\mathscr{G}$ can be chosen with the property that $\cup \mathscr{G} \subset V_n - V_{n+1}$.

REMARK 1.3. In case the set I of isolated points of X is a dense subset of X, $\mathscr{M}$ can be defined as follows. Let $\mathscr{E}$ be a maximal family of countably infinite subsets of I such that (a) the intersection of any two members of $\mathscr{E}$ is finite, and (b) each member of $\mathscr{E}$ is a closed subset of X (for Theorem 1.2, a closed subset of some $Cl_X(V_n - V_{n+1})$). For each $E \in \mathscr{E}$ let $\mathscr{F}(E)$ be the complements in E of finite subsets of E. Take $\mathscr{M} = \{\mathscr{F}(E) | E \in \mathscr{E}\}$.

REMARK 1.4. For the case $X = N$ and $\mathscr{M}$ infinite, the space $X(\mathscr{M})$ is due to J. Isbell (see [5, 5I]).

REMARK 1.5. In general, the space $X(\mathscr{M})$ is not countably compact and hence not weakly normal, for each $\{p(\mathscr{F}) | \mathscr{F} \in \mathscr{M}\} - V^*_n$ is a closed discrete subset of $X(\mathscr{M})$.

COROLLARY 1.6. *Every locally pseudocompact zero-dimensional* $\mathscr{H}$*-space can be embedded densely in a pseudocompaet zero-dimensional* $\mathscr{H}$*-space.*

EXAMPLE 1.7. For the following X, the space $X(\mathscr{M})$ is an $\mathscr{H}$-minimal $\mathscr{C}$-space that is not regular.

Let $T = \{0\} \cup \{1/n \in N\}$, with the usual topology, choose a point a not in the product space $N \times T$, and let $X = \{a\} \cup (N \times T)$, topologized as follows: every open subset of $N \times T$ is open in X; a neighborhood of a is any set of the form $V_n = \{a\} \cup \{(x, y) \in X | x \geqq n$ and $1/y$ is an even integer}, $n \in N$. (X is homeomorphic to $E - \{b\}$, where E is as in [13, p. 268].)

One can take $\mathscr{M}$ to be a maximal family of infinite subsets of $X - Cl\, V_1$ such that the following hold:

(i) For all $M, M' \in \mathscr{M}$, $M \neq M'$ implies $M \cap M'$ is finite;

(ii) For all $M \in \mathscr{M}$ and $n \in N$, $M \cap (\{n\} \times T)$ is finite.

EXAMPLE 1.8. For the following X, the space $X(\mathscr{M})$ (of Theorem 1.2) is an $\mathscr{H}$-minimal $\mathscr{C}$-space that is regular but not completely regular.

Let Y be the set of ordinal numbers less than the first uncountable ordinal, with the order topology, let M be the set of limit ordinals in Y, and denote $Y - M$ by I. Let $Z = I \times \{0\} \cup Y \times N$, topologized as follows: $Y \times N$ has the product topology, and $Y \times N$ is open in Z; a neighborhood of a point $(i, 0) \in Z$ is any subset of Z that contains $(i, 0)$ and all but finitely many elements of $\{i\} \times N$. Let L and R denote the product spaces $Z \times \{1\}$ and $Z \times \{2\}$, and set $U = L \cup R$, with the weak topology generated by $\{L, R\}$. Let S be the relation on U defined by the rule: $(x, i, j) S (y, k, n)$ if (a) $x = y$, $i = k$, and $j = n$, or (b) $x = y \in M$ and $i = k$. Denote the quotient space U/S by T. We shall continue to use the symbols (x, i, j) for the points of T.

On the product space $T \times N$ define $(t, n)\, W (t', n')$ if (a) $t = t'$ and $n = n'$, or (b) $t = (x, 0, j)$, $t' = (x, 0, j')$, and $n' - n = j - j' = 1$ or $n - n' = j' - j = 1$. Let V be the quotient space $(T \times N)/W$. Choose a new point a and let $X = V \cup \{a\}$, topologized as follows: every open subset of V is open in X; a neighborhood of a is any set of the form $V_n = \{a\} \cup \{(t, m) \in V | m \geqq n\}$, $n \in N$.

It is not difficult to see that X is a first countable regular space whose isolated points are dense, and $X - \{a\}$ is zero-dimensional and locally compact. X is not completely regular, because for every $f \in C(X)$ there exists $m \in Y$ such that f is constant on

$$\{(x, 0, j, n) | x \geqq m, j = 1 \text{ or } j = 2, \text{ and } n \in N\} .$$

Thus V_2, for example, contains no zero set neighborhood of a.

REMARK 1.9. The construction above is a modification of Tychonoff's regular but not completely regular space [12].

In [7] F. B. Jones has constructed a $\mathscr{C}$-space that is not completely regular but that is a Moore space. His space cannot be used here, however, because it is neither locally pseudocompact nor zero-dimensional.

In the literature there are many less messy examples of $\mathscr{C}$-closed or $\mathscr{H}$-minimal spaces that are not regular; however, the author does not know of any $\mathscr{C}$-minimal space appearing elsewhere that is not regular (or completely regular).

REMARK 1.10. If one glues together (as in [2]) two copies of the space in Example 1.8, then one gets an example of a regular $\mathscr{H}$-minimal space that is not completely Hausdorff.

2. $\mathscr{C}$-minimal spaces and closed mappings. If $\mathscr{P}$ denotes any one of the usual separation properties, it is known that every $\mathscr{P}$-minimal completely Hausdorff space is compact (e.g., see [6]). Moreover C. T. Scarborough [9] has observed that a completely Hausdorff-minimal space is compact.

One might then expect $\mathscr{C}$-minimal spaces to be well behaved, to be, say, at least countably compact. Of course, Isbell's example or Mrówka's [8] (or ours) shows that this is not the case. The following characterization theorems may, therefore, be of interest.

DEFINITION. (H. E. Hayes) An open filter base $\mathscr{F}$ on a space X is said to be *completely Hausdorff* provided that for every $x \in X$, if x is not an adherent point of $\mathscr{F}$, then there exist $f \in C(X)$ and $F \in \mathscr{F}$ such that $f(F) = 0$ and $f(x) = 1$.

Using usual techniques, one can prove the following.

THEOREM 2.1. *Let X be a $\mathscr{C}$-space. The following are equivalent.*

(i) *X is $\mathscr{C}$-closed.*

(ii) *Every countable completely Hausdorff filter base on X has an adherent potnt.*

(iii) *For every $\mathscr{C}$-space Y and $f \in C(X, Y), f(X)$ is $\mathscr{C}$-closed.*

In order to obtain a $\mathscr{C}$-analogue of Theorem 2.4 of [11], we need a second definition.

DEFINITION. An open filter base $\mathscr{F}$ on a space X is said to be *almost completely Hausdorff* if there exists $p \in X$ so that for every $x \in X - \{p\}$, if x is not an adherent point of $\mathscr{F}$, then there exist $f \in C(X)$ and $F \in \mathscr{F}$ such that $f(F) = 0$ and $f(x) = 1$.

THEOREM 2.2. *Let X be a $\mathscr{C}$-space. The following are equivalent.*

(i) *X is $\mathscr{C}$-minimal.*

(ii) *Every countable completely Hausdorff filter base on X that has a unique adherent point is convergent.*

(iii) *X is semiregular, and every countable almost completely Hausdorff filter base on X has an adherent point.*

The proof is somewhat similar to the proofs needed for Theorems 2.4 and 2.9 in [11].

The next result, to be contrasted with (iii) of Theorem 2.1, is a partial converse to the following well-known theorem: If X is a countably compact space, Y is an $\mathscr{H}$-space (or a space of the type E_1 studied in [1]), and $f \in C(X, Y)$, then f is closed.

We shall call an open filter base $\mathscr{F}$ on X *completely regular* if for each $F \in \mathscr{F}$ there exist $G \in \mathscr{F}$ and $f \in C(X, [0, 1])$ such that f vanishes on G and equals 1 on $X - F$.

THEOREM 2.3. *Let $\mathscr{P}$ denote either completely Hausdorff, completely regular, or zero-dimensional, and suppose that X is a $\mathscr{P}$-space which is also an $\mathscr{H}$-space. The following are equivalent.*

(i) *X is countably compact.*

(ii) *For every $\mathscr{H}$-space Y and $f \in C(X, Y)$, f is closed.*

(iii) *For every $\mathscr{P}$-space Y that is an $\mathscr{H}$-space and $f \in C(X, Y)$, f is closed.*

(iv) *For every closed subset C of X and every countable $\mathscr{P}$-filter base $\mathscr{F}$ on X, if $\mathscr{F}|C$ is a filter base and if $\cap\mathscr{F} = \cap\{\bar{F}|F \in \mathscr{F}\}$, then there is a point $c \in C$ which is in $\cap\mathscr{F}$.*

Proof. (i) $\Rightarrow$ (ii) is known. (ii) $\Rightarrow$ (iii) is obvious. A proof not too different from one in [3] shows that (iii) $\Leftrightarrow$ (iv). We shall prove that (iv) $\Rightarrow$ (i) for the case $\mathscr{P}$ = completely Hausdorff.

Let us suppose then that X is a $\mathscr{C}$-space which contains a countably infinite closed discrete subset C.

Consider a point $c \in C$. Since X is completely Hausdorff and $C - \{c\}$ is countable, there exists $f \in C(X)$ for which $f(c) \notin f(C - \{c\})$. Since $C - \{c\}$ is a closed subset of X and f is closed, we can choose $g \in C((-\infty, \infty))$ with $g(f(c)) = 1$ and $g(f(C - \{c\})) = 0$. Set $h_c = g \circ f$.

Let $\mathscr{F}$ be the family of all finite intersections of

$$\{h_c^{-1}(-1/n, 1/n)|n \in N \text{ and } c \in C\} .$$

Then it is easy to see that $\mathscr{F}$ is a countable completely regular (and hence completely Hausdorff) filter base on X, that $\cap \mathscr{F} = \cap\{\bar{F} \mid F \in \mathscr{F}\}$, and that $\mathscr{F} \mid C$ is a filter base. On the other hand, one also has $C \cap \cap \mathscr{F} = \phi$. This contradicts (iv).

REMARK 2.4. There exists an $\mathscr{H}$-space X that is not countably compact but which has the property: for every Hausdorff space Y and $f \in C(X, Y)$, f is closed. See [3] and [14].

REFERENCES

1. C. E. Aull, *A certain class of topological spaces*, Prace Mat. **11** (1967), 49-53.
2. M. P. Berri and R. H. Sorgenfrey, *Minimal regular spaces*, Proc. Amer. Math. Soc., **14** (1963), 454-458.
3. R. F. Dickman, Jr. and Alan Zame, *Functionally compact spaces*, Pacific J. Math., **31** (1969), 303-311.
4. J. Dugundji, *Topology*, Allyn and Bacon, Boston, Mass., 1966.
5. L. Gillman and M. Jerison, *Rings of continuous functions*, Van Nostrand, New York, 1960.
6. H. Herrlich, *T_ν-Abgeschlossenheit und T_ν-Minimalität*, Math. Z., **88**(1965), 285-294.
7. F. B. Jones, *Moore spaces and uniform spaces*, Proc. Amer. Math. Soc., **9** (1958), 483-486.
8. S. Mrówka, *On completely regular spaces*, Fund. Math., **41** (1954), 105-106.
9. C. T. Scarborough and R. M. Stephenson, Jr., *Minimal topologies*, Colloq. Math., **19** (1968), 215-219.
10. C. T. Scarborough and A. H. Stone, *Products of nearly compact spaces*, Trans. Amer. Math. Soc., **124** (1966), 131-147.
11. R. M. Stephenson, Jr., *Minimal first countable topologies*, Trans. Amer. Math. Soc., **138** (1969), 115-127.
12. A. Tychonoff, *Über die topologische Erweiterung von Räumen*, Math. Ann., **102** (1930), 544-561.
13. P. Urysohn, *Über die Machtigkeit der zusammenhängenden Mengen*, Math. Ann., **94** (1925), 262-295.
14. Giovanni A. Viglino, *C-Compact spaces*, Duke Math. J., **36** (1969), 761-764.

Received February 17, 1970. This research was partially supported by a grant from the University of North Carolina.

UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

THE QUOTIENT ALGEBRA OF A FINITE VON NEUMANN ALGEBRA

MASAMICHI TAKESAKI

We will prove the following: Let M be a finite von Neumann algebra with center Z and A a von Neumann subalgebra of Z. Let Ω be the spectrum space of A and identify A with $C(\Omega)$. Let ε be a σ-weakly continuous linear map of M onto A such that $\varepsilon(x^*x) = \varepsilon(xx^*) \geqq 0$ for every $x \in M$, $\varepsilon(ax) = a\varepsilon(x)$ for every $a \in A$ and $x \in M$, $\varepsilon(1) = 1$ and $\varepsilon(x^*x) \neq 0$ for every nonzero $x \in M$. For each $\omega \in \Omega$, let $\mathfrak{m}_\omega$ denote the set of all $x \in M$ with $\varepsilon(x^*x)(\omega) = 0$. Then $\mathfrak{m}_\omega$ is a closed ideal and the quotient C^*-algebla $M/\mathfrak{m}_\omega$ is a finite von Neumann algebra. Furthermore, if π_ω denote the canonical homomorphism of M onto $M/\mathfrak{m}_\omega$, then $\pi_\omega(N)$ is a von Neumann subalgebra of $M/\mathfrak{m}_\omega$ for every von Neumann subalgebra N containing A.

In [8], [3] and [5] it was shown that the quotient C^*-algebra of a finite von Neumann algebra by any maximal ideal is actually a finite factor. This led us to the algebraic reduction theory for finite von Neumann algebras, which is free from the separability restriction in the direct integral reduction theory. In this paper we will show that the above result still holds for certain ideals, not necessarily maximal. Namely, we will give a straightforward proof for the following.

THEOREM. *Let M be a finite von Neumann algebra with center Z and A a von Neumann subalgebra of Z. Let Ω be spectrum space of A and identify A with $C(\Omega)$. Let ε be a σ-weakly continuous linear map of M onto A such that $\varepsilon(x^*x) = \varepsilon(xx^*) \geqq 0$ for every $x \in M$, $\varepsilon(ax) = a\varepsilon(x)$ for every $a \in A$ and $x \in M$, $\varepsilon(1) = 1$ and $\varepsilon(x^*x) \neq 0$ for every nonzero $x \in M$. For each $\omega \in \Omega$, let $\mathfrak{m}_\omega$ denote the set of all $x \in M$ with $\varepsilon(x^*x)(\omega) = 0$. Then $\mathfrak{m}_\omega$ is a closed ideal and the quotient C^*-algebra $M/\mathfrak{m}_\omega$ is a finite von Neumann algebra. Furthermore, if π_ω denote the canonical homomorphism of M onto $M/\mathfrak{m}_\omega$, then $\pi_\omega(N)$ is a von Neumann subalgebra of $M/\mathfrak{m}_\omega$ for every von Neumann subalgebra N containing A.*

Before going into the proof, we observe that there exists such a map ε if Z is σ-finite. Since M has the $\natural$-operation, it suffices to show that there exists a σ-weakly continuous faithful projection of norm one from Z onto A. If Z is σ-finite, then Z admits a faithful normal state φ. Considering the cyclic representation of Z induced by φ, we may assume that Z acts on a Hilbert space $\mathscr{H}$ containing a vector ξ_0 such that $(x\xi_0 \mid \xi_0) = \varphi(x)$, $x \in Z$. Let e be the projection of $\mathscr{H}$ onto $[A\xi_0]$. Then e is an abelian projection in A' with central support 1. Note that the center of A' is A itself. Then there exists an isomorphism θ of $eA'e$ onto A such that $\theta(xe) = x$ for every $x \in A$ because A is the center of A'. Put $\varepsilon_Z(x) = \theta(exe)$ for every $x \in Z$. Since e is not orthogonal to any nonzero projection in Z, ε_Z has the required properties. As the composed map of this ε_Z and the $\natural$-operation in M, we get a desired map ε. Hence, the situation in the theorem is always presented for any von Neumann subalgebra A of Z if Z is σ-finite.

The proof of theorem. We will prove the assertion for the subalgebra N which implies immediately the former assertion.

Let $\tau_\omega(x) = \varepsilon(x)(\omega)$, $x \in M$. Then τ_ω is a finite trace of M with the left kernel $\mathfrak{m}_\omega$. Let $\{\pi, \mathscr{H}, \xi_0\}$ be the cyclic representation of M induced by τ_ω. Since π has the kernel $\mathfrak{m}_\omega$, π induces a faithful representation $\tilde{\pi}$ of the C^*-algebra $M/\mathfrak{m}_\omega$. Since $\tilde{\pi} \circ \pi_\omega(N) = \pi(N)$, it suffices to show that $\pi(N)$ is a von Neumann algebra. Since the functional $\tau_\omega(x) = (x\xi_0 \mid \xi_0)$, $x \in \pi(M)''$, is a faithful trace on the von Neumann algebra $\pi(M)''$, ξ_0 is a cyclic and separating for $\pi(M)''$. Let S_N denote the unit ball of N. Then by Kaplansky's density theorem $\pi(S_N)$ is strongly dense in the unit ball $S_{\tilde{N}}$ of the von Neumann algebra $\tilde{N} = \pi(N)''$ generated by $\pi(N)$. Since the map $x \in \pi(M)'' \rightarrow x\xi_0$ is injective, if $\pi(S_N)\xi_0 = S_{\tilde{N}}\xi_0$, then we have $\pi(S_N) = S_{\tilde{N}}$; hence $\tilde{N} = \pi(N)$.

Therefore, we shall prove that $\pi(S_N)\xi_0$ is complete. Let $\{x_n\}$ be a sequence in S_N such that

$$\lim_{n,m\to\infty} \| \pi(x_n)\xi_0 - \pi(x_m)\xi_0 \| = 0 .$$

Considering a subsequence of $\{x_n\}$, we may assume that

$$\| \pi(x_n)\xi_0 - \pi(x_{n+1})\xi \| < 2^{-n} , \qquad n = 1, 2, \cdots .$$

In other words,

$$\varepsilon((x_n - x_{n+1})^* (x_n - x_{n+1}))(\omega) < 4^{-n} , \qquad n = 1, 2, \cdots .$$

Let $\{U_n\}$ be a decreasing sequence of neighborhoods of ω in Ω such that

$$\varepsilon((x_n - x_{n+1})^* (x_n - x_{n+1}))(\sigma) < 4^{-n}$$

for every $\sigma \in U_n$, $n = 1, 2, \cdots$. For each $n = 1, 2, \cdots$, let e_n be the projection of A corresponding to the closure of U_n. Then $e_n(\omega) = 1$

for $n = 1, 2, \cdots$. Putting $y_1 = x_1$ and $y_n = e_n x_n + (1 - e_n) y_{n-1}$ for $n = 2, 3, \cdots$ by induction,

$$\varepsilon((y_n - y_{n+1}) * (y_n - y_{n+1})) < 4^{-n} ;$$
$$\pi(y_n)\xi_0 = \pi(x_n)\xi_0 , \qquad n = 1, 2, \cdots .$$

Now, for any normal state φ of A, put $\tau_\varphi(x) = \varphi \circ \varepsilon(x)$, $x \in N$. Then τ_φ is a normal finite trace of N with the support $s(\varphi) \in A$, where $s(\varphi)$ means the support of φ in A. By the inequality:

$$\tau_\varphi((y_n - y_{n+1}) * (y_n - y_{n+1})) = \varphi \circ \varepsilon((y_n - y_{n+1}) * (y_n - y_{n+1})) < 4^{-n} ,$$

$n = 1, 2, \cdots$, $\{y_n s(\varphi)\}$ converges σ-strongly to $y_\varphi \in S_N$ because the σ-strong topology in $S_N \cap Ns(\varphi)$ is induced by the metric d defined by $d(x, y) = \tau_\varphi((x - y) * (x - y))^{1/2}$, $x, y \in S_N \cap Ns(\varphi)$. Let $\{\varphi_i\}_{i \in I}$ be a maximal family of normal states of A with orthogonal supports. Then $\sum_{i \in I} s(\varphi_i) = 1$. Let $y = \sum_{i \in I} y_{\varphi_i} \in S_N$. Since $\{y_n s(\varphi_i)\}$ converges σ-strongly to $s(\varphi_i) y = y_{\varphi_i}$ for each $i \in I$, $\{y_n\}$ converges σ-strongly to y. Now we have, by the triangular inequality,

$$\begin{aligned} \varepsilon((y_n - y_{n+p}) * (y_n - y_{n+p}))^{1/2} &\leqq \sum_{k=n}^{n+p-1} \varepsilon((y_k - y_{k+1}) * (y_k - y_{k+1}))^{1/2} \\ &\leqq \sum_{k=n}^{n+p-1} 2^{-k} \leqq 2^{-n+1} \end{aligned}$$

for n, $P = 1, 2, \cdots$. Hence we have

$$\varepsilon((y_n - y) * (y_n - y))^{1/2} = \lim_{p \to \infty} \varepsilon((y_n - y_{n+p}) * (y_n - y_{n+p}))^{1/2} \leqq 2^{-n+1} ,$$

so that

$$\| \pi(y_n)\xi_0 - \pi(y)\xi_0 \| = \varepsilon((y_n - y) * (y_n - y))(\omega)^{1/2} \leqq 2^{-n+1} ;$$

hence

$$\lim_{n \to \infty} \pi(y_n)\xi_0 = \pi(y)\xi_0 .$$

Therefore, the given Cauchy sequence $\{\pi(x_n)\xi_0\}$ in $\pi(S_N)\xi_0$ converges to $\pi(y)\xi_0 \in \pi(S_N)\xi_0$. Hence $\pi(S_N)\xi_0$ is complete, hence closed in $\mathscr{H}$. This completes the proof.

By [7], we should remind that if M is a von Neumann algebra of type II_1 and if ω is not an isolated point of Ω then $M/\mathfrak{m}_\omega$ does not admit nontrivial representation on a *separable* Hilbert space even if M does have faithful normal representation on a *separable* Hilbert space.

Suppose now A is σ-finite and ω is not an isolated point of Ω.

Suppose that any nonzero projection $e \in N$ majorizes a projection $f \in N$ such that $\varepsilon(f) = \varepsilon(e - f)$. Then we claim that the von Neumann algebra $\pi_\omega(N)$ does not admit a faithful separable normal representation.

Let $\{e_n\}$ be a decreasing sequence of projections in A converging σ-strongly to zero such that $e_n(\omega) = 1$ for $n = 1, 2, \cdots$. Such a sequence does exist by the nonisolatedness of ω and the σ-finiteness of A. Let $f_n = e_n - e_{n+1}$ for $n = 1, 2, \cdots$. By the assumption for N, there exists orthogonal projections $p^n_{1,1}$ and $p^n_{1,2}$ in N such that $f_n = p^n_{1,1} + p^n_{1,2}$ and $\varepsilon(p^n_{1,1}) = \varepsilon(p^n_{1,2}) = \frac{1}{2} f_n$. Suppose we have found projections $\{p^n_{i,j}: i = 1, \cdots, k, j = 1, 2, \cdots, 2^i\}$ such that

(1) for fixed i, $\{p^n_{i,j}: j = 1, \cdots, 2^i\}$ are orthogonal;

(2) $p^n_{i-1,j} = p^n_{i,2j-1} + p^n_{i,2j}$;

(3) $\varepsilon(p^n_{i,j}) = 2^{-i} f_n$.

By the assumption for N, we can find orthogonal projections $\{p^n_{i+1,j}: j = 1, 2, \cdots, 2^{i+1}\}$ such that

$$p^n_{i,j} = p^n_{i+1,2j-1} + p^n_{i+1,2j} \; ;$$
$$\varepsilon(p^n_{i+1,j}) = 2^{-(i+1)} f_n \, , \qquad j = 1, 2, \cdots, 2^{i+1} \, .$$

For each integer i, put

$$u_{n,i} = \sum_{j=1}^{2^n} (-1)^j p^n_{i,,j} \, .$$

Then we have $u^2_{n,k} = f_n$ and for different i_1 and i_2, $u_{n,i_1} u_{n,i_2}$ is the difference of two orthogonal projections p and q such that $\varepsilon(p) = \varepsilon(q) = \frac{1}{2} f_n$; hence $\varepsilon(u_{n,i_1} u_{n,i_2}) = 0$ if $i_1 \neq i_2$.

To each real number s we associate a sequence $\{i_{s,n}\}$ of integers such that

$$\lim_{n \to \infty} \frac{i_{s,n}}{2^n} = s \, .$$

If $s \neq t$, there is an n_0 such that $i_{s,n} \neq i_{t,n}$ for every $n \geqq n_0$. Put

$$u_s = \sum_{n=1}^{\infty} u_{n, i_{s,n}} \, .$$

Then we have $\varepsilon(u_s u_t)(1 - e_{n_0}) = \varepsilon(u_s u_t)$. Therefore we have

$$\tau_\omega(u_s^2) = 1, \, \tau_\omega(u_s u_t) = 0 \quad \text{if} \quad s \neq t \, .$$

Therefore $\{\pi(u_s)\xi_0\}$ is a continuum of orthogonal vectors in $[\pi(N)\xi_0]$. Therefore, the standard representation of the von Neumann algebra $\pi_\omega(N)$ is not separable. Thus $\pi_\omega(N)$ does not admit a faithful normal separable representation.

Now, let A and B be two abelian von Neumann algebras with no minimal projections. Let C be the tensor product $A \otimes B$ of A and B. Then A and B are regarded as subalgebras of C. If B admits a faithful normal state ψ, then there exists a faithful normal projection ε of norm one of C onto A defined by

$$\langle \varepsilon(x), \varphi) \rangle = \langle x, \varphi \otimes \psi \rangle$$

for every $\varphi \in A_*$. This map has the property:

$$\varepsilon(a \otimes b) = \varphi(b)a, a \in A, b \in B .$$

If A is σ-finite, then $C/\mathfrak{m}_\omega$ is an abelian von Neumann algebra, with no separable faithful normal representation. It is easily seen that the map π_ω is σ-weakly continuous on B; hence $\pi_\omega(B)$ is a proper von Neumann subalgebra of $C/\mathfrak{m}_\omega$ if B has a faithful separable normal representation. Therefore, *the pathology that the component algebras are much larger than the synthetic algebra does occur even in the abelian case.*

REFERENCES

1. J. Dixmier *Les algèbras d'òperateurs dans les espace hilbertièn*, 2nd. ed., 1969, Gauthier-Villars.
2. J. Feldman, *Embedding of AW*-algebras*, Duke Math. J., **23** (1956), 303-308.
3. ———, *Nonseparability of certain finite factors*, Proc. Amer. Math. Soc., **7** (1956), 23-26.
4. J. Feldman and J. M. G. Fell, *Separable representations of rings of operators*, Amer. Math. Soc., **65** (1957), 241-249.
5. S. Sakai, *The theory of W*-algebras*, Lecture Notes, Yale Univ., 1962.
6. H. Takemoto, *On the homomorphism of von Neumann algebra*, Tôhoku Math. J., **21** (1969), 152-157.
7. M. Takesaki, *On the nonseparability of singular representations of operator algebra*, Kodai, Math. Sem. Rep. **12** (1960), 102-108.
8. F. B. Wright, *A reduction for algebras of finite type*, Ann. Math. **60** (1954), 560-570.

Received June 8, 1970.

UNIVERSITY OF CALIFORNIA, LOS ANGELES, CALIFORNIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

INTERPOLATION IN $C(\Omega)$

BENJAMIN B. WELLS, JR.

It is known from the work of Bade and Curtis that if $\mathfrak{A}$ is a Banach subalgebra of $C(\Omega)$, Ω a compact Hausdorff space, and if Ω is an F-space in the sense of Gillman and Hendriksen then $\mathfrak{A} = C(\Omega)$. This paper is concerned with the extension of this and similar results to the setting of Grothendieck spaces (G-spaces for short). An important feature of the extension is that emphasis is shifted from the underlying topological structure of Ω to the linear topological character of $C(\Omega)$.

As a corollary we show that if Ω_1 and Ω_2 are infinite compact Hausdorff spaces, then $\Omega_1 \times \Omega_2$ is not a G-space. Consequently if Ω is a G-space then $C(\Omega)$ is not linearly isomorphic to $C(\Omega \times \Omega)$.

If A is a commutative Banach algebra whose spectrum is a totally disconnected G-space, a second corollary of our extension is that the Gelfand homomorphism is onto. This establishes for G-spaces a result due to Seever for N-spaces.

Two definitions of G-space are to be found in the literature.

(A) A Banach space X is a G-space if every weak-* convergent sequence in X^*, the dual of X, is weakly convergent.

(B) A compact Hausdorff space Ω is a G-space if $C(\Omega)$ is a G-space in the sense of (A).

Unless otherwise noted we shall accept (B) as our definition.

It is known from the work of Seever [7] that if Ω is an F-space, i.e., if disjoint open F_σ subsets of Ω have disjoint closures, then Ω is a G-space. A result due to Rudin [3] states that if Ω_1 and Ω_2 are infinite compact Hausdorff spaces then $\Omega_1 \times \Omega_2$ is not an F-space. Corollary 2.6 is an extension of this to G-spaces. Although an example of a G-space which is not an F-space is given in [7], no necessary and sufficient topological characterization of the G property is known.

1. Preliminaries. Let $M(\Omega)$ be the space of regular Borel measures on Ω equipped with the total variation norm. A sequence $\{\mu_n\}$ in $M(\Omega)$ converges for the weak-* topology if for each f in $C(\Omega)$, the space of continuous complex valued functions on Ω, the sequence $\{\mu_n(f)\}$ is convergent. Weak convergence of $\{\mu_n\}$ means convergence of $\{\gamma(\mu_n)\}$ for every γ in $M^*(\Omega)$, the dual of $M(\Omega)$. If Ω is any set $l_1(\Omega)$ will denote the Banach space of point mass measures on Ω with the total variation norm.

A Banach subalgebra (subspace) $\mathfrak{A}$ of $C(\Omega)$ is a subalgebra (subspace) of $C(\Omega)$ under the pointwise operations and is a Banach algebra (space) such that the embedding $\mathfrak{A} \rightarrow C(\Omega)$ is continuous. $\mathfrak{A}$ is said to be normal if for each pair F_1, F_2 of disjoint compact subsets of Ω there is an $f \in \mathfrak{A}$ such that $f = 1$ on F_1 and $f = 0$ on F_2. Following [2] we call $\mathfrak{A}$ ε-normal if for each pair F_1, F_2 of disjoint compact subsets of Ω there exists an $f \in \mathfrak{A}$ satisfying

(i) $|f(\omega) - 1| < \varepsilon, \omega \in F_1$,
(ii) $|f(\omega)| < \varepsilon, \omega \in F_2$.

If Ω_1 and Ω_2 are compact Hausdorff spaces the projective tensor product $V = C(\Omega_1) \widehat{\bigoplus} C(\Omega_2)$ is the set of all functions of the form

$$\sum_{i=1}^{\infty} f_1(x)g_i(y),\ f_i(x) \in C(\Omega_1)$$

and $g_i(y) \in C(\Omega_2)$ such that $\sum_{i=1}^{\infty} \| f_i \|_\infty \| g_i \|_\infty < \infty$. If $h \in V$ then

$$\| h \|_V = \inf\left\{\sum_{i=1}^{\infty} \| f_i \|_\infty \| g_i \|_\infty : h = \sum_{i=1}^{\infty} f_i g_i\right\} .$$

Two Banach spaces X_1 and X_2 are isomorphic if there is a one-to-one continuous linear map from X_1 onto X_2. If X_2 is a closed subspace of X_1, it is said to be complemented in X_1 if there exists a closed subspace Y of X_1 such that $X_2 + Y = X_1$ and $X_2 \cap Y = \{0\}$. We write $X_1 = X_2 \bigoplus Y$.

If D is a discrete space, $C(D)$ will denote the bounded continuous functions on D. It is well known that $C(D)$ is isometrically isomorphic to $C(\beta D)$ where βD is the Stone-Čech compactification of D. A compact Hausdorff space is totally disconnected if there is a basis for the topology consisting of open and closed neighborhoods.

2. We shall need to recall here a criterion due to Grothendieck [5] for relative weak compactness in $M(\Omega)$. Namely, a bounded sequence $\{\mu_n\}$ in $M(\Omega)$ is relatively weakly compact if and only if for every sequence $\{0_i\}$ of pairwise disjoint Borel sets $\lim_{i\to\infty} \mu_n(0_i) = 0$ uniformly in n. By the Eberlein Smulian theorem this is equivalent to every subsequence of $\{\mu_n\}$ having a weakly convergent subsequence.

LEMMA 2.1. *If Ω is a G-space and K is a closed subspace of Ω, then K is a G-space.*

Proof. Suppose $\{\mu_n\}$ in $M(K)$ is weak-* convergent. One may regard $\{\mu_n\}$ as a weak-* convergent sequence in $M(\Omega)$. It is therefore weakly convergent as a sequence in $M(\Omega)$, and so by the Hahn-Banach Theorem it is a weakly convergent sequence in $M(K)$.

LEMMA 2.2. *Let Ω be a G-space and X a dense Banach subspace such that $X \neq C(\Omega)$. Then for every $M > 0$ there is a measure μ with no atomic part such that* $\|\mu\| \geqq M$ *and* $\sup\{|\mu(f)| : f \in X, \|f\|_X \leqq 1\} \leqq 1$.

Proof. We shall write μ_a for the atomic part of μ and μ_c for the continuous part. By a well known theorem of Banach there is a sequence $\{\mu_n\}$ of measures such that $\|\mu_n\| \geqq n$ and sup

$$\{|\mu_n(f)| : f \in X, \ \|f\|_X \leqq 1\}$$

for each n. Since X is dense in $C(\Omega)$ setting $\nu_n = \mu_n/\|\mu_n\|$ we have $\lim_n \nu_n = 0$ weak-* and hence $\lim_n \nu_n = 0$ weakly since Ω is a G-space. The natural projection $p: M(\Omega) \to l_1(\Omega)$ given by $p\mu = \mu_a$ is continuous and hence weakly continuous. Hence $\lim_n \nu_{n,a} = 0$ weakly. Since in $l_1(\Omega)$ weakly convergent sequences are norm convergent, it follows that $\lim_n \|\nu_{n,a}\| = 0$. Thus for an appropriate sequence of scalars $\{c_n\}$ we have $\lim_n \|c_n \nu_{n,c}\| = \infty$ and

$$\sup\{|c_n \nu_{n,c}(f)| : f \in X, \ \|f\|_X \leqq 1\} \leqq 1$$

for every n.

THEOREM 2.3. *Let Ω be a G-space and let X be a dense Banach subspace of $C(\Omega)$. Then there exists a finite open covering $U_1, \cdots, U_n$ of Ω such that $X|\bar{U}_i = C(\bar{U}_i)$, $1 \leqq i \leqq n$.*

Proof. From the compactness of Ω it suffices to show that each point p of Ω has a neighborhood U_p such that $X|\bar{U}_p = C(\bar{U}_p)$. Suppose this fails for some p, and choose U_1 a neighborhood of p. Let X_1 denote the quotient space of X by all functions in X vanishing on $\bar{U}_1$. Applying Lemmas 2.1 and 2.2 it follows that there is a regular Borel measure μ_1 with no atomic part such that $\|\mu_1\| \geqq 1$, supp $\mu_1 \subseteqq \bar{U}_1$ and such that $|\mu_1(f)| \leqq \|f\|_{X_1} \leqq \|f\|_X$ for every $f \in X$.

From the regularity of μ_1 we may choose open $U_2 \subseteqq U_1$, $p \in U_2$ such that $|\mu_1|(\bar{U}_1 - \bar{U}_2) > 1/2 \ \|\mu_1\|$. Since $X|\bar{U}_2 \neq C(\bar{U}_2)$ we may choose in the same way a μ_2 with no atomic part such that supp $\mu_2 \subseteqq \bar{U}_2$, $\|\mu_2\| \geqq 2$ and $|\mu_2(f)| \leqq \|f\|_X$ for all $f \in X$.

Continuing in this fashion, define inductively a sequence of measures $\{\mu_n\}$ with no atomic parts such that $\|\mu_n\| \geqq n$, $|\mu_n(f)| \leqq \|f\|_X$ for every $f \in X$, supp $\mu_n \subseteqq \bar{U}_n$ and $|\mu_n|(\bar{U}_n - \bar{U}_{n+1}) > 1/2 \ \|\mu_n\|$.

Setting $\nu_n = \mu_n/\|\mu_n\|$ we see $\lim_n \nu_n = 0$ weak-* from the density of X. However, since $|\nu_n|(\bar{U}_n - \bar{U}_{n+1}) > 1/2$ for each n, $\{\nu_n\}$ is not weakly convergent by the Grothendieck criterion. This contradiction establishes the theorem.

REMARK. Theorem 2.3 is the sharpest result in the sense that for every compact Hausdorff space Ω there is a dense Banach subspace X of $C(\Omega)$ such that $X \neq C(\Omega)$. By a result of [8] (corollary 3.2 page 201) there are closed subspaces Y, W of $C(\Omega)$ such that $Y + W$ is dense in $C(\Omega)$ but $Y + W \neq C(\Omega)$; in the terminology of that paper every $C(\Omega)$ contains a quasi-complemented uncomplemented subspace. Setting $X = Y \oplus W$ we have the result.

Our next theorem is an extension to G-spaces of a result of [2]. The work is all done by the following:

LEMMA 2.4. [2] *Let Ω be a compact Hausdorff space, and let $\mathfrak{A}$ be a Banach subalgebra of $C(\Omega)$ such that*

(i) *$\mathfrak{A}$ is ε-normal for some $\varepsilon < 1/2$,*

(ii) *There is an open covering $U_1, \cdots, U_n$ of Ω such that $\mathfrak{A} \mid \bar{U}_i = C(\bar{U}_i)$, $1 \leqq i \leqq n$.*

Then $\mathfrak{A} = C(\Omega)$.

Combining this with Theorem 2.3 and the remark that density implies ε-normality we obtain:

THEOREM 2.5. *Let Ω be a G-space, and let $\mathfrak{A}$ be a dense Banach subalgebra of $C(\Omega)$. Then $\mathfrak{A} = C(\Omega)$.*

REMARK. As demonstrated in [2] ε-normality for some $\varepsilon < 1/4$ and density of a Banach subspace of $C(\Omega)$ are equivalent in case Ω is an F-space. We do not know if "dense" may be replaced by "ε-normal" in Theorem 2.5.

COROLLARY 2.6. *If Ω_1 and Ω_2 are infinite compact Hausdorff spaces then $\Omega_1 \times \Omega_2$ is not a G-space.*

Proof. We need only take $\mathfrak{A} = C(\Omega_1) \hat{\otimes} C(\Omega_2)$ *and note that* $\mathfrak{A}$ is a dense Banach subalgebra of $C(\Omega_1 \times \Omega_2)$. ($\mathfrak{A}$ *happens to be normal as well.) But it is well known that* $\mathfrak{A} \neq C(\Omega_1 \times \Omega_2)$.

Let X_1 and X_2 be Banach spaces such that X_2 is a continuous linear image of X_1. It is an easy consequence of the Hahn Banach theorem that if X_1 is a G-space in the sense of definition A, then so is X_2. Consequently if Ω is a G-space then $C(\Omega \times \Omega)$ is not even a continuous linear image of $C(\Omega)$. This is contrasted with a result of Milutin [6, p. 42] which states that if Ω_1 and Ω_2 are uncountable compact metric spaces then $C(\Omega_1)$ is isomorphic to $C(\Omega_2)$. In particular for such Ω, $C(\Omega)$ is isomorphic to $C(\Omega \times \Omega)$.

These notions may be of use in solving complementation problems. Suppose that X_2 is a complemented subspace of X_1. Then if X_1 is a G-space in the sense of definition A, so is X_2. For example, if D

denotes an infinite discrete space, $C(\beta D \times \beta D)$ may be viewed in a natural way as a closed subspace of $C(D \times D)$. Since $\beta(D \times D)$ is a G-space, by the above remarks $C(\beta D \times \beta D)$ has no complement in $C(D \times D)$.

COROLLARY 2.7. [cf. [7] corollary 2 p. 278]. *Let A be a commutative Banach algebra whose spectrum Ω is a totally disconnected G-space. Then the Gelfand homomorphism is onto.*

Proof. By the Šilov idempotent theorem the image of A in $C(\Omega)$ contains the characteristic functions of open closed sets. Hence A is a dense Banach subalgebra of $C(\Omega)$ and the theorem applies.

REMARK. An interesting fact suggested by the proof of Theorem 2.3 is that if Ω is a G-space then no normal subalgebra A of $C(\Omega)$, closed in the uniform norm, is such that $C(\Omega)/A$ has countable (infinite) dimension. To see this suppose to the contrary that $C(\Omega)/A$ has countable dimension. Recall that if A is a normal subalgebra of $C(\Omega)$ such that every point p of Ω has a neighborhood U_p such that $A|\bar{U}_p = C(\bar{U}_p)$ then $A = C(\Omega)$. Thus there is a point $p \in \Omega$ such that for every neighborhood U_p of p, $A|\bar{U}_p \neq C(\bar{U}_p)$. Since A contains the constant functions, by a result of Glicksberg [4 p. 421] we may choose $\mu_1 \in A^\perp$, $||\mu_1|| = 1$ such that $|\mu_1|\ (\bar{U}_1^*) > \delta > 0$ where $\bar{U}_1^*$ is a closed deleted neighborhood of p. By regularity of μ_1 we may choose a neighborhood U_2 of p such that $\bar{U}_2 \subsetneqq U_1$ and $|\mu_1|\ (\bar{U}_2^*) < \delta/2$. Again we may choose $\mu_2 \in A^\perp$, $||\mu_2|| = 1$, such that $|\mu_2|(\bar{U}_2{}^*) > \delta > 0$. Continuing in this fashion we get a sequence of measures $\{\mu_n\} \in A^\perp$, $||\mu_n|| = 1$, and a nested sequence of neighborhoods of p, $\{U_n\}$, $\bar{U}_{n+1} \subsetneqq U_n$ such that $|\mu_n|(\bar{U}_n - \bar{U}_{n+1}) > \delta/2$ for each n. By Grothendieck'e criterion no subsequence of $\{\mu_n\}$ is weakly convergent. Since $C(\Omega)/A$ is separable, the unit ball in $A^\perp$ is weak-* sequentially compact. Thus a subsequence of $\{\mu_n\}$ may be found which is weak-* convergent and hence weakly convergent. This contradiction completes the proof.

In [7] the following theorem is proved.

THEOREM 2.8. *If Ω is an F-space, and if X is a normal Banach subspace of $C(\Omega)$,* then $X = C(\Omega)$.

Question [7]. In Theorem 2.8 can "F" be replaced by "G"? In the terminology of that paper is every G-space an N-space? The following may be of help in giving an answer.

THEOREM 2.9. *Let X be a G-space in the sense of definition A, and let Y be a closed subspace such that X/Y is separable. Then Y*

is a G-space.

Proof. Let $\{y_n^*\}$ denote a sequence in Y^*. It suffices to show that if $\lim_n y_n^* = 0$ weak-* then $\{y_n^*\}$ has a subsequence $\{y_{n_k}^*\}$ such that $\lim_k y_{n_k}^* = 0$ weakly. Let x_n^* be any normpreserving extension of y_n^* to all of X. Since X/Y is separable, a sequence $\{w_n\}$ in X may be found such that $sp\{w_n\} + Y$ is dense in X. By a diagonal argument a subsequence $\{x_{n_k}^*\}$ of $\{x_n^*\}$ may be found such that $\{x_{n_k}^*\}$ converges on each member of $\{w_n\}$ and hence on $sp\{w_n\} + Y$. Since $\{\|x_{n_k}^*\|\}$ is bounded, $\{x_{n_k}^*\}$ is weak-* convergent in X and hence weakly convergent. Thus $\lim_k y_{n_k}^* = 0$ weakly.

Finally the author would like to thank the referee for his helpful suggestions, and in particular for the statement of Theorem 2.9.

References

1. W. G. Bade, *Extensions of Interpolation Sets*, Proc. of Irvine Conference on Functional Analysis, edited by B. R. Gelbaum, Thompson Book Company, Washington 1967.
2. W. G. Bade and P. C. Curtis, *Embedding theorems for Commutative Banach algebras*, Pacific J. Math., **18** (1966), 391-409.
3. P. C. Curtis, *A note concerning certain product spaces*, Arch. Math., **11** (1960), 50-52.
4. I. Glicksberg, *Measures orthogonal to algebras and sets of antisymmetry*, Trans. Amer. Math. Soc., **105** (1962), 415-435.
5. A. Grothendieck, *Sur les applications linéaires faiblement compacte d'espaces du type C(K)*, Canadian J. Math., (1953), 129-173.
6. A. Pelezynski, *Linear extensions, linear averagings, and their application to linear topological classification of spaces of continuous functions*, Dissertationes Mathematicae, Warzawa 1968.
7. G. Seever, *Measures on F spaces*, Trans. Amer. Math. Soc., **133** (1968), 267-280.
8. H. P. Rosenthal, *On Quasi-complemented subspaces of Banach spaces, with an appendix on compactness of operators from $L^p(\mu)$ to $L^r(\nu)$*, J. of Functional Analysis, **4** (1969), 176-214.

Received February 26, 1970.

University of Oregon

Pacific Journal of Mathematics

Vol. 36, No. 3 BadMonth, 1971