Pacific Journal of Mathematics

IN THIS ISSUE—

Vol. 68, No. 2 **April, 1977**

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. 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, may be sent to any one of the four editors. Please classify according to the scheme of Math. Reviews, Index to Vol. **39**. All other communications should be addressed to the managing editor, or Elaine Barth, University of California, Los Angeles, California, 90024.

100 reprints are provided free for each article, only if page charges have been substantially paid. Additional copies may be obtained at cost in multiples of 50.

The *Pacific Journal of Mathematics* is issued monthly as of January 1966. Regular subscription rate: \$72.00 a year (6 Vols., 12 issues). Special rate: \$36.00 a year to individual members of supporting institutions.

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 Jerusalem Academic Press, POB 2390, Jerusalem, Israel.

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No 2, 1977

AN ANALOGUE OF OKA'S THEOREM FOR WEAKLY NORMAL COMPLEX SPACES

WILLIAM A. ADKINS, ALDO ANDREOTTI, J. V. LEAHY

Two well known results concerning normal complex spaces are the following. First, the singular set of a normal complex space has codimension at least two. Second, this property characterizes normality for complex spaces which are local complete intersections. This second result is a theorem of Abhyankar [1] which generalizes Oka's theorem. The purpose of this paper is to prove analogues of these facts for the class of weakly normal complex spaces, which were introduced by Andreotti–Norguet [3] in a study of the space of cycles on an algebraic variety. A weakly normal complex space can have singularities in codimension one, but it will be shown that an obvious class of such singularities is generic.

1. Preliminaries. All complex spaces are assumed to be reduced. If X is a complex space, there is the sheaf $\mathcal{O}_X$ of holomorphic functions on X, and the sheaf $\mathcal{O}_X^c$ of c-holomorphic functions on X. A section of $\mathcal{O}_X^c$ on an open subset U of X is a continuous function $f\colon U \to \mathbf{C}$ such that f is holomorphic on the regular points of U. The complex space X is said to be weakly normal if $\mathcal{O}_X = \mathcal{O}_X^c$. Examples of weakly normal spaces are normal spaces and unions of submanifolds of $\mathbf{C}^m$ in general position.

Let $V_j = \{(x_1, \cdots, x_m) \in \mathbf{C}^m\colon x_k = 0 \text{ for } n \leqq k < j \text{ and } j < k \leqq m\}$ where $n \leqq j \leqq m$. Then V_j is an n-dimensional linear subspace of $\mathbf{C}^m$. Let

$$V_{(n,m)} = \bigcup_{j=n}^{m} V_j = \{(x_1, \cdots, x_m) \in \mathbf{C}^m : x_i x_j = 0 \quad \text{for} \quad n \leqq i < j \leqq m\}$$

and let $S(V_{(n,m)})$ be the singular set of $V_{(n,m)}$.

LEMMA. *$V_{(n,m)}$ is a weakly normal complex space and* $\dim S(V_{(n,m)}) = n - 1$.

Proof. Since $S(V_{(n,m)}) = \{(x_1, \cdots, x_m) \in \mathbf{C}^m\colon x_n = \cdots = x_m = 0\}$, $\dim S(V_{(n,m)}) = n - 1$. Let $f\colon V_{(n,m)} \to \mathbf{C}$ be a continuous function which is holomorphic on the regular points of $V_{(n,m)}$. To prove weak normality of $V_{(n,m)}$, we need to show that f is holomorphic. Let $f_j = f\,|\,_{V_j}$. By the Riemann extension theorem, f_j is holomorphic on the n-plane V_j and thus $f_j = f_j(x', x_j)$ is a convergent power series, where $x' = (x_1, \cdots, x_{n-1})$ and x_j are coordinates on V_j. Since $f_j|_{x_j=0} = f_k|_{x_k=0}$ for $n \leqq j, k \leqq m$, we let $f_0(x') = f_j(x', 0)$ and set $g_j(x', x_j) = f_j(x', x_j) - f_0(x')$ for $n \leqq j \leqq m$. Then $f(x_1, \cdots, x_m) = f_0(x') + \sum_{j=n}^{m} g_j(x', x_j)$ and hence f is holomorphic on $V_{(n,m)}$.

If X is a complex space with $\dim X = n$, let $Sg(X) = S(X) \cup (\bigcup_{0 \leqq k < n} X^{(k)})$ where $S(X)$ is the singular set of X and $X^{(k)}$ is the analytic subset of X defined by $X^{(k)} = \{x \in X : X \text{ has a branch of dimension } k \text{ at } x\}$. If $C_4(X, x)$ denotes the fourth Whitney tangent cone of X at x, then Stutz **[6]** has shown that $W_4 = Sg(X) \cap \{x \in X : \dim C_4(X, x) > n\}$ is an analytic subset of X of codimension at least two.

2. Codimension one singularities of weakly normal spaces. Let X be a complex space. A point $x \in X$ is said to be an elementary point of type (n, m), for $n \leqq m$, if the germ (X, x) is isomorphic to the germ $(V_{(n,m)}, 0)$. Note that if $x \in X$ is an elementary point of type (n, m), then the germ (X, x) is of pure dimension n and the imbedding dimension of (X, x) is m. The set of elementary points of X contains the set of regular points of X, i.e. the elementary points of type (n, n) for some n. In addition, it contains a particularly simple class of singular points of X. If x is an elementary point of type (n, m) with $n < m$, then x is singular and $\dim(S(X), x) = n - 1 = \dim(X, x) - 1$.

If $\dim X = n$, let $Y = \bigcup_{0 \leqq k < n} X^{(k)}$ and let $X_1 = \overline{X \backslash Y}$. By a theorem of Remmert, X_1 is an analytic set of pure dimension n. Let X_s denote the set of all elementary points of X of type (n, m) for some m with $m \geqq n = \dim X$. Hence $X_s \subseteq X_1$ and X_s contains the regular points of X of maximal dimension.

THEOREM 1. *Let X be a weakly normal complex space. Then $A = X_1 \backslash X_s$ is an analytic subset of X_1 of codimension at least 2.*

Proof. Let $n = \dim X$. If $\dim S(X) \leqq n - 2$ then $A = X_1 \cap S(X)$. Hence A is analytic and codimension $A \geqq 2$. Now suppose that $\dim S(X) = n - 1$. We will show that $A = X_1 \cap (Sg(Sg(X)) \cup W_4)$. Since $Sg(Sg(X)) \cup W_4$ is an analytic set of codimension at least 2 in X and since $\dim X = \dim X_1$, this will prove the theorem.

Let $x \in X_s$. If x is a regular point of X, then $x \notin Sg(Sg(X)) \cup W_4$. If x is an elemetary point of type (n, m) where $m > n$, then $\dim C_4(X, x) = n$. Hence $x \notin W_4$. Moreover, $S(X)$ is a manifold of dimension $n - 1$ in a neighborhood of x. Thus $x \notin Sg(Sg(X))$. Hence $X_s \subseteq X_1 \backslash (Sg(Sg(X)) \cup W_4)$ and $X_1 \cap (Sg(Sg(X)) \cup W_4) \subseteq A$.

Now suppose that $x_0 \in X_1 \cap S(X) \cap (X_1 \backslash (\mathrm{Sg}(\mathrm{Sg}(X)) \cup W_4))$. Thus

$x_0 \in \mathrm{Sg}(X) \backslash \mathrm{Sg}(\mathrm{Sg}(X))$ and $\dim C_4(X, x_0) = n$. Note also that the germ (X, x_0) is of pure dimension n. Since the result to be proved is local, we may assume that $X \subseteq \mathbf{C}^t$. By Proposition 4.2 of Stutz **[6]**, there is a neighborhood N of x_0 in X, a polydisc $D \subseteq \mathbf{C}^n$, and a choice of coordinates $x_1, \cdots, x_n$ in $\mathbf{C}^n$ and $y_1, \cdots, y_t$ in $\mathbf{C}^t$ centered at x_0 with the following properties.

If $B_0, \cdots, B_r$ are the global branches of $X \cap N$, then for each j $(0 \leqq j \leqq r)$ there is a holomorphic map $f_j \colon D \to B_j$ such that

(a) f_j is a homeomorphism;

(b) with respect to the coordinates $x_1, \cdots, x_n, y_1, \cdots, y_t$, $f_j(0) = 0$ and

$$f_j(x) = (x_1, \cdots, x_{n-1}, x_n^{p_j}, f_{n+1,j}(x), \cdots, f_{tj}(x))$$

where p_j is a positive integer for $0 \leqq j \leqq r$;

(c) $f_{ij}(x_1, \cdots, x_n) = \sum_{\nu = p_j}^{\infty} f_{ij}^{(\nu)}(x_1, \cdots, x_{n-1}) \cdot x_n^{\nu}$ for $n + 1 \leqq i \leqq t$ and $0 \leqq j \leqq r$.

Let $g_j \colon B_j \to D$ be the continuous inverse of f_j and define a map $h \colon X \cap N \to \mathbf{C}^{n+r}$ by $\pi_j \circ h \,|_{B_j} = g_j$ where $\pi_j \colon \mathbf{C}^{n+r} \to \mathbf{C}_{x_1, \cdots, x_{n-1}, x_{n+j}}$ is the natural linear projection onto the n-plane with coordinates $x_1, \cdots, x_{n-1}, x_{n+j}$, for $0 \leqq j \leqq r$. To see that the map h is well defined, note first that $S(X)$ is an $n-1$ dimensional manifold in a neighborhood of x_0. Furthermore, $B_j \cap B_k \subseteq S(X) \cap N$ for all j, k. But $f_j(x', 0) = (x', 0, \cdots, 0) = f_k(x', 0)$ where $x' = (x_1, \cdots, x_{n-1})$. Therefore, if N is chosen small enough, then $B_j \cap B_k = S(X) \cap N = \{y_n = \cdots = y_t = 0\}$ for $0 \leqq j,\ k \leqq r$. For each $(y_1, \cdots, y_t) \in S(X) \cap N$, it follows that $g_j(y) = (y_1, \cdots, y_{n-1}, 0)$ for $0 \leqq j \leqq r$. Thus h is a well defined continuous map.

Since the jacobian matrix $\partial f_j / \partial x$ is given by

$$\frac{\partial f_j}{\partial x} = \begin{bmatrix} I_{n-1} & 0 \\ 0 & p_j x_n^{p_j - 1} \\ * & * \end{bmatrix}$$

h is holomorphic on the regular points of $X \cap N$. Since X is weakly normal and h is a homeomorphism onto its image, it follows that h is biholomorphic. Therefore x_0 is an elementary singularity of type $(n, n + r)$. Hence $A \subseteq X_1 \cap (\mathrm{Sg}(\mathrm{Sg}(X)) \cup W_4)$ and the theorem is proved.

REMARK. Let X be a weakly normal complex space and suppose that $\operatorname{codim} S(X) = 1$. Theorem 1 shows that there is an elementary singularity of type (n, m) where $m > n$. Since such a singular point is not normal, Theorem 1 implies the well-known theorem that $\operatorname{codim} S(X) \geqq 2$ for a normal complex space X.

THEOREM 2. *Let X be a pure dimensional local complete intersection. Then X is weakly normal if and only if* $\operatorname{codim} X \backslash X_s \geqq 2$.

Proof. Let $A = X \backslash X_s$. If X is weakly normal then $\operatorname{codim} A \geqq 2$ by Theorem 1.

Conversely, suppose $\operatorname{codim} A \geqq 2$. Since $X \backslash A = X_s$, the germ (X, x) is weakly normal for each $x \in X \backslash A$. Since X is a pure dimensional local complete intersection, $pf(\mathcal{O}_{X,x}) = \dim X$ for each $x \in X$, where pf = profondeur. From the Hartog theorem for weak normality **[2]**, we conclude that X is weakly normal.

REMARKS. (1) For the case of curves, the assumption of local complete intersection is not needed. A curve X is weakly normal if and only if $X \backslash X_s = \emptyset$. An algebraic proof of this fact was given by Bombieri **[5]**.

(2) If X is a pure dimensional hypersurface in $\mathbf{C}^{n+1}$, then Theorem 2 can be proved without the use of the Hartog theorem for weak normality. This case follows from the result of Becker in **[4]**.

(3) Let $X \subseteq \mathbf{C}^{n+1}$ be a pure dimensional hypersurface. If X is weakly normal, there is another characterization of $X \backslash X_s$ than that which is given by the proof of Theorem 1. This description is as follows. There is a holomorphic function $f \in \mathcal{O}(\mathbf{C}^{n+1})$ such that $X = V(f) = \{x \in \mathbf{C}^{n+1} : f(x) = 0\}$ and such that there is a sheaf equality $(f) \cdot \mathcal{O} = \mathcal{I}_X$ where $\mathcal{I}_X$ is the sheaf of ideals of X. Then

$$S(X) = \left\{ x \in X : \frac{\partial f}{\partial z_i}(x) = 0 \text{ for } 1 \leqq i \leqq n+1 \right\}.$$

At a point $x_0 \in S(X)$ the Hessian form is defined by

$$H(f)_{x_0}(u) = \sum_{i,j=1}^{n+1} \frac{\partial^2 f}{\partial z_i \partial z_j}(x_0) \cdot u_i u_j.$$

Let $\mu(x_0) = \operatorname{rank} H(f)_{x_0}$ and set $S_2(X) = \{x \in S(X) : \mu(x) \leqq 1\}$.

Claim. If X is weakly normal and $\dim S(X) = n - 1$, then

$$W_4 \cap (S(X) \backslash \operatorname{Sg}(S(X))) = S_2 \cap (S(X) \backslash \operatorname{Sg}(S(X))).$$

Proof. From the proof of Theorem 1, $X \backslash X_s = \operatorname{Sg}(S(X)) \cup W_4$. Suppose $x \in S(X) \backslash \operatorname{Sg}(S(X))$ but $x \notin W_4$. Then the proof of Theorem 1 shows that x is an elementary singular point of type $(n, n+1)$. A proper choice of local coordinates about x shows that (X, x) is isomorphic to $(V(z_1 z_2), 0)$. Hence $\mu(x) = 2$ and $x \notin S_2(X)$.

Now suppose that $x \in S(X)\setminus \mathrm{Sg}(S(X))$ but $x \notin S_2(X)$. Thus $\mu(x) \geqq 2$. If $\mu(x) > 2$ then the implicit function theorem shows that $\dim(S(X), x) \leqq n-2$. Therefore $\mu(x) = 2$ and choosing convenient local coordinates centered at x gives $f(z) = az_1z_2 + 0(3)$ where $a \neq 0$. Hence x is an elementary singular point of type $(n, n+1)$. Therefore, $x \notin W_4$ and the claim is proved.

For weakly normal hypersurfaces this claim gives an easy differential criterion for computing the portion of the set W_4 which is contained in $S(X)\setminus \mathrm{Sg}(S(X))$. This claim is false for hypersurfaces which are not weakly normal.

EXAMPLE. Let $X = \{(x, y, z) \in \mathbf{C}^3 \colon x^2 - zy^2 = 0\}$ be the Cayley umbrella in $\mathbf{C}^3$. Then $X \setminus X_s = \{(0,0,0)\}$ so that X is weakly normal by Theorem 2. Remark (3) then shows that $W_4 = \{0,0,0)\}$.

REFERENCES

1. Shreeram Abhyankar, *Concepts of order and rank on a complex space and a condition for normality*, Math. Annalen, **141** (1960), 171–192.

2. Aldo Andreotti and Per Holm, *Parametric spaces*, preprint.

3. Aldo Andreotti and Francois Norguet, *La convexite holomorphe dans l'espace analytic des cycles d'une variete algebrique*, Ann. Scuola Norm. Sup. Pisa, s. 3, **21** (1967), 31–82.

4. Joseph Becker, *Normal Hypersurfaces*, Pacific J. Math. **61** (1975), 17–19.

5. Enrico Bombieri, *Seminormalite e singolarite ordinarie*, Symposia Mathematica XI, Academic Press, (1973), 205–210.

6. John Stutz, *Analytic sets as branched coverings*, Trans. Amer. Math. Soc., **166** (1972), 241–259.

Received December 20, 1976.

INSTITUTE FOR ADVANCED STUDY
PRINCETON, NJ 08540

OREGON STATE UNIVERSITY
CORVALLIS, OR 97331
AND
UNIVERSITY OF OREGON
EUGENE, OR 97403

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No 2, 1977

ON NONSINGULARLY k-PRIMITIVE RINGS

A. K. BOYLE, M. G. DESHPANDE AND E. H. FELLER

A ring R is called k-primitive if it has a faithful cyclic critical right module C with $|C| = k$. We first show that k-primitive rings with Krull dimension have many properties in common with prime rings. For the case where R is a *PWD* with a faithful critical right ideal, we obtain an internal characterization.

1. Introduction. Let R be a ring with Krull dimension. Then R is prime if and only if R has a faithful compressible right R-module. In this paper we consider a broader class of rings, those which have a faithful cyclic critical right R-module. From **[2]** such a ring is called a *k-primitive* ring where k denotes the Krull dimension of the faithful critical.

In the case where the faithful critical is nonsingular, these rings exhibit many of the properties of prime rings. Not all k-primitive rings have this additional property as an example in §4 shows. We call a k-primitive ring whose faithful critical is nonsingular, a *nonsingularly k-primitive* ring. Section 2 is devoted to showing some of the similarities with prime rings.

In §3 we consider piecewise domains (*PWD*) which are k-primitive rings. An internal characterization of *PWD*'s with faithful critical right ideal is obtained, which is our main result.

All rings will have identity, and the modules are right unital. The singular submodule of a module M_R is denoted $Z(M)$. If X is a subset of R, then ann X or X' denotes the right annihilator of X in R. The Krull dimension of a module M_R is denoted by $|M|$. A certain familiarity with the definitions and basic results concerning Krull dimension is assumed. See **[5]** for reference.

2. Properties of k-primitive rings. If R is a prime ring with Krull dimension then R is nonsingular and has a faithful critical C such that $|C| = |R|$. These conditions are also true for nonsingularly k-primitive rings.

PROPOSITION 2.1. *Let R be a k-primitive ring with faithful cyclic critical C. Then $Z(R) = 0$ and $|C| = |R|$ if and only if R is nonsingularly k-primitive.*

Proof. Suppose $Z(C)=0$. This immediately implies $Z(R)=0$. Let X be the collection of annihilators of finite subsets of C. By [**4**, Theorem 1.24], X satisfies the descending chain condition. Since C is faithful $\bigcap_{x\in C}(x^r)=0$ and there exists a finite subset such that $\bigcap_{i=1}^n(x_i^r)=0$. This implies the existence of an R-monomorphism $R\to\sum_{i=1}^n R/x_i^r\to C^{(n)}$. Thus $|R|\leqq|C^{(n)}|=|C|\leqq|R|$.

Conversely if $Z(R)=0$, then $Z(C)=C$ or $Z(C)=0$. Suppose $Z(C)=C$. Then $C\cong R/K$ where K is a large right ideal. Let $L=\{D\mid D$ is a critical right ideal$\}$. Then $S=\Sigma D$, $D\in L$ is a two sided ideal of R. Since R/K is faithful, $S\not\subset K$. Hence there exists $D\in L$ such that $D\not\subset K$. Since K is large $D\cap K\neq 0$ and $|R/K|=|D+K/K|=|D/D\cap K|<|D|\leqq|R|$. This contradicts the fact that $|C|=|R|$, and therefore $Z(C)=0$.

Let C be the faithful k-critical of a nonsingularly k-primitive ring R. Then $P=\text{ass } C$ is a prime ideal. In the remainder of this paper C and P will be used in this way.

LEMMA 2.2. *If R is a nonsingularly k-primitive ring with faithful critical C, then $P=\text{ass } C$ is a nonessential minimal prime and $|R|=|R/P|$.*

Proof. That P is a nonessential minimal prime is straightforward.

The module C contains a nonzero submodule C^* where ann $C^*=P$. Since C^* is a nonsingular, faithful R/P-module, then by 2.1 $|R|=|C^*|=|R/P|$.

PROPOSITION 2.3. *Let R be a nonsingularly k-primitive ring with faithful, critical C and let $P=\text{ass } C$. Then*

(1) *P contains all nonessential two sided ideals.*
(2) *R has exactly one nonessential prime ideal, namely P.*
(3) *If R is semiprime, then R is prime.*
(4) *Every uniform right ideal which misses P is compressible.*
(5) *Every uniform right ideal is critical and subisomorphic to C.*

Proof. (1) If H is not essential, there exists a right ideal I such that $I\cap H=0$. Then $IH=0$ which implies $H\subset\text{ass } C=P$ by [**2**, Proposition 3.2].

(2) This follows from (1).

(3) If $0=P_1\cap\cdots\cap P_n$ is an irredundant intersection of minimal primes, then P_i is not large and hence $P_i=P$ by (2) for each i. Thus $P=0$.

(4) If U is uniform and $U\cap P=0$, then $U\cong U+P/P\subseteq R/P$. Since uniform right ideals of a prime ring are compressible, the result follows.

(5) Let U be a uniform right ideal. Then $CU \neq 0$. Thus there exists $x \in C$ such that $xU \neq 0$. Since $Z(C)=0$, $U \cong xU \subset C$ and U is critical.

A ring R with Krull dimension is called *very smooth* if every right ideal of R has the same Krull dimension. In a prime ring any two critical right ideals are subisomorphic and thus the ring is very smooth.

PROPOSITION 2.4. *Let R be a nonsingularly k-primitive ring with faithful critical C.*

(1) *If D is a critical right ideal, then* ass $D = P$.
(2) *R is a very smooth ring.*
(3) *If D is a critical right ideal then $D \cap P = 0$ or $D \subseteq P$.*
(4) *If R is not prime and if $C \subset R$, then $C \subset P$.*

Proof. (1) Let $P_0 = ass\ D$. Then D has a submodule D^* such that ann $D^* = P_0$. Since $Z(D^*) = 0$, P_0 is not essential and hence by 2.3(2), $P_0 = P = ass\ C$.

(2) It suffices to show that if D is critical, $|D| = |R|$. By (1) D has a submodule D^* which is nonsingular and faithful as an R/P-module and hence $|D| = |D^*| = |R/P| = |R|$.

(3) If $D \cap P \neq 0$ then $D/D \cap P \cong D + P/P \subset R/P$. Since R/P is very smooth and $|R/P| = |D|$ then $|D/D \cap P| < |D|$ implies $D = D \cap P$.

(4) If $C \subset R$ and $C \cap P = 0$, then $CP = 0$ contradicting the faithfulness of C. Thus $C \cap P \neq 0$. Therefore by (3) $C \subset P$.

PROPOSITION 2.5. *Let R be a nonsingularly k-primitive ring. The compressible right ideals of R are subisomorphic.*

Proof. Since $P =$ ass C is not large, there exists a critical right ideal $D \neq 0$ such that $D \cap P = 0$. By 2.3 (4) D is compressible. Let $K \neq 0$ be any compressible right ideal. Then $KD \neq 0$ since $D \not\subseteq P =$ ass C. Thus there exists $a \in K$ and a monomorphism $D \to aD \subseteq K$. Since K is compressible, K is subisomorphic to aD and hence to D. Since being subisomorphic is a transitive property, any two compressible right ideals are subisomorphic.

COROLLARY 2.6. *Let R be a nonsingularly k-primitive ring with the nonessential prime $P \neq 0$. Then P contains an isomorphic copy of all uniform right ideals.*

3. k-primitive piecewise domains. Let R be a ring with Krull dimension and suppose that R is a piecewise domain with a faithful critical right ideal. Then R is nonsingularly k-primitive and is, in general, not prime. We assume these rings have a faithful critical right ideal. In § 4 we provide an example to show that this need not always be the case. From [7] we have

DEFINITION 3.1. A ring R is a *piecewise domain* (PWD) with respect to a complete set of orthogonal idempotents $e_1, \cdots, e_n$ if $x \in e_i \mathrm{R} e_k$, $y \in e_k \mathrm{R} e_j$ then $xy = 0$ implies $x = 0$ or $y = 0$.

In [7, p. 554] the following criterion is given for R to be a PWD. R is a PWD with respect to the complete set of orthogonal idempotents $\{e_i\}$ if and only if every nonzero element of $\mathrm{Hom}_R(e_iR, R)$ is a monomorphism.

PROPOSITION 3.2. *Let R be a ring with Krull dimension. Then R is a PWD with faithful critical right ideal if and only if $R = \Sigma \oplus e_iR$* where e_iR is critical for every i and e_jR is faithful and nonsingular for some j. In this case R is nonsingularly k-primitive.

Proof. Suppose R is a PWD with faithful critical right ideal. Then $R = \Sigma \oplus e_iR$ where $e_1, \cdots, e_n$ is a complete set of orthogonal idempotents. Since C is faithful, $Ce_iR \neq 0$ for all i. Hence for any given i there exists $c \in C$ such that $ce_iR \neq 0$. We can therefore define a homomorphism θ of e_iR into C using c and by [7, p. 554] the mapping θ is a monomorphism. Thus e_iR is critical.

Now $RC \neq 0$ which implies $e_jRC \neq 0$ for some j. By [8, Lemma 1], $Z(R) = 0$ and thus the mapping determined by the relation $e_jrC \neq 0$ is a monomorphism. So C is subisomorphic to at least one e_jR and $Z(e_jR) = 0$.

Conversely suppose $R = \Sigma \oplus e_iR$ where e_iR is critical. Since e_jR is faithful and nonsingular by 2.4 we know that R is very smooth. So consider a mapping $f: e_iR \to R$. If Ker $f \neq 0$ then $|f(e_iR)| = |e_iR/\mathrm{Ker} f| < |e_iR| = |R|$. Thus necessarily $f = 0$.

The same technique employed in the above proof shows that any two faithful critical right ideals of a PWD are subisomorphic. Furthermore the faithful critical right ideal contains an isomorphic copy of every critical right ideal.

PROPOSITION 3.3. *Let R be a ring with Krull dimension. If R is a PWD with faithful critical right ideal, say $R = \Sigma \oplus e_iR$ where $C = e_1R$ is faithful, then*

(1) $P = \mathrm{ass}\ C = \Sigma_i e_iR$, *where e_iR is not compressible.*
(2) *If Q is a prime ideal not equal to P then $|R/Q| < |R|$.*
(3) *If R is not a prime ring, then not all the e_iR are compressible.*

Proof. (1) and (3). If R is prime, then the e_iR are all compressible and ass $C = P = 0$. If R is not prime, then C cannot be compressible because it is faithful.

Now suppose e_iR are compressible for $i = m+1, \cdots, n$ and e_jR are not compressible for $j = 1, \cdots, m$. Then

$$P = RP = \left(\sum_{i=1}^{n} e_iR\right)P \subseteq \sum_{i=1}^{m} e_iRP \subseteq \sum_{i=1}^{m} e_iR.$$

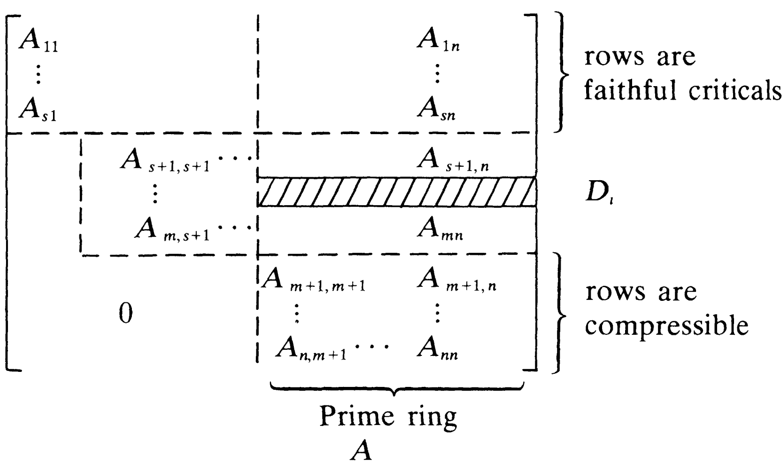
Conversely if D is critical either $D \cap P = 0$ or $D \subset P$ by 2.4(3). Since $e_iR \cap P = 0$ implies by 2.3(4) that e_iR is compressible, necessarily $e_iR \subseteq P$ $1 \leqq i \leqq m$ and hence $P = \Sigma_{i=1}^{m} e_iR$.

(2) If Q is a prime ideal not equal to P, then Q is large. Suppose $|R/Q| = |R|$. As in the proof of 2.4(3) if D is critical $D \subseteq Q$. Thus Q contains all critical right ideals and hence $R \subseteq Q$ which is impossible.

In [**8**, Theorem 2] Gordon obtains an internal characterization of prime right Goldie rings which are PWD's. In the following theorem we obtain an internal characterization for a nonsingularly k-primitive ring with a faithful critical right ideal which is a PWD.

THEOREM 3.4. *Let R be a PWD with Krull dimension, $|R| = k$. Then R is nonsingularly k-primitive with faithful critical right ideal if and only if $R \cong (A_{ij})_{n\times n}$ where for some s, m where $1 \leqq s < m < n$,*

(1) *$A_{ij} = 0$ for $i > s$ and $j \leqq s$ or $i > m$ and $j \leqq m$.*
(2) *$A_{ij} \neq 0$ for $i \leqq s$ or $j > m$.*
(3) *A_{ii} is a domain for $1 \leqq i \leqq n$.*
(4) *If $j \leqq m$ then $|(A_{ij})_{A_{jj}}| < k$.*
(5) *$A = \{(a_{ij}) \in (A_{ij}) \mid a_{ij} = 0 \; 1 \leqq i \leqq m$ or $1 \leqq j \leqq m\}$ is a prime ring of Krull dimension k and $D_i = \{(a_{kj}) \mid a_{ij} = 0 \; 1 \leqq j \leqq m,\ a_{kj} = 0 \; k \neq i\}$ as a right A-module is k-critical for all i.*



Proof. By 3.2 and 3.3 $R = \Sigma_{i=1}^{n} \oplus C_i$ with C_i critical for all i where C_i is faithful for $1 \leqq i \leqq s$ and C_j is compressible for $m+1 \leqq j \leqq n$. Then R is isomorphic to the matrix ring with (i,j) entries in $A_{ij} = \mathrm{Hom}_R(C_j, C_i)$, which are monomorphisms or zero. Since $Z(R) = 0$ by [7], then $Z(C_i) = 0$.

(1) If $i > s$, and $j \leqq s$, then $\mathrm{Hom}(C_j, C_i) = 0$ since these C_i are not faithful. Similarly, $\mathrm{Hom}(C_j, C_i) = 0$ $i > m$, $j \leqq m$, since these C_i are not compressible, and the C_i are compressible (using the fact that submodules of compressible modules are compressible). This proves (1).

(2) If C_i is faithful, then $C_iC_j \neq 0$ for all j and $\mathrm{Hom}(C_j, C_i) \neq 0$ for all j. If C_j is compressible, then $C_iC_j \neq 0$ for any i. Hence $\mathrm{Hom}(C_j, C_i) \neq 0$. This proves (2).

(3) A_{ii} is a domain, since R is *PWD* ring.

(4) For $j \leqq m$, let K_{ij} be an A_{jj} submodule of A_{ij}. Let $K^*_{ij} = \{(a_{rt}) \mid a_{ij} \in K_{ij}$, and $a_{rt} = 0$ otherwise$\}$. Let D_i be as in the theorem, then $D_i \neq 0$ by (2) and D_i is a right ideal of R because of (1). Let $S_i = K^*_{ij}R + D_i/D_i$. Then $S_i \subset C_i/D_i$. Then the mapping $K_{ij} \to S_i$ is a lattice isomorphism of the lattice of submodules of A_{ij} over A_{jj} into the lattice of submodules of C_i/D_i over R. Since C_i is critical and since $D_i \neq 0$, then the Krull dimension of A_{ij} over A_{jj} must be less than k.

(5) The first part follows since ass $C = P$ is equal to $C_1 + \cdots + C_m$ by 3.3 and $A \cong R/P$. Now each D_i is a submodule of a critical module, and hence is critical over R. But the lattice of modules of D_i over R is the same as the lattice of D_i over A, and D_i is critical over A, and since the Krull dimension of D_i over R is k, then $|D_i|$ over A is also k.

Conversely, let R be a *PWD* satisfying these conditions and let C_i denote the ith row of the matrix. We will show that C_i is critical, and that C_i, $i = 1, 2, \cdots, s$ is faithful. Since R is a *PWD*, and using (2), we have that C_i, for $i = 1, 2, \cdots, s$ is faithful. We now show C_1 is k-critical. Let M_1 be a *nonzero* submodule of C_1 over R. Let $M_{1j} = \{a \in A_{1j} \mid a$ is the $(1,j)$ entry of some member of $M_1\}$. Then $M_{1j} \neq 0$ for $j > m$. Now for $j \leqq m$, M_{1j} is an A_{jj} module. Thus by (4), the Krull dimension of A_{1j}/M_{1j} over A_{jj} is less than k. If $N = \Sigma_{i=m+1}^{n} M^*_{1i}$ where $M^*_{1i} = \{(a_{rt}) \mid a_{1i} \in M_{1i}$ and $a_{rt} = 0$ otherwise$\}$, then N is not zero, and by (4), the Krull dimension of D_1/N over A is less than k. Thus with each submodule of C_1/M_1, we can associate an A_{jj} submodule of A_{1j}/M_{1j}, one for each $j \leqq m$, and a factor module D_1/N of D_1 as an A-module. We construct a lattice isomorphism from the submodules of C_1/M_1 into the lattice of submodules of $T = A_{11}/M_{11} \oplus \cdots \oplus A_{1m}/M_{1m} \oplus D_1/N$ over $S = A_{11} \oplus \cdots \oplus A_{mm} \oplus A$, where in T_S we have scalar multiplication defined as coordinate multiplication by elements of the ring S. Now the Krull dimension of A_{1i}/M_{1i} is less than k over A_{ii} for $1 \leqq i \leqq m$, and similarly for D_1/N over A. Hence the Krull dimension of T over S is less than k. Thus $|C_1/M_1| < k$. Since D_1 over A has Krull dimension k, one can show that C_1 has Krull dimension k over R in a similar fashion. Thus C_1 is k-critical. Since $R = C_1 \oplus \cdots \oplus C_n$ and C_1 is faithful, then each C_i is R isomorphic to a submodule of C_1, and hence is critical of Krull dimension k. Thus $|R| = k$.

4. Examples and questions. Using the results of §3, one can easily construct nonsingularly k-primitive rings. Let I a right Noetherian integral domain, where $|I| = k$, and $I[x]$ be the polynomial ring with commuting x. We construct three matrix rings of this type

$$R = \begin{bmatrix} I & I & I[x] & I[x] & I[x][y] & I[x][y] \\ I & I & I[x] & I[x] & I[x][y] & I[x][y] \\ 0 & 0 & I[x] & A & I[x][y] & I[x][y] \\ 0 & 0 & B & I[x] & I[x][y] & I[x][y] \\ 0 & 0 & 0 & 0 & I[x][y] & I[x][y] \\ 0 & 0 & 0 & 0 & I[x][y] & I[x][y] \end{bmatrix}$$

where (1) $A = B = 0$, (2) $A = I[x]$, $B = 0$, (3) $A = I[x]$, $B = I[x]$. These are nonsingularly k-primitive rings of Krull dimension $k + 2$.

In each of these three rings, with the notation of Theorem 3.4, C_1 is the faithful critical right ideal and $s = 2$, $m = 4$. In the case where $A = B = I$, the two nonfaithful, noncompressible modules C_3 and C_4 are (sub)isomorphic. Clearly, in all cases, $P = C_1 + C_2 + C_3 + C_4$.

One can show, in general, that if R is nonsingularly k-primitive, then the complete ring of quotients Q is a simple Artinian ring. If R is PWD, then by **[9]**, R has an Artinian Classical quotient ring Q_{cl}. However, Q_{cl} is never k-primitive unless R is prime. To illustrate this consider for a field F the ring, $R = \begin{bmatrix} F & F[x] \\ 0 & F[x] \end{bmatrix}$. Then $Q(R) = \begin{bmatrix} F(x) & F(x) \\ F(x) & F(x) \end{bmatrix}$ and $Q_{cl}(R) = \begin{bmatrix} F & F(x) \\ 0 & F(x) \end{bmatrix}$.

Question 1. In general does R have a classical quotient ring Q_{cl}? Does Q_{cl} contain a maximal k-primitive subring which contains R?

Question 2. Does $N(R)$ being prime imply R is prime? (True for PWD's.)

Question 3. If R has left Krull dimension, is R a prime ring?

In regard to questions 1 and 2, if our ring R has the regularity condition (See [**6**]), and if N is prime, then R is a prime ring.

To give an example of a nonsingularly k-primitive PWD where the faithful cyclic is not imbeddable in R consider a field F with a derivation $(')$. Let $M = F[x] \oplus F[x]$. Then M is a right $F[x]$ module under $(f, g)h = (fh, gh)$, and M is a left F module under $a(f, g) = (af + a'g, ag)$, $a \in F$. Thus M is a bimodule, and the matrices

$$R = \begin{bmatrix} F & M \\ 0 & F[x] \end{bmatrix} \text{ form a ring.}$$

Let $I = \begin{bmatrix} 0 & N \\ 0 & F[x] \end{bmatrix}$, where $N = \{(0, f(x) \mid f(x) \in F[x]\}$. Then I is not large, and contains no two-sided ideals. In addition $C = R/I$ is critical. In fact, if $C_0 \neq 0$ is a submodule of C, then C/C_0 is Artinian. One can show that C cannot be embedded in R, and $Z(C) = 0$. One can also show that no right ideal of R is faithful and critical. Hence R is a nonsingularly 1-primitive ring without a faithful critical right ideal.

In the case where R is k-primitive but the faithful critical is not nonsingular, then R may not have the properties established in §2. Consider the following example.

Let $A = F[x, (')][z]$ where F is a field with derivative $(')$ as in [**3**, p. 55], and z commutes with x. Let

$$R = \begin{bmatrix} F & A/xA \\ 0 & A \end{bmatrix}.$$

The first row of R is a faithful cyclic critical C which is not nonsingular. Now $|C| = 1$ and $|R| = 2$. Thus R is not very smooth. In addition, R does not satisfy the regularity condition. Hence R does not have an Artinian classical right quotient ring.

References

1. A. K. Boyle and E. H. Feller, *Smooth Noetherian modules*, Comm. Algebras, **4** (1976), 617–637.
2. M. G. Deshpande and E. H. Feller, *The Krull radical*, Comm. Algebras, **3** (1975), 185–193.
3. N. J. Divinsky, *Ring and Radicals*, University of Toronto Press, Mathematical Expositions No. 14, 1965.
4. K. R. Goodearl, *Singular torsion and the splitting properties*, Memoirs of Amer. Math. Soc. No. 124 (1972).
5. R. Gordon and J. C. Robson, *Krull dimension*, Memoirs of Amer. Math. Soc. No. 133 (1973).
6. R. Gordon, *Gabriel and Krull dimension*, Ring Theory, Lecture notes in Pure and Applied Math. Marcel Dekker, New York, (1974), 241–295.

7. R. Gordon and L. W. Small, *Piecewise Domains*, J. Algebra, **23**, No. 3 (1972), 553–564.

8. R. Gordon, *Semiprime right Goldie rings which are a direct sum of uniform right ideals*, Bull. London Math. Soc., **3** (1971), 277–282.

9. ———, *Classical quotient rings of PWD's*, Proc. Amer. Math. Soc., **36** (1972), 39–46.

Received March 24, 1976 and in revised form November 22, 1976.

UNIVERSITY OF WISCONSIN–MILWAUKEE
MARQUETTE UNIVERSITY
AND
UNIVERSITY OF WISCONSIN–MILWAUKEE
MILWAUKEE, WI 53201

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No. 2, 1977

MEASURES INVARIANT UNDER A GROUP OF TRANSFORMATIONS

ROLANDO B. CHUAQUI

The purpose of this article is to present necessary and sufficient conditions of an algebraic character for the existence of a countably additive measure defined on a σ-field of sets vanishing exactly on a given subset of the field, and invariant under a group of transformations.

0. Introduction. Most measures of interest defined on fields of sets are required to be invariant under certain groups of transformations. For any field F of subsets of a set X, a group of transformations G of F is a set of one-one functions closed under composition and inverse, that contains the identity function restricted to X, and such that any function in G transforms elements of F into elements of F. A measure on F is invariant under G if any two elements $A, B \in F$, such that A is the image of B under a function in G, have the same measure.

The problem of finding algebraic conditions for the existence of such measures has been discussed in several places; in particular, in Tarski's book **[14]**, p. 231, where necessary and sufficient conditions are given for the existence of a finitely additive measure invariant under a group of transformations. To my knowledge, no general solution of this type has been published, before this paper, for countably additive measures.

Partial solutions to this problem were obtained in **[11, 2]**. In this second paper a conjecture was formulated which was proved false in **[3]**.

The proof presented here uses extensively the theory of Cardinal Algebras developed in **[14]**. I shall quote theorems and definitions from this book by their number followed by a T. I also use a representation theorem for certain types of Cardinal Algebras obtained in **[6]**.

A related problem is the existence of a countably additive measure on a Boolean σ-algebra when no group of transformations is involved (for the theory of Boolean algebras see **[13]**). The interesting problem, in this case, is to find a strictly positive measure (i.e. a measure that vanishes only on the zero of the algebra). Necessary and sufficient conditions were found in **[11]**, and better conditions in **[10]**. I shall use these latter requirements for the existence of invariant measures.

The conditions obtained for the existence of an invariant σ-measure are a combination of Kelley's requirements and the countably additive version of the main condition of Tarski for the existence of a finitely additive invariant measure: the nonexistence of paradoxical decompositions of the unit set. The main theorem proved (Theorem 5.1) includes as particular cases results obtained by [7] for a group generated by one transformation, and by [9] for a continuous group.

Given a σ-field $\mathfrak{F}$ of sets and a group of transformations of $\mathfrak{F}$, it can be deduced from the main result in this paper mentioned above (5.1), that an invariant σ-measure exists iff an ideal in $\mathfrak{F}$ of a certain type exists. This falls short of determining algebraic necessary and sufficient conditions for the existence of an invariant σ-measure (with no predetermined ideal of null-sets). However, this shortcoming is shared by the solutions available for the existence of a σ-measure on a σ-field of sets when no group of transformations is involved: From Kelley's conditions it can also be deduced that a σ-measure on $\mathfrak{F}$ exists iff an ideal in $\mathfrak{F}$ of a certain type exists.

On the other hand, if we are given the ideal on which the measure vanishes, both Kelley's theorem and mine give algebraic necessary and sufficient conditions for its existence.

In the first section the general setting of fields of sets and groups of transformations is discussed. The second section contains some lemmas about ideals and congruence relations in Boolean Algebras and Cardinal Algebras. Section three studies the countable chain condition. The next section gives the main theorems on invariant measures on Boolean Algebras. Finally, the fifth section applies these theorems to fields of sets.

The measure-theoretic results obtained in this paper were announced without proof in [4], where they were applied to obtain probability measures.

1. Groups of transformations on fields of sets. Throughout this paper we employ the usual set-theoretical terminology. We identify an ordinal number with the set of preceding ordinals, and a cardinal number with the corresponding initial ordinal. In particular ω, the set of natural number is the first infinite ordinal and cardinal; ω_1 is the first uncountable ordinal and the next cardinal after ω. For functions f, g, we use $\mathrm{Do}\, f$, f^{-1}, $f \circ g$, and $f * A$ respectively for the *domain of f*, the *inverse of f*, the *composition of f and g*, and the *image of A under f*. ${}^{A}B$ denotes the *set of functions from A into B*. In particular, ${}^{\omega}A$ is the *set of all denumerably infinite sequences with terms in A*; for $n \in \omega$, ${}^{n}A$ is the *set of all n-termed sequences*; ${}^{\underline{\omega}}A$ denotes the *set of all finite sequences with terms in A*. For arbitrary relations R, we also use $R * A$ for the *image of A under R*.

We shall also study measures on Boolean algebras. By a *measure on a Boolean σ-algebra (a σ-BA)* $\mathfrak{B} = \langle B, V, \Lambda, -, 0, 1\rangle$, we understand a countably additive, nonnegative real function on B that assumes the value one at the unit of the algebra.

If $x \in {}^{\omega}B$, we write $\bigvee\{x_i : i \in \omega\}$ and $\bigwedge\{x_i : i \in \omega\}$ for the *least upper bound* (*l.u.b.*) and *greatest lower bound* (*g.l.b.*) of the sequence x. A *σ-field of subsets of a set X*, $\mathfrak{F} = \langle F, \cup, \cap, -, \emptyset, X\rangle$, is a particular kind of σ-BA in which the universe F consists of subsets of X, and the operations are set-theoretic union, intersection, and complement with respect to X. L.u.b.'s and g.l.b.'s of denumerable sequences coincide with countable unions and intersections. A *group* (or more properly, a *quasi-group*) *of transformations of the σ-field of sets F* is a set G of one-one functions such that:

(i) If $A, B \in F$, $f \in G$, $A \subseteq \mathrm{Do}\, f$, $B \subseteq \mathrm{Do}\, f^{-1}$, then $f * A$, $f^{-1} * B \in F$.

(ii) The identity function restricted to X belongs to G.

(iii) If $f, g \in G$, then f^{-1} and $f \circ g \in G$.

Notice that functions in G are not supposed to have a common domain (see **[14]**, p. 221).

A measure μ on $\mathfrak{F}$ is said to be *invariant under G*, or *G-invariant*, if for any $A, B \in F$ such that there is an $f \in G$ with $A \subseteq \mathrm{Do}\, f$ and $B = f * A$, we have $\mu(A) = \mu(B)$. Our problem, then, is to find necessary and sufficient conditions on $\mathfrak{F}$ and G for the existence of such measures. It is more convenient to work with equivalence relations on BA's thus, we define the equivalence relation $\sim_G$ on F:

$A \sim_G B$ iff there is an $f \in G$ such that $A \subseteq \mathrm{Do}\, f$ and $B = f * A$. It is clear that μ is G-invariant iff:

For any $A, B \in F$, $A \sim_G B$ implies $\mu(A) = \mu(B)$.

In general, for any equivalence relation R on a BA $\mathfrak{B}$ and measure μ on $\mathfrak{B}$, we say that μ is *R-invariant* if for any $a, b \in B$, we have

$$a R b \quad \textit{implies} \quad \mu(a) = \mu(b).$$

If a measure μ on F is G-invariant, then it also has to be $\simeq_G$-invariant for the equivalence relation on F, $\simeq_G$, defined by: $A \simeq_G B$ iff there are sequences of disjoint elements $Y, Z \in {}^{\omega}F$, such that $A = \bigcup\{Y_i : i \in \omega\}$, $B = \bigcup\{Z_i : i \in \omega\}$, and $Y_i \sim_G Z_i$ for every $i < \omega$.

It is easy to see, that if μ is G-invariant, then for any $A, B \in F$ we have,

$$A \underset{G}{\simeq} B \quad \text{implies} \quad \mu(A) = \mu(B).$$

It is convenient to introduce the *disjunctive BA $\dot{\mathfrak{B}}$ associated with a σ-BA $\mathfrak{B}$*. Disjunctive BA's were introduced in Def. 15.14T. For any σ-BA $\mathfrak{B}$, the disjunctive BA associated with $\mathfrak{B}$ is the partial algebra $\dot{\mathfrak{B}} = \langle B, +, \Sigma\rangle$ where $+$ is a binary partial operation and Σ a countable partial operation defined by:

(a) For any $a, b, c \in B$, $a + b = c$ iff $a \vee b = c$ and $a \wedge b = 0$.

(b) For any $x \in {}^{\omega}B$ and $c \in B$, $\Sigma_{i<\omega} x_i = c$ iff $\vee\{x_i : i \in \omega\} = c$ and $x_i \wedge x_j = 0$ for $i < j < \omega$.

This disjunctive BA $\dot{\mathfrak{B}}$ is a generalized cardinal algebra (GCA) by 15.24T. We shall use GCA's and cardinal algebras (CA) throughout this paper. The terminology will be taken from **[14]** with a few exceptions that will be noted in the appropriate places.

A *congruence relation* R on a GCA $\mathfrak{A} = \langle A, +, \Sigma \rangle$ is an equivalence relation that satisfies:

(i) if $a, b, c, d, a + b, c + d \in A$; aRc, and bRd, then $a + b\, R\, c + d$;

(ii) if $x, y \in {}^{\omega}A$, $\Sigma_{i<\omega} x_i$, $\Sigma_{i<\omega} y_i \in B$, and $x_i R y_i$ for every $i < \omega$ then $\Sigma_{i<\omega} x_i\, R\, \Sigma_{i<\omega} y_i$.

Congruence relations are called in **[14]**, infinitely additive equivalence relations (see 6.4T).

A *refining relation* R on a GCA $\mathfrak{A}$ is a relation on A such that:

> if $a, x_0, x_1, b \in A$, $a = x_0 + x_1$, and aRb, then there are $y_0, y_1 \in A$ such that $b = y_0 + y_1$, $x_0 R y_0$, and $x_1 R y_1$.

Refining relations are called finitely refining in **[14]** (see 6.7T).

By 16.6T, $\approx_G$ is a refining congruence relation on the GCA (disjunctive BA) $\dot{\mathfrak{F}}$.

The main purpose of this paper is to give necessary and sufficient conditions for the existence of G-invariant measures on $\mathfrak{F}$. However, it is more convenient to work in a more general setting and find measures μ on a σ-BA, $\mathfrak{B}$ that are R-invariant for R a refining congruence relation on $\mathfrak{B}$. I shall deal with this problem in the following sections, returning to $\mathfrak{F}$ in the last section.

2. Ideals and congruence relations. In this section I shall prove some lemmas, which will be needed later, about ideals and congruence relations in σ-BA's and GCA's. For any σ-BA $\mathfrak{B}$ we have the corresponding disjunctive BA $\dot{\mathfrak{B}}$, which is a GCA. The notion of an ideal in a GCA is defined and discussed in (**[14]**, Chapter 9). I shall call ideals in a GCA *cardinal ideals* to distinguish them from ideals in a BA (see **[13]** for ideals in BA's). There are, then, two notions of ideals in $\mathfrak{B}$: σ-ideals in $\mathfrak{B}$ as a BA and cardinal ideals in $\dot{\mathfrak{B}}$ as a GCA. The first lemma proves that they coincide.

LEMMA 2.1. *Let $\mathfrak{B}$ be a σ-BA, $I \subseteq B$. Then, I is a σ-ideal in $\mathfrak{B}$ iff I is a cardinal ideal in $\dot{\mathfrak{B}}$.*

Proof. Suppose I is a cardinal ideal in $\dot{\mathfrak{B}}$. It is clear that the partial ordering in $\mathfrak{B}$ coincides with that in $\dot{\mathfrak{B}}$. Thus, if $a \in I$ and $b \leqq a$ (in $\mathfrak{B}$), then $b \in I$. Suppose, now that $x \in {}^{\omega}I$. Define

$$x'_0 = x_0$$
$$x'_{n+1} = x_{n+1} - \bigvee\{x_i : i \leqq n\}.$$

It is clear that $x' \in {}^{\omega}I$ and $x'_i \wedge x'_j = 0$ for $i \neq j$. Then,

$$\bigvee\{x_i : i \in \omega\} = \Sigma_{i<\omega} x'_i \in I.$$

Thus, I is a σ-ideal in $\mathfrak{B}$.

The converse implication is obvious.

The next lemma proves that the equivalence relation determined by a σ-ideal I on a σ-BA $\mathfrak{B}$ is the same as the relation determined by I as a cardinal ideal in $\dot{\mathfrak{B}}$ (cf. 9.26T).

LEMMA 2.2. *Let $\mathfrak{B}$ be a σ-BA, I a σ-ideal in $\mathfrak{B}$. Then:*

(i) *for any $a, b \in B$, $a - b \vee b - a \in I$ iff there are $a', b' \in I$ and $c \in B$ such that $a = c + a'$ and $b = c + b'$;*

(ii) *$(\dot{\mathfrak{B}}/I) = \dot{\mathfrak{B}}/I$, i.e. if $x \in {}^{\omega}B$, $c \in B$, then $\Sigma_{i<\omega}(x_i/I) = c/I$ iff there is an $x' \in {}^{\omega}B$ such that $x'_i \wedge x'_j = 0$, $x_i/I = x'_i/I$, and $\Sigma_{i<\omega}(x_i/I) = (\Sigma_{i<\omega} x'_i)/I$, for $i < j < \omega$.*

Proof. (i) (1) Suppose $a - b \vee b - a \in I$. Take $a' = a - b$, $b' = b - a$, $c = a \wedge b$.

(2) Suppose $a = a' + c$, $b = b' + c$, and $b', a' \in I$. Then $a - b \vee b - a \leqq a' \vee b' \in I$.

(ii) Suppose $\Sigma_{i<\omega}(x_i/I) = c/I$. Then $x_i/I \wedge x_j/I = 0/I$ and so, $x_i \wedge x_j \in I$ for $i < j < \omega$. Let $d = \bigvee\{x_i \wedge x_j : i < j < \omega\}$ and $x'_i = x_i - d$ for $i < \omega$. Then $d \in I$, $x'_i \wedge x'_j = 0$, $x'_i/I = x_i/I$, and $\Sigma_{i<\omega}(x'_i/I) = (\Sigma_{i<\omega} x'_i)/I = c/I$ for $i < j < \omega$.

We are interested in refining congruence relations on disjunctive BA's $\dot{\mathfrak{B}}$, and how they behave when we pass to $\dot{\mathfrak{B}}/I$ for an ideal I. This is given in the following lemma:

LEMMA 2.3. *Let $\mathfrak{B}$ be a σ-BA, I a σ-ideal on $\mathfrak{B}$, and R a refining congruence relation on $\dot{\mathfrak{B}}$ such that $R * I \subseteq I$* (i.e. if $a \in I$ and aRb, then $b \in I$). *Define $\bar{R}$ on $\dot{\mathfrak{B}}/I$ by:*

$a/I\,\bar{R}\,b/I$ iff there are $a', b' \in B$ such that $a/I = a'/I$, $b/I = b'/I$, and $a'Rb'$.

Under these conditions, $\bar{R}$ is a refining congruence relation on $\mathfrak{B}/I$.

Proof. It is easy to see that $\bar{R}$ is an equivalence relation on B. We have to prove that $\bar{R}$ is refining and preserves Σ. Let

(1) $a/I\bar{R}\,b/I$,

(2) $a/I = x_1/I + x_2/I$.

By (2) and 2.2(ii), there are disjoint x_1', x_2' such that

(3) $x_1/I = x_1'/I$ and $x_2/I = x_2'/I$.

Also, by (1), there are a', b' such that $a/I = a'/I$, $b/I = b'/I$, and $a'Rb'$. By (3), $a'/I = (x_1' + x_2')/I$. Thus, there are x_1'', x_2'' such that,

(4) $a' = x_1'' + x_2''$, $x_1/I = x_1''/I$, and $x_2/I = x_2''/I$.

Since R is refining and (4), $b' = y_1 + y_2$, $x_1''Ry_1$, and $x_2''Ry_2$. Then, $b/I = y_1/I + y_2/I$, $y_1/I\bar{R}x_1/I$, and $y_2/I\bar{R}x_2/I$. Thus, $\bar{R}$ is refining.

Suppose now that $x, y \in {}^{\omega}B$ with,

(5) $x_i/I\bar{R}\,y_i/I$ for all $i < \omega$;

(6) $\Sigma_{i<\omega}(x_i/I)$ and $\Sigma_{i<\omega}(y_i/I)$ exist in B/I. From (5) we obtain $x', y' \in {}^{\omega}B$ such that,

(7) $x_i'/I = x_i/I$, $y_i'/I = y_i/I$, and $x_i'Ry_i'$, for all $i < \omega$.

From (6) we deduce that $x_i/I \wedge x_j/I = 0/I = y_i/I \wedge y_j/I$ for all i, j, with $i < j < \omega$. Then, the same is true for x', y'. Thus

(8) $x_i' \wedge x_j' \in I$, for all $i, j < \omega$, $i \neq j$.

Take $c = \vee\{x_i' \wedge x_j' : i < j < \omega\}$. Then $c \in I$. Define $x_i'' = x_i' - c$. Then $x_i''/I = x_i'/I = x_i/I$. Also,

(9) $x_i'' \wedge x_j'' = 0$ for $i < j < \omega$.

We have, $x_i' = x_i'' + z_i$ where $z_i = c \wedge x_i'$ and $z_i \in I$. From (7) we obtain $y'', z' \in {}^{\omega}B$ such that,

(10) $y_i' = y_i'' + z_i'$, $y_i''Rx_i''$, and $z_i'Rz_i$.

Since $R * I \subseteq I$, $z_i' \in I$ and

(11) $y_i''/I = y_i'/I = y_i/I$.

Similarly as (7) we get,

(12) $y_i'' \wedge y_j'' \in I$ for $i < j < \omega$.

Take $d = \vee\{y_i'' \wedge y_j'' : i < j < \omega\}$ $(\in I)$ and $y_i''' = y_i'' - d$. We have,

(13) $y_i'''/I = y_i''/I = y_i/I$ and $y_i''' \wedge y_j''' = 0$ for $i < j < \omega$. Also, $y_i'' = y_i''' + u_i$ where $u_i = d \wedge y_i'' \in I$. From (9) we obtain,

(14) $x_i'' = x_i''' + u_i'$, $x_i'''Ry_i'''$, and $u_i'Ru_i$ for all $i < \omega$. Thus, since $R * I \subseteq I$, $u_i' \in I$ and,

(15) $x_i'''/I = x_i''/I = x_i/I$ for all $i < \omega$.

Since $x_i''' \leqq x_i''$, (9) implies:

(16) $x_i''' \wedge x_j''' = 0$ for $i < j < \omega$.

Since R is a congruence relation on $\mathfrak{B}$, (13), (14), and (16) imply that $(\Sigma_{i<\omega}x_i''')\,R\,(\Sigma_{i<\omega}y_i''')$. But then from (15) and (13) we obtain the desired conclusion; i.e.

$$(\Sigma_{i<\omega}x_i/I)\,\bar{R}\,(\Sigma_{i<\omega}y_i/I).$$

The following lemma relates cardinal ideals in the GCA's $\mathfrak{A}$ and $\mathfrak{A}/R$.

LEMMA 2.4. *Let $\mathfrak{A}$ be a* GCA, *R a refining congruence relation on $\mathfrak{A}$, $I \subseteq A$, and $R * I \subseteq I$. Then I is a cardinal ideal in $\mathfrak{A}$ iff I/R is a cardinal ideal in $\mathfrak{A}/R$.*

Proof. (1)Suppose I is a cardinal ideal in $\mathfrak{A}$. Let $a/R \in I/R$, $a \in I$, and $b/R \leqq a/R$. Then, there are $a', b' \in A$ with $b' \leqq a'$, $a R a'$, and $b R b'$. Thus, since $R * I \subseteq I$, we have that $a', b' \in I$. But, then, $b'/R = b/R \in I/R$.

Let, now, $x \in {}^{\omega}I$ and $\Sigma_{i<\omega} (x_i/R) \in A/R$, i.e. $\Sigma_{i<\omega} (x_i/R) = b/R$ for some $b \in A$. Then, there are $x' \in {}^{\omega}A$, and $b' \in A$ such that $\Sigma_{i<\omega} x'_i = b'$, $x_i R x'_i$, and $b R b'$, for all $i < \omega$. Since $R * I \subseteq I$, $x'_i \in I$ for $i < \omega$. Then $\Sigma_{i<\omega} x'_i \in I$ and, thus $b/R = b'/R \in I/R$.

(2) Suppose, now, that I/R is a cardinal ideal in $\mathfrak{A}/R$. Let $a \in I$, and $b \leqq a$. Then $a/R \in I/R$, and $b/R \leqq a/R$. Thus, $b/R \in I/R$, i.e. there is a $b' \in I$ such that $b' R b$. But, then from $R * I \subseteq I$, we get $b \in I$.

Suppose $x \in {}^{\omega}I$ with $\Sigma_{i<\omega} x_i \in A$. Then $x_i/R \in I/R$ for all $i < \omega$, and $\Sigma_{i<\omega} (x_i/R) \in I/R$. Thus, $\Sigma_{i<\omega} (x_i/R) = (\Sigma_{i<\omega} x_i)/R \in I/R$. Therefore, there is a $b \in I$ such that $\Sigma_{i<\omega} x_i R b$. But then, from $R * I \subseteq I$, we get that $\Sigma_{i<\omega} x_i \in I$.

3. The countable chain condition. Let $\mathfrak{A}$ be a GCA. We say that a subset $B \subseteq A$ is *bounded* if there is an $a \in A$ such that $b \leqq a$ for every $b \in B$. *$\mathfrak{A}$ satisfies the countable chain condition* (*ccc*) if every bounded subset $B \subseteq A$ well ordered by the relation $\leqq$ is at most countable. It is clear that if $\mathfrak{B}$ is a σ-BA, then the GCA $\dot{\mathfrak{B}}$ satisfies the *ccc* if and only if $\mathfrak{B}$ satisfies the countable chain condition in the usual Boolean sense. **[5]** calls a GCA that satisfies the *ccc*, separable. In this section I shall prove that the *ccc* is transmitted through several constructions of GCA's.

If $a, b \in A$, we write $a \wedge b$, $a \vee b$ for the g.l.b. and l.u.b. of a, b; if $x \in {}^{\omega}A$, $\wedge_{i\in\omega} x_i$ and $\vee_{i\in\omega} x_i$ stand for the g.l.b. and l.u.b. of the sequence x (see Defs. 3.1T and 3.2T where slightly different symbols are employed).

THEOREM 3.1. *Let $\mathfrak{A}$ be a* GCA *satisfying the ccc, and R a refining, congruence relation on $\mathfrak{A}$. Then $\mathfrak{A}/R$ also satisfies the ccc.*

Proof. Suppose the sequence $y \in {}^{\omega_1}(A/R)$ is such that for $\beta < \alpha < \omega_1$, we have $y_\beta < y_\alpha \leqq a/R$ for some $a \in A$. Choose a sequence $x \in {}^{\omega_1}A$, such that $y_\alpha = x_\alpha/R$ and $x_\alpha \leqq a$ for every $\alpha \in \omega_1$. It is possible to obtain such an x because R is refining.

We shall first prove:

(1) For any $b, c \in A$ with $b, c \leqq a$, we have $b/R < c/R$ iff there is a $d \in A$ such that $d \leqq a$, $b < d$, and $d R c$.

From $b/R < c/R$ we get b' with $b' < c$ and $b R b'$. Thus, we have, $b' < c \leqq a$, $b \leqq a$, $b R b'$, and $a R a$. Using 7.13T (vi) we obtain a d with $b \leqq d \leqq a$ and $d R c$. Now, if $b = d$, we would have $c/R = d/R = b/R$ contradicting $b/R < c/R$. Thus $b < d$.

The converse implication is proved similarly, and, thus, (1) is proved. We pass now, to the proof of the theorem:

Define by transfinite recursion the sequence $z \in {}^{\omega_1}A$ satisfying:

(2) $y_\alpha = x_\alpha/R = z_\alpha/R$, and if $\alpha < \beta < \omega$, then $z_\alpha < z_\beta \leqq a$, as follows:

(a) $z_0 = x_0$.

(b) Suppose z_α is defined. We have $z_\alpha/R < x_{\alpha+1}/R$, and $z_\alpha, x_{\alpha+1} \leqq a$. By (1) there is $d \in A$ such that $z_\alpha < d \leqq a$, and $d R x_{\alpha+1}$. Take $z_{\alpha+1} = d$.

(c) Let α be a limit ordinal, $\alpha < \omega_1$, and z_β defined for all $\beta \in \alpha$. Let $f \in {}^\omega\alpha$ be a strictly increasing sequence of ordinals cofinal in α. We shall prove that

(3) $(\bigvee_{i\in\omega} z_{f(i)})/R = \bigvee_{\beta\in\alpha} y_\beta$.

It is enough to prove,

$$\left(\bigvee_{i\in\omega} z_{f(i)}\right)/R = \bigvee_{i\in\omega} (z_{f(i)}/R).$$

Let $z_{f(i)} + u_i = z_{f(i+1)}$ for every $i < \omega$. Then

$$z_{f(i)}/R + u_i/R = z_{f(i+1)}/R \quad \text{for every } i < \omega.$$

Also, $\bigvee_{i<\omega} z_{f(i)} = z_{f(0)} + \Sigma_{i<\omega} u_i$. Thus,

$$\left(\bigvee_{i\in\omega} z_{f(i)}\right)/R = z_{f(0)}/R + \Sigma_{i<\omega} (u_i/R)$$

$$= \bigvee_{i<\omega} (z_{f(i)}/R).$$

So (3) is proved.

Let now, $b = \bigvee_{i\in\omega} z_{f(i)}$. Then, since $b/R = \bigvee_{\beta\in\alpha} y_\beta \leqq y_\alpha$, and $b \leqq a$, by (1) there is a $d \in A$ with $b \leqq d \leqq a$ and $d R x_\alpha$. Let, then, $z_\alpha = d$.

With this, we complete definition (2) and, thus, obtain a $\leqq$-well-ordered bounded subset of A of type ω_1, contradicting the *ccc* for $\mathfrak{A}$.

We pass now, to show preservation of the *ccc* under another important operation on GCA's: the *cardinal product* (see Def. 6.11T).

THEOREM 3.2. *Let $\mathfrak{A}_i$ be GCA's that satisfy the ccc for every $i \in I$. Then $\Pi_{i\in I} \mathfrak{A}_i$ also satisfies the ccc.*

Proof. Suppose $y \in {}^{\omega_1}(\Pi_{i \in I} A_i)$ and $a \in \Pi_{i \in I} A_i$ be such that $y_\beta \leqq y_\alpha \leqq a$ for $\beta < \alpha < \omega_1$. We then have $y_{\beta i} \leqq y_{\alpha i} \leqq a_i$ for every $i \in I$.

The sets $B_i = \{y_{\alpha i} : \alpha \in \omega_1\} \subseteq A_i$ are bounded and $\leqq$-well-ordered for all $i \in I$; then, $\operatorname{card}(B_i) \leqq \omega$ for all $i \in I$.

Since $\operatorname{card}(B_i) \leqq \omega$ for every $i \in I$, there is an $\alpha \in \omega_1$ such that $y_{\alpha i} = y_{\beta i}$ for all $\beta \geqq \alpha$. Define $\gamma(i)$ as the least such α, i.e.,

$$\gamma(i) = \bigcap \{\alpha : y_{\alpha i} = y_{\beta i} \text{ for all } \beta \geqq \alpha\}.$$

We have, $\gamma(i) \in \omega_1$.

Let $J = \{i : a_i \neq 0_i\}$. From the definition of cardinal product, $\operatorname{card}(J) \leqq \omega$. Also, for every $\alpha \in \omega_1$, $y_{\alpha i} = 0_i$ for all $i \in I - J$. Let δ be the least upper bound of $\gamma(i)$ for $i \in J$; i.e.

$$\delta = \bigcup \{\gamma(i) : i \in J\}.$$

$\delta \in \omega_1$, and $y_\alpha = y_\delta$ for all $\alpha \geqq \delta$. Therefore, $\operatorname{card}\{y_\alpha : \alpha \in \omega_1\} = \omega$.

COROLLARY 3.3. *Let $\mathfrak{A}$ be a* GCA *that satisfies the ccc. Then ${}^{\omega}\mathfrak{A}$ also satisfies the ccc.*

Every GCA $\mathfrak{A}$ can be closed to obtain a CA $\bar{\mathfrak{A}}$ which preserves most of the properties of $\mathfrak{A}$. In [**14**, Ch. 7], these closures are studied. We prove, now, that closures preserve the *ccc*.

THEOREM 3.4. *Let $\mathfrak{A}$ be a* GCA *that satisfies the ccc. Then $\bar{\mathfrak{A}}$, a closure of $\mathfrak{A}$, also satisfies the ccc.*

Proof. By 7.7T, $\bar{\mathfrak{A}}$ is isomorphic to ${}^{\omega}\mathfrak{A}/R$ where R is a refining congruence relation in $\mathfrak{A}$. Thus, from 3.1 and 3.3 we obtain 3.4.

COROLLARY 3.4. *Let $\mathfrak{B}$ be a σ-BA that satisfies the ccc, R a refining congruence relation on $\dot{\mathfrak{B}}$, and $\mathfrak{A} = \overline{\dot{\mathfrak{B}}/R}$ (a closure of $\dot{\mathfrak{B}}/R$). Then $\mathfrak{A}$ is a* CA *such that for any $x \in {}^{\omega}A$ we have $\wedge_{i \in \omega} x_i \in A$.*

Proof. $\mathfrak{A}$ satisfies the *ccc* by 3.1 and 3.4. Hence applying 3.35T we obtain the conclusion.

4. Invariant measures in Boolean algebras. In this section we prove some theorems about the existence of R-invariant measures on a σ-distributive σ-*BA* $\mathfrak{B}$ where R is a refining congruence relation on $\dot{\mathfrak{B}}$. In the next section we apply the theorems to obtain G-invariant measures on σ-fields of sets.

For some of the following definitions see [**10**, **13** p.p. 62, 204].

Let $\mathfrak{B}$ be a σ-BA and $x \in {}^nB$ for some $n \in \omega$. We define $i(x) = m/n$ where m is the largest integer $k \leqq n$ such that

$$x_{i_0} \wedge \cdots \wedge x_{i_{k-1}} \neq 0 \quad \text{for} \quad 0 \leqq i_0 < i_1 < \cdots < i_{k-1} < n.$$

Then, if $A \subseteq B$ we define the *intersection number of* A:

$$i(A) = \inf\{i(x)\colon x \in {}^nA \text{ for some } n \in \omega\}.$$

We say that $\mathfrak{B}$ *has the Kelley property* if $B - \{0\}$ is a countable union of sets with positive intersection number. $\mathfrak{B}$ is σ*-distributive* if for every double sequence $x \in {}^{\omega\times\omega}B$ we have,

$$(*) \qquad \vee\{\wedge\{x_{ij} : j \in \omega\} : i \in \omega\} = \wedge\{\vee\{x_{i,\phi(i)} : i \in \omega\} : \phi \in {}^{\omega}\omega\}.$$

We say that $\mathfrak{B}$ is *weakly* σ*-distributive* if $(*)$ is satisfied for every double sequence $x \in {}^{\omega\times\omega}B$ such that $x_{i,j+1} \leqq x_{ij}$ for every $i, j \in \omega$.

Let $I \subseteq B$ and μ a measure on $\mathfrak{B}$. We say that μ is I*-positive* if we have,

$$\mu(a) = 0 \quad \text{iff} \quad a \in I, \quad \text{for every } a \in B.$$

Let $a \in B$; we say that a is R*-negligible* if there is a sequence of disjoint elements of B, $x \in {}^{\omega}B$, such that $x_i R a$ for every $i \in \omega$. Let $N_R = \{a\colon a \text{ is } R\text{-negligible}\}$. It is clear that if the measure μ is R-invariant, then if $a \in N_R$, we must have $\mu(a) = 0$.

We need the following lemma about N_R:

LEMMA 4.1. *Let $\mathfrak{B}$ be a σ-BA, and R a refining congruence relation on $\dot{\mathfrak{B}}$. Then N_R is a σ-ideal in $\mathfrak{B}$.*

Proof. By 2.1 it is enough to prove that N_R is a cardinal ideal in $\dot{\mathfrak{B}}$. Consider $\mathfrak{A} = \dot{\mathfrak{B}}/R$; $\mathfrak{A}$ is a GCA. Let $h = 1/R$ and

$$A(h) = \{c \in \mathfrak{A}\colon c + h = h\}.$$

Then, by 9.15T, $A(h)$ is a cardinal ideal. It is easy to see that $R * (N_R) \subseteq N_R$. Hence by 2.4 it is enough to prove that $A(h) = N_R/R$.

Since $c \in A(h)$ iff $\infty c \leqq h$ (by 1.29T) the proof is reduced to:

$$a \in N_R \quad \text{iff} \quad \infty(a/R) \leqq h.$$

Suppose, first, that $a \in N_R$. Let $x \in {}^{\omega}B$ be such that $x_i \wedge x_j = 0$ and $x_i R a$ for every $i < j < \omega$.

We have, $a/R = x_i/R$, for every $i < \omega$. But $\Sigma_{i<\omega} x_i \leqq 1$. Thus, $\infty(a/R) = \Sigma_{i<\omega}(x_i/R) \leqq 1/R = h$.

Suppose, now, that $\infty(a/R) \leqq h$. Since $h \in \mathfrak{A}$, $\infty(a/R) \in \mathfrak{A}$ by 7.4T. Then, $\infty(a/R) = b/R$ for some $b \in B$. From the definition of the coset algebra (Def. 6.3T), there is a sequence of disjoint elements $x \in {}^{\omega}B$ and a $c \in B$ such that $\Sigma_{i<\omega} x_i = c$ and $x_i R a$ for all $i < \omega$. Thus, $a \in N_R$.

We formulate, now, the main theorem of this section:

THEOREM 4.2. *Let $\mathfrak{B}$ be a σ-distributive σ-BA, R a refining congruence relation on $\dot{\mathfrak{B}}$, and I a subset of B. The following conditions are necessary and jointly sufficient for the existence of a countably additive, I-positive, and R-invariant measure on $\mathfrak{B}$:*

(i) *I is a proper σ-ideal;*
(ii) $N_R \subseteq I$;
(iii) $R * I \subseteq I$;
(iv) *$\mathfrak{B}/I$ has the Kelley property and is weakly σ-distributive.*

For the proof we need a result of **[6]**. We have to introduce some definitions. Let X be the Stone space of a σ-BA and $\bar{\mathbf{R}}$ the nonnegative real numbers with ∞. Then $\mathscr{C}(X, \bar{\mathbf{R}})$ denotes the set of continuous functions on X with compact support and values in $\bar{\mathbf{R}}$. $\langle \mathscr{C}(X, \bar{\mathbf{R}}), +, \Sigma \rangle$ is a CA (cf **[6]**, p. 31) where $+$ is pointwise addition, and $\Sigma_{i<\omega} f_i$ is the continuous limit of the partial sums $\Sigma_{i<n} f_i$. This limit differs from the pointwise limit in a set of first category.

Proof of 4.2. The necessity of the conditions is easy to prove. We must use Kelley's necessary and sufficient condition for the existence of a strictly positive measure on a BA (see **[10]**, Th. 9, and **[8]**, Th. 3.7).

We proceed, now, to construct the desired measure. For the rest of this section let $\mathfrak{B}$, R, and I satisfy (i)–(iv). Define the relation $\bar{R}$ on $\mathfrak{B}/I$ by:

$$a/I \, \bar{R} \, b/I \text{ iff these are } a', b' \in B \text{ such that } a/I = a'/I,$$

$$b/I = b'/I, \quad \text{and} \quad a' R b'.$$

By 2.3, $\bar{R}$ is refining congruence relation on $\dot{\mathfrak{B}}/I$. Let $\mathfrak{A} = \overline{(\dot{\mathfrak{B}}/I)/\bar{R}}$, i.e. is a CA that is a closure of the GCA $(\dot{\mathfrak{B}}/I)/\bar{R}$.

Let $h = (1/I)/\bar{R}$. We shall prove that:

LEMMA 4.3. *h is finite* (Def. 4.10T: $h + a = h$ implies that $a = 0$ for all $a \in A$). *In fact, we have that $h + a = h$ and $a = c/I/\bar{R}$, implies that $c/I = 0/I$.*

Proof. Suppose $h + a = h$. Then by 1.29T, $\infty a \leqq h$. Let $a = (c/I)/\bar{R}$ with $c \in B$. By the refining property of $\bar{R}$ and 2.2 we obtain a sequence of disjoint elements $x \in {}^\omega B$ such that $c/I \bar{R} x_i/I$ for all $i < \omega$. From the definition of $\bar{R}$ we get $x', z \in {}^\omega B$ such that $c/I = z_i/I$, $x_i/I = x'_i/I$, and $z_i R x'_i$ for every $i < \omega$. Thus $c = u_i + v_i$, $z_i = u_i + t_i$, $x_i = y_i + y'_i$, and $x'_i = y_i + y''_i$ with $v_i, t_i, y'_i, y''_i \in I$. Let $d = \wedge\{u_i : i \in \omega\}$. Then as $\mathfrak{B}$ is σ-distributive, $c = d + e$ with $e \in I$. Now, $d R s_i$ with $s_i \leqq x'_i$ for every $i < \omega$. Thus, by 2.4T, $s_i = s'_i + s''_i$ with $s'_i \leqq y_i$ and $s''_i \leqq y''_i$. Then, $s''_i \in I$ for every $i < \omega$. Therefore, $d = r_i + r'_i$ with $r_i R s'_i$ and $r'_i R s''_i$. Since $R * I \subseteq I$, $r'_i \in I$ for every $i < \omega$. Thus by σ-distributivity, $d = \wedge\{r_i : i \in \omega\} + f$ where $f \in I$. We have that s' is a sequence of disjoint elements of ${}' B$. Thus, $\wedge\{r_i : i \in \omega\} R p_i \leqq s'_i$ for every $i < \omega$; so, $\wedge\{r_i : i \in \omega\}$ is R-negligible and, therefore it belongs to I. Then $d \in I$ and $c \in I$. Therefore, $c/I = 0/I$ and $a = 0$. Thus, we have proved 4.3.

$\mathfrak{A}$, also, has the following two properties:

(1) $a \wedge b \in A$, for every $a, b \in A$.

This is obtained from 3.4, because $\mathfrak{B}/I$ satisfies the *ccc*.

(2) For every $a \in A$, $a \leqq \infty h$.

(2) is obtained from 7.1T.

Let $\mathfrak{E}$ be the σ-BA of idemmultiple elements of $\mathfrak{A}$ (a is idemmultiple if $a = a + a$; see Def. 4.1T and 8.3T for this algebra). Let X be the Stone space of $\mathfrak{E}$. 3.11 of Fillmore 1965 implies that $\mathfrak{A}$ is isomorphic to a subalgebra of $\langle \mathscr{C}(X, \bar{\mathbf{R}}), +, \Sigma \rangle$, say by a function F. For each element $a \in A$, $F(a)$ has support $E(a)$, the open-closed set corresponding to ∞a. Also, $F(h)$ is the characteristic function of $E(h) = X$.

There is a strictly positive measure μ on $\mathfrak{E}$ with $\mu(\infty h) = 1$. In order to prove this, we need the following two lemmas:

LEMMA 4.4. *If $\mathfrak{B}/I$ has the Kelley property, then $\mathfrak{E}$ also has the Kelley property.*

Proof. Suppose $B/I - \{0/I\} = \cup\{B_n : n \in \omega\}$ where each B_n has a positive intersection number. Let,

$$E_n = \{a : a \in E \text{ and there is a } b \in B_n \text{ such that } b/\bar{R} \leqq a\}.$$

In this definition, $\leqq$ is the partial ordering of $\mathfrak{A}$.

Let $x \in {}^n E_m$ and $y \in {}^n B_m$ be such that $y_k/\bar{R} \leqq x_k$ for all $k < n$. Let $i(x) = m_x/n$ where m_x is the largest k such that

$$x_{i_0} \wedge \cdots \wedge x_{i_{k-1}} \neq 0 \quad \text{for} \quad i_0 < \cdots < i_{k-1} < n; \quad i(y) = m_y/n$$

where m_y is defined similarly. We shall prove that $m_x \geqq m_y$.

Suppose that $k \leqq m_y$. Then, there are $y_{i_0}, \cdots, y_{i_{k-1}}$ such that $y_{i_0} \wedge \cdots \wedge y_{i_{k-1}} \neq 0$. Thus,

$$\infty(y_{i_0} \wedge \cdots \wedge y_{i_{k-1}})/\bar{R} \neq 0.$$

But

$$\begin{aligned}\infty(y_{i_0} \wedge \cdots \wedge y_{i_{k-1}})/\bar{R} &\leqq \infty(y_{i_0}/\bar{R}) \wedge \cdots \wedge \infty(y_{i_{k-1}}/\bar{R})\\ &\leqq x_{i_0} \wedge \cdots \wedge x_{i_{k-1}}\end{aligned}$$

so, $x_{i_0} \wedge \cdots \wedge x_{i_{k-1}} \neq 0$.

Thus, we have proved that $i(x) \geqq i(y)$. But

$$i(E_m) = \inf\{i(x)\colon x \in {}^{\omega}E_m\} \geqq \inf\{i(y)\colon y \in {}^{\omega}B_m\} = i(B_m) > 0.$$

Therefore, the lemma is proved.

LEMMA 4.5. *If $\mathfrak{B}/I$ is weakly σ-distributive, then $\mathfrak{E}$ is also weakly σ-distributive.*

Proof. In (**[13]**, Th. 30.1 (2)), it is proved that a σ-BA is weakly σ-distributive iff for every $x \in {}^{\omega\times\omega}E$ we have that:

$$\wedge\{\vee\{x_{ts} : s \in \omega\}: t \in \omega\} \neq 0$$

implies that there is a function ϕ from ω into the finite subsets of ω such that,

$$\wedge\{\vee\{x_{ts} : s \in \phi(t)\}: t \in \omega\} \neq 0.$$

Suppose, then, that $x \in {}^{\omega\times\omega}\mathscr{E}$ and

$$\wedge\{\vee\{x_{ts} : s \in \omega\}: t \in \omega\} \neq 0.$$

So, there is a $c \in B/I$, $c \neq 0$, such that,

$$c/\bar{R} \leqq \wedge\{\vee\{x_{ts} : s \in \omega\}: t \in \omega\}.$$

Since the elements x_{ts} of $\mathscr{E}$ are idemmultiple in $\mathfrak{A}$, by 4.7T we get,

$$c/\bar{R} \leqq \Sigma_{s\in\omega} x_{ts}, \quad \text{for every } t \in \omega.$$

From 2.2T we obtain,

$$c/\bar{R} = \Sigma_{s\in\omega} y_{ts}, \quad \text{with} \quad y_{ts} \leqq x_{ts} \quad \text{for every } t, s \leqq \omega.$$

We define, by recursion, for each $t \in \omega$, a sequence $z_t \in {}^{\omega}(B/I)$ as follows: We have,

$$c\,\bar{R}\,\Sigma_{s\in\omega} z'_{ts}, \quad \text{with} \quad z'_{ts}/\bar{R} = y_{ts}, \quad \text{for every } t, s \in \omega.$$

Then, as $\bar{R}$ is refining,

$$c = z_{t0} + u_{t0}, \quad z_{t0}\,\bar{R}\,z'_{t0}, \quad \text{and} \quad u_{t0}\,\bar{R}\,\Sigma_{s\in\omega} z'_{t,s+1},$$

for every $t \in \omega$.

Repeat the same procedure to obtain $z_{t,s+1}$ given $z_{t,s}$. We have, since $\bar{R}$ is a congruence relation,

$$(\Sigma_{s\in\omega} z_{ts})\,\bar{R}\,(\Sigma_{s\in\omega} z'_{ts})\,\bar{R}\,c.$$

Let $d = c - \Sigma_{s\in\omega} z_{ts}$. Then $c/\bar{R} + d/\bar{R} = c/\bar{R}$. Thus, $h + d/\bar{R} = h$ by 1.30T. By 4.3 we get that $d = 0/I$.

Therefore, $c = \Sigma_{s\in\omega} z_{ts}$ and $z_{ts}/\bar{R} = y_{ts}$ for all $t, s \in \omega$. Now, $\wedge\{\vee\{z_{ts} : s \in \omega\} : t \in \omega\} = \wedge\{c : t \in \omega\} = c \neq 0$. Since $\mathfrak{B}/I$ is weakly σ-distributive, there is a function ϕ from ω into the finite subsets of ω such that,

$$\wedge\{\vee\{z_{ts} : s \in \phi(t)\} : t \in \omega\} \neq 0.$$

So,

$$\begin{aligned} 0 \neq \wedge\{\vee\{z_{ts} : s \in \phi(t)\} : t \in \omega\}/\bar{R} &\leqq \wedge\{\Sigma_{s\in\phi(t)} y_{ts} : t \in \omega\} \\ \leqq \wedge\{\Sigma_{s\in\phi(t)} x_{ts} : t \in \omega\} &= \wedge\{\vee\{x_{ts} : s \in \phi(t)\} : t \in \omega\}. \end{aligned}$$

Therefore, the lemma is proved. We now continue with the proof of 4.2:

From (iv), 4.4 and 4.5 we obtain that $\mathfrak{C}$ has the Kelley property and is weakly σ-distributive. Then using (Th. 9 of **[10]**,), we obtain a strictly positive measure μ on $\mathfrak{C}$.

Let $\mathrm{Co}(X)$ be the family of open-closed sets of X. Define the measure $\bar{\mu}$ on $\mathrm{Co}(X)$ by $\bar{\mu}(E(a)) = \mu(\infty a)$. Since X is compact, $\bar{\mu}$ is countably additive as a measure on the field of sets $\mathrm{Co}(X)$. Extend $\bar{\mu}$ to the σ-field $\mathbf{B}(X)$ of subsets of X generated by $\mathrm{Co}(X)$ ($\mathbf{B}(X)$ are the Borel sets in X). Using normal measure-theoretic procedures, define an integral Π on all $\bar{\mu}$-measurable bounded functions. All functions in $\mathscr{C}(X, \bar{\mathbf{R}})$ are bounded $\bar{\mu}$-measurable.

Define λ on $\mathfrak{A}$ by, $\lambda(a) = \Pi(F(a))$. λ satisfies the following properties:

(3) $\lambda(h) = 1$.

(3) is proved by,

$$\lambda(h) = \Pi(F(h)) = \bar{\mu}(E(h)) = \mu(\infty h) = 1.$$

(4) $\lambda(a+b) = \lambda(a) + \lambda(b)$, for all $a, b \in A$.
(4) is proved by,

$$\lambda(a+b) = \Pi(F(a)+F(b)) = \Pi(F(a)) + \Pi(F(b)) = \lambda(a) + \lambda(b).$$

(5) If $x \in {}^{\omega}A$ is a nondecreasing sequence, and $a = \bigvee_{i\in\omega} x_i$, then $\lim_{i\to\infty} \lambda(x_i) = \lambda(a)$.

Proof of (5). Let $x \in {}^{\omega}A$, $x_i \leqq x_{i+1}$ for all $i < \omega$, and $a = \bigvee_{i<\omega} x_i$. Then

$$\lim_{i\to\infty} \lambda(x_i) = \lim_{i\to\infty} \Pi(F(x_i)) = \Pi(\lim_{i\to\infty} F(x_i)), \quad \text{where} \quad \lim_{i\to\infty} F(x_i) = f$$

is the pointwise limit of the functions $F(x_i)$'s (this limit exists, because $F(x_i) \leqq F(x_{i+1})$ for all $i < \omega$). We have that $E(x_i)$ is the support of $F(x_i)$. Thus, if $y \notin \bigcup\{E(x_i): i \in \omega\}$, then $f(y) = 0$.

We also have that $\bigvee_{i\in\omega} F(x_i) = F(a)$, where the l.u.b. is taken in $\langle \mathscr{C}(X, \bar{\mathbf{R}}), +, \Sigma\rangle$. The support of $F(a)$ is $E(a)$. Now, $\mu(\infty a) = \lim_{i\to\infty} \mu(\infty x_i)$, because $\infty a = \bigvee_{i\in\omega} \infty x_i$ and $\infty x_i \leqq \infty x_{i+1}$ for all $i \in \omega$. Thus, $\bar{\mu}(E(a)) = \lim_{i\to\infty} \bar{\mu}(E(x_i))$. Also,

$$\bar{\mu}(\bigcup\{E(x_i): i \in \omega\}) = \lim_{i\to\infty} \bar{\mu}(E(x_i)). \quad \text{Then}$$

$\bar{\mu}(E(a) - \bigcup\{E(x_i): i \in \omega\}) = 0$, and, thus, $\lambda(a) = \Pi(F(a)) = \Pi(f)$. Therefore, $\lim_{i\to\infty} \lambda(x_i) = \lambda(a)$ and (5) is proved.

(6) If $a \neq 0$. $a \in A$, then $\lambda(a) > 0$.

Proof of (6). $\lambda(a) = \Pi(F(a))$ and $F(a)$ is a continuous nonnegative function, which is positive somewhere. Then, there is an open-closed set C such that $F(a) \geqq \epsilon > 0$ on C for some $\epsilon > 0$. But, as μ is strictly positive, $\bar{\mu}(C) > 0$. Thus, $\lambda(a) = \Pi(F(a)) > 0$.

Since λ satisfies (3), (4), (5) and (6), we apply 16.11T and obtain a strictly positive measure on $\mathfrak{B}/I$. Transfering the measure to $\mathfrak{B}$ we obtain the desired properties.

5. Invariant measures on fields of sets. In this section, we apply Theorem 4.2 to measures on σ-fields of sets. Thus, let $\mathfrak{F}$ be a σ-field of sets and G a group of transformations of $\mathfrak{F}$. If μ is a G-invariant measure, μ has to vanish on all $\approx_G$-negligible sets. We call these sets *G-negligible* i.e. $A \in F$ is G-negligible if there is a sequence of disjoint elements $Y \in {}^{\omega}F$, such that $A \simeq_G Y_i$ for every $i < \omega$. N_G is the set of all G-negligible sets. From 4.2 we obtain immediately:

THEOREM 5.1. Let $\mathfrak{F}$ be a σ-field of subsets of a set X, G a group of transformations of $\mathfrak{F}$, and I a subset of F. The following conditions are necessary and jointly sufficient for the existence of a countably additive, G-invariant, and I-positive measure on $\mathfrak{F}$:

(i) I is a proper σ-ideal in $\mathfrak{F}$.

(ii) If $A \in I$ and $B \sim_G A$, then $B \in I$.

(iii) $N_G \subseteq I$.

(iv) $\mathfrak{F}/I$ has the Kelley property and is weakly σ-distributive.

It is easy to generalize 5.1 to the case when the measure μ is required to be equal to one, not on X, but on another set $C \in F$. Instead of considering G-negligible sets, we have to consider G-negligible sets relative to C (i.e. $A \in F$ is G-negligible relative to C if there is a sequence of disjoint elements $Y \in {}^{\omega}F$, such that $Y_i \subseteq C$ and $B \simeq_G Y_i$ for every $i < \omega$). Also, C should not belong to I. The conjecture that the only necessary and sufficient condition for the existence of a G-invariant measure is that X is not G-negligible, was proposed in **[2]**, at least for the case when F is the field of all subsets of X. However in **[3]**, the following counterexample was indicated:

Let $X = \omega_1$ and G be the group of all permutations f of X, such that $f(x) \neq x$ for at most denumerable x in X. It is easy to see that G is a group of transformations on the field of all subsets of X. The ideal of G-negligible sets contains all sets that are at most denumerable. Thus, X is not G-negligible. However, the existence of a G-invariant measure on this field would imply that ω_1 is a measurable cardinal.

When we want measures on BA's we are mainly interested in strictly positive measures. For G-invariant measures μ on a σ-field of sets, it is hardly ever possible to obtain strictly positive measures, since μ must vanish on the G-negligible sets and, by a result of (**[1]**, p. 194), nonempty G-negligible sets exist in most cases of interest. In particular, these sets exist when for every $n < \omega$, there are disjoint $Y_0, \cdots, Y_{n-1} \in F$ such that $X = \Sigma_{i<n} Y_i$ and $Y_i \simeq_G Y_j$ for every $i, j < n$. What we can get are measures that only vanish at G-negligible sets. We call these measures *G-strictly positive* (i.e. μ is G-strictly positive iff μ is N_G-positive). Using 4.1 we get as a particular case of 5.1:

THEOREM 5.2. *Let $\mathfrak{F}$ be a σ-field of subsets of a set X and G a group of transformations of $\mathfrak{F}$. The following conditions are necessary and jointly sufficient for the existence of countably additive, G-strictly positive, and G-invariant measure on $\mathfrak{F}$:*

(i) $X \notin N_G$;

(ii) *$\mathfrak{F}/N_G$ has the Kelley property and is weakly σ-distributive.*

REFERENCES

1. R. Bradford, *Cardinal addition and the axiom of choice*, Ann. Math. Logic, **3**, (1971), 111–196.

2. R. Chuaqui, *Cardinal algebras and measures invariant under equivalence relations*, Trans. Amer. Math. Soc., **142** (1969), 61–79.

3. ———, *The existence of an invariant measure and paradoxical decompositions*, Notices Amer. Math. Soc., **20** (1973), A-636, Abstract #73T-B313.

4. ———, *A semantical definition of probability*, in Non-Classical Logics, Model Theory and Computability, Arruda, da Costa and Chuaqui (eds), North-Holland Pub. Co. (1977), 135–167.

5. P. C. Deliyannis, *Group representations and cardinal algebras*, Canad. J. Math., **22** (1970), 759–772.

6. P. A. Fillmore, *The dimension theory of certain cardinal algebras*, Trans. Amer. Math. Soc. **117** (1965), 21–36.

7. E. Hopf, *Theory of measures and invariant integrals*, Trans. Amer. Math. Soc., **34** (1932), 373–393.

8. A. Horn and A. Tarski, *Measures in Boolean algebras*, Trans. Amer. Math. Soc., **64** (1948), 467–497.

9. O. G. Jorsboe, *Set transformations and invariant measures*, Various Publications Series No. 3, Matematisk Institut, Aarhus Universitet. (1966).

10. J. L. Kelley, *Measures on Boolean algebras*, Pacific J. Math., **9** (1959), 1165–1178.

11. D. Maharam, *On measures in abstract sets*, Trans. Amer. Math. Soc., **51** (1942), 413–433.

12. ———, *An algebraic characterization of measure algebras*, Amer. Math., **48** (1947), 154–167.

13. R. Sikorski, *Boolean algebras*, 3rd. ed. Springer-Verlag, Berlin–Heidelberg (1969).

14. A. Tarski, *Cardinal algebras*, Oxford U. Press, New York, (1949).

Received August 16, 1976.

UNIVERSIDADE ESTADUAL DE CAMPINAS (BRASIL)
AND
UNIVERSIDAD CATOLICA DE CHILE
VICIMA MACKENNA 4860
SANTIAGO 10, CHILE

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No 2, 1977

GAUGE GROUPS AND CLASSIFICATION OF BUNDLES WITH SIMPLE STRUCTURAL GROUP

W. D. CURTIS AND F. R. MILLER

Suppose π_i $i = 1,2$ are principal K-bundles which are C^r-isomorphic in the sense that there exists a K-equivariant C^r-diffeomorphism $f: \mathcal{P}_1 \to \mathcal{P}_2$. If h belongs to the gauge group H_2 of $\mathcal{P}_2$ then $h \circ f$ lies in H_1 and we have a group isomorphism $H_2 \to H_1$ which is C^∞. It is the purpose of this paper to investigate the converse in the case where K is a simple Lie group. (If K is abelian the gauge group of every K bundle over X is $C^r(X, K)$ so there is no hope of a converse. However for simple groups the situation is much better).

0. Introduction. Let K be a compact connected Lie group with Lie algebra $\mathcal{K}$. Let $\pi: \mathcal{P} \to X$ be a principal K-bundle of class C^∞ where X is a compact, connected C^∞-manifold.

Throughout this paper r will be a positive integer which is chosen at this time and remains unchanged from here on.

We denote by H the subgroup of $C^r(\mathcal{P}, K)$ consisting of all those h for which $h(pk) = k^{-1}h(p)k$ for all p in $\mathcal{P}$ and $k \in K$. H is naturally isomorphic to the group of all C^r-bundle automorphisms of $\mathcal{P}$ which cover the identity on X **[1, 2]**. The group H will be called the gauge group of π the terminology being motivated by current usage in theoretical physics. $C^r(\mathcal{P}, K)$ is a Banach Lie group and H is a sub-manifold and so H is a Banach Lie group **[2]**. The Lie algebra of H can be identified as $\mathcal{H} = \{h: \mathcal{P} \to \mathcal{K} \mid h \text{ is } C^r \text{ and } h(pk) = Ad(k^{-1})h(p)$ for $p \in \mathcal{P}$, $k \in K\}$.

The bracket in $\mathcal{H}$ and the exponential map exp: $\mathcal{H} \to H$ are the natural pointwise operations.

1. Ideals in $\mathcal{H}$. Suppose $\mathcal{J} \subset \mathcal{H}$ is an ideal. For $p \in \mathcal{P}$ $e_p: \mathcal{H} \to \mathcal{K}$ is defined by $e_p(h) = h(p)$ for $h \in \mathcal{H}$. e_p is a Lie algebra epimorphism so $e_p(\mathcal{J})$ is an ideal in $\mathcal{K}$.

LEMMA 1.1. *If $p \in \mathcal{P}$ and $k \in K$ then $e_p(\mathcal{J}) = e_{pk}(\mathcal{J})$.*

Proof. $e_{pk}(h) = h(pk) = Ad(k^{-1})h(p) = Ad(k^{-1})e_p(h)$. Thus $e_{pk}(\mathcal{J}) = Ad(k^{-1})e_p(\mathcal{J})$. But $e_p(\mathcal{J})$ is an ideal in $\mathcal{K}$ so $Ad(k^{-1})e_p(\mathcal{J}) = e_p(\mathcal{J})$.

DEFINITION 1.2. If $x \in X$ let $\mathcal{K}_x = e_p(\mathcal{J})$ where $p \in \pi^{-1}(x)$.

DEFINITION 1.3. If $\mathcal{I}$ is an ideal in $\mathcal{H}$ we say $\mathcal{I}$ has property s if $[\mathcal{I}, \mathcal{H}] = \mathcal{I}$.

We recall that $[\mathcal{I}, \mathcal{H}]$ is the Lie subalgebra of $\mathcal{H}$ generated by all elements of the form $[a, b]$ where $a \in \mathcal{I}$, $b \in \mathcal{H}$. $[\mathcal{I}, \mathcal{H}]$ consists exactly of all finite sums $\Sigma_i [a_i, b_i]$, $a_i \in \mathcal{I}$, $b_i \in \mathcal{H}$.

We denote by $\mathcal{F}(X)$ the algebra of C^r, real valued functions on X. $\mathcal{H}$ is a module over $\mathcal{F}(X)$ for if $f \in \mathcal{F}(X)$ and $h \in \mathcal{H}$ define $fh: \mathcal{P} \to \mathcal{K}$ by $(fh)(p) = f(\pi(p))h(p)$. One easily sees fh lies in $\mathcal{H}$ so we have a module.

LEMMA 1.4. *If the ideal $\mathcal{I} \subset \mathcal{H}$ has property s then $\mathcal{I}$ is a $\mathcal{F}(X)$-submodule of $\mathcal{H}$.*

Proof. Let $h \in \mathcal{I}$, $\phi \in \mathcal{F}(X)$. We show $\phi h \in \mathcal{I}$. $\mathcal{I}$ has property s so we may write $h = \Sigma_i [h_i, f_i]$ where $h_i \in \mathcal{I}$ and $f_i \in \mathcal{H}$. Then $\phi h = \Sigma_i \phi[h_i, f_i] = \Sigma_i [h_i, \phi f_i] \in \mathcal{I}$ where we used the pointwise nature of the bracket to get the last equation.

LEMMA 1.5. *If $\mathcal{H}_1$ and $\mathcal{H}_2$ correspond to bundles π_1 and π_2 and $\psi: \mathcal{H}_1 \to \mathcal{H}_2$ is a Lie algebra isomorphism then if $\mathcal{I}$ has property s in $\mathcal{H}_1$ then $\psi(\mathcal{I})$ has property s in $\mathcal{H}_2$.*

Before proving the final lemma of this section we make a preliminary construction. Suppose U is open in X and ξ is a section of π over U. Suppose $h \in \mathcal{H}$ and h has support in $\pi^{-1}(U)$. Define $\bar{h}: X \to \mathcal{K}$ by,

$$\bar{h}(x) = \begin{cases} h(\xi(x)) & x \in U \\ 0 & x \notin U. \end{cases}$$

$\bar{h} \in C^r(X, \mathcal{K})$ has support in U. Conversely if we start with $\bar{h}: X \to \mathcal{K}$ having support in U we can define $h \in \mathcal{H}$ as follows. There is a unique C^∞-map $\theta: \pi^{-1}(U) \to \mathcal{K}$ such that $\xi(\pi(p))\theta(p) = p$ for $p \in \pi^{-1}(U)$. We define

$$h(p) = \begin{cases} Ad(\theta(p)^{-1})\bar{h}(\pi(p)) & p \in \pi^{-1}(U) \\ 0 & p \notin \pi^{-1}(U). \end{cases}$$

It is easily checked that $h \in \mathcal{H}$.

If $x_0 \in X$ we have:

$$H_{x_0} = \{f \in H \mid f(p) = e \quad \text{for all} \quad p \in \pi^{-1}(x_0)\}.$$
$$\mathcal{H}_{x_0} = \{h \in \mathcal{H} \mid h(p) = 0 \quad \text{for all} \quad p \in \pi^{-1}(x_0)\}.$$

LEMMA 1.6. *Assume $\mathcal{K}$ is semisimple. Then $\mathcal{H}_{x_0}$ has property s.*

Proof. Let $(\phi_i)_i$ be a finite partition of unity on X subordinate to an open cover $(U_i)_i$ such that π is trivial over each U_i. Then if $h \in \mathcal{H}_{x_0}$ we have $h = \Sigma_i \phi_i h$ and each $\phi_i h \in \mathcal{H}_{x_0}$. Therefore the problem is reduced to proving the following: If $U \subset X$ is open such that π has a local section ξ defined on U and if $h \in \mathcal{H}_{x_0}$ has support in $\pi^{-1}(U)$ then h can be written as $h = \Sigma_\nu [g_\nu, \phi_\nu]$ where $g_\nu \in \mathcal{H}_{x_0}$, $\phi_\nu \in \mathcal{H}$.

Let $\bar{h}: X \to \mathcal{K}$ correspond to h using the section ξ as above. Let $(E_i)_i$ be a basis for $\mathcal{K}$. Write $\bar{h} = \Sigma_i \bar{h}^i E_i$ where $\bar{h}^i$ are real valued. Since $\mathcal{K}$ is semisimple we may write $E_i = \Sigma_j [F_{ij}, G_{ij}]$ where F_{ij}, G_{ij} are in $\mathcal{K}$. Therefore $h = \Sigma_{i,j} \bar{h}^i [F_{ij}, G_{ij}] = \Sigma_{i,j} [\bar{h}^i F_{ij}, G_{ij}] = \Sigma_\nu [\bar{g}_\nu, \bar{\phi}_\nu]$ where $\bar{g}_\nu$ and $\bar{\phi}_\nu: X \to \mathcal{K}$ are C^r with $\bar{g}_\nu(x_0) = 0$. We can easily arrange that $\bar{g}_\nu$ and $\bar{\phi}_\nu$ have support in U. Then let g_ν, ϕ_ν be the corresponding functions on $\mathcal{P}$. Then if $p \in \mathcal{P}$ with $\pi(p) = x$ we have,

$$h(p) = Ad(\theta(p)^{-1})\bar{h}(x) = Ad(\theta(p)^{-1})\Big(\sum_\nu [\bar{g}_\nu(x), \bar{\phi}_\nu(x)]\Big)$$

$$= \sum_\nu [Ad(\theta(p)^{-1})\bar{g}_\nu(x), Ad(\theta(p)^{-1})\bar{\phi}_\nu(x)]$$

$$= \sum_\nu [g_\nu(p), \phi_\nu(p)] = \Big(\sum_\nu [g_\nu, \phi_\nu]\Big)(p).$$

2. A classification theorem. In this section, in addition to the assumptions made in the introduction, we assume K is a simple Lie group with trivial center. We first make some observations.

Given a principal K-bundle $\pi: \mathcal{P} \to X$ we construct the associated fiber bundle $\mathcal{A} \to X$ with fiber $\mathcal{K}$ where K acts on $\mathcal{K}$ via the adjoint representation of K. Each $p \in \mathcal{P}$ with $\pi(p) = x$ gives a linear isomorphism $\phi_p: \mathcal{K} \to \mathcal{A}_x$. Since $Ad: K \to \mathrm{Lis}(\mathcal{K})$ actually takes values in $\mathrm{Aut}(\mathcal{K})$ we see $\mathcal{A}$ is a bundle of Lie algebras. Therefore $\Gamma^r(\mathcal{A})$, the space of C^r-sections of $\mathcal{A}$, is a Lie algebra with pointwise bracket. There is a natural isomorphism $\mathcal{H} \to \Gamma^r(\mathcal{A})$ given by $h \to \tilde{h}$ where $\tilde{h}(x) = \phi_p(h(p))$ for each $x \in X$ where $p \in \pi^{-1}(x)$ **[3]**. This isomorphism is an isomorphism of $\mathcal{F}(X)$-modules and is a homeomorphism with respect to the C^r-topologies.

Now suppose $\pi_i: \mathcal{P}_i \to X$ are principal K-bundles, $i = 1,2$, with gauge groups H_i and $\mathcal{H}_i$ the Lie algebra of H_i. For $x_0 \in X$ the ideal $\mathcal{H}_{ix_0}$

is closed. Let $\psi: H_1 \to H_2$ be a C^1-group isomorphism. There is an induced Lie algebra isomorphism $\psi_*: \mathcal{H}_1 \to \mathcal{H}_2$ given by

$$\psi_*(h)(p) = \frac{d}{dt}\bigg|_{t=0} [\psi(\exp(th))](p)$$

ψ_* is a topological isomorphism and so for each $x_0 \in X$ $\psi_*(\mathcal{H}_{1x_0})$ is a closed ideal having property s in $\mathcal{H}_2$. If we write $\mathcal{I} = \psi_*(\mathcal{H}_{1x_0})$ and refer to the discussion of section 1 we have ideals $\mathcal{K}_x \subset \mathcal{K}$ for each $x \in X$. There are apparently two possible cases.

Case 1. $\mathcal{K}_x = \mathcal{K}$ for all $x \in X$.

We argue this cannot occur. Since $\mathcal{I}$ is an ideal with property s $\mathcal{I}$ is an $\mathcal{F}(X)$-submodule. If $\mathcal{K}_x = \mathcal{K}$ for all x in X we shall show $\mathcal{I} = \mathcal{H}_2$ which is impossible since $\mathcal{H}_{1x_0} \neq \mathcal{H}_1$. To show $\mathcal{I} = \mathcal{H}_2$ we regard $\mathcal{I}$ as a closed $\mathcal{F}(X)$-submodule of $\Gamma^r(\mathcal{A}_2)$. Then for $x \in X$, $v \in \mathcal{A}_{2x}$ there is $h \in \mathcal{I}$ for which $h(x) = v$. One now uses the $\mathcal{F}(X)$-module structure to show for any $x \in X$ and for any r-jet $\xi \in j^r_x\mathcal{A}_2$ there is an $h \in \mathcal{I}$ for which $j^r_x h = \xi$. Since $\mathcal{I}$ is a closed submodule we conclude $\mathcal{I} = \Gamma^r(\mathcal{A}_2)$ by applying a "global" version of a well-known theorem of Whitney. We refer to **[5]**, Corollary 1.6, p. 25.

Case 2. $\mathcal{K}_x = \mathcal{K}$ for some x.

In this case there is some x_1 for which $\mathcal{K}_{x_1} = (0)$ since K is simple. We claim there cannot be an $x_2 \neq x_1$, for which $\mathcal{K}_{x_2} = 0$. For if there were then we would have $\mathcal{I} \subset \mathcal{H}_{2x_1} \cap \mathcal{H}_{2x_2}$. But the codimension of $\mathcal{I}$ in $\mathcal{H}_2$ equals the codimension of $\mathcal{H}_{1x_0}$ in $\mathcal{H}_1$ which equals the codimension of $\mathcal{H}_{2x_1}$ in $\mathcal{H}_2$ so $\mathcal{I} \subset \mathcal{H}_{2x_1} \cap \mathcal{H}_{2x_2}$ is not possible. Therefore in the present case we see there is a unique $x_1 \in X$ for which $\mathcal{I} = \mathcal{H}_{2x_1}$.

Thus we see that a C^1 isomorphism $\psi: H_1 \to H_2$ gives rise to a bijection $\bar{\psi}: X \to X$ defined by

$$\psi_*(\mathcal{H}_{1x}) = \mathcal{H}_{2\bar{\psi}(x)}.$$

Now let $h \in \mathcal{H}_1$, $f \in \mathcal{F}(X)$. We have $\bar{\psi}: X \to X$ and we write $\bar{\psi}_*(f) = f \circ \bar{\psi}^{-1}$.

LEMMA 2.1. $\psi_*(fh) = \bar{\psi}_*(f)\psi_*(h)$.

Proof. Let $p_2 \in \mathcal{P}_{2x}$ let $\lambda = \bar{\psi}_*(f)(x)$. Then

$$\begin{aligned}\psi_*(fh)(p_2) &= \psi_*(fh - \lambda h)(p_2) + \psi_*(\lambda h)(p_2)\\ &= \psi_*((f-\lambda)h)(p_2) + \lambda\psi_*(h)(p_2).\end{aligned}$$

Let $x' = \bar{\psi}^{-1}(x)$ and let $p_1 \in \mathscr{P}_{1x'}$. Then $(f - \lambda)h(p_1) = (f(x') - \lambda)h(p_1) = 0$ by choice of λ. Thus $(f - \lambda)h \in \mathscr{H}_{1x'}$ and so $\psi_*((f - \lambda)h) \in \mathscr{H}_{2x}$ so $\psi_*((f - \lambda)h)(p_2) = 0$. Thus

$$\psi_*(fh)(p_2) = \lambda\psi_*(h)(p_2) = (\bar{\psi}_*(f) \cdot \psi_*(h))(p_2)$$

as desired.

LEMMA 2.2. *The map $\bar{\psi}$: $X \to X$ is a C^r-diffeomorphism.*

Proof. We need only show $\bar{\psi}^{-1}$ is C^r. It is enough to show that if $f \in \mathscr{F}(X)$ then $f \circ \bar{\psi}^{-1}$ is C^r. Choose $x_0 \in X$, U a neighborhood of x_0 $\mathscr{P}_2$ trivial over U. Then let V be a neighborhood of x_0 with $\bar{V} \subset U$. Let k be a section of $\mathscr{A}_2$ over U which in the local trivialization has constant principal part. We can then cut k down to get a new section, again called k, defined on all of X and agreeing with the original k on V. Then choose $h \in \Gamma^r(\mathscr{A}_1)$ such that $\psi_*(h) = k$. (We are identifying $\mathscr{H}_i$ and $\Gamma(\mathscr{A}_i)$). Now by Lemma we have $\psi_*(fh) = (f \circ \bar{\psi}^{-1})\psi_*(h) = (f \circ \bar{\psi}^{-1})k$. When we view the C^r-section $(f \circ \bar{\psi}^{-1})k$ in our local trivialization we conclude $f \circ \bar{\psi}^{-1}$ is C^r on V. So we conclude $f \circ \bar{\psi}^{-1}$ is C^r and hence $\bar{\psi}^{-1}$ is C^r.

We now define a bundle isomorphism $\tilde{\psi}$ such that the following commutes:

$$\begin{array}{ccc} \mathscr{A}_1 & \xrightarrow{\tilde{\psi}} & \mathscr{A}_2 \\ \downarrow & & \downarrow \\ X & \xrightarrow{\bar{\psi}} & X \end{array}$$

Let $\alpha_x \in \mathscr{A}_{1x}$. Choose a section $h \in \Gamma^r(\mathscr{A}_1)$ such that $h(x) = \alpha_x$. Define $\tilde{\psi}(\alpha_x)$ by $\tilde{\psi}(\alpha_x) = \psi_*(h)(\bar{\psi}(x))$. This is independent of the choice of h for if h_1 were another section with $h_1(x) = \alpha_x$ then $h - h_1$ vanishes at x. Hence $\psi_*(h - h_1)$ vanishes at $\bar{\psi}(x)$ so $\psi_*(h)(\bar{\psi}(x)) = \psi_*(h_1)(\bar{\psi}(x))$. It is clear that the diagram commutes and that $\tilde{\psi}$ mapping $\mathscr{A}_{1x}$ to $\mathscr{A}_{2\bar{\psi}(x)}$ is a Lie algebra isomorphism.

LEMMA 2.3. *$\tilde{\psi}$ is C^r.*

Proof. We work locally trivializing $\mathscr{A}_1$. Let U be open in X, $V \subset U$ also open, γ: $U \times R^m \to \mathscr{A}_1 | U$ be a trivialization of $\mathscr{A}_1$ over U. Using this we see there are C^r-sections $h_1, \cdots, h_m \in \Gamma^r(\mathscr{A}_1)$ such that for each x in the subset V, $h_1(x), \cdots h_m(x)$ give a basis for the fiber over x which corresponds to the standard basis of R^m under γ. We claim

$\tilde{\psi} \circ \gamma : V \times R^m \to \mathscr{A}_2$ is given by

$$\tilde{\psi} \circ \gamma (x, \xi^1, \cdots, \xi^m) = \sum_{i=1}^{m} \xi^i \psi_*(h_i)(\bar{\psi}(x)).$$

If so then $\tilde{\psi}$ is C^r. But given $\xi^1, \cdots, \xi^m$ choose $f^i \in \mathscr{F}(X)$, $f^i(x) = \xi^i$. Then by Lemma 2.1 we see

$$\begin{aligned}\tilde{\psi}(\gamma(x, \xi^1, \cdots, \xi^m)) &= \tilde{\psi}\Big(\sum_{i=1}^{m} \xi^i h_i(x)\Big) = \tilde{\psi}\Big(\Big(\sum_{i=1}^{m} f^i h_i\Big)(x)\Big) \\ &= \psi_*\Big(\sum_{i=1}^{m} f^i h_i\Big)(\bar{\psi}(x)) \\ &= \sum_{i=1}^{m} \bar{\psi}_*(f^i)(\bar{\psi}(x))\psi_*(h_i)(\bar{\psi}(x)) \\ &= \sum_{i=1}^{m} \xi^i \psi_*(h_i)(\bar{\psi}(x)).\end{aligned}$$

Let $p \in \mathscr{P}_{1x}$. Then $\phi_p^1 : \mathscr{K} \to \mathscr{A}_{1x}$ is a Lie algebra isomorphism. If $q \in \mathscr{A}_{2\bar{\psi}(x)}$ then we have a Lie algebra isomorphism $\phi_q^2 : \mathscr{K} \to \mathscr{A}_{2\bar{\psi}(x)}$. (Note the superscripts tell which bundle is being used).

Now $(\phi_q^2)^{-1} \circ \tilde{\psi} \circ \phi_p^1 : \mathscr{K} \to \mathscr{K}$ lies in $\mathrm{Aut}(\mathscr{K})$. Let $\mathscr{E} = \{(p, q) \mid p \in \mathscr{P}_{1x}$ and $q \in \mathscr{P}_{2\bar{\psi}(x)}$ for some $x \in X\}$. $\mathscr{E}$ is the total space of the fiber product of $\mathscr{P}_1$ and $\bar{\psi}^* \mathscr{P}_2$. We have a map $\rho : \mathscr{E} \to \mathrm{Aut}(\mathscr{K})$, $\rho(p, q) = (\phi_q^2)^{-1} \circ \tilde{\psi} \circ \phi_p^1$. ρ is continuous and $\mathscr{E}$ is connected so ρ takes values in one of the connected components of $\mathrm{Aut}(\mathscr{K})$. Since K is a simple group the identity component of $\mathrm{Aut}(\mathscr{K})$ is $\mathrm{Aut}^\circ(\mathscr{K}) = Ad(K)$. Suppose $\sigma \in \mathrm{Aut}(\mathscr{K})$ and that $\rho(E) \subset \mathrm{Aut}^\circ(\mathscr{K})\sigma = Ad(K)\sigma$. Let $q \in \mathscr{P}_2$, $k \in K$. Then $\phi_{qk}^2 = \phi_q^2 \circ Ad(k)$. So $\rho(p, qk) = Ad(k^{-1}) \circ \rho(p, q)$. We conclude that for each $p \in \mathscr{P}_{1x}$ there is a unique $\mu(p)$ in $\mathscr{P}_{2\bar{\psi}(x)}$ for which $\rho(p, \mu(p)) = \sigma$. We then have a map $\mu : \mathscr{P}_1 \to \mathscr{P}_2$ covering $\bar{\psi}$. K acts freely on the right of both $\mathscr{P}_1$ and $\mathscr{P}_2$. We now show there is an automorphism $\bar{\sigma}$ of K, induced by σ, such that if a new action of K on $\mathscr{P}_2$ is defined by $q * k = q\bar{\sigma}(k)$, (the right side being the original action) then μ becomes K-equivariant. We have $\sigma \in \mathrm{Aut}(\mathscr{K})$. $\tau \to \sigma\tau\sigma^{-1}$ is an automorphism of $\mathrm{Aut}(\mathscr{K})$ and hence restricts to an automorphism of $\mathrm{Aut}^\circ(\mathscr{K}) = Ad(K)$. Using the isomorphism $Ad : K \to Ad(K)$ we see a unique automorphism $\bar{\sigma}$ is induced. $\bar{\sigma}$ satisfies the equation $Ad(\bar{\sigma}(k)) = \sigma Ad(k)\sigma^{-1}$. Now we show $\mu(pk) = \mu(p) * k$ for $p \in \mathscr{P}_1$, $k \in K$. We need only show $\rho(pk, \mu(p) * k) = \sigma$. But

$$\begin{aligned}\rho(pk, \mu(p) * k) &= \rho(pk, \mu(p)\bar{\sigma}(k)) = Ad(\bar{\sigma}(k))^{-1} \circ \rho(p, \mu(p)) \circ Ad(k) \\ &= Ad(\bar{\sigma}(k))^{-1} \circ \sigma \circ Ad(k) = \sigma Ad(k)^{-1}\sigma^{-1}\sigma Ad(k) = \sigma\end{aligned}$$

so we are done.

DEFINITION 2.4. Let $\pi: \mathcal{P} \to X$ be a principal K-bundle, τ an automorphism of K. The principal K-bundle $\pi^\tau: \mathcal{P}^\tau \to X$ is defined by introducing the new action $*: \mathcal{P} \times K \to P$, $p * k = p\tau(k)$. We say π^τ is conjugate to π by τ.

Considering the previous discussion we have now proved

THEOREM 2.5. *Under the assumptions made above if $\psi: H_1 \to H_2$ is a C^1 isomorphism then there is a C^r-diffeomorphism $\bar{\psi}: X \to X$ and an automorphism $\bar{\sigma}$ of K such that $\pi_1 \cong \bar{\psi}^*(\pi_2^{\bar{\sigma}})$.*

REMARK. Of course if $\bar{\sigma}$ is an inner automorphism we get $\pi_2^{\bar{\sigma}} \cong \pi_2$ and $\bar{\sigma}$ can be dropped.

3. Classical groups. We apply the results of §2 to the groups $SO(2n+1)$ $n \geqq 1$, $U(n)$ $n \geqq 2$, and $SO(2n)$ $n \geqq 3$. Since the center of $SO(2n+1)$ is trivial and the automorphism group of its Lie algebra is connected [**6**, pages 285–6] we get

THEOREM 3.1. *Let $\pi_i: \mathcal{P}_i \to X$ be principal $SO(2n+1)$ bundles with gauge groups H_i, $i = 1, 2$. Suppose $\psi: H_1 \to H_2$ is a C^1 (local) isomorphism. Then there is a C^r-diffeomorphism $\bar{\psi}: X \to X$ so that $\pi_1 \cong \bar{\psi}^*(\pi_2)$.*

Now let K be $SO(2n)$ $n \geqq 3$ or $U(n)$ $n \geqq 2$, $\pi_i: \mathcal{P}_i \to X$ be principal K bundles with gauge groups H_i and $\psi: H_1 \to H_2$ a C^r local isomorphism. Let Z denote the center of K. Now $\hat{\mathcal{P}}_i = \mathcal{P}_i/Z$ is a principal K/Z bundle over X. Let $\hat{H}_i$ be the gauge group of $\hat{\mathcal{P}}_i$. In both cases ($SO(2n)$ and $U(n)$) one can show that the Lie algebra isomorphism $\psi_*: \mathcal{H}_1 \to \mathcal{H}_2$ gives Lie algebra isomorphism $\hat{\psi}_*: \hat{\mathcal{H}}_1 \to \hat{\mathcal{H}}_2$ and also that the center of K/Z is trivial. Thus the results of §2 give a C^r diffeomorphism $\phi: X \to X$ and an automorphism σ of K/Z so that $\hat{\pi}_1 \cong \phi^*(\hat{\pi}_2^\sigma)$. Note that if σ is an inner automorphism $\hat{\pi}_2^\sigma \cong \hat{\pi}_2$ so that σ can be dropped. The form of σ not inner is given in [**6**, page 287]. It can be seen that σ lifts to $\sigma: K \to K$ and that $(\mathcal{P}_i/Z)^\sigma = \mathcal{P}_i^\sigma/Z$. We thus get

THEOREM 3.2. *Let K be $SO(2n)$ $n \geqq 3$ or $U(n)$ $n \geqq 2$, $\pi_i: \mathcal{P}_i \to X$ be principal K bundles with gauge groups H_i, $i = 1, 2$. Suppose $\psi: H_1 \to H_2$ is a (local) C^r isomorphism. Then there is a C^r diffeomorphism $\bar{\psi}: X \to X$ and automorphism $\sigma: K \to K$, so that $\mathcal{P}_1/Z \cong \bar{\psi}^*(\mathcal{P}_2 | Z)^\sigma \cong \bar{\psi}^*(\mathcal{P}_2^\sigma)/Z$ where Z is the center of K.*

One can show that $\mathcal{P}_1$ is a "tensor product" of $\bar{\psi}^*(\mathcal{P}_2^\sigma)$ with a principal Z-bundle over X. One way to see this is to use the classification for bundles as given in **[4]**. We state the result in terms of associated vector bundles.

THEOREM 3.3. *Let π_i: $\mathcal{P}_i \to X$ be principal* SO$(2n)$ $n \geqq 3$ $(U(n)$ $n \geqq 2)$ *bundles with gauge groups H_i, $i = 1, 2$. Let ξ_i be the real* (complex) *vector bundle associated with $\mathcal{P}_i$ using the usual representation of* SO$(2n)(U(n))$. *Suppose ψ: $H_1 \to H_2$ is a* (local) *C^1-isomorphism then there is a C^r diffeomorphism $\bar{\psi}$: $X \to X$, σ an automorphism of* SO$(2n)(U(n))$, *and η a real* (complex) *line bundle so that ξ_1 is* SO$(2n)(U(n))$ *isomorphic to $\psi^*(\xi_2^\sigma) \otimes \eta$.*

Final remark. We need not have assumed that $\mathcal{P}_1$ and $\mathcal{P}_2$ were bundles over the same manifold X. We could have considered π_1: $\mathcal{P}_1 \to X$ and π_2: $\mathcal{P}_2 \to Y$. If the gauge groups H_1 and H_2 are (locally) C^1 isomorphic we get a C^r-diffeomorphism $\bar{\psi}$: $X \to Y$.

REFERENCES

1. W. D. Curtis, *The automorphism group of a compact group action*, Trans. Amer. Math. Soc., **203** (1975), 45–54.
2. W. D. Curtis, Y-L. Lee and F. R. Miller, *A class of infinite dimensional subgroups of Diff$^r(x)$ which are Banach Lie groups*, Pacific J. Math., **47** (1973), 59–65.
3. Kobayashi and Nomizu, *Foundations of Differential Geometry Vol. I*, Interscience, New York 1963.
4. R. Lashof, *Classification of Fibre bundles by the loop space of the base*, Annals of Math., **64** (1965), 436–446.
5. B. Malgrange, *Ideals of Differentiable Functions*, Oxford Univ. Press 1966.
6. J. A. Wolf, *Spaces of Constant Curvature*, Third Edition, Publish or Perish Inc., Boston Mass., 1974.

Received June 8, 1976.

KANSAS STATE UNIVERSITY
MANHATTAN, KS 66506

PACIFIC JOURNAL OF MATHEMATICS
Vol 68 No 2, 1977

THE ESSENTIAL UNIQUENESS OF BOUNDED NONOSCILLATORY SOLUTIONS OF CERTAIN EVEN ORDER DIFFERENTIAL EQUATIONS

G. J. ETGEN AND W. E. TAYLOR, JR.

Let n be a positive integer, let p be a positive continuous function on $[0,\infty)$, and consider the $2n$th order linear differential equation

$$u^{(2n)} - p(x)u = 0. \tag{1}$$

It is well known that this equation has a solution $w = w(x)$ satisfying

$$(-1)^k w^{(k)}(x) > 0, \qquad k = 0, 1, \cdots, 2n-1, \tag{2}$$

on $[0,\infty)$, and it is clear that w is positive and bounded. The purpose of this paper is to investigate the essential uniqueness of the solution w, where the statement "w is essentially unique" means that if y is any other solution of (1) which satisfies (2), then $y = kw$ for some nonzero constant k.

In addition to having solutions which satisfy (2), it is easy to show that equation (1) has solutions $z = z(x)$ satisfying

$$z^{(k)}(x) > 0, \qquad k = 0, 1, \cdots, 2n-1, \tag{3}$$

on $[a,\infty)$ for some $a \geqq 0$. For some recent results concerning the behavior of solutions of (1) satisfying either (2) or (3), the reader is referred to the work of D. L. Lovelady **[6]**, and T. T. Read **[7]**.

A solution of (1) which satisfies (2) is said to be *strongly decreasing*, and a solution satisfying (3) is said to be *strongly increasing*. If y is a nontrivial solution of (1), then y is *oscillatory* if it has infinitely many zeros on $[0,\infty)$. Equivalently, y is oscillatory if the set of zeros of y is not bounded above. The differential equation (1) is oscillatory if it has at least one nontrivial oscillatory solution. Hereafter, the term "solution of (1)" shall be interpreted to mean "nontrivial solution." A solution of (1) which is not oscillatory is called *nonoscillatory*. Clearly, any solution satisfying either (2) or (3) is nonoscillatory. We shall say that equation (1) has *property* (H) if every nonoscillatory, eventually positive solution satisfies either (2) or (3).

S. Ahmad **[1]** has studied (1) in the case $n = 2$, and he has shown that (1) is oscillatory if and only if it has property (H). While this result is not known in general, Lovelady [**6**, Theorem 2] has shown that property (H) implies the oscillation of (1). Read [**7**] and G. W. Johnson [**4**] have obtained some results on the asymptotic properties of solutions of (1). In particular, they have obtained criteria which imply that any solution w satisfying (2) has the property $\lim_{x\to\infty} w(x) = 0$. Finally, we refer to the work of G. D. Jones and S. M. Rankin [**5**] where the problem of the essential uniqueness of a solution w satisfying (2) was considered for the case $n = 2$.

2. Preliminary results. Let $\mathscr{S}$ denote the $2n$-dimensional vector space of solutions of equation (1). Our first result is essential in the work which follows. Since the proof is straightforward, using well known techniques, it will be omitted.

LEMMA 2.1. *If $y \in \mathscr{S}$ and $y^{(k)}(a) \geqq 0$, $k = 0, 1, \cdots, 2n-1$, for some $a \geqq 0$, with at least one inequality being strict, then $y^{(k)}(x) > 0$, $k = 0, 1, \cdots, 2n-1$, on (a, ∞) and*

$$\lim_{x\to\infty} y^{(k)}(x) = \infty, \qquad k = 0, 1, \cdots, 2n-2.$$

If $z \in \mathscr{S}$ and $(-1)^k z^{(k)}(b) \geqq 0$, $k = 0, 1, \cdots, 2n-1$, for some $b > 0$, with at least one inequality being strict, then $(-1)^k z^{(k)}(x) > 0$ on $[0, b)$.

Let J be the function defined on $\mathscr{S} \times \mathscr{S}$ by

$$J(u, v)(x) = \sum_{k=0}^{2n-1} (-1)^k v^{(k)}(x) u^{(2n-k-1)}(x) \tag{4}$$

For any pair of functions $u, v \in \mathscr{S}$, it is easy to verify by differentiating $J(u, v)$ that $J'(u, v)(x) = 0$ for all $x \in [0, \infty)$. Thus $J(u, v) \equiv c$, a constant on $[0, \infty)$. The case where $J(u, v) \equiv 0$ shall be denoted by $u \perp v$. Fix $y \in \mathscr{S}$. Following the ideas introduced by J. M. Dolan in [**2**], we define the subset $\mathscr{S}(y)$ of $\mathscr{S}$ by

$$\mathscr{S}(y) = \{z \in \mathscr{S} \mid y \perp z\}$$

Let $u_1, u_2, \cdots, u_{2n-1}$ be $2n-1$ solutions of equation (1), and let $W(u_1, u_2, \cdots, u_{2n-1})$ denote their Wronskian. It is well known that W is a solution of (1), and that W is nontrivial if and only if the solutions are linearly independent. Let $y \in \mathscr{S}$ and let $T[y, u_1, u_2, \cdots, u_{2n-1}]$ denote the Wronskian of the $2n$ solutions. Then, by expanding T along its first column, we get the following relationship between T, W and the function J

$$T[y, u_1, u_2, \cdots, u_{2n-1}] = J[y, W(u_1, u_2, \cdots, u_{2n-1})] \tag{5}$$

THEOREM 2.2. *Let* $y \in \mathcal{S}$. *Then the following hold.*

(i) $\mathcal{S}(y)$ *is a* $(2n-1)$*-dimensional subspace of* $\mathcal{S}$ *and* $y \in \mathcal{S}(y)$.

(ii) *If* $z \in \mathcal{S}(y)$, *and* y *and* z *are linearly independent, then there exists a solution* $u \in \mathcal{S}(y)$ *such that* $J(u, z) \neq 0$.

(iii) *If* $\{u_1, u_2, \cdots, u_{2n-1}\}$ *is a basis for* $\mathcal{S}(y)$, *then* $W(u_1, u_2, \cdots, u_{2n-1}) = ky$ *for some nonzero constant* k.

(iv) *If* $v \in \mathcal{S}$, *then* $\mathcal{S}(y) \cap \mathcal{S}(v)$ *has dimension* $2n-1$ *if and only if* y *and* v *are linearly dependent*; *otherwise* $\mathcal{S}(y) \cap \mathcal{S}(v)$ *has dimension* $2n-2$.

Proof. Part (i) is easy to verify using (4) and the definition of $\mathcal{S}(y)$.

(ii) Let $z \in \mathcal{S}(y)$ be independent of y. Suppose z has a zero of multiplicity k, $1 \leqq k \leqq 2n-1$, at some point $c \geqq 0$. Since $\mathcal{S}(y)$ has dimension $2n-1$ we can construct a solution $u \in \mathcal{S}(y)$ such that

$$u(c) = u'(c) = \cdots = u^{(2n-k-2)}(c) = 0 = u^{(2n-k+1)}(c) = \cdots = u^{(2n-1)}(c) = 0,$$

$$u^{(2n-k-1)}(c) = 1, \quad u^{(2n-k)}(c) = \gamma,$$

where γ is some constant. Then, from (4), $J(u, z) = z^{(k)}(c) \neq 0$. If $z \neq 0$ on $[0, \infty)$, then choose a point c such that $y(c) \neq 0$, and choose $m \neq 0$ such that $y(c) - mz(c) = 0$. Let $v = y - mz$. Then $v \in \mathcal{S}(y)$ and $v \neq 0$ since y and z are independent. Now, we can repeat the argument above to determine a solution $u \in \mathcal{S}(y)$ such that $J(u, v) \neq 0$. Since $J(u, v) = J(u, y - mz) = -mJ(u, z)$, we conclude that $J(u, z) \neq 0$.

(iii) Let $\{u_1, u_2, \cdots, u_{2n-1}\}$ be a basis for $\mathcal{S}(y)$. Since $y \in \mathcal{S}(y)$, $y = \sum_{i=1}^{2n-1} c_i u_i$ and thus

$$0 = T[y, u_1, u_2, \cdots, u_{2n-1}] = J[y, W(u_1, u_2, \cdots, u_{2n-1})]$$

Hence the solution $W(u_1, u_2, \cdots, u_{2n-1})$ is an element of $\mathcal{S}(y)$. The same reasoning shows that

$$J[z, W(u_1, u_2, \cdots, u_{2n-1})] = 0$$

for all $z \in \mathcal{S}(y)$, and we can conclude, from (ii), that $W(u_1, u_2, \cdots, u_{2n-1}) = ky$.

Part (iv) is an immediate consequence of either (ii) or (iii). This completes the proof of the theorem.

We now consider the properties of the subspace $\mathcal{S}(w)$ in the case where w satisfies (2).

THEOREM 2.3. *Assume that equation* (1) *has property* (H), *and suppose* $w \in \mathcal{S}$ *satisfies* (2). *Then*:

(i) *If $y \in \mathscr{S}(w)$, then either y satisfies* (2), *or y is oscillatory*

(ii) *If $y \in \mathscr{S}(w)$ and $y^{(k)}(a) = 0$ for some $a \geqq 0$ and some nonnegative integer k, $0 \leqq k \leqq 2n-1$, then y is oscillatory*

(iii) *If $z \in \mathscr{S}$ and $z \notin \mathscr{S}(w)$, then z is unbounded.*

Proof. (i) Let $y \in \mathscr{S}(w)$ and assume that y is nonoscillatory with $y > 0$ on $[a, \infty)$, $a \geqq 0$. Suppose y does not satisfy (2). Then y satisfies (3) and there is a number $b \geqq a$ such that $y^{(k)}(x) > 0$, $k = 0, 1, \cdots, 2n-1$, on $[b, \infty)$. By evaluating $J(w, y)$ at any $x \geqq b$, we have that $J(w, y) \neq 0$, contradicting the fact that $y \in \mathscr{S}(w)$.

Part (ii) follows immediately from (i).

(iii) Let $z \in \mathscr{S}$ and suppose $z \notin \mathscr{S}(w)$. Fix any point $a \geqq 0$. Since $\mathscr{S}(w)$ has dimension $2n-1$ we can construct a basis for $\mathscr{S}(w)$ consisting of w and $2n-2$ solutions $u_1, u_2, \cdots, u_{2n-2}$ such that u_k has a zero of multiplicity k at $x = a$, $k = 1, 2, \cdots, 2n-2$. By (ii) every linear combination of the solutions $u_1, u_2, \cdots, u_{2n-2}$ is oscillatory. Let y be the solution of (1) determined by the initial conditions $y(a) = y'(a) = \cdots = y^{(2n-2)}(a) = 0$, $y^{(2n-1)}(a) = 1$. Then y satisfies (3) on $[b, \infty)$ for every $b \geqq a$. Thus $y \notin \mathscr{S}(w)$ and the set $\{y, w, u_1, u_2, \cdots, u_{2n-2}\}$ is a basis for $\mathscr{S}$. Now

$$z = cy + dw + \sum_{i=1}^{2n-2} c_i u_i,$$

where $c \neq 0$. Since w is bounded, and $\Sigma_{i=1}^{2n-2} c_i u_i$ is oscillatory, we can conclude that z is unbounded.

Our next result has appeared in [**5**, Lemma 4] for the case $n = 2$. The proof is straightforward and, consequently, it will be omitted.

LEMMA 2.4. *Let $\{u_1, u_2, \cdots, u_{2n}\}$ be a basis for $\mathscr{S}$. Then there exists a basis $\{z_1, z_2, \cdots, z_{2n}\}$ for $\mathscr{S}$ and $2n$ nonzero constants $k_1, k_2, \cdots, k_{2n}$, such that*

$$u_i \equiv k_i W(z_1, z_2, \cdots, z_{i-1}, z_{i+1}, \cdots, z_{2n}), \qquad i = 1, 2, \cdots, 2n.$$

3. Main results. It is easy to see that equation (1) has no oscillatory solutions when $n = 1$. Also, it is easy to show that the nonoscillatory solution w satisfying (2) is essentially unique in this case. Our first result shows that this situation holds in general.

THEOREM 3.1. *If equation* (1) *has no oscillatory solutions, then the nonoscillatory solution w satisfying* (2) *is essentially unique.*

Proof. Suppose that (1) has two linearly independent solutions w and v satisfying (2). Fix any $a \geqq 0$ and choose k such that

$w(a) - kv(a) = 0$. Let y be the solution given by $y(x) = w(x) - kv(x)$. Since y is nonoscillatory, we shall assume that $y > 0$, and that $\Pi_{k=1}^{2n-1} y^{(k)} \neq 0$ on $[b, \infty)$, $b > a$. Then $y^{(2n)} = py > 0$. Since each of w and v is bounded on $[0, \infty)$, y is bounded and we can conclude that no two consecutive derivatives $y^{(k)}$, $y^{(k+1)}$, $1 \leqq k \leqq 2n - 2$, can have the same sign on $[b, \infty)$. But this implies

$$sgn y = sgn y'' = \cdots = sgn y^{(2n)} \neq sgn y' = sgn y''' = \cdots = sgn y^{(2n-1)}$$

on $[b, \infty)$ and, with Lemma 2.1, contradicts the fact that $y(a) = 0$.

We now consider the case where equation (1) is oscillatory. The next result gives a connection between the essential uniqueness of the solution w satisfying (2) and the maximum number of linearly independent oscillatory solutions in $\mathcal{S}$.

THEOREM 3.2. *Assume that equation* (1) *has property* (H). *The following two statements are equivalent*:

(a) *The solution* w *of* (1) *satisfying* (2) *is essentially unique.*

(b) *Equation* (1) *has at most* $2n - 1$ *linearly independent oscillatory solutions.*

Proof. To show that (a) implies (b) we use a simple extension of the proof of the corresponding result for the case $n = 2$ in [**5**, Theorem 4]. In particular, assume that w is essentially unique, and suppose $\mathcal{S}$ has a basis consisiting of $2n$ oscillatory solutions $u_1, u_2, \cdots, u_{2n}$. Using Lemma 2.4, let $\{z_1, z_2, \cdots, z_{2n}\}$ be a basis for $\mathcal{S}$ such that for each i, $1 \leqq i \leqq 2n$,

$$W(z_1, \cdots, z_{i-1}, z_{i+1}, \cdots, z_{2n}) = k_i u_i.$$

Consider the solution $u_1 = k_1 W(z_2, z_3, \cdots, z_{2n})$. Since u_1 is oscillatory, there is an increasing sequence $\{x_i\}_{i=1}^{\infty}$ such that $\lim_{i\to\infty} x_i = \infty$ and $u_1(x_i) = 0$ for all i. Therefore, for each positive integer i there are $2n - 1$ constants $c_{2i}, c_{3i}, \cdots, c_{2n,i}$ such that $\sum_{j=2}^{2n} c_{ji}^2 = 1$ and the solution v_{1i},

$$v_{1i} = \sum_{j=2}^{2n} c_{ji} z_j,$$

has a zero of order $2n - 1$ at $x = x_i$. Because the sequences $\{c_{ji}\}$, $j = 2, 3, \cdots, 2n$, are bounded, we can assume, without loss of generality, that $\lim_{i\to\infty} c_{ji} = c_j$, $j = 2, 3, \cdots, 2n$, and $\sum_{j=2}^{2n} c_j^2 = 1$. By using an argument similar to the one used in [**1**, Theorem 1],

$$\lim_{i\to\infty} v_{1i} = v_1 = c_2 z_2 + c_3 z_3 + \cdots + c_{2n} z_{2n}$$

is a bounded nonoscillatory solution of (1) satisfying (2). Repeating this process $2n-1$ more times with the solutions $u_2, u_3, \cdots, u_{2n}$, we obtain the bounded nonoscillatory solutions

$$
\begin{aligned}
v_2 &= d_{21}z_1 + d_{23}z_3 + \cdots + d_{2,2n}z_{2n}, \quad \sum_{\substack{j=1\\ j\neq 2}}^{2n} d_{2j}^2 = 1,\\
v_3 &= d_{31}z_1 + d_{32}z_2 + d_{34}z_4 + \cdots + d_{3,2n}z_{2n}, \quad \sum_{\substack{j=1\\ j\neq 3}}^{2n} d_{3j}^2 = 1,\\
&\vdots\\
v_{2n} &= d_{2n,1}z_1 + d_{2n,2}z_2 + \cdots + d_{2n,2n-1}z_{2n-1}, \quad \sum_{j=1}^{2n-1} d_{2n,j}^2 = 1.
\end{aligned}
$$

The solution v_1 must be independent of at least one of the other v_i's, because, if not, then it is easy to show that $c_2 = c_3 = \cdots = c_{2n} = 0$ which contradicts $\Sigma_{j=2}^{2n} c_j^2 = 1$. Thus $\mathscr{S}$ cannot have more than $2n-1$ linearly independent oscillatory solutions.

Now assume that $\mathscr{S}$ contains at most $2n-1$ linearly independent oscillatory solutions. Let $w \in \mathscr{S}$ satisfy (2). As seen in the proof of Theorem 2.3 (iii), we can construct a solution basis for $\mathscr{S}(w)$ consisting of w and $2n-2$ oscillatory solutions $u_1, u_2, \cdots, u_{2n-2}$ such that u_k has a zero of multiplicity k at $x = a$, $k = 1, 2, \cdots, 2n-2$, $a \geqq 0$ fixed. Choose a point $b > a$ such that $u_1(b) \neq 0$ and let m be chosen such that $u_1(b) - mw(b) = 0$. Then $y = u_1 - mw \in \mathscr{S}(w)$, y is oscillatory, and $y, u_1, u_2, \cdots, u_{2n-2}$ are linearly independent. Suppose there exists a solution v satisfying (2) such that w and v are linearly independent. Then, from Theorem 2.2 (iv) $\mathscr{S}(w) \neq \mathscr{S}(v)$ and there exists a solution $z \in \mathscr{S}(v)$ such that $z \notin \mathscr{S}(w)$. Since $z \in \mathscr{S}(v)$ and v satisfies (2), z cannot satisfy (3). Since $z \notin \mathscr{S}(w)$, z must be unbounded. Therefore z is an unbounded oscillatory solution and it, together with the $2n-1$ independent oscillatory solutions in $\mathscr{S}(w)$ found above, constitute a solution basis for $\mathscr{S}$. This contradicts the hypothesis that $\mathscr{L}$ has at most $2n-1$ linearly independent oscillatory solutions, and completes the proof of the theorem.

COROLLARY 3.3. *Assume that equation* (1) *has property* (H). *If all the oscillatory solutions of* (1) *are bounded, then the solution w of* (1) *satisfying* (2) *is essentially unique.*

Proof. As seen in the proof of the theorem, if w is not essentially unique, then there exists an unbounded oscillatory solution $z \notin \mathscr{S}(w)$.

Our final result requires the concept introduced by Dolan and Klaasen in [3]. In particular, if $\mathscr{R}$ and $\mathscr{Q}$ are subsets of $\mathscr{S}$, then $\mathscr{R}$ is said to dominate $\mathcal{Q}$, denoted $\mathcal{R} > \mathcal{Q}$, if for each $y \in \mathcal{R}$ and $z \in Q$, $y + \lambda z \in \mathcal{R}$ for all real numbers λ.

Let $\mathcal{U}$ denote the unbounded nonoscillatory solutions of equation (1), $\mathcal{B}$ the set of bounded nonoscillatory solutions, and $\mathcal{O}$ the set of oscillatory solutions. When equation (1) has property (H), the sets $\mathcal{U}$ and $\mathcal{B}$ are easy to describe since $z \in \mathcal{U}$ implies either z or $-z$ is strongly increasing and $w \in \mathcal{B}$ implies either w or $-w$ is strongly decreasing.

THEOREM 3.4. *Assume that equation* (1) *has property* (H). *The following statements are equivalent*

(a) $\mathcal{U} > \mathcal{O}$

(b) $\mathcal{O} > \mathcal{B}$

(c) *The solution* w *of* (1) *satisfying* (2) *is essentially unique.*

Proof. Suppose (a) holds and suppose there is a number $k \neq 0$ such that $y + kw$ is nonoscillatory where $y \in \mathcal{O}$ and $w \in \mathcal{B}$, i.e., w satisfies (2). It is clear that the solution $v = y + kw$ does not satisfy (3), and so, by property (H), v satisfies (2). Obviously w and v are linearly independent. Fix any $a \geqq 0$. Let $u_1, u_2, \cdots, u_{2n-2}$ be the $2n-2$ linearly independent oscillatory solutions in $\mathcal{S}(w)$ such that u_k has a zero of multiplicity k, $k = 1, 2, \cdots, 2n-2$, at $x = a$. Let $z \in \mathcal{S}(v)$ such that $z \notin \mathcal{S}(w)$. We may assume that $z(a) = 0$ (which implies z oscillates), for if $z(a) \neq 0$, then choose $m \neq 0$ such that $z_1 = z - mw$ has a zero at a. Clearly $z_1 \in \mathcal{S}(v)$ and $z_1 \notin \mathcal{S}(w)$. Let y be the solution of (1) determined by the initial conditions $y(a) = y'(a) = \cdots = y^{(2n-2)}(a) = 0$, $y^{(2n-1)}(a) = 1$. From Lemma 2.1, $y \in \mathcal{U}$. The set $\{u_1, u_2, \cdots, u_{2n-2}, y\}$ forms a basis for the set of solutions of (1) having a zero at a. Therefore

$$z = \sum_{i=1}^{2n-2} c_i u_i + cy = u + cy.$$

Since $u(a) = 0$ and $u \in \mathcal{S}(w)$, u is oscillatory. Also, since $z \notin \mathcal{S}(w)$, $c \neq 0$. Thus $\bar{z} = (1/c)z = y + (1/c)u$ is oscillatory and contradicts the fact that $\mathcal{U} > \mathcal{O}$.

Suppose (b) holds and w is not essentially unique. Then there exists a solution v of (1) satisfying (2) which is independent of w. Let $u_1, u_2, \cdots, u_{2n-2}$ be the $2n-2$ linearly independent oscillatory solutions in $\mathcal{S}(w)$ such that u_k has a zero of multiplicity k, $k = 1, 2, \cdots, 2n-2$, at $x = a$, $a \geqq 0$ fixed. Then $\{w, u_1, u_2, \cdots, u_{2n-2}\}$ is a basis for $\mathcal{S}(w)$, and every linear combination of $u_1, u_2, \cdots, u_{2n-2}$ is oscillatory. Since v is bounded, we must have $v \in \mathcal{S}(w)$ by Theorem 2.3 (iii). Thus

$$v = \sum_{i=1}^{2n-2} c_i u_i + cw$$

where not all the c_i's are zero, that is, $v = u + cw$ is nonoscillatory where $u \in \mathcal{O}$ and $w \in \mathcal{B}$. This contradicts (b).

Finally, assume that (c) holds and suppose that $\mathcal{U}$ does not dominate $\mathcal{O}$. Then there exists $z \in \mathcal{U}$, $y \in \mathcal{O}$ and a nonzero number k such that $z + ky$ is oscillatory. It follows from Theorem 3.2 that $\mathcal{S}$ contains at most $2n - 1$ linearly independent oscillatory solutions. Since $\mathcal{S}(w)$ has a basis consisting of $2n - 1$ oscillatory solutions (see the proof of Theorem 3.2), we can conclude that both y and $z + ky$ are in $\mathcal{S}(w)$. But this implies $z \in \mathcal{S}(w)$ which is impossible since either z or $-z$ is strongly increasing. This completes the proof of the theorem.

References

1. S. Ahmad, *On the oscillation of solutions of a class of linear fourth order differential equations*, Pacific J. Math., **34** (1970), 289–299.

2. J. M. Dolan, *On the relationship between the oscillatory behavior of a linear third-order differential equation and its adjoint*, J. Differential Equations, **7** (1970), 367–388.

3. J. M. Dolan and G. A. Klaasen, *Dominance of n-th order linear equations*, Rocky Mountain J. Math. **5** (1975), 263–270.

4. G. W. Johnson, *A bounded nonoscillatory solution of an even order linear differential equation*, J. Differential Equations, **15** (1974), 172–177.

5. G. D. Jones and S. M. Rankin, *Oscillation properties of certain selfadjoint differential equations of the fourth order*, Pacific J. Math., (to appear).

6. D. L. Lovelady, *An asymptotic analysis of an even order differential equation*, submitted for publication.

7. T. T. Read, *Growth and decay of solutions of $y^{(2n)} - py = 0$*, Proc. Amer. Math. Soc., **43** (1974), 127–132.

Received September 17, 1976. The second author was partially supported by NSF grant No. HES 75-03730.

University of Houston
Houston, TX 77004

and

Prairie view A & M University
Houston, TX 77004

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No. 2, 1977

ON A REPRESENTATION THEORY FOR IDEAL SYSTEMS

PAUL EZUST

In widely divergent branches of mathematics, objects emerge which bear sufficient formal resemblance to the ideals of rings for them to be called "ideals". In a series of papers, Karl E. Aubert developed an axiomatic theory of ideal systems which subsumes most of the existing "ideal" theories. The goal of this paper is a representation theory for ideal systems in commutative monoids which will allow the formation of a cohomology theory for these systems. One of the results is a theorem which gives at once a monadic (co)homology for each ideal system. The base category in the monad includes PTOP, the category of pointed topological spaces and basepoint-preserving continuous maps, as a full subcategory and, for each ideal system, the category of algebras associated with the monad consists of the module systems over the ideal system. It is the module systems which are the principal objects of this study.

Described below are some of the basic notions of Aubert's theory of ideal systems. For simplicity in connection with our own work we assume that S is a commutative monoid (written multiplicatively) with an annihilating zero element (denoted 0).

DEFINITION. A *closure operation* x on a set W is a function which assigns to each subset $A \subseteq W$ a unique subset $A_x \subseteq W$ subject to the following conditions:

(i) $A \subseteq A_x$ for all $A \subseteq W$

(ii) $A \subseteq B_x \Rightarrow A_x \subseteq B_x$ for all $A, B \subseteq W$

NOTE. We do *not* assume that a closure operation x satisfies the (topological) condition: $(A \cup B)_x = A_x \cup B_x$. In general this condition will not be satisfied.

DEFINITION. A pair (S, x) is an *ideal system* if S is a commutative monoid with zero and x is a closure operation on S which satisfies the following axioms.

x.1 $\{0\}_x = \{0\}$

x.2 $AB_x \subseteq B_x$ for all $A, B \subseteq S$ ["multiplicative ideal property"]

x.3 $AB_x \subseteq (AB)_x$ for all $A, B \subseteq S$ ["continuity axiom"].

TERMINOLOGY. The sets $A_x \subseteq S$ are called the *x-ideals* of S.

NOTATION. $A + B = (A \cup B)_x$

$A : B = \{s \in S \mid sb \in A \ \forall b \in B\}$

$a \equiv b(A_x)$ iff $A_x + \{a\} = A_x + \{b\}$.

Several examples of particular ideal systems are discussed in Aubert's extensive survey paper **[2]** and the reader is referred to that paper for definitions, etc.

In a brief note **[8]**, Aubert and Hansen introduced the notion of "module system" over an ideal system as an ancillary device to the theory of ideal systems. Despite the pessimism expressed in that paper, it is our purpose to show that the theory of module systems over ideal systems yields a representation theory analogous to the theory of modules over rings.

Throughout this paper the terminology and notations of category theory have been used as are found in such standard texts as Herrlich and Strecker **[14]** and Mitchell **[22]**. The author originally became interested in the problems discussed herein during a course given by Professor Karl E. Aubert at Tufts University during the academic year, 1969–70.

2. Axioms for module-systems.

DEFINITION. Let (S, x) be a fixed ideal system. A left *S-set* is a set M together with a map $S \times M \to M$, denoted by $(s, m) \to sm$, satisfying

(i) $s(tm) = (st)m \ \forall s, t \in S, \forall m \in M$

(ii) $1m = m \ \forall m \in M$ (where 1 denotes the identity element of the monoid S).

DEFINITION. A pair (M, y), where M is a (left) S-set and y is a closure operation on M, is a *module-system* over (S, x) if the following are satisfied:

y.1 $\exists \theta \in M$ such that $0m = \theta \ \forall m \in M$, and $\{\theta\}_y = \{\theta\}$. We shall denote $\theta = 0$.

y.2 $AU_y \subseteq U_y \ \forall A \subseteq S, \forall U \subseteq M$

y.3 $AU_y \subseteq (AU)_y \ \forall A \subseteq S, \forall U \subseteq M$

y.4 $A_xU \subseteq (AU)_y \ \forall A \subseteq S, \forall U \subseteq M$.

NOTATION. Let (M, y) be a module-system, let $U, V \subseteq M$, $A \subseteq S$, $u, v, w \in M$, and $s \in S$. Then,

$$U + V = (U \cup V)_y$$

$$U : V = \{s \in S \mid sv \in U \ \forall v \in V\}$$

$$U: A = \{m \in M \mid am \in U \ \forall a \in A\}$$

$$\mathrm{Ann}(u) = 0: u \ (= \{0\}: \{u\}) \text{ in } S$$

$$\mathrm{Ann}(a) = 0: a \ (= \{0\}: \{a\}) \text{ in } M$$

$$w \equiv v(U_y) \text{ iff } U_y + \{w\} = U_y + \{v\}.$$

OBSERVATION. Frequent use shall be made of the following two equivalences which were established by Aubert and Hansen **[8]**.

1. Axiom y.3 is equivalent to the following statement:

$$(U_y: s)_y = U_y: s \ \ \forall U \subseteq M, \ \forall s \in S.$$

2. Axiom y.4 is equivalent to the following statement:

$$(U_y: v)_x = U_y: v \ \ \forall U \subseteq M, \ \forall v \in M.$$

DEFINITION. Let (S, x) be a fixed ideal system. The category *MS* consists of objects which are module systems (M, y) over (S, x) and morphisms $\xi: (M_1, y_1) \to (M_2, y_2)$ which are set functions that satisfy the following conditions:

(i) $\xi(su) = s\xi(u) \ \forall s \in S, \ \forall u \in M_1$

(ii) $\xi(U_{y_1}) \subseteq (\xi(U))_{y_2} \ \forall U \subseteq M_1$.

REMARK. Morphism condition (ii), above, is equivalent to:

$$(\xi^{-1}(V_{y_2}))_{y_1} = \xi^{-1}(V_{y_2}) \forall V \subseteq M_1.$$

EXAMPLES.

1. For any fixed ideal system (S, x), let $M = A_x$ for some $A \subseteq S$, and $y = x$. Thus, for $B \subseteq M$, $B_y = B_x$, and (M, y) is an object of *MS*.

2. Let S be the multiplicative semigroup of a commutative ring with identity, and let x be the classical ideal closure, $A_x = A_d = \langle A \rangle$ $\forall A \subseteq S$. Then any module M over the ring, with the classical submodule closure, $U_y = \langle U \rangle$, is an object of *MS*.

3. Let S be a commutative monoid with 0 and for each $A \subseteq S$, let $A_x = SA$ [this closure is called the *s-closure*]. For any *S-set*, M, and any $U \subseteq M$, define $U_y = SU$ [this closure will be referred to as the s-closure also]. Then (M, y) is an object of *MS*.

4. Let (S, x) be an ideal system and let M be an S-set. For any $U \subseteq M$, define $U_y = U \cup \{0\}$ [this closure will be referred to as the *discrete* closure on M]. Then (M, y) is an object of *MS*.

5. Let $S = \{1/n \mid n \in Z, \ n > 0\} \cup \{0\}$. For each $A \subseteq S$, define $A_x = \{s \in S \mid s \leqq \sup A\}$; i.e., $A_x = [0, \bar{a}]$, where $\bar{a} = \sup A$. Then

(S, x) is an example of an ideal system for which the inclusion x.3 is proper.

3. The morphisms of MS.

DEFINITION. An S-set M with 0 is called an *(S, x)-set* provided $(0: u)_x = 0: u$ for all $u \in M$. A map $\varphi: M_1 \to M_2$ from one S-set to another is called an S-map if it satisfies (i) above.

PROPOSITION 1. *Let M be an (S, x)-set and $\{f_j: M \to M_j \mid j \in J\}$ a family of S-maps, where $\{(M_j, y_j) \mid j \in J\}$ is a family of objects of S_I. Then there exists a closure operator y such that (M, y) is an object and $f_j: M \to M_j$ is a morphism for all $j \in J$. The coarsest such system y is said to be induced in M by the family $\{f_j: M \to M_j \mid j \in J\}$.*

Proof. Let M be an (S, x)-set and $F = \{f_j: M \to M_j \mid j \in J\}$ be a family of S-maps into objects (M_j, y_j), for $j \in J$. Let $Q = \{f_j^{-1}(U^j_{y_j}) \mid U^j \subseteq M_j,\ j \in J\} \cup \{0\}$. For any $V \subseteq M$ define $V_y = \cap \{W \in Q \mid V \subseteq W\}$.

DEFINITION. Let M be an (S, x)-set and $G = \{g_j: M_j \to M \mid j \in J\}$ be a family of S-maps from objects (M_j, y_j) to M. The finest closure system, y, on M (if one exists) such that (M, y) is an object of MS and such that each g_j is a morphism, will be called the closure system which is *coinduced* in M by the family G. Let $P = \{U \subseteq M \mid (g_j^{-1}(U))_{y_j} = g_j^{-1}(U)$ for all $j \in J\}$. G is called a *covering family* of S-maps into M if (1) for each $U \in P$, $\exists j \in J$ such that $g_j(g_j^{-1}(U)) = U$; and (2) $0 \in P$.

PROPOSITION 2. *Let M be an (S, x)-set and let $G = \{g_j: M_j \to M \mid j \in J\}$ be a covering family of S-maps from objects (M_j, y_j) to M. Then there exists a coinduced closure system y for M* (with respect to G).

Proof. Let M and G be as described above and let P be as defined above. Let $Q = \{U \in P \mid (U: m)_x = U: m \ \forall m \in M\}$ and, for each $V \subseteq M$, define $V_y = \cap \{U \in Q \mid V \subseteq U\}$.

DEFINITION. An equivalence relation $\sim$ on an object (M, y) is a *congruence* if $u \sim v \Rightarrow su \sim sv \, \forall s \in S$. Let $[v] = \{u \in M \mid u \sim v\}$. A congruence $\sim$ is *admissible* if $[0]_y = [0]$.

PROPOSITION 3. *Let (M, y) be an object of MS and $\sim$ an admissible congruence on M. Then $(M/\sim, \tilde{y})$ is an object of MS, where $M/\sim$ is* *the set of* $\sim$ *classes in* M *and* $\tilde{y}$ *is coinduced by the map* $\pi: M \to M/\sim$ *defined by* $\pi(u) = [u]\ \forall u \in M$.

Proof. By Proposition 2, one need only show that for any admissible congruence $\sim$ on an object (M, y), the set $M/\sim$ is an (S, x)-set and the map $\pi: M \to M/\sim$ is a covering S-map.

PROPOSITION 4. *Let* (M, y) *be an object and* $U_y \subseteq M$. *Then*

(a) U_y *determines an admissible congruence on* M *given by the rule*: $u \equiv v\ (U_y)$ *iff* $U_y + \{u\} = U_y + \{v\}$. *Denote the set of congruence classes "modulo* U_y*" by* M/U_y.

(b) *The inclusion map* $i: U_y \to M$ *induces a system* y' *on* U_y *given by the rule*: $V_{y'} = V_y \cap U_y = V_y\ \forall V \subseteq U_y$. *Thus,* U_y *is a subobject of* M. *The prime is generally omitted.*

PROPOSITION 5. *The Zero object,* $M = \{0\}$, *is both initial and terminal in* MS *and will be denoted, simply,* 0.

THEOREM 6. *Let* $\varphi: (M_1, y_1) \to (M_2, y_2)$ *be a morphism. Then*

(a) φ *is a monomorphism iff* φ *is injective.*

(b) φ *is an epimorphism iff* φ *is surjective.*

(c) *If* φ *is monic then* $(\varphi^{-1}(U))_{y_1} \subseteq \varphi^{-1}(U_{y_2})\ \forall U \subseteq M_2$.

Proof. (a) Suppose φ is a monomorphism such that $\varphi(u) = \varphi(v)$ for some $u, v \in M_1$. Define (M_3, y_3) by: $M_3 = S \vee S$, the disjoint union of two copies of S (labeled with u and v, respectively) with the zero elements identified, and $U_{y_3} = (U \cap S_u)_x \cup (U \cap S_v)_x\ \forall U \subseteq M_3$. In fact, this construction is a special case of the more general construction of the coproduct of S with itself, which is discussed in §4. Let $\psi_1: M_3 \to M_1$ be defined by the rule: $\psi_1(s_u) = su$ and $\psi_1(s_v) = sv\ \forall s \in S$. Define $\psi_2: M_3 \to M_1$ by the rule: $\psi_2(s_u) = sv$ and $\psi_2(s_v) = su\ \forall s \in S$. ψ_1 and ψ_2 are morphisms such that $\varphi\psi_1 = \varphi\psi_2$. Since φ is monic, it follows that $\psi_1 = \psi_2$; i.e., $u = v$.

(b) Suppose φ is an epimorphism. Then $\varphi(M_1)$ is an S-set.

Claim. $\varphi(M_1) = M_2$. Let $M_3 = M_2/\varphi(M_1)$ be the S-set of congruence classes in M_2 modulo the S-set $\varphi(M_1)$; i.e., for $u, v \in M_2$, $u \equiv v\ (\varphi(M_1))$ means $Su \cup \varphi(M_1) = Sv \cup \varphi(M_1)$. For any $U \subseteq M_3$, define $U_{y_3} = SU$. Let $\pi: M_2 \to M_3$ be the S-map $\pi(u) = [u]$ and let $M = \{(u, [u]) \mid u \in M_2\} \cup \{(u, [0]) \mid u \in M_2\}$. For each $s \in S$, $s(u, [u]) = (su, s[u]) = (su, [su])$ and $s(u, [0]) = (su, [0])$. Also, for each $u \in M_2$, $s(u, [0]) = (0, [0])$ iff $su = 0$ and $s(u, [u]) = (0, [0])$ iff $su = 0$, so that $(0, [0]): (u, [0]) = 0: u = (0: u)_x$. Hence, M is an (S, x)-set. Define $\xi_1: M_2 \to M$ by the rule: $\xi_1(u) = (u, [u])\ \forall u \in M_2$. Define $\xi_2: M_2 \to M$

by the rule: $\xi_2(u) = (u, [0])\ \forall u \in M_2$. Then $\{\xi_1, \xi_2\}$ is a covering family of S-maps into M. Let y be coinduced on M by $\{\xi_1, \xi_2\}$ and note that $\xi_1\varphi = \xi_2\varphi$; hence, $\xi_1 = \xi_2$. Thus, $\pi(u) = [0]\ \forall u \in M_2$; i.e., $\varphi(M_1) = M_2$.

(c) Suppose φ is a monomorphism. Then, by (a) above, φ is injective. Thus, $\varphi((\varphi^{-1}(U))_{y_1}) \subseteq U_{y_2}$; hence, $(\varphi^{-1}(U))_{y_1} \subseteq \varphi^{-1}(U_{y_2})$.

THEOREM 7. *MS has* (a) *Kernels,* (b) *Images,* (c) *Cokernels, and* (d) *Coimages.*

Proof. Let $\varphi: (M_1, y_1) \to (M_2, y_2)$ be a morphism. (a) $\operatorname{Ker}\varphi = \varphi^{-1}(0)$. (b) $\operatorname{Im}\varphi = (\varphi(M_1), y\varphi)$, where the closure operator $y\varphi$ is coinduced by the (surjective) map $\varphi': M_1 \to \varphi(M_1)$ defined by the rule: $\varphi'(u) = \varphi(u)\ \forall u \in M_1$. (c) Define the congruence $\sim$ by the rule: $u \sim u$ $\forall u \in M_2$ and, for $u \neq v$, $u \sim v$ iff $\{u, v\} \subseteq (\varphi(M_1))_{y_2}$. In forming $M_2/\sim$, the S-set of $\sim$ classes, $(\varphi(M_1))_{y_2}$ is compressed down to $[0]$ and the rest of M_2 remains unchanged. Let $\pi: M_2 \to M_2/\sim$ be the projection $u \to [u]$. Note that $[u] = [0]$ for $u \in (\varphi(M_1))_{y_2}$ and $[u] = u$ for $u \notin (\varphi(M_1))_{y_2}$. Also note that $M_2/\sim$ is an (S, x)-set and let $\tilde{y}$ be coinduced by $\{\pi\}$. Then $\operatorname{Coker}\varphi = (M_2/\sim, \tilde{y})$. (d) For each $u \in M_1$, let $\bar{u} = \varphi^{-1}(\varphi(u))$ and let $M_1/\varphi = \{\bar{u} \mid u \in M_1\}$. Let $\pi: M_1 \to M_1/\varphi$ be the projection, $u \to \bar{u}$. For each subset $\pi(U) \subseteq M_1/\varphi$, define $(\pi(U))_{\varphi y} = \hat{\varphi}^{-1}((\varphi(U))_{y_2})$, where $\hat{\varphi}: M_1/\varphi \to M_2$ is the map, $\hat{\varphi}(\bar{u}) = \varphi(u)$, for all $\bar{u} \in M_1/\varphi$. Then $\operatorname{Coim}\varphi = (M_1/\varphi, \varphi y)$.

REMARKS. (1) φ monic $\Rightarrow$ $\operatorname{Im}\varphi \cong M_1$.

(2) In any exact category (e.g., the category of modules over a commutative ring with unity), for any morphism $\varphi: M_1 \to M_2$, $\operatorname{Im}\varphi \cong \operatorname{Coim}\varphi$. The following example shows that this is not generally true in *MS*.

EXAMPLE. Let $M = \{0, a, b, c\}$, $S = \{0, 1\}$, with the obvious multiplication. Let (M_1, y_1) and (M_2, y_2) be defined as follows. $M_1 = M_2 = M$. y_1 is the s-system on M_1, and y_2 is the *indiscrete* system on M_2: $\{0\}_{y_2} = \{0\}$, and $U \neq \{0\} \Rightarrow U_{y_2} = M_2$. Let $\varphi: M_1 \to M_2$ be the identity map. Then $(M_2, y\varphi) = \operatorname{Im}\varphi \ncong \operatorname{Coim}\varphi = (M_1, \varphi y)$.

PROPOSITION 8. *Let* $\varphi: M_1 \to M_2$ *be a morphism. If* $\varphi(U_{y_1}) = (\varphi(U))_{y_2}$ *for all* $U \subseteq M_1$ *then* $\operatorname{Im}\varphi \cong \operatorname{Coim}\varphi$.

OBSERVATION. The example which precedes Proposition 8 also illustrates the fact that a morphism in *MS* might be both monic and epic and yet fail to be an isomorphism; i.e., *MS* is not balanced. Another way of characterizing this situation is to note that the forgetful functor $F: MS \to SET$ does not reflect isomorphisms. It follows (Proposition 32.5 **[14]**) that *MS* is not an algebraic category.

4. Categorical constructions in *MS*.

THEOREM 9. *MS has Products.*

Proof. Let $\{(M_j, y_j) \mid j \in J\}$ be a family of objects of *MS*. Let ΠM_j denote the cartesian product of the sets M_j ($j \in J$). For each $(m_j) \in \Pi M_j$ and each $s \in S$, define $s(m_j) = (sm_j)$. Let 0 denote (0_j) and observe that, for all $(m_j) \in \Pi M_j$, $0\colon (m_j) = \cap \{0\colon m_j \mid j \in J\}$, the latter being an intersection of x-ideals in S. Thus, ΠM_j, is an (S, x)-set. For each $k \in J$, define $\pi_k\colon \Pi M_j \to M_k$ by the rule, $\pi_k((m_j)) = m_k$ (this is the canonical projection map from the cartesian product to its factors). Let Πy_j be the system induced in ΠM_j by the family of projections, $\{\pi_j \mid j \in J\}$. Then, for each $U \subseteq \Pi M_j$, $U_{\Pi y_j} = \cap \{\pi_j^{-1}((\pi_j(U))_{y_j}) \mid j \in J\} = \times \{(\pi_j(U))_{y_j} \mid j \in J\}$. It is easy to verify that $(\Pi M_j, \Pi y_j)$ is the product.

NOTATION. $M_1 \times M_2$ will frequently be used to denote the product, $\Pi\{M_j \mid j = 1, 2\}$, of two objects of *MS*. The corresponding closure system will be denoted, $y_1 \times y_2$.

THEOREM 10. *MS has Coproducts.*

Proof. Let $\{(M_j, y_j) \mid j \in J\}$ be a family of objects of *MS* and let ΣM_j denote the disjoint union, $\vee \{M_j \mid j \in J\}$ with all zeros identified. For each $k \in J$, let $\delta_k\colon M_k \to \Sigma M_j$ be the natural inclusion map. Let Σy_j be defined on ΣM_j as follows. For any $U \subseteq \Sigma M_j$, $U_{\Sigma y_j} = \vee \{(\delta_j^{-1}(U))_{y_k} \mid k \in J\}$. Note that $U_{\Sigma y_j} = \cup \{(U \cap M_k)_{y_k} \mid k \in J\}$ if we identify M_k with its set-theoretic image, $\delta_k(M_k)$ in ΣM_j. Clearly $(\Sigma M_j, \Sigma y_j)$ is an object of *MS* and each map δ_k is a morphism. Note that Σy_j is the closure system coinduced in ΣM_j by the family of inclusions, $\{\delta_j \mid j \in J\}$. It is not hard to verify that $(\Sigma M_j, \Sigma y_j)$ is the coproduct.

DEFINITION. An object of *MS* is *free* if it is of the form ΣM_j ($j \in J$), where for each $j \in J$, $(M_j, y_j) \cong (S, x)$. We denote such an object $(F(J), y_*)$ and refer to the index set J as the *basis* for the free object $(F(J), y_*)$.

REMARK. In particular, (S, x) is free with basis $\{1\}$.

PROPOSITION 11. [Universal Mapping Property of Free Objects]. *$(F(U), y_*)$ is a free object with basis U iff for any object (M, y) and any set map $\sigma\colon U \to M$, there is a unique morphism $\varphi\colon F(U) \to M$ such that $\varphi\eta = \sigma$, where $\eta\colon U \to F(U)$ is the inclusion, $u \to 1_u$ for all $u \in U$.*

DEFINITION. The morphism φ described above is called the *lift* of σ.

PROPOSITION 12. *Let* (M, y) *be an object of* MS *and let* $\varphi: F(M) \to M$ *be the epimorphism that lifts the identity morphism* $1: M \to M$. *Then* $M \cong \operatorname{Coim}(\varphi)$.

DEFINITION. An object (M, y) is *projective* if for any morphism $\theta: M \to M_2$ and any epimorphism $\psi: M_1 \to M_2$ [where (M_1, y_1) and (M_2, y_2) are objects] there exists a morphism $\xi: M \to M_1$ such that $\psi\xi = \theta$.

$$\begin{tikzcd}
& M \arrow[dl, "\xi"'] \arrow[d, "\theta"] \\
M_1 \arrow[r, "\psi"'] & M_2
\end{tikzcd}$$

REMARK. It follows immediately from the above definition that if (M, y) is projective, then (M, y') is projective for any closure system y' (on M) which is finer than y. Thus, since the s-system is the finest possible closure system for M, each projective object in the category ENS-S of all S-sets determines a family of projective objects of MS and, conversely, each projective object of MS determines a projective object of ENS-S.

PROPOSITION 13. *Let* (M, y) *be an object of* MS. *Then* M *is projective iff* M *is a retract of a free object of* MS [In particular, each free object of MS is projective.]

REMARK. In the category R-Mod, of left R-modules, an object is a retract of a free iff it is a direct summand of a free. The following example demonstrates that this is not the case in general in MS.

EXAMPLE. Let $S = \{0, 1, a, b\}$ with multiplication defined as follows: $aa = bb = ab = ba = a$. Let $M = \{0, a\}$ and let S and M each have the s-system closure. Then (M, y) is a projective object of MS and M is not a direct summand of S since $(S - M)_y \neq (S - M) \cup \{0\}$. Since a free object of MS must be a coproduct of copies of S it follows that M is not a direct summand of any free object.

PROPOSITION 14. *Let* $\{(M_j, y_j) \mid j \in J\}$ *be a family of objects of* MS. *Then* $(\Sigma M_j, \Sigma y_j)$ *is projective iff* (M_j, y_j) *is projective* $\forall j \in J$.

REMARK. In view of Theorems 9 and 10, it is clear that MS is not an additive category since finite products are not isomorphic to finite coproducts.

5. Completeness and cocompleteness of *MS*.

PROPOSITION 15. *MS is locally and colocally small.*

PROPOSITION 16. *MS has Intersections.*

Proof. Let $\{\alpha_j : (M_j, y_j) \to (M, y) \mid j \in J\}$ be a family of subobjects of (M, y). Since α_j monic $\Rightarrow$ $M_j \cong \operatorname{Im} \alpha_j$, for each $j \in J$ we take $M' = \bigcap \{\operatorname{Im} \alpha_j \mid j \in J\}$, a set-theoretic intersection of subsets of M. For each $j \in J$, let $\beta_j : M' \to \operatorname{Im} \alpha_j$ be the natural inclusion map. Then M' is an (S, x)-set and β_j is an S-map for each $j \in J$. Let y' be the system induced on M' by the family $\{\beta_j \mid j \in J\}$. Let $\alpha : M' \to M$ be the natural inclusion map. Then $\alpha : (M', y') \to (M, y)$ is the intersection of the subobjects $\{\alpha_j \mid j \in J\}$.

PROPOSITION 17. *MS has Equalizers.*

Proof. Let $\varphi, \theta : M_1 \to M_2$ be morphisms, and let $E = \{u \in M_1 \mid \varphi(u) = \theta(u)\}$. Then $\operatorname{Equ}(\varphi, \theta) = (E, y_e)$, where y_e is induced by the inclusion $\eta : E \to M_1$.

The following Theorem follows from Theorem 23.8 **[14]**.

THEOREM 18. *MS has the following properties*:
(a) *MS is complete* (in particular, *MS* has inverse limits).
(b) *MS has* (multiple) *pullbacks.*
(c) *MS has inverse images.*

From Theorems 10 and 18 and Proposition 15 we obtain the hypotheses of Theorem 23.12 **[14]**, and using the dual of Theorem 23.8 **[14]** we obtain the following

THEOREM 19. *MS has the following properties*:
(a) *MS is cocomplete* (in particular, *MS* has direct limits).
(b) *MS has* (multiple) *pushouts.*
(c) *MS has direct images.*
(d) *MS has coequalizers.*
(e) *MS has cointersections.*

6. Properties of the hom functor *MS* $\to$ *MS*.

THEOREM 20. *For each pair of objects* (M_1, y_1), (M_2, y_2) *of MS*, $(\hom_S(M_1, M_2), \hat{y})$ *is an object of MS, where, for* $\varphi \in \hom_S(M_1, M_2)$ *and* $s \in S$, $s\varphi$ *is defined by the rule*: $(s\varphi)(u) = s(\varphi(u))\ \forall u \in M_1$, *and, for any*

$W \subseteq \hom_S(M_1, M_2)$, $W_{\hat{y}} = \cap\{[m, U_{y_2}] \mid W \subseteq [m, U_{y_2}]\}$, *where* $[m, U_{y_2}] = \{\xi \in \hom_S(M_1, M_2) \mid \xi(m) \in U_{y_2}\}$.

PROPOSITION 21. *For any object* (M, y) *of* MS, $(M, y) \cong (\hom_S(S, M), \hat{y})$, *where* (S, x) *is considered as an object of* MS.

THEOREM 22. *MS has an internal Hom functor*, Hom: $MS^{op} \times MS \to MS$.

Proof. By Theorem 20 it will suffice to verify that $\hom_S(\varphi, \theta)\colon \hom_S(M_1, M_2) \to \hom_S(M_1', M_2')$ is a morphism for all $\varphi \in \hom_S(M_1', M_1)$ and all $\theta \in \hom_S(M_2, M_2')$.

$$\begin{tikzcd} \hom_S(M_1, M_2) \arrow[d, "\hom_S(\varphi,\theta)"] \\ \hom_S(M_1', M_2') \end{tikzcd} \qquad \begin{tikzcd} M_1 \arrow[r, "f"] & M_2 \arrow[d, "\theta"] \\ M_1' \arrow[u, "\varphi"] \arrow[r, "\theta f \varphi"'] & M_2' \end{tikzcd}$$
$$= \hom_S(\varphi, \theta)(f)$$

Indeed, it is true in any category that the corresponding construction yields a well defined set map. Thus, with $\hom_S(\varphi, \theta)(f) = \theta f \varphi$, we have the following equations:

$$\begin{aligned} \hom_S(\varphi, \theta)^{-1}([u, U_{y_2'}]) &= \{f \in \hom_S(M_1, M_2) \mid \theta f \varphi(u) \in U_{y_2'}\} \\ &= \{f \in \hom_S(M_1, M_2) \mid f(\varphi(u)) \in \theta^{-1}(U_{y_2'})\} = [\varphi(u), \theta^{-1}(U_{y_2'})]. \end{aligned}$$

NOTATION. Since MS has an internal Hom functor, we will follow the practice of Herrlich and Strecker **[14]** and others and write it with a capital H. Also, we will suppress the subscript S when no confusion will result.

PROPOSITION 23. *For any family* $\{(M_j, y_j) \mid j \in J\}$ *of objects of MS*, $\mathrm{Hom}(\Sigma M_j, M) \cong \Pi \mathrm{Hom}(M_j, M)$ *for any object* (M, y).

PROPOSITION 24. *The functor* $\mathrm{Hom}(M, \_)\colon MS \to MS$ (for fixed object (M, y)) *preserves products*; *i.e., for any family* $\{(M_j, y_j) \mid j \in J\}$ *of objects*, $\mathrm{Hom}(M, \Pi M_j) \cong \Pi \mathrm{Hom}(M, M_j)$.

PROPOSITION 25. *The functor* $\mathrm{Hom}(M, \_)\colon MS \to MS$ *preserves equalizers.*

Proof. Let $f, g \in \mathrm{Hom}(M_1, M_2)$.

$$\begin{tikzcd} E \arrow[r, "\eta"] & M_1 \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & M_2 \\ & M \arrow[ul, "\alpha"] \arrow[u, "\delta"'] & \end{tikzcd}$$

Then, by Prop. I.17, $\eta: E \to M_1$ is the equalizer of f and g, where $E = \{u \in M_1 \mid f(u) = g(u)\}$ and y_e, the closure on E, is induced by the canonical inclusion, $\eta: E \to M_1$.

To prove that $\mathrm{Hom}(M, \_\,)$ preserves equalizers we shall show that $\mathrm{Hom}(M, E) \cong \mathrm{Equ}(\hat{f}, \hat{g})$, where $\hat{f} = \mathrm{Hom}(M, f)$ and $\hat{g} = \mathrm{Hom}(M, g)$, and $\mathrm{Equ}(\hat{f}, \hat{g}) = \{\xi \in \mathrm{Hom}(M, M_1) \mid \hat{f}(\xi) = \hat{g}(\xi)\}$.

$$\begin{tikzcd} \mathrm{Equ}(\hat{f}, \hat{g}) \arrow[r, "\lambda"] & \mathrm{Hom}(M, M_1) \arrow[r, shift left, "\hat{f}"] \arrow[r, shift right, "\hat{g}"'] & \mathrm{Hom}(M, M_2) \\ & \mathrm{Hom}(M, E) \arrow[ul, "\sigma"] \arrow[u, "\hat{\eta}"'] & \end{tikzcd}$$

Let $\lambda: \mathrm{Equ}(f, g) \to \mathrm{Hom}(M, M_1)$ be the canonical inclusion and $\hat{\eta} = \mathrm{Hom}(M, \eta)$: $\mathrm{Hom}(M, E) \to \mathrm{Hom}(M, M_1)$; i.e., $\hat{\eta}(k) = \eta k$. Then $\hat{f}\hat{\eta} = \hat{g}\hat{\eta}$, hence there exists a morphism, $\sigma: \mathrm{Hom}(M, E) \to \mathrm{Equ}(\hat{f}, \hat{g})$ such that $\hat{\eta} = \lambda\sigma$. σ is the required isomorphism.

The next Proposition follows from Theorem 24.3 **[14]**.

PROPOSITION 26. $\mathrm{Hom}(M, \_\,)$: $MS \to MS$ *preserves pullbacks, multiple pullbacks, terminal objects, inverse images, finite intersections, and limits.*

THEOREM 27. $\mathrm{Hom}(M, \_\,)$: $MS \to MS$ *has a left adjoint.*

Proof. Consider the functor diagram, where $G = \mathrm{Hom}(M, \_\,)$, $U = \mathrm{hom}_S(M, \_\,)$, and $V = \text{Forgetful}$.

$$\begin{tikzcd} MS \arrow[rr, "G"] \arrow[dr, "U"'] & & MS \arrow[dl, "V"] \\ & SET & \end{tikzcd}$$

Clearly this diagram commutes. By Propositions 15, 18, and 19, MS is complete, cocomplete, locally small, and colocally small. By Proposition 26, G preserves limits. By Theorem 30.20 **[14]**, U has a left adjoint. Clearly V is faithful. The result follows from Theorem 28.12 **[14]**.

7. The tensor product in MS.

DEFINITION. We denote the left adjoint of $\mathrm{Hom}(M, \_): MS \to MS$ $\_ \otimes M$, and we refer to $M' \otimes M$ as the *tensor product* of M' and M. The closure system on $M' \otimes M$ is denoted $y' \otimes y$.

REMARKS. The adjoint situation, $(\eta, \delta): \_ \otimes M \dashv \mathrm{Hom}(M, \_)$, gives, for each object M_1 of MS a morphism $\eta_{M_1}: M_1 \to \mathrm{Hom}(M_2, M_1 \otimes M_2)$. Define $\psi: M_1 \times M_2 \to M_1 \otimes M_2$ by the rule, $\psi((u_1, u_2)) = (\eta_{M_1}(u_1))(u_2)$ and denote $\psi((u_1, u_2)) = u_1 \otimes u_2$. Note that $s(u_1 \otimes u_2) = su_1 \otimes u_2 = u_1 \otimes su_2$, for all $s \in S$. In fact, ψ is *bilinear*, in the sense that both $\psi(u_1, \_): M_2 \to M_1 \otimes M_2$ and $\psi(\_, u_2)$: $M_1 \to M_1 \otimes M_2$ are morphisms (defined in the obvious ways). Indeed, $\psi(u_1, \_) = \eta_{M_1}(u_1) \in \mathrm{Hom}(M_2, M_1 \otimes M_2)$ by definition. To see that $\psi(\_, u_2) \in \mathrm{Hom}(M_1, M_1 \otimes M_2)$, note that

$$\begin{aligned}
\psi(_, u_2)^{-1}(U_{y_1 \otimes y_2}) &= \{u_1 \in M_1 \mid \psi(u_1, u_2) \in U_{y_1 \otimes y_2}\} \\
&= \{u_1 \in M_1 \mid (\eta_{M_1}(u_1))(u_2) \in U_{y_1 \otimes y_2}\} \\
&= \{u_1 \in M_1 \mid \eta_{M_1}(u_1) \in [u_2, U_{y_1 \otimes y_2}]\} \\
&= \eta_{M_1}^{-1}([u_2, U_{y_1 \otimes y_2}]).
\end{aligned}$$

DEFINITION. Let $G: \underline{A} \to \underline{B}$ be a functor and let M be an object of $\underline{B}$ a pair (μ, N), where N is an object of $\underline{A}$ and $\mu: M \to G(N)$, is called a *universal map for M with respect to G* (or a G-universal map for M) provided that for each N' (object of $\underline{A}$) and each $f: M \to G(N')$, there is a unique $\underline{A}$-morphism $\bar{f}: N \to N'$ such that the triangle commutes.

$$\begin{tikzcd}
M \arrow[r, "\mu"] \arrow[dr, "f"'] & G(N) \arrow[d, "G(\bar{f})"] & N \arrow[d, "\bar{f}"] \\
& G(N') & N'
\end{tikzcd}$$

NOTATION. Given objects (M_j, y_j), for $j = 1, 2, 3$, let $\mathrm{Bihom}(M_1 \times M_2, M_3)$, denote the set of all bilinear maps $M_1 \times M_2 \to M_3$.

PROPOSITION 28. *The map* $\theta: \mathrm{Bihom}(M_1 \times M_2, M_3) \to \mathrm{Hom}(M_1, \mathrm{Hom}(M_2, M_3))$ *given by,* $\theta(\sigma) = \bar{\sigma}$, *where* $(\bar{\sigma}(u_1))(u_2) = \sigma(u_1, u_2)$, *is a bijection.*

THEOREM 29. *Let* $a \in \mathrm{Bihom}(M_1 \times M_2, M_3)$. *Then there exists a unique* $\bar{\bar{\sigma}} \in \mathrm{Hom}(M_1 \otimes M_2, M_3)$ *such that* $\bar{\bar{\sigma}}\psi = \sigma$ [where

ψ: $M_1 \times M_2 \to M_1 \otimes M_2$ *is the canonical map,* $(m_1, m_2) \to m_1 \otimes m_2$; i.e., $\theta(\psi) = \eta_{M_1}$].

Proof.

$$\begin{tikzcd}
M_1 \times M_2 \arrow[rr, "\psi"] \arrow[dr, "\sigma"'] & & M_1 \otimes M_2 \arrow[dl, "\bar{\bar{\sigma}}"] \\
& M_3 &
\end{tikzcd}$$

To complete the first diagram with a morphism $\bar{\bar{\sigma}}$, we make use of the fact that, by Theorem 27.3 [14] $(\eta_{M_1}, M_1 \otimes M_2)$ is a universal map for M_1 with respect to $\mathrm{Hom}(M_2, \_\ )$.

$$\begin{tikzcd}
M_1 \arrow[r, "\eta_{M_1}"] \arrow[d, "\bar{\sigma}"] & \mathrm{Hom}(M_2, M_1 \otimes M_2) \arrow[dl, "\mathrm{Hom}(M_2, \bar{\bar{\sigma}})"] \\
\mathrm{Hom}(M_2, M_3) &
\end{tikzcd}
\qquad
\begin{tikzcd}
M_1 \otimes M_2 \arrow[d, "\bar{\bar{\sigma}}"] \\
M_3
\end{tikzcd}$$

Thus, there exists a unique $\bar{\bar{\sigma}} \in \mathrm{Hom}(M_1 \otimes M_2, M_3)$ such that $\bar{\sigma} = \mathrm{Hom}(M_2, \bar{\bar{\sigma}})\eta_{M_1}$ i.e., $\bar{\sigma} = \bar{\bar{\sigma}}\eta_{M_1}$. Note that $\bar{\bar{\sigma}}$ makes the first diagram commute.

PROPOSITION 30. *$M_1 \otimes M_2 = \{m_1 \otimes m_2 \mid m_1 \in M_1, \ m_2 \in M_2\}$ and $y_1 \otimes y_2$ is the closure operation coinduced on $M_1 \otimes M_2$ by the family,*

$$F = \{\eta_{M_1}(m_1) \mid m_1 \in M_1\} \cup \{\eta_{M_2}(m_2) \mid m_2 \in M_2\}.$$

Proof. Let $M = \{m_1 \otimes m_2 \mid m_1 \in M_1, \ m_2 \in M_2\}$. Then $M \subseteq M_1 \otimes M_2$. Although F is not a covering family, we can form the coinduced closure, y as follows: Let $Q_1 = \{U \subseteq M \mid (\eta_{M_1}(m_1)^{-1}(U))_{y_2} = \eta_{M_1}(m_1)^{-1}(U) \ \ \forall m_1 \in M_1\} \cap \{U \subseteq M \mid (\eta_{M_2}(m_2)^{-1}(U))_{y_1} = \eta_{M_2}(m_2)^{-1}(U) \ \ \forall m_2 \in M_2\}$. Let $Q = \{U \in Q_1 \mid (U: (u_1 \otimes u_2))_x = U: (u_1 \otimes u_2), \ \forall u_1 \otimes u_2 \in M\}$ and note that $Q_1 = Q$. For each $V \subseteq M$, let $V_y = \cap \{U \in Q \mid V \subseteq U\}$. Then (M, y) is an object of MS and y is the finest closure system on M which permits all the S-maps in F to be morphisms into M. Define $\xi: M_1 \times M_2 \to M$ by the rule: $\xi(m_1, m_2) = m_1 \otimes m_2$. Then ξ is bilinear and surjective.

$$\begin{tikzcd}
M_1 \times M_2 \arrow[r, "\xi"] \arrow[d, "\sigma"] & M \arrow[dl, "\bar{\bar{\sigma}}"] \\
M_3 &
\end{tikzcd}$$

Let $\sigma: M_1 \times M_2 \to M_3$ be a bilinear map. Define $\bar{\bar{\sigma}}: M \to M_3$ by the rule: $\bar{\bar{\sigma}}(m_1 \otimes m_2) = \sigma(m_1, m_2)$. Then $\bar{\bar{\sigma}}$ is a morphism and the diagram commutes. In fact, $\bar{\bar{\sigma}}$ is the identity morphism, $m_1 \otimes m_2 \to m_1 \otimes m_2$ and, by Theorem 29, its inverse is also a morphism; hence, $M_1 \otimes M_2 = M$ and $y_1 \otimes y_2 = y$.

PROPOSITION 31. *For any objects* (M_1, y_1), (M_2, y_2) *in* MS, $M_1 \otimes M_2 \cong M_2 \otimes M_1$.

PROPOSITION 32. *For any object* (M, y) *of* MS, $S \otimes M \cong M \cong M \otimes S$.

Proof. Let $\mu: S \otimes M \to M$ be the map given by $\mu(s \otimes m) = sm$. Note that $\mu = \eta_S(1)^{-1}$. μ is the required isomorphism.

PROPOSITION 33. $\otimes$ *is associative.*

Proof. By Theorem 10 **[17]** it is enough to show that $\mathrm{Hom}(M_1 \otimes M_2, M_3) \cong \mathrm{Hom}(M_1, \mathrm{Hom}(M_2, M_3))$. By Theorem 27.9 **[14]**, the adjoint situation, $(\eta, \delta): \_ \otimes M \dashv \mathrm{Hom}(M, \_)$, gives a bijection $\alpha: \mathrm{Hom}(M_1 \otimes M_2, M_3) \to \mathrm{Hom}(M_1, \mathrm{Hom}(M_2, M_3))$ defined by the rule, $(\alpha(f)(m_1))(m_2) = f(m_1 \otimes m_2)$, for all $f \in \mathrm{Hom}(M_1 \otimes M_2, M_3)$. α is the required isomorphism since, for all $m_1 \otimes m_2 \in M_1 \otimes M_2$ and all $U_{y_3} \subseteq M_3$, $\alpha([m_1 \otimes m_2, U_{y_3}]) = [m_1, [m_2, U_{y_3}]]$.

PROPOSITION 34. $\_ \otimes M$ *preserves colimits. In particular,* $\_ \otimes M$ *preserves coproducts.*

PROPOSITION 35. *Let* $\varphi \in \mathrm{Hom}(M_1, M_2)$. *Then, for any object* (M, y) *in* MS $\varphi \otimes M: M_1 \otimes M \to M_2 \otimes M$ *is the map,* $u_1 \otimes u \to \varphi(u_1) \otimes u$.

Proof.

$$\begin{tikzcd} M_1 \arrow[r, "\eta_{M_1}"] \arrow[d, "\varphi"'] & \mathrm{Hom}(M, M_1 \otimes M) \arrow[d, "\mathrm{Hom}(M, \varphi \otimes N)"'] \\ M_2 \arrow[r, "\eta_{M_2}"] & \mathrm{Hom}(M, M_2 \otimes M) \end{tikzcd}$$

The adjoint situation $(\eta, \delta): \_ \otimes M \dashv \mathrm{Hom}(M, \_)$ makes the diagram commute for each object M. Thus, for each $m_1 \in M_1$, $\mathrm{Hom}(M, \varphi \otimes M)(\eta_{M_1}(m_1)) = \eta_{M_2}(\varphi(m_1))$; i.e., for all $m \in M$,

$$(\varphi \otimes M)(\eta_{M_1}(m_1))(m) = (\varphi \otimes M)(m_1 \otimes m)$$
$$= \eta_{M_2}(\varphi(m_1))(m) = \varphi(m_1) \otimes m.$$

NOTATION. $\varphi \otimes 1$ will sometimes be written instead of $\varphi \otimes M$ in cases where no confusion will result.

PROPOSITION 36. *For any object (M, y), the functor $\_ \otimes M$ preserves epimorphisms.*

DEFINITION. An object (m, y) is *Flat* if the functor $\_ \otimes M$ preserves monomorphisms.

PROPOSITION 37. *S is a flat object of MS.*

PROPOSITION 38. *Let $\{(M_j, y_j) \mid j \in J\}$ be a family of objects of MS. Then $(\Sigma M_j, \Sigma y_j)$ is flat iff (M_j, y_j) is flat for each $j \in J$.*

PROPOSITION 39. *Every projective object of MS is flat.*

8. Restriction and extension of scalars.

REMARKS. Let $\varphi: (S, x) \to (S', x')$ be a morphism of ideal systems; i.e., $\varphi(st) = \varphi(s)\varphi(t)$, for all $s, t \in S$, and $\varphi(A_x) \subseteq (\varphi(A))_{x'}$, for all $A \subseteq S$. Then any object (M', y') of MS' can be considered as an object of MS in the following manner: for each $s \in S$, $u' \in M'$, define $su' = \varphi(s)u'$. It is easy to verify that, with this S-set structure, (M', y') is an object of MS (the closure system y' does not change). This process is usually referred to as *restriction of scalars.* Let $\xi' \in \mathrm{Hom}_{S'}(M_1', M_2')$. If we restrict scalars as described above, we can consider both objects M_1' and M_2' as objects of MS and then ξ' becomes an S-morphism with its S-map structure defined by the rule, $\xi'(su') = \xi'(\varphi(s)u')$ for all $s \in S$.

PROPOSITION 40. *Let $\varphi: (S, x) \to (S', x')$ be a morphism of ideal systems. Then the process of restriction of scalars determines a faithful, covariant functor, $R_\varphi: MS' \to MS$, which preserves monomorphisms and epimorphisms.*

DEFINITION. A functor which preserves monomorphisms and epimorphisms shall be called *exact.*

PROPOSITION 41. *Let $\varphi: (S, x) \to (S', x')$ be a morphism of ideal systems. Then the functor $R_\varphi: MS' \to MS$ has a left adjoint $E_\varphi: MS \to MS'$ given by the rule, $E_\varphi(M) = M \otimes R_\varphi S'$ for all objects*

(M, y) of MS [$E_\varphi(M)$ is given S'-set structure by defining for each $s' \in S'$ and each $u \otimes t' \in E_\varphi(M)$, $s'(u \otimes t') = u \otimes s't'$] and $E_\varphi(\delta) = \delta \otimes R_\varphi S'$ for any morphism $\delta \in$ Hom$_S(M_1, M_2)$.

Proof. By Theorem 27.9 **[14]**, it is sufficient to show that the two set-valued bifunctors, $\mathrm{hom}_{S'}(E_\varphi\,\_\,,\_\,)$ and $\mathrm{hom}_S(\_\,, R_\varphi\,\_\,)$ are naturally isomorphic. Thus, let (M, y) be an object of MS and (M', y') be an object of MS', and define $\beta\colon \mathrm{hom}(E_\varphi M, M') \to \mathrm{hom}(M, R_\varphi M')$ by the rule: $\beta(f)(m) = f(m \otimes 1')$ $\forall m \in M$. Then β is a bijection.

REMARK. The functor $E_\varphi\colon MS \to MS'$ is usually referred to as *extension of scalars.*

PROPOSITION 42. *Let $\varphi\colon (S, x) \to (S', x')$ be a morphism of ideal systems. Then the functor $R_\varphi\colon MS' \to MS$ has a right adjoint $H_\varphi\colon MS \to MS'$ given by the rule: $H_\varphi(M) =$ Hom$_S(R_\varphi S', M)$ $\forall$ objects (M, y) of MS [$H_\varphi(M)$ becomes an object of MS' by defining for each $s' \in S'$ and each $\sigma \in H_\varphi(M)$, $(s'\sigma)(t') = \sigma(s't')$ $\forall t' \in R_\varphi S'$] and $H_\varphi(\lambda) =$* Hom$_S(R_\varphi S', \lambda)$ $\forall \lambda \in$ Hom$_S(M_1, M_2)$.

Proof. By Theorem 27.9 **[14]**, it is sufficient to show that the two set-valued bifunctors, $\mathrm{hom}_S(R_\varphi\,\_\,,\_\,)$ and $\mathrm{hom}_{S'}(\_\,, H_\varphi\,\_\,)$ are naturally isomorphic. Thus, let (M, y) be an object of MS and (M', y') be an object of MS', and define $\gamma\colon \mathrm{hom}(R_\varphi M', M) \to \mathrm{hom}(M', H_\varphi M)$ by the rule: $(\gamma(g)(m'))(s') = g(s'm)$ $\forall s' \in S'$ and $\forall g \in \mathrm{hom}(R_\varphi M', M)$. Then γ is a bijection.

REMARK. Let $\varphi\colon (S, x) \to (S', x')$ be a morphism of ideal systems. Then for any object (M, y) of MS and any object (M', y') of MS', $M \otimes R_\varphi M'$ may be regarded as an object of MS' if it is given S'-set structure in the following manner: $s'(m \otimes m') = m \otimes s'm'$ $\forall s' \in S'$ and $\forall m \otimes m' \in M \otimes R_\varphi M'$.

PROPOSITION 43. *Let $\varphi\colon (S, x) \to (S', x')$ be a morphism of ideal systems. Let (M, y) be an object of MS and let (M', y') be an object of MS'. Then (in MS) $M \otimes R_\varphi M' \cong R_\varphi(E_\varphi M \otimes' M')$* [where $\otimes'$ indicates that the tensor product is formed in MS'].

Proof. Let $\alpha\colon M \times R_\varphi M' \to R_\varphi(E_\varphi M \otimes' M')$ be defined by the rule: $\alpha(u, u') = (u \otimes 1') \otimes' u'$. Then α is S-bilinear; hence, there exists an S-morphism $\bar{\alpha}\colon M \otimes R_\varphi M' \to R_\varphi(E_\varphi M \otimes' M')$ such that the diagram commutes; i.e., $\bar{\alpha}(m \otimes m') = \alpha(m, m') = (m \otimes 1') \otimes' m'$ [ψ is the canonical bilinear map]. Note $\bar{\alpha}(m \otimes s'm') = (m \otimes s') \otimes' m'$.

$$\begin{tikzcd} M \otimes R_\varphi M' \arrow[r, "\bar{\alpha}"] & R_\varphi(E_\varphi M \otimes' M') \\ M \times R_\varphi M' \arrow[u, "\psi"] \arrow[ur, "\alpha"'] & \end{tikzcd}$$

Let $\hat{\alpha}: M \otimes R_\varphi M' \to E_\varphi M \otimes' M'$ denote $\bar{\alpha}$ regarded as an S'-morphism. Now, for each $m' \in R_\varphi M'$, define $\underline{m}': M \times R_\varphi S' \to M \otimes R_\varphi M'$ by the rule: $\underline{m}'(m, s') = m \otimes s'm'$. Then $\underline{m}'$ is S-bilinear; hence, there exists an S-morphism $\hat{\underline{m}}': M \otimes R_\varphi S' \to M \otimes R_\varphi M'$ such that $\hat{\underline{m}}'(m \otimes s') = m \otimes s'm'$.

$$\begin{tikzcd} M \otimes R_\varphi S' \arrow[r, "\hat{\underline{m}}'"] & M \otimes R_\varphi M' \\ M \times R_\varphi S' \arrow[u, "\psi"] \arrow[ur, "\underline{m}'"'] & \end{tikzcd}$$

Actually, $\hat{\underline{m}}'$ is an S'-morphism with domain $E_\varphi M$ and codomain $M \otimes R_\varphi M'$, the latter regarded as an object of MS'. Let $\beta: E_\varphi M \times M' \to M \otimes R_\varphi M'$ be defined by the rule: $\beta(m \otimes s', m') = m \otimes s'm'$. Note that β is well defined since, for each fixed $m' \in R_\varphi M'$, $\beta(\_\,, m') = \hat{\underline{m}}'$ and, hence, does not depend upon the choice of representative of $m \otimes s'$. Since β is S'-bilinear, it follows that there is an S'-morphism $\hat{\beta}: (M \otimes R_\varphi S') \otimes M' \to M \otimes R_\varphi M'$ such that $\hat{\beta}((m \otimes s') \otimes' m') = m \otimes s'm'$.

$$\begin{tikzcd} (M \otimes R_\varphi S') \otimes' M' \arrow[r, "\hat{\beta}"] & M \otimes R_\varphi M' \\ (M \otimes R_\varphi S') \times M' \arrow[u, "\psi"] \arrow[ur, "\beta"'] & \end{tikzcd}$$

Thus, we have produced S'-morphisms, $\hat{\alpha}: M \otimes R_\varphi M' \to E_\varphi M \otimes' M'$ and $\hat{\beta}: E_\varphi M \otimes' M' \to M \otimes R_\varphi M'$ which are inverses of one another; i.e., $M \otimes R_\varphi M' \cong E_\varphi M \otimes M'$ in MS'. It follows that $M \otimes R_\varphi M' \cong R_\varphi(E_\varphi M \otimes' M')$ in MS.

PROPOSITION 44. *Let $\varphi: (S, x) \to (S', x')$ be a morphism of ideal systems. Then the two MS-valued bifunctors, $\_ \otimes R_\varphi \_ : MS \times MS' \to MS$ and $R_\varphi(E_\varphi \_ \otimes' \_)$: $MS \times MS' \to MS$ are naturally isomorphic.*

PROPOSITION 45. *Let $\varphi: (S, x) \to (S', x')$ be a morphism of ideal systems. Suppose that $R_\varphi S'$ is a flat object of MS. Then $R_\varphi M'$ is flat in MS for all flat objects (M', y') in MS'.*

Proof. The functor $\_\otimes R_\varphi M'\colon MS \to MS$ preserves monomorphisms whenever (M', y') is a flat object of MS'.

PROPOSITION 46. *Let $\varphi\colon (S,x)\to(S',x')$ be a morphism of ideal systems. Suppose that (M,y) is a flat object of MS. Then $E_\varphi M$ is flat in MS'.*

Proof. Let $\xi\colon M_1'\to M_2'$ be an S'-monomorphism. Then by Proposition 40, $R_\varphi\xi\colon R_\varphi M_1'\to R_\varphi M_2'$ is an S-monomorphism; hence, since M is flat in MS, $M\otimes R_\varphi\xi\colon M\otimes R_\varphi M_1'\to M\otimes R_\varphi M_2'$ is an S-monomorphism. By Proposition 44, we have that $R_\varphi(E_\varphi M\otimes'\xi)\colon R_\varphi(E_\varphi M\otimes' M_1')\to R_\varphi(E_\varphi M\otimes' M_2')$ is an S-monomorphism; hence, $E_\varphi M\otimes'\xi\colon E_\varphi M\otimes' M_1'\to E_\varphi M\otimes' M_2'$ is an S'-monomorphism (since R_φ is faithful).

9. Monads and algebras in $M\{0,1\}$.

NOTATION. We shall denote the category $M\{0,1\}$, simply, $\mathfrak{B}$. For any ideal system, (S,x), $\tau\colon\{0,1\}\to S$ will denote the map, $\tau(0)=0$, $\tau(1)=1$. Clearly τ is a morphism of ideal systems [$\{0,1\}$ is given the obvious (s-system) closure system]. In the sequel we will denote $R_\tau S$, simply, S.

THEOREM 47. *For any ideal system (S,x), $\bar{K}_S=(K_S,\eta,\mu)$ is a monad in $\mathfrak{B}$, where $K_S\colon\mathfrak{B}\to\mathfrak{B}$ is the functor, $\_\otimes S$, and $\eta\colon 1_{\mathfrak{B}}\to K_S$ is the natural transformation given by, $\eta_M(m)=m\otimes 1$, and $\mu\colon K_SK_S\to K_S$ is the natural transformation, $\mu_M\colon (M\otimes S)\otimes S\to M\otimes S$ given by $\mu_M((m\otimes s)\otimes t)=m\otimes st$.*

Proof. The "unit," η, and the "multiplication," μ, make the following diagrams commute:

$$\begin{tikzcd} K_S \arrow[r, "\eta K_S"] \arrow[dr, equal] & K_SK_S \arrow[d, "\mu"] & K_S \arrow[l, "K_S\eta"'] \arrow[dl, equal] \\ & K_S & \end{tikzcd} \qquad \begin{tikzcd} K_SK_SK_S \arrow[r, "\mu K_S"] \arrow[d, "K_S\mu"'] & K_SK_S \arrow[d, "\mu"] \\ K_SK_S \arrow[r, "\mu"'] & K_S \end{tikzcd}$$

where

$$(\eta K_S)_M=\eta_{M\otimes S}\colon M\otimes S\to(M\otimes S)\otimes S$$

$$(K_S\eta)_M=\eta_M\otimes 1\colon M\otimes S\to(M\otimes S)\otimes S$$

$$(\mu K_S)_M=\mu_{M\otimes S}\colon ((M\otimes S)\otimes S)\otimes S\to(M\otimes S)\otimes S$$

$$(K_S\mu)_M=\mu_M\otimes 1\colon ((M\otimes S)\otimes S)\otimes S\to(M\otimes S)\otimes S.$$

THEOREM 48. *For any ideal system*, (S, x), $\bar{H}_S = (H_S, \epsilon, \delta)$ *is a comonad in* $\mathfrak{B}$, *where* $H_S : \mathfrak{B} \to \mathfrak{B}$ *is the functor*, $\operatorname{Hom}(S, \_)$, *and* $\epsilon : H_S \to 1_{\mathfrak{B}}$ *is the natural transformation*, $\epsilon_M : \operatorname{Hom}(S, M) \to M$, *given by*, $\epsilon_M(f) = f(1)$, *and* $\delta : H_S \to H_S H_S$ *is the natural transformation*, $\delta_M : \operatorname{Hom}(S, M) \to \operatorname{Hom}(S, \operatorname{Hom}(S, M))$, *given by*, $(\delta_M(f)(s))(t) = (sf)(t)$ *for all* $t \in S$.

Proof. One must verify here that the following diagrams commute:

$$
\begin{tikzcd}
H_S & H_S H_S \arrow[l, "\epsilon H_S"'] \arrow[r, "H_S\epsilon"] & H_S \\
& H_S \arrow[u, "\delta"] \arrow[ul, equal] \arrow[ur, equal] &
\end{tikzcd}
\qquad
\begin{tikzcd}
H_S H_S H_S & H_S H_S \arrow[l, "\delta H_S"'] \\
H_S H_S \arrow[u, "H_S\delta"] & H_S \arrow[l, "\delta"] \arrow[u, "\delta"']
\end{tikzcd}
$$

where

$$(\epsilon H_S)_M = \epsilon_{\operatorname{Hom}(S,M)} : \operatorname{Hom}(S, \operatorname{Hom}(S, M)) \to \operatorname{Hom}(S, M)$$

$$(H_S \epsilon)_M = \operatorname{Hom}(S, \epsilon_M) : \operatorname{Hom}(S, \operatorname{Hom}(S, M)) \to \operatorname{Hom}(S, M)$$

$$(\delta H_S)_M = \delta_{\operatorname{Hom}(S,M)} : \operatorname{Hom}(S, \operatorname{Hom}(S, M)) \\ \to \operatorname{Hom}(S, \operatorname{Hom}(S, \operatorname{Hom}(S, M)))$$

$$(H_S \delta)_M = \operatorname{Hom}(S, \delta_M) : \operatorname{Hom}(S, \operatorname{Hom}(S, M)) \\ \to \operatorname{Hom}(S, \operatorname{Hom}(S, \operatorname{Hom}(S, M))).$$

REMARKS. Let (S, x) be an ideal system and let $\mathfrak{B}^S$ denote the category of $\bar{K}_S$-algebras. Let $G : MS \to \mathfrak{B}^S$ be defined as follows: For each object (M, y) of MS, $G(M) = \langle R_\tau M, h \rangle$, where $h : R_\tau M \otimes S \to R_\tau M$ is the $\mathfrak{B}$-morphism $m \otimes s \to sm$. For each S-morphism $f : M \to M'$, $G(f) = R_\tau f : R_\tau M \to R_\tau M'$. Then $G(f)$ is a $\mathfrak{B}^S$-morphism and, hence, G is a (covariant) functor.

Now let $F : \mathfrak{B}^S \to MS$ be defined as follows: For each object $\langle M, h \rangle$ of $\mathfrak{B}^S$ (where (M, y) is an object of $\mathfrak{B}$), $F(\langle M, h \rangle) = (\bar{M}, y)$, where $\bar{M} = M$, equipped with the S-multiplication, $sm = h(s \otimes m)$ for all $s \in S$, $m \in M$. For any $\mathfrak{B}^S$-morphism $g : \langle M, h \rangle \to \langle M', h' \rangle$, $F(g) = \bar{g}$, where $\bar{g} = g$, converted into an S-map by taking $\bar{g}(sm) = g(h(s \otimes m))$ for all $s \in S$, $m \in M$.

THEOREM 49. [Monadicity]. *For any ideal system*, (S, x), $\mathfrak{B}^S$ *is isomorphic to* MS.

Proof. With notation as in the remarks above we need only show that $FG = 1_{MS}$ and $GF = 1_{\mathfrak{B}^s}$. To show $FG = 1_{MS}$: since for any object (M, y) of MS, $FG(M) = F(\langle R_\tau M, h\rangle) = (\overline{R_\tau M}, y)$, where $\overline{R_\tau M} = R_\tau M$, endowed with an S-multiplication which is derived from the map h ; i.e., $(\overline{R_\tau M}, y) = (M, y)$. For each object $\langle M, h\rangle$ of $\mathfrak{B}^s$, $GF(\langle M, h\rangle) = G((\bar{M}, y)) = \langle R_\tau\bar{M}, \bar{h}\rangle$, where $\bar{h}: R_\tau\bar{M} \otimes S \to R_\tau\bar{M}$ is defined by the rule, $\bar{h}(m \otimes s) = sm = h(m \otimes s)$. Thus, $\bar{h} = h$, and, since it is clear that $R_\tau\bar{M} = M$, it follows that $GF(\langle M, h\rangle) = \langle M, h\rangle$.

Concluding remarks. The monads and comonads constructed above provide the tools with which resolutions and derived functors can be constructed which, in turn, lead to a (co)homology theory for $\mathfrak{B}$. The category of pointed topological spaces and basepoint preserving maps, PTOP, can be found in $\mathfrak{B}$. In fact, the inclusion functor PTOP $\to \mathfrak{B}$ is a full, faithful embedding.

References

1. M. F. Atiyah, and I. G. MacDonald, *Introduction to Commutative Algebra*, Addison-Wesley, (1969).
2. Karl E. Aubert, *Theory of X-ideals*, Acta Mathematica, **107** (1962), 1–52.
3. ———, *Ideal systems and lattice theory I*, Algebra Universalis, **1** (1971), No. 2, 204–213.
4. ———, *Ideal Systems and Lattice Theory II*, Tufts University (mimeographed notes) 1971.
5. Karl E. Aubert, and G. Gismarvik, *Chinese Remainder Theorems in Ideal Systems*, Tufts University (mimeographed notes) 1971.
6. Karl E. Aubert, *Un Nouvel Axiome dans la Théorie, des Systèmes d'idéaux*, C. R. Acad. Sci. Paris, t, **272** (1971), 422–425.
7. ———, *Localisation Dans les Systèmes D'Idéaux*, C. R. Acad. Sci. Paris, t, **272** (1971), 465–468.
8. Karl E. Aubert, and Erling R. Hansén, *Systèmes De Modules*, C. R. Acad. Sci. Paris, t, **272** (1971), 525–528.
9. ———, *Additive Ideal Systems*, J. Algebra, **18** (1971), 511–528.
10. M. Barr, and J. Beck, *Homology and Standard Constructions.* Seminar on Triples and Categorical Theory. Springer-Verlag, **80** (1969), 245–335.
11. Henri Cartan, and Samuel Eilenberg, *Homological Algebra*, Princeton University, (1956).
12. J. Duskin, *$K(\pi, n)$-torsors and the interpretation of "triple" cohomology*, Proc. Nat. Acad. Sci. USA, **71** (1974), 2554–2557.
13. Paul Ezust, *Systems of Modules over Systems of Ideals*, Thesis, Tuft University, 1975.
14. Horst Herrlich, and G. E. Strecker, *Category Theory*, Allyn and Bacon, (1973).
15. P. J. Hilton, and U. Stammbach, *A Course in Homological Algebra*, Springer-Verlag, (1970).
16. Peter Hilton, *Lectures in Homological Algebra*, Regional Conference Series in Mathematics #8. Am. Math. Soc., (1970).
17. G. M. Kelly, *Tensor products in categories*, J. Algebra, **2** (1965), 15–37.
18. F. E. J. Linton, *Applied Functorial Semantics*, Seminar on Triples and Categorical Theory, Springer-Verlag, **80** (1969), 53–74.
19. ———, *An outline on Functorial Semantics*, Seminar on Triples and Categorical Theory, Springer-Verlag, **80** (1969).
20. Saunders MacLane, *Homology*, Springer-Verlag, (1963).

21. ———, *Categories for the Working Mathematician*, Springer-Verlag, (1971).

22. Barry Mitchell, *Theory of Categories*, Academic Press, (1965).

23. Bodo Pareigis, *Categories and Functors*, Academic Press, (1970).

24. Bo Strenstrom, *Flatness and Localization over Monoids*, Math. Nachr., **48** (1971), 315–334.

25. F. Umer, *On Cotriples and Andre (co)Homology, Their Relationship with Classical Homological Algebra*, Seminar on Triples and Category Theory, Springer-Verlag, **80** (1969), 376–398.

Received June 30, 1975 and in revised form November 30, 1976. The results presented herein are drawn from the author's doctoral dissertation which was written at Tufts University under the direction of Professor George F. Leger.

SUFFOLK UNIVERSITY
BOSTON, MA 02114

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No. 2, 1977

THE DEFICIENCY INDEX OF A THIRD ORDER OPERATOR

RICHARD C. GILBERT

Let L be a formally selfadjoint third order linear ordinary differential operator defined on $[r,\infty)$. Using a method of Fedorjuk, asymptotic formulas are found for the solutions of $Ly = i\sigma y$, $\sigma \neq 0$. These formulas are used to determine the deficiency index of L when L has polynomial coefficients. As a consequence, the deficiency index is determined for values of the parameters involved for which it has not previously been determined.

1. Introduction. The general form of a third order formally selfadjoint linear ordinary differential operator L can be written

$$(1)\quad Ly = (ib_2y'')' + [(2^{-1}ib_2' + a_1)y']' + ib_1y' + (2^{-1}ib_1' + a_0)y,$$

where a_0, a_1, b_1, b_2 are real functions of x and $b_2(x) \neq 0$. (See [**4**, Ch. 1, §1.5]. We have assumed sufficient differentiability on the coefficients so that the Dunford and Schwartz form can be written in the form (1).) Unsworth [**12**] considered the case that $b_2(x) = 2$, $b_1(x) = 2ax^{\alpha}$, $a_1(x) = bx^{\beta}$, $a_0(x) = cx^{\gamma}$, $1 \leqq x < \infty$. Using the asymptotic methods of Devinatz [**3**], Unsworth deduced the deficiency index of L for various values of the parameters a, b, c, α, β, γ. Pfeiffer [**10**] considered the case $b_2(x) = 1$, $b_1(x) = ax^{\alpha}$, $a_1(x) = 0$, $a_0(x) = cx^{\gamma}$. The purpose of the present article is to obtain by the method of Fedorjuk [**6**] asymptotic formulas for the solutions of $Ly = i\sigma y$, $\sigma \neq 0$, and to apply these formulas to finding the deficiency index of L for the case $b_2(x) = 1$, $b_1(x) = ax^{\alpha}$, $a_1(x) = bx^{\beta}$, $a_0(x) = cx^{\gamma}$. Although Fedorjuk applied his method only to even order operators, it can be used for odd order operators as well. Shirikyan [**11**] applied the Fedorjuk method to a certain class of odd order operators. It turns out that the Fedorjuk method applied to the above case yields the deficiency index for values of the parameters different from Unsworth and Pfeiffer.

It is known that, except for a first order operator, a differential operator of order n cannot have deficiency index (n,p) or (p,n), where $p < n$. (See Atkinson [**1**] or Kogan and Rofe-Beketov [**7**], [**8**].) Further, for an operator of order $n = 2\nu - 1$ it is known that the deficiency numbers n_+ and n_- satisfy the inequalities $\nu \leqq n_+ \leqq 2\nu - 1$, $\nu - 1 \leqq n_- \leqq 2\nu - 1$, or the same inequalities with n_+ and n_- interchanged. (See

Everitt **[5]** or Kogan and Rofe-Beketov **[8]**.) It follows that the deficiency indices (2, 1), (1, 2), (2, 2) and (3, 3), obtained in this paper and by Unsworth and Pfeiffer, are the only possible deficiency indices for a third order operator.

2. Asymptotic formulas for the solutions of $Ly = i\sigma y$. We shall make the following assumptions on the coefficients a_0, a_1, b_1, b_2 of L. The need for the various assumptions will be seen as we go along.

In all that follows in this article, it will be necessary in various places to require that x is sufficiently large. We shall therefore assume once and for all that x_0 is chosen so large that if $x \geqq x_0$, then x is sufficiently large in all places where this is needed. We shall also often omit the stipulation $x \geqq x_0$ when it is clear from the context that this is needed.

ASSUMPTION I. $b_1(x), b_2(x) \in C^3[r, \infty)$. $a_0(x), a_1(x) \in C^2[r, \infty)$. $b_2(x) \neq 0$ for $x \geqq r$, $b_2(x) = 1 + o(1)$ as $x \to +\infty$. $a_0(x) \neq 0$ for $x \geqq r$. Either $a_0(x) \to +\infty$ and $a_0'(x) > 0$ for $x \geqq x_0$, or else $a_0(x) \to -\infty$ and $a_0'(x) < 0$ for $x \geqq x_0$.

ASSUMPTION II. $\lim_{x\to\infty} a_1/a_0^{1/3} = d \neq 3/2^{2/3}$, $b_1/a_0^{2/3} = o(1)$, $b_1'/a_0 = o(1)$, $b_2'/a_0^{1/3} = o(1)$.

ASSUMPTION III. $b_2''/a_0^{2/3} = o(1)$, $a_1'/a_0^{2/3} = o(1)$, $b_1''/a_0^{4/3} = o(1)$, $a_0'/a_0^{4/3} = o(1)$.

ASSUMPTION IV. b_2' and $b_1'/a_0^{2/3}$ are absolutely integrable on $[r, \infty)$.

Let

$$f(\lambda, x) = -\lambda^3 + im(x)b_2^{-1}(x)\lambda^2 - b_1(x)b_2^{-1}(x)\lambda + in(x)b_2^{-1}(x), \tag{2}$$

where

$$m(x) = 2^{-1}ib_2'(x) + a_1(x), \tag{3}$$

$$n(x) = 2^{-1}ib_1'(x) + a_0(x) - i\sigma. \tag{4}$$

Here σ is a real constant, $\sigma \neq 0$.

Let

$$\tau(x) = [a_0(x)b_2^{-1}(x)]^{1/3}[1 + (b_1'(x) - 2\sigma)(2a_0(x))^{-1}i]^{1/3}, \tag{5}$$

where if $z = \rho e^{i\theta}$, $-\pi < \theta \leqq \pi$, then we take $z^{1/3} = \rho^{1/3}e^{i\theta/3}$. Then,

$\tau^3 = nb_2^{-1}$, and $\tau(x) \neq 0$ for $x \geqq r$.

Putting

$$\lambda = i\eta\tau(x), \tag{6}$$

then

$$f(\lambda, x) = 0 \tag{7}$$

becomes

$$h(\eta, x) = 0, \tag{8}$$

where

$$h(\eta, x) = \eta^3 - m(x)[b_2(x)\tau(x)]^{-1}\eta^2 - b_1(x)[b_2(x)\tau^2(x)]^{-1}\eta + 1. \tag{9}$$

An essential part of the Fedorjuk method is that we should have

$$\lim_{x\to\infty} m(x)[b_2(x)\tau(x)]^{-1} = d + ie_1, \tag{10}$$

$$\lim_{x\to\infty} b_1(x)[b_2(x)\tau^2(x)]^{-1} = d_2 + ie_2, \tag{11}$$

where $d + ie_1$ and $d_2 + ie_2$ are complex constants. Then, as $x \to \infty$, $h(\eta, x)$ approaches a polynomial $h_0(\eta)$ with constant coefficients. We also want $h_0(\eta) = 0$ to have distinct roots. For reasons that will appear later we further want as $x \to \infty$ that $|a_0(x)| \to \infty$ and that

$$\tau(x) = a_0^{1/3}(x)[1 + o(1)]. \tag{12}$$

In I and II we have assumed $a_0(x) \to \pm\infty$, $b_2 = 1 + o(1)$, $b_1'/a_0 = o(1)$ in order that (12) and $|a_0(x)| \to \infty$ might be true. In order to explain the remaining assumptions in I and II, let us note that if (10) and (11) are to be true, we must have

$$\lim_{x\to\infty} (b_2'/a_0^{1/3}) = 2e_1, \tag{13}$$

$$\lim_{x\to\infty} (a_1/a_0^{1/3}) = d, \tag{14}$$

$$\lim_{x\to\infty} (b_1/a_0^{2/3}) = d_2, \tag{15}$$

and $e_2 = 0$. But then (13) and our assumptions that $|a_0| \to \infty$ and $b_2 = 1 + o(1)$ imply that $e_1 = 0$. Further, (15) and the assumptions on a_0 in I and the assumption that $b_1'/a_0^{2/3}$ is absolutely integrable on $[r, \infty)$ in IV imply that $d_2 = 0$. Thus, we have explained the reasons for all the limit assumptions in I and II.

From Assumptions I and II we have that

$$m(x)[b_2(x)\tau(x)]^{-1} = d + f_1(x), \tag{16}$$

$$b_1(x)[b_2(x)\tau^2(x)]^{-1} = f_2(x), \tag{17}$$

where $f_1(x) = o(1)$, $f_2(x) = o(1)$, and $f_1(x)$ and $f_2(x)$ are continuously differentiable on $[r, \infty)$. It follows that

$$h(\eta, x) = h_0(\eta) - \eta^2 f_1(x) - \eta f_2(x), \tag{18}$$

where

$$h_0(\eta) = \eta^3 - d\eta^2 + 1. \tag{19}$$

Since we have assumed in II that $d \neq 3/2^{2/3}$, $h_0(\eta) = 0$ has three distinct nonzero roots. If $d < 3/2^{2/3}$, then $h_0(\eta) = 0$ has one real negative root and two complex conjugate nonreal roots. If $d > 3/2^{2/3}$, then $h_0(\eta) = 0$ has three distinct real roots, one of which is negative and the other two positive. We denote the roots by η_{01}, η_{02}, η_{03}, where $\eta_{01} < \eta_{02} < \eta_{03}$ in the case of three real roots, and η_{01} is real and $\operatorname{Im} \eta_{02} > 0$, $\operatorname{Im} \eta_{03} < 0$ in the case of one real root. In the case of three real roots, $h'(\eta_{01}) > 0$, $h'(\eta_{02}) < 0$, $h'(\eta_{03}) > 0$. In the case of one real root, $h'(\eta_{01}) > 0$. In every case, $h'(\eta_{0k}) \neq 0$, $k = 1, 2, 3$.

According to Bellman [**2**, p. 26], for $x \geqq x_0$, (8) has three distinct roots $\eta_k(x)$, $k = 1, 2, 3$ which are given by the formula

$$\eta_k(x) = (2\pi i)^{-1} \int_{C_k} \eta h_\eta(\eta, x)[h(\eta, x)]^{-1}\, d\eta, \tag{20}$$

where C_k is a small circle around η_{0k}. $\eta_k(x)$ is continuously differentiable, and

$$\eta_k(x) = \eta_{0k}[1 + o(1)]. \tag{21}$$

We have that $h_\eta(\eta_k(x), x) \neq 0$, and that $\eta_k(x) \neq 0$, for $x \geqq x_0$. From (6) one sees that (7) has for $x \geqq x_0$, three distinct continuously differentiable nonzero roots $\lambda_k(x)$ given by

$$\lambda_k(x) = i\eta_k(x)\tau(x), \qquad k = 1, 2, 3, \tag{22}$$

and

(23) $$\lambda_k(x) = ia_0^{1/3}(x)\eta_{0k}[1 + o(1)].$$

We have that $f_\lambda(\lambda_k(x), x) \neq 0$.

ASSUMPTION V. $(b_2')^2/a_0^{1/3}$, $(b_2'')^2/a_0$, $(a_1')^2/a_0$, $(b_1')^2/a_0^{5/3}$, $(b_1'')^2/a_0^{7/3}$, $(a_0')^2/a_0^{7/3}$, $b_2''/a_0^{1/3}$, $b_2'''/a_0^{2/3}$, $a_1''/a_0^{2/3}$, b_1''/a_0, $b_1'''/a_0^{4/3}$, $a_0''/a_0^{4/3}$ are all absolutely integrable on $[r, \infty)$.

ASSUMPTION VI. For each pair j, k, one of the following is true:
(a) $\mathrm{Re}(\lambda_j(x) - \lambda_k(x)) \geqq 0$ for $x \geqq x_0$;
(b) $\mathrm{Re}(\lambda_j(x) - \lambda_k(x)) \leqq 0$ for $x \geqq x_0$, and

$$\int_{x_0}^{\infty} \mathrm{Re}(\lambda_j(x) - \lambda_k(x))\, dx = -\infty;$$

(c) $\int_{x_0}^{\infty} \mathrm{Re}(\lambda_j(x) - \lambda_k(x))\, dx$ is convergent.

Using Assumptions I–VI, it is now possible to obtain asymptotic formulas for the solutions of the equation

(24) $$Ly = i\sigma y.$$

Let w be the column vector with components $w_1 = y$, $w_2 = y'$, $w_3 = ib_2y'' + my'$. (24) is then equivalent to the system

(25) $$w' = A(x)w,$$

where

(26) $$A(x) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & imb_2^{-1} & -ib_2^{-1} \\ -n & -ib_1 & 0 \end{pmatrix}.$$

The eigenvalues of $A(x)$ are the roots of (7), i.e., $\lambda_k(x)$, $k = 1, 2, 3$.

Let us now make the transformation

(27) $$w = T_0(E + T_2)z,$$

where z is a column vector with components z_1, z_2, z_3, and T_0 and T_2 are matrices to be determined, and E is the identity matrix. Then, (25) becomes

$$z' = \Lambda_0 z + (\Lambda_0 T_2 - T_2\Lambda_0 - T_0^{-1}T_0')z + B(x)z, \tag{28}$$

where

$$\begin{aligned} B(x) = (E + T_2)^{-1}[(T_2^2\Lambda_0 + T_2T_0^{-1}T_0')(E + T_2) - T_2'] \\ - T_2\Lambda_0T_2 - T_0^{-1}T_0'T_2, \end{aligned} \tag{29}$$

and

$$\Lambda_0 = T_0^{-1}AT_0.$$

We shall show that we can choose T_0 and T_2 such that for $x \geqq x_0$, T_0^{-1} and $(E + T_2)^{-1}$ exist, and $T_0^{-1}AT_0$ and $\Lambda_0T_2 - T_2\Lambda_0 - T_0^{-1}T_0'$ are diagonal. To that end, we choose T_0 to be a matrix whose columns are eigenvectors for A, namely,

$$T_0 = \begin{pmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ [ib_2\lambda_1 + m]\lambda_1 & [ib_2\lambda_2 + m]\lambda_2 & [ib_2\lambda_3 + m]\lambda_3 \end{pmatrix}, \tag{30}$$

$$T_0^{-1} = \begin{pmatrix} n/\lambda_1F_\lambda(\lambda_1, x) & -i\lambda_1b_2/F_\lambda(\lambda_1, x) & -1/F_\lambda(\lambda_1, x) \\ n/\lambda_2F_\lambda(\lambda_2, x) & -i\lambda_2b_2/F_\lambda(\lambda_2, x) & -1/F_\lambda(\lambda_2, x) \\ n/\lambda_3F_\lambda(\lambda_3, x) & -i\lambda_3b_2/F_\lambda(\lambda_3, x) & -1/F_\lambda(\lambda_3, x) \end{pmatrix}, \tag{31}$$

where

$$F(\lambda, x) = ib_2f(\lambda, x). \tag{32}$$

Then, for $x \geqq x_0$,

$$T_0^{-1}AT_0 = \Lambda_0 = \text{diagonal}[\lambda_j]. \tag{33}$$

We note that

$$\lim_{x\to\infty} a_0^{-2/3}(x)F_\lambda(\lambda_j(x), x) = ih_0'(\eta_{0j}) = \rho_{0j}\exp[i\theta_{0j}], \tag{34}$$

where $-\pi < \theta_{0j} \leqq \pi$, $\rho_{0j} > 0$. Let

$$a_0^{-2/3}(x)F_\lambda(\lambda_j(x), x) = \rho_j(x)\exp[i\theta_j(x)], \tag{35}$$

where $\rho_j(x)$ and $\theta_j(x)$ are chosen so that $\lim_{x\to\infty}\rho_j(x) = \rho_{0j}$, and $\lim_{x\to\infty}\theta_j(x) = \theta_{0j}$. We choose that branch of log such that for $x \geqq x_0$,

$$\log F_\lambda(\lambda_j(x), x) = (2/3)\,\mathrm{Log}\,|a_0(x)| + \mathrm{Log}\,\rho_j(x) + i\theta_j(x). \tag{36}$$

Then, for $t, x \geqq x_0$,

$$(d/dx)\log F_\lambda(\lambda_j(x),x) = [F_{\lambda\lambda}(\lambda_j(x),x)\lambda_j'(x) + F_{\lambda x}(\lambda_j(x),x)][F_\lambda(\lambda_j(x),x)]^{-1}, \tag{37}$$

$$\int_t^x [(d/ds)\log F_\lambda(\lambda_j(s),s)]\,ds = \log F_\lambda(\lambda_j(x),x) - \log F_\lambda(\lambda_j(t),t), \tag{38}$$

$$\exp[(1/2)\log F_\lambda(\lambda_j(x),x)] = [1+o(1)]\,|\,a_0^{1/3}(x)\,|\,\rho_{0j}^{1/2}\exp(i\theta_{0j}/2). \tag{39}$$

Now we note that the elements $(T_0^{-1}T_0')_{jk}$ of the matrix $T_0^{-1}T_0'$ are given for $x \geqq x_0$, by

$$(T_0^{-1}T_0')_{jj} = (1/2)[F_{\lambda\lambda}(\lambda_j(x),x)\lambda_j'(x) + F_{\lambda x}(\lambda_j(x),x) + ib_2'(x)\lambda_j^2(x) + ib_1'(x)][F_\lambda(\lambda_j(x),x)]^{-1}, \tag{40}$$

or,

$$(T_0^{-1}T_0')_{jj} = (1/2)(d/dx)\log F_\lambda(\lambda_j(x),x) + 2^{-1}i[b_2'(x)\lambda_j^2(x) + b_1'(x)][F_\lambda(\lambda_j(x),x)]^{-1}; \tag{41}$$

and

$$(T_0^{-1}T_0')_{jk} = [\lambda_k(ib_2'\lambda_k\lambda_j + ib_1') + m'\lambda_k\lambda_j + n'] \times [(\lambda_k - \lambda_j)F_\lambda(\lambda_j(x),x)]^{-1}, \qquad k \neq j. \tag{42}$$

Let

$$\lambda_j^{(1)} = -(T_0^{-1}T_0')_{jj}, \tag{43}$$

$$\Lambda_1 = \text{diagonal}\,[\lambda_j^{(1)}]. \tag{44}$$

We note that the $\lambda_j^{(1)}(x)$ are continuous for $x \geqq x_0$. Let the matrix T_2 be defined by the equations

$$(T_2)_{jj} = 0, \tag{45}$$

$$(T_2)_{jk} = -(T_0^{-1}T_0')_{jk}(\lambda_k - \lambda_j)^{-1}, \qquad k \neq j. \tag{46}$$

T_2 has been defined so that $\Lambda_0 T_2 - T_2\Lambda_0 - T_0^{-1}T_0'$ is a diagonal matrix; indeed,

$$\Lambda_0 T_2 - T_2\Lambda_0 - T_0^{-1}T_0' = \Lambda_1. \tag{47}$$

Thus, T_0 and T_2 in the transformation (27) have been chosen so that for $x \geqq x_0$, equation (28) is

$$z' = (\Lambda_0 + \Lambda_1)z + Bz. \tag{48}$$

We shall now show that for $x \geqq x_0$, $B(x)$ exists and is continuous, and $\|B(x)\|$ is integrable on $[x_0, \infty)$. To do this will require a series of lemmas whose proofs are mostly straightforward or else contained in Fedorjuk [6] and are therefore omitted. For $x \geqq x_0$, let

$$\lambda(x) = \max_j |\lambda_j(x)|. \tag{49}$$

Then,

$$\lambda(x) = |\tau(x)| \max_j |\eta_j(x)| = |a_0^{1/3}(x)|[1 + o(1)] \max_j |\eta_j| > 0.$$

In the following, the capital letters C and D denote suitably chosen positive constants.

LEMMA 1. $D_1 |a_0^{1/3}(x)| \leqq \lambda(x) \leqq D_2 |a_0^{1/3}(x)|$.

LEMMA 2. $C_1\lambda(x) \leqq |\lambda_j(x) - \lambda_k(x)| \leqq C_2\lambda(x)$, $j \neq k$.

Let

$$\alpha(x) = \max\{|b_2'|, |m'|/\lambda, |b_1'|/\lambda^2, |n'|/\lambda^3\}, \tag{50}$$

$$\beta(x) = \max\{|b_2''|, |m''|/\lambda(x), |b_1''|/\lambda^2(x), |n''|/\lambda^3(x)\}, \tag{51}$$

$$\delta(x) = \max\{|b_2'|, |m'|/|a_0^{1/3}|, |b_1'|/|a_0^{2/3}|, |n'|/|a_0|\}, \tag{52}$$

$$\gamma(x) = \max\{|b_2''|, |m''|/|a_0^{1/3}|, |b_1''|/|a_0^{2/3}|, |n''|/|a_0|\}. \tag{53}$$

LEMMA 3. $\alpha(x) \leqq C\delta(x)$.

LEMMA 4. $\beta(x) \leqq C\gamma(x)$.

LEMMA 5. $C_1\lambda^2(x) \leqq |F_\lambda(\lambda_j(x), x)| \leqq C_2\lambda^2(x)$.

LEMMA 6. $|F_x(\lambda_j(x), x)| \leqq C\lambda^3(x)\alpha(x)$.

LEMMA 7. $|[\lambda_k(x) - \lambda_j(x)]F_\lambda(\lambda_j(x), x)| \geqq C\lambda^3(x)$.

LEMMA 8. $|[\lambda_k(x) - \lambda_j(x)]^2 F_\lambda(\lambda_j(x), x)| \geqq C\lambda^4(x)$.

LEMMA 9. $|F_{\lambda\lambda}(\lambda_j(x), x)| \leqq C\lambda(x)$.

LEMMA 10. $|F_{\lambda x}(\lambda_j(x), x)| \leqq C\lambda^2(x)\alpha(x)$.

LEMMA 11. $|\lambda_j'(x)| \leqq C\alpha(x)\lambda(x)$.

If $A = (A_{jk})_{j,k=1}^{n}$ is an $n \times n$ matrix, we define the norm $\|A\|$ by $\|A\| = n \max_{j,k} |A_{jk}|$.

LEMMA 12. $\|\Lambda_1(x)\| \leqq C\alpha(x)$.

LEMMA 13. $\|T_0^{-1}(x) T_0'(x)\| \leqq C\alpha(x)$.

LEMMA 14. $\|T_2(x)\| \leqq C\alpha(x)/\lambda(x)$.

LEMMA 15. $\|T_2'(x)\| \leqq C\{\alpha^2(x) + \beta(x)]/\lambda(x)$.

LEMMA 16. *$[E + T_2(x)]^{-1}$ exists and is continuous for $x \geqq x_0$, and $\|[E + T_2(x)]^{-1}\| \leqq C$.*

LEMMA 17. *$B(x)$ exists and is continuous for $x \geqq x_0$, and $\|B(x)\| \leqq C[\alpha^2(x) + \beta(x)]/\lambda(x)$.*

We note that Lemmas 16 and 17 depend on the fact that $\lim_{x\to\infty} \alpha(x)/\lambda(x) = 0$, which follows from Assumptions II and III.

LEMMA 18. *$\|B(x)\|$ is integrable on $[x_0, \infty)$.*

We note that Lemma 18 follows from Lemma 17, and Assumption V.

It is now possible to show that (48) has three linearly independent solutions which satisfy certain specified boundary conditions at infinity. To that end, we observe that a fundamental matrix $Z_0(x_0, x)$ for the homogeneous equation

$$z' = (\Lambda_0 + \Lambda_1)z, \qquad x \geqq x_0, \tag{54}$$

is given by

$$Z_0(x_0, x) = \text{diagonal}\left[\exp \int_{x_0}^{x} (\lambda_j(t) + \lambda_j^{(1)}(t))\, dt\right]. \tag{55}$$

Putting

$$Z(x) = U(x) Z_0(x_0, x), \tag{56}$$

we find that $Z(x)$ is a matrix solution of (48) for $x \geqq x_0$ if $U(x)$ satisfies

$$U(x) = C + (KU)(x), \qquad x \geqq x_0, \tag{57}$$

where C is an arbitrary constant matrix, and K is a linear operator on matrices $U(x)$ such that

$$(KU(x))_{jk} = \int_{x_{jk}}^{x} (Z_0(t,x)B(t)U(t)Z_0(x,t))_{jk}\,dt, \tag{58}$$

x_{jk} being an arbitrary number in the interval $[x_0, \infty]$.

Let M be the Banach space of continuous matrices $V(x)$ on $[x_0, \infty)$, with $\|V\|_M = \sup_{x \geqq x_0} \|V(x)\| < \infty$. For reasons that will appear in Lemmas 19 and 20 below, if Assumption VI (a) or (c) holds, we take $x_{jk} = \infty$; if Assumption VI (b) holds, we take $x_{jk} = x_0$. Also, we take $C = E$.

LEMMA 19. *If x_0 is sufficiently large, then K: $M \to M$, and $\|K\|_M \leqq 1/2$.*

Proof. From (58) it follows that if $V \in M$ and if $x \geqq x_0$, then

$$\begin{aligned} |((KV)(x))_{jk}| \leqq \Bigg| \int_{x_{jk}}^{x} & \left[\exp \int_t^x \mathrm{Re}(\lambda_j(s) - \lambda_k(s))\,ds\right] \\ & \times \left[\exp \int_t^x \mathrm{Re}(\lambda_j^{(1)}(s) - \lambda_k^{(1)}(s))\,ds\right] \|B(t)\|\,dt \Bigg| \, \|V\|_M. \end{aligned} \tag{59}$$

By (41), (43) and (38),

$$\begin{aligned} \int_t^x & [\lambda_j^{(1)}(s) - \lambda_k^{(1)}(s)]\,ds \\ &= (1/2)[\log F_\lambda(\lambda_j(x), x) - \log F_\lambda(\lambda_j(t), t)] \\ &\quad - (1/2)[\log F_\lambda(\lambda_k(x), x) - \log F_\lambda(\lambda_k(t), t)] \\ &\quad + (i/2) \int_t^x b_2'(s)\{\lambda_j^2(s)[F_\lambda(\lambda_j(s), s)]^{-1} - \lambda_k^2(s)[F_\lambda(\lambda_k(s), s)]^{-1}\}\,ds \\ &\quad + (i/2) \int_t^x b_1'(s)\{[F_\lambda(\lambda_j(s), s)]^{-1} - [F_\lambda(\lambda_k(s), s)]^{-1}\}\,ds. \end{aligned} \tag{60}$$

It now follows from (36), (49), Lemma 1, Lemma 5, and Assumption IV that $\left| \int_t^x (\lambda_j^{(1)}(s) - \lambda_k^{(1)}(s))\,ds \right|$ is bounded for $t, x \geqq x_0$. Hence, if $V \in M$, then

$$|((KV)(x))_{jk}| \tag{61}$$
$$\leqq C\left|\int_{x_{jk}}^{x}\left[\exp\int_{t}^{x}\operatorname{Re}(\lambda_j(s)-\lambda_k(s))\,ds\right]\|B(t)\|\,dt\right|\|V\|_M,\quad x\geqq x_0.$$

By our choice of x_{jk}, if Assumption VI (a) or (b) holds, $\exp\int_t^x \operatorname{Re}(\lambda_j(s)-\lambda_k(s))\,ds \leqq 1$. If Assumption VI (c) holds, then $\left|\int_t^x \operatorname{Re}(\lambda_j(s)-\lambda_k(s))\,ds\right| \leqq C_1$ for t, $x \geqq x_0$, and therefore $\exp\int_t^x \operatorname{Re}(\lambda_j(s)-\lambda_k(s))\,ds \leqq \exp C_1$. It follows from (61) that

$$|((KV)(x))_{jk}| \leqq C\int_{x_0}^{\infty}\|B(t)\|\,dt\,\|V\|_M,\qquad x\geqq x_0. \tag{62}$$

Hence,

$$\|(KV)(x)\| = 3\max_{j,k}|((KV)(x))_{jk}| \leqq 3C\int_{x_0}^{\infty}\|B(t)\|\,dt\,\|V\|_M. \tag{63}$$

If we now choose x_0 so large that $\int_{x_0}^{\infty}\|B(t)\|\,dt \leqq 1/6C$, then $\|K\|_M \leqq 1/2$. This proves Lemma 19.

LEMMA 20. *If x_0 is sufficiently large, equation* (57) *has a unique solution $U(x)\in M$. It is true that $\|(KU)(x)\| = o(1)$ as $x\to\infty$. $U(x)$ can be written in the form*

$$U(x) = E + o(1),\qquad x\geqq x_0. \tag{64}$$

Proof. The existence and uniqueness of $U(x)$ follows from Lemma 18 and Banach's contraction mapping theorem or successive approximations. To prove that $\|(KU)(x)\| = o(1)$, we observe that if Assumption VI (a) or (c) holds (so that we take $x_{jk} = \infty$), then from (61), $|((KU)(x))_{jk}| \leqq C\int_x^{\infty}\|B(t)\|\,dt\,\|U\|_M = o(1)$. If Assumption VI (b) holds (so that we take $x_{jk} = x_0$), then from (61),

$$|((KU)(x))_{jk}| \leqq C\left\{\int_{x_0}^{x_1}\left[\exp\int_{x_1}^{x}\operatorname{Re}(\lambda_j(s)-\lambda_k(s))\,ds\right]\|B(t)\|\,dt\right.$$
$$\left.+\int_{x_1}^{x}\|B(t)\|\,dt\right\}\|V\|_M,$$

where $x\geqq x_1\geqq x_0$. From this inequality it is seen that $|((KU)(x))_{jk}| =$

$o(1)$ also when Assumption VI (b) holds. (64) follows from (57) and the fact that $\|(KU)(x)\| = o(1)$. This completes the proof of Lemma 20.

THEOREM 1. *Under Assumptions* I–VI, *the equation* $Ly = i\sigma y$, $x \geqq r$, $\sigma \neq 0$, *has three linearly independent solutions* y_k, $k = 1,2,3$, *of the form*

$$(65) \qquad y_k = [1 + o(1)]a_0^{-1/3}(x)\exp\int_{x_0}^{x} \lambda_k(t)\,dt, \qquad x \geqq x_0,$$

where the $\lambda_k(t)$ *are given by equation* (22).

Proof. By (56) and (64), there is a solution matrix $Z(x)$ for (48) of the form

$$(66) \qquad Z(x) = [E + o(1)]Z_0(x_0, x), \qquad x \geqq x_0.$$

If x_0 is sufficiently large, $\det[E + o(1)] \neq 0$ for $x \geqq x_0$ and therefore $Z(x)$ is a fundamental matrix for (48). By (66) and (27) a solution matrix for (25) is given by

$$(67) \qquad W(x) = T_0(x)[E + T_2(x)][E + o(1)]Z_0(x_0, x), \qquad x \geqq x_0.$$

Since $[E + T_2(x)]^{-1}$ exists by Lemma 16 and $T_0^{-1}(x)$ exists by (31), $W(x)$ is a fundamental matrix. By Lemma 14 and the fact that $\lim_{x\to\infty} \alpha(x)/\lambda(x) = 0$, we see that

$$(68) \qquad W(x) = T_0(x)[E + o(1)]Z_0(x_0, x), \qquad x \geqq x_0.$$

Let $y_k(x) = w_{1k}(x)$, $k = 1,2,3$, where $w_{1k}(x)$ is the element in the first row and kth column of $W(x)$. Then, by the equivalence of (24) and (25), y_k is a solution of (24), and by (68) and (30),

$$(69) \qquad y_k = [1 + o(1)]\exp\int_{x_0}^{x} [\lambda_k(t) + \lambda_k^{(1)}(t)]\,dt, \qquad x \geqq a.$$

From the equations $y_k = w_{k1}$, $y' = w_{k2}$, $y_k'' = -(ib_2)^{-1}mw_{k2} + (ib_2)^{-1}w_3$, we see that $W(y_1, y_2, y_3)(x) = \det W(x) \neq 0$, $x \geqq x_0$, where $W(y_1, y_2, y_3)$ is the Wronskian of y_1, y_2, y_3. Hence, y_1, y_2, y_3 are linearly independent for $x \geqq x_0$. By (43), (41), (38), (39), (49), Lemma 5, Lemma 1 and Assumption IV we see that

$$(70) \quad \exp\int_{x_0}^{x} \lambda_k^{(1)}(t)\,dt = C_k\,|a_0^{-1/3}(x)|\,[1 + o(1)], \quad x \geqq x_0, \quad C_k \neq 0.$$

(65) now follows from (69), (70) and the fact that $|a_0| \to \infty$, so that $a_0(x) > 0$ or $a_0(x) < 0$ for $x \geqq x_0$. This finishes the proof of Theorem 1.

3. Asymptotic formulas for the $\lambda_k(x)$. In this section we take the coefficients of the operator L of equation (1) to be the following on the interval $[1, \infty)$:

$$b_2(x) \equiv 1, \tag{71}$$

$$b_1(x) = ax^{\alpha}, \qquad \alpha < 2\gamma/3, \tag{72}$$

$$a_1(x) = bx^{\gamma/3}, \tag{73}$$

$$a_0(x) = cx^{\gamma}, \qquad \gamma > 0, \quad c \neq 0. \tag{74}$$

LEMMA 21. *If $b/c^{1/3} \neq 3/2^{2/3}$, then the coefficients of L given by* (71)–(74) *satisfy Assumptions* I–V *with*

$$d = b/c^{1/3}. \tag{75}$$

The proof is straightforward. We note that it is required in (74) that $\gamma > 0$ and $c \neq 0$ in order that $a_0(x) \to +\infty$ or $a_0(x) \to -\infty$ (Assumption I). The exponent $\gamma/3$ occurs in (73) in order that $\lim_{x\to\infty} a_1/a_0^{1/3} = d$ (Assumption II) with the possibility that $d \neq 0$. The inequality $\alpha < 2\gamma/3$ is required in (72) in order that $b_1/a_0^{2/3} = o(1)$ (Assumption II).

LEMMA 22. *If $b/c^{1/3} < 3/2^{2/3}$, the coefficients of L given by* (71)–(74) *satisfy Assumptions* I–VI.

Proof. Since $d = b/c^{1/3} < 3/2^{2/3}$, $h_0(\eta) = 0$ has one real negative root and two complex conjugate nonreal roots. Suppose $\eta_{02} = p + iq$, $\eta_{03} = p - iq$, $q > 0$. Then from (23) one sees that Assumption VI is satisfied; in fact, (a) or (b) is true for each pair j, k. This proves the lemma.

If $d > 3/2^{2/3}$, then $h_0(\eta) = 0$ has three real roots. In this case in order to check Assumption VI it is necessary to have asymptotic formulas for the $\lambda_k(x)$ which are more precise than (23). We obtain these by use of (20).

LEMMA 23. *Suppose the coefficients of L are given by* (71)–(74) *and that $b/c^{1/3} \neq 3/2^{2/3}$. Then the roots $\lambda_k(x)$ of* (7) *are given by*

$$
\begin{aligned}
\lambda_k(x) = ia_0^{1/3}\Big\{&\eta_{0k} + [\eta_{0k} - v_{11}d](6c)^{-1}(iD) \\
&+ ac^{-2/3}v_{10}x^{-\nu} + [-(\eta_{0k} - v_{11}d) + v_{22}d^2](6c)^{-2}(iD)^2 \\
&- ac^{-2/3}[v_{10} + v_{21}d](6c)^{-1}(iD)x^{-\nu} \\
&+ (ac^{-2/3})^2 v_{20}x^{-2\nu} \\
&+ [(5/3)(\eta_{0k} - v_{11}d) - 3v_{22}d^2 - v_{33}d^3](6c)^{-3}(iD)^3 \\
&+ O(D^2x^{-\nu}) + O(Dx^{-2\nu}) + (ac^{-2/3})^3 v_{30}x^{-3\nu} \\
&+ \sum_{j=4}^{n+2}\sum_{s=1}^{j} O(D^s x^{-(j-s)\nu}) + O(Dx^{-(n+2)\nu}) \\
&+ \sum_{j=4}^{n} v_{j0}(ac^{-2/3})^j x^{-j\nu} \\
&+ w_{n+1,0}(x)(ac^{-2/3})^{n+1}x^{-(n+1)\nu} \\
&+ w_{n+2,0}(x)(ac^{-2/3})^{n+2}x^{-(n+2)\nu}\Big\},
\end{aligned} \tag{76}
$$

where n is an integer, $n \geqq 4$, *the* v_{js} *are constants which depend on* η_{0k} *and are real when* η_{0k} *is real,* $w_{n+1,0}(x)$ *and* $w_{n+2,0}(x)$ *are complex functions which are bounded as* $x \to \infty$,

$$\nu = 2\gamma/3 - \alpha > 0, \tag{77}$$

$$
\begin{aligned}
D &= a\alpha x^{-(\nu+1+\gamma/3)} - 2\sigma x^{-\gamma} \\
&= o(1) \quad as \quad x \to \infty.
\end{aligned} \tag{78}
$$

If η_{0k} *is real,*

$$
\begin{aligned}
\operatorname{Re}\lambda_k(x) = a_0^{1/3}\Big\{&[v_{11}d - \eta_{0k}](6c)^{-1}D \\
&+ ac^{-2/3}[v_{10} + dv_{21}](6c)^{-1}Dx^{-\nu} \\
&- [(5/3)(v_{11}d - \eta_{0k}) + d^2(3v_{22} + dv_{33})](6c)^{-3}D^3 \\
&+ O(D^2x^{-\nu}) + O(Dx^{-2\nu}) \\
&+ \sum_{j=4}^{n+2}\sum_{s=1}^{j} O(D^s x^{-(j-s)\nu}) + O(Dx^{-(n+2)\nu}) \\
&+ O(x^{-(n+1)\nu}) + O(x^{-(n+2)\nu})\Big\}.
\end{aligned} \tag{79}
$$

It is true that

$$v_{11} = \eta_{0k}^2[h_0'(\eta_{0k})]^{-1}, \tag{80}$$

$$v_{11}d - \eta_{0k} = 3[h_0'(\eta_{0k})]^{-1}, \tag{81}$$

$$v_{10} = \eta_{0k}[h_0'(\eta_{0k})]^{-1}, \tag{82}$$

(83) $v_{21} = \eta_{0k}^2[3h_0'(\eta_{0k}) - \eta_{0k}h_0''(\eta_{0k})][h_0'(\eta_{0k})]^{-3}$,

(84) $v_{22} = 2^{-1}\eta_{0k}^3[4h_0'(\eta_{0k}) - \eta_{0k}h_0''(\eta_{0k})][h_0'(\eta_{0k})]^{-3}$,

(85) $v_{33} = 2^{-1}\eta_{0k}^4\{[3h_0'(\eta_{0k}) - \eta_{0k}h_0''(\eta_{0k})]^2 + [h_0'(\eta_{0k})]^2\}[h_0'(\eta_{0k})]^{-2}$.

Proof. From (5) and (71)–(74) we see that

(86) $$\tau(x) = a_0^{1/3}(x)t(x),$$

(87) $$t(x) = [1 + (2c)^{-1}(iD)]^{1/3}.$$

As $x \to \infty$,

(88) $$t(x) = 1 + (6c)^{-1}(iD) - (6c)^{-2}(iD)^2 + (5/3)(6c)^{-3}(iD)^3 + O(D^4).$$

The functions $f_1(x)$ and $f_2(x)$ of (16)–(18) are given for $x \to \infty$ by

(89) $$f_1(x) = d[-(6c)^{-1}(iD) + 2(6c)^{-2}(iD)^2 - (14/3)(6c)^{-3}(iD)^3 + O(D^4)],$$

(90) $$\begin{aligned} f_2(x) = ac^{-2/3}x^{-\nu}[1 &- 2(6c)^{-1}(iD) \\ &+ 5(6c)^{-2}(iD)^2 - (40/3)(6c)^{-3}(iD)^3 + O(D^4)]. \end{aligned}$$

Now, $h^{-1} = h_0^{-1}[1 - (\eta/h_0)(\eta f_1 + f_2)]^{-1}$. Let n be a positive integer. For $\eta \in C_k$ and for $x \geqq x_0$,

(91) $$\begin{aligned} h^{-1} &= h_0^{-1}\Big\{1 + \sum_{j=1}^{n} (\eta/h_0)^j(\eta f_1 + f_2)^j + (\eta/h_0)^{n+1}(\eta f_1 + f_2)^{n+1} \\ &\qquad \times [1 - (\eta/h_0)(\eta f_1 + f_2)]^{-1}\Big\} \\ &= h_0^{-1} + \sum_{j=1}^{n}\left(\sum_{s=0}^{j} a_{js}(\eta)f_1^s f_2^{j-s}\right) \\ &\quad + \sum_{s=0}^{n+1} a_{n+1,s}(\eta)f_1^s f_2^{n+1-s}[1 - (\eta/h_0)(\eta f_1 + f_2)]^{-1}. \end{aligned}$$

Hence,

(92) $$\begin{aligned} \eta h_\eta h^{-1} &= \eta h_0' h_0^{-1} + \sum_{j=1}^{n+1}\sum_{s=0}^{j} b_{js}(\eta)f_1^s f_2^{j-s} \\ &\quad + \left[\sum_{s=0}^{n+1} c_{n+1,s}(\eta, x)f_1^s f_2^{n+1-s} + \sum_{s=0}^{n+2} c_{n+2,s}(\eta, x)f_1^s f_2^{n+2-s}\right] \\ &\quad \times [1 - (\eta/h_0)(\eta f_1 + f_2)]^{-1}. \end{aligned}$$

Substituting (92) into (20),

$$\eta_k(x) = \eta_{0k} + \sum_{j=1}^{n+1} \sum_{s=0}^{j} v_{js} f_1^s f_2^{j-s} + \sum_{s=0}^{n+1} w_{n+1,s}(x) f_1^s f_2^{n+1-s} + \sum_{s=0}^{n+2} w_{n+2,s}(x) f_1^s f_2^{n+2-s}, \tag{93}$$

where the v_{js} are constants which are real if η_{0k} is real, and the functions $w_{n+1,s}(x)$ and $w_{n+2,s}(x)$ are bounded as $x \to \underset{r}{+}\infty$. If we substitute (93) into (22), we obtain for $x \geqq x_0$,

$$\begin{aligned} \lambda_k(x) \doteq ia_0^{1/3}\Big\{ & t\eta_{0k} + \sum_{j=1}^{3} \sum_{s=0}^{j} v_{js} t f_1^s f_2^{j-s} \\ & + \sum_{j=4}^{n+1} \Big[v_{j0} t f_2^j + \sum_{s=1}^{j} v_{js} t f_1^s f_2^{j-s} \Big] \\ & + w_{n+1,0} t f_2^{n+1} + \sum_{s=1}^{n+1} w_{n+1,s} t f_1^s f_2^{n+1-s} \\ & + w_{n+2,0} t f_2^{n+2} + \sum_{s=1}^{n+2} w_{n+2,s} t f_1^s f_2^{n+2-s} \Big\}. \end{aligned} \tag{94}$$

We now use (88), (89), (90) to calculate asymptotic expansions for each of the terms $t\eta_{0k}$, $tf_1^s f_2^{j-s}$. We obtain

$$\begin{aligned} t\eta_{0k} &= \eta_{0k} + \eta_{0k}(6c)^{-1}(iD) - \eta_{0k}(6c)^{-2}(iD)^2 + (5/3)\eta_{0k}(6c)^{-3}(iD)^3 \\ &\quad + O(D^4), \\ tf_1 &= d[-(6c)^{-1}(iD) + (6c)^{-2}(iD)^2 - (5/3)(6c)^{-3}(iD)^3 + O(D^4)], \\ tf_2 &= ac^{-2/3}x^{-\nu}[1 - (6c)^{-1}(iD) + O(D^2)], \quad \text{etc.} \end{aligned}$$

Substituting into (94), we obtain (76). (79) follows immediately from (76). From the way in which (93) was derived, we see that $v_{11} = (2\pi i)^{-1} \int_{C_k} [\eta^3 h_0' - 2\eta^2 h_0] h_0^{-2}\, d\eta$. Hence,

$$\begin{aligned} v_{11} &= (2\pi i)^{-1} \int_{C_k} [\eta^2 h_0^{-1} - (d/d\eta)(\eta^3 h_0^{-1})]\, d\eta \\ &= (2\pi i)^{-1} \int_{C_k} \eta^2 h_0^{-1}\, d\eta = \eta_{0k}[h_0'(\eta_{0k})]^{-1}. \end{aligned}$$

This proves (80). (82)–(85) are proved similarly. (81) follows from (80) and the fact that $d = (\eta_{0k}^3 + 1)\eta_{0k}^{-2}$. This proves Lemma 23.

Let

$$\mu = \min\{\nu + 1 + \gamma/3, \gamma\} \quad \text{if} \quad a\alpha \neq 0$$
$$= \gamma \quad \text{if} \quad a\alpha = 0. \tag{95}$$

Then, as $x \to \infty$,

$$D = O(x^{-\mu}). \tag{96}$$

In the following we shall consider three cases. Case 1 is the case that $\nu = 2\mu$, which occurs if $\alpha = -4\gamma/3$. Case 2 is the case $\nu > 2\mu$, which occurs if $\alpha < -4\gamma/3$. Case 3 is the case $\nu < 2\mu$, which occurs if $-4\gamma/3 < \alpha < 2\gamma/3$.

LEMMA 24. *Suppose the coefficients of L are given by* (71)–(74) *and that $b/c^{1/3} \neq 3/2^{2/3}$. If η_{0k} is real,* $\operatorname{Re}\lambda_k(x)$ *has the following asymptotic expansions*:

Case 1. $\nu = 2\mu$ (*i.e.*, $\alpha = -4\gamma/3$). *Then,*

$$\begin{aligned} \operatorname{Re}\lambda_k(x) = a_0^{1/3}\{&[h_0'(\eta_{0k})]^{-1}(2c)^{-1}D \\ &+ ac^{-2/3}[v_{10} + dv_{21}](6c)^{-1}Dx^{-2\mu} \\ &- [5(h_0'(\eta_{0k}))^{-1} + d^2(3v_{22} + dv_{33})](6c)^{-3}D^3 \\ &+ O(x^{-4\mu})\}. \end{aligned} \tag{97}$$

Case 2. $\nu > 2\mu$ (i.e., $\alpha < -4\gamma/3$). *Then,*

$$\begin{aligned} \operatorname{Re}\lambda_k(x) = a_0^{1/3}\{&[h_0'(\eta_{0k})]^{-1}(2c)^{-1}D \\ &- [5(h_0'(\eta_{0k}))^{-1} + d^2(3v_{22} + dv_{33})](6c)^{-3}D^3 + O(x^{-3\mu-\epsilon})\}, \end{aligned} \tag{98}$$

where $\epsilon > 0$.

Case 3. $\nu < 2\mu$ (i.e., $-4\gamma/3 < \alpha < 2\gamma/3$). *Then,*

$$\begin{aligned} \operatorname{Re}\lambda_k(x) = a_0^{1/3}\{&[h_0'(\eta_{0k})]^{-1}(2c)^{-1}D \\ &+ ac^{-2/3}[v_{10} + dv_{21}](6c)^{-1}Dx^{-\nu} \\ &+ O(x^{-\mu-\nu-\epsilon})\}, \end{aligned} \tag{99}$$

where $\epsilon > 0$.

Proof. (97) and (98) follow directly from (79). If we choose n so large that $n\nu > \mu$, then we also see that (99) follows from (79). This proves Lemma 24.

LEMMA 25. *If* $b/c^{1/3} > 3/2^{2/3}$, $b/c^{1/3} \neq 3/2^{1/3}$, *and* $\sigma \neq a\alpha/2$, *then the coefficients of* L *given by* (71)–(74) *satisfy Assumptions* I–VI.

Proof. Since $d = b/c^{1/3} > 3/2^{2/3}$, $h_0(\eta) = 0$ has three real roots. Because $d \neq 3/2^{1/3}$, $h_0'(\eta_{01})$, $h_0'(\eta_{02})$, $h_0'(\eta_{03})$ are all distinct. From (78) and (95) we see that $D = C_1 x^{-\mu}[1 + o(1)]$, where $C_1 \neq 0$ because $\sigma \neq a\alpha/2$. By Lemma 24,

$$\begin{aligned} \operatorname{Re}[\lambda_j(x) - \lambda_k(x)] = C_1(2c)^{-1}a_0^{1/3}\{[h_0'(\eta_{0j})]^{-1} - [h_0'(\eta_{0k})]^{-1}\} \\ \times x^{-\mu}[1 + o(1)]. \end{aligned} \tag{100}$$

From (100) and (74) it follows that Assumption VI is satisfied. This proves Lemma 25.

LEMMA 26. *Suppose the coefficients of* L *are given by* (71)–(74) *and that* $b/c^{1/3} = 3/2^{1/3}$. *Then the roots of* $h_0(\eta) = 0$ *are* $\eta_{01} = 2^{-1/3}(1 - 3^{1/2})$, $\eta_{02} = 2^{-1/3}$, $\eta_{03} = 2^{-1/3}(1 + 3^{1/2})$, *and*

$$h_0'(\eta_{01}) = h_0'(\eta_{03}) \neq h_0'(\eta_{02}), \tag{101}$$

$$v_{10}(\eta_{01}) + dv_{21}(\eta_{01}) = 3^{-1}2^{-2/3}(-2 + 3^{1/2}), \tag{102}$$

$$v_{10}(\eta_{03}) + dv_{21}(\eta_{03}) = 3^{-1}2^{-2/3}(-2 - 3^{1/2}), \tag{103}$$

$$3v_{22}(\eta_{01}) + dv_{33}(\eta_{01}) = 3^{-1}2^{-2/3}[250 - (143)3^{1/2}], \tag{104}$$

$$3v_{22}(\eta_{03}) + dv_{33}(\eta_{03}) = 3^{-1}2^{-2/3}[250 + (143)3^{1/2}]. \tag{105}$$

The proof follows immediately from (80)–(85) and the fact that $h_0(\eta) = \eta^3 - (3/2^{1/3})\eta^2 + 1$.

LEMMA 27. *Suppose that* $b/c^{1/3} = 3/2^{1/3}$, $\alpha < -4\gamma/3$. *Then, the coefficients of* L *given by* (71)–(74) *satisfy Assumptions* I–VI.

Proof. Since $\alpha < -4\gamma/3$, $\nu + 1 + \gamma/3 > \gamma$. By (95), $\mu = \gamma$. By (78), $D = -2\sigma x^{-\gamma}(1 + o(1))$. From (101) and (98) it follows that $\operatorname{Re}[\lambda_2(x) - \lambda_1(x)]$ and $\operatorname{Re}[\lambda_2(x) - \lambda_3(x)]$ satisfy (a), (b) or (c) of Assumption VI. From (98), (101), (104), (105),

$$\operatorname{Re}[\lambda_3(x) - \lambda_1(x)] = C_1 x^{-8\gamma/3}(1 + o(1)),$$

where $C_1 \neq 0$. Thus, $\mathrm{Re}[\lambda_3(x) - \lambda_1(x)]$ also satisfies (a), (b) or (c) of Assumption VI. This proves Lemma 27.

LEMMA 28. *Suppose that* $b/c^{1/3} = 3/2^{1/3}$, $-4\gamma/3 < \alpha < 2\gamma/3$, $\sigma \neq a\alpha/2$, $a \neq 0$. *Then, the coefficients of* L *given by* (71)–(74) *satisfy Assumptions* I–VI.

Proof. It follows from (78) that $D = C_1 x^{-\mu}(1 + o(1))$, where $C_1 \neq 0$ because $\sigma \neq a\alpha/2$. By (101) and (99), $\mathrm{Re}[\lambda_2(x) - \lambda_1(x)]$ and $\mathrm{Re}[\lambda_2(x) - \lambda_3(x)]$ satisfy (a), (b) or (c) of Assumption VI. From (99) and (101)–(103), $\mathrm{Re}[\lambda_3(x) - \lambda_1(x)] = C_2 x^{-\mu-\nu+\gamma/3}(1 + o(1))$, where $C_2 \neq 0$ because $a \neq 0$. Hence, $\mathrm{Re}[\lambda_3(x) - \lambda_1(x)]$ satisfies (a), (b) or (c) of Assumption VI. This proves Lemma 28.

LEMMA 29. *Suppose that* $b/c^{1/3} = 3/2^{1/3}$, $\alpha = -4\gamma/3$, $\sigma^2 \neq -2^{2/3}ac^{4/3}/143$. *Then, the coefficients of* L *given by* (71)–(74) *satisfy Assumptions* I–VI.

Proof. Since $\alpha = -4\gamma/3$, $\mu = \gamma$. Hence, $D = -2\sigma x^{-\gamma}(1 + o(1))$ by (78). From (101) and (97) it follows that $\mathrm{Re}[\lambda_2(x) - \lambda_1(x)]$ and $\mathrm{Re}[\lambda_2(x) - \lambda_3(x)]$ satisfy (a), (b) or (c) of Assumption VI. From (97) and (101)–(105), $\mathrm{Re}[\lambda_3(x) - \lambda_1(x)] = C_1 x^{-8\gamma/3}(1 + o(1))$, where $C_1 \neq 0$ because $\sigma^2 \neq -2^{2/3}ac^{4/3}/143$. Hence, $\mathrm{Re}[\lambda_3(x) - \lambda_1(x)]$ satisfies (a), (b) or (c) of Assumption VI. This proves Lemma 29.

LEMMA 30. *Suppose the coefficients of* L *are given by* (71)–(74) *and that* $b/c^{1/3} \neq 3/2^{2/3}$. *If* η_{0k} *is real,* $\mathrm{Re}\,\lambda_k(x)$ *has the following asymptotic expansions:*

Case A. *Suppose* $a = 0$. *Then,*

$$\mathrm{Re}\,\lambda_k(x) = -\sigma[h_0'(\eta_{0k})]^{-1}c^{-2/3}x^{-2\gamma/3}(1 + o(1)). \tag{106}$$

Case B. *Suppose* $a \neq 0$.
(i) *Suppose* $1 < 2\gamma/3$.
(a) *If* $1 < \alpha < 2\gamma/3$, *then*

$$\mathrm{Re}\,\lambda_k(x) = a\alpha c^{-2/3}[2h_0'(\eta_{0k})]^{-1}x^{\alpha-1-2\gamma/3}(1 + o(1)). \tag{107}$$

(b) *If* $\alpha = 1$ *and* $\sigma \neq a/2$,

$$\mathrm{Re}\,\lambda_k(x) = (a - 2\sigma)[2h_0'(\eta_{0k})]^{-1}c^{-2/3}x^{-2\gamma/3}(1 + o(1)). \tag{108}$$

(c) *If* $\alpha < 1$, (106) *is valid.*
(ii) *If* $\alpha < 2\gamma/3 \leqq 1$, (106) *is valid.*

The proof follows directly from Lemma 24 with calculation of μ and D in the various cases.

4. The deficiency index of the operator *L*. In the following, L_2 will denote the space $L_2[1,\infty)$, i.e., the space of complex-valued functions on $[1,\infty)$ which have Lebesgue square integrable absolute values.

LEMMA 31. *Suppose the coefficients of L are given by* (71)–(74) *and that* $b/c^{1/3} < 3/2^{2/3}$, *so that* $\eta_{0k} = u_k + iv_k$, *where* $v_2 > 0$ *and* $v_3 < 0$. *Then the function* $f_k(x) = a_0^{-1/3}(x)\exp\int_{x_0}^{x} \lambda_k(t)\,dt$, $x \geqq x_0$, *has the following properties*:

(i) *If* $k = 2$ *and* $c > 0$ *or if* $k = 3$ *and* $c < 0$, *then* $f_k \in L_2$ *for* $\sigma > 0$ *and for* $\sigma < 0$.

(ii) *If* $k = 2$ *and* $c < 0$ *or if* $k = 3$ *and* $c > 0$, *then* $f_k \notin L_2$ *for* $\sigma > 0$ *and for* $\sigma < 0$.

Proof. We shall give an intuitive proof which can be made precise as in Naimark [**9**, §23]. We have by (23) that

$$\begin{aligned} |f_k(x)| &\approx |c|^{-1/3}x^{-\gamma/3}\exp\left[-v_k c^{1/3}\int_{x_0}^{x} t^{\gamma/3}\,dt\right] \\ &= |c|^{-1/3}x^{-\gamma/3}\exp\left[-v_k c^{1/3}(\gamma/3+1)^{-1}(x^{\gamma/3+1} - x_0^{\gamma/3+1})\right] \\ &\to +\infty \quad \text{if} \quad v_k c^{1/3} < 0. \end{aligned}$$

This proves (ii). Also,

$$\begin{aligned} |f_k(x)|^2 &\approx |c|^{-2/3}x^{-2\gamma/3}\exp\left[-2v_k c^{1/3}\int_{x_0}^{x} t^{\gamma/3}\,dt\right] \\ &\leqq |c|^{-2/3}x^{\gamma/3}\exp\left[-2v_k c^{1/3}\int_{x_0}^{x} t^{\gamma/3}\,dt\right] \\ &= (-2v_k c)^{-1}(d/dx)\exp\left[-2v_k c^{1/3}\int_{x_0}^{x} t^{\gamma/3}\,dt\right]. \end{aligned}$$

This proves (i).

LEMMA 32. *Suppose the coefficients of L are given by* (71)–(74) *and that* $b/c^{1/3} \neq 3/2^{2/3}$. *If* η_{0k} *is real, the function* $f(x) = a_0^{-1/3}(x)$

$\exp \int_{x_0}^{x} \lambda_k(t)\,dt$, $x \geqq x_0$, *has the following properties*:

(I) *If* $2\gamma/3 > 1$ *and* $\sigma \neq a/2$, *then* $f \in L_2$ *for* $\sigma > 0$ *and for* $\sigma < 0$.

(II) *If* $2\gamma/3 \leqq 1$, *then* $f \in L_2$ *for* $\sigma/h_0'(\eta_{0k}) > 0$, *and* $f \notin L_2$ *for* $\sigma/h_0'(\eta_{0k}) < 0$.

Proof. Case A. Suppose $a = 0$. By (106),

$$\begin{aligned} |f(x)|^2 &\approx c^{-2/3}x^{-2\gamma/3}\exp\left\{-2\sigma c^{-2/3}[h_0'(\eta_{0k})]^{-1}\int_{x_0}^{x} t^{-2\gamma/3}\,dt\right\} \\ &= (-2\sigma)^{-1}h_0'(\eta_{0k})(d/dx)\exp\left\{-2\sigma c^{-2/3}[h_0'(\eta_{0k})]^{-1}\int_{x_0}^{x} t^{-2\gamma/3}\,dt\right\}. \end{aligned}$$

From this last expression we see that (I) and (II) are true for Case A.

Case B. Suppose $a \neq 0$. If $1 < \alpha < 2\gamma/3$, then by (107),

$$\begin{aligned} |f(x)|^2 &\approx c^{-2/3}x^{-2\gamma/3}\exp\left\{a\alpha c^{-2/3}[h_0'(\eta_{0k})]^{-1}\int_{x_0}^{x} t^{\alpha-1-2\gamma/3}\,dt\right\} \\ &\leqq c^{-2/3}x^{\alpha-1-2\gamma/3}\exp\left\{a\alpha c^{-2/3}[h_0'(\eta_{0k})]^{-1}\int_{x_0}^{x} t^{\alpha-1-2\gamma/3}\,dt\right\} \\ &= (a\alpha)^{-1}h_0'(\eta_{0k})(d/dx)\exp\left\{a\alpha c^{-2/3}[h_0'(\eta_{0k})]^{-1}\int_{x_0}^{x} t^{\alpha-1-2\gamma/3}\,dt\right\}. \end{aligned}$$

Since $\int_{x_0}^{x} t^{\alpha-1-2\gamma/3}\,dt$ converges, we see that (I) is true if $1 < \alpha < 2\gamma/3$. If $\alpha = 1 < 2\gamma/3$ and $\sigma \neq a/2$, then by (108),

$$\begin{aligned} |f(x)|^2 &\approx c^{-2/3}x^{-2\gamma/3}\exp\left\{(a-2\sigma)c^{-2/3}[h_0'(\eta_{0k})]^{-1}\int_{x_0}^{x} t^{-2\gamma/3}\,dt\right\} \\ &= (a-2\sigma)^{-1}h_0'(\eta_{0k}) \\ &\qquad \times (d/dx)\exp\left\{(a-2\sigma)c^{-2/3}[h_0'(\eta_{0k})]^{-1}\int_{x_0}^{x} t^{-2\gamma/3}\,dt\right\}. \end{aligned}$$

Since $\int_{x_0}^{x} t^{-2\gamma/3}\,dt$ converges, we see that (I) is true for $\alpha = 1 < 2\gamma/3$ and $\sigma \neq a/2$. If $\alpha < 1 < 2\gamma/3$ or if $\alpha < 2\gamma/3 \leqq 1$, then by Lemma 30, (106) is valid and therefore (I) and (II) follow as in Case A. This proves Lemma 32.

Let n_+ denote the dimension of the space of solutions of $Ly = i\sigma y$, $x \geqq r$, which are in $L_2[r, \infty)$ for $\sigma > 0$. It is known that n_+ is independent of σ. Let n_- denote the same number for $\sigma < 0$. We shall call n_+ and

n_- the deficiency numbers of L, and we shall call the pair (n_+, n_-) the deficiency index.

THEOREM 2. *Suppose that the coefficients of L are given by* (71)–(74) *and that* $b/c^{1/3} < 3/2^{2/3}$. *If* $2\gamma/3 > 1$, $n_+ = n_- = 2$. *If* $2\gamma/3 \leqq 1$, $n_+ = 2$, $n_- = 1$.

Proof. By Lemma 22, the coefficients of L satisfy Assumptions I–VI. By Theorem 1, $Ly = i\sigma y$, $x \geqq 1$, $\sigma \neq 0$, has three linearly independent solutions y_k given by (65). By Lemma 31, for $c > 0$, $y_2 \in L_2$ and $y_3 \notin L_2$ for $\sigma > 0$ and for $\sigma < 0$; for $c < 0$, $y_2 \notin L_2$ and $y_3 \in L_2$ for $\sigma > 0$ and for $\sigma < 0$. By Lemma 32, if $2\gamma/3 > 1$, $y_1 \in L_2$ for $\sigma > 0$ and for $\sigma < 0$, $\sigma \neq a/2$; if $2\gamma/3 \leqq 1$, $y_1 \in L_2$ for $\sigma > 0$, and $y_2 \notin L_2$ for $\sigma < 0$, because $h_0'(\eta_{01}) > 0$. It follows that if $2\gamma/3 > 1$, then $n_+ = n_- = 2$, and if $2\gamma/3 \leqq 1$, then $n_+ = 2$. It also follows that if $2\gamma/3 \leqq 1$, then $n_- = 1$, provided we can show that for $c > 0$ and $\sigma < 0$ no nontrivial linear combination of y_1 and y_3 is in L_2, and for $c < 0$ and $\sigma < 0$ no nontrivial linear combination of y_1 and y_2 is in L_2. We deal with the case $c > 0$, $\sigma < 0$; the case $c < 0$ and $\sigma < 0$ is similar. It is sufficient to show that $y_1 + By_3 \notin L_2$ if $B \neq 0$. By Theorem 1, (23), and Lemma 30,

$$\begin{aligned} |y_1/y_3| &= [1 + o(1)] \exp \int_{x_0}^{x} [\operatorname{Re} \lambda_1(t) - \operatorname{Re} \lambda_3(t)]\, dt \\ &= [1 + o(1)] \exp c^{1/3} v_3 \int_{x_0}^{x} t^{\gamma/3} [1 + o(1)]\, dt \to 0 \quad \text{as} \quad x \to +\infty. \end{aligned}$$

Hence, for $x \geqq x_1$, $|y_1/y_3 + B|^2 \geqq K$, where K is a constant. Thus

$$\int_{x_1}^{\infty} |y_1 + By_3|^2\, dx = \int_{x_1}^{\infty} |y_3|^2 |y_1/y_3 + B|^2\, dx \geqq K \int_{x_1}^{\infty} |y_3|^2\, dx.$$

It follows that $y_1 + By_3 \notin L_2$. This completes the proof of Theorem 2.

THEOREM 3. *Suppose that the coefficients of L are given by* (71)–(74) *and that* $b/c^{1/3} > 3/2^{2/3}$.

Case A. *Suppose* $b/c^{1/3} \neq 3/2^{1/3}$. *If* $2\gamma/3 > 1$, $n_+ = n_- = 3$. *If* $2\gamma/3 \leqq 1$, $n_+ = 2$, $n_- = 1$.

Case B. *Suppose* $b/c^{1/3} = 3/2^{1/3}$ *and* $\alpha \leqq -4\gamma/3$. *If* $2\gamma/3 > 1$, $n_+ = n_- = 3$. *If* $2\gamma/3 \leqq 1/4$, $n_+ = 2$, $n_- = 1$.

Case C. *Suppose* $b/c^{1/3} = 3/2^{1/3}$, $-4\gamma/3 < \alpha < 2\gamma/3$, $a \neq 0$. *If* $2\gamma/3 > 1$, $n_+ = n_- = 3$. *If* $4\gamma/3 - 1 \leqq \alpha < 2\gamma/3 < 1$, *then* $n_+ = 2$, $n_- = 1$.

Proof. By Lemmas 25–29, the coefficients of L satisfy Assumptions I–VI in all three cases, provided $\sigma \neq a\alpha/2$ and $\sigma^2 \neq -2^{2/3}ac^{4/3}/143$. Hence, if we avoid these values of σ, $Ly = i\sigma y$, $x \geqq 1$, $\sigma \neq 0$, has three linearly independent solutions y_k given by (65). By Lemma 32 we have the following: (I) If $2\gamma/3 > 1$ and $\sigma \neq a/2$, then $y_1, y_2, y_3 \in L_2$ for $\sigma > 0$ and for $\sigma < 0$; (II) if $2\gamma/3 \leqq 1$, then for $\sigma > 0$, $y_1, y_3 \in L_2$ and $y_2 \notin L_2$, while for $\sigma < 0$, $y_2 \in L_2$ and $y_1, y_3 \notin L_2$. By (I) we see that if $2\gamma/3 > 1$, then $n_+ = n_- = 3$ in all three cases. If $2\gamma/3 \leqq 1$, then $n_+ = 2$ and $n_- = 1$, provided we can show that no non-trivial linear combination of y_1 and y_3 is in L_2. Using (106), this can be proved for Case A as in the proof of Theorem 2. In Cases B and C it is necessary to use (97)–(99). The assumptions in Cases B and C enable one to do this as in the proof of Theorem 2. This completes the proof of Theorem 3.

THEOREM 4. *Suppose that the coefficients of L are given by* (71–74) (*without the requirements that* $\alpha < 2\gamma/3$, $\gamma > 0$). *Then the deficiency index of L is as follows for the indicated values of the parameters* γ, α:

I. $\gamma > 3/2$, $\alpha < 2\gamma/3$: (2,2) *if* $b/c^{1/3} < 3/2^{2/3}$; (3,3) *if* $b/c^{1/3} > 3/2^{2/3}$, $b/c^{1/3} \neq 3/2^{1/3}$.

II. $0 < \gamma \leqq 3/2$, $\alpha < 2\gamma/3$: (2,1) *if* $b/c^{1/3} \neq 3/2^{2/3}$ *and* $b/c^{1/3} \neq 3/2^{1/3}$.

III. $\gamma \leqq 0$, $\alpha \leqq 0$: (2,1).

IV. $0 < \alpha \leqq 1$, $\alpha > 2\gamma/3$: (2,1).

V. $1 < \alpha$, $\alpha > 2\gamma/3$: (3,3) *if* $a > 0$; (2,2) *if* $a < 0$.

Proof. The statements for regions I and II follow from Theorems 2 and 3. III follows from the fact that $n_+ + n_- = 3$ by Dunford and Schwartz [**4**, XIII. 10. E.II(5)] and from the fact that $2 \leqq n_+$ and $1 \leqq n_-$ by Everitt [**5**] or Kogan and Rofe-Beketov [**8**]. IV and V follow from Unsworth [**12**]. This proves Theorem 4.

REMARK 1. Note that $\alpha = 2\gamma/3$, $\gamma > 0$, is the only portion of the (γ, α)-plane not included in Theorem 4.

REMARK 2. The results of §7 of Pfeiffer [**5**] are included in Theorem 4 except for the case $c = 0$.

REFERENCES

1. F. V. Atkinson, *Discrete and Continuous Boundary Problems*, Academic Press, New York, 1964.
2. R. Bellman, *Stability Theory of Differential Equations*, McGraw-Hill Book Company, New York, 1953.
3. A. Devinatz, *The deficiency index of certain fourth-order ordinary self-adjoint differential operators*, Quart. J. Math. (Oxford) (2), **23** (1972), 267–286.

4. N. Dunford and J. T. Schwartz, *Linear Operators. Part II: Spectral Theory*, Interscience Publishers, New York, 1963.

5. W. N. Everitt, *Integrable-square solutions of ordinary differential equations*, Quart. J. Math. (Oxford) (2), **10** (1959), 145–155.

6. M. V. Fedorjuk, *Asymptotic methods in the theory of one-dimensional singular differential operators*, Transactions of the Moscow Mathematical Society, **15** (1966), 333–386. English translation, Amer. Math. Soc.

7. V. I. Kogan and F. S. Rofe-Beketov, *The deficiency indices of differential operators of odd order*, Kharkov Ped. Inst. named after Lenin, Materials of the sci. and techn. conference on results of sci. work in 1969, publ. by Kharkov State Univ., 1970, No. 7, 93–95.

8. ———, *On the question of the deficiency indices of differential operators with complex coefficients*, Proc. Roy. Soc. Edinburgh Sect. A72, part 4, (1975), 281–298.

9. M. A. Naimark, *Linear Differential Operators*, English translation, Frederick Ungar Pub. Co., New York, 1968.

10. G. W. Pfeiffer, *Deficiency indices of a third order equation*, Differential Equations, **11** (1972), 474–490.

11. R. A. Shirikyan, *Asymptotic methods in the theory of odd-order one-dimensional singular differential operators*, Differentsial'nye Uravneniya, **3** (1967), 1942–1956.

12. K. Unsworth, *Asymptotic expansions and deficiency indices associated with third-order self-adjoint differential operators*, Quart. J. Math. (Oxford) (2), **24** (1973), 177–188.

Received September 4, 1974 and in revised form March 1, 1976. This work was supported in part by the National Science Foundation, Grants GP 36640 and GP 36640 A-1.

California State University — Fullerton

Fullerton, CA 92634

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No 2, 1977

S-SPACES IN COUNTABLY COMPACT SPACES USING OSTASZEWSKI'S METHOD

JOHN GINSBURG

A method adapted from that used by A. J. Ostaszewski is used to construct S-spaces as subspaces of given spaces. Assuming the set-theoretic principle $\diamond$, it is shown that every countably compact space containing no nontrivial convergent sequences contains a perfect S-space. As a corollary, assuming $\diamond$, if X is a countably compact F-space, then X contains a hereditarily extremally disconnected, hereditarily normal, perfect S-space.

1. Introduction. The set-theoretic principle $\diamond$, due to Jensen **[3]**, has found many interesting applications in topology, particularly the construction of Souslin lines and various S-spaces. The basic technique for constructing S-spaces from $\diamond$ is due to A. J. Ostaszewski **[6]**, and has been modified and applied in constructing other interesting topological spaces, notably in **[5]** and **[8]**. Roughly speaking, the method involves constructing a space having desired properties by defining its topology inductively over more and more of the space (and in some cases refining a given topology) using some principle of enumeration.

Here we will show how the method can be used to construct S-spaces as subspaces of given spaces. That is, rather than building up a space by inductively defining its topology, the desired examples will be obtained by working within a given topological space and extracting a subspace.

Our principal topological references are **[2]**, **[7]** and **[10]**. For set-theoretic notions we refer to **[4]**.

For the reader's convenience we now recall a few notions from topology which we will employ.

A space X is an *S-space* if X is regular, hereditarily separable and not Lindelöf.

X is *countably compact* if every countable covering of X by open sets has a finite subcover.

For a completely regular space X, βX denotes the Stone–Čech compactification of X.

A subset A of X is *C^*-embedded* in X if every bounded, continuous real-valued function on A admits a continuous extension to X. A *cozero-set* in X is a set of the form $\{p \in X: f(p) \neq 0\}$ where f is a continuous real-valued function on X. X is an *F-space* if X is completely regular and every cozero-set in X is C^*-embedded in X. A completely regular space X is *extremally disconnected* if the closure of every open subset of X is open.

For the basic information on F-spaces and extremally disconnected spaces, the reader is referred to **[2]** and **[10]**. We will make use of the following two facts, established in 1.62 and 1.64 of **[10]**.

1.1. If X is σ-compact and locally compact, then $\beta X - X$ is a compact F-space.

1.2. If X is an F-space then every countable subspace of X is C^*-embedded in X.

For the consistency of $\diamond$ with the axioms of set theory the reader is referred to **[3]**. We will not need a precise statement of $\diamond$, rather we will use the following consequence of $\diamond$ derived in **[6]**.

1.3. Let $\lim \omega_1$ denote the set of limit ordinals less than ω_1. Then there is a family $\{S_\gamma: \gamma \in \lim \omega_1\}$ of subsets of ω_1 such that each S_γ is a cofinal subset of γ and such that for every uncountable subset S of ω_1 there is a $\gamma \in \lim \omega_1$ with $S_\gamma \subseteq S$.

It is clear we may assume that each S_γ is a simple ω-sequence increasing to γ in 1.3. This is the form in which we will apply 1.3. (the conclusion of 1.3 is often referred to as "club"; see **[7]**)

2. S-subspaces of countably compact spaces. We now assume the conclusion of 1.3. This assumption will enable us to construct S-spaces in certain countably compact spaces. It is apparently not yet known whether 1.3 is equivalent to $\diamond$ or whether it is strictly weaker. It is known that $\diamond$ is equivalent to the conjunction of 1.3 and the continuum hypothesis, and so this question amounts to whether or not 1.3 implies the continuum hypothesis. (see **[7]**)

All hypothesized spaces are assumed to be infinite.

2.1. THEOREM. *If X is a regular, countably compact Hausdorff space containing no nontrivial convergent sequences, then X contains a perfect S-space.*

Proof. Let $\{S_\gamma: \gamma \in \lim \omega_1\}$ satisfy 1.3 where each S_γ is an ω-sequence increasing to γ. Let X satisfy the hypotheses of the theorem. We inductively select points $(x_\xi: \xi \in \omega_1)$ in X, and open sets $(G_\xi: \xi \in \omega_1)$ in X so that

(i) for all ξ, $x_\xi \in G_\xi$

(ii) $\xi < \eta \to x_\eta \notin G_\xi$

(iii) for all limit ordinals γ and all $n \in \omega$, $x_{\gamma+n} \in \mathrm{cl}\{x_\xi: \xi \in S_\gamma\}$.

To get the desired sequences $(x_\xi\colon \xi \in \omega_1)$ and $(G_\xi\colon \xi \in \omega_1)$ we construct $(x_\xi\colon \xi < \gamma)$ and $(G_\xi\colon \xi < \gamma)$ by induction on the limit ordinal γ. To start the construction, we choose a countable discrete subset $(x_n\colon n \in \omega)$ of X, (X is assumed infinite), and a sequence of open sets $(G_n\colon n \in \omega)$ in X such that $x_n \in G_n$ and $m \neq n \to x_m \notin G_n$.[1]

Now suppose $\sigma \in \lim \omega_1$ and for every limit ordinal $\gamma < \sigma$ we have chosen the sequences $(x_\xi\colon \xi < \gamma)$ and $(G_\xi\colon \xi < \gamma)$ satisfying (i), (ii), and (iii). If σ is a limit of limits, we simply gather together all the x_ξ's and G_ξ's previously constructed to form $(x_\xi\colon \xi < \sigma)$ and $(G_\xi\colon \xi < \sigma)$, clearly satisfying (i), (ii), and (iii). So we need only consider the case where $\sigma = \gamma + \omega$ for some limit ordinal γ. Thus, having the sequences $(x_\xi\colon \xi < \gamma)$ and $(G_\xi\colon \xi < \gamma)$ we must define the points $(x_{\gamma+n}\colon n \in \omega)$ and the open sets $(G_{\gamma+n}\colon n \in \omega)$. Consider the infinite set $R_\gamma = \{x_\xi\colon \xi \in S_\gamma\}$. Since X is countably compact, every countable subset of X has a limit point in X. But since X contains no nontrivial convergent sequences, every countable set has infinitely many (in fact uncountably many) limit points. Thus $\operatorname{cl} R_\gamma - R_\gamma$ is infinite, and so contains a countable discrete subspace $(x_{\gamma+n}\colon n \in \omega)$. Choose a sequence of open sets $(G_{\gamma+n}\colon n \in \omega)$ which witnesses this discreteness, that is, with $x_{\gamma+n} \in G_{\gamma+n}$ and such that $m \neq n \to x_{\gamma+m} \notin G_{\gamma+n}$.

We now check (i), (ii), and (iii) for $(x_\xi\colon \xi < \gamma + \omega)$ and $(G_\xi\colon \xi < \gamma + \omega)$. (i) is clear, as is (iii), by virtue of the induction hypothesis and the selection of the points $x_{\gamma+n}$ in $\operatorname{cl} R_\gamma$. To verify (ii), because of the induction hypothesis and the choice of $(x_{\gamma+n}\colon n \in \omega)$ and $(G_{\gamma+n}\colon n \in \omega)$, it is sufficient to check the following:

If $\xi < \gamma$ and $n \in \omega$, then $x_{\gamma+n} \notin G_\xi$. But S_γ is an ω-sequence increasing to γ, and so there are at most finitely many ordinals in S_γ which are less than ξ. By property (ii) of the induction hypothesis, this means there are at most finitely many x_η with $\eta \in S_\gamma$ which lie in G_ξ. But $x_{\gamma+n}$ is a limit point of R_γ, so every neighborhood of $x_{\gamma+n}$ contains infinitely many x_η with $\eta \in S_\gamma$. In particular, $x_{\gamma+n} \notin G_\xi$.

This completes the inductive construction, and results in sequences $(x_\xi\colon \xi \in \omega_1)$ and $(G_\xi\colon \xi \in \omega_1)$ satisfying (i), (ii), and (iii).

We now claim that $Y = \{x_\xi\colon \xi \in \omega_1\}$ is a perfect S-space. The verification of this is essentially identical with the argument given in **[6]**, so we will be content to sketch that argument here. That Y is not Lindelöf is immediate from (ii) and (i). Any countable subspace of Y is separable, and if $\{x_\xi\colon \xi \in S\}$ is an uncountable subspace of Y, there is, by 1.3, a $\gamma \in \lim \omega_1$ such that $S_\gamma \subseteq S$. Using (iii) we see that $\{x_\xi\colon \xi \in S$ and $\xi < \gamma\}$ is a countable dense subset of $\{x_\xi\colon \xi \in S\}$. This proves Y is hereditarily separable. Since $\gamma < \eta \to x_\eta \in \operatorname{cl}\{x_\xi\colon \xi \in S_\gamma\}$, the same ar-

[1] The fact that every infinite Hausdorff space contains a countably infinite discrete subspace is well-known and easy to prove. A proof may be found in 0.13 of **[2]**.

gument shows that every closed subset of Y is either countable or co-countable, from which it is immediate that every closed subset of Y is a G_δ in Y, that is, Y is perfect.

2.2. COROLLARY. *If X is a countably compact F-space then X contains a hereditarily extremally disconnected, hereditarily normal, perfect S-space.*

Proof. Using 1.2 it is easy to see there are no nontrivial convergent sequences in an F-space, so the hypotheses of 2.1 apply. We show that the S-space Y obtained in 2.1 is hereditarily extremally disconnected and hereditarily normal under the present assumptions on X. Now, as is well-known, a space is extremally disconnected if and only if each of its open subsets is C^*-embedded (see 1H in **[2]**), and a space is normal if and only if each of its closed subsets is C^*-embedded (see 3D in **[2]**). So to verify that Y is normal and extremally disconnected hereditarily, it is sufficient to prove that every subspace of Y is C^*-embedded in Y. So, let $Z \subseteq Y$, and let f be a bounded, continuous real-valued function on Z. Since Y is hereditarily separable, Z contains a countable dense subset D. By 1.2, D is C^*-embedded in X, and so the function $f \mid D$ admits a continuous extension F to all of X. Clearly $F \mid Y$ is the desired extension of f.

REMARK. 2.3. There is a large number of spaces to which these results can be applied. One class of such spaces is furnished by 1.1. So assuming 1.3 we see for example that $\beta\mathbf{R} - \mathbf{R}$ and $\beta N - N$ contain interesting S-spaces.

REMARK. 2.4. The fact that $\diamondsuit$ implies the existence of S-spaces which are extremally disconnected was previously observed by M. Wage **[9]**. Wage's construction, like Ostaszewski's original method, involves inductively defining a topology to get the desired example.

One significant difference between the S-spaces obtained in 2.2 and the original S-space described in **[6]** is countable compactness. The S-space in **[6]** is, in addition, countably compact, while the S-spaces in 2.2 are never countably compact. If CH is true this follows from the results in **[11]** which imply that, assuming CH, every countably compact, separable normal F-space is compact, and therefore Lindelöf. If CH is false, we argue as follows: A slight modification of the argument in **[1]** shows that a countably compact space of cardinality $< c$ is sequentially compact. Since our S-spaces have cardinality ω_1 and contain no convergent sequences, they cannot be countably compact if CH fails either. Thus our S-spaces constructed using 1.3 are not countably compact.

REFERENCES

1. S. P. Franklin, *On two questions of Moore and Mrowka*, Proc. Amer. Math. Soc. 21 (1969), 597–599.

2. L. Gillman and M. Jerison, *Rings of continuous functions*, University Series in Higher Math., Van Nostrand, Princeton, N.J., (1960).

3. R. B. Jensen, *The fine structure of the constructible hierarchy*, Ann. Math. Logic, **4** (1972), 229–308.

4. I. Juhász, *Cardinal functions in topology*, Math. Centre Tracts 34, Mathematical Centre, Amsterdam, (1971).

5. I. Juhász, K. Kunen, and M. E. Rudin, *Two more hereditarily separable non-Lindelöf spaces*, Canad. J. Math., **28** No. 5 (1976), 998–1005.

6. A. J. Ostaszewski, *On countably compact, perfectly normal spaces*, J. London Math. Soc., (2), **14** (1976), 505–516.

7. M. E. Rudin, *Lectures on set-theoretic topology*, CBMS series, **23**, Amer. Math. Soc., 1975.

8. E. K. Van Douwen, *A technique for constructing honest locally compact, submetrizable examples*, to appear.

9. M. Wage, *Extremally disconnected S-spaces*, to appear.

10. R. C. Walker, *The Stone–Čech compactification.* Ergebnisse der Mathematik und ihrer Grenzgebiete, Bd. 83, Berlin-Heidelberg-New York, Springer 1974.

11. R. G. Woods, *The structure of small normal F-spaces*, Proc. Auburn Top. Conf. (1976).

Received November 29, 1976 and in revised form January 19, 1977.

UNIVERSITY OF MANITOBA,

WINNIPEG, MANITOBA, CANADA

PACIFIC JOURNAL OF MATHEMATICS
Vol 68 No 2, 1977

ON A THEOREM OF DELAUNAY AND SOME RELATED RESULTS

B. GORDON AND S. P. MOHANTY

Dedicated to the memory of Professor T. S. Motzkin

Delaunay has proved that if $\epsilon = ap\phi^2 + bp\phi + c$ is a unit in the ring $Z[\theta]$, where $\theta^3 - P\theta^2 + Q\theta - R = 0$, p is an odd prime, $\phi = p^t\theta$, $t \geqq 0$ and $p \nmid a$, then no power ϵ^m (m positive) can be a binorm, i.e. $\epsilon^m = u + v\theta$ is impossible for m a positive integer. Hemer has pointed out that in the above situation, $\epsilon^m = u + v\theta$ is also impossible for m a negative integer.

In this paper the above result is extended as follows.

THEOREM 1. *If $\epsilon = a\theta^2 + b\theta + c$ is a unit in $Z[\theta]$, where $\theta^3 = d\theta^2 + e\theta + f$ and $p^\alpha \| a$, $p^\beta \| b$, p being a prime, then $\epsilon^n = u + v\theta$ is impossible for $n \neq 0$ in the following cases:*
(i) *When $1 \leqq \alpha \leqq \beta$ and p is odd,*
(ii) *When $2 \leqq \alpha \leqq \beta$ and $p = 2$,*
(iii) *When $\beta \leqq \alpha < 2\beta$ and p is odd,*
(iv) *When $\beta \leqq \alpha < 2\beta - 1$ and $p = 2$.*

As an application of this and some other similar theorems, all integer solutions of the equation $y^2 = x^3 + 113$ are determined.

First we prove two simple lemmas.

LEMMA 2. *If $p^\alpha \| \binom{n}{p^q}$ then $p^\alpha \mid \binom{n}{i}$, where the prime p satisfies $p^q < i < p^{q+1}$ and $p^{\alpha-1}$ a $\binom{n}{p^{q+1}}$. Furthermore if $p \mid n$ and $p \nmid i$ then $p^{\alpha+1} \mid \binom{n}{i}$.*

Proof. Let $i = p^q + r$. Then $0 < r < p^{q+1} - p^q$. Hence

$$\binom{n}{i} = \binom{n}{p^q}\binom{n - p^q}{r}\frac{r!}{\prod_{j=1}^{r}(p^q + j)}.$$

Since $\Pi_{j=1}^r (p^q + j)/r!$ is an integer not divisible by p and $p^\alpha \| \binom{n}{p^q}$, we have $p^\alpha \mid \binom{n}{i}$.

If $p \mid n$ and $p \nmid i$ then $p \nmid r$ for $i = p^q + r$. Then

$$\binom{n - p^q}{r} = \left(\frac{n - p^q}{r}\right)\binom{n - p^q - 1}{r - 1}$$

is divisible by p. Hence $p^{\alpha+1} \mid \binom{n}{i}$.

Again from

$$\binom{n}{p^{q+1}} = \binom{n}{p^q}\binom{n-p^q}{p^{q+1}-p^q}\frac{s!}{\prod_{j=1}^{s}(p^{q+1}-j)}\left(\frac{p^{q+1}-p^q}{p^{q+1}}\right),$$

where $s = p^{q+1} - p^q - 1$, we see that $p^{\alpha-1} \mid \binom{n}{p^{q+1}}$, and the lemma is proved.

LEMMA 3. *Let $\epsilon = a\theta^2 + b\theta + c$ be a unit in $Z[\theta]$, where $\theta^3 = d\theta^2 + e\theta + f$, and $\epsilon^{-1} = a'\theta^2 + b'\theta + c'$. If $p^\alpha \| a$, $p^\beta \| b$, where p is a prime and $\alpha\beta \neq 0$, then $p^\alpha \| a'$ and $p^\beta \| b'$ in the following cases:*

(i) $\alpha \leqq \beta < 2\alpha$

(ii) $\beta \leqq \alpha < 2\beta$

For $\alpha \leqq \beta$ we have $p^\alpha \| a'$ and $p^\alpha \mid b'$.

Proof. Since $(a\theta^2 + b\theta + c)(a'\theta^2 + b'\theta + c') = 1$, we have,

(1) $$aa'd^2 + ab'd + a'bd + aa'e + ac' + ca' + bb' = 0,$$

(2) $$aa'f + aa'de + ab'e + a'be + bc' + b'c = 0,$$

and

(3) $$aa'df + ab'f + a'bf + cc' = 1.$$

From (3) it follows that $p \nmid c'$.

Case (i). From (1) we have $ca' \equiv 0 \pmod{p^\alpha}$ as $\alpha \leqq \beta$. Since $p \nmid c$ we get $a' \equiv 0 \pmod{p^\alpha}$. From (2) we obtain $b'c \equiv 0 \pmod{p^\alpha}$ for $\alpha \leqq \beta$, whence $b' \equiv 0 \pmod{p^\alpha}$. If $\beta < 2\alpha$, then (2) gives $b'c \equiv 0 \pmod{p^\beta}$, or $b' \equiv 0 \pmod{p^\beta}$. If $p^{\alpha+1} \mid a'$, then from (1) we have $ac' \equiv 0 \pmod{p^{\alpha+1}}$. Since $p \nmid c'$ we get $a \equiv 0 \pmod{p^{\alpha+1}}$, a contradiction. Hence $p^\alpha \| a'$. Similarly if $p^{\beta+1} \mid b'$, then from (2) we get $bc' \equiv 0 \pmod{p^{\beta+1}}$ when $\beta < 2\alpha$. Again we arrive at a contradiction since $p \nmid c'$ and $p^\beta \| b$. Hence $p^\beta \| b'$.

Case (ii). Since $\beta \leqq \alpha$, (2) yields $b'c \equiv 0 \pmod{p^\beta}$. Then we have $b' \equiv 0 \pmod{p^\beta}$ for $p \nmid c$. Using $\alpha < 2\beta$, we get $a'(bd + c) \equiv 0 \pmod{p^\alpha}$ from (1). Then $a' \equiv 0 \pmod{p^\alpha}$ as $p \nmid (bd + c)$. If $b' \equiv$

$0 \pmod{p^{\beta+1}}$, then from (2) we see that $bc' \equiv 0 \pmod{p^{\beta+1}}$, a contradiction. Hence $p^\beta \| b'$. If $a' \equiv 0 \pmod{p^{\alpha+1}}$ we have from (1) $ac' + bb' \equiv 0 \pmod{p^{\alpha+1}}$. We get a contradiction for $\alpha < 2\beta$. Hence $p^\alpha \| a'$.

Proof of Theorem 1. Let $n > 0$. Case (i) and (ii). Let $1 \leqq \alpha \leqq \beta$. Since ϵ is a unit, $p \nmid c$. Moreover $\epsilon = a\theta^2 + b\theta + c = p^\alpha(r\theta^2 + s\theta) + c$ where $p \nmid r$. Let $(r\theta^2 + s\theta)^i = a_i\theta^2 + b_i\theta + c_i$, with a_i, b_i and c_i rational integers. Then

$$\epsilon^n = (a\theta^2 + b\theta + c)^n = [c + p^\alpha(r\theta^2 + s\theta)]^n = c^n + \binom{n}{1} c^{n-1}p^\alpha(r\theta^2 + s\theta)$$

$$+ \binom{n}{2} c^{n-2}p^{2\alpha}(a_2\theta^2 + b_2\theta + c_2) + \cdots + p^{n\alpha}(a_n\theta^2 + b_n\theta + c_n) = u + v\theta.$$

Comparing the coefficients of θ^2, we have

$$nc^{n-1}p^\alpha r + \binom{n}{2} c^{n-2}p^{2\alpha}a_2 + \cdots + p^{n\alpha}a_n = 0. \tag{4}$$

If p is an odd prime, we see using Lemma 2 that the first term of (4) is divisible by a lower power of p than the others. If $p = 2$ and $\alpha \geqq 2$ the same conclusion holds. Hence (4) can never be satisfied. So ϵ^n can never be of the form $u + v\theta$ in these cases.

Cases (iii) *and* (iv). Now $\epsilon = p^\beta(r\theta^2 + s\theta) + c$, where $p^{\alpha-\beta} \| r$. Then the coefficient of θ^2 in $\epsilon^n = [c + p^\beta(r\theta^2 + s\theta)]^n$ is

$$nc^{n-1}p^\beta r + \binom{n}{2} c^{n-2}p^{2\beta}a_2 + \cdots + p^{n\beta}a_n, \tag{5}$$

where $(r\theta^2 + s\theta)^i = a_i\theta^2 + b_i\theta + c_i$ with a_i, b_i and c_i rational integers. Again using Lemma 2 and the fact that $\alpha < 2\beta$, we see that the first term of (5) is divisible by a lower power of p than the others if p is an odd prime.

In case $p = 2$ and $\alpha < 2\beta - 1$ the same conclusion holds. Hence (5) can never be zero, i.e. $\epsilon^n = u + v\theta$ is impossible. This proves the theorem for $n > 0$.

We next consider $\epsilon^n = u + v$ for $n < 0$.

Let $n = -m$ and $\epsilon^{-1} = a'\theta^2 + b'\theta + c'$. Then we have $\epsilon^n = (\epsilon^{-1})^m = (a'\theta^2 + b'\theta + c')^m$ where $m > 0$. From Lemma 3, we see that $p^\alpha \| a'$, $p^\alpha | b'$ for $\alpha \leqq \beta$, and $p^\alpha \| a'$, $p^\beta \| b'$ for $\beta \leqq \alpha < 2\beta - 1$, $\alpha \leqq \beta < 2\alpha$ and $\beta \leqq \alpha < 2\beta$. Hence $(a'\theta^2 + b'\theta + c')^m = u + v\theta$ is impossible for $m > 0$. Combining these results we see that $\epsilon^n = u + v\theta$ is impossible for $n \neq 0$, and the theorem is proved.

We note that if the conditions of Theorem 1 are not fulfilled, then $\epsilon^n = u + v\theta$ is possible for $n > 3$; examples are given in [**2**, page 417].

Very often the following theorem is useful.

THEOREM 4. *Let* $\epsilon = a_1\theta^2 + b_1\theta + c_1$ *be a unit in* $Z[\theta]$, *where* $\theta^3 - p_1\theta - q_1 = 0$. *If* $p_1 \equiv 0 \pmod 3$, *then*

$$\epsilon^n = u + v\theta \tag{6}$$

is impossible for $n \neq 0$ *provided* $a_1 \not\equiv 0 \pmod 3$, $b_1^2 + 2a_1c_1 \not\equiv 0 \pmod 3$, *and* $b_1^2c_1 + a_1c_1^2 + a_1^2b_1q_1 \not\equiv 0 \pmod 3$.

Proof. Let $\epsilon^n = a_n\theta^2 + b_n\theta + c_n$. Then we have

$$a_{n+1} = a_n(a_1p_1 + c_1) + b_nb_1 + c_na_1,$$
$$b_{n+1} = a_n(a_1q_1 + b_1p_1) + b_n(c_1 + a_1p_1) + c_nb_1,$$

and

$$c_{n+1} = a_nb_1q_1 + b_na_1q_1 + c_nc_1.$$

Hence we get $a_2 = a_1^2p_1 + b_1^2 + 2a_1c_1$, $b_2 = a_1^2q_1 + 2b_1c_1 + 2a_1b_1p_1$, and $c_2 = c_1^2 + 2a_1b_1q_1$. Then $a_3 = a_1^3p_1^2 + 3a_1b_1^2p_1 + 3a_1^2c_1p_1 + 3b_1^2c_1 + 3a_1c_1^2 + 3a_1^2b_1q_1$, $b_3 = 2a_1^3p_1q_1 + 3a_1b_1^2q_1 + 3a_1^2c_1q_1 + 3a_1^2b_1p_1^2 + b_1^3p_1 + 6a_1b_1c_1p_1 + 3b_1c_1^2$, and $c_3 = 3a_1^2b_1p_1q_1 + b_1^3q_1 + 6a_1b_1c_1q_1 + a_1^3q_1^2 + c_1^3$. Suppose $p_1 \equiv 0 \pmod 3$. Then $a_3 \equiv 0 \pmod 3$, $b_3 \equiv 0 \pmod 3$, and $c_3 \equiv b_1q_1 + a_1q_1^2 + c_1 \pmod 3$.

Since ϵ^3 is a unit, $c_3 \not\equiv 0 \pmod 3$ as $a_3 \equiv b_3 \equiv 0 \pmod 3$. Hence we have $c_3 \equiv 1$ or $2 \pmod 3$.

Suppose $n \equiv 1 \pmod 3$, and put $n = 1 + 3m$ in (6). We get

$$\epsilon \cdot (\epsilon^3)^m = u + v\theta,$$

or

$$(a_1\theta^2 + b_1\theta + c_1)(\pm 1)^m \equiv u + v\theta \pmod 3.$$

This congruence is impossible unless $a_1 \equiv 0 \pmod 3$. Hence if $a_1 \not\equiv 0 \pmod 3$, then $n \not\equiv 1 \pmod 3$. Suppose $n \equiv 2 \pmod 3$, and let $n = 2 + 3m$. Then (6) gives

$$(a_2\theta^2 + b_2\theta + c_2)(\pm 1)^m \equiv u + v\theta \pmod 3.$$

This is impossible unless $a_2 \equiv 0 \pmod 3$, i.e. $b_1^2 + 2a_1c_1 \equiv 0$

(mod 3). Hence if $b_1^2+2a_1c_1 \not\equiv 0 \pmod 3$, then $n \equiv 2 \pmod 3$ is impossible. Finally suppose $n = 3m$ in (6). Then we get

$$(a_3\theta^2 + b_3\theta + c_3)^m = u + v\theta. \tag{7}$$

Now $a_3 \equiv b_3 \equiv 0 \pmod 3$, and $a_3 \equiv 3b_1^2c_1 + 3a_1c_1^2 + 3a_1^2b_1q_1 \pmod 9$. If $b_1^2c_1 + a_1c_1^2 + a_1^2b_1q_1 \not\equiv 0 \pmod 3$, then $a_3 \not\equiv 0 \pmod 9$ and hence by Theorem 1, (7) is impossible for m an integer, positive or negative.

Therefore $n = 0$ is the only solution to (6).

LEMMA 5 (Delaunay [2, page 385]). *If $b\theta + c$, where $b \neq 0, \pm 1$, is a positive unit of $Z[\theta]$ where $\theta^3 - P\theta^2 + Q\theta - R = 0$, then no power > 1 of $b\theta + c$ can be a binomial unit.* (In other words all the positive powers of the positive unit $b\theta + c$ are of the form $L\theta^2 + M\theta + N$, where $L \neq 0$).

We prove two theorems which are useful when $b = \pm 1$.

THEOREM 6. *Let $\epsilon = \pm\theta + c$ be a unit in $Z[\theta]$, where $\theta^3 - P\theta^2 + Q\theta - R = 0$. If $\theta^3 \equiv 0 \pmod{p^2}$, where p is a prime, then $p \nmid c$ and $\epsilon^n = u + v\theta$ is impossible for $n > 1$.*

Proof. We have $(\epsilon - c)^3 \equiv 0 \pmod{p^2}$. If $p \mid c$ then $\epsilon^3 \equiv 0 \pmod p$ where $p^3 \mid N(\epsilon^3) = \pm 1$. Hence $p \nmid c$. Let $\epsilon^n = u + v\theta$, $n > 1$. Then

$$(c \pm \theta)^n = c^n + \binom{n}{1} c^{n-1}(\pm\theta) + \binom{n}{2} c^{n-2}\theta^2 + \binom{n}{3} c^{n-3}(\pm\theta)^3 + \cdots$$
$$+ (\pm\theta)^n = u + v\theta.$$

Let $\theta^n = r_n\theta^2 + s_n\theta + t_n$. Then

$$\binom{n}{2} c^{n-2} + \binom{n}{3} c^{n-3}(\pm r_3) + \cdots + (\pm r_n) = 0. \tag{8}$$

As $\theta^3 \equiv 0 \pmod{p^2}$, we have $r_i \equiv 0 \pmod{p^{2[i/3]}}$. Since $p \nmid c$, $p \mid \binom{n}{2}$. Suppose $p^k \,\|\, \binom{n}{2}$. If $p = 2$ then $2^k \,\|\, \binom{n}{2}$. If $p \neq 2$ then $p^k \,\Big\|\, \binom{n}{2}, \binom{n}{3} \cdots \binom{n}{p-1}$ and $p^{k-1} \,\Big\|\, \binom{n}{p}$. Using Lemma 2, we see that each term of (8) except the first is divisible by at least p^{k+1}. Hence $p^{k+1} \,\Big|\, \binom{n}{2}$, a contradiction.

THEOREM 7. *Let $\epsilon = \pm\theta + c_1$ be a unit of the ring $Z[\theta]$, where $\theta^3 - 3P\theta^2 + 3Q\theta - R = 0$. If $c_1 + P \not\equiv 0 \pmod 3$ and $c_1^2 + 2c_1P + Q \not\equiv 0$ (mod 3), then $\epsilon^n = u + v\theta$ is impossible for $n > 1$.*

Proof. Let $\varepsilon = \theta + c_1$. Then $\theta = \epsilon - c_1$. So from

$$\theta^3 - 3P\theta^2 + 3Q\theta - R = 0,$$

we get

$$(\epsilon - c_1)^3 - 3P(\epsilon - c_1)^2 + 3Q(\epsilon - c_1) - R = 0,$$

or

$$\epsilon^3 = 3(c_1 + P)\epsilon^2 - 3(c^2 + 2c_1P + Q)\epsilon + (c_1^3 + 3c_1^2P + 3c_1Q + R).$$

Now $N(\epsilon) = c_1^3 + 3c_1^2P + 3c_1Q + R = \pm 1$.

For convenience we write $\epsilon^3 = 3r\epsilon^2 - 3s\epsilon \pm 1$. Now by hypothesis $3 \nmid r$ and $3 \nmid s$. Let $\epsilon^n = u + v\theta$. Then $\epsilon^n = u + v(\epsilon - c_1) = u_1 + v_1\epsilon$, say. Suppose $n \equiv 2 \pmod 3$. Then $\epsilon^2(\epsilon^3)^m = u_1 + v_1\epsilon$, where $n = 2 + 3m$. As $\epsilon^3 \equiv \pm 1 \pmod 3$, we have $\pm\epsilon^2 \equiv u_1 + v_1\epsilon \pmod 3$, which is impossible. Let $n \equiv 0 \pmod 3$ and $n \neq 0$. Putting $n = 3m$, we get

$$(3r\epsilon^2 - 3s\epsilon \pm 1)^m = u_1 + v_1\epsilon. \tag{9}$$

But this is impossible by Theorem 1, whether m is a positive or a negative integer, for $3 \nmid r$. Hence if $n \neq 0$, the only possibility is $n \equiv 1 \pmod 3$.

Let $n = 1 + 3m$, where $m > 0$. Then

$$\epsilon(3r\epsilon^2 - 3s\epsilon \pm 1)^m = u_1 + v_1\epsilon,$$

or

$$(3r\epsilon^2 - 3s\epsilon \pm 1)^m = v_1 \pm u_1(\epsilon^2 - 3r\epsilon + 3s).$$

Let $(r\epsilon^2 - s\epsilon)^i = r_i\epsilon^2 + s_i\epsilon + t_i$, where r_i, s_i, t_i are rational integers. Then

$$(\pm 1)^m + \binom{m}{1}(\pm 1)^{m-1}3(r\epsilon^2 - s\epsilon) + \binom{m}{2}(\pm 1)^{m-2}3^2(r_2\epsilon^2 + s_2\epsilon + t_2)$$
$$+ \cdots + 3^m(r_m\epsilon^2 + s_m\epsilon + t_m) = \pm u_1\epsilon^2 \mp 3ru_1\epsilon + (v_1 \pm 3su_1).$$

On equa ing coefficients of ϵ^2 and ϵ, we obtain

$$(\pm 1)^{m-1}3mr + (\pm 1)^{m-2}3^2\binom{m}{2}r_2 + (\pm 1)^{m-3}3^3\binom{m}{3}r_3 + \cdots + 3^m r_m = \pm u_1, \tag{10}$$

and

$$(11)\quad -(\pm 1)^{m-1}3ms + (\pm 1)^{m-2}3^2\binom{m}{2}s_2 + (\pm 1)^{m-3}3^3\binom{m}{3}s_3 + \cdots + 3^m s_m$$
$$= \mp 3ru_1.$$

Multiplying both sides of (10) by $3r$ and then adding to (11), we obtain

$$(\pm 1)^{m-1}3m(3r^2 - s) + (\pm 1)^{m-2}3^2\binom{m}{2}(3r_2 r + s_2)$$
$$+ (\pm 1)^{m-3}3^3\binom{m}{3}(3r_3 r + s_3) + \cdots + 3^m(3r_m r + s_m) = 0.$$

We see from this that $3 \mid m(3r^2 - s)$. As $3 \nmid s$, we have $3 \mid m$. Suppose $3^k \parallel m$. Using Lemma 2, we easily see that all the terms except the first are divisible by 3^{k+2}, while the first is exactly divisible by 3^{k+1}, which is impossible. Hence $m = 0$, i.e. $n = 1$.

So if n is a nonnegative integer and $\epsilon^n = u + v\theta$, then $n = 0$ or $n = 1$. The proof for $\epsilon = -\theta + c$, is completely analogous.

THEOREM 8. *If $\epsilon = b_1\theta + c_1$ is a positive unit in $Z[\theta]$, where $\theta^3 - P\theta^2 + Q\theta - R = 0$ with $D(\theta)$ negative and $\neq -23$, then $\epsilon^n = u + v\theta$ implies that $n \geqq 0$.*

To prove this theorem we need the following well-known result.

LEMMA 9 (Nagell **[8]**). *If η is a unit, $D(\eta) < 0$, $0 < \eta < 1$, then $\eta^n = x + y\eta$ implies that $n \geqq 0$, except in the case when $\eta^3 + \eta^2 - 1 = 0$. In this case $\eta^{-2} = 1 + \eta$ and $D(\eta) = -23$.*

Proof of Theorem 8. Let $\epsilon = b_1\theta + c_1$ be a positive unit in $Z[\theta]$. Then $0 < \epsilon < 1$. Since ϵ is contained in $Z[\theta]$, we get $D(\epsilon) = \delta^2 \cdot D(\theta)$. Hence $D(\epsilon) < 0$ and $\neq -23$.

Let $\epsilon^n = u + \theta$. Since $\epsilon = b_1\theta + c_1$ we have

$$(b_1\theta + c_1)^n = u + v\theta.$$

Then $b_1 \mid v$ when n is a positive integer. In case n is negative, we put $n = -m$ where m is positive. Let $\epsilon^{-1} = a'\theta^2 + b'\theta + c'$. Then $\theta^3 = P\theta^2 - Q\theta + R$ and $\epsilon\epsilon^{-1} = 1$ imply

$$(12)\qquad b_1 a'P + b_1 b' + c_1 a' = 0,$$

$$(13)\qquad -b_1 a'Q + b_1 c' + c_1 b' = 0,$$

and

$$b_1 a' R + c_1 c' = 1. \tag{14}$$

Since $(b_1, c_1) = 1$, $\epsilon = b_1\theta + c_1$ being a unit, we conclude that $b_1 | a'$ and $b_1 | b'$ from (12) and (13) respectively. Then from

$$(b_1\theta + c_1)^n = (a'\theta^2 + b'\theta + c')^m = u + v\theta,$$

we see that $b_1 | v$.
Since $\epsilon = b_1\theta + c_1$, we have $\theta = (\epsilon - c_1)/b_1$, and hence $\epsilon^n = u + v\theta$ can be written as

$$\epsilon^n = u + \frac{v(\epsilon - c_1)}{b_1} = (u - vc_1/b_1) + v\epsilon/b_1 = x + y\epsilon,$$

where x and y are rational integers. Then by Lemma 9, $n \geqq 0$. For binorms in fields of degree higher than three, one can see **[9]**. Recently Bernstein **[1]** has shown that units of the form $\epsilon = 1 + xw + yw^2$, $x, y \in Q$ exist for infinitely many algebraic number fields $Q(w)$ of degree $n \geqq 4$.

Now we solve $y^2 - 113 = x^3$ to show the application of some of the above theorems. The above equation is a special case of the well-known Mordell Equation $y^2 - k = x^3$, which has interested mathematicians for more than three centuries, and has played an important role in the development of number theory. In the range $0 < k \leqq 100$ it is known that $y^2 - k = x^3$, $k = 17$ has the maximum number of solutions. In the range $100 < k \leqq 200$ it is found **[6]** that $y^2 - k = x^3$, $k = 113$ has the maximum number of solutions. The complete solution of this equation is given below.

The fundamental unit of $Q(\sqrt{113})$ is $\eta = 776 + 73\sqrt{113}$, and $h(Q\sqrt{113}) = 1$. 2 splits into two different prime ideals in the field $Q(\sqrt{113})$. Hence by Theorem 5 of Hemer **[4]**, all the integral solutions of $y^2 - 113 = x^3$ can be obtained from the following equations:

$$\pm y + \sqrt{113} = \left(\frac{a + b\sqrt{113}}{2}\right)^3, \quad x = \frac{a^2 - 113b^2}{4},$$

$$\pm y + \sqrt{113} = (776 + 73\sqrt{113})\left(\frac{a + b\sqrt{113}}{2}\right)^3, \quad x = (113b^2 - a^2)/4,$$

$$\frac{1}{2}(\pm y + \sqrt{113}) = \left(\frac{11 + \sqrt{113}}{2}\right)\left(\frac{a + b\sqrt{113}}{2}\right)^3, \quad x = (a^2 - 113b^2)/2,$$

$$\frac{1}{2}(\pm y+\sqrt{113})=\left(\frac{11+\sqrt{113}}{2}\right)(776+73\sqrt{113})\left(\frac{a+b\sqrt{113}}{2}\right)^3,$$
$$x=(113b^2-a^2)/2,$$

$$\frac{1}{2}(\pm y+\sqrt{113})=\left(\frac{11+\sqrt{113}}{2}\right)(776-73\sqrt{113})\left(\frac{a+b\sqrt{113}}{2}\right)^3,$$
$$x=(113b^2-a^2)/2.$$

On equating irrational parts we have respectively

(15) $$3a^2b+113b^3=8,$$

(16) $$73(a^3+3\cdot 113ab^2)+776(3a^2b+113b^3)=8,$$

(17) $$(a^3+3\cdot 113ab^2)+11(3a^2b+113b^3)=8,$$

(18) $$1579(a^3+3\cdot 113ab^2)+16785(3a^2b+113b^3)=8,$$

(19) $$-27(a^3+3\cdot 113ab^2)+287(3a^2b+113b^3)=8.$$

Clearly (15) has no solution in integers. From (16) it is easily seen that a and b are both even. Putting $a=2u_1$, $b=2v_1$ in (16), we obtain

(20) $$73(u_1^3+3\cdot 113u_1v_1^2)+776(3u_1^2v_1+113v_1^3)=1.$$

The substitution $u_1=21u-52v$, $v_1=-2u+5v$ in (20) yields

(21) $$F(u,v)=u^3-33uv^2+76v^3=1.$$

This corresponds to the ring $Z[\theta]$, where $\theta^3-33\theta-76=0$. In this ring the fundamental unit is $\epsilon=4\theta^2-16\theta-71$. By Theorem 1,

$$(4\theta^2-16\theta-71)^n=u+v\theta$$

is only possible for $n=0$. Then $u=1$, $v=0$, and so $a=42$, $b=-4$. Hence $x=11$, $y=\pm 38$.

The substitution $a=u_1-11v_1$, $b=v_1$ in (17) gives

(22) $$u_1^3-24u_1v_1^2+176v_1^3=8.$$

Hence $u_1\equiv 0 \pmod 2$. Putting $u_1=2u$, $v_1=v$ in (22), we get

(23) $$F(u,v)=u^3-6uv^2+22v^3=1.$$

This corresponds to the ring $Z[\theta]$, where $\theta^3 - 6\theta - 22 = 0$; $Z[\theta]$ has fundamental unit $\epsilon = 2\theta - 7$.

Now we consider

$$(2\theta - 7)^n = u + v\theta. \tag{24}$$

By Theorem 8, $n \geqq 0$ and by Lemma 5, $n \leqq 1$. Therefore (24) has only the two solutions $n = 0$, $n = 1$. These solutions correspond to $x = 2$, $y = \pm 11$ and $x = 422$, $y = \pm 8669$ respectively.

Substituting $a = -21u_1 + 53v_1$, $b = 2u_1 - 5v_1$ in (18), we get

$$8v_1^3 + 12v_1^2u_1 - 42v_1u_1^2 + 27u_1^3 = 8. \tag{25}$$

We put $u_1 = 2v$, $v_1 = u - v$ in (25), since $u_1 \equiv 0 \pmod 2$. This gives

$$F(u, v) = u^3 - 24uv^2 + 50v^3 = 1. \tag{26}$$

This corresponds to the ring $Z[\theta]$, where $\theta^3 - 24\theta - 50 = 0$, with the fundamental unit $\epsilon = -3\theta^2 + 10\theta + 41$. We see that $\epsilon \equiv 2\theta^2 + 1 \pmod 5$ and $\epsilon^2 \equiv 1 \pmod 5$ while $\epsilon^2 \equiv -5\theta^2 + 5\theta + 6 \pmod{25}$. Hence $\epsilon^2 = a_1\theta^2 + b_1\theta + c_1$ implies that $5 \| a_1$, $5 \| b_1$. Hence, by Theorem 1, $\epsilon^n = u + v\theta$ is impossible for an even integer $n \neq 0$. When n is odd we have

$$2\theta^2 + 1 \equiv u + v\theta \pmod 5.$$

This is impossible. So we have $n = 0$. Then $u = 1$, $v = 0$ and hence $x = 8$, $y = \pm 25$.

The substitution $a = 111u_1 + 10v_1$, $b = 11u_1 + v_1$ in (19) yields

$$v_1^3 - 312v_1u_1^2 - 2128u_1^3 = 8. \tag{27}$$

Since (27) implies $v_1 \equiv 0 \pmod 2$, we put $v_1 = 12u + 10v$, $u_1 = -u - v$ and get

$$F(u, v) = v^3 + 12vu^2 + 14u^3 = 1. \tag{28}$$

The fundamental unit of the ring $Z[\theta]$, where $\theta^3 + 12\theta - 14 = 0$, is $\epsilon = \theta - 1$, satisfying $\epsilon^3 + 3\epsilon^2 + 15\epsilon - 1 = 0$.

Then by Theorems 8 and 6,

$$\epsilon^n = (\theta - 1)^n = v + u\theta$$

has only two solutions, viz. $n = 0$ and 1.

Incidentally, we cannot reach this conclusion by using the standard criterion of Hemer **[4]**, which is as follows:

Let $\epsilon = \pm\theta + c$ be a unit in a cubic ring, and let the odd prime p be a divisor of $N(\epsilon' + \epsilon'')$. Suppose further that $\epsilon^m = a_m\epsilon^2 + b_m\epsilon + c_m$ is the least power of ϵ with $m > 0$ such that $a_m \equiv b_m \equiv 0 \pmod p$. Then $\epsilon^n = u + v\epsilon$ has no even solution except $n = 0$ if $a_m \not\equiv 0 \pmod{p^2}$, and no odd solution except $n = 1$ if $c_{m+2} \not\equiv 0 \pmod{p^2}$.

Now $N(\epsilon' + \epsilon'') = N(-3 - \epsilon) = -46$ has only the odd prime divisor $p = 23$. The least exponent m such that $a_m \equiv b_m \equiv 0 \pmod{23}$ is $m = 22$, and $a_m \not\equiv 0 \pmod{23^2}$. But unfortunately $c_{24} \equiv 0 \pmod{23^2}$.

When $n = 0$, $u = 0$, $v = 1$; $a = -11$, $b = -1$; $x = -4$, $y = \pm 7$.

When $n = 1$, $u = 1$, $v = -1$; $a = 20$, $b = 2$; $x = 26$, $y = \pm 133$.

Hence the Diophantine equation $y^2 - 113 = x^3$ has exactly 6 solutions in integers. They are $(x, y) = (11, \pm 38)$, $(8, \pm 25)$, $(2, \pm 11)$, $(-4, \pm 7)$, $(422, \pm 8669)$ and $(26, \pm 133)$.

ACKNOWLEDGEMENT. We are thankful to the referee for comments for the improvement of the paper.

REFERENCES

1. Leon Bernstein, *Truncated units in infinitely many algebraic number fields of degree $n \geqq 4$*, Math. Ann., **213** (1975), 275–279.

2. B. N. Delaunay and D. K. Faddeev, *The theory of irrationalities of the third degree*, Amer. Math. Soc., Providence, Rhode Island (1964).

3. R. Finkelstein and H. London, *On Mordell's Equations $y^2 - k = x^3$*, Bowling Green State University Press.

4. O. Hemer, *On the Diophantine equation $y^2 - k = x^3$*, Diss. Uppsala (1952).

5. L. J. Mordell, *Diophantine equations*, Pure and Appl. Math., **30**, Academic Press, New York, (1969), 238–254.

6. S. P. Mohanty, *On the Diophantine equation $y^2 - k = x^3$*, Diss. UCLA (1971).

7. ———, *On consecutive integer solutions for $y^2 - k = x^3$*, Proc. Amer. Math. Soc., **48** (1975), 281–285.

8. T. Nagell, *Darstellung ganzer Zahlen durch binäre kubische Formen mit negativer Diskriminante*, Ibid Bd. **28** (1928).

9. Hans-Joachim Stender, *Lösbare Gleichungen $ax^n - by^n = C$ and Grundeinheiten für einige algebraische Zahlkörper vom Grade n, $n = 3, 4, 6$*; Habilitation paper, University of Cologne (1975).

Received November 3, 1975. The preparation of this paper was partly supported by NSF grant GP-23113.

UNIVERSITY OF CALIFORNIA, LOS ANGELES
AND
I. I. T. KANPUR, KANPUR-16, INDIA

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No 2, 1977

TOPOLOGICAL GROUPS WHICH SATISFY AN OPEN MAPPING THEOREM

DOUGLASS L. GRANT

Let $\mathscr{C}$ be a category of Hausdorff topological groups. A Hausdorff topological group G is called a $B(\mathscr{C})$ group if every continuous and almost open homomorphism from G onto a group in $\mathscr{C}$ is open. An internal characterization of such groups is obtained. For certain $\mathscr{C}$, the permanence properties of $B(\mathscr{C})$ groups and related categories are investigated, with some positive results pertaining to products and subobjects, and several counterexamples. Forms of the closed graph theorem for topological groups are then obtained which generalize results of T. Husain.

1. Definitions and permanence properties. Given a topological group G with topology u, we shall denote the filter of neighbourhoods of the identity by $\mathscr{V}(G)$ or $\mathscr{V}(u)$, and closures by Cl_G or Cl_u, depending on the emphasis desired. If u and v are two group topologies on a group G, then $v(u)$ will denote that group topology on G having as a fundamental system of unit neighbourhoods the collection $\{\mathrm{Cl}_v U: U \in \mathscr{V}(u)\}$. The set of closed normal subgroups of a topological group G will be denoted by $\mathscr{N}(G)$. A homomorphism $f: G \to H$ of topological groups is said to be almost open (resp., almost continuous) if the image (resp., inverse image) of a unit neighbourhood is dense in a unit neighbourhood. An isomorphism of topological groups is a group isomorphism which is both continuous and open.

Let $\mathscr{C}$ be a category of Hausdorff topological groups. After **[8]**, we say that a Hausdorff group G is a $B(\mathscr{C})$ group if every continuous and almost open homomorphism from G onto a group in $\mathscr{C}$ is open, and that G is a $B_r(\mathscr{C})$ group if every homomorphism with these properties which is also one-to-one is open. We reserve the symbol $\mathscr{A}$ for the category of all Hausdorff topological groups.

Husain **[8]** showed that locally compact groups and complete metrizable groups are $B(\mathscr{A})$ groups, while Brown [**2**, Theorem 4] showed that any topological group complete in the sense of Čech has the $B(\mathscr{A})$ property. A minimal topological group (i.e., one with its coarsest compatible Hausdorff topology) is easily seen to be a $B_r(\mathscr{A})$ group. Other examples will be mentioned later.

Husain also observed [**8**, Theorem 31.4] that a topological group (G, u) is a $B_r(\mathscr{A})$ group iff, for every Hausdorff group topology v on G such that $v \subseteq u$ and $v(u) = v$, it follows that $u = v$. We give analogous statements for $B_r(\mathscr{C})$ and $B(\mathscr{C})$ groups, where $\mathscr{C}$ satisfies a very mild condition. For a topological group (G, u) and $H \in \mathscr{N}(G)$, let uH denote the group topology on G having the collection $\{UH: U \in \mathscr{V}(G)\}$ as a subbasis of unit neighbourhoods, and u/H the quotient topology on G/H. We require a lemma which follows directly from Proposition 30.3 of **[8]**.

LEMMA 1.1. *A topological group G is a $B(\mathscr{C})$ group iff G/H is a $B_r(\mathscr{C})$ group for every $H \in \mathscr{N}(G)$.*

The following definition is adapted from Isbell [**10**, p. 119]. Let $\mathscr{X}$ be a category, $\mathscr{Y}$ a subcategory of $\mathscr{X}$, $\mathscr{F}$ a class of morphisms in $\mathscr{X}$. Then $\mathscr{Y}$ is said to be right fitting with respect to $\mathscr{F}$ if $X \in \mathscr{X}$, $Y \in \mathscr{Y}$, $f\colon Y \to X$ a morphism in $\mathscr{F}$ together imply $X \in \mathscr{Y}$. Let $\mathscr{H}$ denote the class of isomorphisms of Hausdorff topological groups. (More extensive use of this notion, involving other classes of maps, will be made in §2.)

THEOREM 1.2. *Let $\mathscr{C}$ be a subcategory of $\mathscr{A}$ which is right fitting with respect to $\mathscr{H}$.*

(a) *A topological group (G, u) is a $B_r(\mathscr{C})$ group iff, for every group topology v on G such that $(G, v) \in \mathscr{C}$, $v \subseteq u$, and $v(u) = v$, it follows that $v = u$.*

(b) *A topological group (G, u) is a $B(\mathscr{C})$ group iff, for every $H \in \mathscr{N}(G)$ and every group topology v on G such that $(G/H, v/H) \in \mathscr{C}$, $v \subseteq uH$, and $v(uH) = v$, it follows that $v = uH$.*

Proof. Part (a) follows in a manner similar to Theorem 31.4 of **[8]**. One then obtains (b) by invoking Lemma 1.1, applying (a) to the quotient groups, and observing that every group topology on a group G/H coarser than the quotient topology arises from a group topology on G coarser than uH.

Investigation of some permanence properties of $B(\mathscr{A})$ and $B_r(\mathscr{A})$ groups was carried out by L. J. Sulley **[15]**, who gave criteria for the inheritance of these properties by dense subgroups and by completions, in the Abelian case. His assumption of commutativity can be removed quite painlessly, however. The proof of the next lemma proceeds in a fashion nearly identical to that of the corresponding result in **[15]**.

LEMMA 1.3. *Let E be a Hausdorff group, G a dense subgroup of E, $H \in \mathscr{N}(E)$, $q\colon E \to E/H$ the natural map. Then the map $r\colon G \to q(G)$ obtained by restricting q is continuous and almost open. Furthermore, r is open iff $H \cap G$ is dense in H.*

THEOREM 1.4. *Let G be a Hausdorff group, E its completion with respect to its two-sided uniformity.*

(a) *G is a $B(\mathscr{A})$ group iff E is a $B(\mathscr{A})$ group and $G \cap H$ is dense in H for every $H \in \mathscr{N}(E)$.*

(b) *G is a $B_r(\mathscr{A})$ group iff E is a $B_r(\mathscr{A})$ group and $G \cap H$ is nontrivial for every nontrivial $H \in \mathscr{N}(E)$.*

Proof. The "only if" parts of (a) and (b) follow as in **[15]**, using Lemma 1.3. For the "if" part of (a), let F be any Hausdorff group, F' its completion with respect to its two-sided uniformity, f a continuous, almost open homomorphism of G onto F. By Proposition 5, p. 246 of **[1]**, f has a unique extension $f'\colon E \to F'$, which can be shown to be almost open onto its range. The balance follows as in **[15]**, using Lemma 1.3. The proof of the "if" part of (b) is similar, with the additional observation that the extension f' of the one-to-one homomorphism f is also one-to-one.

It follows from this criterion **[15, 16]** that the group U of complex roots of unity is a $B(\mathscr{A})$ group, while, for instance, neither the group Q of rationals nor the group U_p of p-power roots of unity is a $B_r(\mathscr{A})$ group.

Clearly, if a product of groups has the $B(\mathscr{A})$ property, then each factor has this property. Using Theorem 1.4, however, we can show that neither the class of $B(\mathscr{A})$ groups nor that of $B_r(\mathscr{A})$ groups is closed even under finite Cartesian products.

EXAMPLE 1. Let R denote the reals with usual topology and T the circle group, and let U be as above. All of these groups are $B(\mathscr{A})$ groups, but $R \times U$ is not even a $B_r(\mathscr{A})$ group. The Hausdorff completion of this group is $R \times T$, which is locally compact and so a $B(\mathscr{A})$ group. Let β represent any irrational number. Then $R \times T$ has a non-trivial subgroup

$$H = \{(n, \exp 2n\pi\beta i)\colon n \in Z\},$$

which is discrete and therefore closed, and whose intersection with $R \times U$ is trivial. It follows from Theorem 1.4(b) that $R \times U$ is not a $B_r(\mathscr{A})$ group, and perforce not a $B(\mathscr{A})$ group. The same argument, applied to the product of U with the discrete group of integers, shows that this product also fails to have the $B_r(\mathscr{A})$ property.

The following example shows that certain special products retain the property, however.

EXAMPLE 2. We show that any finite power of U is a $B(\mathscr{A})$ group. Soundararajan **[13]** has called a subgroup H of a topological group G totally dense if $H \cap L$ is dense in L for every closed subgroup L of G. For Abelian groups G, this coincides with the property described in Theorem 1.4(a). Letting $\langle x \rangle$ denote the subgroup generated by an element $x \in G$, he asserts that H is totally dense in G iff $H \cap \mathrm{Cl}_G \langle x \rangle$ is dense in $\mathrm{Cl}_G \langle x \rangle$ for all $x \in G$. Since the completion T^n of U^n is compact and so a $B(\mathscr{A})$ group, it is therefore sufficient to show that

$$(\Delta) \qquad \mathrm{Cl}_{T^n} (U^n \cap \mathrm{Cl}_{T^n} \langle x \rangle) = \mathrm{Cl}_{T^n} \langle x \rangle$$

for all $x = (x_1, x_2, \cdots, x_n) \in T^n$.

We may assume without loss of generality that for some nonnegative integers r, s, the entries $x_1, \cdots, x_r$ are elements of U, that $x_{r+1}, \cdots, x_{r+s}$ are images under the exponential map of irrational numbers $\alpha_1, \cdots, \alpha_s$ which, together with 1, form a linearly independent set over Q, and that the balance of the x_i are images of linear combinations over Q of the α_j and 1. By the linear independence of $\{1, \alpha_1, \cdots, \alpha_s\}$, it follows from Theorem 443 of **[5]** that $H = \mathrm{Cl}_{T^n} \langle x \rangle = F \times T^s \times C(M)$, where F is a finite subgroup of T^r (and hence of U^r), M is an $(s+1) \times (n-r-s)$ matrix with rational entries, and

$$C(M) = \{(\exp)^{n-r-s}(k(\alpha_1, \cdots, \alpha_s, 1)M) \colon k \in Z\}.$$

It can be seen that the intersection of H with U^n is dense in H, whence (Δ) is satisfied. Hence, U^n is a $B(\mathscr{A})$ group for any positive integer n.

If G_1 is totally dense in G and $G_1 \subseteq G_2 \subseteq G$, then G_2 is totally dense in G. It then follows that $U^n \times T^m$ is totally dense in T^{n+m} for any positive integers n and m, and so is a $B(\mathscr{A})$ group. Since Stephenson [**14**, Theorem 2] has shown that totally dense subgroups of compact groups are minimal, it also follows that U^n is a minimal topological group. Since the product of a minimal group with a compact group is minimal **[14]**, if further follows that $U^n \times K$ is a minimal topological group, and so a $B_r(\mathscr{A})$ group, for any compact group K.

As to subobjects, it does not appear to be true, in general, that our two properties are inherited by closed or even closed normal subgroups. However, some partial results of a positive nature have been obtained. A subgroup H of a topological group G is said to be a *retract* of G if there is a continuous homomorphism $r \colon G \to H$ whose restriction to H is the identity. By [**6**, p. 59] and [**17**, pp. 20 and 95], H is normal and a retract of G iff there exists a subgroup H' of G such that the multiplication map $m \colon H \times H' \to G$ is an isomorphism.

PROPOSITION 1.5. *Let H be a normal subgroup of G and a retract thereof. If G is a $B(\mathscr{A})$* (resp., *$B_r(\mathscr{A})$*) *group, then H is a $B(\mathscr{A})$* (resp., *$B_r(\mathscr{A})$*) *group.*

Proof. The case for $B(\mathscr{A})$ groups follows at once, since the projection map $H \times H' \to H$ is continuous and open. Now, let G be a

$B_r(\mathcal{A})$ group, A any Hausdorff group, m as above, j the continuous inverse of m, and $f: H \to A$ a continuous, almost open, bijective homomorphism. Then $h = j(f \times id_{H'})$ is also bijective, and, since f is almost open and G is a $B_r(\mathcal{A})$ group, it follows that h is open. Hence, $mh = f \times id_{H'}$ is open, whence f is open.

The following two lemmas are required to establish the next permanence property, and will also be used extensively in §2.

DEFINITION. *If $f: G \to H$ is a homomorphism* (not necessarily continuous) *of topological groups, let* $\mathcal{V}_f = \{V^* = Vf[\mathrm{Cl}_G f^{-1}(V)]: V \in \mathcal{V}(H)\}$.

LEMMA 1.6. *Let $f: (G, u) \to (H, v)$ be a homomorphism. If $\mathcal{V}_f$ is a subbasis for the unit neighbourhood filter of a group topology v_f and the graph of f is closed in $G \times H$, then (H, v_f) is a Hausdorff topological group.*

Husain defined a topology w related to v_f in **[9]**, and established a similar result concerning it. The proof of Lemma 1.6 follows in a manner parallel to his, with certain simplifications arising from the elimination of one closure operator from the definition of the sets in the unit neighbourhood basis.

DEFINITION. *For a subgroup K of a group G, let* $\mathrm{Cent}_G K$ *denote the centralizer of K in G, and* $S_G(K) = K\,\mathrm{Cent}_G K$.

LEMMA 1.7. (a) *If $f(G)$ is dense in H, then $\mathcal{V}_f$ is a subbasis for the unit neighbourhood filter of a group topology v_f, and* $\mathcal{W}_f = \{\mathrm{Cl}_H f[\mathrm{Cl}_G f^{-1}(V)] = \hat{V}: V \in \mathcal{V}(H)\}$ *is also a subbasis of unit neighbourhoods for v_f.*

(b) *If K is an open subgroup of a topological group (G, u) such that $S_G(K)$ is dense in G and w is a group topology on K, then $\mathcal{V}(w)$ is also a fundamental system of unit neighbourhoods for a group topology on G.*

Proof. (a) Clearly $\hat{V} \subseteq V^*$, for every $V \in \mathcal{V}(H)$. Now $f[f^{-1}(V)] = V \cap f(G)$, which is dense in V, and so $V \subseteq \hat{V}$. Now, if $V_1^2 \subseteq V$, then

$$V_1^* = V_1 f[\mathrm{Cl}_G f^{-1}(V_1)] \subseteq V_1 \hat{V}_1 \subseteq (\hat{V}_1)^2 \subseteq \hat{V}.$$

Thus, $\mathcal{W}_f$ generates the same filter as $\mathcal{V}_f$.

To show that v_f is a group topology, we show that $\mathcal{W}_f$ satisfies (GV'_{I})–(GV'_{III}) of **[1**, p. 222–3**]**. The first two follow immediately, since $(\hat{V})^2 \subseteq (V^2)^\wedge$ and $(V^{-1})^\wedge = (\hat{V})^{-1}$. As for (GV'_{III}), let $V, V_1 \in \mathcal{V}(H)$, $a \in H$, $t \in G$ such that $V_1 = V_1^{-1}$, $V_1^3 \subseteq V$, and $a \in V_1 f(t)$. Then

$$a[f(t)^{-1}V_1f(t)]^\wedge a^{-1} \subseteq V_1f(t)\cdot f(t)^{-1}V_1f(t)\cdot f(t)^{-1}V_1$$
$$= V_1\hat{V}_1V_1 \subseteq (\hat{V}_1)^3 \subseteq (V_1^3)^\wedge \subseteq \hat{V}.$$

(b) It is necessary only to show $\mathscr{V}(w)$ satisfies (GV'_{III}). *Let* $U, U_1 \in \mathscr{V}(G)$, $x \in G$, $k \in K$, $c \in \mathrm{Cent}_G K$ such that $U_1 = U_1^{-1}$, $U_1^3 \subseteq U$ and $x \in U_1ck$. Then

$$x(k^{-1}U_1k)x^{-1} \subseteq U_1ck(k^{-1}U_1k)k^{-1}c^{-1}U_1 = U_1(cU_1c^{-1})U_1 = U_1^3 \subseteq U.$$

PROPOSITION 1.8. *Let $\mathscr{C}$ be a category of Hausdorff groups such that every $B_r(\mathscr{C})$ group is in $\mathscr{C}$. Then every closed central subgroup of a $B_r(\mathscr{C})$* (resp., *$B(\mathscr{C})$*) *group is a $B_r(\mathscr{C})$* (resp., *$B(\mathscr{C})$*) *group.*

Proof. Let (G, u) be a $B_r(\mathscr{C})$ group, K a closed central subgroup of G, w any group topology on K such that $(K, w) \in \mathscr{C}$, $w \subseteq u \mid K$ and $w(u \mid K) = w$. Since K is central in G, it is routine to show that $\mathscr{V}(v) = \{UW\colon U \in \mathscr{V}(u),\ W \in \mathscr{V}(w)\}$ is a subbasis for the unit neighbourhood filter of a group topology v on G. The graph of the natural injection $K \to G$ is closed in $(K, w) \times (K, w)$, and so in $(K, w) \times (G, u)$, since K is u-closed. From this and the fact that $w(u \mid K) = w$, it follows from Lemma 1.6 that v is Hausdorff. By Proposition 31.8 of **[8]**, one has $v \subseteq v(u)$. However, if $U, U_1 \in \mathscr{V}(u)$ such that $U_1^2 \subseteq U$, then $\mathrm{Cl}_v U \subseteq (\mathrm{Cl}_v U_1)^2 \subseteq U_1 \mathrm{Cl}_w (U_1 \cap K) \in \mathscr{V}(v)$. Hence, $v(u) = v$, and the identity map $(G, u) \to (G, v)$ is continuous and almost open. Then (G, v) is a $B_r(\mathscr{C})$ group, and so is in $\mathscr{C}$. Hence, $v = u$, whence $u \mid K = v \mid K = w$. Therefore, K is a $B_r(\mathscr{C})$ group.

The case for $B(\mathscr{C})$ groups is proved in a similar fashion, letting $H \in \mathscr{N}(K)$, w a topology on K such that $(K/H, w/H) \in \mathscr{C}$, $w \subseteq (u \mid K)H$ and $w[(u \mid K)H] = w$. One can then define v_H by $\mathscr{V}(v_H) = \{UW\colon U \in \mathscr{V}(H),\ W \in \mathscr{V}(w)\}$ and show that $v_H = uH$, whence $uH \mid K = v_H \mid K = w$, and $(K, u \mid K)$ is a $B(\mathscr{C})$ group.

PROPOSITION 1.9. *Let $\mathscr{C}$ be as in Proposition* 1.8, *(G, u) a $B_r(\mathscr{C})$ group* (resp., *$B(\mathscr{C})$ group*) *with equal left and right uniform structures. Then any closed subgroup K of G such that $S_G(K)$ is dense in G is a $B_r(\mathscr{C})$* (resp., *$B(\mathscr{C})$*) *group.*

Proof. Let the topologies $u \mid K$, w and v be as in Proposition 1.8. Without loss of generality, we may assume that an element of $\mathscr{V}(u)$ is fixed under all inner automorphisms of G [**7**, p. 22]. For such a neighbourhood U and any $A \subseteq G$, we then have $AU = UA$. It is then easy to see that $\mathscr{V}(v)$ satisfies (GV'_{I}) and (GV'_{II}) of [**1**, p. 222–3]. To see that (GV'_{III}) is also satisfied, let $x \in G$, and $UW, U_1W_1 \in \mathscr{V}(v)$ such that U_1, W_1 are symmetric and $(U_1W_1)^3 \subseteq UW$. Then there exist elements

a, b of K and of its centralizer, respectively, such that $x \in U_1 W_1 ab$. Then

$$\begin{aligned} xU_1(a^{-1}W_1a)x &\subseteq U_1W_1abU_1(a^{-1}W_1a)(b^{-1}a^{-1}U_1W_1) \\ &= U_1W_1U_1bW_1b^{-1}U_1W_1 = (U_1W_1)^3 \subseteq UW. \end{aligned}$$

As in Proposition 1.8, it follows that v is Hausdorff and that $v(u) = v$, whence (G, v) is a $B_r(\mathscr{C})$ group and so is in $\mathscr{C}$. Then $v = u$, and so $u \mid K = v \mid K = w$. Hence, $(K, u \mid K)$ is a $B_r(\mathscr{C})$ group.

Now let (G, u) be a $B(\mathscr{C})$ group, then choose H and w and construct v_H as in the analogous case of Proposition 1.8. Since (G, u) has equal uniformities and H is normal in K, for $UW \in \mathscr{V}(v_H)$, $U_1U_1^{-1} \subseteq U$, $W_1W_1^{-1} \subseteq W$, we have $(U_1W_1)(U_1W_1)^{-1} \subseteq UW$. The continuity of the conjugation maps follows in a manner similar to the $B_r(\mathscr{C})$ case. It then follows that $v_H = uH$, as in Proposition 1.8.

REMARKS. (i) A closed subgroup K of G such that $S_G(K)$ is dense in G is necessarily normal in G, since $S_G(K)$ is a subgroup of the normalizer of K in G, and the normalizer is closed **[4]**.

(ii) The condition that G have equal uniformities can be replaced by the slightly weaker condition that G has a fundamental system of unit neighbourhoods fixed under all conjugations by elements of K.

(iii) Clearly, $\mathscr{A}$ satisfies the condition imposed on $\mathscr{C}$ in Propositions 1.8 and 1.9. Indeed, this condition is satisfied by any category right fitting with respect to isomorphisms, if one were to modify the definition of $B(\mathscr{C})$ (resp., $B_r(\mathscr{C})$) groups to require the existence of at least one continuous, almost open (resp., and one-to-one) homomorphism onto a group in $\mathscr{C}$, thus precluding a vacuous satisfaction of the definition from **[8]**.

Let $\mathscr{E}$ denote the class of morphisms in $\mathscr{A}$ which are almost open.

PROPOSITION 1.10. *Let $\mathscr{C}$ be such that either* (i) *every $B_r(\mathscr{C})$ group is in $\mathscr{C}$, or* (ii) *$\mathscr{C}$ is right fitting with respect to $\mathscr{E}$. Then any open subgroup K of a $B_r(\mathscr{C})$* (resp., *$B(\mathscr{C})$*) *group G such that $S_G(K)$ is dense in G is a $B_r(\mathscr{C})$* (resp., *$B(\mathscr{C})$*) *group.*

Proof. Under either condition, let w be a group topology on K such that $(K, w) \in \mathscr{C}$, $w \subseteq u \mid K$ and $w(u \mid K) = w$. Let v be the topology on G having as its unit neighbourhood filter $\mathscr{V}(v) = \{UW\colon U \in \mathscr{V}(u),\ W \in \mathscr{V}(w)\}$. Let $j\colon (K, w) \to (G, u)$ be the natural injection. By our assumptions on w, it follows that v induces the v_j topology on K, and, by Lemma 1.7 (a) and (b), $\mathscr{V}(v)$ generates a group topology on G. As in Proposition 1.8, we have $v \subseteq u$ and $v(u) = v$.

Now, if $\mathscr{C}$ satisfies (i), we observe that the identity map

$(G, u) \to (G, v)$ is continuous and almost open, whence (G, v) is a $B_r(\mathscr{C})$ group and so in $\mathscr{C}$. If $\mathscr{C}$ satisfies (ii), we observe that the natural injection $(K, w) \to (G, v)$ is continuous and almost open, so $(G, v) \in \mathscr{C}$.

It then follows that $v = u$, since (G, u) is a $B_r(\mathscr{C})$ group, and so $u \mid K = v \mid K = w$. Therefore, $(K, u \mid K)$ is a $B_r(\mathscr{C})$ group.

The analogous statement for the $B(\mathscr{C})$ case is proved by first letting $H \in \mathscr{N}(K)$ and w be a group topology such that $(K/H, w/H) \in \mathscr{C}$, $w \subseteq (u \mid K)H$ and $w[(u \mid K)H] = w$, and proceeding as above.

To demonstrate some more perverse properties of these categories, we now display counterexamples, concerned with join topologies, direct limits and quotients.

EXAMPLE 3. Let (R, u) denote the reals with the usual topology, g a discontinuous automorphism of the reals, and $(R, g(u))$ the reals endowed with the topology consisting of images under g of u-open sets. Then g is a homeomorphism from (R, u) to $(R, g(u))$, whence $(R, g(u))$ is locally compact and so a $B(\mathscr{A})$ group.

The identity map $j\colon (R, u \vee g(u)) \to (R, u)$ is clearly continuous, and is also almost open, since the image under g of any u-open set is u-dense in R **[7**, p. 49]. However, j is plainly not open. Hence, $(R, u \vee g(u))$ is not even a $B_r(\mathscr{A})$ group. This example also shows, of course, that the join of two locally compact group topologies is not necessarily locally compact. Thanks are due to E. Dubinsky for suggesting the above example in the latter context.

EXAMPLE 4. Let (R, d) denote the reals with discrete topology, (R, u) as in Example 3. Let $G_1 = (R, u) \times (R, d)$, $G_2 = (R, d) \times (R, u)$, and let $f\colon G_1 \to G_2$ be defined by $(x, y) \mapsto (y, x)$, and let this system be ordered by $1 < 2$. Its inductive limit in the category of topological spaces is then R^2 endowed with the topology $(u \times d) \wedge (d \times u)$. It is proved in **[12]**, however, that this is not even a group topology, although the groups involved are locally compact and hence $B(\mathscr{A})$ groups.

EXAMPLE 5. Let T be as in Example 1, G the subgroup of T consisting of those elements of squarefree order. It is shown in **[15]** that G is a $B_r(\mathscr{A})$ group which is not a $B(\mathscr{A})$ group. It then follows from Lemma 1.1 that not all quotients of G can be $B_r(\mathscr{A})$ groups. This counterexample shows that the $B_r(\mathscr{A})$ property is not divisible, and thus that the portion of Proposition 31.7 of **[8]** which refers to $B_r(\mathscr{C})$ groups is false. Gaps are thereby created in the proofs of Theorems 32.8 and 32.9 of **[8]**. A corrected version of the former appears in §2.

A sixth example, which follows, shows that, for the class $\mathscr{C}_1$ of first countable Hausdorff groups, the $B(\mathscr{A})$ groups form a proper subclass of the $B(\mathscr{C}_1)$ groups. We first observe that, since a countably compact subspace of a first countable space is closed [**3**, p. 230], it follows that a continuous, almost open homomorphism of a locally countably compact group into a first countable group is open. Therefore, every locally countably compact group, and hence every countably compact Hausdorff group, is a $B(\mathscr{C}_1)$ group.

EXAMPLE 6. Let S be any uncountable set, and let G be any compact Hausdorff group with nontrivial centre. Let $B = G^S$, and define

$$P = \{(x_\alpha): x_\alpha \neq e \text{ for at most countably many } \alpha \in S\}.$$

By [**11**, p. 127], P is countably compact and a proper dense subgroup of the group B, which is compact and so a $B_r(\mathscr{A})$ group. For each $g \in G$, let (g) denote the element (x_α) of B such that $x_\alpha = g$ for all $\alpha \in S$. It is easy to see that $H = \{(g): g \in \operatorname{Cent} G\}$ is a nontrivial closed normal subgroup of B, and that $H \cap P = \{(e)\}$. By Theorem 1.4(b) , it follows that P is not a $B_r(\mathscr{A})$, and perforce not a $B(\mathscr{A})$, group.

2. Closed graph theorems.

In [**8**], Husain announced a quite general form of the closed graph theorem for topological groups (Theorem 32.5), and drew an extensive list of corollaries therefrom. However, the proof of this theorem contained a serious flaw, acknowledged by Husain in [**9**], where he salvaged some of the results from [**8**]. In this section, we salvage more results from [**8**] by weakening the assumption of commutativity of the codomain imposed by Husain in [**9**].

Let us recall the definition of "right fitting" from §1, and agree to denote the graph of a mapping f by $R(f)$ throughout the balance of the paper.

THEOREM 2.1. *Let $\mathscr{C}$ be a category of Hausdorff groups which is right fitting with respect to $\mathscr{E}$. Let $(G, u) \in \mathscr{C}$, (H, v) be a $B_r(\mathscr{C})$ group, $f\colon G \to H$ an almost continuous, almost open homomorphism with closed graph such that $f(G)$ is dense in H. Then f is continuous.*

Proof. By Lemmas 1.7(a) and 1.6, v_f is a Hausdorff group topology. Since $v_f \subseteq v$, the identity map $j\colon (H, v) \to (H, v_f)$ is continuous. By Proposition 31.8 of [**8**], we have that $v_f \subseteq v_f(v) \subseteq v$. Letting $U, V \in \mathscr{V}(H)$ such that $V^2 \subseteq U$, we proceed as in Theorem 32.5 of [**8**] to show that $V^* \subseteq \mathrm{Cl}_{v_f} U$, and conclude that $v_f(v) = v_f$, whence j is almost open. Then $g = jf$ is almost open, and also continuous, since $g^{-1}(U) \supseteq g^{-1}(jf[\mathrm{Cl}_G f^{-1}(U)]) \supseteq \mathrm{Cl}_G f^{-1}(U)$, which is in $\mathscr{V}(G)$. Thus, $g \in$

$\mathscr{C}$, $(H, v_f) \in \mathscr{C}$, and j is open, since (H, v) is a $B_r(\mathscr{C})$ group. Hence, $v_f = v$, $g = f$, and f is continuous.

COROLLARY 2.2. *Let $\mathscr{C}$ be as in Theorem 2.1. Let $(G, u) \in \mathscr{C}$, (H, v) be a $B_r(\mathscr{C})$ group, $f: G \to H$ an almost continuous, almost open homomorphism with closed graph such that $S_H[\mathrm{Cl}_H f(G)]$ is dense in H. Then f is continuous.*

Proof. By Proposition 1.10, $K = \mathrm{Cl}_H f(G)$ is a $B_r(\mathscr{C})$ group, and it follows that the corestriction of f to K is continuous, by Theorem 2.1. Hence, f is continuous.

THEOREM 2.3. *Let $\mathscr{C}$ be any category of Hausdorff groups such that every $B_r(\mathscr{C})$ group is in $\mathscr{C}$. Let $(G, u) \in \mathscr{C}$, (H, v) be a $B_r(\mathscr{C})$ group and $f: G \to H$ an almost continuous, almost open homomorphism with closed graph such that $f(G)$ is dense in H. Then f is continuous.*

Proof. This proof parallels that of Theorem 2.1, except that the fact that (H, v_f) is in $\mathscr{C}$ is deduced by observing that $j: (H, v) \to (H, v_f)$ continuous and almost open implies (H, v_f) is a $B_r(\mathscr{C})$ group.

The next corollary follows in a manner similar to Corollary 2.2.

COROLLARY 2.4. *Let $\mathscr{C}$ be as in Theorem 2.3. Let $(G, u) \in \mathscr{C}$, (H, v) be a $B_r(\mathscr{C})$ group, $f: G \to H$ an almost continuous, almost open homomorphism with closed graph such that $S_H[\mathrm{Cl}_H f(G)]$ is dense in H. Then f is continuous.*

For categories $\mathscr{C}$ of groups which satisfy the condition of Theorem 2.3, we can remove the "almost open" hypothesis on the map at the cost of adding certain qther hypotheses. A preliminary lemma is required.

LEMMA 2.5. *If either* (i) $f(G) \subseteq \mathrm{Cent}\, H$, *or* (ii) *$H$ has equal uniformities and $S_H[f(G)]$ is dense in H, then v_f is a group topology.*

Proof. The proof in case (i) is obvious. For (ii), we once again use (GV'_{I})–(GV'_{III}) of [**1**, p. 222–3], obtaining the first two in a manner parallel to that of Proposition 1.8. To obtain (GV'_{III}), let $V^* \in \mathscr{V}_f$, and select $W^* \in \mathscr{V}_f$ such that W is symmetric and invariant under conjugations, and $(W^*)^3 \subseteq V^*$. Let $x \in H$, $t \in G$ and $b \in \mathrm{Cent}_H[f(G)]$ such that $x \in f(t)bW^*$. Then

$$xW^*x^{-1} \subseteq f(t)b(W^*)^3b^{-1}f(t)^{-1} = [f(t)W^*f(t)^{-1}]^3.$$

But $f(t)f[\mathrm{Cl}_G f^{-1}(W)]f(t)^{-1} = f(\mathrm{Cl}_G[f^{-1}(f(t)Wf(t)^{-1})]) = f[\mathrm{Cl}_G f^{-1}(W)]$. By virtue of this and the invariance of W under conjugations, it follows that

$$xW^*x^{-1} \subseteq [f(t)W^*f(t)^{-1}]^3 = (W^*)^3 \subseteq V^*.$$

THEOREM 2.6. *Let $\mathscr{C}$ be as in Theorem* 2.3. *Let $(G, u) \in \mathscr{C}$, (H, v) be a $B_r(\mathscr{C})$ group $f\colon G \to H$ an almost continuous homomorphism with closed graph such that either* (i) *$f(G) \subseteq \mathrm{Cent}\, H$, or* (ii) *$H$ has equal uniformities and $S_H[f(G)]$ is dense in H. Then f is continuous.*

Proof. By Lemma 2.5, v_f is a group topology in either case. As in Theorem 2.1, it follows that $v_f(v) = v$, and as in Theorem 2.3, (H, v_f) is a $B_r(\mathscr{C})$ group and so in $\mathscr{C}$. Thus, $v_f = v$, and f is continuous.

As with previous results, we point out that Theorem 2.6 holds, in particular, for $\mathscr{C} = \mathscr{A}$. The conditions on the homomorphism can be further relaxed if additional topological conditions are imposed on the groups involved.

DEFINITION. A group G is called weakly separable **[2]** if, for every $V \in \mathscr{V}(G)$, there exists a countable subset X_V of G such that $VX_V = G$. (This property clearly generalizes both separability and the Lindelöf property.)

The proof of the next lemma parallels that of Proposition 32.11(b) of **[8]**, which this result generalizes.

LEMMA 2.7. *Any homomorphism from a Hausdorff group with the Baire property to a weakly separable group is almost continuous.*

THEOREM 2.8. *Let $\mathscr{C}$ be as in Theorem* 2.3. *Let G be a group in $\mathscr{C}$ with the Baire property, H a weakly separable $B_r(\mathscr{C})$ group. Then a homomorphism $f\colon G \to H$ with closed graph is continuous if either* (i) *$f(G) \subseteq \mathrm{Cent}\, H$, or* (ii) *$H$ has equal uniformities and $S_H[\mathrm{Cl}_H f(G)]$ is dense in H.*

Proof. By Lemma 2.7, f is almost continuous. Then f is continuous by Theorem 2.6.

These considerations also allow us to prove a form of the open mapping theorem which corrects and extends Theorem 32.8 of **[8]**.

THEOREM 2.9. *Let $\mathscr{C}$ be as in Theorem* 2.3. *Let G be a $B(\mathscr{C})$ group with equal uniformities, H any Hausdorff group. Then any almost open homomorphism g of G onto H with closed graph is open.*

Proof. Let K be the kernel of g and $n\colon G \to G/K$ the quotient map. Let $g = fn$. Since $R(g)$ is closed and contains $K \times \{e_H\}$, which is normal in $G \times H$, it follows by Corollary 24.4 of **[8]** that $R(f)$ is closed.

By Proposition 30.3 of **[8]**, f is almost open, whence f^{-1} is almost continuous, f^{-1} also has closed graph, clearly. Now G/K is a $B(\mathscr{C})$ group, by Proposition 31.7 of **[8]**, so f^{-1} is continuous, by Theorem 2.6. Hence, f is open and so is g, by Proposition 30.3 of **[8]**.

Finally, let $\mathscr{D}$ denote the class of morphisms in $\mathscr{A}$ which have image dense in the codomain.

THEOREM 2.10. *Let $\mathscr{C}$ be a category of Hausdorff groups which is right fitting with respect to $\mathscr{D}$, $G \in \mathscr{C}$, (H, v) a $B_r(\mathscr{C})$ group. Then an almost continuous homomorphism $f\colon G \to (H, v)$ with closed graph is in $\mathscr{D}$ if $f(G)$ is dense in H.*

Proof. As before, we form the v_f topology and observe, by Lemmas 1.6 and 1.7(a) that v_f is a Hausdorff group topology. Letting $j\colon (H, v) \to (H, v_f)$ be the identity map, we further observe that $g = jf$ is continuous, as in Theorem 2.1. Since $v_f \subseteq v$, $g(G)$ is dense in (H, v_f), so $g \in \mathscr{D}$, and $(H, v_f) \in \mathscr{C}$. Also as in Theorem 2.1, it follows that j is continuous and almost open, whence j is open, since (H, v) is a $B_r(\mathscr{C})$ group. Therefore, $v_f = v$, f is continuous, and $f \in \mathscr{D}$.

In closing, we note that the "right fitting" properties mentioned above are by no means exotic. Among the categories of Hausdorff groups which are right fitting with respect to $\mathscr{D}$ are the compact, precompact, Abelian, connected and separable groups, and among those right fitting with respect to $\mathscr{E}$ are the locally compact, locally precompact, metrizable and locally connected groups. Groups with equal uniformities, second countable groups and Abelian profinite groups are right fitting with respect to $\mathscr{E} \cap \mathscr{D}$.

Portions of this paper were included in the author's Ph.D. thesis, written at McMaster University under the direction of Dr. Taqdir Husain. His guidance was greatly appreciated, as were the most helpful comments and suggestions of the referee. The author also gratefully acknowledges the support of the government of the Province of Ontario while he was a graduate student, and, subsequently, of the National Research Council of Canada and of the Council for Research of St. Francis Xavier University.

REFERENCES

1. N. Bourbaki, *Elements of Mathematics, General Topology, Part* 1, Addison-Wesley, Reading, 1966.

2. L. G. Brown, *Topologically complete groups*, Proc. Amer. Math. Soc., **35** (1972), 593.
3. J. Dugundji, *Topology*, Allyn and Bacon, Boston, 1966.
4. D. L. Grant, *Centralizers and normalizers in Hausdorff groups*, Math. Mag., **48** (1975), 218.
5. G. H. Hardy and E. M. Wright, *The Theory of Numbers*, 3rd Edition, Oxford, 1954.
6. E. Hewitt and K. A. Ross, *Abstract Harmonic Analysis*, Springer-Verlag, 1963.
7. E. Hewitt and K. Stromberg, *Real and Abstract Analysis*, Spring-Verlag, 1965.
8. T. Husain, *Introduction to Topological Groups*, Saunders, Philadelphia, 1966.
9. ———, *On a closed graph theorem for topological groups*, Proc. Japan Acad., **44** (1968), 446.
10. J. R. Isbell, *Uniform Spaces*, Amer. Math. Soc. Colloquium Publications, Providence, 1964.
11. L. Pontryagin, *Topological Groups*, Princeton University Press, Princeton, 1946.
12. P. Samuel, *Ultrafilters and compactifications of uniform spaces*, Trans. Amer. Math. Soc., **64** (1948), 100.
13. T. Soundararajan, *Totally dense subgroups of topological groups, in general topology and its applications to modern analysis and algebra*, Proceedings of the Kanpur Topological Conference, 1968, Academia, Prague, 1968.
14. R. M. Stephenson, Jr., *Minimal topological groups*, Math. Ann., **192** (1971), 193.
15. L. J. Sulley, *A note on B- and B_r-complete topological Abelian groups*, Proc. Camb. Phil. Soc., **66** (1969), 275.
16. C. Sunyach, *Compléments au théorème du graphe fermé*, C. R. Acad. Sci. Paris, **267** (1968), 215.
17. A. Weil, *L'integration dans les groupes topologiques et ses applications*, Actualités Sci. et Ind. **869**, **1145**, Hermann et Cie., Paris, 1941 and 1951.

Received June 24, 1974 and in revised form January 7, 1977.

COLLEGE OF CAPE BRETON,
SYDNEY, NOVA SCOTIA, CANADA

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No 2, 1977

A CHARACTERIZATION OF SOLENOIDS

CHARLES L. HAGOPIAN

Suppose M is a homogeneous continuum and every proper subcontinuum of M is an arc. Using a theorem of E. G. Effros involving topological transformation groups, we prove that M is circle-like. This answers in the affirmative a question raised by R. H. Bing. It follows from this result and a theorem of Bing that M is a solenoid. Hence a continuum is a solenoid if and only if it is homogeneous and all of its proper subcontinua are arcs. The group G of homeomorphisms of M onto M with the topology of uniform convergence has an unusual property. For each point w of M, let G_w be the isotropy subgroup of w in G. Although G_w is not a normal subgroup of G, it follows from Effros' theorem and Theorem 2 of this paper that the coset space G/G_w is a solenoid homeomorphic to M and, therefore, a topological group.

1. Introduction. Let $\mathscr{S}$ be the class of all homogeneous continua M such that every proper subcontinuum of M is an arc. It is known that every solenoid belongs to $\mathscr{S}$. It is also known that every circle-like element of $\mathscr{S}$ is a solenoid. In fact, in 1960 R. H. Bing **[4**, Theorem 9, p. 228**]** proved that each homogeneous circle-like continuum that contains an arc is a solenoid. At that time Bing **[4**, p. 219**]** asked whether every element of $\mathscr{S}$ is a solenoid. In this paper we answer Bing's question in the affirmative by proving that every element of $\mathscr{S}$ is circle-like.

2. Definitions and related results. We call a nondegenerate compact connected metric space a *continuum*.

A *chain* is a finite sequence $L_1, L_2, \cdots, L_n$ of open sets such that $L_i \cap L_j \neq \emptyset$ if and only if $|i-j| \leqq 1$. If L_1 also intersects L_n, the sequence is called a *circular chain*. Each L_i is called a *link*. A chain (circular chain) is called an ϵ-*chain* (ϵ-*circular chain*) if each of its links has diameter less than ϵ. A continuum is said to be *arc-like* (*circle-like*) if for each $\epsilon > 0$, it can be covered by an ϵ-chain (ϵ-circular chain).

A space is *homogeneous* if for each pair p, q of its points there exists a homeomorphism of the space onto itself that takes p to q. Bing **[2] [3]** proved that a continuum is a pseudo-arc if and only if it is homogeneous and arc-like. L. Fearnley **[9]** and J. T. Rogers, Jr. **[20]** independently showed that every homogeneous, hereditarily indecomposable, circle-like continuum is a pseudo-arc **[11]**. However, there are many topologically different homogeneous circle-like continua that have decomposable subcontinua **[24]** **[25]**.

Let $n_1, n_2, \cdots$ be a sequence of positive integers. For each positive integer i, let G_i be the unit circle $\{z \in R^2 : |z| = 1\}$, and let f_i be the map of G_{i+1} onto G_i defined by $f_i(z) = z^{n_i}$. The inverse limit space of the sequence $\{G_i, f_i\}$ is called a *solenoid.* Since each G_i is a topological group and each f_i is a homomorphism, every solenoid is a topological group [**13**, Theorem 6.14, p. 56] and therefore homogeneous. Each solenoid is circle-like since it is an inverse limit of circles with surjective bonding maps [**17**, Lemma 1, p. 147].

A solenoid can be described as the intersection of a sequence of solid tori $M_1, M_2, \cdots$ such that M_{i+1} runs smoothly around inside M_i exactly n_i times longitudinally without folding back and M_i has cross diameter of less than i^{-1}. The sequence $n_1, n_2, \cdots$ determines the topology of the solenoid. If it is $1, 1, \cdots$ after some place, the solenoid is a simple closed curve. If it is $2, 2, \cdots$, the solenoid is the dyadic solenoid defined by D. van Dantzig **[7]** and L. Vietoris **[23]**. Other properties involving the sequence $n_1, n_2, \cdots$ are given in [**4**, p. 210]. From this description we see that every proper subcontinuum of a solenoid is an arc.

Solenoids appear as invariant sets in the qualitative theory of differential equations. In **[21]** E. S. Thomas proved that every compact 1-dimensional metric space that is minimal under some flow and contains an almost periodic point is a solenoid.

Every homogeneous plane continuum that contains an arc is a simple closed curve **[4]** **[10]** **[15]**. Hence each planar solenoid is a simple closed curve.

Each of the three known examples of homogeneous plane continua (a circle, a pseudo-arc **[2]** **[18]**, and a circle of pseudo-arcs **[5]**) is circle-like. If one could show that every homogeneous plane continuum is circle-like, it would follow that there does not exist a fourth example **[6]** **[12]** [**14**, p. 49] and a long outstanding problem would be solved.

A *topological transformation group* (G, M) is a topological group G together with a topological space M and a continuous mapping $(g, w) \rightarrow gw$ of $G \times M$ into M such that $ew = w$ (e denotes the identity of G) and $(gh)w = g(hw)$ for all elements g, h of G and w of M.

For each point w of M, let G_w be the isotropy subgroup of w in G (that is, the set of all elements g of G such that $gw = w$). Let G/G_w be the left coset space with the quotient topology. The mapping φ_w of G/G_w onto Gw that sends gG_w to gw is one-to-one and continuous. The set Gw is called the *orbit* of w.

Assume M is a continuum and G is the topological group of homeomorphisms of M onto M with the topology of uniform convergence [**16**, p. 88]. E. G. Effros [**8**, Theorem 2.1] proved that each orbit is a set of the type G_δ in M if and only if for each point w of M, the mapping φ_w is a homeomorphism.

Suppose M is a homogeneous continuum. Then the orbit of each point of M is M, a G_δ-set. According to Effros' theorem, for each point w of M, the coset space G/G_w is homeomorphic to M. By Theorem 2 of §4, if M has the additional property that all of its proper subcontinua are arcs, then G/G_w is a solenoid and, therefore, a topological group. Note that G_w is not a normal subgroup of G.

Throughout this paper R^2 is the Cartesian plane. For each real number r, we shall denote the horizontal line $y = r$ and the vertical line $x = r$ in R^2 by $H(r)$ and $V(r)$ respectively.

Let P and Q be subsets of R^2. The set P is said to *project horizontally* into Q if every horizontal line in R^2 that meets P also meets Q.

We shall denote the boundary and the closure of a given set Z by $\mathrm{Bd}\, Z$ and $\mathrm{Cl}\, Z$ respectively.

3. Preliminary results. In this section M is a homogeneous continuum (with metric ρ) having only arcs for proper subcontinua.

Let p and q be two points of the same arc component of M. The union of all arcs in M that have p as an endpoint and contain q is called a *ray* starting at p.

The following two lemmas are easy to verify.

LEMMA 1. *Each ray is dense in M.*

LEMMA 2. *If an open subset Z of M is not dense in M, then each component of Z is an arc segment with both endpoints in* $\mathrm{Bd}\, Z$.

Let ϵ be a positive number. A homeomorphism h of M onto M is called an *ϵ-homeomorphism* if $\rho(v, h(v)) < \epsilon$ for each point v of M.

LEMMA 3. *Suppose ϵ is a given positive number and w is a point of M. Then w belongs to an open subset W of M with the following property. For each pair p, q of points of W, there exists an ϵ-homeomorphism h of M onto M such that $h(p) = q$.*

Proof. Define G, G_w, and φ_w as in §2. Since M is homogeneous, the orbit of each point of M is M. Therefore φ_w is a homeomorphism of G/G_w onto M [**8**, Theorem 2.1].

Let π_w be the natural open mapping of G onto G/G_w that sends g to gG_w. Define T_w to be the mapping of G onto M that sends g to $g(w)$. Since $T_w = \varphi_w \pi_w$, it follows that T_w is an open mapping [**22**, Theorem 3.1]. Note that the following diagram commutes.

$$\begin{tikzcd} G \arrow[rr, "T_w"] \arrow[dr, "\pi_w"'] & & M \\ & G/G_w \arrow[ur, "\varphi_w"'] & \end{tikzcd}$$

Let U be the open subset of G consisting of all $\epsilon/2$-homeomorphisms of M onto M. Define W to be the open set $T_w[U]$. Since the identity e belongs to U and $T_w(e) = w$, the set W contains w.

Assume p and q are points of W. Let f and g be elements of U such that $T_w(f) = p$ and $T_w(g) = q$. Since $f(w) = p$ and $g(w) = q$, the mapping $h = gf^{-1}$ of M onto M is an ϵ-homeomorphism with the property that $h(p) = q$.

For each positive integer i, let A_i be an arc with endpoints p_i and q_i. The sequence $A_1, A_2, \cdots$ is said to be *folded* if it converges to an arc A and the sequence $p_1, q_1, p_2, q_2, \cdots$ converges to an endpoint of A.

LEMMA 4. (Bing [**4**, Theorem 6, p. 220]). *There does not exist a folded sequence of arcs in M.*

Lemma 4 follows from a simple argument (shorter than Bing's) involving Lemma 3 and the fact that M does not contain a triod.

A chain $L_1, L_2, \cdots, L_n$ in M is said to be *free* if $\operatorname{Cl} L_1 \cap \operatorname{Cl} L_n = \varnothing$ and $\operatorname{Bd} \cup \{L_i : 1 \leqq i \leqq n\}$ is a subset of $\operatorname{Cl}(L_1 \cup L_n)$.

LEMMA 5. (Bing [**4**, Property 17, p. 219]). *Let A be an arc in M with endpoints p and q. For each positive number ϵ, there exists a free ϵ-chain $L_1, L_2, \cdots, L_n$ in M covering A such that p and q belong to L_1 and L_n respectively.*

A continuum is *decomposable* if it is the union of two proper subcontinua; otherwise, it is *indecomposable.*

LEMMA 6. *If M is decomposable, then M is a simple closed curve.*

Proof. Since M is the union of two proper subcontinua (arcs), M is locally connected. Since M is homogeneous, it does not have a separating point. Hence M contains a simple closed curve [**19**, Theorem 13, p. 91]. It follows that M is a simple closed curve.

4. Principal results.

THEOREM 1. *If M is a homogeneous continuum and every proper subcontinuum of M is an arc, then M is circle-like.*

Proof. According to Lemma 6, if M is decomposable, then M is a simple closed curve and therefore circle-like. Hence we assume that M is indecomposable.

By Lemmas 4 and 5, there exists a free chain $L_1, L_2, \cdots, L_\alpha$ $(\alpha > 5)$ in M such that $N = \mathrm{Cl} \cup \{L_i : 1 \leqq i \leqq \alpha\}$ is a proper subset of M and $N - \mathrm{Cl} \cup \{L_i : 3 \leqq i \leqq \alpha - 2\}$ contains every arc in N that has both of its endpoints in $\mathrm{Cl}\, L_1$ or $\mathrm{Cl}\, L_\alpha$. (This chain is formed from another free chain by unioning links to make L_2 and $L_{\alpha-1}$ sufficiently long and narrow.) Let B be the union of all components of N that meet $\mathrm{Cl}\,(L_3 \cup L_{\alpha-2})$. By Lemma 2, each component of B is an arc with one endpoint in $\mathrm{Bd}\, L_1$ and the other endpoint in $\mathrm{Bd}\, L_\alpha$. Note that B is a closed set. Since M is indecomposable, each component of B is a continuum of condensation.

Since B contains no folded sequence of arcs, we can assume that B is the intersection of M and the plane R^2 and that the following conditions are satisfied:

I. A component C of B is $\{(x, y) : 0 \leqq x \leqq 6 \text{ and } y = 0\}$.

II. Each component of $B - C$ is a horizontal interval above $H(0)$ (the x-axis) and below $H(1)$ that crosses both $V(1)$ and $V(5)$.

III. The sets $\mathrm{Cl}\,(L_1 \cup L_2 \cup L_{\alpha-1} \cup L_\alpha)$ and $\{(x, y) : 1 \leqq x \leqq 5\}$ are disjoint.

(Bing's theorem [**2**, Theorem 11], involving sequences of refining covers that induce a homeomorphism, can be used to define this embedding of B in R^2. Each cover of B consists of finitely many free chains that correspond to disjoint straight horizontal chains with rectangular links in R^2.) Note that $B \cap \{(x, y) : 1 < x < 5\}$ is an open subset of M.

Let ρ be a metric on M whose restriction to B agrees with the Euclidean metric on R^2 [**1**, Theorems 4 and 5].

There exists a positive number d less than 1 such that $M \cap H(d) = \varnothing$ and the following condition is satisfied:

Property 1. Every arc in M that has its endpoints in $\{(x, y) : x = 3 \text{ and } 0 \leqq y < d\}$ meets both $\{(x, y) : x = 1 \text{ and } 0 \leqq y < d\}$ and $\{(x, y) : x = 5 \text{ and } 0 \leqq y < d\}$.

To see this we assume Property 1 does not hold for any positive number d. For each positive integer i, let W_i be an open set in

$M \cap \{(x, y): 1 < x < 5\}$ that contains (3,0) such that for each pair p, q of points of W_i, there exists an i^{-1}-homeomorphism of M onto M that takes p to q (Lemma 3). For each i, there exists an arc A_i in M with endpoints p_i and q_i in $W_i \cap V(3)$ such that the horizontal interval Γ_i from p_i to $V(1)$ is in A_i if and only if the horizontal interval Δ_i from q_i to $V(1)$ is in A_i.

For each i, let h_i be an i^{-1}-homeomorphism of M onto M such that $h_i(p_i) = q_i$. Since each h_i maps Γ_i approximately onto Δ_i, for each i, there exists a point a_i of A_i such that $h_i(a_i) = a_i$.

For each i, let B_i be the arc in A_i from p_i to a_i. Note that for each i, the diameter of B_i is greater than 1 and $B_i \cap h_i[B_i]$ consists of the point a_i.

Let a be a limit point of the sequence $\{a_i\}$. Assume without loss of generality that $\{a_i\}$ is a convergent sequence in $E = \{v \in M : \rho(v, a) < 1/2\}$.

For each i, let E_i be an arc in $B_i \cap \mathrm{Cl}\, E$ that goes from a point b_i of $\mathrm{Bd}\, E$ to a_i. Assume without loss of generality that $\{b_i\}$ converges to a point of $\mathrm{Bd}\, E$ and $\{E_i\}$ converges to an arc F in $\mathrm{Cl}\, E$. Since each h_i is an i^{-1}-homeomorphism, $\{E_i \cup h_i[E_i]\}$ is a folded sequence of arcs converging to F. This contradiction of Lemma 4 completes our argument for Property 1.

For $i = 1$ and 2, let

$$D_i = M \cap \{(x, y): i \leqq x \leqq 6 - i \quad \text{and} \quad 0 \leqq y < d\}.$$

Let ϵ be a given positive number less than $\rho(D_2, M - D_1)$. We shall complete this proof by defining an ϵ-circular chain that covers M.

By Lemma 1, there exists an arc A in M that is irreducible with respect to the property that it contains $\{(5,0), (6,0)\}$ and intersects $\{(x, y): x = 5 \text{ and } 0 < y < d\}$. According to Property 1, A intersects $\{(x, y): x = 4 \text{ and } 0 < y < d\}$.

Let W be an open set in $D_1 - A$ containing (4,0) such that for each pair p, q of points of W, there exists an $\epsilon/50$-homeomorphism of M onto M that takes p to q (Lemma 3).

Let c be a number $(0 < c < \epsilon/50)$ such that $M \cap H(c) = \varnothing$ and $M \cap \{(x, y): x = 4 \text{ and } 0 \leqq y < c\}$ is in W. Since W and A are disjoint, c is less than d.

For $i = 1$ and 2, let

$$C_i = M \cap \{(x, y): i \leqq x \leqq 6 - i \quad \text{and} \quad 0 \leqq y < c\}.$$

Let δ be the minimum of ϵ and $\rho(C_2, M - C_1)$. Let U be an open subset of C_1 containing (2,0) such that for each point q of U, there exists a δ-homeomorphism of M onto M that takes (2,0) to q (Lemma 3).

Define S to be the ray in M that starts at (2,0) and contains A. Let $\{s_i\}$ be the sequence consisting of all points of $S \cap \{(x, y): x = 3$ and $0 \leqq y < d\}$ and having the property that for each i, the points s_i precedes s_{i+1} with respect to the linear order on S.

Define T_1 to be an arc containing A in S that starts at the point $t_1 = (2,0)$ and ends at a point t_2 of $U \cap V(2)$. Let h be a δ-homeomorphism of M onto M that takes t_1 to t_2.

We proceed inductively. Assume an arc T_n is defined in S with endpoints t_n and t_{n+1} in $C_2 \cap V(2)$. Let y be the number such that $h(t_{n+1})$ belongs to $H(y)$. Define T_{n+1} to be the arc in S with endpoints t_{n+1} and $t_{n+2} = (2, y)$. Since h is a δ-homeomorphism, t_{n+2} belongs to C_2. Note that since each T_n has diameter greater than 1, the ray S is the union of $\{T_n: n = 1, 2, \cdots\}$.

Define β to be the largest integer such that $\{s_i : 1 \leqq i \leqq \beta\}$ is a subset of T_1. The δ-homeomorphism h maps each T_n approximately onto T_{n+1}. Hence, for each n, the arc T_n contains $\{s_i : (n-1)\beta < i \leqq n\beta\}$. Furthermore, β has the following property:

Property 2. For each positive integer i, the point s_i belongs to C_2 if and only if $s_{i+\beta}$ belongs to C_2.

Define γ to be the least positive integer that has Property 2. Note that since s_2 does not belong to C_2, the integer γ is greater than 1.

Let K be $\{s_i : i = j\gamma + 1$ and $j = 0, 1, 2, \cdots\}$, and let L be $(S \cap D_2 \cap V(3)) - K$.

Property 3. The sets $\mathrm{Cl}\, K$ and $\mathrm{Cl}\, L$ are disjoint.

To establish Property 3, we assume there is a point z in $\mathrm{Cl}\, K \cap \mathrm{Cl}\, L$. Let Z be an open subset of M containing z such that for each pair p, q of points of Z, there exists a δ-homeomorphism of M onto M that takes p to q (Lemma 3).

Let s_i and s_n be points of $Z \cap K$ and $Z \cap L$, respectively, and let f be a δ-homeomorphism of M onto M such that $f(s_i) = s_n$. Let θ be the smallest positive integer such that $s_{n-\theta}$ belongs to K. The existence of f implies that θ has Property 2. Since θ is less than γ, this is a contradiction and Property 3 is established.

Note that since $M = \mathrm{Cl}\, S$ (Lemma 1), $\mathrm{Cl}(K \cup L) = D_2 \cap V(3)$.

Let I be the arc in S that goes from s_1 to $s_{\gamma+1}$. By an argument similar to Bing's [**4**, Property 17, p. 219], there exists a free $\epsilon/50$-chain $P_1, P_2, \cdots, P_\lambda$ in M covering I such that

(i) s_1 and $s_{\gamma+1}$ belong to P_1 and P_λ respectively,

(ii) $P_1 \cup P_\lambda$ is in C_2,

(iii) each component of $H = \bigcup \{P_j : 1 \leqq j \leqq \lambda\}$ that meets $\mathrm{Cl}\, P_1$ also meets P_1 and $V(5)$, and

(iv) each component of H that meets $\mathrm{Cl}\, P_\lambda$ meets P_λ and $V(1)$.

From Property 1 we get the following:

Property 4. Each component of H meets both P_1 and P_λ.

Let P_μ be an element of $P_1, P_2, \cdots, P_\lambda$ that contains the point (4,0). Since W intersects each component of C_2, there exists a finite sequence $g_1, g_2, \cdots, g_\sigma$ of $\epsilon/50$-homeomorphisms of M onto M such that $\mathrm{Cl}\, K$ projects horizontally into $\bigcup \{g_i[P_\mu] : 1 \leqq i \leqq \sigma\}$. Assume without loss of generality that no proper subsequence of $g_1, g_2, \cdots, g_\sigma$ has this horizontal projection property.

Note that each $g_i[P_\mu]$ is a subset of D_1.

From Properties 1 and 4 we get the following:

Property 5. For each i $(1 \leqq i \leqq \sigma)$, if T is a component of $g_i[H]$, then $T \cap g_i[\mathrm{Cl}\, P_\mu]$ is a nonempty set that projects horizontally to a point of $D_2 \cap V(3)$.

For each i $(1 \leqq i \leqq \sigma)$, let X_i be the set consisting of all points in $g_i[P_\mu]$ that project horizontally into $\mathrm{Cl}\, K$, and let Y_i be the union of all components of $g_i[H]$ that meet X_i.

For each i $(1 \leqq i \leqq \sigma)$, the set Y_i is open in M. To see this assume that for some i, a point u of Y_i is in $\mathrm{Cl}\,(M - Y_i)$. According to Property 3, u does not belong to $g_i[P_\mu]$. By Property 5, there exists a sequence $\{J_n\}$ of arcs in $g_i[H]$ that meet $g_i[P_\mu]$ such that the limit superior J of $\{J_n\}$ is an arc in $g_i[H]$ that contains u and for each n, the set $J_n \cap g_i[P_\mu]$ projects horizontally to a point of $\mathrm{Cl}\, L$. It follows that $J \cap g_i[\mathrm{Cl}\, P_\mu]$ is a nonempty set that projects horizontally to a point of $\mathrm{Cl}\, L$. Since J is in the u-component of Y_i, this is a contradiction of Property 5. Hence Y_i is an open subset of M.

For each i $(1 \leqq i \leqq \sigma)$ and j $(1 \leqq j \leqq \lambda)$, let $Q_{i,j} = Y_i \cap g_i[P_j]$. It follows from an argument similar to the one given in the preceding paragraph that for each i, the set $\mathrm{Cl}\,(Q_{i,1} \cup Q_{i,\lambda})$ contains $\mathrm{Bd} \bigcup \{Q_{i,j} : 1 \leqq j \leqq \lambda\}$. Hence, for each i, the sequence $Q_{i,1}, Q_{i,2}, \cdots, Q_{i,\lambda}$ is a free chain in M.

Property 6. For each i $(1 \leqq i \leqq \sigma)$, the set $Q_{i,1} \cup Q_{i,\lambda}$ projects horizontally into $\mathrm{Cl}\, K$.

Obviously, $Q_{i,1}$ projects horizontally into $\mathrm{Cl}\, K$. Therefore, to establish Property 6, we assume there is a point t of $Q_{i,\lambda}$ that projects horizontally into $\mathrm{Cl}\, L$. By Property 3, there exists a positive number η less than ϵ such that $Q = \{v \in M : \rho(v, t) < \eta\}$ projects horizontally in $\mathrm{Cl}\, L$.

Let T denote the t-component of Y_i, and let w be a point of $T \cap Q_{i,1}$ (Property 4). Since g_i is an $\epsilon/50$-homeomorphism, T crosses $D_1 \cap V(1)$ exactly γ times (Property 1). Since w belongs to $Q_{i,1}$, it projects horizontally into $\mathrm{Cl}\, K$.

By Lemma 3, there exists an η-homeomorphism g of M onto M such that $g(w)$ belongs to $Q_{i,1}$ and projects horizontally into K. Since the $g(w)$-component of Y_i is an arc segment in S that crosses $D_1 \cap V(1)$ exactly γ times and is mapped approximately onto T by g^{-1}, the point $g(t)$ of Q projects horizontally into K. This contradiction of the definition of Q completes our argument for Property 6.

Let π be an integer $(5 < \pi < \mu)$ such that P_π contains the point $(3 + \epsilon/10, 0)$. Let ω be an integer $(\mu < \omega < \lambda - 4)$ such that P_ω contains the point of $V(3 - \epsilon/10)$ that projects horizontally to $s_{\gamma+1}$.

Property 7. For each n $(1 \leqq n \leqq \sigma)$, the set $Q_{n,1} \cup Q_{n,\lambda}$ does not intersect $\cup\{Q_{i,j} : 1 \leqq i \leqq \sigma \text{ and } \pi \leqq j \leqq \omega\}$.

To see this assume there exist integers i, j, and n such that $\pi \leqq j \leqq \omega$ and a point p belongs to $Q_{i,j} \cap (Q_{n,1} \cup Q_{n,\lambda})$. According to Property 6, $\{p\} \cup Q_{i,1} \cup Q_{i,\lambda}$ projects horizontally into $\mathrm{Cl}\, K$. By Property 3, there exists a positive number χ less than ϵ such that $\{v \in M : \rho(v, p) < \chi\}$ projects horizontally into $\mathrm{Cl}\, K$.

Let P be the p-component of Y_i. Let Y be an arc in P that goes from a point q of $Q_{i,1}$ to p. Since g_i and g_n are $\epsilon/50$-homeomorphisms and $\pi \leqq j \leqq \omega$, the set $Q_{i,1} \cup Q_{i,\lambda}$ and the p-component of $P \cap D_1$ are disjoint. Hence Y crosses $D_1 \cap V(1)$ exactly ι times where ι is a positive integer less than γ.

By Lemma 3, there exists a χ-homeomorphism k of M onto M such that $k(q)$ belongs to $Q_{i,1}$ and projects horizontally into K. The arc $k[Y]$ crosses $D_1 \cap V(1)$ exactly ι times. Since $k[Y]$ is in S and $\rho(p, k(p)) < \chi$, the point $k(p)$ projects horizontally into K. It follows from the definition of K that ι is a multiple of γ, and this is a contradiction. Hence Property 7 holds.

For each i $(1 \leqq i \leqq \sigma)$ and j $(1 \leqq j \leqq \lambda)$, let $P_{i,j} = Q_{i,j} - \mathrm{Cl} \cup \{Y_n : 1 \leqq n < i\}$. By Property 7, for each i, the subchain of $P_{i,1}, P_{i,2}, \cdots, P_{i,\lambda}$ that has $P_{i,\pi}$ and $P_{i,\omega}$ as end links is free in M.

For each j $(1 \leqq j \leqq \lambda)$, let $U_j = \cup \{P_{i,j} : 1 \leqq i \leqq \sigma\}$. The subchain $\mathscr{C}$ of $U_1, U_2, \cdots, U_\lambda$ that has U_π and U_ω as end links is a free $\epsilon/16$-chain in M.

Let D be the union of all components of $C_2 \cap \{(x, y) : 3 - \epsilon/5 < x < 3 + \epsilon/5\}$ that meet $\mathrm{Cl}\, K$. According to Property 3, D is open in M. The diameter of D is less than $\epsilon/2$. Each point of $U_\pi \cup U_\omega$ is within $\epsilon/5$ of $V(3)$. By Property 6, $U_\pi \cup U_\omega$ projects horizontally into $\mathrm{Cl}\, K$. Hence $U_\pi \cup U_\omega$ is in D.

Let τ be the largest integer less than μ such that U_τ intersects

D. Let ψ be the smallest integer greater than μ such that U_ψ intersects D. For each j $(1 \leqq j < \psi - \tau)$, let $Z_j = U_{\tau+j}$. Note that $Z_1, Z_2, \cdots, Z_{\psi-\tau-1}$ is a free ϵ-chain in M.

Define $Z_{\psi-\tau}$ to be the union of D and all elements of $\mathcal{D} = \{U_j : \pi \leqq j \leqq \tau \text{ or } \psi \leqq j \leqq \omega\}$. Since $\mathrm{Cl}\, K$ projects horizontally into U_μ and $\mathcal{C}$ is a free chain in M, each element of $\mathcal{D}$ intersects D. Thus $Z_{\psi-\tau}$ is an open set in M of diameter less than ϵ. Note that $Z_{\psi-\tau}$ meets both Z_1 and $Z_{\psi-\tau-1}$.

Since $\mathcal{C}$ is free and $U_\pi \cup U_\omega$ is in D, the boundary of $\cup \{Z_j : 1 \leqq j < \psi - \tau\}$ is in $Z_{\psi-\tau}$. Since $\mathrm{Cl}\, K$ projects horizontally into U_μ, the set Z_1 contains every boundary point of $Z_{\psi-\tau}$ that is to the right of $V(3)$ in R^2.

Furthermore, each point of $\mathrm{Bd}\, Z_{\psi-\tau}$ that is to the left of $V(3)$ is in $Z_{\psi-\tau-1}$. To see this let s be such a point. Let X be the arc in M that intersects $V(1)$ and is irreducible between s and $\mathrm{Cl}\, U_\mu$ (Lemma 1). By Property 1, X does not meet $U_\pi \cup U_\omega$. Since U_μ is an interior link in the free chain $\mathcal{C}$, the arc X is covered by $\mathcal{C}$ and s belongs to $Z_{\psi-\tau-1}$.

It follows that $\mathrm{Bd}\, Z_{\psi-\tau}$ is in $Z_1 \cup Z_{\psi-\tau-1}$. Therefore $Z_1, Z_2, \cdots, Z_{\psi-\tau}$ is an ϵ-circular chain that covers M. Hence M is circle-like.

Since every homogeneous circle-like continuum that contains an arc is a solenoid [**4**, Theorem 9, p. 228], Theorem 1 implies the following:

THEOREM 2. *A continuum M is a solenoid if and only if M is homogeneous and every proper subcontinuum of M is an arc.*

REFERENCES

1. R. H. Bing, *Extending a metric*, Duke Math. J., **14** (1947), 511–519.
2. ———, *A homogeneous indecomposable plane continuum*, Duke Math. J., **15** (1948), 729–742.
3. ———, *Each homogeneous nondegenerate chainable continuum is a pseudo-arc*, Proc. Amer. Math. Soc., **10** (1959), 345–346.
4. ———, *A simple closed curve is the only homogeneous bounded plane continuum that contains an arc*, Canad. J. Math., **12** (1960), 209–230.
5. R. H. Bing and F. B. Jones, *Another homogeneous plane continuum*, Trans. Amer. Math. Soc., **90** (1959), 171–192.
6. C. E. Burgess, *A characterization of homogeneous plane continua that are circularly chainable*, Bull. Amer. Math. Soc., **75** (1969), 1354–1355.
7. D. van Dantzig, *Ueber topologisch homogene Kontinua*, Fund. Math., **15** (1930), 102–125.
8. E. G. Effros, *Transformation groups and C^*-algebras*, Ann. of Math., **81** (1965), 38–55.
9. L. Fearnley, *The pseudo-circle is not homogeneous*, Bull. Amer. Math. Soc., **75** (1969), 554–558.
10. C. L. Hagopian, *Homogeneous plane continua*, Houston J. Math., **1** (1975), 35–41.
11. ———, *The fixed-point property for almost chainable homogeneous continua*, Illinois J. Math., **20** (1976), 650–652.
12. ———, *Indecomposable homogeneous plane continua are hereditarily indecomposable*, Trans. Amer. Math. Soc., **224** (1976), 339–350.

13. E. Hewitt and K. A. Ross, *Abstract harmonic analysis*, Volume 1, Academic Press, New York, 1963.

14. F. B. Jones, *Homogeneous plane continua*, Proceedings of the Auburn Topology Conference, Auburn Univ., Auburn, Ala., (1969), pp. 46–56.

15. ———, *Use of a new technique in homogeneous continua*, Houston J. Math., **1** (1975), 57–61.

16. K. Kuratowski, *Topology*, Volume 2, 3rd ed., Monografie Mat., Tom 21, PWN, Warsaw, (1961); English transl., Academic Press, New York; PWN, Warsaw, 1968.

17. S. Mardesic and J. Segal, *ϵ-mappings onto polyhedra*, Trans. Amer. Math. Soc., **109** (1963), 146–164.

18. E. E. Moise, *A note on the pseudo-arc*, Trans. Amer. Math. Soc., **67** (1949), 57–58.

19. R. L. Moore, *Foundations of point set theory*, rev. ed., Amer. Math. Soc. Colloq. Publ., vol. **13**, Amer. Math. Soc., Providence, R. I., 1962.

20. J. T. Rogers, Jr., *The pseudo-circle is not homogeneous*, Trans. Amer. Math. Soc., **148** (1970), 417–428.

21. E. S. Thomas, *One-dimensional minimal sets*, Topology, **12** (1973), 233–242.

22. G. S. Ungar, *On all kinds of homogeneous spaces*, Trans. Amer. Math. Soc., **212** (1975), 393–400.

23. L. Vietoris, *Ueber den höheren Zusummenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen*, Math. Ann. **97** (1927), 454–472.

24. C. L. Hagopian and J. T. Rogers, Jr., *A classification of homogeneous, circle-like continua*, Houston J. Math., to appear.

25. J. T. Rogers, Jr., *Solenoids of pseudo-arcs*, Houston J. Math., to appear.

Received April 2, 1976 and in revised form February 15, 1977.

California State University
Sacramento, CA 95819

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No. 2, 1977

ON COMPLETENESS OF THE BERGMAN METRIC AND ITS SUBORDINATE METRICS, II

KYONG T. HAHN

Let M be a complex manifold of dimension n furnished with both the Bergman metric and the Carathéodory distance. The main result of the present paper is to prove that the Bergman metric is always greater than or equal to the Carathéodory distance on M. The case where M is a bounded domain in the space C^n was already considered by the author in Proc. Nat. Acad. Sci. (U.S.A.), 73 (1976), 4294.

1. Introduction. The main purpose of the present paper is to prove the following

THEOREM A. *Let M be a complex manifold which admits both the Bergman metric s_M and the Carathéodory differential metric α_M. For each $z \in M$ and each holomorphic tangent vector ξ,*

$$\alpha_M(z, \xi) \leqq s_M(z, \xi). \tag{1}$$

Let ρ_M and d_M denote the integrated metrics on M which are induced from α_M and s_M, respectively. Then the Carathéodory distance c_M ([2]) satisfies

$$c_M \leqq \rho_M \leqq d_M \tag{2}$$

and there are cases when ρ_M differs from c_M and d_M.

From this observation and Theorem A, we obtain

THEOREM B. *Let M be a complex manifold given as in Theorem A. Then the Bergman metric is complete in M whenever the Carathéodory distance is complete.*

If in particular M is a bounded domain in the complex Euclidean space C^n $(n \geqq 1)$, M always admits the Bergman metric and the Carathéodory differential metric.

Theorems A and B have a number of interesting consequences.

In [4], C. Earle has proved the completeness of the Carathéodory distance in the Teichmüller space $T(g)$ of a compact Riemann surface of genus $g \geqq 2$. Therefore, Theorem B immediately implies the following

THEOREM C. *In the Teichmüller space $T(g)$ of any compact Riemann surface of genus $g \geqq 2$, the Bergman metric is complete.*

Recently, S. Wolpert **[11]** and T. Chu have independently proved that the Weil–Petersson metric is not complete in $T(g)$. Therefore, we have the following

THEOREM D. *In the Teichmüller space $T(g)$ of any compact Riemann surface of genus $g \geqq 2$, the Weil–Petersson metric is not uniformly equivalent to the Bergman metric.*

Finally we have

THEOREM E. *Let G be a bounded open connected subset of a separable complex Hilbert space X of finite or infinite dimension, and let M be a complex manifold of finite dimension which admits s_M. If G is homogeneous, then there exists a constant, depending only on G, such that for any holomorphic mapping $f: M \to G$*

$$\alpha_G(f(z), Df(z)\xi) \leqq k(G) s_M(z, \xi) \qquad (z \in M,\ \xi \in C^n), \tag{3}$$

where $Df(z)$ denotes the Frechét derivative of f at $z \in M$.

If in particular G is a ball, B, in X, then

$$\alpha_B(f(z), Df(z)\xi) \leqq s_M(z, \xi). \tag{4}$$

Theorem E contains Theorem A as a special case when B is the unit disc in the complex plane C.

2. The kernel form and invariant metric of Bergman. The theory of the Bergman kernel function and invariant metric on a bounded domain in the space C^n has been extended to a complex manifold by S. Kobayashi **[7]** and also by A. Lichnerowicz **[8]**.

Let $\mathscr{F}(M)$ be the set of holomorphic n-forms

$$\alpha = a\,dz_1 \wedge \cdots \wedge dz_n$$

on M such that

$$\left| \int_M \alpha \wedge \bar{\alpha} \right| < \infty. \tag{1}$$

Then $\mathscr{F}(M)$ is a separable complex Hilbert space with an inner product given by

$$(\alpha, \beta) = i^{n^2} \int_M \alpha \wedge \bar{\beta} \qquad (\alpha, \beta \in \mathscr{F}(M)). \tag{2}$$

Let $\{\varphi_0, \varphi_1, \cdots\}$ be an orthonormal basis for $\mathscr{F}$. Then every $\alpha \in \mathscr{F}$ may be represented uniquely by the convergent series

$$\alpha(z) = \sum_{\nu=0}^{\infty} c_\nu \varphi_\nu(z), \qquad c_\nu = (\alpha, \varphi_\nu), \tag{3a}$$

or

$$a_U(z) = \sum_{\nu=0}^{\infty} c_\nu (\Phi_\nu)_U(z), \tag{3b}$$

where $\varphi_\nu = (\Phi_\nu)_U dz_1 \wedge \cdots \wedge dz_n$, in a local coordinate neighborhood U of $z \in M$.

Moreover,

$$(\alpha, \alpha) = \|\alpha\|^2 = \sum_{\nu=0}^{\infty} |c_\nu|^2. \tag{4}$$

Let V be a local coordinate neighborhood of $\zeta \in M$ in which $\varphi_\nu(\zeta) = (\Phi_\nu)_V(\zeta) d\zeta_1 \wedge \cdots \wedge d\zeta_n$. Then the series

$$i^{n^2} \sum_{\nu=0}^{\infty} \varphi_\nu(z) \wedge \overline{\varphi_\nu(\zeta)} = \tag{5}$$

$$i^{n^2} \sum_{\nu=0}^{\infty} (\Phi_\nu)_U(z) \overline{(\Phi_\nu)_V(\zeta)} dz_1 \wedge \cdots \wedge dz_n \wedge \overline{d\zeta_1} \wedge \cdots \wedge d\zeta_n$$

converges absolutely and uniformly on every compact subset of $M \times \bar{M}$, where $\bar{M}$ is the complex manifold conjugate to M, and hence, represents a holomorphic $2n$-form on $M \times \bar{M}$. Moreover, the sum (5) is independent of choice of orthonormal basis. The Bergman kernel form is defined by the sum (5) and written as

$$\kappa(z, \bar{\zeta}) = \kappa_\zeta(z) = i^{n^2} k(z, \bar{\zeta}) dz_1 \wedge \cdots \wedge dz_n \wedge \overline{d\zeta_1} \wedge \cdots \wedge \overline{d\zeta_n} \tag{5a}$$

with a locally defined Bergman kernel function:

$$k(z, \bar{\zeta}) = \sum_{\nu=0}^{\infty} (\Phi_\nu)_U(z) \overline{(\Phi_\nu)_V(\zeta)}, \qquad (z, \zeta) \in U \times V. \tag{5b}$$

Further we define the reduced kernel form by

$$K_\zeta(z) = k(z, \zeta)\, dz_1 \wedge \cdots \wedge dz_n. \tag{6}$$

As in the classical case, see [1], the reduced kernel form has the reproducing property of n-forms in $\mathcal{F}$. More precisely,

LEMMA 1. *For any $\alpha \in \mathcal{F}$ with $\alpha(z) = a_U(z)\, dz_1 \wedge \cdots \wedge dz_n$,*

$$a_U(z) = (\alpha, K_z) = i^{n^2} \int_M \alpha(t) \wedge K(z, \bar{t}) \qquad (z \in M). \tag{7}$$

Proof. First we observe that for each fixed $z \in M$, $K_z(t)$ is a holomorphic n-form in M. From the uniform convergence of the series (3a) and (5),

$$\begin{aligned}(\alpha, K_z) &= \Big(\sum_\nu c_\nu \varphi_\nu, \sum_\mu \overline{\Phi_\mu(z)}\, \varphi_\mu\Big) \\ &= \sum_\nu c_\nu \Big(\varphi_\nu, \sum_\mu \overline{\Phi_\mu(z)}\, \varphi_\mu\Big) \\ &= \sum_\nu c_\nu \sum_\mu \Phi_\mu(z)(\varphi_\nu, \varphi_\mu) = \sum_\nu c_\nu \Phi_\nu(z) = a_U(z).\end{aligned}$$

Setting in Lemma 1 $\alpha = K_\zeta$, $\zeta \in M$, we have

$$k_\zeta(z) = (K_\zeta, K_z) = \overline{(K_z, K_\zeta)} = \overline{k_z(\zeta)}. \tag{8}$$

In particular, $k_z(z) \geqq 0$. $k_z(z) > 0$ holds whenever M satisfies

(A1) For any z in M, there is an $\alpha \in \mathcal{F}(M)$ such that $\alpha(z) \neq 0$. In this case,

$$s^2(z, \xi) = \sum_{\alpha,\beta=1}^{n} \frac{\partial^2 \log k(z, \bar{z})}{\partial z_\alpha \partial \bar{z}_\beta} \xi_\alpha \bar{\xi}_\beta \qquad (z \in M,\ \xi \in C^n) \tag{9}$$

is a well-defined positive semidefinite hermitian form which is invariant under biholomorphic mappings of M. In fact $s^2(z, \xi)$ is positive definite if and only if M satisfies

(A2) For every holomorphic tangent vector ξ at $z \in M$, there is an $\alpha \in \mathcal{F}(M)$ such that $\alpha(z) = 0$ and

$$da \cdot \xi = \sum_{\mu=1}^{n} \frac{\partial a}{\partial z_\mu}(z) \xi_\mu \neq 0,$$

where $\alpha = a\, dz_1 \wedge \cdots \wedge dz_n$.

Therefore, any complex manifold M with properties (A1) and (A2) is entitled to an invariant Kähler metric s_M of Bergman.

3. An extension of Schwarz inequality. Let $\mathcal{M}(\Omega)$ be the set of square integrable n-forms defined on a measurable subset Ω of a complex manifold M of dimension n. Then $\mathcal{M}(\Omega)$ is a separable complex Hilbert space with respect to the inner product:

$$(\alpha, \beta)_\Omega = i^{n^2} \int_\Omega \alpha \wedge \bar{\beta} \qquad (\alpha, \beta \in \mathcal{M}(\Omega)). \tag{1}$$

We need the following extension of the Schwarz inequality.

LEMMA 2. *Let $\{\alpha_\nu\}$ and $\{\beta_\nu\}$ be two sequences* (finite or infinite) *from $\mathcal{M}(\Omega)$ such that*

$$\sum_\nu (\alpha_\nu, \alpha_\nu)_\Omega < \infty, \qquad \sum_\nu (\beta_\nu, \beta_\nu)_\Omega < \infty. \tag{2}$$

Then

$$M^* M \leqq N \cdot \sum_\nu (\alpha_\nu, \alpha_\nu)_\Omega \tag{3}$$

where "$\leqq$" denotes the matrix inequality, i.e., $A \leqq B$ if and only if $B - A$ is positive semidefinite, M and N the matrices whose entries are $M_{\mu\nu} = (\alpha_\mu, \beta_\nu)_\Omega$ and $N_{\mu\nu} = (\beta_\mu, \beta_\nu)_\Omega$ $(\mu, \nu = 0, 1, 2, \cdots)$, respectively, and M^* the adjoint of M.

Proof. It is enough to prove the case where $\{\alpha_\nu\}$ and $\{\beta_\nu\}$ are infinite sequences. The other cases can be proved in the same way. Let $u = (u_0, u_1, \cdots)$ be any non-zero constant vector in $\ell^2(C)$. Then

$$\begin{aligned} u^* M^* M u &= \sum_{\mu=0}^\infty \left(\sum_{\nu=0}^\infty M_{\mu\nu} u_\nu \right)^* \left(\sum_{\nu=0}^\infty M_{\mu\nu} u_\nu \right) \\ &= \sum_{\mu=0}^\infty \left| \left(\alpha_\mu, \sum_{\nu=0}^\infty \beta_\nu \bar{u}_\nu \right)_\Omega \right|^2 \end{aligned} \tag{4}$$

By the Schwarz inequality in $\mathcal{M}(\Omega)$, (4) becomes

$$
\begin{aligned}
u^*M^*Mu &\leqq \sum_{\mu=0}^{\infty} (\alpha_\mu, \alpha_\mu)_\Omega \left(\sum_{\nu=0}^{\infty} \beta_\nu \bar{u}_\nu, \sum_{\tau=0}^{\infty} \beta_\tau \bar{u}_\tau \right)_\Omega \\
&\leqq u^*Nu \sum_{\mu=0}^{\infty} (\alpha_\mu, \alpha_\mu)_\Omega
\end{aligned} \tag{5}
$$

from which (3) follows, since u was arbitrary.

In the case where $M = C^n$ and Ω is a measurable subset of C^n, we define $\mathcal{M}(\Omega)$ to be the set of square integrable functions on Ω. Lemma 2 then holds in this case. We shall state it separately for the future use.

COROLLARY 1. *Let* $\{a_\nu\}$ *and* $\{b_\nu\}$ *be two sequences* (finite or infinite) *from* $\mathcal{M}(\Omega)$, $\Omega \subset C^n$, *such that*

$$
\sum_\nu (a_\nu, a_\nu)_\Omega < \infty, \qquad \sum_\nu (b_\nu, b_\nu)_\Omega < \infty \tag{6}
$$

Then

$$
M^*M \leqq N \cdot \sum_\nu (a_\nu, a_\nu)_\Omega, \tag{7}
$$

where M *and* N *are matrices whose entries are* $(a_\mu, b_\nu)_\Omega$ *and* $(b_\mu, b_\nu)_\Omega$ $(\mu, \nu = 0, 1, 2, \cdots)$, *respectively.*

4. The main theorems.

THEOREM 1. *Let* $f = (f_0, f_1, \cdots)$ *be a holomorphic mapping from a complex manifold* M *satisfying properties* (A1) *and* (A2) *of* §2 *into a separable complex Hilbert space* X *of finite or infinite dimension such that*

$$
\|f(z)\|_X \leqq Q \quad \textit{for some} \quad Q > 0. \tag{1}
$$

Then

$$
\|Df(z)\xi\|_X \leqq Q s_M(z, \xi) \qquad (z \in M,\ \xi \in C^n), \tag{2}
$$

where $\| \ \|_X$ *denotes the usual norm in* X.

Proof. For each $z \in M$, let

$$
\alpha_\mu(t) = f_\mu(t) K_z(t) = f_\mu(t) k_z(t)\, dt_1 \wedge \cdots \wedge dt_n \qquad (\mu = 0, 1, 2, \cdots) \tag{3}
$$

$$
\beta_\nu(t) = \frac{\partial}{\partial \bar{z}_\nu}\left(\frac{K_z(t)}{k_z(z)}\right) = \frac{1}{k_z^2(z)}\left(k_z(z)\frac{\partial K_z(t)}{\partial \bar{z}_\nu} - K_z(t)\frac{\partial k_z(z)}{\partial \bar{z}_\nu}\right) \tag{4}
$$

$$
(\nu = 1, 2, \cdots, n)
$$

where

$$\frac{\partial K_z(t)}{\partial \bar{z}_\nu} = \frac{\partial k(t, \bar{z})}{\partial \bar{z}_\nu} dt_1 \wedge \cdots \wedge dt_n. \tag{5}$$

In view of the reproducing property of the kernel form, see Lemma 1, we obtain

$$\begin{aligned}\sum_\mu (\alpha_\mu, \alpha_\mu) = \sum_\mu (f_\mu K_z, f_\mu K_z) = \Big(\sum_\mu f_\mu \bar{f}_\mu K_z, K_z\Big) &\leqq Q^2(K_z, K_z) \\ &= Q^2 k(z, \bar{z}),\end{aligned} \tag{6}$$

$$\begin{aligned}(\beta_\mu, \beta_\nu) = \frac{1}{k_z^4(z)} \Big\{ k_z^2(z) \Big(\frac{\partial K_z}{\partial \bar{z}_\mu}, \frac{\partial K_z}{\partial \bar{z}_\nu} \Big) - k_z(z) \frac{\partial k_z(z)}{\partial z_\nu} \Big(\frac{\partial K_z}{\partial \bar{z}_\mu}, K_z \Big) \\ - k_z(z) \frac{\partial k_z(z)}{\partial \bar{z}_\mu} \Big(K_z, \frac{\partial K_z}{\partial \bar{z}_\nu} \Big) + \frac{\partial k_z(z)}{\partial \bar{z}_\mu} \frac{\partial k_z(z)}{\partial z_\nu} (K_z, K_z) \Big\}.\end{aligned} \tag{7a}$$

From Lemma 1, we also have

$$\Big(\frac{\partial K_z}{\partial \bar{z}_\mu}, \frac{\partial K_z}{\partial \bar{z}_\nu} \Big) = \frac{\partial^2}{\partial z_\nu \partial \bar{z}_\mu} (K_z, K_z) = \frac{\partial^2 k(z, \bar{z})}{\partial z_\nu \partial \bar{z}_\mu} \tag{8a}$$

$$\Big(\frac{\partial K_z}{\partial \bar{z}_\mu}, K_z \Big) = \frac{\partial}{\partial \bar{z}_\mu} (K_z, K_z) = \frac{\partial}{\partial \bar{z}_\mu} k(z, \bar{z}). \tag{8b}$$

Therefore, (7a) becomes

$$\begin{aligned}(\beta_\mu, \beta_\nu) &= \frac{1}{k^3(z, \bar{z})} \Big[k(z, \bar{z}) \frac{\partial^2 k(z, \bar{z})}{\partial z_\nu \partial \bar{z}_\mu} - \frac{\partial k(z, \bar{z})}{\partial z_\nu} \frac{\partial k(z, \bar{z})}{\partial \bar{z}_\mu} \Big] \\ &= \frac{1}{k(z, \bar{z})} \frac{\partial^2}{\partial z_\nu \partial \bar{z}_\mu} \log k(z, \bar{z}).\end{aligned} \tag{7b}$$

$$\begin{aligned}(\alpha_\mu, \beta_\nu) &= \Big(\alpha_\mu, \frac{\partial}{\partial \bar{z}_\nu} \Big(\frac{K_z}{k_z(z)} \Big) \Big) \\ &= \frac{\partial}{\partial z_\nu} \Big(\alpha_\mu, \frac{K_z}{k_z(z)} \Big) = \Big(\frac{\partial f_\mu}{\partial z_\nu} \Big)(z).\end{aligned} \tag{9}$$

From Lemma 2 applied to $\mathscr{F}(M)$, together with (6), (7b), (9) and (9) of §2, Theorem 1 follows.

Let $\mathscr{H}(M, B^m)$ be the set of all holomorphic mappings f of a complex manifold M into the unit ball B^m in the space C^m $(1 \leqq m \leqq \omega)$. Following H. Reiffen **[10]** we define

$$\alpha_M^{(m)}(z, \xi) = \sup\{\| Df(z)\xi \|_m : f \in \mathscr{H}(M, B^m)\} \tag{10}$$

for $(z, \xi) \in M \times C^m$, where B^ω denotes the unit ball in the Hilbert space $C^\omega = \ell^2(C)$ with the usual ℓ^2-norm.

It is easy to see that $\alpha_M^{(m)}$ is a pseudo differential metric in the sense of Grauert and Reckziegel **[5]**, and that $\alpha_M^{(m)}$ becomes a differential metric whenever M satisfies the properties (A1) and (A2) of §2 by bounded mappings in the class $\mathscr{H}(M, B^m)$. We note that $\alpha_M^{(1)} = \alpha_M$ is the Carathéodory differential metric of H. Reiffen **[10]**. However, it turns out that for all m, $1 \leqq m \leqq \omega$, $\alpha_M^{(m)}$ coincide with α_M, as it is seen in the following.

LEMMA 3. *Let M be a complex manifold of dimension n. For each $z \in M$ and each $\xi \in C^n$,*

$$\alpha_M^{(m)}(z, \xi) = \alpha_M^{(\omega)}(z, \xi) \quad \textit{for all} \quad m \geqq 1. \tag{11}$$

Proof. Suppose that $f = (f_1, f_2, \cdots, f_m) \in \mathscr{H}(M, B^m)$. Then $\tilde{f} = (f, 0) = (f_1, \cdots, f_m, 0, 0, \cdots)$ is a holomorphic mapping of M into B^ω. Let

$$\tilde{\mathscr{H}}(M, B^\omega) = \{\tilde{f}: \tilde{f} = (f, 0),\ f \in \mathscr{H}(M, B^m)\}.$$

Then

$$\tilde{\mathscr{H}}(M, B^\omega) \subset \mathscr{H}(M, B^\omega) \quad \text{and} \quad \|Df(z)\xi\|_m = \|D\tilde{f}(z)\xi\|_\omega.$$

Therefore,

$$\begin{aligned} \alpha_M^{(m)}(z, \xi) &= \sup\{\|Df(z)\xi\|_m : f \in \mathscr{H}(M, B^m)\} \\ &= \sup\{\|D\tilde{f}(z)\xi\|_\omega : \tilde{f} \in \tilde{\mathscr{H}}(M, B^\omega)\} \\ &\leqq \alpha_M^{(\omega)}(z, \xi). \end{aligned} \tag{12}$$

The opposite inequality follows from the following observation.

$$\begin{aligned} \|Df(z) \cdot \xi\|_\omega &= \sup\{|\ell(Df(z) \cdot \xi)| : \ell \in \ell^2(C)^*, \|\ell\| = 1\} \\ &= \sup\{|D(\ell \cdot f)(z) \cdot \xi| : \ell \in \ell^2(C)^*, \|\ell\| = 1\} \\ &\leqq \alpha_M(z, \xi), \end{aligned} \tag{13}$$

where $\ell^2(C)^*$ denotes the dual of $\ell^2(C)$.

The second half of Lemma 3 is due to Clifford Earle (by communication) to whom the author is indebted.

It should be pointed out that the method of the proof of Theorem 1 is essentially due to K. H. Look **[9]**. In fact, he has proved Theorem 1 for the case when M is a bounded domain in C^n and $X = C^n$. However,

K. H. Look did not seem to realize Lemma 3 which enabled us to relate Theorem 1 to the Carathéodory distance.

Theorem A is now an immediate corollary of Theorem 1, or rather a special case of Theorem 1.

Proof of Theorem A. Set $X = C$ and $Q = 1$ in Theorem 1. Then (2) becomes

$$|Df(z)\xi| \leqq s(z,\xi) \qquad (z \in M,\ \xi \in C) \tag{14}$$

for all $f \in \mathcal{H}(M, B^1)$, and Theorem A follows.

Proof of Theorem E. Let x_0 be any fixed point in G and let $\gamma: G \to G$ be a holomorphic automorphism of G such that $\gamma(x) = x_0$, where $x = f(z)$, $z \in M$. Then $\gamma \cdot f$ is a holomorphic mapping of M into G such that $(\gamma \cdot f)(z) = x_0$. Let Q be the radius of the smallest ball in X which contains G. We may assume that the center of this ball lies at the origin. By Theorem 1,

$$\|D(\gamma \cdot f)(z)\xi\|_X \leqq Qs(z,\xi), \qquad (z \in M,\ \xi \in C^n). \tag{15a}$$

It is known **[3]** that if G is bounded then there are two positive continuous functions λ and Λ in G such that

$$\lambda(x)\|\xi\|_X \leqq \alpha_G(x,\xi) \leqq \Lambda(x)\|\xi\|_X \qquad (x \in G), \tag{16}$$

for each $\xi \in X$. Set $\eta = Df(z)\xi$. Then (15a) becomes

$$\|D\gamma(x)\eta\|_X \leqq Qs(z,\xi), \qquad x = f(z). \tag{15b}$$

By the invariant property of the Carathéodory differential metric α_G under biholomorphic mappings of G, see **[3]**,

$$\alpha_G(x,\eta) = \alpha_G(\gamma(x), D\gamma(x)\eta) = \alpha_G(x_0, D\gamma(x)\eta). \tag{17}$$

From the second half of (16), (17), and (15b),

$$\alpha_G(f(z), Df(z)\xi) \leqq \Lambda(x_0)Qs(z,\xi). \tag{15c}$$

The first half of Theorem E follows from (15c) when we set

$$k(G) = Q \inf_{x \in G} \Lambda(x). \tag{18}$$

If in particular G is a ball, say $B=\{x \in X: \|x\|_X < R\}$, $R>0$, then $Q=R$ and inequalities (16) may be reduced to

$$(19) \qquad \frac{\|\xi\|_X}{\sqrt{R^2-\|x\|_X^2}} \leqq \alpha_B(x,\xi) \leqq \frac{R\|\xi\|_X}{R^2-\|x\|_X^2} \qquad (x \in B, \quad \xi \in X),$$

see **[3]**. Therefore, $k(G)=1$ in (18) which proves the rest of Theorem E.

References

1. S. Bergman, *Über die Kernfunktion eines Bereiches und ihr Verhalten am Rande*, J. Reine Angew. Math., **169** (1933), 1–42; **172** (1935), 89–128.

2. C. Carathéodory, *Über das Schwarzsche Lemma bei analytischen Funktionen von zwei komplexen Veränderlichen*, Math. Ann., **97** (1926), 76–98.

3. C. J. Earle and R. S. Hamilton, *A fixed point theorem for holomorphic mappings*, Global Analysis, Proc. of Symposium. in Pure Math. XVI, Amer. Math. Soc., Providence, R.I., 1965.

4. C. J. Earle, *On the Carathéodory metric in Teichmüller spaces*, Ann. of Math. Studies, **79**, Princeton Univ. Press, Princeton, N.J., 1974.

5. H. Grauert and H. Reckziegel, *Hermitesche Metriken und normale Familien holomorpher Abbildungen*, Math. Z., **89** (1965), 108–125.

6. K. T. Hahn, *On completeness of the Bergman metric and its subordinate metrics*, Proc. Nat. Acad. Sci. (U.S.A.) **73** (1976), 4294.

7. S. Kobayashi, *Geometry of bounded domains*, Trans. Amer. Math. Soc., **92** (1959), 267–290.

8. A. Lichnerowicz, *Variétés complexes et tenseur de Bergman*, Ann. Inst. Fourier (Grenoble), **15** (1965), 345–407.

9. K. H. Look, *Schwarz lemma and analytic invariants*, Scientia Sinica, **7** (1958), 453–504.

10. H. J. Reiffen, *Die differential geometrischen Eigenschaften der invarianten Distanz Funktion von Carathéodory*, Schr. Math. Inst. Univ. Münster, **26** (1963).

11. S. Wolpert, *Non-completeness of the Weil-Petersson metric for Teichmüller space*, to appear in Pacific J. Math., **61** (1975), 573–577.

Received June 20, 1976.

THE PENNSYLVANIA STATE UNIVERSITY

Current address: 818 Webster Dr.
State College, PA 16801

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No 2, 1977

ABSTRACTLY SPLIT GROUP EXTENSIONS

G. HOCHSCHILD AND D. WIGNER

1. Introductory survey. Consider a group extension $1 \to A \to E \xrightarrow{\pi} G \to 1$ in some category of groups with superstructure (topological, analytic, algebraic). Suppose it is split in the category of abstract groups, i.e., there is a homomorphism $\sigma: G \to E$ of abstract groups such that $\pi \circ \sigma$ is the identity map on G. We are concerned with the question of when it is possible to conclude that the extension is split as an extension in the given category.

The most surprising known result in this connection is due to C. Moore [**2**, Th. 2.3]. It says that if the given category is that of locally compact separable topological groups, if A lies in the center of G and G coincides with its commutator group $[G, G]$, then every σ as above is necessarily continuous.

A more transparent situation in which our question has a positive answer is the following. Suppose the given extension is in the category of locally compact separable topological groups, that A is a finite-dimensional real vector group and that G has a discrete subgroup K such that the coset space G/K is compact. Then, if an abstract split σ exists, it follows that there is also a continuous split. In fact, our assumptions on A and the topology of G imply that the given extension has a continuous cross-section [**3**, Th. 12.2], and [**0**, Th. 4.2]. This yields a continuous 2-cocycle f for G in A such that the extension is split if and only if f is the coboundary of a continuous map from G to A. By a well-known result due to van Est [**4**, §4], the restriction map from the continuous cohomology of G in A to that of K in A is injective. Our assumption that the extension has an abstract split evidently implies that the cohomology class of f is in the kernel of the restriction map. Therefore, it must be the 0-class, so that our extension has a continuous split.

Another positive case is that of an extension in the category of connected (real or complex) Lie groups in which the image G is simply connected and the kernel A is a central vector group. The existence of an abstract split σ evidently implies that $A \cap [E, E] = 1$, so that the given extension yields the extension

$$1 \to A \to E/[E, E] \to G/[G, G] \to 1$$

in the natural way. Since G is simply connected, $[G, G]$ is closed in G, and $G/[G, G]$ is simply connected [**1**, Ch. XII, Th. 1.2]. As before, we have from [**3**, Th. 12.2] that the given extension has a continuous cross-section, whence we find that E is simply connected. As just now, it follows that $[E, E]$ is closed in E, and $E/[E, E]$ is simply connected. Thus, the above extension is simply an extension of vector groups and therefore has a continuous (linear) split $\tau: G/[G, G] \to E/[E, E]$. Let P denote the inverse image of $\tau(G/[G, G])$ in E, so that P is a closed subgroup of E containing $[E, E]$, and $P/[E, E] = \tau(G/[G, G])$. Now it is easy to check that E is the direct product $A \times P$, which shows that the given group extension is split in the category of connected Lie groups.

The simplest example of an abstractly split extension of topological groups that is not continuously split, which must be known to many, is the following. Let $\mathscr{R}$, $\mathscr{Q}$, $\mathscr{Z}$ denote the additive groups of real numbers, rational numbers, integers, respectively. Let $\pi: \mathscr{R} \times (\mathscr{Q}/\mathscr{Z}) \to \mathscr{R}/\mathscr{Z}$ be defined by $\pi(x, y) = (x + \mathscr{Z}) + y$. Clearly, π is a continuous open homomorphism. The only compact subgroups of $\mathscr{R} \times (\mathscr{Q}/\mathscr{Z})$ are the finite subgroups of $\mathscr{Q}/\mathscr{Z}$, whence it is clear that the group extension given by π has no continuous split. However, from a $\mathscr{Q}$-space decomposition $\mathscr{R} = S \oplus \mathscr{Q}$, we obtain the group decomposition $\mathscr{R}/\mathscr{Z} = S \times (\mathscr{Q}/\mathscr{Z})$, and hence an abstract group split $\sigma: \mathscr{R}/\mathscr{Z} \to \mathscr{R} \times (\mathscr{Q}/\mathscr{Z})$.

This example is not satisfactory, because of the lack of connectedness. In the categories of connected Lie groups and connected affine algebraic groups, our question leads to interesting subquestions by various further specialisations. In the positive direction, we shall make some progress for unipotent affine algebraic groups over fields of characteristic 0. In the negative direction, we shall see how abstractly, but not continuously, split extensions of connected Lie groups arise from the fundamental group of the image group. The question of the existence of such examples, with simply connected image group, remains unresolved.

With regard to the above and to what follows, it is a pleasure to acknowledge the benefits we had from exploratory discussions with Brian Peterson and Chih-Han Sah.

2. Deflated extensions. A source of examples of the kind alluded to just above is the following construction. Let H be a topological group, and let K be a discrete central subgroup of H. Let τ be a homomorphism from K to an abelian topological group A. Let C be the subgroup of the direct product $A \times H$ consisting of the elements $(\tau(k), k)$, with k in K. Clearly, C is a discrete central subgroup of $A \times H$. Write E for $(A \times H)/C$, and let $\eta: H \to H/K$ be the canonical homomorphism. The composite with η of the projection $A \times H \to H$ induces a continuous, open and surjective homomorphism $\pi: E \to H/K$ whose kernel may be identified with A in the evident way. It is easy to verify that the group extension determined by π has an *abstract* split $H/K \to E$ if and only if τ is the restriction to K of an *abstract* group homomorphism $H \to A$, and that it has a *continuous* split if and only if τ is the restriction of a *continuous* group homomorphism $H \to A$.

Here is the simplest specific example arising in this way. Let M be the group of real matrices

$$[\alpha, \beta, \gamma] = \begin{pmatrix} 1 & 0 & 0 \\ \alpha & 1 & 0 \\ \gamma & \beta & 1 \end{pmatrix}.$$

Put $H = \mathcal{R} \times M$. Now fix an irrational real number μ, and let K be the discrete central subgroup of H consisting of the elements $(a + \mu b, [0,0,b])$, where a and b range over $\mathcal{Z}$. Finally, let $A = \mathcal{R}$, and define the homomorphism $\tau: K \to \mathcal{R}$ by

$$\tau(a + \mu b, [0,0,b]) = a + b$$

We claim that the resulting extension of topological groups

$$0 \to \mathcal{R} \to E \xrightarrow[\pi]{} H/K \to 1$$

has an abstract split, but does not have a continuous split. It is evident that τ can be extended to an abstract group homomorphism $H \to \mathcal{R}$ (annihilating M). Therefore, our group extension has an abstract split.

Now suppose that, contrary to our claim, there is a continuous split of our group extension. As stated above, this yields a continuous homomorphism $f: H \to \mathcal{R}$ whose restriction to K coincides with τ. We have $K \subset \mathcal{R} \times [M, M]$. Now note that f annihilates $[M, M]$ and is linear on the factor $\mathcal{R}$. It follows that there is a real number ρ such that

$$a + b = \tau(a + \mu b, [0,0,b]) = \rho a + \rho\mu b.$$

This gives the contradiction $\mu = 1$, so that our claim is established.

3. Unipotent groups. We consider the category of unipotent affine algebraic F-groups, where F is a field of characteristic 0. Our results will automatically hold also in the category of simply connected nilpotent real or complex analytic groups. We denote the Lie algebra of a group G by $\mathcal{L}(G)$. Recall that there are mutually inverse polynomial maps $\exp_G: \mathcal{L}(G) \to G$ and $\log_G: G \to \mathcal{L}(G)$, through which the categories of unipotent affine algebraic F-groups and of finite-dimensional nilpotent F-Lie algebras are equivalent. Our question can be transferred to the category of nilpotent Lie algebras by virtue of the following proposition.

PROPOSITION. *Let G and H be unipotent affine algebraic F-groups, with F of characteristic 0. Let $\gamma: G \to H$ be an abstract group homomorphism, and define the map $\gamma': \mathscr{L}(G) \to \mathscr{L}(H)$ by $\gamma' = \log_H \circ \gamma \circ \exp_G$. Then γ' is a morphism of $\mathscr{Q}$-Lie algebras. In this way, the abstract group homomorphisms $G \to H$ are in bijective correspondence with the morphisms of $\mathscr{Q}$-Lie algebras $\mathscr{L}(G) \to \mathscr{L}(H)$.*

Proof. If u and v are Lie algebra elements such that $[u, v] = 0$ then $\exp(u)\exp(v) = \exp(u + v)$. Using this with $\exp_G$ and $\exp_H$, we find that if n is an integer and x an element of $\mathscr{L}(G)$ then $\exp_H(\gamma'(nx)) = \exp_H(n\gamma'(x))$. Hence $\gamma'(nx) = n\gamma'(x)$. It follows that $\gamma'(qx) = q\gamma'(x)$ for every rational number q.

Since $\mathscr{L}(G)$ and $\mathscr{L}(H)$ are nilpotent, we can express products of exponentials in G and H by means of the Campbell–Hausdorff formula. This formula provides a set of rational numbers, indexed by finite sequences of 0's and 1's, such that the following holds. If u and v are given elements of $\mathscr{L}(G)$ or $\mathscr{L}(H)$, one attaches to each finite sequence of 0's and 1's a certain multiple commutator of u and v, according to the following recipe. To the sequence 0 we attach u, to the sequence 1 we attach v. Generally, if $[s]$ denotes the commutator attached to the sequence s, then $[0s] = [u, [s]]$ and $[1s] = [v, [s]]$. Since our Lie algebras are nilpotent, we have $[s] = 0$ whenever the length of s exceeds a certain bound. Therefore, if $q(s)$ is the rational number corresponding to s in the Campbell–Hausdorff formula, the sum $\Sigma_s q(s)[s]$ is defined as an element of the Lie algebra. Denoting this by $\eta(u, v)$, we have $\exp(u)\exp(v) = \exp(\eta(u, v))$. We recall that if η_n is the part of η coming from the sequences of length n, then $\eta_1(u, v) = u + v$ and $\eta_2(u, v) = \frac{1}{2}[u, v]$.

There is a sequence

$$(0) = Z_0 \subset \cdots \subset Z_n = \mathscr{L}(G)$$

of ideals of $\mathscr{L}(G)$ such that $[\mathscr{L}(G), Z_{k+1}] \subset Z_k$ for $k = 0, \cdots, n-1$. Now suppose we have already shown that $\gamma'(u + v) = \gamma'(u) + \gamma'(v)$ and $\gamma'([u, v]) = [\gamma'(u), \gamma'(v)]$ whenever u is in $\mathscr{L}(G)$ and v is in Z_k. Let q be a rational number, v an element of Z_{k+1} and u any element of $\mathscr{L}(G)$. From the definitions, we have

$$\gamma'(\eta(qu, qv)) = \eta(\gamma'(qu), \gamma'(qv))$$

This may be written

$$\gamma'(\Sigma_k q^k \eta_k(u, v)) = \Sigma_k q^k \eta_k(\gamma'(u), \gamma'(v))$$

Since, for $k > 1$, $\eta_k(u, v)$ lies in Z_k we may apply our inductive hypothesis to expand the left side, and we obtain

$$\Sigma_k q^k \gamma'(\eta_k(u, v)) = \Sigma_k q^k \eta_k(\gamma'(u), \gamma'(v))$$

Since this holds for all rational numbers q, the coefficients of q^k on the two sides must be equal. In particular, taking $k = 1$ and $k = 2$, we obtain $\gamma'(u + v) = \gamma'(u) + \gamma'(v)$ and $\gamma'([u, v]) = [\gamma'(u), \gamma'(v)]$. This proves, inductively, that γ' is a morphism of $\mathcal{Q}$-Lie algebras.

Conversely, suppose that $\delta: \mathcal{L}(G) \to \mathcal{L}(H)$ is a morphism of $\mathcal{Q}$-Lie algebras. Put $\gamma = \exp_H \circ \delta \circ \log_G$. Applying the Campbell–Hausdorff formula and noting that η has rational coefficients, one verifies directly that $\gamma: G \to H$ is a homomorphism of abstract groups. Clearly, $\gamma' = \delta$. This completes the proof of the proposition.

If L is an F-Lie algebra, K a finite algebraic extension field of $\mathcal{Q}$ contained in F and M a K-Lie algebra such that $L = M \otimes_K F$, then we call M an *absolutely algebraic form* of L.

THEOREM. *Let F be a field of characteristic* 0, *and let G be a unipotent affine algebraic F-group. Suppose that $\mathcal{L}(G)$ has an absolutely algebraic form. Let $1 \to A \to E \to G \to 1$ be a group extension of unipotent affine algebraic F-groups having an abstract split. Then this group extension is split in the category of affine algebraic F-groups.*

Proof. Write $\mathcal{L}(G) = L = M \otimes_K F$, as above. Viewing M as a $\mathcal{Q}$-Lie algebra, consider the extension of K-Lie algebras

$$0 \to P \to M \otimes_{\mathcal{Q}} K \xrightarrow[\pi]{} M \to 0$$

coming from the K-space structure of M. Write U for $M \otimes_{\mathcal{Q}} K$, and note that U is a two-sided K-module, with

$$c \cdot (m \otimes k) = (cm) \otimes k \quad \text{and} \quad (m \otimes k) \cdot c = m \otimes (kc)$$

for c and k in K and m in M. The kernel P of π is clearly the two-sided K-submodule consisting of all sums of elements of the form $c \cdot u - u \cdot c$, with u in U and c in K. Now K is a finite-dimensional separable $\mathcal{Q}$-algebra, so that $K \otimes_{\mathcal{Q}} K$ is a finite-dimensional semisimple $\mathcal{Q}$-algebra, whence every two-sided K-module is semisimple. Let S be a two-sided K-module complement of P in U. Clearly, $c \cdot s = s \cdot c$ for every s in S and every c in K. Let T denote the two-sided K-submodule of U consisting of all elements u for which $c \cdot u = u \cdot c$ for all c in K. We claim that $S = T$. In order to prove this, it suffices to show that $P \cap T = (0)$. By the semisimplicity, P is a direct two-sided K-module sum $(P \cap T) \oplus R$. Since $U = P + T$, it follows from the definition of P that every element of P is a sum of elements $c \cdot p - p \cdot c$ with c in K and p in P (not only in U), and the above decomposition of P shows that we may even take the elements p to be in R. But this gives $P = R$, so that $P \cap T = (0)$.

Thus, $U = P \oplus T$. Evidently, T is an ideal of U, so that this is a direct K-Lie algebra decomposition. The restriction of π to T is an isomorphism of K-Lie algebras $T \to M$. Let $\mu: M \to T$ be its inverse. By tensoring with F and evident identifications, μ yields a morphism of F-Lie algebras

$$\mu^*: L = M \otimes_K F \to T \otimes_K F \subset U \otimes_K F = M \otimes_{\mathcal{Q}} F \subset L \otimes_{\mathcal{Q}} F$$

If $\tau: L \otimes_{\mathcal{Q}} F \to L$ is the morphism of F-Lie algebras coming from the F-space structure of L, it follows from the definition of μ^* that $\tau \circ \mu^*$ is the identity map on L.

Now let $1 \to A \to E \xrightarrow[\rho]{} G \to 1$ be as in the statement of the theorem. This yields the extension of F-Lie algebras

$$0 \to \mathscr{L}(A) \to \mathscr{L}(E) \xrightarrow[\rho']{} \mathscr{L}(G) \to 0$$

By our above proposition, an abstract split of the given group extension yields a morphism of $\mathcal{Q}$-Lie algebras

$$\sigma: \mathscr{L}(G) \to \mathscr{L}(E)$$

such that $\rho' \circ \sigma$ is the identity map on $\mathscr{L}(G)$. By tensoring with F, we obtain the morphism of F-Lie algebras

$$\sigma^*: \mathscr{L}(G) \otimes_{\mathcal{Q}} F \to \mathscr{L}(E) \otimes_{\mathcal{Q}} F$$

Let $\gamma: \mathscr{L}(E) \otimes_{\mathcal{Q}} F \to \mathscr{L}(E)$ be the morphism of F-Lie algebras coming from the F-space structure of $\mathscr{L}(E)$. Then, if μ^* is the morphism of F-Lie algebras obtained above, the composite

$$\gamma \circ \sigma^* \circ \mu^*: \mathscr{L}(G) \to \mathscr{L}(E)$$

is a split of our above extension of F-Lie algebras. Via $\log_G$ and $\exp_E$, this yields a split of the given group extension in the category of affine algebraic F-groups, so that our theorem is established.

References

0. A. M. Gleason, *Spaces with a compact Lie group of transformations*, Proc. Am. Math. Soc., vol. **1** (1950), pp. 35–43.

1. G. Hochschild, *The structure of Lie groups*, Holden-Day, San Francisco, 1965.

2. C. C. Moore, *Group extensions of p-adic and adelic linear groups*, I.H.E.S., No. **35** (1968).

3. N. Steenrod, *The topology of fibre bundles*, Princeton University Press, 1951.

4. W. T. van Est, *A generalization of the Cartan–Leray spectral sequence*, II, Indag. Math. **20**, No. 4 (1958).

Received December 16, 1976.

UNIVERSITY OF CALIFORNIA
BERKELEY, CA 94720

AND

UNIVERSITY OF MICHIGAN
ANN ARBOR, MI 48104

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No. 2, 1977

INNER INVARIANT SUBSPACES

GARY S. ITZKOWITZ

We single out a special subclass of the invariant subspaces which we call the inner invariant (i.i.) subspaces. A closed subspace K of a Hilbert space H is said to be *i.i.* for a linear operator T (with domain D) if: (1) $T(K \cap D) \subseteq K$, (2) $\{T(K \cap D) + (K \cap D)\}^{-} = K$, and (3) $x \in D \backslash K \Rightarrow Tx \notin K$. This generalizes subspaces invariant for both T and T^{-1} when the latter exists.

Some of the results in this paper are:

1. Let $\lambda \in \mathbf{C}$. If $|\lambda| < 1$ then K is *i.i.* for $U - \lambda$ where U is the shift on Hardy space H^p iff $K = gH^p$ where g is inner and $g(\lambda) \neq 0$. If $|\lambda| \geqq 1$, then K is *i.i.* for $U - \lambda$ iff $K = gH^p$ where g is inner. 2. There is an isometry J from H^2 onto $L^2(0, \infty)$ such that the *i.i.* subspaces of $V + 1$ (where $Vf(x) = \int_0^x f(y)\,dy$) are precisely the subspaces $J(gH^2)$ for g an inner function. 3. Any skew-symmetric simple operator with defect indices $(0, 1)$ is isomorphic with V and V^{-1}.

1. Introduction. In §2 below we present the definitions of invariance and inner invariance for (not necessarily bounded) operators. Then their fundamental properties are analyzed.

In §3 we calculate the inner invariant subspaces of the shift operator in several settings. In one setting we describe the inner invariant subspaces of the shift on the Hardy space H^p. Then we generalize this result.

In §4 we consider the unbounded Volterra operator V. We first characterize this operator abstractly and then use this to get the result that on $L^2(0, \infty)$ integration and differentiation (i.e., V and V^{-1}) are isometrically isomorphic.

Finally, in §5, we describe the inner invariant subspace structure of the unbounded operator $V + 1$.

2. Definitions and basic properties. We make the following conventions. We work in a Hilbert space H and closed linear subspaces K. A linear, though not necessarily bounded, operator on H will be denoted by T with linear domain $D = D(T)$. If T^{-1} exists, we write D^{-1} for $D(T^{-1}) \equiv T(D(T))$.

If R and S are linear subspaces of H then $R + S$ is the linear subspace generated by the elements of R and S. The closure of R in H is denoted by $\bar{R}$ or $\{R\}^{-}$.

DEFINITION 2.1. A closed subspace K of H is *invariant* for T if:
(i) $T(K \cap D) \subseteq K$;
(ii) $\{T(K \cap D) + (K \cap D)\}^{-} = K$.

This definition allows for the extreme that neither $T(K \cap D)$ nor $K \cap D$ alone is dense in K yet K is still an invariant subspace for T.

DEFINITION 2.2. A closed subspace K is *inner invariant* for T if it is invariant for T and in addition it satisfies the following property:
(iii) $x \in D \setminus K \Rightarrow Tx \notin K$.

The following example shows that an invariant subspace is not necessarily inner invariant and hence that invariance and inner invariance are different.

EXAMPLE 2.3. Consider the shift operator s on l^2 where $s([a_0, a_1, a_2, \cdots]) = [0, a_0, a_1, a_2, \cdots]$. Let $K = \{[a_n]_{n=0}^{\infty} \mid a_0 = 0\}$; then it is a straightforward matter to show that K is an invariant subspace for s but is not inner invariant for s.

LEMMA 2.4. *If a subspace K is inner invariant for T then $T(K \cap D) = K \cap T(D)$.*

Proof. By the invariance of K we get trivially that $T(K \cap D) \subseteq K \cap T(D)$. To show $K \cap T(D) \subseteq T(K \cap D)$ let $y \in K \cap T(D)$ so that $y = Tx$ for some x in D. If $x \in D \setminus K$ and $Tx \in K$ then we are contradicting (iii) in Definition 2.2.

THEOREM 2.5. *If T is one-to-one then the following are equivalent*:
(i) *K is inner invariant for T*;
(ii) *K is inner invariant for T^{-1}*;
(iii) *K is invariant for both T and T^{-1}.*

Proof. We will just prove (i) $\Rightarrow$ (ii), the other cases being clear. Since K is invariant, Lemma 2.4 implies that $T(K \cap D) = K \cap T(D)$. Since T is one-to-one, T^{-1} exists and $D^{-1} \equiv T(D)$. Also $D = T^{-1}(D^{-1})$.

Thus we have the following equalities:

$$T(K \cap D) = K \cap T(D) = K \cap D^{-1} \tag{1}$$

and

$$K \cap D = T^{-1}(K \cap T(D)) = T^{-1}(K \cap D^{-1}) \tag{2}$$

and

$$K \cap T^{-1}(D^{-1}) = T^{-1}(K \cap D^{-1}). \tag{3}$$

Now, by the inner invariance of K with respect to T we have, by using (1) and (2), that

$$\begin{aligned} K &= \{T(K \cap D) + (K \cap D)\}^- \\ &= \{(K \cap D^{-1}) + T^{-1}(K \cap D^{-1})\}^-. \end{aligned} \tag{4}$$

Hence condition (ii) in the definition of inner invariance (for T^{-1}) is satisfied.

We now use (3) and (4) to show K is inner invariant for T^{-1}. First, from (4) we get (automatically)

$$T^{-1}(K \cap D^{-1}) \subseteq K$$

so that (i) of the definition of invariance is satisfied for T^{-1}.

Assume $x \in D^{-1}$ and $T^{-1}x \in K$. In order for condition (iii) in the definition of inner invariance to hold for T^{-1} we must show $x \in K$. If x did not belong to K, then this, with the assumptions that $x \in D^{-1}$ and $T^{-1}x \in K$, would be a contradiction to condition (iii) and therefore to inner invariance.

Now $T^{-1}x \in K$ and $T^{-1}x \in T^{-1}(D^{-1})$ since $x \in D^{-1}$. Therefore $T^{-1}x \in K \cap T^{-1}(D^{-1}) = T^{-1}(K \cap D^{-1})$ by (3). The operator T^{-1} is one-to-one so that x must belong to $K \cap D^{-1}$ and hence is in K. Thus K is inner invariant for T^{-1}.

The next two examples exhibit (1) operators, all of whose invariant subspaces are inner invariant and (2) operators without inner invariant subspaces. The latter settles the inner invariant subspace problem easily in contradistinction with the long standing but recently solved problem concerning the existence of invariant subspaces of bounded operators. (At the August, 1976 meeting of the American Mathematical Society in Toronto, Per Enflo announced that he had solved the invariant subspace problem.)

EXAMPLE 2.6. Consider the bounded Volterra operator V on $L^2(0,1)$ defined by $Vf(x) = \int_0^x f(y)\,dy$. Kalisch **[10]** proved that the proper invariant subspaces of this operator are of the form $L^2(a,1)$ for $0 < a < 1$. We show that these subspaces are also inner invariant for V. Since $L^2(a,1)$ is invariant, we need only demonstrate the last condition in the definition of inner invariance. Thus let a be fixed and assume that $0 \neq f \in L^2(0,1)$ with $f \notin L^2(a,1)$. We must show that $Vf \notin L^2(a,1)$ but this is obvious since f must have some of its support on $(0,a)$ and therefore $Vf(x) = \int_0^x f(y)\,dy$ must also have support on $(0,a)$. Thus $Vf \notin L^2(a,1)$.

EXAMPLE 2.7. Consider the Hardy space H^2 in the unit disk (i.e., $\{z \in \mathbf{C} \mid |z| < 1\}$). On this space we will be concerned with the weighted shift operator S defined by $Sf(z) = zf(z/2)$ for $f \in H^2$. Donoghue **[3]** showed that the proper invariant subspaces of S are of the form $z^n H^2$ for $n = 1, 2, 3, \cdots$. It is then trivial to show that none of these subspaces is inner invariant.

DEFINITION 2.8. The closed subspace K of H is said to be *reducing* for T if:

(i) $D(T) = (D \cap K) \oplus (D \cap K^\perp)$;

(ii) $T(K \cap D) \subseteq K$ and $T(K^\perp \cap D) \subseteq K^\perp$;

(iii) $\{T(K \cap D) + (K \cap D)\}^- = K$ or $\{T(K^\perp \cap D) + (K^\perp \cap D)\}^- = K^\perp$.

This definition is a natural extension of the definition of invariance (Definition 2.1) to the concept of reducing, but it is not the standard definition used for unbounded operators. In Akhiezer and Glazman **[1]**, page 82, a closed subspace K is said to reduce a linear operator T if only conditions (i) and (ii) of Definition 2.8 hold. In this case we will say that K is *A-reducing* for T.

PROPOSITION 2.9. *Let $\overline{D(T)} = H$. Then K reduces T iff it A-reduces T.*

Proof. Straightforward.

REMARK. In general, A-reducing (in the absence of the density assumption) does not imply reducing.

PROPOSITION 2.10. *Let T be one-to-one with $\overline{D(T)} = H$. Then if K reduces T, both K and $K^\perp$ are inner invariant for T.*

Proof. Since $\overline{D(T)} = H$ we get trivially that $\{K \cap D\}^- = K$ and $\{K^\perp \cap D\}^- = K^\perp$. Hence both statements in condition (iii) of the definition of reducing are true so that we can conclude that both K and $K^\perp$ are invariant for T. To prove inner invariance, let $x \in D \backslash K$ so that $x = k + k^\perp$ with $k \in K \cap D$ and $k^\perp \in K^\perp \cap D$. Then $Tx = Tk + Tk^\perp$ where $Tk \in K$ and $Tk^\perp \in K^\perp$ since K reduces T. We know $x \notin K$ so that $k^\perp \neq 0$. If $Tx \in K$ then we must have $Tk^\perp = 0$ which would imply $k^\perp = 0$, a contradiction. Hence $Tx \notin K$. A similar argument works for $K^\perp$.

REMARK. If $D(T)$ were not dense in H then only one of the two conditions in (iii) of Definition 2.8 might hold, in which case only one of K or $K^\perp$ would be inner invariant.

PROPOSITION 2.11. *Let T be one-to-one with $\overline{D(T)} = H$. Then if K reduces T, K is inner invariant for T and so is invariant for T. Furthermore, in general, these implications do not hold in the other direction.*

Proof. Proposition 2.10 states that the first implication is true and the second one is a simple consequence of the definition of inner invariance. To see that the implications do not hold in the other direction, consider Examples 2.6 and 2.3. In Example 2.6 we saw that the subspaces $L^2(a, 1)$ for $0 < a < 1$ are inner invariant for the Volterra operator V, but they do not reduce it since $L^2(a, 1)^\perp = L^2(0, a)$ is not invariant for V. In Example 2.3, we exhibited an operator with an invariant subspace that was not inner invariant.

EXAMPLE 2.12. We show that if T is not one-to-one then Propositions 2.10 and 2.11 may actually be false. Let T_1 be a linear operator on a finite dimensional Hilbert space H_1 with nonzero kernel. Let T_2 be a nonsingular linear operator on a finite dimensional Hilbert space H_2. Form the Hilbert space $H = H_1 \oplus H_2$ and the operator $T = T_1 \oplus T_2$. The operator T is not one-to-one since T_1 is singular. The subspace H_2 clearly reduces T but is not inner invariant for T. To see this consider $x = h_1 + h_2$ with $h_1 \in H_1$, $h_2 \in H_2$ and $0 \neq h_1 \in$ kernel of T_1; then $x \notin H_2$ but $Tx = T_1h_1 + T_2h_2 = T_2h_2 \in H_2$ so that H_2 is invariant but not inner invariant for T.

EXAMPLE 2.13. We show several things here. First we exhibit an operator that has inner invariant and (non inner) invariant subspaces. Second, the inner invariant subspaces will be totally ordered. Third, the (non inner) invariant subspaces are examples of subspaces for which condition (ii) in the definition of invariance holds nontrivially. This example extends Example 2.6 and so we use the same notation.

Since the point spectrum of V is empty, V^{-1} exists. We write L^2 for $L^2(0, 1)$. Since $D(V) = L^2$ we conclude that $D(V^{-1}) = \{f \in L^2 \mid f$ is absolutely continuous, $f' \in L^2$, and $f(0) = 0\}$ and $V^{-1}f(x) = d/dx\, f(x)$. The operator V^{-1} is a closed unbounded operator (Stone **[15]**, Theorem 10.7, Page 428).

As was shown in Example 2.6, the inner invariant subspaces of V and V^{-1} are the subspaces $L^2(a, 1)$ for $0 < a < 1$ and they are totally ordered. These do not constitute all the invariant subspaces for V^{-1}. It is straightforward to show that the spaces P_n, where P_n is the set of all polynomials of degree less than or equal to the positive integer n, are invariant. As a matter of fact $P_n \cap D(V^{-1})$ which is the linear subspace generated by $\{x, x^2, \cdots, x^n\}$ is properly contained in P_n, and $V^{-1}(P_n \cap D(V^{-1})) = P_{n-1}$ is also properly contained in P_n but $\{(P_n \cap D^{-1}) + V^{-1}(P_n \cap D^{-1})\} = P_n$ so that condition (ii) in the definition of invariance is indeed satisfied nontrivially.

We close this section with some propositions giving us certain conditions under which some or all of the concepts of invariance, inner invariance, and reducing coincide. In the following propositions the linear operator A is assumed to be a *bounded* operator defined everywhere on the Hilbert space H.

PROPOSITION 2.14. *If A is self-adjoint and one-to-one then the following three conditions are equivalent*:

(i) *K is inner invariant for A*;
(ii) *K is reducing for A*;
(iii) *K is invariant for A.*

Proof. By Proposition 2.11 all we need do is show that K being invariant for A implies K is reducing for A.

If $x \in K$ and $y \in K^\perp$ then $0 = (Ax, y)$ since K is invariant. But $(Ax, y) = (x, Ay)$ since $A = A^*$. Thus Ay is perpendicular to x. Since x is arbitrary in K we conclude that $Ay \in K^\perp$ so that $K^\perp$ is also invariant for A and thus reduces A.

PROPOSITION 2.15. *Let A be a unitary operator on H. Then K is inner invariant for A iff K is reducing for A.*

Proof. ($\Leftarrow$). Trivial.

($\Rightarrow$). Theorem 2.5 tells us that if K is inner invariant for A then it is invariant for both A and A^{-1} and conversely. Let $x \in K$ and $y \in K^\perp$. Then since K is invariant for A^{-1}, and A is unitary, we get the following:

$$0 = (A^{-1}x, y) = (A^*x, y) = (x, Ay).$$

Thus Ay is perpendicular to K, i.e., $Ay \in K^\perp$ so that $K^\perp$ is invariant for A. Since both K and $K^\perp$ are invariant for A, the subspace K reduces A.

PROPOSITION 2.16. *If A is bounded and $A-1$ is generalized nilpotent then A and A^{-1} have the same closed invariant subspaces. Thus the inner invariant subspaces of A (and A^{-1}) coincide with the invariant subspaces of A.*

Proof. Consider the infinite series $\Sigma_{n=0}^{\infty}(1-A)^n$; this series converges if $\Sigma_{n=0}^{\infty}\|(1-A)^n\|$ converges and this series does converge since $A-1$ is generalized nilpotent. Since $\Sigma_{n=0}^{\infty}(1-A)^n=(1-(1-A))^{-1}=A^{-1}$ we conclude that A^{-1} exists and is bounded.

Hence if K is invariant for A then it is invariant for $1-A$ and therefore for $(1-A)^n$ for all positive integers n. Thus K is invariant for $A^{-1}=\Sigma(1-A)^n$.

In the other direction we have

$$A=(A^{-1})^{-1}=(1-(1-A^{-1}))^{-1}=\sum_{n=0}^{\infty}(1-A^{-1})^n.$$

This is a valid expression for A since

$$(1-A^{-1})^n=[A^{-1}(A-1)]^n=A^{-n}(A-1)^n$$

and thus the generalized nilpotency of $A-1$ insures the convergence of $\Sigma_{n=0}^{\infty}(1-A^{-1})^n$.
Thus, if K is invariant for A^{-1} then K is invariant for $(1-A^{-1})^n$ for all n, so that K is invariant for A.

3. The shift operator. In this section we describe the inner invariant subspace structure of the shift operator in several settings. In the first setting, the spaces considered are the Hardy spaces H^p for $1 \leqq p \leqq \infty$. For background on Hardy spaces, the reader is referrred to Hoffman **[9]**.

Briefly, the Hardy spaces are the Banach spaces of p-integrable analytic functions in the unit disk $\{z \mid |z|<1\}$, or equivalently, the subspace of L^p of the unit circumference with no negative Fourier coefficients.

DEFINITION 3.1. The *shift operator* U on H^p is defined by $Uf(z)=zf(z)$ for $f \in H^p$.

LEMMA 3.2. *The function $z-a$ $(a \in \mathbf{C})$ is an outer function for all a of modulus* 1.

Proof. We show that $z-a$ is outer by showing that $\log|e^{i\theta}-a|$ is integrable and that

$$0 \equiv \log|-a| = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log|e^{i\theta} - a|\, d\theta.$$

It is clear that $z - a$ belongs to H^1 since it is bounded and analytic in the unit disk. Thus $\log|e^{i\theta} - a|$ must be integrable.

Let $a = 1$ and set $F(z) = z - 1$ so that $\log|F(0)| = 0$. With a little calculation we get

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} \log|F(e^{i\theta})|\, d\theta &= \frac{1}{4\pi}\int_{-\pi}^{\pi} \log[2(1-\cos\theta)]\, d\theta \\ &= \log 2 + \frac{1}{2\pi}\int_{0}^{2\pi} \log\left(\sin\frac{\theta}{2}\right) d\theta \\ &= 0 = \log|F(0)|.\end{aligned}$$

Thus $z - 1$ is outer. The case $F(z) = z - a$ ($|a| = 1$) may then be reduced to the case $a = 1$ by a simple change of variable and thus yields the same result.

We now describe the inner invariant subspaces of translates of the shift, i.e., inner invariant subspaces of $U - a$ for $a \in \mathbf{C}$.

THEOREM 3.3. *The nonzero closed inner invariant subspaces S of $U - a$ on H^p $(1 \leqq p \leqq \infty)$ are the following*:

(i) *If $|a| < 1$ then S is inner invariant for $U - a$ iff $S = gH^p$ where g is an inner function and $g(a) \neq 0$.*

(ii) *If $|a| \geqq 1$ then S is inner invariant for $U - a$ iff $S = gH^p$ where g is an inner function; i.e., S is inner invariant iff S is invariant.*

Proof. Since Beurling **[2]** showed that the invariant subspaces of U (and therefore of $U - a$) are of the form gH^p for g inner, all we need do is test which of these satisfy condition (iii) in the definition of inner invariance. We divide this into several cases.

Case 1. Let $|a| < 1$ and assume $g(a) \neq 0$. Given f in $H^p \backslash gH^p$ we must show that $(U - a)f \notin gH^p$ in order for gH^p to be inner invariant for $U - a$. Equivalently, if $(U - a)f \in gH^p$ we must show that $f \in gH^p$. Thus assume that there is a function $h \in H^p$ such that $(U - a)f(z) = g(z)h(z)$, i.e.,

$$(z - a)f(z) = g(z)h(z). \tag{5}$$

Therefore

$$f(z) = g(z)h(z)/z - a. \tag{6}$$

Now the left hand side of (5) is zero when $z = a$ so that $g(a)h(a) = 0$, but $g(a) \neq 0$ by assumption, hence $h(a) = 0$. In other words $h(z)/z - a$ is analytic. Then, by using (6), in order for f to belong to gH^p we need only show

$$\lim_{r\to 1} \int_{-\pi}^{\pi} \left| \frac{h(re^{i\theta})}{re^{i\theta} - a} \right|^p d\theta < \infty.$$

Now, the inequality

$$\frac{1}{|re^{i\theta} - a|^p} \leqq \frac{1}{|r - |a||^p}$$

implies that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|h(re^{i\theta})|^p}{|re^{i\theta} - a|^p} d\theta \leqq \frac{1}{|r - |a||^p} \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} |h(re^{i\theta})|^p d\theta$$

and therefore

$$\begin{aligned} \lim_{r\to 1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{h(re^{i\theta})}{re^{i\theta} - a} \right|^p d\theta &\leqq \frac{1}{(1 - |a|)^p} \lim_{r\to 1} \frac{1}{2\pi} \int_{-\pi}^{\pi} |h(re^{i\theta})|^p d\theta \\ &= \frac{1}{(1 - |a|)^p} \|h\|_p^p < \infty \end{aligned}$$

since $h \in H^p$. From this we conclude that $h/z - a$ is in H^p so that $f = gh$ is in gH^p. This means gH^p is inner invariant for $U - a$ provided $g(a) \neq 0$.

Case 2. Let $|a| < 1$ and assume $g(a) = 0$. Since $g(a) = 0$ and $g \in H^p$, the proof in Case 1 above (with h replaced by g) shows that $g/z - a$ is in H^p. At the same time $g/z - a$ is not in gH^p since $1/z - a$ is not in H^p. Thus $g/z - a$ is in $H^p \backslash gH^p$ and $(U - a)g(z)/z - a = g(z)$ is in gH^p which contradicts the definition of inner invariance. Therefore gH^p cannot be inner invariant for $U - a$ when $g(a) = 0$.

These first two cases prove part (i).

Case 3. Let $|a| > 1$. In this case it is obvious that all invariant subspaces of $U - a$ are inner invariant.

Case 4. Let $|a| = 1$. We assume that $(z - a)f \in gH^p$ or

$$(z - a)f(z) = g(z)h(z) \tag{7}$$

for some h in H^p. Since f and h are in H^p which is contained in H^1, we know that there is a decomposition of f and h unique up to constants of modulus 1 such that

$$f = g_1F_1 \quad \text{and} \quad h = g_2F_2 \tag{8}$$

where g_1 and g_2 are inner functions and F_1 and F_2 are outer functions in H^p.

Substituting (8) into (7) we get

$$(z - a)g_1F_1 = gg_2F_2. \tag{9}$$

Then since $z - a$ is outer (Lemma 3.2) and bounded we can conclude that $(z - a)F_1$ is outer. Since the decomposition of an H^p function is unique up to constants of modulus 1, we can conclude from (9) that $g_1 = cgg_2$ and $(z - a)F_1 = bF_2$ for some c, b in $\mathbf{C}$ with $|c| = 1 = |b|$. Therefore $F_1 = bF_2/z - a \in H^p$ and this implies that $F_2/z - a \in H^p$. Hence $h/z - a = g_2F_2/z - a \in H^p$ so that $f = gh/z - a$ is in gH^p. This means that gH^p is inner invariant for $U - a$.

PROPOSITION 3.4. (i) *If* $|a| < 1$ *then* $\{(z - a)H^2\}^- \subsetneq H^2$.
(ii) *If* $|a| = 1$ *then* $\{(z - a)H^2\}^- = H^2$.
(iii) *If* $|a| > 1$ *then* $(z - a)H^2 = H^2$.

Proof. If $|a| > 1$ then $1/z - a$ is in H^2 so that $f = (z - a)f/z - a$ is in $(z - a)H^2$, i.e., $(z - a)H^2 = H^2$. If $|a| = 1$ then $z - a$ is an outer function (Lemma 3.2). We know that $(z - a)H^2 \supseteq \{(z - a)z^n\}_{n=0}^{\infty}$ but it is also true that this sequence spans H^2 since $z - a$ is outer. Therefore $\{(z - a)H^2\}^- = H^2$. If $|a| < 1$ then the function

$$g(z) = -\frac{\bar{a}}{|a|} \cdot \frac{z - a}{1 - \bar{a}z}$$

is an inner function and

$$-\frac{\bar{a}}{|a|} \frac{1}{1 - \bar{a}z}$$

is analytic in the unit disk. Therefore $(z - a)H^2 = g(z)H^2$ and $g(z)H^2$ is a closed proper subspace of H^2.

So far we have discussed the unilateral shift operator, i.e., the shift on H^p. Now we investigate the inner invariant subspace structure of the bilateral shift on $L^2(K)$ where K is itself a Hilbert space.

We start by considering the special case where $K = \mathbf{C}$ so that $L^2(K) = L^2$ of the unit circumference, which we will abbreviate simply as L^2. Then the bilateral shift U on either L^2 or $L^2(K)$ is defined by $Uf(\theta) = e^{i\theta}f(\theta)$. We use χ_A to denote the characteristic function of a subset A of the unit circumference.

First, we present the following lemma.

LEMMA 3.5. *Let* $1 \leqq p \leqq \infty$, *let* $c > 0$ *and let* $g \in H^p$. *A necessary and sufficient condition that the function*

$$h(z) = \frac{g(z)}{(z-a)^c}$$

(with $|a| = 1$) *be in* H^p *is that* $h(e^{i\theta})$ *belong to* L^p *of the unit circumference.*

Proof. This is a slight generalization of a well known result and we leave the straightforward calculation to the interested reader.

THEOREM 3.6. *The closed nonzero inner invariant subspaces* S *of* $U - a$ $(a \in \mathbf{C})$ *on* L^2 *are of the following form:*

(i) *If* $|a| < 1$, *then* S *is inner invariant for* $U - a$ *iff* $S = \chi_A L^2$ *where* A *is a Baire set of the unit circumference.*

(ii) *If* $|a| \geqq 1$, *then* S *is inner invariant for* $U - a$ *iff* $S = \chi_A L^2$ (for A a Baire set) *or* $S = FH^2$ (for F a measurable function on the unit circumference with modulus 1); i.e., *iff* S *is invariant for* U.

Proof. It is well known that the invariant subspaces of U are either of the form FH^2 or $\chi_A L^2$. We now divide the proof into several cases.

Case 1. Let $|a| < 1$. It is clear that invariant subspaces of the form FH^2 are not inner invariant for $U - a$ since $1/(e^{i\theta} - a) \in L^2$ but not to H^2.

Now consider subspaces of the form $\chi_A L^2$ but these subspaces are always inner invariant since $1/(e^{i\theta} - a)$ is bounded in L^2 (for $|a| > 1$ too).

Case 2. Let $|a| > 1$. Here $1/(e^{i\theta} - a)$ belongs not only to L^2 but also to H^2 (since it is analytic). Thus all invariant subspaces of the form FH^2 (for F with modulus 1) are inner invariant for $U - a$. The space $\chi_A L^2$ is inner invariant for $U - a$ when $|a| > 1$, as was mentioned previously.

Case 3. Let $|a| = 1$. In this case $1/(e^{i\theta} - a) \notin L^2$. We consider, for f in L^2 and g in H^2, the following:

$$f(\theta) = F(\theta)g(\theta)/(e^{i\theta} - a). \tag{10}$$

Since $f \in L^2$ and $|F(\theta)| = 1$, equation (10) implies that $g(\theta)/(e^{i\theta} - a) \in L^2$ with g in H^2. Lemma 3.5 then tells us that $g(\theta)/(e^{i\theta} - a)$ is in H^2 so that $f \in FH^2$ and thus FH^2 is inner invariant.

Lastly, for some f and g in L^2 we consider $f(\theta) = \chi_A(\theta)g(\theta)/(e^{i\theta} - a) = \chi_A(\theta)(\chi_A(\theta)g(\theta)/e^{i\theta} - a) = \chi_A(\theta)f(\theta)$ so that $f \in \chi_A L^2$ and thus $\chi_A L^2$ is inner invariant when $|a| = 1$.

For background material on $L^2(K)$ the reader is referred to Fillmore **[5]**, pages 31–44. We will use theorems from that work. Loosely speaking $L^2(K)$, for K a Hilbert space, consists of all measurable functions f from the unit circumference into K such that

$$\|f\|_{L^2(K)} = \frac{1}{2\pi}\int_0^{2\pi} \|f(\theta)\|_K^2\, d\theta < \infty.$$

Our goal is to describe the inner invariant subspaces of $U - a$ on $L^2(K)$. We will write M for a reducing subspace of U on $L^2(K)$. It is known that M has the form $M = \{f \mid f(\theta) \in M(\theta) \text{ a.e.}\}$ where $M(\theta) = P(\theta)K$ and P is the projection operator from $L^2(K)$ onto M. If S is an invariant subspace of U, it is known that $S = M \oplus N$ where M reduces U, the space N is invariant for U and $\bigcap_{k=0}^{\infty} U^k(N) = \{0\}$. More specifically, we can state that $N = VH^2(R)$ where R is a closed subspace of K and V is a partial isometry on $L^2(K)$, with initial space $L^2(R)$, that commutes with U.

We now prove the following series of lemmas.

LEMMA 3.7. *If g is in $H^2(K)$ and $|a| = 1$ then a necessary and sufficient condition for $h(\theta) = g(\theta)/(e^{i\theta} - a)$ to belong to $L^2(K)$ is that $h(\theta)$ be in $H^2(K)$.*

Proof. This is clearly a further extension of Lemma 3.5. If h is in $H^2(K)$ then it automatically belongs to $L^2(K)$; so assume g is in $H^2(K)$ and $g/(e^{i\theta} - a)$ is in $L^2(K)$. Let $\{b_n\}_{n=0}^{\infty}$ be an orthonormal basis for K and let $g_n(\theta) = (g(\theta), b_n)$; i.e., the inner product of g with b_n. Then it is known that $g(\theta) = \Sigma_n g_n(\theta)b_n$ (convergence of this sum being in the norm of K) and that g_n is in H^2. Since $\{b_n\}_{n=0}^{\infty}$ is an orthonormal basis for K we get $\|g(\theta)\|_K^2 = \Sigma_{n=0}^{\infty}|g_n(\theta)|^2$. From this we get (using Fatou's Lemma) that

$$\|g/(e^{i\theta}-a)\|^2_{L^2(K)} = \frac{1}{2\pi}\int_0^{2\pi} \|g(\theta)/(e^{i\theta}-a)\|^2_K\, d\theta$$
$$= \frac{1}{2\pi}\int_0^{2\pi} \sum_{n=0}^{\infty} |g_n(\theta)/(e^{i\theta}-a)|^2\, d\theta$$
$$= \sum_{n=0}^{\infty} \frac{1}{2\pi}\int_0^{2\pi} |g_n(\theta)/(e^{i\theta}-a)|^2\, d\theta$$
$$= \sum_{n=0}^{\infty} \|g_n/(e^{i\theta}-a)\|^2_{L^2}.$$

Since $g/(e^{i\theta}-a)$ is in $L^2(K)$ we can conclude that $g_n/(e^{i\theta}-a)$ is in L^2 for all n. Thus we now have g_n in H^2 with $g_n/(e^{i\theta}-a)$ in L^2 (for $|a|=1$) and so we can conclude, by Lemma 3.5, that $g_n/(e^{i\theta}-a)$ is in H^2 for all n. From this we get finally that $g/(e^{i\theta}-a)$ is in $H^2(K)$.

LEMMA 3.8. (i) *If* $|a|<1$, *then* N *is never inner invariant for* $U-a$. (ii) *If* $|a|\geqq 1$, *then* N *is always inner invariant for* $U-a$.

Proof. We are assuming $N = VH^2(R)$ with V and R as described above.

Case 1. Let $|a|<1$ and x be in R. Then $x/(e^{i\theta}-a)$ is in $L^2(R)\backslash H^2(R)$. Now V maps $L^2(R)$ into $L^2(K)$ so that $V(x/e^{i\theta}-a)$ is in $L^2(K)$. Thus

$$(e^{i\theta}-a)V(x/e^{i\theta}-a) = Vx$$

since V commutes with U and hence with $U-a$. Since x is in R and R is contained in $H^2(R)$ we get Vx in $H^2(R)$. This contradicts the inner invariance of N.

Case 2. Let $|a|>1$. Assume there is an f in $L^2(K)$ such that

$$(e^{i\theta}-a)f(\theta) = V(\theta)g(\theta) \tag{11}$$

for some g in $H^2(R)$. Since $|a|>1$, the function $1/(e^{i\theta}-a)$ is analytic and bounded in the unit disk so that $g/(e^{i\theta}-a)$ is in $H^2(R)$. Thus (11) implies that

$$f(\theta) = \frac{1}{e^{i\theta}-a}V(\theta)g(\theta) = \frac{1}{e^{i\theta}-a}V(\theta)\left[(e^{i\theta}-a)\frac{g(\theta)}{e^{i\theta}-a}\right]$$
$$= V(\theta)\frac{g(\theta)}{e^{i\theta}-a}$$

since V commutes with $U-a$ on $L^2(R)$. We already know that $g(\theta)/(e^{i\theta}-a)$ is in $H^2(R)$ so that f is in $VH^2(R)$. Thus N is inner invariant for $U-a$ when $|a|>1$.

Case 3. Let $|a|=1$. Assume f is in $L^2(K)$ and g is in $H^2(R)$. We must show

$$f(\theta)=\frac{1}{e^{i\theta}-a}V(\theta)g(\theta)\in VH^2(R), \tag{12}$$

but

$$\begin{aligned}\|f\|^2_{L^2(K)}&=\frac{1}{2\pi}\int_0^{2\pi}\left\|\frac{1}{e^{i\theta}-a}V(\theta)g(\theta)\right\|^2_K d\theta\\&=\frac{1}{2\pi}\int_0^{2\pi}\left|\frac{1}{e^{i\theta}-a}\right|^2\|V(\theta)g(\theta)\|^2_K d\theta\\&=\frac{1}{2\pi}\int_0^{2\pi}\left|\frac{1}{e^{i\theta}-a}\right|^2\|g(\theta)\|^2_R d\theta\end{aligned}$$

since $V(\theta)$ is a partial isometry a.e. from R into K.

Thus

$$\begin{aligned}\|f\|^2_{L^2(K)}&=\frac{1}{2\pi}\int_0^{2\pi}\|g(\theta)/(e^{i\theta}-a)\|^2_R d\theta\\&=\|g/(e^{i\theta}-a)\|^2_{L^2(R)}.\end{aligned}$$

Since f is in $L^2(K)$ this means $g/(e^{i\theta}-a)$ is in $L^2(R)$. Since V commutes with $U-a$ on $L^2(K)$ we can rewrite (12) as

$$f(\theta)=V(\theta)g(\theta)/(e^{i\theta}-a).$$

Since g is in $H^2(R)$ and $g/(e^{i\theta}-a)$ is in $L^2(R)$, an application of Lemma 3.7 tells us that $g/(e^{i\theta}-a)$ is in $H^2(R)$ so that f must be in $VH^2(R)$.

These lemmas now allow us to completely describe the inner invariant subspaces of the bilateral shift.

THEOREM 3.9. *The closed nonzero inner invariant subspaces S of $U-a$ on $L^2(K)$ are*:

(i) *If $|a|<1$ then S is inner invariant for $U-a$ iff $S=M$ where M is a reducing subspace for U*, i.e., *iff S is reducing for $U-a$.*

(ii) *If $|a|\geqq 1$ then S is inner invariant for $U-a$ iff $S=M\oplus N$ with $N=VH^2(R)$*, i.e., *iff S is invariant for $U-a$.*

Proof. Lemma 3.8 (i) tells us N is not inner invariant for $U-a$ when $|a|<1$, but M being a reducing subspace for U is clearly also a reducing subspace for $U-a$. Since $U-a$ is one-to-one, Proposition 2.10 tells us M is inner invariant for $U-a$. The question is, can $M \oplus N$ be inner invariant in these circumstances?

Since N is not inner invariant, there is an x in $L^2(K)\backslash N$ such that $(U-a)x \in N$. We know $L^2(K) = M \oplus M^\perp$ with $N \subseteq M^\perp$. Thus we can write $x = m_1 + m_2$ where $m_1 \in M$ and $m_2 \in M^\perp$, and $(U-a)x = (U-a)m_1 + (U-a)m_2 \in N$. We know $(U-a)m_1$ is in M since M is reducing for $U-a$. Since N is orthogonal to M this means $(U-a)m_1 = 0$ which in turn means $m_1 = 0$ since $U-a$ is one-to-one. Thus $x = m_2 \in M^\perp \backslash N$ so that $x \notin M \oplus N$ yet $(U-a)x \in M \oplus N$ so that $M \oplus N$ cannot be inner invariant.

If $|a| \geqq 1$ then Lemma 3.8 (ii) tells us N is inner invariant. It is also clear that M is inner invariant. We will now show that $M \oplus N$ must also be inner invariant.

Assume $x \in L^2(K)$ and

$$(U-a)x \in M \oplus N.$$

The element x can be expressed as $x = m_1 + m_2$ as before, so that $(U-a)x$ is again equal to $(U-a)m_1 + (U-a)m_2$. Since M reduces $U-a$, we know

$$(U-a)m_2 \in M^\perp \quad \text{and} \quad (U-a)m_1 \in M,$$

so that $(U-a)m_2 = (U-a)x - (U-a)m_1 \in M \oplus N$. Therefore $(U-a)m_2 \in M^\perp \cap (M \oplus N) = N$. Since N is inner invariant for $U-a$ we can conclude that m_2 is in N.

Let us look a bit more closely at part (ii) of the previous theorem. The subspace S has the form $M \oplus N$. If K were one dimensional, i.e., $K = \mathbf{C}$ then this theorem should reduce to Theorem 3.6 (ii). It is not difficult to show that in this case the subspace M reduces to $\chi_A L^2$ (for A a Baire set) and N reduces to FH^2 (for F a measurable function of modulus 1). At this point there seems to be an inconsistency. In Theorem 3.6 we have $S = FH^2$ *or* $S = \chi_A L^2$ (but not both together) while in Theorem 3.9 we seem to allow the possibility that $S = FH^2 \oplus \chi_A L^2$. That there is no inconsistency follows from the following:

Assume $S = FH^2 \oplus \chi_A L^2$. This means FH^2 is contained in the orthogonal complement of $\chi_A L^2$, i.e., if B is the complement of A on the unit circumference, then $FH^2 \subseteq \chi_B L^2$. If A has measure zero then there is no problem since, in this case, $\chi_A L^2 = \{0\}$ and $S = FH^2$. Thus, let us assume that the measure of A is positive. Let $f = Fg \in FH^2$ with g in H^2. Then since $FH^2 \subseteq \chi_B L^2$ we must necessarily conclude that $f = 0$ on A, i.e., $f = 0$ on a set of positive measure. Since F has modulus 1 a.e., it is necessary that $g = 0$ on a set of positive measure. This is impossible if g is not identically zero.

Since $g \in H^2$, $g \in H^1$ and the log of an H^1 function is integrable. But $\log|g| = -\infty$ on a set of positive measure which contradicts the integrability of $\log|g|$. Hence g must be identically zero. This implies $FH^2 = \{0\}$ so that $S = \chi_A L^2$. Thus when $K = \mathbf{C}$ it cannot happen that $S = M \oplus N$ nontrivially.

Lastly, there is a question that still remains open, namely, what are the inner invariant subspaces of the unilateral shift and its translates on $H^2(K)$? The invariant subspaces we know: they are of the form $N = VH^2(R)$ where V is a partial isometry on $H^2(K)$ commuting with the unilateral shift and R is a closed subspace of K. The argument used in Lemma 3.8, Cases 2 and 3, with minor modification, suffices to tell us that all invariant subspaces of U are inner invariant for $U - a$ when $|a| \geqq 1$, but what happens when $|a| < 1$? Things are not too clear in this case. In the one-dimensional case $S = FH^2$ and S is inner invariant for $U - a$ iff $F(a) \neq 0$. What condition on the partial isometry V on $H^2(R)$ reduces to $V(a) \neq 0$ when K is one-dimensional?

4. A characterization of the Volterra operator.

We characterize integration abstractly on an arbitrary Hilbert space. Then we apply the work of the previous section to obtain the inner invariant subspace structure for a translate of the integration operation.

For the sake of convenience we will write $A \sim B$ to mean A is unitarily equivalent to B.

DEFINITION 4.1. The *Volterra operator* V is defined by $Vf(x) = \int_0^x f(y)dy$ for f, in $D(V)$ where $D(V) = \{f \in L^2(0,\infty) \mid Vf \in L^2(0,\infty)\}$.

A straightforward application of the Fubini Theorem tells us that

$$(V - a)^{-1}f(x) = -\frac{1}{a}f(x) - \frac{1}{a^2}\int_0^x e^{(x-t)/a}f(t)\,dt$$

for a in $\mathbf{C} - \{0\}$ where f is in $D(V - a)^{-1}$. It is then not difficult to show that $(V - a)^{-1}$ is bounded and defined everywhere in $L^2(0,\infty)$ for $\operatorname{Re}(a) < 0$.

PROPOSITION 4.2. *The Volterra operator V is a closed densely defined unbounded operator.*

Proof. Since $(V+1)^{-1}$ is bounded and defined everywhere, it is closed. Therefore $V+1$ and V are also closed.

Sarason **[13]** showed that $(V+1)^{-1}$ on $L^2(0,\infty)$ was unitarily equivalent to $(1+U)/2$ on H^2 where U is the shift operator. Since $(V+1)^{-1} \sim (U+1)/2$ we get $V+1 \sim 2(U+1)^{-1}$. Therefore $V+1$ (and hence V) is densely defined if $(U+1)^{-1}$ is, but $D(U+1)^{-1} = (U+1)D(U) = (z+1)H^2$ and Proposition 3.4 (ii) assures us that $(z+1)H^2$ is dense in H^2.

To show V is unbounded, define $f_a(x)$ to be 1 if $0 \leqq x \leqq a$ and $-a(a+1)/(x+1)^2$ if $x > a$ where $1 < a < \infty$. Then $\| Vf_a \| > a \| f_a \|$ for $a = 2,3,4,\cdots$ so that V is indeed unbounded.

In order to characterize the Volterra operator it will be useful to generalize the concept of a symmetric operator and its Cayley transform.

DEFINITION 4.3. A linear operator A on a Hilbert space H is called *b-symmetric* if $e^{ib}A \subset A^*$ where b is real. If $b = \pi$ then A is called skew-symmetric.

We now note that virtually all results about symmetric operators and their Cayley transforms also hold for b-symmetric operators. Simply substitute b-symmetric for symmetric in their proofs. A good reference for symmetric operators is Akhiezer and Glazman **[1]**.

Recall that if B is a symmetric operator then its Cayley transform C is defined by $C = (B-z)(B-\bar{z})^{-1}$ where $\mathrm{Im}(z) > 0$ and C is a partial isometry.

We now define a Cayley transform for a b-symmetric operator. If A is b-symmetric then $B = e^{ib/2}A$ is symmetric and has a Cayley transform C where (for $\mathrm{Im}(z) > 0$)

$$\begin{aligned} C &= (B-z)(B-\bar{z})^{-1} \\ &= (e^{ib/2}A - z)(e^{ib/2}A - \bar{z})^{-1} \\ &= (A - e^{-ib/2}z)(A - e^{-ib/2}\bar{z})^{-1}. \end{aligned}$$

Now let $w = e^{-ib/2}z$ so that

$$C = C(A) \equiv (A-w)(A - e^{-ib}\bar{w})^{-1} \tag{13}$$

for $\mathrm{Im}(e^{ib/2}w) > 0$.

DEFINITION 4.4. We call the partial isometry $C = C(A)$, i.e., (13) above, the *Cayley transform* of the b-symmetric operator A. The domain $D(C) = (A - e^{-ib}\bar{w})D(A)$ and the range $R(C) = (A-w)D(A)$ where $\mathrm{Im}(e^{ib/2}w) > 0$.

DEFINITION 4.5. If A is a b-symmetric operator, we define the *defect spaces* H_w^+ and H_w^- by

$$H_w^+ = D(C)^\perp = \{(A - e^{-ib}\bar{w})D(A)\}^\perp;$$
$$H_w^- = R(C)^\perp = \{(A - W)D(A)\}^\perp$$

where $\operatorname{Im}(e^{ib/2}w) > 0$.

DEFINITION 4.6. The *defect indices* of a b-symmetric operator A are (p, q) where $p =$ dimension H_w^+ and $q =$ dimension H_w^-. From Definition 4.6 it is clear that (p, q) are also the defect indices of the Cayley transform $C(A)$.

PROPOSITION 4.7. *Let A be b-symmetric and* $\operatorname{Im}(e^{ib/2}w) > 0$. *Then* $H_w^+ = \{x \in H \mid A^*x = e^{ib}wx\}$, $H_w^- = \{x \in H \mid A^*x = \bar{w}x\}$ *and* $D(A) \cap H_w^+ = D(A) \cap H_w^- = H_w^+ \cap H_w^- = \{0\}$.

THEOREM 4.8. *Let A be b-symmetric, then A is closed iff* $D(A^*) = D(A) \oplus H_w^+ \oplus H_w^-$ *where* $\operatorname{Im}(e^{ib/2}w) > 0$. *The algebraic direct sum above is not necessarily orthogonal.*

Proof. $(\Rightarrow)$ This direction is well known and its proof can be found in Akhiezer and Glazman **[1]**, Volume II, Page 98.

$(\Leftarrow)$ This part is new and its proof is due to Robert Waterman. We assume that A is not closed and that $C = C(A)$ is the Cayley transform of A. We will also make use of the following equivalent statements:

A is not closed $\Leftrightarrow$ $C(A)$ is not closed $\Leftrightarrow$ $D(C)$ is not closed $\Leftrightarrow$ $R(C)$ is not closed.

Since C is not closed we can get a smallest closed extension $\hat{C}$ of C by considering the naturally induced partial isometry on $\overline{D(C)}$ that extends C. It is then clear that $(C-1)D(C) = D(A)$ and $D(A)$ is dense in H. Thus $(\hat{C}-1)D(\hat{C})$ is also dense in H since $(\hat{C}-1)D(\hat{C}) = (\hat{C}-1)\overline{D(C)} \supset (C-1)D(C)$. In this case $\hat{A} \equiv (e^{-ib}\bar{w}\hat{C} - w)(\hat{C}-1)^{-1}$ is a closed b-symmetric operator. We know $\hat{A} \supsetneq A$ since $\hat{C} \supsetneq C$ (see **[1]**, Volume II, page 96). Since $\hat{A}$ is a closed b-symmetric operator, the first half of this theorem tells us $D(\hat{A}^*) = D(\hat{A}) \oplus H_w^+(\hat{A}) \oplus H_w^-(\hat{A})$. From Definition 4.5 we get $H_w^+(\hat{A}) = D(\hat{C})^\perp = \overline{D(C)}^\perp = D(C)^\perp = H_w^+(A)$ and $H_w^-(\hat{A}) = R(\hat{C})^\perp = \overline{R(C)}^\perp = R(C)^\perp = H_w^-(A)$. Therefore

$$D(\hat{A}^*) = D(\hat{A}) \oplus H_w^+(A) \oplus H_w^-(A) \supsetneq D(A) \oplus H_w^+(A) \oplus H_w^-(A) \tag{14}$$

since $A \subsetneqq \hat{A}$. But $A \subsetneqq \hat{A}$ implies $\hat{A}^* \subsetneqq A^*$ so that $D(\hat{A}^*) \subsetneqq D(A^*)$. This completes the proof since by (14) $D(A^*) \supsetneqq D(\hat{A}^*) \supsetneqq D(A) \oplus H_w^+ \oplus H_w^-$.

DEFINITION 4.9. A b-symmetric (or partial isometry) A is *simple* if there does not exist a closed subspace K of H invariant under A such that A restricted to K is b-adjoint (or unitary). An operator F is *b-adjoint* if $e^{ib}F = F^*$.

REMARK. A b-symmetric operator A is simple iff $C(A)$ is simple.

Let H be a separable Hilbert space and let $\{e_i\}_{i=1}^{\infty}$ be an orthonormal basis for H. We define an operator $C^{\#}$ by

$$C^{\#} e_k = e_{k+1} \quad \text{for} \quad k = 1, 2, 3, \cdots.$$

It is clear that $C^{\#}$ is a partial isometry with defect indices $(0, 1)$. As a matter of fact $C^{\#}$ is the shift operator. A short calculation then shows that $(C^{\#} - 1)D(C^{\#})$ is dense in H so that

$$A^{\#} = (e^{-ib}\bar{w}C^{\#} - w)(C^{\#} - 1)^{-1}$$

for $\operatorname{Im}(e^{ib/2}w) > 0$ is a b-symmetric operator.

THEOREM 4.10. (J. von Neumann) *If a simple b-symmetric operator A on H has defect indices* $(0, 1)$ *then H is separable and A is unitarily equivalent to $A^{\#}$.*

REMARK. This important theorem was originally proved for symmetric operators. Its present form is a direct generalization of this and its proof is similar.

THEOREM 4.11. *The Volterra operator V on $L^2(0, \infty)$ is a simple skew-symmetric operator with defect indices* $(0, 1)$.

Proof. In Sarason **[13]**, the following was shown: If U is the shift operator on H^2 then $(V + 1)^{-1} \sim (U + 1)/2$. A little calculation then gives $U \sim (1 - V)(1 + V)^{-1} = -C(V)$. Therefore V is simple with defect indices $(0, 1)$ since U is, and V is skew-symmetric since it is the "anti"-Cayley transform of U (i.e., $b = \pi$ and $w = 1$ in (13) above).

As a corollary of the last two theorems we now have the abstract characterization of the Volterra operator.

COROLLARY 4.12. *Any skew-symmetric operator A, defined on a*

Hilbert space H, that has defect indices $(0,1)$ is unitarily equivalent to the Volterra operator V on $L^2(0,\infty)$.

Proof. By Theorems 4.11 and 4.10 we have $V \sim A^{\#}$ and $A \sim A^{\#}$ so that $V \sim A$.

Since $Vf(x) = \int_0^x f(y)\,dy$ for f in $D(V) = \{f \in L^2 \mid Vf \in L^2\}$, we easily get that $V^{-1}g(x) = d/dx\, g(x)$ for g in $D(V^{-1}) = \{f \text{ in } L^2 \mid f' \in L^2,$ $f \in A.C.$ and $f(0) = 0\}$ where $f'(x) = d/dx\, f(x)$ and $A.C.$ is the set of absolutely continuous functions on $L^2 = L^2(0,\infty)$.

THEOREM 4.13. *The operator V is unitarily equivalent to V^{-1}.*

Proof. In the proof of Theorem 4.11 we showed that $C(V) \sim -U$. Thus since $U \sim e^{ib}U$ for real b we get that

$$\begin{aligned} V &= (-C(V)-1)(C(V)-1)^{-1} \\ &\sim (U-1)(-U-1)^{-1} \\ &= (1-U)(1+U)^{-1} \\ &\sim (1-e^{ib}U)(1+e^{ib}U)^{-1} \\ &= (1+U)(1-U)^{-1} \quad \text{if} \quad b = \pi \\ &= \{(1-U)(1+U)^{-1}\}^{-1} \\ &\sim \{V\}^{-1} = V^{-1}. \end{aligned}$$

We finish this section with some remarks about the nature of the adjoint V^* of the Volterra operator V. On $L^2(0,1)$, $Vf = \int_0^x f(y)\,dy$ and $V^*f = \int_x^1 f(y)\,dy$. At a first glance it might seem natural to suppose that V^* on $L^2(0,\infty)$ would be similarly defined, that is, $V^*f = \int_x^\infty f(y)\,dy$. This is not the case.

Since by Theorems 4.11 and 4.2, V is a closed skew-symmetric operator, we can use Theorem 4.8 to deduce that $D(V^*) = D(V) \oplus H_1^+ \oplus H_1^-$. The defect indices of V are $(0,1)$ so that $H_1^+ = \{0\}$. We now need to describe the defect space H_1^-. The function e^{-t} in $L^2(0,\infty)$ is not in $D(V)$ since $V(e^{-t}) = 1 - e^{-t}$ and the constant 1 does not belong to $L^2(0,\infty)$. On the other hand e^{-t} does belong to $D(V^*)$ since

$$(Vf, e^{-t}) = \int_0^\infty Vf(t)e^{-t}\,dt = \lim_{b\to\infty} \int_0^b f(t)e^{-t}\,dt = (f, e^{-t})$$

for f in $D(V)$. Thus e^{-t} is in $D(V^*)$ and $V^*e^{-t} = e^{-t}$ so that e^{-t} is in H_1^-. Since H_1^- is one dimensional, we conclude $H_1^- = [e^{-t}]$.

PROPOSITION 4.17. *The adjoint V^* can be described by $V^*f(x) = \lim_{b\to\infty} \int_x^b f(y)\,dy$ for f in $D(V^*) = D(V) \oplus [e^{-t}]$.*

Proof. From above we know $V^*e^{-x} = e^{-x} = \lim_{b\to\infty} \int_x^b e^{-t}\,dt$. If f is in $D(V)$ then (using integration by parts) we have $(Vf, f) = \lim_{b\to\infty} |Vf(b)|^2 - (f, Vf)$ which implies that $\lim_{b\to\infty} Vf(b) = 0$ for otherwise $\|Vf\| = \infty$ which contradicts f being in $D(V)$. Hence

$$0 = \lim_{b\to\infty} \int_0^b f(t)\,dt = \int_0^x f(t)\,dt + \lim_{b\to\infty} \int_x^b f(t)\,dt$$

so that

$$\lim_{b\to\infty} \int_x^b f(t)\,dt = -Vf(x) = V^*f(x)$$

since V is skew-symmetric.

We can actually describe $D(V^*)$ in slightly different terms, namely $D(V^*) = \{f \in L^2(0,\infty) \mid V^*f \in L^2(0,\infty)\}$. It is obvious that $D(V^*)$ is contained in this set, which we will call D^*. We need only show $D^* \subset D(V^*)$. To do this consider f in $D(V)$ and g in D^*. Then

$$\begin{aligned}(Vf, g) &= \int_0^\infty Vf(x)\overline{g(x)}\,dx = \lim_{a\to\infty} \int_0^a Vf(x)\overline{g(x)}\,dx \\ &= \lim_{a\to\infty} \int_0^a f(x) \int_x^a \overline{g(y)}\,dy\,dx\end{aligned}$$

and

$$\lim_{a\to\infty} \int_0^a f(x) \int_x^a \overline{g(y)}\,dy\,dx = \int_0^\infty f(x)\overline{Wg(x)}\,dx$$

where $W = V^*$ with $D(W) = D$, so that

$$(Vf, g) = (f, Wg)$$

for f in $D(V)$ and g in D^*.

5. Inner invariant subspaces for Volterra type operators. We are now ready to apply the results of the past few sections to the unbounded Volterra operator.

THEOREM 5.1. *There exists a surjective isometry* (i.e., a unitary map) *I from H^2 to $L^2(0,\infty)$ such that the closed nonzero inner invariant subspaces of $V+1$ are precisely $\{I(gH^2) \mid g$ is an inner function$\}$. The same result holds for $V^{-1}+1$ in place of $V+1$.*

Proof. Recall that V is a skew-symmetric simple operator with defect indices $(0,1)$. Thus $(1-V)(1+V)^{-1}$ is the shift operator since this is the Cayley transform of V. Therefore there is a unitary map I from H^2 onto $L^2(0,\infty)$ such that $IU=(1-V)(1+V)^{-1}I$ where U is the shift operator on H^2. Thus $IU+I=2(1+V)^{-1}I$ so that $I\frac{1}{2}(U+1)=(1+V)^{-1}I$. Thus K is an inner invariant subspace of $\frac{1}{2}(U+1)$ iff $I(K)$ is an inner invariant subspace of $(1+V)^{-1}$, and $\frac{1}{2}(U+1)$ has the same inner invariant subspaces as $U+1$. Theorem 3.3 then tells us that K must equal gH^2 with g an inner function. Thus $\{I(gH^2) \mid g \text{ inner}\}$ is the set of inner invariant subspaces for $(V+1)^{-1}$, and by Theorem 2.5 for $V+1$ too.

To get the second part of this theorem, notice that by Theorem 4.13 we have $V \sim V^{-1}$ so that $V+1 \sim V^{-1}+1$.

REMARK. It should be pointed out that while the unitary maps from H^2 onto $L^2(0,\infty)$ may be different for $V+1$ and $V^{-1}+1$ it happens that it is also possible to pick a single unitary map I in such a way that the spaces $I(gH^2)$ are inner invariant for both $V+1$ and $V^{-1}+1$ simultaneously. For an example of such a unitary map see Sarason **[13]**.

If we now look back at Example 2.6 we will notice that the spaces $L^2(a,1)$ for $0 \leqq a \leqq 1$ are inner invariant for the Volterra operator V on $L^2(0,1)$. Since our Volterra operator V on $L^2(0,\infty)$ is a natural extension of this it seems appropriate to inquire if the spaces $L^2(a,\infty)$ are inner invariant for V.

PROPOSITION 5.2. *The subspaces $L^2(a,\infty)$, for $0<a<\infty$, of $L^2(0,\infty)$ are inner invariant for $V+1$.*

Proof. We will show that $L^2(a,\infty)$ is invariant for both $V+1$ and $(V+1)^{-1}$. Then an application of Theorem 2.5 shows that $L^2(a,\infty)$ is inner invariant.

Since $(V+1)^{-1}$ is bounded and defined everywhere, we need only show that $(V+1)^{-1}(L^2(a,\infty)) \subseteq L^2(a,\infty)$ but $(V+1)^{-1}f(x)=f(x)-\int_0^x e^{t-x}f(t)\,dt$ for f in $L^2(a,\infty)$. Since $\int_0^x e^{t-x}f(t)\,dt$ has its support in $[a,\infty)$ if f does, we can conclude that $(V+1)^{-1}f$ is indeed in $L^2(a,\infty)$.

To show $L^2(a,\infty)$ is invariant for $V+1$, we have

$$\begin{aligned} D(V+1)\cap L^2(a,\infty) &= D(V)\cap L^2(a,\infty) \\ &= \left\{f\in L^2(0,\infty) \,\middle|\, \int_0^x f(y)\,dy \in L^2(0,\infty)\right\} \cap L^2(a,\infty) \\ &= \left\{f\in L^2(a,\infty) \,\middle|\, \int_a^x f(y)\,dy \in L^2(a,\infty)\right\} \end{aligned}$$

and it is clear that this is dense in $L^2(a,\infty)$ iff $D(V)$ is dense in $L^2(0,\infty)$. Thus condition (ii) in the definition of invariance is satisfied.

As in the first part of the proof, $\int_0^x f(y)\,dy$ has its support in $[a,\infty)$ if f does. Thus if f is in $D(V)\cap L^2(a,\infty)$ then $(V+1)f$ must be in $L^2(a,\infty)$.

REMARK. From Theorem 5.1, there must be an inner function g_a such that $I(g_aH^2)=L^2(a,\infty)$. Using the isometry I given by Sarason **[13]** we see that $g_a(z)=\exp(a(z+1)/(z-1))$ is precisely the inner function we want.

PROPOSITION 5.3. *The inner invariant subspaces of $V+1$ are all inner invariant for V.*

Proof. In Proposition 4.2 we showed that $(V+1)^{-1}\sim(1+U)/2$ so that $V+1\sim 2(1+U)^{-1}$ or $V\sim 2(1+U)^{-1}-1$. Thus $I(gH^2)$ is inner invariant for V iff gH^2 is inner invariant for $2(1+U)^{-1}-1$. Since gH^2 is inner invariant for $(1+U)/2$ it is inner invariant for $2(1+U)^{-1}$ and therefore necessarily invariant for $2(1+U)^{-1}$. We must now show that gH^2 is also invariant for $2(1+U)^{-1}-1$.

Clearly $D(2(1+U)^{-1}-1)=D(1+U)^{-1}=(z+1)H^2$. By Proposition 3.4 we know that $(z+1)H^2$ is dense in H^2. Also since gH^2 is inner invariant for $U+1$ we get $(z+1)H^2\cap gH^2=(z+1)gH^2$ (we leave the proof of this fact to the interested reader). Now condition (ii) in the definition of invariance tells us that we must show $\{((z+1)H^2\cap gH^2)+[2(U+1)^{-1}-1]((z+1)H^2\cap gH^2)\}^-=gH^2$. Since $(z+1)H^2\cap gH^2=g(z+1)H^2$ we know that this subspace is dense in gH^2 because $(z+1)H^2$ is dense in H^2. Thus condition (ii) will be satisfied if we can show $[2(U+1)^{-1}-1](g(z+1)H^2)\subseteq gH^2$ (this is condition (i) of the definition of invariance) but this is true iff $2(U+1)^{-1}g(z+1)H^2\subseteq gH^2$ which is clearly true. Thus $I(gH^2)$ is invariant for V.

To complete this proof we now show that $I(gH^2)$ is also invariant for V^{-1} and then call on Theorem 2.5. We do this by showing $V^{-1}+1\sim$

$2(U-1)^{-1}$ so that $V^{-1} \sim 2(U-1)^{-1}-1$. Then we use the argument given above to show that gH^2 is indeed invariant for $2(U-1)^{-1}-1$.

To show $V^{-1}+1 \sim 2(U-1)^{-1}$, consider the fact that $(1+V)^{-1} = 1-K$ where $Kg(x) = \int_0^x e^{t-x} g(t)\,dt$. We then consider the differential equation $(iV^{-1}+i)f = g$, that is, $if'(x)+if(x) = g(x)$. The solution of this equation is $f(x) = iKg(x)$. Thus $f(x) = (iV^{-1}+i)^{-1}g(x) = iKg(x)$ or $-K = (V^{-1}+1)^{-1}$. Since $1+V \sim 2(1+U)^{-1}$ we have $1-K = (1+V)^{-1} \sim (U+1)/2$ or $-K \sim (U-1)/2$. Since $(V^{-1}+1)^{-1} = -K$ we conclude $(V^{-1}+1) \sim 2(U-1)^{-1}$.

We now examine one difference between the bounded Volterra operator on $L^2(0,1)$ and the unbounded Volterra operator on $L^2(0,\infty)$. In reference **[10]**, Kalisch showed that V on $L^2(0,1)$ is a unicellular operator. For the unbounded Volterra operator, it is not.

It is not difficult to show this. We need only show that the spaces $I(gH^2)$ for g an inner function are not totally ordered. This is true iff the spaces gH^2 are themselves not totally ordered. Consider two subspaces g_1H^2 and g_2H^2 of H^2 with

$$g_1(z) = \frac{\bar{a}}{|a|}\cdot\frac{a-z}{1-\bar{a}z} \quad\text{and}\quad g_2(z) = \frac{\bar{b}}{|b|}\cdot\frac{b-z}{1-\bar{b}z}$$

where a and b are nonzero complex numbers such that $a \neq b$. It is then a routine calculation to show that neither $g_1H^2 \subseteq g_2H^2$ nor $g_2H^2 \subseteq g_1H^2$. Thus the unbounded Volterra operator is not unicellular.

We do know all of the inner invariant subspaces for $V+1$ but do we know all of them for V? The answer at present is no, though we can (and will) exhibit a rather large set of inner invariant subspaces of V that are not inner invariant for $V+1$.

EXAMPLE 5.4. Let P_n be the set of functions of the form $p(t)e^{-t}$ where $p(t)$ is a polynomial of degree at most n for n a positive integer. Thus P_n is an $n+1$ dimensional subspace of $L^2(0,\infty)$. We will show that P_n is inner invariant for V^{-1} and therefore also for V. Now $V^{-1} = d/dt$ with $D(V^{-1}) = \{f \in L^2(0,\infty) \mid f \in A.C.,\ f' \in L^2(0,\infty) \text{ and } f(0) = 0\}$.

Claim 1. $P_n \cap D(V^{-1}) = \{p(t)e^{-t} \mid p(0) = 0$ and p is a polynomial of degree at most $n\}$.

Proof of claim. Straightforward.

Claim 2. $V^{-1}(P_n \cap D(V^{-1})) \subseteq P_n$.

Proof of claim. We have $V^{-1}p(t)e^{-t} = (p'(t) - p(t))e^{-t}$ which is in P_n whether or not $p(0) = 0$.

Claim 3. $P_n = \{(P_n \cap D(V^{-1})) + V^{-1}(P_n \cap D(V^{-1}))\}$.

Proof of claim. Straightforward.

These three claims taken together tell us that P_n is invariant for V^{-1}. We now show that P_n is inner invariant.

If $V^{-1}f = f'$ is in P_n then $f'(t) = p(t)e^{-t}$ with p a polynomial of degree at most n. Therefore

$$\begin{aligned} f(t) &= \int_0^t p(y)e^{-y}\,dy \\ &= -(p(t) + p'(t) + p''(t) + \cdots + p^{(n)}(t))e^{-t} \\ &\quad + (p(0) + p'(0) + \cdots + p^{(n)}(0)). \end{aligned}$$

Since f is in $L^2(0,\infty)$ we conclude $\Sigma_{k=0}^n p^{(k)}(0) = 0$; hence $f(t) = -\Sigma_{k=0}^n p^{(k)}(t)e^{-t}$. Thus $f(t) = g(t)e^{-t}$ where $g(t)$ is a polynomial of degree at most n, and so f belongs to P_n and P_n is inner invariant for V^{-1} and also for V.

We will show that the subspaces P_n (for $n = 1, 2, \cdots$) are invariant for $V + 1$ and $V^{-1} + 1$ but not inner invariant for either. Keep in mind that $D(V^{-1} + 1) = D(V^{-1})$ so that $D(V^{-1} + 1) \cap P_n = \{p(t)e^{-t} \mid p(0) = 0\}$, where p is understood to be a polynomial of degree at most n.

Let f be in $D(V^{-1} + 1) \cap P_n$ so that $f(t) = p(t)e^{-t}$ with $p(0) = 0$. Then

$$\begin{aligned} (V^{-1} + 1)f(t) &= (V^{-1} + 1)p(t)e^{-t} \\ &= p'(t)e^{-t}. \end{aligned} \tag{15}$$

Thus, because of (15) we have $(V^{-1} + 1)(D(V^{-1}) \cap P_n) = P_{n-1}$. Further

$$\begin{aligned} &(V^{-1} + 1)(D(V^{-1}) \cap P_n) + (D(V^{-1}) \cap P_n) \\ &\quad = P_{n-1} + \{p(t)e^{-t} \mid p(0) = 0\} = P_n \end{aligned} \tag{16}$$

since $t^n e^{-t} \in \{p(t)e^{-t} \mid p(0) = 0\}$ and $P_{n-1} + [t^n e^{-t}] = P_n$. Thus (15) and (16) imply that P_n is invariant for $V^{-1} + 1$. On the other hand $t^{n+1}e^{-t}$ is in $D(V^{-1} + 1)\backslash P_n$ while $(V^{-1} + 1)(t^{n+1}e^{-t}) = (n+1)t^n e^{-t}$ belongs to P_n so that P_n is not inner invariant for $V^{-1} + 1$.

We will use L^2 for $L^2(0,\infty)$ below. Consider $D(V + 1) = D(V) = \{f \in L^2 \mid Vf \in L^2\}$. We look at $D(V) \cap P_n$. Let f be in P_n. Thus

$f(t)=p(t)e^{-t}$. If f is to belong to $D(V)$ then $\int_0^t p(y)e^{-y}\,dy$ must belong to L^2; but $\int_0^t p(y)e^{-y}\,dy = -e^{-t}(\Sigma_{k=0}^n p^{(k)}(t)) + \Sigma_{k=0}^n p^{(k)}(0)$. If this is to belong to L^2 then $\Sigma_{k=0}^n p^{(k)}(0)=0$ in which case $(V+1)(p(t)e^{-t}) = -e^{-t}(\Sigma_{k=0}^n p^{(k)}(t)) + p(t)e^{-t} = -e^{-t}(\Sigma_{k=1}^n p^{(k)}(t))$. Thus $(V+1)\cdot(P_n \cap D(V)) \subseteq P_n$ and $P_n \cap D(V) = \{p(t)e^{-t} \mid \Sigma_{k=0}^n p^{(k)}(0)=0\}$.

Claim 4. For $k=1,2,\cdots,n$, $(t^n - kt^{k-1})e^{-t} \in P_n \cap D(V)$.

Proof of Claim. A simple calculation gives us

$$\begin{aligned}\sum_{i=0}^{k} p^{(i)}(t) &= \sum_{i=0}^{k} \frac{d^i}{dt^i}(t^i - it^{i-1})\\ &= t^k.\end{aligned}$$

Therefore $\Sigma_{i=0}^k p^i(0) = 0^k = 0$.

Claim 5. $P_{n-1} \subseteq (V+1)(P_n \cap D(V))$.

Proof of Claim. If we do some more calculation we get

$$(V+1)(t^k - kt^{k-1})e^{-t} = -e^{-t}(kt^{k-1}) \tag{17}$$

for $k=1,2,\cdots,n$. Therefore $(V+1)(P_n \cap D(V))$ contains the space generated by e^{-t}, te^{-t}, $t^2e^{-t},\cdots$, $t^{n-1}e^{-t}$ which is P_{n-1}.

Now let $p(t)=t^n - n!$ then clearly $\Sigma_{k=0}^n p^k(0) = -n! + n! = 0$ so that $(t^n - n!)e^{-t}$ belongs to $P_n \cap D(V)$. From Claim 5 we get $n!e^{-t}$ in $(V+1)(P_n \cap D(V))$. Thus $t^n e^{-t} = (t^n - n!)e^{-t} + n!e^{-t}$ and this belongs to $(P_n \cap D(V)) + (V+1)(P_n \cap D(V))$ so that $P_n = (P_n \cap D(V)) + (V+1)(P_n \cap D(V))$. This means P_n is invariant for $V+1$. From Claim 4 we know that $(t^{n+1} - (n+1)t^n)e^{-t}$ belongs to $D(V)\backslash P_n$ but $(V+1)\cdot(t^{n+1} - (n+1)t^n)e^{-t} = -(n+1)t^n e^{-t}$ by (17) above, which is in P_n. Thus P_n is not inner invariant for $V+1$.

A few comments about the inner invariant subspaces $I(gH^2)$ are now in order. We have shown that these subspaces fill out the set of inner invariant subspaces for $V+1$. We can also show that these subspaces are inner invariant for $V-a$ where a is in the resolvent set of V. Since -1 belongs to the resolvent set and we do know all of the inner invariant subspaces for $V-(-1) = V+1$, is it possible that the spaces $I(gH^2)$ also fill out the inner invariant subspace structure of $V-a$? If not, what other inner invariant subspaces are there?

PROPOSITION 5.5. *The operators $V - a$ and $V^{-1} - a$, for $a \in \mathbf{C}$, have no nontrivial reducing subspaces.*

Proof. It is well known that the shift operator on H^2 has no nontrivial reducing subspaces (see Hoffman **[9]**, Page 110). Also, for a b-symmetric operator A, a subspace K reduces A iff it reduces $C(A)$ (see Akhiezer and Glazman **[1]**). Since $C(V) = -U$ where V is the Volterra operator and U is the shift, we can conclude that V has no nontrivial reducing subspaces. Since $V \sim V^{-1}$, we also know that V^{-1} has no nontrivial reducing subspaces. We know

$$D(V-a) = D(V) \quad \text{and} \quad \overline{D(V)} = L^2(0,\infty). \tag{18}$$

Let us assume that K is a reducing subspace for $V - a$ where a is a nonzero complex number. Therefore

$$D(V-a) = (D(V) \cap K) \oplus (D(V) \cap K^{\perp})$$

by Definition 2.8 and (18) above. We therefore conclude that $\{K \cap D(V)\}^{-} = K$ and $\{K^{\perp} \cap D(V)\}^{-} = K\perp$. In this case $(V-a)(K \cap D(V)) \subseteq K$ iff $V(K \cap D(V)) \subseteq K$, and $(V-a)(K^{\perp} \cap D(V)) \subseteq K^{\perp}$ iff $V(K^{\perp} \cap D(V)) \subseteq K^{\perp}$ so that K must also reduce V. Thus $K = \{0\}$ or $K = L^2(0,\infty)$. Since $V^{-1} \sim V$ we get the same result for $V^{-1} - a$.

6. Some applications.

As in Goldberg's book, "Unbounded Linear Operators", it is possible to define a natural induced linear operator on a quotient space. To do this consider a linear operator T with domain D contained in a Hilbert space H. If K is a closed subspace of H we can consider the quotient space H/K. The elements of H/K are equivalence classes of the form $x + K$ for x in H. We will denote this equivalence class by $[x]_k$, or simply $[x]$ when no ambiguity results.

DEFINITION 6.1. We define the *induced operator* $\hat{T}: H/K \to H/K$ by $\hat{T}_k[x]_k = [Tx]_k$ where $D(\hat{T}_k) = \{[x]_k \mid \exists x_0 \in [x]_k \text{ with } x_0 \in D(T)\}$. When no ambiguity results we will use $\hat{T}$ for $\hat{T}_k$.

PROPOSITION 6.2. *A closed invariant subspace K of T is inner invariant for T iff $\hat{T}_k$ is* 1–1.

Proof.
$(\Rightarrow)$ Assume $\hat{T}[x] = \hat{T}[y]$. From Definition 6.1 we know $Tx =$

$Ty + k$ for some k in K. Since $k = T(x - y)$ we conclude $k \in T(D) \cap K$, but the inner invariance of K then implies $k \in T(K \cap D)$ so that there is a k_1 in $K \cap D$ such that $Tk_1 = k$. This tells us $T(x - y - k_1) = 0$ so that $n = x - y - k_1$ is in the null space of T. But for an inner invariant subspace K, we always have the null space contained in $K \cap D$. Therefore $n \in K \cap D$ so that $x = y + (n + k_1)$ with $n + k_1 \in K$. This means $[x] = [y]$.

($\Leftarrow$) Clear.

Let G be a subspace of H that contains K. We will let $\hat{G}$ denote the subspace G/K of H/K. Let q_k be the natural homomorphism from H onto H/K. Then $\hat{G} = q_k(G)$. It is clear that $q_k: G \to \hat{G}$ is 1–1. Denote $D(\hat{T})$ by $\hat{D}$.

THEOREM 6.3. *Let K be a closed invariant subspace of T. Let G be closed with $G \supset K$. Then G is (inner) invariant for T* iff *$q_k(G) = \hat{G}$ is (inner) invariant for $\hat{T}_k$.*

Proof.

($\Rightarrow$) Suppose $[g] \in \hat{G} \cap \hat{D} = (G \cap D)/K$. We can assume $g \in G \cap D$. Since G is invariant for T, we get $Tg \in G$. Thus $\hat{T}[g] = [Tg]$ is in $\hat{G}$, that is, $\hat{T}(\hat{G} \cap \hat{D}) \subset \hat{G}$. We must now show that $\hat{G} = \{(\hat{G} \cap \hat{D}) + \hat{T}(\hat{G} \cap \hat{D})\}^-$ but through elementary calculations we get

$$
\begin{aligned}
\hat{G} = G/K &= \{(G \cap D) + T(G \cap D)\}^-/K \\
&= \{[(G \cap D) + T(G \cap D)]/K\}^- \\
&= \{(G \cap D)/K + T(G \cap D)/K\}^- \\
&= \{(G \cap D)/K + \hat{T}[(G \cap D)/K]\}^- \\
&= \{(\hat{G} \cap \hat{D}) + \hat{T}(\hat{G} \cap \hat{D})\}^-.
\end{aligned}
$$

($\Leftarrow$) Since q_k is 1–1 and onto from the subspaces of H containing K to the subspaces of H/K, we know that given $\hat{G}$ a subspace of H/K, there must exist a subspace G containing K such that $q_k(G) = \hat{G}$. We will assume $\hat{G}$ is invariant for $\hat{T}$ and show that G must then be invariant for T.

Let $x \in G \cap D$ so that $[x] \in \hat{G} \cap \hat{D}$. Thus $\hat{T}[x] \in \hat{G}$ since $\hat{G}$ is invariant for $\hat{T}$. Therefore there is a $y \in [Tx]$ such that $y \in G$. This means $Tx - y = k \in K$ or that $Tx = y + k \in G + K \subset G$. Thus $T(G \cap D) \subset G$.

Since $\hat{G}$ is invariant for $\hat{T}$ we know

$$\begin{aligned} G/K = \hat{G} &= \{(\hat{G} \cap \hat{D}) + \hat{T}(\hat{G} \cap \hat{D})\}^- \\ &= \{(G \cap D)/K + T(G \cap D)/K\}^- \\ &= \{[(G \cap D) + T(G \cap D)]/K\}^- \\ &= \{(G \cap D) + T(G \cap D)\}^-/K. \end{aligned}$$

Since q_k is 1–1 we can conclude that $G = \{(G \cap D) + T(G \cap D)\}^-$ as soon as we know that $\{(G \cap D) + T(G \cap D)\}^- \supset K$. This is true since K being invariant for T implies

$$K = \{(K \cap D) + T(K \cap D)\}^- \subset \{(G \cap D) + T(G \cap D)\}^-.$$

For inner invariance, assume G is not inner invariant for T. Then there is an x in $D \backslash G$ with Tx in G. This says $[x] \in \hat{D} \backslash \hat{G}$ with $\hat{T}[x] \in \hat{G}$ which contradicts the inner invariance of $\hat{G}$.

In the other direction we assume $\hat{G}$ not inner invariant for $\hat{T}$. Thus there is an $[x]$ in $\hat{D} \backslash \hat{G}$ with $\hat{T}[x] \in \hat{G}$. This means there is a y in $[x]$ with $y \in D \backslash G$. Now $Ty \in [Ty] = \hat{T}[y] = \hat{T}[x] \in \hat{G}$. Thus there is a $z \in [Ty]$ with z in G. Hence $Ty - z = k$ for some k in K so that $Ty = z + k \in G$. Since $y \in D \backslash G$ this contradicts the inner invariance of G.

REMARK. Since q_k is 1–1 and onto, this theorem established a 1–1 correspondence between the (inner) invariant subspaces of T containing K and the (inner) invariant subspaces of $\hat{T}_k$.

We state the following straightforward result.

COROLLARY 6.4. *Let K be inner invariant for T and let G be a subspace containing K. Then the following conditions are equivalent*:

(i) *G is inner invariant for T*

(ii) *$\hat{G}$ is inner invariant for $\hat{T}_k$*

(iii) *$\hat{G}$ is inner invariant for $\hat{T}_k^{-1}$*

(iv) *$\hat{G}$ is invariant for both $\hat{T}_k$ and $\hat{T}_k^{-1}$.*

REFERENCES

1. N. I. Akhiezer, and I. M. Glazman, *Theory of Linear Operators in Hilbert Space*, Vols. I and II, Frederick Unger Publishing Company, New York, N.Y., Vol. I (1966), Vol. II (1963).

2. A. Beurling, *On two problems concerning linear transformations in Hilbert space*, Acta Mathematica, **81** (1949).

3. W. F. Donoghue, Jr., *The lattice of invariant subspaces of a completely continuous quasi-nilpotent transformation*, Pacific J. Math., **7**, (1957).

4. N. Dunford, and J. T. Schwartz, *Linear Operators, Part I*, Interscience Publishers, Inc., New York, N.Y., 1967.

5. Peter A. Fillmore, *Notes on Operator Theory*, Van Nostrand Reinhold Company, New York, N.Y., 1970.

6. I. M. Glazman, and N. I. Akhiezer, see **[1]** above.

7. Henry Helson, *Lectures on Invariant Subspaces*, Academic Press, New York, N.Y., 1964.

8. E. Hewitt, and K. Stromberg, *Real and Abstract Analysis*, Springer-Verlag, New York, N.Y., 1965.

9. Kenneth Hoffman, *Banach Spaces of Analytic Functions*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962.

10. G. K. Kalisch, *On similarity, reducing manifolds and unitary equivalence of certain Volterra operators*, Ann. of Math., **66**, No. 3, 1957.

11. ———, *Direct proofs of spectral representation theorems*, J. Math. Anal. Appl., **8**, 1964, 351–363.

12. H. L. Royden, *Real Analysis*, The Macmillan Company, New York, N.Y., 1963.

13. Donald Sarason, *A remark on the Volterra operator*, J. Math. Anal. Appl., **12**, 1965, 244–246.

14. J. T. Schwartz, and N. Dunford, see **[4]** above.

15. M. H. Stone, *Linear Transformations in Hilbert Space*, Amer. Math. Soc., Providence, R.I., 1966.

16. K. Stromberg, and E. Hewitt, see **[8]** above.

Received May 30, 1975. Supported in Part by N.S.F. Grants # 13288 and 21334 and by A.F.O.S.R. Grant # 70–1870.

GLASSBORO STATE COLLEGE
GLASSBORO, NJ 08028

PACIFIC JOURNAL OF MATHEMATICS
Vol 68 No 2, 1977

A COMMUTATIVITY THEOREM FOR NON-ASSOCIATIVE ALGEBRAS OVER A PRINCIPAL IDEAL DOMAIN

JIANG LUH AND MOHAN S. PUTCHA

Let A be an algebra (not necessarily associative) over a principal ideal domain R such that for all $a, b \in A$, there exist $\alpha, \beta \in R$ such that $(\alpha, \beta) = 1$ and $\alpha ab = \beta ba$. It is shown that A is commutative.

Throughout this paper N will denote the set of natural numbers and Z^+ the set of positive integers. A will denote an algebra with identity 1 over a Principal Ideal Domain R. If $a, b \in A$ then $[a, b] = ab - ba$. If $\alpha, \beta \in R$, then (α, β) denotes the greatest common divisor of α and β. If $a \in A$, then the *order* of a, $o(a)$ is the generator of the ideal $I = \{\alpha \mid a \in R, \alpha a = 0\}$ of R. $o(a)$ is unique up to associates. As a generalization of concepts in **[1]**, **[2]**, **[3]**, **[4]**, **[5]** we consider the following:

(*) For all $a, b \in A$, there exist $\alpha, \beta \in R$ such that $(\alpha, \beta) = 1$ and $\alpha ab = \beta ba$.

We will show that if A satisfies (*), then A is commutative. This generalizes [**3**; Theorem 3.5].

LEMMA 1. *Let p be a prime in R, $m \in Z^+$ such that $p^m A = (0)$. If A satisfies (*), then A is commutative.*

Proof. Let C denote the center of A. Let $x \in A$, $o(x) = p^k$, $k \in N$. We prove by induction on k that $x \in C$. If $k = 0$, then $x = 0$. So let $k > 0$. Let $y \in A$. First we show

$$[x, y] \neq 0 \quad \text{implies} \quad [yx, y] = 0. \tag{1}$$

If $yx = 0$, this is trivial. So let $yx \neq 0$. Now for some $\alpha_1, \alpha_2 \in R$,

$$\begin{aligned} &\alpha_1 xy = \alpha_2 yx, (\alpha_1, \alpha_2) = 1 \\ &\beta_1(x+1)y = \beta_2 y(x+1), (\beta_1 \beta_2) = 1. \end{aligned} \tag{2}$$

So $\alpha_1\beta_1(x+1)y = \alpha_1\beta_2 y(x+1)$. Thus substituting the above, we get

$$(\alpha_2\beta_1 - \alpha_1\beta_2)yx = (\alpha_1\beta_2 - \alpha_1\beta_1)y. \tag{3}$$

We claim that $(\alpha_2\beta_1 - \alpha_1\beta_2)yx \neq 0$. For otherwise $(\alpha_1\beta_2 - \alpha_1\beta_1)y = 0$. Since $y \neq 0$, we get $p \mid \alpha_1\beta_2 - \alpha_1\beta_1$.

Also $(\alpha_1\beta_2 - \alpha_1\beta_1)yx = 0$. Since $(\alpha_2\beta_1 - \alpha_1\beta_2)yx = 0$, we get $(\alpha_2 - \alpha_1)\beta_1 yx = 0$. Since $yx \neq 0$, $p \mid \beta_1(\alpha_2 - \alpha_1)$. So

$$p \mid \alpha_1(\beta_2 - \beta_1), p \mid \beta_1(\alpha_2 - \alpha_1).$$

Case 1. $p \nmid \alpha_1$. Then since $\alpha_1(\beta_2 - \beta_1)y = 0$, we get $(\beta_2 - \beta_1)y = 0$. So by (2), $\beta_1[x, y] = 0 = \beta_2[x, y]$. Since $[x, y] \neq 0$, we get $p \mid \beta_1$, $p \mid \beta_2$, contradicting (2).

Case 2. $p \mid \alpha_1$. Then $p \nmid \alpha_2$ and so $p \nmid \alpha_2 - \alpha_1$. Thus $p \mid \beta_1$. So $p \nmid \beta_2$, $p \nmid \beta_2 - \beta_1$. Since $\alpha_1(\beta_2 - \beta_1)y = 0$ we get $\alpha_1 y = 0$. So $\alpha_1 xy = 0$. By (2), $\alpha_2 yx = 0$. Since $yx \neq 0$, we get $p \mid \alpha_2$, a contradiction.

Hence by (3)

$$(\alpha_2\beta_1 - \alpha_1\beta_2)yx \neq 0.$$

In particular

$$\alpha_2\beta_1 - \alpha_1\beta_2 \neq 0.$$

So

$$\alpha_2\beta_1 - \alpha_1\beta_2 = p^t\delta,\ t \in N,\ \delta \in R,\ (\delta, p) = 1.$$

If $t \geqq k$, then $(\alpha_2\beta_1 - \alpha_1\beta_2)yx = 0$, a contradiction. So $t < k$. Hence

$$p^{k-t}(\alpha_1\beta_2 - \alpha_1\beta_1)y = p^{k-t}p^t\delta yx = 0.$$

Let $o(y) = p^i$, $i \in N$. If $i < k$, then $y \in C$, a contradiction. So $i \geqq k$. Hence

$$p^k \mid p^i \mid p^{k-t}(\alpha_1\beta_2 - \alpha_1\beta_1).$$

So $p^t \mid \alpha_2\beta_2 - \alpha_1\beta_1$ and $\alpha_1\beta_2 - \alpha_1\beta_1 = p^t\gamma$, $\gamma \in R$. Then $p^t\delta yx = p^t\gamma y$. Hence $p^t(\delta yx - \gamma y) = 0$. By induction hypothesis, $\delta yx - \gamma y \in C$. So $[\delta yx - \gamma y, y] = 0$. Thus $\delta[yx, y] = 0$. Since $(\delta, p) = 1$, $[yx, y] = 0$. This establishes (1).

Now let $u \in A$ and suppose $[x, u] \neq 0$. Then also $[x, u + 1] \neq 0$. By (1), $[ux, u] = 0 = [(u + 1)x, u]$. So $[x, u] = 0$, a contradiction. So $x \in C$ and the lemma is proved.

LEMMA 2. *Suppose A satisfies* (*). *Let $a, b \in A, o(b) = 0$. If $ba = 0$, then $ab = 0$.*

Proof. Suppose $ab \neq 0$. Then there exist $\beta_1, \beta_2, \gamma_1, \gamma_2 \in R$ such that

$$\begin{aligned} &\beta_1(a+1)b = \beta_2 b(a+1),\ (\beta_1, \beta_2) = 1, \\ &\gamma_1 a(b+1) = \gamma_2 b(a+1),\ (\gamma_1, \gamma_2) = 1. \end{aligned} \tag{4}$$

So

$$\beta_1 ab = (\beta_2 - \beta_1)b \quad \text{and} \quad (\gamma_2 - \gamma_1)a = \gamma_1 ab. \tag{5}$$

If $\beta_2 = \beta_1$, then β_1, β_2 are units and by (5) $ab = ba = 0$, a contradiction. So $\beta_2 - \beta_1 \neq 0$. Similarly $\gamma_2 - \gamma_1 \neq 0$. Since $o(b) = 0$, we get by (5) that $o(ab) = 0$. So $o(a) = 0$. Hence by (5), $\beta_1 \neq 0$, $\gamma_1 \neq 0$. Also by (5) $[\beta_1 ab, b] = 0$.

So

$$\begin{aligned} (\gamma_2 - \gamma_1)\beta_1 ab &= \gamma_1\beta_1(ab)b \\ &= \gamma_1\beta_1 b(ab) \\ &= \beta_1(\gamma_2 - \gamma_1)ba \\ &= 0. \end{aligned}$$

So $o(ab) \neq 0$, a contradiction. This proves the lemma.

LEMMA 3. *Suppose A satisfies $(*)$. Let $b \in A$, $o(b) = 0$. Then $b \in C$, the center of A.*

Proof. Let $a \in A$. There exist $\alpha_1, \alpha_2, \beta_1, \beta_2 \in R$ such that

$$\begin{aligned} &\alpha_1 ab = \alpha_2 ba,\ (\alpha_1, \alpha_2) = 1, \\ &\beta_1(a+1)b = \beta_2 b(a+1),\ (\beta_1, \beta_2) = 1. \end{aligned} \tag{6}$$

Multiplying the second equation by α_1 and substituting the first we obtain

$$b[(\alpha_2\beta_1 - \alpha_1\beta_2)a - (\alpha_1\beta_2 - \alpha_1\beta_1)\cdot 1] = 0.$$

By Lemma 2,

$$[(\alpha_2\beta_1 - \alpha_1\beta_2)a - (\alpha_1\beta_2 - \alpha_1\beta_1)\cdot 1]b = 0.$$

Let $\mu = \alpha_2\beta_1 - \alpha_1\beta_2$. Then $\alpha_1(\beta_2 - \beta_1)b = \mu ab = \mu ba$. By (6) $\alpha_1\mu ab = \alpha_2\mu ba = \alpha_2\mu ab$. So

$$(\alpha_2 - \alpha_1)\alpha_1(\beta_2 - \beta_1)b = 0.$$

Since $o(b) = 0$, we obtain by (6) that either $\alpha_1 = \alpha_2$ is a unit, $\beta_1 = \beta_2$ is a unit or else $\alpha_1 = 0$. The first two cases imply by (6) that $ab = ba$. So let $\alpha_1 = 0$. Then $\alpha_2 ba = 0$ and α_2 is a unit by (6). So $ba = 0$. By Lemma 2, $ab = 0$. Thus in any case $ab = ba$ and we are done.

THEOREM 4. *Suppose A satisfies* (*). *Then A is commutative.*

Proof. Suppose A is not commutative. We will obtain a contradiction. There exists $x \in A$ such that $x \notin C$, the center of A. So $x + 1 \notin C$. By Lemma 3 $o(x) \neq 0$ and $o(x+1) \neq 0$. Hence $o(1) \neq 0$. Let $o(1) = d \neq 0$. Then d is not a unit and hence $d = p_1^{\alpha_1} \cdots p_t^{\alpha_t}$ for some primes $p_1, \cdots, p_t \in A$ and some positive integers $\alpha_1, \cdots, \alpha_t$. Let $A_i = \{a \mid a \in A, p_i^{\alpha_i} a = 0\}$. Then each A_i is a nonzero subalgebra of A and $A = A_1 \oplus \cdots \oplus A_t$. Being subalgebras of A, the A_i's also satisfy (*). Being homomorphic images of A, all the A_i's have identity elements. By Lemma 1 each A_i and hence A is commutative, a contradiction. This proves the theorem.

REFERENCES

1. R. Coughlin and M. Rich, *On scalar dependent algebras*, Canad. J. Math., **24** (1972), 696–702.
2. R. Coughlin, E. Kleinfeld, and M. Rich, *Scalar dependent algebras*, Proc. Amer. Math. Soc., **39** (1973), 69–73.
3. K. Koh, J. Luh and M. S. Putcha, *On the associativity and commutativity of algebras over commutative rings*, Pacific J. Math., **63** (1976), 423–430.
4. M. Rich, *A commutativity theorem for algebras*, Amer. Math. Monthly, **82** (1975), 377–379.
5. J. C. K. Wang and J. Luh, *The structure of a certain class of rings*, Math. Japon., **20** (1975), 149–157.

Received October 5, 1976. The second author was partially supported by NSF Grant MCS 76–05784.

NORTH CAROLINA STATE UNIVERSITY

RALEIGH, NC 27607

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No 2, 1977

ADDENDUM TO "RATIONAL APPROXIMATION OF e^{-x} ON THE POSITIVE REAL AXIS"

D. J. NEWMAN AND A. R. REDDY

Our aim in this addendum is to improve Theorem 3 of Newman and Reddy (*Pacific J. Math.*, **64** (1976), 227–232). We also take this opportunity to correct some misprints occurring in Theorem 6 of the above paper. For convenience we refer the above note to **[1]**. We follow here notation and numbering as in **[1]**.

THEOREM 3*. $\lambda^*_{0,4n}(e^{-x}) \leqq 4n^{-4}, n \geqq 1$.

Proof. It is easy to verify that $1 + x + x^2/2! + x^3/3! + x^4/4!$ has zeros only in the left hand plane. As far as we know this is the largest partial sum of e^x which has zeros only in the left half plane. Now using this in the proof of Theorem 3 of **[1]** instead of $1 + x + x^2/2!$, and by following the same approach we can get the required result.

We would like to point out now that the cases $n = 1, 2, 3$ of Theorem 5 follows from (12) and (14).

In the proof of Theorem 6 of **[1]**, the following changes are necessary.

$$\text{Change } \frac{v^2}{2} \text{ to } \frac{v^2}{2.25}, \frac{1}{\binom{2m}{m}\sqrt{m}} \text{ to } \frac{1.9}{\binom{2m}{m}\sqrt{m}}, \text{ and } \frac{n}{\sqrt{m}} \text{ to } \frac{(1.9)n}{\sqrt{m}}.$$

Then we get for all $n \geqq 8$, $\epsilon \geqq e^{-5n^{2/3}}$. By choosing $A = 3n^{2/3}$, $m = [n^{2/3}]$, we get for $1 \leqq n \leqq 7$, $\epsilon \geqq e^{-5n^{2/3}}$.

Received January 21, 1977

TEMPLE UNIVERSITY
AND
THE INSTITUTE FOR ADVANCED STUDY, PRINCETON
PRINCETON, NJ 08540

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No 2, 1977

ON THE DISTRIBUTION OF a-POINTS OF A STRONGLY ANNULAR FUNCTION

AKIO OSADA

This paper gives an example of a strongly annular function which omits 0 near an arc I on the unit circle C and which omits 1 near the complementary arc $C-I$. This example affirmatively answers the following question of Bonar: Does there exist any annular function for which we can find two or more complex numbers w such that the limiting set of its w-points does not cover C?

1. Introduction. The purpose of this paper is to study the distribution of a-points of annular functions. We recall that a holomorphic function in the open unit disk $D:|z|<1$ is said to be annular [**1**] if there is a sequence $\{J_n\}$ of closed Jordan curves about the origin in D, converging out to the unit circle $C:|z|=1$, such that the minimum modulus of $f(z)$ on J_n increases to infinity as n increases. When the J_n can be taken as circles concentric with C, $f(z)$ will be called strongly annular. Given a finite complex number a, the minimum modulus principle guarantees that every annular function f has infinitely many a-points in D and hence their limit points form a nonempty closed subset, say $Z'(f, a)$, of C. On the other hand, by virtue of the Koebe–Gross theorem concerning meromorphic functions omitting three points, it follows from the annularity of f that open sets $C-Z'(f, a)$ and $C-Z'(f, b)$ on the circle can not overlap if $a \neq b$ and consequently that the set of all values a for which $Z'(f, a) \neq C$ must be at most countable. Therefore we may well say such a to be singular for f.

For this reason we will be concerned with the set $S(f)=\{a: Z'(f, a) \neq C\}$ in this paper. We denote by $|S(f)|$ the cardinality of $S(f)$ and then, from the simple fact observed above, we have that $0 \leqq |S(f)| \leqq \aleph_0$, which in turn conversely tempt us to raise the following question: Given a cardinality $N (0 \leqq N \leqq \aleph_0)$, can we find any annular function f for which $|S(f)|=N$? ([**1**], [**2**]).

We know many examples of strongly annular functions such that $|S(f)|=0$ [**4**]. In particular if an annular function f belongs to the MacLane class, i.e., the family of all nonconstant holomorphic functions in D which have asymptotic values at each point of everywhere dense subsets of C, the set $S(f)$ becomes necessarily empty. As for $N=1$, Barth and Schneider [**3**] constructed an example of an annular function f for which $|S(f)|=1$. The example involved in their construction, however, did not appear to be strongly annular. An example of a strongly annular f with $|S(f)|=1$ was constructed independently by Barth, Bonar and Carroll **[2]** and the author **[5]**. The aim of this paper is to give an example of a strongly annular function f for which $|S(f)|=2$.

2. For this purpose we consider a class of functions holomorphic in D. Let I_0 and I_1 be a pair of complementary open arcs on the unit circle C and choose a Jordan arc J_j connecting the end points of I_j, which is contained, except for its end points, in the open sector

$$\{z: 0<|z|<1,\ z/|z|\in I_j\} \qquad (j=0,1).$$

Further denote by G_j the Jordan domain surrounded by I_j and J_j and consider

$$S(G_0, G_1)=\{g\in H(D): g \text{ is bounded away from 0 (or 1) in } G_0 \text{ (or } G_1)\}$$

where $H(D)$ denotes the set of all functions holomorphic in D. In terms of this notation our purpose is in amount to find a strongly annular function which is locally a uniform limit of a sequence in $S(G_0, G_1)$. To construct such a function, we make essential use of the approximation theorem of Runge, which asserts that if K is a compact set with connected complement relative to the plane and a function g is holomorphic in an open set containing K, for any $\rho>0$, there is a polynomial P such that

$$|P(z)-g(z)|<\rho \qquad (z\in K).$$

We call such P an approximating polynomial with respect to the triple (K, g, ρ). In our arguments to follow we may restrict ourselves to the special pair of G_0 and G_1 such that

$$G_0=\{z=x+iy: |z|<1,\ 2x+|y|>1\} \quad \text{and} \quad G_1=\{z: -z\in G_0\}$$

with no loss of generality, which serves to simplify the geometric formulation. Then the Runge theorem, in cooperation with our previous lemma, yields the following:

LEMMA. *Let there be given positive numbers ϵ and k, numbers a and b with $0<a<b<1$, and a function f in $S(G_0, G_1)$ (simply S), which is bounded in G_1. Then there exists a function g in S, which is also bounded in G_1, such that*

$$(1) \qquad |g(z)|>k \qquad (|z|=b)$$

and

(2) $$|g(z)-f(z)|<\epsilon \qquad (|z|\leqq a).$$

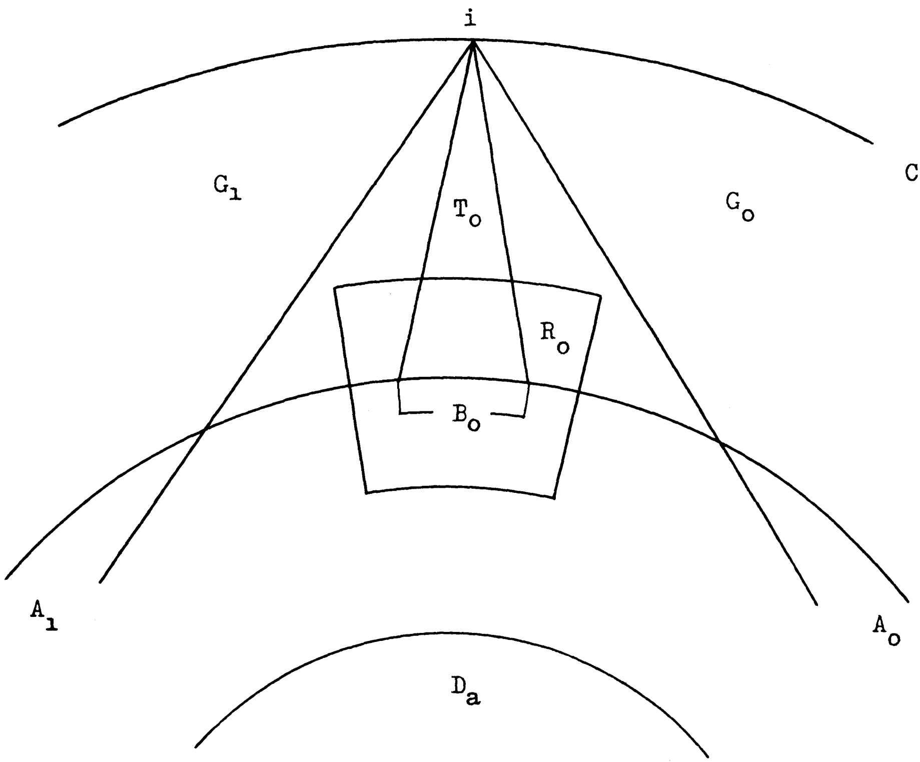
Proof. We first divide the circle $|z|=b$ into 4 closed arcs as follows:

$$A_0=[-bie^{it},\ bie^{-it}], \qquad A_1=\{z: -z\in A_0\}$$
$$B_0=[bie^{-it},\ bie^{it}], \qquad B_1=\{z: \bar z\in B_0\}.$$

Here $t(>0)$ should be chosen so small that we may apply our lemma **[5]** to an appropriately small open annular sector R_0, which is contained in

$$\{z=x+iy: y>0, |z|>a, 2|x|+|y|<1\}$$

and contains the arc B_0. Set $R_1=\{z:\bar z\in R_0\}$.



Next, to make use of the Runge theorem, we prepare two triples, which are defined, except for c_j and ρ_j, by the following:

$$(3)\quad \begin{cases} K_j = \bar{G}_j \cup A_j \cup A_{1-j} \cup \bar{D}_a,\ \bar{D}_a = \{z : |z| \leqq a\} \\ g_j(z) = 0 \qquad (z \in \bar{G}_j \cup A_j \cup \bar{D}_a) \\ g_j(z) = c_j (>0) \qquad (z \in A_{1-j}) \end{cases} \qquad (j = 0, 1).$$

As for c_j (or ρ_j) we shall later choose positive numbers large (or small) enough to satisfy our requirements. Obviously these definitions allow us to apply the Runge theorem to (K_j, g_j, ρ_j) $(j = 0, 1)$ and hence we can find an approximating polynomial P_j. On the other hand, if necessary, adding a small vector we may assume that $f(z) \neq 0, 1$ on the circle $|z| = b$. Combining these functions, define a function F holomorphic in D by

$$F(z) = \{(f(z) - 1) \exp(P_0(z)) + 1\} \exp(P_1(z)).$$

Then carefully observing (3) and suitably choosing values of c_j and ρ_j, we can conclude that the function F is a member of S, bounded in G_1 and has the following properties:

$$(4)\qquad |F(z)| > 2k \qquad (z \in \{z : |z| = b\} - B_0 - B_1)$$

$$(5)\qquad |F(z) - f(z)| < \epsilon/2 \qquad (z \in \bar{D}_a).$$

In addition it may be supposed that F does not vanish on $B_0 \cup B_1$.

Thus the last step in our construction of g is to make $|F(z)|$ large on the remaining arcs B_0 and B_1 without losing the properties described above of F. Given $c_2 > 0$ and $\rho_2 > 0$, applying our lemma **[5]** to the annular sectors R_0 and R_1 previously chosen, and successively using the standard "pole sweeping" method for the resulting rational functions, we can find a holomorphic function H_j in D such that

$$(6)\qquad |H_j(z)| > c_2 \qquad (z \in B_j),$$

$$(7)\qquad \operatorname{Re} H_j(z) > -\rho_2 \qquad (z \in R_j \cap \{z : |z| = b\} - B_j)$$

and

$$(8)\qquad |H_j(z)| < 2\rho_2 \qquad (z \in D - T_j)$$

where T_0 (or T_1) denotes an appropriate "pole sweeping route" ending at $z = i$ (or $-i$) which is contained in

$$E_0 = \{z = x + iy : y > 0,\ |z| > b,\ 2|x| + |y| < 1\}$$

(or $E_1 = \{z : \bar{z} \in E_0\}$) (see Figure 1). Using these functions and F defined above, set

$$F(z)\{1 + H_0(z)\}\{1 + H_1(z)\} = g(z).$$

Since F does not vanish on $B_0 \cup B_1$, if we appropriately choose a large (or small) positive number as a value of c_2 (or ρ_2), by virtue of (4) and (5) together with (6), (7) and (8), we can show that the function g belongs to the class S, is bounded in G_1 and further satisfies (1) and (2). This proves Lemma.

3. The following result is immediate from Lemma in 2.

THEOREM. *Let $\{r_n\}$ and $\{k_n\}$ be two sequences of positive numbers with $r_n \uparrow 1$ and $1 < k_n \uparrow +\infty$. Then there exists a function f, which is locally a uniform limit of a sequence in S and which furthermore satisfies that $|f(z)| \geqq k_n$ on the circle $|z| = r_n$.*

Proof. It is sufficient to construct a sequence $\{f_n(z)\}$ in S such that

(9) $$|f_n(z)| > k_j \quad \text{if} \quad 1 \leqq j \leqq n \qquad (z \in C_j = \{z : |z| = r_j\}),$$

(10) $$|f_n(z) - f_{n-1}(z)| < \epsilon_{n-1} \qquad (|z| \leqq r_{n-1},\ n \geqq 2)$$

and

(11) f_n is bounded in G_1

where $\{\epsilon_n\}$ is a preassigned sequence of positive numbers with $\Sigma\epsilon_n < +\infty$. In order to construct $\{f_n\}$ inductively, let $f_1(z) = 2k_1$ and suppose that $f_1, \cdots, f_{n-1}$ have already been defined. In Lemma in 2, on setting $f = f_{n-1}$, $a = r_{n-1}$, $b = r_n$, $k = k_n$ and $\epsilon = \min\{\epsilon_{n-1}, m_1, \cdots, m_{n-1}\}$ where $m_j = \min\{|f_{n-1}(z)| - k_j : z \in C_j\}$, we can find a function f_n in S satisfying (9), (10) and (11). Thus we obtain a sequence $\{f_n\}$ in S, which, by virtue of (10), converges uniformly on any compact subset of D. Obviously its limit f is a desired function in Theorem. Hence our proof is complete.

The author is grateful for the valuable comments and suggestions of the referee.

REFERENCES

1. F. Bagemihl and P. Erdös, *A problem concerning the zeros of a certain kind of holomorphic function in the unit circle*, J. Reine Angew. Math., **214/215** (1964), 340–344.

2. K. Barth, D. D. Bonar and F. W. Carroll, *Zeros of strongly annular functions*, Math. Z., **144** (1975), 175–179.
3. K. Barth and W. Schneider, *On a problem of Bagemihl and Erdös concerning the distribution of zeros of an annular function*, J. Reine Angew. Math., **234** (1969), 179–183.
4. D. D. Bonar, *On annular functions*, VEB Deutscher Verlag der Wissenschaften, Berlin, (1971).
5. A. Osada, *On the distribution of zeros of a strongly annular function*, Nagoya Math. J., **56** (1974), 13–17.

Received January 28, 1976 and in revised form June 6, 1976.

GIFU COLLEGE OF PHARMACY
MITAHORA, GIFU, JAPAN

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No. 2, 1977

A CHARACTERIZATION OF THE GAUSSIAN DISTRIBUTION IN A HILBERT SPACE

JEFFREY L. SPIELMAN

In this paper we consider the case in which random variables X_j take values in a real, separable Hilbert space $\mathscr{H}$. We look at a linear form $\Sigma A_j X_j$ where each A_j is a bounded linear operator in $\mathscr{H}$. We then assume that this linear form is identically distributed with a monomial and form conditions under which it is possible to deduce that the common distribution of the random variables is the Gaussian distribution.

The study of identically distributed linear forms of independent and identically distributed random variables has been undertaken by several authors. J. Marcinkiewicz studied linear forms in which all moments of the random variables are assumed to exist. He then proved that the common distribution of the random variables was the Normal distribution. R. G. Laha and E. Lukacs have considered the case where one of the linear forms is a monomial. They have obtained characterizations of the Normal distribution for both the case when the variance is assumed finite and when no assumption is made concerning the variance.

1. Statement of the main result. Suppose now that $X_1, X_2, \cdots$ is a sequence (possibly finite) of independent, identically distributed, nondegenerate $\mathscr{H}$-valued random variables, where X_1 has a finite variance (i.e. $\operatorname{Var} X_1 < +\infty$). Let $A_1, A_2, \cdots$ be a sequence of 1–1 bounded linear operators in $\mathscr{H}$, with the following two properties:

$$\sum_j \|A_j\|^2 < +\infty \quad \text{and} \quad \sum_j A_j^* A_j \geqq I \tag{1}$$

and

$$\sup_j \|A_j\| < 1. \tag{2}$$

We note that in the above A_j^* represents the adjoint of A_j and that the inequality $\Sigma_j A_j^* A_j \geqq I$ is true in the sense of positive-definiteness. (For example, see page 313 of [7].)

Our goal is to prove the following theorem.

THEOREM 1. *Suppose that* $\Sigma_j A_j X_j$ *converges with probability one. If* $\Sigma_j A_j X_j$ *has the same distribution as* X_1, *then* X_1 *has a Gaussian distribution.*

In §2 we will prove an important preliminary result (Theorem 2). Then in §3 we will present the proof of Theorem 1.

2. A preliminary result. In this section we will prove the following result.

THEOREM 2. *Let* $X_1, X_2, \cdots$ *be a sequence* (possibly finite) *of independent, identically distributed, nondegenerate,* $\mathcal{H}$*-valued random variables. Suppose that the sum* $\Sigma_j A_j X_j$ *exists with probability one, where* $A_1, A_2, \cdots$ *are bounded linear operators in* $\mathcal{H}$, *with* $\sup_j \|A_j\| < 1$.

If $\Sigma_j A_j X_j$ has the same distribution as X_1, then X_1 has an infinitely divisible distribution.

Note. The hypotheses of Theorem 2 are somewhat weaker than the hypotheses of Theorem 1.

Before beginning the proof of Theorem 2, let us fix some notation.

Let $\varphi(y)$ be the common characteristic functional of $X_1, X_2, \cdots$. Then $\varphi(y) = \mathscr{E} e^{i\langle X_1, y\rangle}$ for all $y \in \mathcal{H}$, where $\mathscr{E}$ denotes mathematical expectation.

The characteristic functional of $A_j X_j$ is then given by:

$$\mathscr{E} e^{i\langle A_j X_j, y\rangle} = \mathscr{E} e^{i\langle X_j, A_j^* y\rangle} = \varphi(A_j^* y) \tag{3}$$

where A_j^* denotes the adjoint operator of A_j.

Now, suppose that $\Sigma_j A_j X_j$ has the same distribution as X_1. Then equation (3) gives us:

$$\varphi(y) = \prod_j \varphi(A_j^* y), \qquad \text{for all} \qquad y \in \mathcal{H}, \tag{4}$$

where the product converges uniformly on bounded spheres. (See Theorem 4.4, pg. 171 of **[5]**.)

Since $\Sigma_j A_j X_j$ converges, then $\Sigma_{j=n}^\infty A_j X_j$ converges, with probability one, to the origin of $\mathcal{H}$ as $n \to \infty$. (Of course, if $X_1, X_2, \cdots$ is a finite sequence, the preceding statement is unnecessary.)

Thus, it is possible to choose N_0 for any $\epsilon > 0$, such that $P\{\|\Sigma_{j=N+1}^\infty A_j X_j\| > \epsilon\} < \epsilon$, whenever $N \geqq N_0$. Let $\varphi_N(y)$ denote the characteristic functional of $\Sigma_{j=N+1}^\infty A_j X_j$. Then using equation (4), we have:

$$\varphi(y) = \varphi(A_1^* y) \cdots \varphi(A_N^* y)\varphi_N(y). \tag{5}$$

Proof of Theorem 2. We assume that $\Sigma_j A_j X_j$ has the same distribution as X_1. Then equation (5) holds. If we replace y by $A_j^* y$ in equation (5), we obtain:

$$\varphi(A_j^* y) = \varphi(A_1^* A_j^* y) \cdots \varphi(A_N^* A_j^* y)\varphi_N(A_j^* y) \tag{6}$$

for each $j = 1, 2, \cdots, N$.

Combining equations (5) and (6) we have:

$$\varphi(y) = \prod_{j=1}^{N} \varphi((A_j^*)^2 y) \cdot \prod_{j \neq k} \varphi(A_j^* A_k^* y) \prod_{j=1}^{N} \varphi_N(A_j^* y)\varphi_N(y).$$

If we repeat the above process n times, we get the following result:

$$\varphi(y) = \Pi \varphi(A_{j_1}^* \cdots A_{j_n}^* y) \prod_{k=1}^{n-1} \Pi \varphi_N(A_{j_1}^* \cdots A_{j_{n-k}}^*)\varphi_N(y). \tag{7}$$

The product on the right hand side of equation (7) consists of $N^n + N^{n-1} + \cdots + N + 1$ factors, where each of the subscripts $j_1, \cdots, j_n$ can take any of the values $1, \cdots, N$ with repetitions allowed.

Thus, equation (7) says that X_1 is distributed as the sum of $k_n = \Sigma_{k=0}^n N^k$ independent, $\mathcal{H}$-valued random variables, $Y_{n,k}$ ($k = 1, 2, \cdots, k_n$), for any positive integer n.

We will now show that $Y_{n,k}$ is a *uniformly infinitessimal* collection of random variables. That is, we will show that for any $\epsilon > 0$, $\sup_{1 \leqq k \leqq k_n} P\{\| Y_{nk} \| > \epsilon\} \to 0$ as $n \to \infty$. Once this has been established, the infinite divisibility of X_1 will follow from Corollary 6.2, page 199 of **[5]**.

Consider the factors on the right hand side of equation (7). Let $\epsilon > 0$ be given. By definition $\varphi_N(y)$ is the characteristic functional of $\Sigma_{j=N+1}^\infty A_j X_j$ and $P\{\| \Sigma_{j=N+1}^\infty A_j X_j \| > \epsilon\} < \epsilon$, for all $N \geqq N_0$.

Consider now a factor of the form $\varphi_N(A_{j_1}^* A_{j_2}^* \cdots A_{j_{n-k}}^* y)$. This is the characteristic functional of

$$A_{j_{n-k}} \cdots A_{j_2} A_{j_1} \sum_{j=N+1}^{\infty} A_j X_j.$$

Also,

$$\begin{aligned} P\left\{\left\| A_{j_{n-k}} \cdots A_{j_2} A_{j_1} \sum_{j=N+1}^{\infty} A_j X_j \right\| > \epsilon\right\} & \\ \leqq P\left\{\| A_{j_{n-k}} \| \cdots \| A_{j_2} \| \cdot \| A_{j_1} \| \cdot \left\| \sum_{j=N+1}^{\infty} A_j X_j \right\| > \epsilon\right\} & \\ \leqq P\left\{\left\| \sum_{j=N+1}^{\infty} A_j X_j \right\| > \epsilon\right\} < \epsilon \quad \text{whenever} \quad N \geqq N_0, & \end{aligned}$$

since $\sup_j \| A_j \| < 1$.

Finally, we consider a factor of the form $\varphi(A^*_{j_1}\cdots A^*_{j_n}y)$, which is the characteristic functional of $A_{j_n}\cdots A_{j_1}X_1$. Set $\alpha = \sup_j \|A_j\|$.
Then

$$\begin{aligned} P\{\|A_{j_n}\cdots A_{j_1}X_1\| > \epsilon\} & \\ &\leqq P\{\|A_{j_n}\|\cdots\|A_{j_1}\|\cdot\|X_1\| > \epsilon\} \\ &\leqq P\left\{\|X_1\| > \frac{\epsilon}{\alpha^n}\right\}. \end{aligned}$$

Now choose an integer N' such that $P\{\|X_1\| > \epsilon/\alpha^n\} < \epsilon$, whenever $n \geqq N'$. (This is possible because $0 < \alpha < 1$). Set $n_0 = \max\{N_0, N'\}$.
Hence, we have shown that $P\{\|Y_{nk}\| > \epsilon\} < \epsilon$, for all $k = 1, 2, \cdots, k_n$, whenever $n \geqq n_0$. Therefore, the collection Y_{nk} is uniformly infinitesimal and X_1 is infinitely divisible. This completes the proof of the theorem.

3. Proof of the main result. For convenience, we now will make the assumption that $X_1, X_2, \cdots$ are symmetric random variables. Since the common distribution of these random variables is infinitely divisible, the common characteristic functional, $\varphi(y)$, has a unique Levy–Khintchine representation given by:

$$\ln \varphi(y) = -\tfrac{1}{2}\langle Sy, y\rangle + \int (\cos\langle x, y\rangle - 1)\,dL(x) \tag{8}$$

where S is an S-operator (a nonnegative, self-adjoint compact operator on $\mathcal{H}$, with a finite trace), and L is a σ-finite measure with finite mass outside every neighborhood of the origin and with the property that

$$\int_{\|x\|\leqq 1} \|x\|^2\,dL(x) < +\infty.$$

(see **[5]**, page 181.)
Furthermore, since $X_1, X_2, \cdots$ have finite variance, $\varphi(y)$ has a unique Kolmogorov representation, given by:

$$\ln \varphi(y) = -\tfrac{1}{2}\langle Sy, y\rangle + \int_{\mathcal{H}\setminus\{0\}} \frac{\cos\langle x, y\rangle - 1}{\|x\|^2}\,dK(x) \tag{9}$$

where S is an S-operator and K is a finite measure on $\mathcal{H}$. (See **[6]**.)
By equations (4) and (8) we have:

$$\begin{aligned} -\frac{1}{2}\sum_j \langle SA^*_j y, A^*_j y\rangle + \sum_j \int (\cos\langle x, A^*_j y\rangle - 1)\,dL(x) & \\ = -\tfrac{1}{2}\langle Sy, y\rangle + \int (\cos\langle x, y\rangle - 1)\,dL(x).& \end{aligned}$$

Also,

$$-\frac{1}{2}\sum_j \langle SA_j^*y, A_j^*y\rangle + \sum_j \int (\cos\langle x, A_j^*y\rangle - 1)dL(x)$$

$$(10) \qquad = -\frac{1}{2}\Big\langle \sum_j A_jSA_j^*y, y\Big\rangle + \sum_j \int (\cos\langle A_jx, y\rangle - 1)dL(x)$$

$$= -\frac{1}{2}\Big\langle \sum_j A_jSA_j^*y, y\Big\rangle + \sum_j \int (\cos\langle x, y\rangle - 1)dLA_j^{-1}(x).$$

It is not difficult to show that $\Sigma_j A_jSA_j^*$ is an S-operator. Also, it is clear that LA_j^{-1} is the σ-finite measure which occurs in the Levy–Khintchine representation of A_jX_j, for each j.

We denote by $\mathcal{B}$, the class of Borel sets in $\mathcal{H}$. Then the measure K_j, defined by:

$$K_j(D) = \int_D \|x\|^2 dLA_j^{-1}(x), \quad \text{for all} \quad D \in \mathcal{B}$$

is the finite measure which occurs in the Kolmogorov representation of A_jX_j, for each j.

Since $X_1, X_2, \cdots$ have finite variance,

$$(11) \qquad \int \|x\|^2 dL(x) < +\infty. \quad \text{(See [6].)}$$

By equation (10), $\ln \Pi_j \varphi(A_j^*y)$

$$(12) \qquad = -\frac{1}{2}\Big\langle \sum_j A_jSA_j^*y, y\Big\rangle + \sum_j \int_{\{x\neq 0\}} \frac{(\cos\langle x, y\rangle - 1)}{\|x\|^2} dK_j(x).$$

We note that

$$\sum_j \int_{\{x\neq 0\}} \left|\frac{\cos\langle x, y\rangle - 1}{\|x\|^2}\right| dK_j(x) \leqq \sum_j \int_{\{x\neq 0\}} \frac{\|x\|^2\|y\|^2}{\|x\|^2} dK_j(x)$$

$$= \|y\|^2 \sum_j \int_{\{x\neq 0\}} dK_j(x) = \|y\|^2 \sum_j \int \|x\|^2 dLA_j^{-1}(x)$$

$$= \|y\|^2 \sum_j \int \|A_jx\|^2 dL(x) \leqq \Big(\sum_j \|A_j\|^2\Big)\|y\|^2 \int \|x\|^2 dL(x) < \infty$$

because of relations (1) and (11).

Thus we may interchange the integral and summation signs in equation (12) to obtain:

$$\ln \prod_j \varphi(A_j^* y) = -\frac{1}{2}\left\langle \sum_j A_j S A_j^* y, y \right\rangle + \int_{\{x \neq 0\}} \frac{\cos\langle x, y\rangle - 1}{\|x\|^2} d\left(\sum_j K_j(x)\right).$$

Then, by the uniqueness of the Kolmogorov representation, we have:

$$\sum_j A_j S A_j^* = S \quad \text{and} \quad \sum_j K_j = K.$$

From the second of these relations, $\Sigma_j K_j(\mathscr{H}) = K(\mathscr{H})$, which leads to the following sequence of equations.

$$\sum_j \int \|x\|^2 dLA_j^{-1}(x) = \int \|x\|^2 dL(x)$$

$$\sum_j \int \|A_j x\|^2 dL(x) = \int \|x\|^2 dL(x)$$

$$\int \left[\sum_j \|A_j x\|^2 - \|x\|^2\right] dL(x) = 0$$

$$\int \left[\left\langle \sum_j A_j^* A_j x, x \right\rangle - \langle x, x\rangle\right] dL(x) = 0.$$

In view of relation (4), it must then be true that

$$L\left\{x : \sum_j \|A_j x\|^2 - \|x\|^2 > 0\right\} = 0. \tag{13}$$

We note that for n a positive integer, $\Sigma_{j=1}^n A_j X_j$ has characteristic functional $\Pi_{j=1}^n \varphi(A_j^* y)$, and

$$\ln \prod_{j=1}^n \varphi(A_j^* y) = -\frac{1}{2}\left\langle \sum_{j=1}^n A_j S A_j^* y, y \right\rangle + \int (\cos\langle x, y\rangle - 1) d\left(\sum_{j=1}^n LA_j^{-1}(x)\right). \tag{14}$$

Thus $\Sigma_{j=1}^n LA_j^{-1}$ converges weakly, outside closed neighborhoods of $0 \in \mathscr{H}$, to L, as $n \to \infty$. (See **[5]**, page 189).

It now becomes necessary to state and prove two technical lemmas.

LEMMA 1. *For any* $\epsilon > 0$,

$$\int_{\|x\|>\epsilon} \|x\|^2 dL(x) = \sum_j \int_{\|A_j x\|>\epsilon} \|A_j x\|^2 dL(x).$$

Proof. Let ϵ_1 and ϵ_2 be positive constants with $\epsilon_1 < \epsilon_2$. Define a function $f(x)$ by:

$$f(x) = \begin{cases} \|x\|^2, & \epsilon_1 < \|x\| \leqq \epsilon_2 \\ (\epsilon_2)^2, & \|x\| > \epsilon_2. \end{cases}$$

Then $f(x)$ is bounded and continuous. Thus by comment (14),

$$\int_{\|x\|>\epsilon_1} f(x)dL(x) = \lim_{n\to\infty} \sum_{j=1}^{n} \int_{\|x\|>\epsilon_1} f(x)dLA_j^{-1}(x)$$

which implies that

$$\sum_{j=1}^{n} \int_{\epsilon_1<\|x\|\leqq\epsilon_2} \|x\|^2\, dLA_j^{-1}(x) + (\epsilon_2)^2 \sum_{j=1}^{n} LA_j^{-1}\{\|x\| > \epsilon_2\}$$

converges to

$$\int_{\epsilon_1<\|x\|\leqq\epsilon_2} \|x\|^2 dL(x) + (\epsilon_2)^2 L\{\|x\| > \epsilon_2\}$$

as $n \to \infty$.

But, again because of comment (14), $\Sigma_{j=1}^{n} LA_j^{-1}\{\|x\| > \epsilon_2\}$ converges to $L\{\|x\| > \epsilon_2\}$ as $n \to \infty$.

Therefore, $\Sigma_{j=1}^{n} \int_{\epsilon_1<\|x\|\leqq\epsilon_2} \|x\|^2 dLA_j^{-1}(x)$ converges to

$$\int_{\epsilon_1<\|x\|\leqq\epsilon_2} \|x\|^2 dL(x),$$

whenever we choose $0 < \epsilon_1 < \epsilon_2$.

Let $\epsilon > 0$ be given. Let ϵ_n be a strictly increasing sequence of positive numbers, $\epsilon_n \uparrow +\infty$, with $\epsilon_1 > \epsilon$. For convenience we set $\epsilon = \epsilon_0$.

Then

$$\begin{aligned} \int_{\|x\|>\epsilon_0} \|x\|^2 dL(x) &= \sum_{k=0}^{\infty} \int_{\epsilon_k<\|x\|\leqq\epsilon_{k+1}} \|x\|^2 dL(x) \\ &= \sum_{k=0}^{\infty} \sum_{j} \int_{\epsilon_k<\|x\|<\epsilon_{k+1}} \|x\|^2 dLA_j^{-1}(x) \\ &= \sum_{j} \sum_{k=0}^{\infty} \int_{\epsilon_k<\|x\|\leqq\epsilon_{k+1}} \|x\|^2 dLA_j^{-1}(x) \\ &= \sum_{j} \int_{\|x\|>\epsilon} \|x\|^2 dLA_j^{-1}(x) = \sum_{j} \int_{\|A_jx\|>\epsilon} \|A_jx\|^2 dL(x). \end{aligned}$$

This completes the proof.

LEMMA 2. $L(\{x : \|A_kx\|^2 \geqq \|x\|^2,$ *for some* $k = 1, 2, \cdots\}) = L(\{0\})$.

Proof. Let k be a fixed positive integer.

Set $E_k = \{x : \|A_k x\|^2 = \|x\|^2\}$. Then, using equation (13), $L(E_k) = L(E_k \cap \{x : \Sigma_j \|A_j x\|^2 = \|x\|^2\})$.

Thus, $L(E_k) = L\{x : \Sigma_{j \neq k} \|A_j x\|^2 = 0\} = L(\{0\})$, since each operator A_j is 1–1.

$$\text{Similarly,} \quad L\left(\bigcup_k E_k\right) = L(\{0\}). \tag{15}$$

Using the same type of argument, it is easy to show that for all $k = 1, 2, \cdots$

$$L\{x : \|A_k x\|^2 > \|x\|^2\} = 0. \tag{16}$$

Combining equations (15) and (16) we are done.

From relations (1) and (13), we see that $L\{x : \|x\|^2 \neq \Sigma_j \|A_j x\|^2\} = 0$. Hence, referring to Lemma 1, it is true that, for all $\epsilon > 0$,

$$\sum_j \int_{\|x\|>\epsilon} \|A_j x\|^2 dL(x) = \sum_j \int_{\|A_j x\|>\epsilon} \|A_j x\|^2 dL(x),$$

and this implies that

$$\sum_j \left[\int_{\|x\|>\epsilon} \|A_j x\|^2 \, dL(x) - \int_{\|A_j x\|>\epsilon} \|A_j x\|^2 \, dL(x)\right] = 0, \text{ for all } \epsilon > 0. \tag{17}$$

But

$$L\{x : x \neq 0 \quad \text{and} \quad \|A_j x\| \geqq \|x\|\} = 0, \quad \text{for all} \quad j. \tag{18}$$

Thus, each term in the sum of equation (17) must be nonnegative, which yields:

$$\int_{\|x\|>\epsilon} \|A_j x\|^2 dL(x) = \int_{\|A_j x\|>\epsilon} \|A_j x\|^2 dL(x), \quad \text{for all} \quad \epsilon > 0 \quad \text{and all} \quad j.$$

Or, using equation (18),

$$\int_{\{\|x\|>\epsilon\}\cap F_j} \|A_j x\|^2 dL(x) = \int_{\{\|A_j x\|>\epsilon\}\cap F_j} \|A_j x\|^2 dL(x),$$

for all $\epsilon > 0$ and all j, where $F_j = \{x : \|A_j x\| < \|x\|\}$, for each $j = 1, 2, \cdots$.

The above implies that

$$\int_{\{\|x\|>\epsilon\}\cap\{\|A_jx\|\leqq\epsilon\}\cap F_j} \|A_jx\|^2 dL(x) = 0, \quad \text{for all } \epsilon > 0 \text{ and all } j,$$

or,

$$\int_{\{\|x\|>\epsilon \text{ and } \|A_jx\|\leqq\epsilon\}} \|A_jx\|^2 dL(x) = 0, \quad \text{for all } \epsilon > 0 \text{ and all } j.$$

So, we must have that

$$L\{x : \|x\| > \epsilon \quad \text{and} \quad \|A_jx\| \leqq \epsilon\} = 0, \quad \text{for all } \epsilon > 0 \text{ and all } j. \tag{19}$$

Consider the set Q^+ of positive rational numbers. Let k be a fixed positive integer.

$$L\left[\bigcup_{r\in Q^+} \{x : \|x\| > r \quad \text{and} \quad \|A_kx\| \leqq r\}\right] = L[\{x : \|A_kx\| < \|x\|\}] = L[\mathcal{H}\backslash\{0\}].$$

Therefore, $L[\mathcal{H}\backslash\{0\}] \leqq \Sigma_{r\in Q^+} L\{x : \|x\| > r \text{ and } \|A_kx\| \leqq r\} = 0$, by equation (19).

This last relation says that L is degenerate at $0 \in \mathcal{H}$, which means that the common characteristic functional of $X_1, X_2, \cdots$ is given by:

$$\ln \varphi(y) = -\tfrac{1}{2}\langle Sy, y\rangle \quad (\text{see Eq. (8)}).$$

Hence $X_1, X_2, \cdots$ have a common Gaussian distribution.

Recall that we have assumed $X_1, X_2, \cdots$ to be symmetric, but it is now easy to extend our result to the general case by using Cramer's Theorem (see page 141 of **[1]**).

The proof of Theorem 1 is now completed.

References

1. U. Grenander, *Probabilities on Algebraic Structures*, Wiley, New York (1963).

2. R. G. Laha, *On a property of infinitely divisible distributions in a Hilbert space*, Acta Math. Acad. Sci. Hungar. **20** (1969), 143–147.

3. R. G. Laha and E. Lukacs, *On a linear form whose distribution is identical with that of a monomial*, Pacific J. Math., **15** (1965), 207–214.

4. J. Marcinkiewicz, *Sur une propriété de la loi de Gauss*, Math. Zeitschr., **44** (1949), 612–618.

5. K. R. Parthasarathy, *Probability Measures on Metric Spaces*, Academic Press, New York (1967).

6. B. L. S. Prakasa Rao, *Some characterization theorems for Wiener processes in a Hilbert space*, Z. Wahr. verw. Geb., **19** (1971), 103–116.

7. W. Rudin, *Functional Analysis*, McGraw–Hill, New York (1973).

Received April 2, 1976.

CHICAGO STATE UNIVERSITY
CHICAGO, IL 60628

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No 2, 1977

SYMMETRIZABLE-CLOSED SPACES

R. M. STEPHENSON, JR.

Symmetrizable-closed, semimetrizable-closed, minimal symmetrizable, and minimal semimetrizable spaces are characterized. G. M. Reed's theorem that every Moore-closed space is separable is extended to: Every Baire, semimetrizable-closed space is separable. Several examples are given.

If P is a topological property, a Hausdorff P-space will be called *P-closed* provided that it is a closed subset of every Hausdorff P-space in which it can be embedded. A Hausdorff P-space $(X, \mathscr{T})$ will be called *minimal P* if there exists no Hausdorff P-topology on X strictly weaker than $\mathscr{T}$.

In **[3]** J. W. Green characterized and studied Moore-closed and minimal Moore spaces. In this paper we obtain some analogous results for semimetrizable spaces and symmetrizable spaces.

A *symmetric* for a topological space X is a mapping $d: X \times X \to [0, \infty)$ such that

(1) For all $x, y \in X$, $d(x, y) = d(y, x)$, and $d(x, y) = 0$ if and only if $x = y$.

(2) A set $V \subset X$ is open if and only if for each $x \in V$ there exists $n \in \boldsymbol{N}$ such that V contains the set $B(n, x) = \{y \in X \mid d(x, y) < 1/n\}$.

A space X which admits a symmetric is said to be *symmetrizable*, and if, in addition, each $B(n, x)$ is a neighborhood of x, then X is said to be *semimetrizable* and d is called a *semimetric* for X. Equivalently, X is semimetrizable via d provided that for $x \in X$, $A \subset X$, and $d(x, A) = \inf\{d(x, a) \mid a \in A\}$, the condition $x \in \bar{A}$ if and only if $d(x, A) = 0$ is satisfied.

A number of the techniques used here are not new; for example, see **[2]**. The terminology used is standard. One perhaps not too familiar concept is that of θ-adherence. A point p of a topological space is said to be a *θ-adherent point* (or be in the *θ-adherence*) of a filter base $\mathscr{F}$ provided that for every set $F \in \mathscr{F}$ and neighborhood V of p, one has $F \cap \bar{V} \neq \emptyset$.

Our first two theorems are characterization theorems.

THEOREM 1. *Let $(X, \mathscr{T})$ be a symmetrizable Hausdorff space. The following are equivalent.*

(i) *The space $(X, \mathscr{T})$ is minimal symmetrizable.*

(ii) *Every countable filter base on $(X, \mathscr{T})$ which has a unique θ-adherent point is convergent.*

Proof. (ii) implies (i). Suppose that $(X, \mathcal{S})$ is symmetrizable and Hausdorff and $\mathcal{S} \subset \mathcal{T}$. Let d be a symmetric for $(X, \mathcal{S})$. For each point $p \in X$ the filter base

$$\mathcal{B}_p = \{\{x: d(x,p) < 1/n\}: n \in \boldsymbol{N}\}$$

has a unique θ-adherent point in $(X, \mathcal{S})$, namely p, and so $\mathcal{B}_p$ also has at most one θ-adherent point in $(X, \mathcal{T})$. By (ii) and the relation $p \in \bigcap \mathcal{B}_p$, it follows that each $\mathcal{B}_p$ must converge to p in $(X, \mathcal{T})$. Thus for every $T \in \mathcal{T}$ and $p \in T$ there exists $n \in \boldsymbol{N}$ such that $T \supset \{x: d(x,p) < 1/n\}$, that is, $T \in \mathcal{S}$. Therefore, $\mathcal{T} \subset \mathcal{S}$ and $(X, \mathcal{T})$ is minimal symmetrizable.

(i) implies (ii). Assume that there exist a point $q \in X$ and filter base $\mathcal{F} = \{F_n: n \in \boldsymbol{N}\}$ on X such that:

(a) for each $n \in \boldsymbol{N}$, $F_n \supset F_{n+1}$;
(b) q is the unique θ-adherent point of $\mathcal{F}$ in $(X, \mathcal{T})$;
(c) $\mathcal{F}$ fails to be convergent; and
(d) $F_1 = X$.

We will prove that $(X, \mathcal{T})$ cannot be minimal symmetrizable.

Let $\mathcal{V} = \{V \in \mathcal{T}: \text{if } q \in V \text{ then } V \text{ contains some member of } \mathcal{F}\}$. Then $\mathcal{V}$ is a topology on X with $\mathcal{V} \subset \mathcal{T}$, and because $\mathcal{F}$ has no θ-adherent point other than q, the space $(X, \mathcal{V})$ is Hausdorff. By (c), $\mathcal{V} \neq \mathcal{T}$.

Now consider any symmetric d for $(X, \mathcal{T})$. Define $d^*: X \times X \to [0, \infty)$ by the rule

$$d^*(y,x) = d^*(x,y) = \begin{cases} d(x,y) & \text{if } x \neq q \neq y \\ 0 & \text{if } x = q = y \\ \min\{d(x,y), 1/n\} & \text{if } y = q \text{ and } x \in F_n \backslash F_{n+1}. \end{cases}$$

Clearly, d^* is a symmetric for the space $(X, \mathcal{V})$, and so $(X, \mathcal{T})$ cannot be minimal symmetrizable.

THEOREM 2. *Let X be a symmetrizable Hausdorff space. The following are equivalent.*

(i) *X is symmetrizable-closed.*
(ii) *Every countable filter base on X has a θ-adherent point.*

Proof. (ii) implies (i). Suppose that there exists a symmetrizable Hausdorff space Y such that X is a subspace of Y but $X \neq \bar{X}$. Because $\bar{X}$ is a closed subset of Y, $\bar{X}$ is symmetrizable (e.g., see [**5**, p. 93]). Let d be a symmetric for $\bar{X}$. Since X fails to be a closed subset of $\bar{X}$, there must exist a point $p \in \bar{X} \backslash X$ with $0 = \inf\{d(p,x): x \in X\}$. Thus for each $n \in \boldsymbol{N}$,

$$F_n = \{x \in X: d(p, x) < 1/n\}$$

is nonempty, and so $\mathcal{F} = \{F_n: n \in N\}$ is a countable filter base on X. Obviously $\mathcal{F}$ has no θ-adherent point in X.

(i) implies (ii). Assume that there exists a filter base $\mathcal{G} = \{G_n: n \in N\}$ on X such that $G_1 = X$, each $G_n \supset G_{n+1}$, and $\mathcal{G}$ has no θ-adherent point in X. Choose a new point $q \notin X$, let $E = X \cup \{q\}$, and call a subset V of E open if and only if (a) $V \cap X$ is open in X and (b) if $q \in V$ then for some $n \in N$, $V \supset G_n$. Then E is a Hausdorff space in which X is embedded as a proper dense subspace. E is also symmetrizable, for if d is any symmetric for X, then the function $d^*: E \times E \to [0, \infty)$ determined by the rule

$$d^*(x, y) = d^*(y, x) = \begin{cases} d(x, y) & \text{if } x, y \in X \\ 0 & \text{if } x = q = y \\ 1/n & \text{if } x \in G_n \backslash G_{n+1} \text{ and } y = q, \end{cases}$$

is easily seen to be a symmetric for E.

For many properties P, P-minimality is a sufficient condition for P-closedness. For P = symmetrizable, the same is true.

COROLLARY 3. *Every minimal symmetrizable Hausdorff space* $(X, \mathcal{T})$ *is symmetrizable-closed.*

Proof. If d is a symmetric for $(X, \mathcal{T})$ and $\mathcal{F}$ is a descending sequence of nonempty sets having no θ-adherent point in $(X, \mathcal{T})$, with $X \in \mathcal{F}$, then for any point $q \in X$, the function d^* defined in the proof of Theorem 1 is a symmetric for a strictly weaker symmetrizable Hausdorff space $(X, \mathcal{V})$.

COROLLARY 4. *Every regular, symmetrizable-closed space is compact.*

Proof. In a regular space θ-adherence and adherence are equivalent concepts, so by Theorem 2, every regular symmetrizable-closed space is countably compact. By a result of Nedev [7] every countably compact symmetrizable Hausdorff space is compact.

For various properties P, topologists have often been interested in the question as to whether or not there exists a non-compact P-space in which every closed subset is P-closed. If P = Hausdorff or completely Hausdorff, the answer is known to be no, but if P = regular, the question is open. For P = symmetrizable, the following result holds.

COROLLARY 5. *Let X be a symmetrizable Hausdorff space in which every closed subset is symmetrizable-closed. Then X is compact.*

Proof. Obviously no infinite discrete space can be symmetrizable-closed, so every infinite closed subset of X must have a limit point, that is, X must be countably compact.

Let us now give some examples of these concepts.

EXAMPLE 6. In **[1]** N. Bourbaki pointed out that a certain space X due to Urysohn is a minimal Hausdorff space that fails to be compact. We will describe this space and show that it is also semimetrizable, in order to show that there exist noncompact, Hausdorff minimal symmetrizable spaces.

Let

$$X = \mathbf{N} \cup \{n \pm 1/m : n, m \in \mathbf{N}, m > 2\} \cup \{\pm\pi\}.$$

Define $d: X \times X \to [0, \infty)$ by the rule

$$d(x,y) = d(y,x) = \begin{cases} 0 & \text{if } x = y; \\ |x-y| & \text{if } x, y \notin \{\pm\pi\}; \\ 1 & \begin{array}{l}\text{if } x \in \mathbf{N} \text{ and } y \in \{\pm\pi\}, \text{ or} \\ \text{if } x = n + 1/m \text{ and } y = -\pi, \text{ or} \\ \text{if } x = n - 1/m \text{ and } y = \pi, \text{ or} \\ \text{if } x = \pi \text{ and } y = -\pi, \text{ where} \\ m, n \in \mathbf{N} \text{ and } m > 2;\end{array} \\ 1/n & \begin{array}{l}\text{if } x = n + 1/m \text{ and } y = \pi, \text{ or} \\ \text{if } x = n - 1/m \text{ and } y = -\pi, \\ \text{where } m, n \in \mathbf{N} \text{ and } m > 2.\end{array} \end{cases}$$

Call a subset V of X open if and only if for each point $v \in V$ there exists $e > 0$ with $\{x: d(x, v) < e\} \subset V$. Then d is a semimetric for the space X, and X is homeomorphic with the space in **[1]** (X is also described in **[2**, p. 101**]**).

EXAMPLE 7. If X is as in Example 6, then its subspace

$$Y = \mathbf{N} \cup \{n + 1/m : n, m \in \mathbf{N}, m > 2\} \cup \{\pi\}$$

is well known to be Hausdorff-closed but not minimal Hausdorff. Since Y is a subspace of X, it is also semimetrizable. If Y' denotes the space which has the same points as those of Y but which is topologized so that it is the one-point compactification of the space $Y \backslash \{\pi\}$, then Y' is metrizable, and so one sees that Y is not minimal semimetrizable. Thus

Y is an example of a Hausdorff semimetrizable, symmetrizable-closed space that is not minimal semimetrizable.

For P = semimetrizable, the results one can obtain concerning the concepts P-closed and P-minimal are much more similar to those in **[3]**. Since the proofs are not too different from some of the ones above and in **[3]** and **[9]**, the details are omitted. First two definitions are needed.

A topological space is called *feebly compact* if every countable open filter base has an adherent point. A space is called *semiregular* if it has a base consisting of regular open sets, i.e., sets having the form $V = (\bar{V})°$.

THEOREM 8. *Let X be a semimetrizable Hausdorff space. The following are equivalent.*

(i) *X is semimetrizable-closed.*

(ii) *X is feebly compact.*

THEOREM 9. *Let X be a semimetrizable Hausdorff space. The following are equivalent.*

(i) *X is minimal semimetrizable.*

(ii) *Every countable open filter base on X with a unique adherent point is convergent.*

(iii) *X is semiregular and semimetrizable-closed.*

For semimetrizable spaces, it is easy to show that the concepts semimetrizable-closed and symmetrizable-closed are distinct. For example, let X be any noncompact, regular, semimetrizable-closed space (such as one of the spaces discussed in **[3]**). By Corollary 4, X cannot be symmetrizable-closed.

Not too much is known concerning the density character and cardinality of semimetrizable-closed and symmetrizable-closed spaces. G. M. Reed **[8]** has proved that every Moore-closed space is separable, but I do not know if an analogous result holds for all semimetrizable or symmetrizable spaces. (A proof is given in **[10]** that a feebly compact symmetrizable space is separable if it has a dense set of isolated points.) In our final theorem it is shown that Reed's condition Moore-closed space, or, equivalently, feebly compact Moore space (see **[3]**), can be weakened.

We recall that a topological space X is said to be a *Baire* space provided that for every countable family $\mathscr{C}$ of dense open subsets of X, the set $\cap \mathscr{C}$ is also dense. It is known **[6]** that every regular, feebly compact space is a Baire space.

THEOREM 10. *Every Baire, feebly compact, semimetrizable space X is separable.*

Proof. The proof will consist of two parts. We will first prove that $(*)$ every family of pairwise disjoint nonempty open subsets of X is countable. Next, using $(*)$, we will construct a countable dense subset for X.

Let d be a semimetric for X. For $x \in X$ and $n \in \boldsymbol{N}$, $\{y \in X: d(x, y) < 1/n\}$ will be denoted by $B(n, x)$, and the interior of $B(n, x)$ will be denoted by $I(n, x)$.

Proof of $(*)$: Suppose that there exists an uncountable family $\mathscr{V}$ of pairwise disjoint nonempty open subsets of X. For each $V \in \mathscr{V}$ and $m \in \boldsymbol{N}$ let

$$V_m = \{x \in V: B(m, x) \subset V\}^-,$$

and note that since $V = \cup\{V_m: m \in \boldsymbol{N}\}$, it follows from the Baireness of X that one can select an integer $m(V)$ for which $V_{m(V)}$ has nonempty interior. Choose $i \in \boldsymbol{N}$ such that $\mathscr{W} = \{V \in \mathscr{V}: m(V) = i\}$ is uncountable, and for each $W \in \mathscr{W}$ let J_W denote the interior of W_i. By the feeble compactness of X, there must exist a point $p \in X$ at which $\mathscr{J} = \{J_W: W \in \mathscr{W}\}$ fails to be locally finite. But consider any set $J_W \in \mathscr{J}$ with $\phi \neq K = J_W \cap I(i, p)$. Because K is a nonempty open subset of W_i, there must exist a point $q \in W$ with $B(i, q) \subset W$ and with $q \in K$. Then $d(p, q) < 1/i$ and so $p \in B(i, q) \subset W$. This latter relation, however, shows that $\mathscr{J}$ must be locally finite at p, for given any $J_V \in \mathscr{J}$ with $V \neq W$, we have $W \cap J_V = \emptyset$. Thus we have obtained a contradiction, and the proof of $(*)$ is complete.

For the remainder of the proof, if $n \in \boldsymbol{N}$ let

$$\mathscr{B}_n = \{I(k, x): x \in X, k \in \boldsymbol{N}, \text{ and } k \geqq n\},$$

and let $\mathscr{D}_n$ be a maximal family of pairwise disjoint members of $\mathscr{B}_n$. Once the sequence $\{\mathscr{D}_n: n \in \boldsymbol{N}\}$ has been determined, choose, for each $n \in \boldsymbol{N}$ and $D \in \mathscr{D}_n$, one point np_D such that $D = I(k, np_D)$ for some $k \in \boldsymbol{N}$ with $k \geqq n$, and let

$$C_n = \{np_D: D \in \mathscr{D}_n\}.$$

Then $C = \cup\{C_n: n \in \boldsymbol{N}\}$ is a countable subset of X, because by $(*)$, each $\mathscr{D}_n$ is countable. We will conclude the proof by proving that C is also dense in X.

Because each $\cup \mathscr{D}_n$ is an open dense subset of X, the set $E = \cap\{\cup \mathscr{D}_n: n \in \boldsymbol{N}\}$ is also a dense subset of X.

Now consider an arbitrary point $e \in E$. For each $n \in \boldsymbol{N}$ there exists a set $I(k, np_D) \in \mathscr{D}_n$ which contains e. Thus each $d(e, np_D) < 1/n$, which shows that $e \in \bar{C}$.

Therefore, $E \subset \bar{C}$ and so $X = \bar{E} = \bar{C}$.

While not every Baire semimetrizable-closed space is regular (e.g., Example 6), R. W. Heath has informed the author that he can prove every regular, semimetrizable-closed space is a Moore space — to verify Heath's result, appeal to the characterizations A and B' in **[4]** and the well known fact that in a regular feebly compact space any countable open filter base with a unique adherent point is convergent.

Since every separable first countable Hausdorff space has cardinality $\leqq c$, it follows from Theorem 10 that every Baire semimetrizable-closed space has cardinality $\leqq c$. We will conclude by showing that if the conditions "Hausdorff, semimetrizable, and Baire" are deleted, then the bound c may be exceeded.

EXAMPLE 11. Let m be an arbitrary infinite cardinal number, let $\mathcal{M}_m$ be a maximal family of countably infinite subsets of m such that the intersection of any two members is finite. Denote by $\{p_M : M \in \mathcal{M}_m\}$ a set of distinct point not in m, and let $X_m = m \cup \{p_M : M \in \mathcal{M}_m\}$. For each $M \in \mathcal{M}_m$ let $g_M : M \to N$ be one-to-one. Define $d: X_m \times X_m \to [0, \infty)$ by the rule

$$d(x,y) = d(y,x) = \begin{cases} 1 & \text{if } x, y \in m \text{ and } x \neq y; \\ 1 & \text{if } x = p_M \text{ and } y \notin \{p_M\} \cup M; \\ 1/g_M(y) & \text{if } x = p_M \text{ and } y \in M; \text{ and} \\ 0 & \text{if } x = y. \end{cases}$$

Topologize X_m by declaring a set V to be open if and only if for each point $v \in V$ there exists $e > 0$ with $\{x \in X_m : d(x, v) < e\} \subset V$. *Then the space X_m is a feebly compact symmetrizable space of cardinality $\geqq m$.*

REFERENCES

1. N. Bourbaki, *Espaces minimaux et espaces complètement séparés*, C. R. Acad. Sci. Paris, **212** (1941), 215–218.

2. M. P. Berri, J. R. Porter, and R. M. Stephenson, Jr., *A survey of minimal topological spaces*, in: J. Novak, ed., *General Topology and its Relation to Modern Analysis and Algebra*, Proc. Kanpur Topology Conf., 1968 (Academic Press, New York, 1971), 93–114.

3. J. W. Green, *Moore-closed spaces completeness and centered bases*, General Topology and Appl., **4** (1974), 297–313.

4. R. W. Heath, *Arc-wise connectedness in semi-metric spaces*, Pacific J. Math., **12** (1962), 1301–1319.

5. P. W. Harley, III and R. M. Stephenson, Jr., *Symmetrizable and related spaces*, Trans. Amer. Math. Soc., **219** (1976), 89–111.

6. R. A. McCoy, *A filter characterization of regular Baire spaces*, Proc. Amer. Math. Soc., **40** (1973), 268–270.

7. S. Nedev, *Symmetrizable spaces and final compactness*, Soviet Math. Dokl., **8** (1967), 890–892.

8. G. M. Reed, *On chain conditions in Moore spaces*, General Topology and Appl., **4** (1974), 255–267.

9. R. M. Stephenson, Jr., *Minimal first countable topologies*, Trans. Amer. Math. Soc., **138** (1969), 115–127.

10. ———, *Symmetrizable, $\mathscr{F}$, and weakly first countable spaces*, Canad. J. Math., **29** (1977), 480–488.

Received December 6, 1976.

UNIVERSITY OF SOUTH CAROLINA
COLUMBIA, SC 29208

PACIFIC JOURNAL OF MATHEMATICS
Vol. 68, No. 2, 1977

COMPLETELY SEMISIMPLE INVERSE Δ-SEMIGROUPS ADMITTING PRINCIPAL SERIES

P. G. TROTTER AND TAKAYUKI TAMURA

A Δ-semigroup is a semigroup whose lattice of congruences is a chain with respect to inclusion. A completely semisimple inverse Δ-semigroup that admits a principal series is characterized here as a semigroup that results from a particular series of ideal extensions of Brandt semigroups by Brandt semigroups. A characterization is given of finite inverse Δ-semigroups in terms of groups, Brandt semigroups, and one to one partial transformations of sets.

1. Introduction. A *Δ-semigroup* is a semigroup whose lattice of congruences is a chain with respect to inclusion. Schein **[8]** and Tamura **[11]** showed that a commutative Δ-semigroup is either a quasi-cyclic group A, or a commutative nil semigroup B with the divisibility chain condition, or A^0, or B^1. We study here the structure of completely semisimple inverse Δ-semigroups with principal series. Such semigroups will be characterized in terms of Δ-groups, idempotent properties, and ideal extensions of Brandt semigroups by Brandt semigroups.

In **[11]** it was shown that the least semilattice congruence on a Δ-semigroup has at most two classes. We begin by characterizing completely semisimple inverse semigroups admitting principal series and having this property.

In the final section we show that each finite inverse Δ-semigroup determines a set of structure data that involves groups, Brandt semigroups and one to one partial transformations of sets. Conversely the semigroup can be reconstructed from the structure data.

2. Preliminaries. We call a semigroup S an $\mathscr{S}_1$-, or $\mathscr{S}_2$-*semigroup* if the smallest semilattice congruence on S has one, or two congruence classes respectively. S is a Δ-semigroup only if it is an $\mathscr{S}_1$- or $\mathscr{S}_2$-semigroup. In this section we characterize completely semisimple inverse $\mathscr{S}_1$-, or $\mathscr{S}_2$-semigroups that admit principal series.

A subsemigroup H of a semigroup S is $\mathscr{S}$-*unitary* if and only if whenever $HxyH \subseteq H$ for $x, y \in S^1$ then $Hx, yH \subseteq H$. Notice that if E is a semilattice and $efg = e$ in E then $ef = e = ge$. Hence, any class of a semilattice congruence on S is $\mathscr{S}$-unitary. Let $\mathscr{J}^*$ denote the least congruence on S containing Green's relation $\mathscr{J}$. For $a \in S$ let J_a be the $\mathscr{J}$-class of a and $J(a) = S^1aS^1$.

THEOREM 2.1. *Let S be a regular semigroup. The following are equivalent*:

(i) *S is an $\mathcal{S}_1$-semigroup.*

(ii) $\mathcal{J}^* = S \times S$.

(iii) *Each $\mathcal{S}$-unitary subsemigroup of S that is a union of $\mathcal{J}$-classes is an ideal.*

Proof. Howie and Lallement **[2]** have shown that $\mathcal{J}^*$ is the least semilattice congruence on S. Hence (i) and (ii) are equivalent.

(ii) *implies* (iii). Let H be an $\mathcal{S}$-unitary subsemigroup that is a union of $\mathcal{J}$-classes but is not an ideal. Suppose $xay \in H$ for some $x, y \in S^1$, $a \in S$. Then $HxayH \subseteq H$ so $Hxa \subseteq H$. Hence $HxaH \subseteq H$ so $aH \subseteq H$ and $HaH \subseteq H$. If $J_a = J_b$, $b \in S$, then there exists $r, s \in S^1$ so that $HxrbsyH \subseteq H$ and similarly $HbH \subseteq H$. So $HJ_aH \subseteq H$. Since H is not an ideal and is $\mathcal{S}$-unitary there is a $d \in S$ so that $HdH \not\subseteq H$. Define

$$(1)\qquad \begin{aligned} C_h &= \{a \in S;\ xJ_ay \cap H \neq \Box \text{ for some } x, y \in S^1\} \text{ and} \\ \bar{C}_H &= \{d \in S;\ d \notin C_H\}. \end{aligned}$$

Let ρ_H denote the equivalence relation on S with classes C_H and $\bar{C}_H$. If $a \in C_H$ then we have $HJ_aH \subseteq H \subseteq C_H$. Furthermore, since H is $\mathcal{S}$-unitary, $HabH \subseteq H$ if and only if $HaH, HbH \subseteq H$, for $a, b \in S^1$. Hence C_H is a unitary semigroup, $\bar{C}_H$ is an ideal, and ρ_H is a nonuniversal semilattice congruence.

(iii) *implies* (ii). Since a $\mathcal{J}^*$-class is $\mathcal{S}$-unitary, it is an ideal. But ideals of S intersect nontrivially.

The next theorem is an immediate consequence of results in **[5]**, **[6]** or **[9]**.

THEOREM 2.2. *For any semigroup S the following are equivalent*:

(i) *S is an $\mathcal{S}_1$-semigroup.*

(ii) *Each ideal of S is an $\mathcal{S}_1$-semigroup.*

(iii) *S is an ideal extension of an $\mathcal{S}_1$-semigroup I by an $\mathcal{S}_1$-semigroup T.*

Note that T has zero divisors.

COROLLARY 2.3. *Let S be a regular semigroup with a principal series. S is a $\mathcal{S}_1$-semigroup if and only if each 0-simple principal factor of S has a zero divisor.*

Proof. By Theorem 2.2, the condition is clearly necessary.

Conversely let $S_0 \subset S_1 \subset \cdots \subset S_n = S$ be a principal series. Since S_0 is simple it is an $\mathscr{S}_1$-semigroup. Continuing by induction, assume S_{i-1} is an $\mathscr{S}_1$-semigroup and S_i/S_{i-1} has zero divisors for some i, $1 \leqq i \leqq n$. S_i/S_{i-1} is 0-simple so is an $\mathscr{S}_1$-semigroup. Hence, by Theorem 2.2 (iii), S_i is an $\mathscr{S}_1$-semigroup.

Let $B(G, I)$ denote the Brandt semigroup that is a Rees matrix semigroup over the group with zero G^0 and with the identity $I \times I$ sandwich matrix. We call G the *basic group* of $B(G, I)$. $B(G, I)$ has zero divisors if $|I| > 1$ and is isomorphic to G^0 if $|I| = 1$. Since an inverse semigroup is completely [0]-simple if and only if it is a group [Brandt semigroup], we have from Corollary 2.3:

COROLLARY 2.4. *Let S be a completely semisimple inverse semigroup with principal series $S_0 \subset S_1 \subset \cdots \subset S_n = S$. S is an $\mathscr{S}_1$-semigroup if and only if* (i) *S_0 is a group, and* (ii) *$S_i/S_{i-1} \cong B(G_i, I_i)$ with $|I_i| > 1$ for $1 \leqq i \leqq n$.*

We conclude this section with a similar result for inverse $\mathscr{S}_2$-semigroups.

THEOREM 2.5. *Let S be a completely semisimple inverse semigroup with principal series $S_0 \subset S_1 \subset \cdots \subset S_n = S$. S is an $\mathscr{S}_2$-semigroup if and only if* (i) *S_0 is a group and* (ii) *$S_i/S_{i-1} \cong B(G_i, I_i)$ for $1 \leqq i \leqq n$ where $|I_r| = 1$ for exactly one r, $1 \leqq r \leqq n$.*

Proof. We first observe that if $|I_r| = 1$ and $J_r = S_r \backslash S_{r-1}$, $1 \leqq r \leqq n$, then J_r is an $\mathscr{S}$-unitary subgroup of S that is a $\mathscr{J}$-class but not an ideal. As in the proof of Theorem 2.1 there is a semilattice congruence ρ_{J_r} with classes C_{J_r}, $\bar{C}_{J_r}$ defined as in (1).

Assume that S is an $\mathscr{S}_2$-semigroup then S_0 is a group and by Corollary 2.4 there exists an r, $1 \leqq r \leqq n$, so that $|I_r| = 1$. Suppose also that $|I_t| = 1$, $1 \leqq t \leqq n$. Then $\rho_{J_r} = \mathscr{J}^* = \rho_{J_t}$. Hence $C_{J_r} = C_{J_t}$ and since J_r, J_t are $\mathscr{J}$-classes, $r = t$.

Conversely assume (i) and (ii). As in the proof of Theorem 2.1, C_{J_r} is a unitary subsemigroup and $\bar{C}_{J_r}$ is an ideal of S. Then the $\mathscr{J}$-classes of C_{J_r} and the $\mathscr{J}$-classes of $\bar{C}_{J_r}$ are $\mathscr{J}$-classes of S. Since $S_0 \subseteq \bar{C}_{J_r}$ and $J_r \subseteq C_{J_r}$ are the only $\mathscr{J}$-classes that are groups then $\bar{C}_{J_r}$, C_{J_r} are $\mathscr{S}_1$-semigroups. Hence $\rho_{J_r} = \mathscr{J}^*$.

3. Characterization.

In this section completely semisimple inverse Δ-semigroups with principal series are characterized.

The following Lemma is an immediate consequence of results of Preston **[7]**. Parts (i) and (ii) are also corollaries of Tamura **[10]**.

LEMMA 3.1. *Let $S = B(G, I)$ be a Brandt semigroup.*
(i) *S is Δ-semigroup if and only if G is a Δ-group.*
(ii) *Each congruence of S is idempotent separating or universal.*
(iii) *S is primitive.*

We need some further results.

LEMMA 3.2. *Let S be an inverse semigroup with ideal I. Any congruence ρ' on I extends to a congruence ρ on S so that*

$$a\rho = \begin{cases} a\rho' & \text{if } a \in I \\ \{a\} & \text{if } a \in S \setminus I. \end{cases}$$

In particular, any ideal of an inverse Δ-semigroup is an inverse Δ-semigroup.

Proof. Let A and B be congruence classes of ρ'. Suppose $xay \in B$ for some $x, y \in S^1$, $a \in A$. Since $xaa^{-1}aa^{-1}ay \in B$ and $xaa^{-1}, a^{-1}ay \in I$ then $xaa^{-1}Aa^{-1}ay \subseteq B$. If $c \in A$ then $aa^{-1}ca^{-1}a \in A$ and $xcc^{-1}aa^{-1}ca^{-1}ac^{-1}cy = xaa^{-1}ca^{-1}ay \in B$ so $xcc^{-1}Ac^{-1}cy \subseteq B$. In particular $xcy \in B$. Hence $xAy \subseteq B$. Since I is an ideal the result follows.

If S is an inverse semigroup with semilattice E, let $C(E)$ denote the centralizer of E in S.

LEMMA 3.3. *Let S be a completely semisimple inverse semigroup with principal series $\{0\} \subset S_1 \subset S$ and with semilattice E. Then on S;*

(i) *Each non idempotent separating congruence has S_1 or S as a congruence class if and only if for any $e, f \in E$ so that $e \in S \backslash S_1$, $f \in S_1 \backslash 0$ there exists $a \in S$ so that $a^{-1}ea = f$ and so that $fa = 0$ if $e > f$.*

(ii) *Each idempotent separating congruence is the identity equivalence on $S \setminus S_1$ if and only if $C(E) \cap (S \setminus S_1) \subseteq E$.*

Proof. (i) Suppose the non idempotent separating congruences have S_1 or S as congruence classes. If $a \in S_1 \backslash 0$ then $S_1 = J(a)$. If $b \notin J(a)$ then considering the Rees congruence modulo $J(b)$ we see that $J(a) \subset J(b) = S$. Hence the principal ideals of S are chain ordered. Let τ be the least congruence so that for some $e \neq f$ in $E \backslash 0$, $(e, f) \in \tau$. Assume that $e \in S \backslash S_1$ and $f \in S_1$. Then τ is universal. Since $0 \in f\tau$ then by Teissier **[12]** there exists $x_1, y_1, \cdots, x_n, y_n \in S^1$ so that

$$f = x_1 i_1 y_1, \quad x_1 j_1 y_1 = x_2 i_2 y_2, \cdots, \quad x_n j_n y_n = 0$$

where $i_p, j_p \in \{e, f\}$, $p = 1, \cdots, n$. But $(x_p i_p y_p)^{-1}(x_p i_p y_p) = z_p^{-1} i_p z_p$ where $z_p = x_p^{-1} x_p y_p$. So

$$f = z_1^{-1} i_1 z_1, \quad z_1^{-1} j_1 z_1 = z_2^{-1} i_2 z_2, \cdots, \quad z_n^{-1} j_n z_n = 0.$$

Deleting repetitious terms we may assume that $z_p^{-1} i_p z_p \neq z_p^{-1} j_p z_p$, and that $z_p^{-1} i_p z_p > z_p^{-1} j_p z_p$ (otherwise replace z_q by $z_q z_p^{-1} i_p z_p$ for $p \leqq q \leqq n$). If $e > f$, then $z_p^{-1} e z_p \geqq z_p^{-1} f z_p$, so $i_p = e$, $j_p = f$. Furthermore, by Lemma 3.1 (iii) we have $f = z_1^{-1} e z_1 > z_1^{-1} f z_1 = 0$. Hence $(fz_1)^{-1}(fz_1) = 0$ so $fz_1 = 0$.

Conversely, suppose $e \neq f$ in E. If $e \in S \backslash S_1$, and $f \in S_1 \backslash 0$, then $a^{-1} e a = f$ for some $a \in S$, so $J(f) \subset J(e) = S$. Let τ be the least congruence with $(e, f) \in \tau$. By Lemma 3.1 (ii), if $e, f \in S_1$, then $e\tau \supseteq S_1$. If $e, f \in S \backslash S_1$, then, by Lemma 3.1 (iii), $ef \in S_1$ and $ef \in e\tau$. If $e \in S \backslash S_1$ and $f \in S_1$, then $0 \in e\tau$ since either $e > f$ and $0 = a^{-1} f a \leqq a^{-1} e a = f$ for some $a \in S$, or $ef = 0$ by Lemma 3.1 (iii). Then $e\tau = J(e) = S$.

(ii) By [1] the greatest idempotent separating congruence on S has group kernel normal system $\{H_e \cap C(E);\ e \in E\}$ where H_e is the $\mathcal{H}$-class of e.

LEMMA 3.4. *Let S be a completely semisimple inverse $\mathcal{S}_1$-semigroup with principal series $\{0\} \subset S_1 \subset S$. S is a Δ-semigroup if and only if*

(i) *the Brandt semigroups S/S_1 and S_1 have Δ-basic groups,*

(ii) *each non idempotent separating congruence of S has S_1 or S as a congruence class, and*

(iii) *each idempotent separating congruence of S is the identity equivalence on $S \backslash S_1$.*

Proof. Let S be a Δ-semigroup. By Lemmas 3.1 (i) and 3.2, (i) is satisfied. Comparing congruences with the Rees congruence modulo S_1 we see that (iii) holds and that any non universal congruence has its classes in S_1 or $S \backslash S_1$. Hence, applying Lemma 3.1 (ii) to S_1, we see that (ii) holds.

Conversely, by (i), (iii) and Lemma 3.1 (i) applied to S_1, the idempotent separating congruences are chain ordered. By (ii) the other non universal congruences have S_1 as a class and are then chain ordered since, by (i), S/S_1 is a Δ-semigroup. Hence, by (iii), S is a Δ-semigroup.

LEMMA 3.5. *Let S be a completely semisimple inverse $\mathcal{S}_2$-semigroup with principal series $\{0\} \subset S_1 \subset S$. S is a Δ-semigroup if and only if S_1 is an $\mathcal{S}_1$-Δ-semigroup, $S \backslash S_1$ is a Δ-group and S satisfies conditions* (ii) *and* (iii) *of Lemma* 3.4.

Proof. By Theorem 2.5 just one of $S_1 \backslash 0$ or $S \backslash S_1$ is a group. Assume S is a Δ-semigroup. The $\mathcal{J}$-classes of S are chain ordered **[11]**. If $S_1\backslash 0$ is a group then, as in the proof of Theorem 2.5, $\mathcal{J}^*$ has classes $\{0\}$, $S\backslash 0$. But then $\mathcal{J}^*$ is not comparable with the Rees congruence modulo S_1. Hence $S\backslash S_1$ is a group while $S_1\backslash 0$ is not. The remainder of the proof is as for Lemma 3.4.

The following theorem is the main result. Together with the results 2.4, 2.5, 3.1(i), 3.3, 3.4 and 3.5, it provides a characterization of completely semisimple inverse Δ-semigroups with principal series in terms of Δ-groups and idempotent properties.

THEOREM 3.6. *Let S be a completely semisimple inverse semigroup with principal series $S_0 \subset S_1 \subset \cdots \subset S_n = S$. S is a Δ-semigroup if and only if*

(i) S_0 *is a Δ-group*; $S_0 = \{0\}$ *if* $n > 0$,
(ii) S_1 *is a Brandt semigroup with Δ-basic group if* $n > 0$,
(iii) S_i/S_{i-2} *is an* $\mathcal{S}_1$-Δ-*semigroup for* $i = 2, \cdots, n-1$,
(iv) S_n/S_{n-2} *is an* $\mathcal{S}_1$-, *or* $\mathcal{S}_2$-Δ-*semigroup.*

Proof. Say S is a Δ-semigroup. S and S_0 have the same maximal group homomorphic image and if $n > 0$ the only such group is trivial **[11]**. Hence, by Lemmas 3.2, 3.1(i), (i), 3.4 and 3.5, we see that (i), $\cdots$(iv) are satisfied.

Conversely we prove that for any congruence ρ on S and some i, $0 \leqq i \leqq n$, then $a\rho = S_i$ for $a \in S_i$, $e\rho \cap E = \{e\}$ for $e \in (S_{i+1}\backslash S_i) \cap E$ where E is the semilattice of S, and $a\rho = \{a\}$ for $a \in S\backslash S_{i+1}$. Then S will be a Δ-semigroup. The result holds for $n = 0$ or 1, by Lemma 3.1 (ii). Continue by induction, assuming the result for $n = t$. Since S_{t+1}/S_{t-1} is a Δ-semigroup, then the congruences of S_{t+1} that have their classes in S_t or $S_{t+1}\backslash S_t$ are of the required form by Lemmas 3.4, 3.5. Suppose ρ is a congruence on S_{t+1} with $(a, b) \in \rho$, $a \in S_{t+1}\backslash S_t$, $b \in S_t$. Then the congruence on the Δ-semigroup S_{t+1}/S_{t-1} induced by ρ is universal by Lemmas 3.4, 3.5. Hence there exists $h \in S_t\backslash S_{t-1}$, $k \in S_{t-1}$ so that $(h, k) \in \rho$. But then, by the induction assumption, $S_t \subseteq h\rho$. Since S_{t+1}/S_{t-1} is a Δ-semigroup, then, by Lemmas 3.4, 3.5, $a\rho = S_{t+1}$.

4. Further study of finite case.

We now investigate circumstances under which the extensions of Theorem 3.6 are possible for finite inverse semigroups. Some further information is required.

EXAMPLE 1. Let H_X be the subgroup of the symmetric group P_X whose elements displace a finite number of elements of the set X. The alternating group A_X is a simple normal subgroup of H_X with index 2 (see **[3]**) for $|X| \neq 4$. Hence H_X is a Δ-group. In particular if $|X|$ is finite then P_X is a Δ-group.

EXAMPLE 2. The symmetric inverse semigroup $\mathcal{I}_X$, $|X|$ finite, is a Δ-semigroup. To see this, let $D(\alpha)$, $R(\alpha)$ denote the domain and range of $\alpha \in \mathcal{I}_X$ respectively. Any ideal of $\mathcal{I}_X$ is of the form $I_n = \{\alpha \in \mathcal{I}_X; |D(\alpha)| \leqq n\}$. Since $\mathcal{I}_X$ is finite it has a principal series and is completely semisimple. If α is an idempotent its $\mathcal{H}$-class is $\{\beta \in \mathcal{I}_X;\ D(\beta) = R(\beta) = D(\alpha)\}$ (see **[4]**), which is the symmetric group on $D(\alpha)$. So for some α a non group principal factor has Δ-basic group isomorphic to $P_{D(\alpha)}$. If α, γ are idempotents so that $|D(\alpha)| > |D(\gamma)| \geqq 1$ then there is a $\beta \in \mathcal{I}_X$ so that $\beta^{-1}\alpha\beta = \gamma$ and $|D(\beta^{-1}\gamma\beta)| < |D(\gamma)|$. If α is not an idempotent there is an idempotent β so that $|D(\alpha)| - |D(\beta)| \leqq 1$ and $\beta\alpha \neq \alpha\beta$. $\mathcal{I}_X$ can now be seen to satisfy the requirements of Lemmas 3.3, 3.4 and Theorem 3.6.

Let Z_n denote the set $\{1, 2, \cdots, n\}$. If $X = Z_n$ write $P_n = P_X$ and $\mathcal{I}_n = \mathcal{I}_X$.

Suppose S is a finite Δ-semigroup with $\{0\} \subset S_1 \subset S$, $S_1 \cong B(G, Z_n)$ and $S/S_1 \cong B(H, Z_r)$. Let $(G \times Z_n \times Z_n) \cup \{0\}$ denote the set of elements of S_1, with the binary operation $(x, i, j)(y, h, k) = (xy, i, k)$ if $j = h$, and 0 if $j \neq h$.

Denote the semigroup of right translations of S_1 by $P(S_1)$ and for $a \in S$ define $\rho^a \in P(S_1)$ by $b\rho^a = ba$ for all $b \in S_1$. Since inverse semigroups are left reductive there is a unique homomorphism $\theta: S \to P(S_1)$ so that the restriction of θ to S_1 is the regular representation of S_1 (by **[6**; III.1.12]). θ is given by $a\theta = \rho^a$, $a \in S$, and $(S_1)\theta \cong S_1$. Since S is a Δ-semigroup then, by Lemma 3.4 or 3.5, θ is injective. Call θ the *extension homomorphism of S*.

Let 1 denote the identity of G. For $u \in S$, $i \in Z_n$ define $D(u) = \{j \in Z_n;\ (1, i, j)u\theta \neq 0\}$. By **[6**; V.3.6 and V.5.4] there exists $\phi_u \in \mathcal{I}_n$ with domain $D(\phi_u) = D(u)$ and a map $a_u: D(u) \to G$ so that

$$(x, i, j)u\theta = \begin{cases} (x(ja_u), i, j\phi_u) & \text{if } j \in D(u), \\ 0 & \text{if } j \notin D(u). \end{cases}$$

Furthermore the map given by $u\theta \to (a_u, \phi_u)$ defines an isomorphism between $(S)\theta$ and the semigroup $\{(a_u, \phi_u);\ u \in S\}$ with the binary operation $(a_u, \phi_u)(a_v, \phi_v) = (a_u \cdot a_v, \phi_u\phi_v)$ where $j(a_u \cdot a_v) = (ja_u)(j\phi_u a_v)$. Since θ is an isomorphism then $(a_{uv}, \phi_{uv}) = (a_u \cdot a_v, \phi_u\phi_v)$. Note that with the operations $\cdot$ and composition of maps, the sets $\{a_u;\ u \in S\}$ and $\{\phi_u;\ u \in S\}$ respectively are homomorphic images of S. For convenience we will identify $u\theta$ and (a_u, ϕ_u) for each $u \in S$.

Clearly $v \in S_1$ if and only if $|D(\phi_v)| \leqq 1$. Since ϕ_u is a bijection for $u \in S$ then $\{v \in S;\ |D(\phi_v)| \leqq |D(\phi_u)|\}$ is an ideal of S. Hence for $u, v \in S \backslash S_1$, $|D(\phi_u)| = |D(\phi_v)|$; call this number the *rank* of S/S_1. Clearly e is an idempotent of $S \backslash 0$ if and only if $(D(\phi_e))a_e = \{1\}$

and ϕ_e is an identity map. A product of distinct idempotents $e, f \in S$ is in S_1 so $|D(\phi_e) \cap D(\phi_f)| \leqq 1$. Hence if $S \backslash S_1$ is not a group then the rank of S/S_1 is bounded above by $[(n+1)/2]$.

DEFINITION 4.1. For integers m and n, $1 < m \leqq n$, let ${}^n\Gamma_m$ denote the largest number so that $\{Y_i;\ |Y_i| = m,\ i = 1, \cdots, {}^n\Gamma_m\}$ is a family of subsets of Z_n with $|Y_i \cap Y_j| \leqq 1$ for $i \neq j$. For an integer r, $1 < r \leqq {}^n\Gamma_m$, let $\mathscr{A} = \{X_i;\ |X_i| = m,\ i \in Z_r\}$ be a family of subsets of Z_n with $|X_i \cap X_j| \leqq 1$ for $i \neq j$. Let

$$\mathscr{A}^* = \{\alpha \in \mathscr{I}_n;\ \alpha = 0 \quad \text{or} \quad D(\alpha), R(\alpha) \in \mathscr{A}\}$$

with a binary operation $*$ so that

$$\alpha * \beta = \begin{cases} \alpha\beta & \text{if } R(\alpha) = D(\beta) \\ 0 & \text{if } R(\alpha) \neq D(\beta). \end{cases}$$

LEMMA 4.2. *Let S be a Δ-semigroup with principal series $\{0\} \subset S_1 \subset S$ so that $S_1 \cong B(G, Z_n)$ and $S/S_1 \cong B(H, Z_r)$ has rank m. Then*

(i) *either $1 < m \leqq [(n+1)/2]$ and $1 < r \leqq {}^n\Gamma_m$ or $1 < m \leqq n$ and $r = 1$,*

(ii) *H is embeddable in the symmetric group P_m.*

Proof. Part (i) follows from the preceding commets and definition. Let $Q = (S \backslash S_1) \cup \{0\}$ and define a binary operation $*$ so that $u * v = uv$ if $uv \in S \backslash S_1$, and 0 otherwise. Then $Q \cong B(H, Z_r)$. Let $\mathscr{A} = \{D(\phi_u);\ u \in S \backslash S_1\}$. The map $\delta\colon Q \to \mathscr{A}^*$ given by $u\delta = \phi_u$ is a homomorphism. If $u \neq v$ in $S \backslash S_1$ and $\phi_u = \phi_v$ then $|D(\phi_{uv^{-1}})| > 1$ so uv^{-1} is a non idempotent element of $S \backslash S_1$. But then for any idempotent $e \in S$, it can be readily shown that $(euv^{-1})\theta = (uv^{-1}e)\theta$ so $euv^{-1} = uv^{-1}e$. This contradicts Lemmas 3.3 and 3.4 or 3.5. Hence δ is injective. If $e \neq 0$ is an idempotent of Q then it can be easily shown that the $\mathscr{H}$-class of e in Q is $H_e = \{u \in Q;\ D(\phi_u) = R(\phi_u) = D(\phi_e)\}$. Then $(H_e)\delta \cong H_e \cong H$. Part (ii) follows since the elements of $(H_e)\delta$ are permutations of $D(\phi_e)$.

Let $\mathscr{B}^* = \{a_u;\ u \in S \backslash S_1\} \cup \{0\}$ with a binary operation $*$ so that $a_u * a_v = a_u \cdot a_v$ if $uv \in S \backslash S_1$, and 0 otherwise. Since δ is injective, if $\phi_u = \phi_v$ for $u, v \in S \backslash S_1$ then $u = v$ so $a_u = a_v$. Hence there is a homomorphism $\lambda\colon (Q)\delta \to \mathscr{B}^*$ given by $\phi_u\lambda = a_u$ if $u \in S \backslash S_1$ and $0\lambda = 0$. The set $\bar{H} = \{(u\delta\lambda, u\delta);\ u \in Q\}$ with the binary operation so that $(u\delta\lambda, u\delta)(v\delta\lambda, v\lambda) = ((u * v)\delta\lambda, (u * v)\delta)$ is then a semigroup isomorphic to Q.

DEFINITION 4.3. A *structure data set* is a set $\{n, r, m, G, \bar{H}\}$ defined as follows:

(i) n, r and m are integers so that either $1 < m \leqq [(n+1)/2]$ and $1 < r \leqq {}^n\Gamma_m$, or $1 < m \leqq n$ and $r = 1$.

(ii) G is a Δ-group.

(iii) Let $\mathscr{A} = \{X_i; |X_i| = m,\ i \in Z_r\}$ be a family of subsets of Z_n so that $|X_i \cap X_j| \leqq 1$ if $i \neq j$. Let H be a Δ-subgroup of the symmetric group P_m. Let K be a subsemigroup of $\mathscr{A}^*$ so that $K \cong B(H, Z_r)$. Let $\lambda: K \to \mathscr{B}^*$ be a surjective homomorphism so that for $\phi \in K$, $\phi\lambda: D(\phi) \to G$ is a map and so that for $j \in D(\phi * \psi)$ then $j(\phi * \psi)\lambda = (j(\phi\lambda))(j\phi(\psi\lambda))$. Define $\bar{H} = \{(\phi\lambda, \phi); \phi \in K\}$ with a binary operation so that $(\phi\lambda, \phi)(\psi\lambda, \psi) = ((\phi * \psi)\lambda, \phi * \psi)$. Write $\phi\lambda * \psi\lambda = (\phi * \psi)\lambda$.

Notice that in the terminology of **[6]**, $\bar{H}$ satisfies this definition if and only if $\bar{H}$ is a subsemigroup of the right wreath product of G and K so that the map $\bar{H} \to K$ given by $(\phi\lambda, \phi) \to \phi$ is an isomorphism.

We have seen that any finite inverse Δ-semigroup S with principal series $\{0\} \subset S_1 \subset S$ determines a structure data set $\{n, r, m, G, \bar{H}\}$. Call this a *structure data set of* S. We say that structure data sets $\{n, r, m, G, \bar{H}\}$ and $\{n', r', m', G', \bar{H}'\}$ are *equivalent* if and only if $n = n'$, $r = r'$, $m = m'$ and there exists an isomorphism $\alpha: G° \to (G')°$ and a bijection $\beta: Z_n \to Z_n$ so that the map $\gamma: \bar{H} \to \bar{H}'$ given by $(a, \phi)\gamma = (\beta^{-1}a\alpha, \beta^{-1}\phi\beta)$ is a bijection.

LEMMA 4.4. *Let S and T be finite inverse Δ-semigroups with principal series $\{0\} \subset S_1 \subset S$ and $\{0\} \subset T_1 \subset T$ respectively. Then $S \cong T$ if and only if the structure data sets of S and T are all equivalent.*

Proof. Lable the elements of S_1 and T_1 so that $S_1 = (G \times Z_n \times Z_n) \cup \{0\}$ and $T_1 = (G' \times Z_{n'} \times Z_{n'}) \cup \{0\}$, with binary operations as defined after Example 2. Then structure data sets $\{n, r, m, G, \bar{H}\}$ and $\{n', r', m', G', \bar{H}'\}$ of S and T respectively can be uniquely determined by the method described above. Depending on the labelling of the elements of S_1, each structure data set of S can be so determined. Let θ_S and θ_T be the extension homomorphisms of S and T respectively and let $\eta: S \to T$ be an isomorphism. Then $n = n'$, $r = r'$ and the restriction of η to S_1 determines an isomorphism $\alpha: G° \to (G')°$ and a bijection $\beta: Z_n \to Z_n$ so that $(x, i, j)\eta = (x\alpha, i\beta, j\beta) \in T_1$. The map $u\theta_S$, $u \in S$, is given by $v(u\theta_S) = vu$ for all $v \in S_1$. Let $u\eta\theta_T = (b_{u\eta}, \psi_{u\eta})$ and $v = (x, i, j)$ then

$$((x(ja_u))\alpha, i\beta, j\phi_u\beta) = (v(u\theta_s))\eta = (v\eta)(u\eta\theta_T) = ((x\alpha)j\beta b_{u\eta}, i\beta, j\beta\psi_{u\eta}).$$

So $(x(ja_u))\alpha = (x(j\beta b_{u\eta})\alpha^{-1})\alpha$. Since $D(u)\beta = D(u\eta)$ then $a_u =$

$\beta b_{u\eta}\alpha^{-1}$. Likewise $\phi_u\beta = \beta\psi_{u\eta}$. Thus $m = m'$ and since $\bar{H} = \{(a_u, \phi_u);\ u \in S\backslash S_1\} \cup \{0\}$ then the structure data sets are equivalent.

Conversely, given that $\{n, r, m, G, \bar{H}\}$ and $\{n', r', m', G', \bar{H}'\}$ are equivalent structure data sets, let $\alpha\colon G^\circ \to (G')^\circ$ be an isomorphism, $\beta\colon Z_n \to Z_n$ be a bijection and $\gamma\colon \bar{H} \to \bar{H}'$ be the bijection so that $(a, \phi)\gamma = (\beta^{-1}a\alpha, \beta^{-1}\phi\beta)$. Define $\eta_1\colon S_1 \to T_1$ by $(x, i, j)\eta_1 = (x\alpha, i\beta, j\beta)$. Then η_1 is an isomorphism. As in the first part of the proof we get for $v \in S_1$ that $v\eta_1\theta_T = (\beta^{-1}a_v\alpha, \beta^{-1}\phi_v\beta)$. Hence there is a bijection $\gamma'\colon (S)\theta_S \to (T)\theta_T$ given by

$$(a_u, \phi_u)\gamma' = (\beta^{-1}a_u\alpha, \beta^{-1}\phi_u\beta).$$

Define $\eta\colon S \to T$ by $u\eta\theta_T = u\theta_S\gamma'$. Then

$$\begin{aligned}(x, i, j)\eta_1(u\eta\theta_T) &= (x\alpha(j\beta\beta^{-1}a_u\alpha), i\beta, j\beta\beta^{-1}\phi_u\beta) = ((x(ja_u))\alpha, i\beta, j\phi_u\beta)\\ &= (x, i, j)u\theta_S\eta_1.\end{aligned}$$

So $u\eta\theta_T = \eta_1^{-1}(u\theta_S)\eta_1$ and clearly η is an isomorphism.

THEOREM 4.5. *Each finite inverse Δ-semigroup S with principal series $\{0\} \subset S_1 \subset S$ has a structure data set $\{n, r, m, G, \bar{H}\}$. A semigroup is isomorphic to S if and only if its structure data sets are equivalent to $\{n, r, m, G, \bar{H}\}$. Conversely, each structure data set $\{n, r, m, G, \bar{H}\}$ is a structure data set of some finite inverse Δ-semigroup T with principal series $\{0\} \subset T_1 \subset T$.*

Proof. The first two statements have been proved. Suppose $\{n, r, m, G, \bar{H}\}$ is a structure data set. Let $T_1 = (G \times Z_n \times Z_n) \cup \{0\}$ with binary operation as defined after Example 2. Then $T_1 \cong B(G, Z_n)$. Let $T = T_1 \cup \bar{H}\backslash(0,0)$. For $(a, \phi), (b, \psi) \in \bar{H}\backslash(0,0)$ and $(x, i, j), (y, h, k) \in T_1$ define a binary operation on T so that:

$$(a, \phi)(b, \psi) = \begin{cases} (a * b, \phi\psi) \in \bar{H} & \text{if} \quad D(\phi\psi) = D(\phi) \\ (la(l\phi a), l, l\phi\psi) \in T_1 & \text{if} \quad D(\phi\psi) = \{l\} \\ 0 & \text{if} \quad D(\phi\psi) = \square, \end{cases}$$

$$(x, i, j)(a, \phi) = \begin{cases} (x(ja), i, j\phi) \in T_1 & \text{if} \quad j \in D(\phi) \\ 0 & \text{if} \quad j \notin D(\phi), \end{cases}$$

$$(a, \phi)(x, i, j) = \begin{cases} (i\phi^{-1}ax, i\phi^{-1}, j) \in T_1 & \text{if} \quad i \in R(\phi) \\ 0 & \text{if} \quad i \notin R(\phi), \end{cases}$$

$$(x, i, j)(y, h, k) = \begin{cases} (xy, i, k) \in T_1 & \text{if } j = h \\ 0 & \text{if } j \neq h. \end{cases}$$

Since $\bar{H} \cong B(H, Z_r)$ it can be routinely checked that T is an inverse semigroup. It can also be checked, using Lemmas 3.3 and 3.4 or 3.5, that S is a Δ-semigroup. Since $(x, i, j)(a, \phi) = (x(ja), i, j\phi)$ for $(x, i, j) \in T_1$, $(a, \phi) \in T/T_1$ and $j \in D(\phi)$ we see that $\{n, r, m, G, \bar{H}\}$ is a structure data set of T.

Let S be a finite inverse Δ-semigroup with principal series $S_0 \subset S_1 \subset \cdots \subset S_q = S$ where $q > 1$. We can uniquely determine, up to equivalence, the structure data sets of the semigroups S_i/S_{i-2} for $i = 2, \cdots, q$. Conversely, let $\{\{n_i, r_i, m_i, G_i, \bar{H}_i\};\ i = 2, \cdots, q\}$ be a family of structure data sets so that $n_j = r_{j-1}$ and $G_j \cong H_{j-1}$ for $j = 3, \cdots, q$, where H_{j-1} is the basic group of $\bar{H}_{j-1}$. Then, by Theorem 3.6 and the proof of Theorem 4.5, we can construct a finite inverse Δ-semigroup T with principal series $T_0 \subset T_1 \subset \cdots \subset T_q = T$ so that $\{n_i, r_i, m_i, G_i, \bar{H}_i\}$ is a structure data set of T_i/T_{i-2}. Any finite inverse Δ-semigroup that is not a group or a Brandt semigroup can be so constructed.

The authors are grateful for the referee's kind suggestions on this paper.

References

1. A. H. Clifford, and G. B. Preston, *The algebraic theory of semigroups*, Math. Surveys No. 7, Amer. Math. Soc., Providence, Vol. I (1961), Vol. II (1967).
2. J. M. Howie, and G. Lallement, *Certain fundamental congruences on a regular semigroup*, Proc. Glasgow Math. Assoc., **7** (1966), 145–156.
3. M. Hall Jr., *The Theory of Groups*, Macmillan, New York (1959).
4. W. D. Munn, *Fundamental inverse semigroups*, Quart. J. Math. Oxford (2), **21** (1970), 157–170.
5. M. Petrich, *The maximal semilattice decomposition of a sèmigroup*, Math. Zeitschr., **85** (1964), 68–82.
6. ———, *Introduction to semigroups*, Merrill, Columbus (1973).
7. G. B. Preston, *Congruences on Brandt semigroups*, Math. Ann., **139** (1959), 91–94.
8. B. M. Schein, *Commutative semigroups whose congruences form a chain*, Polonaise des Sciences, **9** (1969), 523–527.
9. T. Tamura, *The theory of construction of finite semigroups I*, Osaka Math. J., **8** (1956), 243–261.
10. ———, *Decomposition of a completely simple semigroup*, Osaka Math. J., **12** (1960), 269–275.
11. ———, *Commutative semigroups whose lattice of congruences form a chain*, Bull. Soc. Math. France, **97** (1969), 369–380.
12. M. Teissier, *Sur les équivalences regulières dans les demi-groupes*, C. R. Acad. Sci., Paris, **232** (1951), 1987–1989.

Received November 7, 1975 and in revised form September 13, 1976.

THE UNIVERSITY OF TASMANIA, HOBART, AUSTRALIA
AND
THE UNIVERSITY OF CALIFORNIA,
DAVIS, CA 95616

PACIFIC JOURNAL OF MATHEMATICS
Vol. 68, No 2, 1977

TORSION FREE ABELIAN GROUPS QUASI-PROJECTIVE OVER THEIR ENDOMORPHISM RINGS

C. VINSONHALER AND W. J. WICKLESS

Certain classes of torsion free abelian groups which are quasi-projective as modules over their endomorphism rings are characterized. The main results concern completely decomposable and strongly indecomposable groups.

1. Preliminaries. Abelian groups which are quasi-projective over their endomorphism rings have been characterized by Fuchs in the torsion case. His methods have been extended by Longtin to the algebraically compact and cotorsion groups **[5]**. In this paper, we investigate some other classes of groups with this property. Specifically:

DEFINITION. A (left) module M over a ring R is quasi-projective provided the natural map $\mathrm{Hom}_R(M, M) \to \mathrm{Hom}_R(M, M/K)$ is epic for every submodule K, of M.

An abelian group G will be called E-quasi-projective (Eqp) provided G is quasi-projective as a module over $E = \mathrm{End}(G)$. Henceforth, the word group will denote a torsion free abelian group. Other notation follows Fuchs **[4]**, in particular, $t(G) = $ type G for any group G of rank 1.

The following simple lemmas will be quite useful.

LEMMA 1.1. *Let G be* Eqp *and K a fully invariant subgroup of G. Then G/K is a quasi-projective E-module.*

Proof. See Proposition 2.1 in Wu and Jans **[9]**.

LEMMA 1.2. *Let G be* Eqp *and K a fully invariant subgroup. Then ZE, the center of E, maps onto* $\mathrm{Hom}_E(G/K, G/K)$.

Proof. Let Π: $G \to G/K$ be the factor map. Since G is Eqp, for every $\theta \in \mathrm{Hom}_E(G/K, G/K)$, there exists $\alpha \in \mathrm{Hom}_E(G, G) = ZE$ such that $\Pi\alpha = \theta\Pi$.

LEMMA 1.3. *Let G be* Eqp *and K a fully invariant subgroup such that G/K is torsion. Then if ZE is countable, G/K is bounded.*

Proof. If G/K is unbounded and torsion, then $\mathrm{Hom}_E(G/K, G/K)$ is uncountable: it must contain either a copy of Q_p^* (the ring of p-adic integers) for some prime p, or a direct product of an infinite number of cyclic groups. Hence the result follows from Lemma 1.2.

2. Decomposable groups. In this section, some classes of decomposable Eqp groups are characterized, including completely decomposable and homogeneous separable groups.

We begin with completely decomposable groups, those groups G isomorphic to a direct sum of rank one groups.

LEMMA 2.1. *If $G = \bigoplus_{i \in I} \Sigma A_i$ is a direct sum of rank one groups, then G is indecomposable as an E-module if and only if given any two summands A_i and A_j, there is a finite sequence $A_i = A_{i_1}, A_{i_2}, \cdots, A_{i_n} = A_j$ such that $t(A_{i_k})$ is comparable to $t(A_{i_{k+1}})$ for $k = 1, 2, \cdots, n-1$.*

Proof. If S is a subset of the set $\{A_i\}_{i \in I}$, define $I(S) = \{A_l \mid t(A_l)$ is comparable to $t(A_m)$ for some A_m in $S\}$ and $I^n(S) = I(I^{n-1}(S))$. Then it is easy to see that for a fixed A_{i_0}, $\bigoplus \Sigma\{A_i \mid A_i \in \bigcup_{n=1}^{\infty} I^n(A_{i_0})\}$ and $\bigoplus \Sigma\{A_i \mid A_i \notin \bigcup_{n=1}^{\infty} I^n(A_{i_0})$ are E-submodules whose sum is G. The lemma follows immediately.

LEMMA 2.2. *If $G = \bigoplus \Sigma A_i$ is completely decomposable and indecomposable as an $E = \mathrm{End}(G)$ module, then $ZE \subseteq Q$.*

Proof. Maps in ZE must commute with projections and maps $A_i \to A_j$. The fact that G is E-indecomposable and Lemma 2.1 imply that any map in ZE multiplies each A_i by the same rational number.

THEOREM 2.1. *Let $G = \bigoplus_{i \in I} A_i$ be a direct sum of rank one groups such that G is indecomposable as an E-module. Then the following are equivalent:*

1. *G is* Eqp.
2. *The type set $T = \{t_i \mid t_i = t(A_i)$ for some $i\}$ satisfies:*

(a) *If $t_i, t_j \in T$ and $t_i, t_j \leqq t_k$ for some $t_k \in T$, then $t_i, t_j \geqq t_l$ for some $t_l \in T$;*

(b) *Countable descending chains in T are bounded below;*

(c) *If $t_i \in T$ is finite at an infinite set of positions $\{p_j\}$, then $\exists t_k \in T$ such that t_k is 0 at an infinite subset of $\{p_j\}$.*

3. If K is a fully invariant subgroup of G such that G/K is torsion, then G/K is bounded.

Proof. (1) $\Rightarrow$ (2). Let $t_i, t_j, t_k \in T$ such that $t_i, t_j \leqq t_k$. Suppose there is no t_l such that $t_l \leqq t_i, t_j$. Let $K = \bigoplus \Sigma\{A_m \mid t_m \not\leqq t_i$ and $t_m \not\leqq t_j\}$. Then K is a fully invariant subgroup of G, and G/K is a direct sum of two E-modules, $G/K = B_1 \bigoplus B_2$ where $B_1 = \bigoplus \Sigma\{A_m \mid t_m \leqq t_i\} + K$ and $B_2 = \bigoplus \Sigma\{A_m \mid t_m \leqq t_j\} + K$. But multiplication by integers n_1 on B_1 and $n_2 \neq n_1$ on B_2 is an E-map of G/K to G/K which is not induced by a map in ZE. By Lemma 1.2, this is a contradiction.

Now suppose $t_{i1} \leqq t_{i2} \gneqq \cdots$ is a countable descending chain of types in T which is not bounded below. Let p be a prime not dividing A_{i1} and define K to be sum of $\{A_i \mid t_i \not\leqq t_{i1}\}$ and $\{p^k A_j \mid t_j \leqq t_{ik}, t_j \not\leqq t_{ik+1}, k = 1, 2, \cdots\}$. Then K is fully invariant and G/K is an unbounded torsion group. Since G is E-indecomposable, by Lemmas 2.2 and 1.3, this is a contradiction.

Finally, assume $t_0 \in T$ is finite at an infinite set of positions $\{p_j\}$, and suppose no $t_i \leqq t_0$ is zero on an infinite subset of $\{p_j\}$. Then for each $t_i \leqq t_0$, choose $x_i \in A_i$ such that p_j-height $x_i \geqq 1$ *for all* p_j. Now let H be the minimal fully invariant subgroup containing the x_i. Since homomorphisms do not decrease height, $(1/p_j)x_0 \not\in H$ for any p_j. Thus $A_0/A_0 \cap H$ contians a copy of $Z(p_j)$ for each p_j. By Lemma 1.3 this is a contradiction.

(2) $\Rightarrow$ (3). Let H be fully invariant in G such that G/H is torsion. Suppose first that for some A_i, $A_i/A_i \cap H$ is unbounded, with nonzero p_k-component for an infinite set $P = \{p_k\}$ of primes. Note that $A_i/A_i \cap H$ contains no $Z(p^\infty)$ since rank $A_i = 1$ and $A_i \cap H$ is fully invariant in A_i. We may, therefore, assume that t_i is finite and positive at all $p_k \in P$. By condition 2(c), there exists $t_j < t_i$ such that t_j is zero at an infinite subset of P. Since H is fully invariant and $t_j < t_i$, $A_i \cap H \subseteq p_k A_i$ implies $A_j \cap H \subseteq p_k A_j$ for all $p_k \in P$. This is impossible since t_j is zero at infinitely many p_k.

Now if G/H is unbounded, choose a countable sequence $A_{i1}, A_{i2}, \cdots$ such that $\bigoplus \Sigma_{r=1}^{\infty} A_{ir}/A_{ir} \cap H$ is unbounded. By conditions 2(a) and 2(b) there exists a fixed A_i with $t_i \leqq t_{ir}$ for all $r \geqq 1$. It follows that $A_i/A_i \cap H$ must be unbounded. This is impossible, as above.

(3) $\Rightarrow$ (1). It is easy to show that if G/K is bounded for all fully invariant K with G/K torsion, then $\mathrm{Hom}_E(G, G/K) = \{n\Pi \mid n \in Z\}$ where $\Pi : G \to G/K$ is the natural factor map. It follows that $\mathrm{Hom}_g(G, G/K) = \{n\Pi \mid n \in Z\}$, for any fully invariant K.

The above theorem characterizes the completely decomposable Eqp groups since any completely decomposable group G may be expressed as a direct sum $\bigoplus \Sigma G_i$ of E-indecomposable subgroups which are completely decomposable, and in this decomposition $\mathrm{End}(G_i) = E|_{G_i}$.

COROLLARY 2.1. *Let G be completely decomposable of finite rank with type set T. Then G is* Eqp *iff T satisfies* 2(*a*) *and minimal types in T are idempotent.*

Proof. T is finite so that minimal types are idempotent iff 2(c) holds. Since 2(b) holds vacuously, the result follows.

COROLLARY 2.2. *Let $G = \bigoplus_{i \in I} \Sigma A_i$ with $\{A_i \mid i \in I\}$ rigid* **[4]**. *Then G is* Eqp *iff $t(A_i)$ is idempotent for all $i \in I$.*

Proof. If $\{A_i\}$ is rigid, (a) and (b) hold vacuously and (c) holds iff each $t(A_i)$ is idempotent.

REMARK. Since E is commutative if $\{A_i\}$ is rigid, Corollary 2.2 can also be derived from a trivial modification of a result of Arnold (**[1]**, Theorem 1.1).

EXAMPLE. The following is a nontrivial (uncountable E-indecomposable) example of a completely decomposable group satisfying 2(a), 2(b) and 2(c) of 2 in Theorem 2.1.

Define a relation on the set I of all infinite subsets of the natural numbers by $S \leqq T$ iff $S \setminus T$ is finite. Let $\{S_\alpha\}$ be a maximal chain in I. It is easy to see that $\{S_\alpha\}$ is uncountable. For each α, define a type t_α by $t_\alpha = [\langle x_i^\alpha \rangle]$; $x_i^\alpha = 1$, $i \in S_\alpha$; $x_i^\alpha = 0$, $i \notin S_\alpha$. It is easy to see that $\{t_\alpha\}$ satisfies 2(a) and (b) of Theorem 2.1. By the maximality of the chain $\{S_\alpha\}$, $\{t_\alpha\}$ also satisfies 2(c). Let $A = \bigoplus \Sigma_\alpha A_\alpha$, where A_α is of rank one and type t_α. Then A is Eqp by Theorem 2.1.

We next characterize homogeneous separable Eqp groups (**[4]**, §87).

LEMMA 2.3. *Let G be homogeneous and separable. Then $ZE \subseteq Q$.*

Proof. This is an easy exercise. (See **[4]**, Problem 12, page 235.)

LEMMA 2.4. *Let G be homogeneous and separable. Then, for all nonzero fully invariant $K \subseteq G$, we have G/K is a torsion group.*

Proof. Let $0 \neq K$ be fully invariant in G. Choose $0 \neq x \in K$. Since G is homogeneous separable we can write $G = \langle x \rangle_* \oplus G'$, where $\langle x \rangle_*$ denotes the pure subgroup generated by x. If $G' \subseteq K$, then $G/K = \langle x \rangle_* / \langle x \rangle_* \cap K$ and G/K is torsion. Otherwise, choose $y \in G' \setminus G' \cap K$. Since G' is also homogeneous separable, write $G = \langle x \rangle_* \oplus \langle y \rangle_* \oplus G''$. Since G is homogeneous, there exists $\alpha \in E$, $n \in Z^+$ with $\alpha(x) = ny$. Thus, $ny \in K$. Since y was an arbitrary element of $G' \setminus G' \cap K$, we have $G/K = \langle x \rangle_* / \langle x \rangle_* \cap K \oplus G'/G' \cap K$ is torsion.

REMARK. The claim made in Problem 13, page 235 of **[4]** is incorrect. Any rank one group of nil type will sei e as a counterexample.

A group G is called strongly irreducible iff for all nonzero fully invariant $K \subseteq G$, G/K is bounded. (See **[7]**.)

THEOREM 2.2. *Let G be a homogeneous separable group. Then G is* Eqp *iff G is strongly irreducible.*

Proof. If G is homogeneous, separable and Eqp, Lemmas 1.3, 2.3 and 2.4 show that G is strongly irreducible.

Conversely, let G be strongly irreducible, homogeneous and separable. Let $K \neq (0)$ be fully invariant in G and $\theta \in \mathrm{Hom}(G, G/K)$. Write G/K in its primary decomposition, $G/K = \bigoplus \Sigma_{i=1}^{N}(G/K)_{p_i}$. Say, for some fixed $p_j \in \{p_i \mid i = 1 \cdots N\}$, we have $(G/K)_{p_j} = \bigoplus \Sigma_{\alpha \in A_j} \langle \bar{a}_\alpha \rangle$ with order $(\bar{a}_\alpha) = p_j^{s_\alpha}$, $s_\alpha \leqq s_j$ (Here $\bar{a} = a + K$). Since θ is an E map and G is homogeneous separable, it is easy to show that, for some fixed $m_j \in Z^+$, we must have $\theta(a_\alpha) = m_j \bar{a}_\alpha$ for all $\alpha \in A_j$. Choose $m \in Z^+$ with $m \equiv m_j(p_j^{s_j}), j = 1 \cdots N$. Then $\Pi m = \theta$.

The final results of this section deal with groups G which can be written as a sum of two groups related in a special way. We will need the notions of outer type (OT) and inner type (IT) of a group as defined in Warfield **[8]**.

THEOREM 2.3. *Let $G = A \oplus B$ where $IT(A) > OT(B)$ and let $\bar{E} =$* End(B). *Then G is* Eqp *iff B is* $\bar{\mathrm{E}}$qp *and rank $Z\bar{E} = 1$.*

Proof. $(\Rightarrow)$ Let K be an $\bar{E}$-submodule of B. Then $A \oplus K$ is an E-submodule of G since $\mathrm{Hom}(A, B) = 0$. Therefore, any $\bar{E}$-map $\theta: B \to B/K$, induces an E-map $0 \oplus \theta: A \oplus B \to A \oplus B/A \oplus K$ which must lift to a map in ZE of the form $\alpha \oplus \beta$, where $\alpha: A \to A$, $\beta: B \to B$. It follows that β is an $\bar{E}$-map which lifts θ.

Now suppose rank $Z\bar{E} > 1$. Choose $\gamma \in Z\bar{E}$ and $b \in B$, such that b, $\gamma(b)$ are independent. Then $0 \oplus \gamma: A \oplus B \to (A \oplus B)/A$ is an E map and lifts as above to a map of the form $\alpha \oplus \beta$ in ZE. Since $IT(A) > OT(B)$, there exists $\delta \in \mathrm{Hom}(B, A)$ such that $\delta(b) = 0$ and $\delta(\gamma(b)) \neq 0$. But then $0 = \alpha\delta(b) = \delta\beta(b) = \delta\gamma(b) \neq 0$, a contradiction.

$(\Leftarrow)$ Let K be a fully invariant subgroup of G, and $\theta: G \to G/K$ an E-map. Then $K = K \cap A \oplus K \cap B$ and $\theta(B) \subseteq B/B \cap K$ so that θ restricted to B may be lifted to a map $\alpha \in Z\bar{E} \subseteq Q$. Since $IT(A) > OT(B)$, $A = \bigcup_{f.B \to A} \mathrm{Image}\, f$. It follows that $\alpha: A \to A$ must be a lifting of $\theta|_A$.

REMARK. This theorem may be generalized slightly to the case $IT(A) \geqq OT(B)$.

COROLLARY. *If $G = D \oplus R$ where D is divisible and R is reduced, then G is* Eqp *iff R is $E(R)qp$ and rank $ZE(R) = 1$.*

3. Strongly indecomposable groups. In this section we characterize the strongly indecomposable Eqp groups of finite rank. We start by characterizing the strongly indecomposable, strongly irreducible ones. Recall that a group G is called strongly indecomposable if it admits no nontrivial quasi-decompositions ([**4**], §92).

THEOREM 3.1. *Let G be strongly indecomposable, strongly irreducible of finite rank. Then G is* Eqp *iff G/P^kG is a cyclic E module for all nonzero prime ideals $P \subseteq E$.*

Proof. Suppose G is Eqp. Since G is strongly indecomposable and strongly irreducible, we can conclude that E is a subring of an algebraic number field F with $QE = F$. (See [**7**].) Note that E is Noetherian and $P \neq (0)$ prime in E implies P is maximal. (Since $QE = F$, every nonzero ideal $I \subseteq E$ contains a nonzero rational integer. Thus, E/I is finite.)

We show G/P^kG is a cyclic E module for all nonzero prime ideals $P \subseteq E$. If not, let $X = \{\bar{x}_1 \cdots \bar{x}_n\}$ be a minimal set of E generators for G/P^kG, where $\bar{x}_i = x_i + P^kG$. Let H be given by $E\bar{x}_1 \cap \Sigma_{i=2}^n E\bar{x}_i = H/P^kG$. Then H is fully invariant and $G/H = A \oplus B$ with $A = E\bar{\bar{x}}_1$, $B = \Sigma_{i=2}^n E\bar{\bar{x}}_i$, where $\bar{\bar{x}}_i = x_i + H$. This is a nontrivial direct sum decomposition because of the minimality of X.

Let f, g be the projections from G/H onto A, B and Π: $G \to G/H$ the natural map. Let $\bar{f}, \bar{g} \in E$ be such that $\Pi\bar{f} = f\Pi$, $\Pi\bar{g} = g\Pi$. Finally, let $I = \{\alpha \in E \mid \alpha(G) \subseteq H\}$. Then $P^k \subseteq I$, so $I \subseteq p$ (primes in E are maximal). Clearly $\bar{f}\bar{g} \in I \subseteq P$, so $\bar{f} \in P$ or $\bar{g} \in P$. If $\bar{f} \in P$, $PA = A$. Thus, $P^kA = A$, so $IA = A$. But $IA = (0)$ and $A \neq (0)$ — a contradiction. A similar contradiction arises from the assumption $\bar{g} \in P$. Thus G/P^kG is cyclic.

Conversely, let G be strongly indecomposable strongly irreducible of finite rank with G/P^kG cyclic for all nonzero primes $P \subseteq E$. We show, for all positive rational integers n, G/nG is E cyclic. Let $n \in Z^+$. Since $(0) \neq (n) \subseteq E$ and E is Noetherian we have $(n) \supseteq P_1^{k_1} \cdots P_s^{k_s}$ with the P_i's nonzero prime ideals in E ([**10**], page 200). Now the ideals $P_i^{k_i}$, $i = 1 \cdots s$, are co-maximal in E ([**9**], page 176) and, by assumption, $G/P_i^{k_i}G$ is E-cyclic. It is easy to show (using the Chinese Remainder Theorem in E) that $G/(\Pi P_i^{k_i})G$ is E-cyclic. Thus, G/nG is E-cyclic.

Now let θ: $G \to G/K$ be an E map, $(0) \neq K$ a fully invariant subgroup of G. Since G is strongly irreducible, $nG \subseteq K$ for some positive integer n. Thus G/K is E-cyclic, say $G/K =$

$E(g+K)$. Choose $\alpha \in E$ with $\theta(g) = \alpha(g+K)$. We claim that α is a lifting of θ. To show this, it only remains to show that $\theta(K) = (0)$. Let $G/nG = E(h+nG)$. Then for any $k \in K$, $k = \beta h + nx$ for some $\beta \in E$, $x \in G$. Now $\beta h \in K$, so, since E is commutative, $\beta(G) \subseteq K$. Thus, $\theta\beta(G) = \beta\theta(G) = (0)$ in G/K. Finally, $\theta(k) = \theta\beta(h) + n\theta(x) = 0 + K$. This shows that G is Eqp and completes the proof.

We now consider the general case, and begin with a more general definition of quasi-projectivity which is invariant under quasi-isomorphism.

DEFINITION. If R is a ring, an R-module M is *almost* quasi-projective, if there exists a positive integer t such that the image of $\mathrm{Hom}_R(M, M)$ in $\mathrm{Hom}_R(M, M/N)$ is bounded by t for every submodule N of M.

LEMMA 3.1. *If $M \sim N$ are (quasi-isomorphic) R modules, and M is almost quasi-projective, then N is almost quasi-projective.*

Proof. Without loss of generality, assume $nM \subseteq N \subseteq M$ for some positive integer n. Let K be a submodule of N and $f: N \to N/K$. Then $nf: M \to M/K$ lifts to a map $\bar{f} \in \mathrm{Hom}_R(M, M)$ such that $\Pi\bar{f} = tnf$ where $\Pi: M \to M/K$. Then $n\bar{f} \in \mathrm{Hom}_R(N, N)$ and is a lifting of $tn^2 f$. Hence N is almost quasi-projective.

LEMMA 3.2. *Let G be strongly indecomposable and almost* Eqp. *Then there is a $g \in G$ such that G/Eg is bounded.*

Proof. Choose $\{g_1, \cdots, g_k\}$ of minimal cardinality with respect to $G/Eg_1 + Eg_2 + \cdots + Eg_k$ is bounded. This is possible by Lemma 1.3. If $k > 1$, let $H = Eg_1 \cap \Sigma_{i=2}^k Eg_i$. Then H is fully invariant and $\Sigma_{i=1}^k Eg_i/H = Eg_1 + H \oplus \Sigma_{i=2}^k Eg_i + H$. Furthermore, $Eg_1 + H$ is not torsion since $ng_1 \in H \Rightarrow nEg_1 \subseteq H$, contradicting the minimality of k. Since G is strongly indecomposable, any $\alpha \in E$ is either monic or nilpotent (see **[6]**). But if t is a positive integer such that $tG \subseteq \Sigma_{i=1}^k Eg_i$, then t followed by projection onto $\Sigma_{i=2}^k Eg_i + H$ is a map from G to G/H which cannot be lifted, as the lifting could be neither monic nor nilpotent. Thus $k = 1$, proving the lemma.

THEOREM 3.2. *Let G be strongly indecomposable of finite rank. Then G is* Eqp *iff G is strongly irreducible and G/P^kG is a cyclic E module for all nonzero prime ideals $P \subseteq E$.*

Proof. In view of Theorem 3.1, we only need show that strongly indecomposable Eqp groups of finite rank are strongly irreducible.

By the preceding Lemma $nG \subseteq Eg \subseteq G$ for some n and $Eg \cong E/L$ (as E-modules) for some left ideal $L \subseteq J(E)$, the Jacobson radical of E. Therefore by Lemma 3.1, E/L is almost Eqp, with associated integer t, for some $t > 0$.

Now for any $x \in E$ consider

$$\begin{tikzcd} & E/L \arrow[dl, dashed, "\alpha"'] \arrow[d, "x_l"] & \\ E/L \arrow[r, "\pi"] & E/Lx+L \arrow[r] & 0 \end{tikzcd}$$

x_l = left multiplication by x
$Lx + L \subseteq J(E) \neq E$.

Then $n\alpha$ is an E endomorphism of G, hence in ZE. Furthermore $n\alpha - tx_l \colon E \to Lx + L$, so that $n\alpha - tx \in Lx + L$. Hence $tx \in ZE + Lx + L$. This implies $Ltx \subseteq L + L^2x$, so that $t^2x \in Z(E) + L^2x + L$. Continuing inductively $t^kx \in Z(E) + L^kx + L$. Since L is nilpotent ($L \subseteq J(E)$), for some $m > 0$, $L^m = 0$ and $t^mx \in ZE + L$. Thus $t^mE \subset Z(E) + L$, and $G \sim E/L \sim Z(E) + L/L \cong ZE/L \cap ZE$, a commutative ring with identity. By ([2], Th. 1.4, Cor. 3.6, Th. 1.13) G must be strongly irreducible.

COROLLARY 3.1. *Let G be finite rank strongly indecomposable with rank $E <$* rank *G. Then G is not* Eqp.

Proof. For any $0 \neq g \in G$, Eg is a fully invariant subgroup of G with rank $Eg \leqq$ rank $E <$ rank G. Thus, G is not strongly irreducible, so G cannot be Eqp.

4. Groups of rank two. In this section we use the results of §§1–3 to survey the Eqp property for groups of rank two. This is most conveniently done by considering the six possibilities for the quasi-endomorphism algebra, $QE(G) = Q \otimes_Z E(G)$. (See [3].) If $QE(G) \cong [Q]_{2\times 2}$ or $QE(G) \cong \left\{ \begin{pmatrix} x & 0 \\ y & z \end{pmatrix} \middle| \; x, y, z \in Q \right\}$ then G is completely decomposable. In the first case we have $G = A \oplus A$, and in the second case $G = A \oplus B$ with A, B of rank one, $t(A) < t(B)$. In either case Corollary 2.1 applies; G is Eqp iff $t(A)$ is idempotent. If $QE(G) \cong Q \oplus Q$, then G is quasi-decomposable $G \sim A \oplus B$ with $t(A)$, $t(B)$ incomparable. A slight modification of the arguments of Theorem 2.1 prove that G is Eqp iff $t(A)$ and $t(B)$ are idempotent.

We next consider the strongly indecomposable cases. If $QE(G) \cong$

Q or $QE(G) \cong \left\{ \begin{pmatrix} x & 0 \\ y & x \end{pmatrix} \middle| \; x, y \in Q \right\}$, then G is strongly indecomposable but not strongly irreducible, so G is not Eqp by Theorem 3.2. We settle the final possibility, $QE(G) \cong Q(\sqrt{N})$, in the following theorem.

THEOREM 4.1. *Let G be of rank two with* $QE(G) \cong Q(\sqrt{N})$. *Then g is* Eqp *iff G is strongly irreducible.*

Proof. If G is Eqp, G is strongly irreducible by Theorem 3.2. Conversely, let G be strongly irreducible and K any nonzero fully invariant subgroup of G. Write the finite group G/K in its primary decomposition: $G/K = \bigoplus_{i=1}^{n} (G/K)_{p_i}$. Since rank $G = 2$, K is fully invariant, and $QE(G) = Q(\sqrt{N})$, it is easy to show, for each p_i, either $(G/K_i)_{p_i} = Z(p_i^{s_i})$ for some $s_i \geqq 0$ in Z, or $(G/K_i)_{p_i} = Z(p_i^{t_i}) \oplus Z(p_i^{t_i})$ for some $t_i \geqq 0$ in Z. Moreover, in the latter case we can choose $a \in G$ so that $a + K_i$ and $\sqrt{N}a + K_i$ are generators of $(G/K_i)_{p_i}$. It is now easy to check that G/K is a cyclic E module. Thus, Theorem 3.2 applies and G is Eqp.

REFERENCES

1. D. M. Arnold and E. L. Lady, *Endomorphism Rings and Direct Sums of Torsion Free Abelian Groups*, Trans. Amer. Math. Soc., **211** (1975), 225–237.

2. R. A. Beaumont and R. S. Pierce, *Torsion free rings*, Illinois J. Math., **5** (1961), 61–98.

3. ———, *Torsion Free Groups of Rank Two*, Amer. Math. Soc., Mem. 38, A. M. S., Providence, 1961.

4. L. Fuchs, *Infinite Abelian Groups, Vol. II*, Academic Press, New York, 1973.

5. ———, *On Torsion Abelian Groups Quasi-Projective over Their Endomorphism Rings*, Proc. Amer. Math. Soc., **42**, No. 1, Jan. (1974), 13–15.

6. J. D. Reid, *On the Ring of Quasi-Endomorphisms of a Torsion Free Group*, Topics in Abelian Groups, Chicago, 1963, 51–58.

7. ———, *Remarks on irreducible groups*, preprint.

8. R. B. Warfield, Jr., *Homomorphisms and duality for torsion-free Abelian groups*, Math. Z., **107** (1968), 189–200.

9. L. E. T. Wu and J. P. Jans, *On Quasi Projectives*, Illinois J. Math., **11** (1967), 439–448.

10. O. Zariski and P. Samuel, *Commutative Algebra*, Vol. I, Van Nostrand, Princeton, 1958.

Received May 3, 1976.

UNIVERSITY OF CONNECTICUT
STORRS, CT 06268

PACIFIC JOURNAL OF MATHEMATICS
Vol. 68, No. 2, 1977

TOPOLOGIES ON THE SET OF CLOSED SUBSETS

FRANK WATTENBERG

In this paper the techniques of Nonstandard Analysis are used to study topologies on the set $\Gamma(X)$ of closed subsets of a topological space X. The first section of the paper investigates the "compact" topology developed by Narens and constructs a variant of that topology which is particularly useful for non locally compact spaces X. (When X is locally compact this variant is shown to be identical with Naren's original "compact" topology.) This new topology is a natural extension to $\Gamma(X)$ of the one point compactification of X embedded in $\Gamma(X)$ in the obvious way with the point at infinity corresponding to the empty set. The second section shows that the techniques developed by Narens can be used to obtain a natural characterization of the Vietoris Topology by considering monads of non nearstandard points. The final section uses this same approach to construct a topological analog of the Hausdorff metric for normal spaces.

0. Introduction. Suppose that X is a topological space and that Γ denotes the set of closed subsets of X. It is frequently desirable to endow Γ with a topology of its own. Various topologies on Γ have been proposed and studied by several mathematicians. If X is a metric space, Hausdorff (see [2], [6], [7]) defined a metric on Γ in a natural way. With this metric X is embedded isometrically as a closed subset of Γ by the mapping $x \mapsto \{x\}$. One drawback of this metric, however, is that it depends in an essential way on the metric on X. That is, d and d' may be two metrics for the same topology on X, but induce Hausdorff metrics which do not give the same topology on Γ. In [10] E. Michaels investigates among other topologies the Vietoris or Finite topology on Γ. This topology also has the property that X is embedded as a closed subset of Γ by the mapping $x \mapsto \{x\}$. Both the Hausdorff metric on Γ and the Vietoris topology on Γ make Γ into a compact space if and only if X was originally compact. More recently, L. Narens [12] has introduced an interesting topology on Γ using the techniques of Nonstandard Analysis. This topology always makes Γ a compact set with the empty set $\varnothing \in \Gamma$ acting (see Theorem I.8) somewhat like the point at infinity of the one-point compactification of X.

Nonstandard Analysis provides a particularly nice framework for investigating topological questions. Intuitively, a topological space is a set together with some notion of "nearness"

If X is any set and $^*X$ is a nonstandard extension of X then a topology on X can be described by a relationship of "infinitely close" on some points of $^*X$ (see **[9]**, **[13]**, **[14]**). If X is a topological space, $x \in X$ and $y \in {}^*X$ we say y is *infinitely close* to x, written $y \sim x$ or $y \in \mu(x)$, provided for every standard open set $\mathcal{O}$ if $x \in \mathcal{O}$ then $y \in {}^*\mathcal{O}$. In this case x is called the *standard part* of y, denoted $x = \mathrm{St}(y)$. If A is an internal subset of $^*X$, let $\mathrm{St}(A) = \{x \in X \mid \mu(x) \cap A \neq \emptyset\}$. Under suitable conditions on $^*X$ $\mathrm{St}(A)$ is always closed. Now, if $A, B \in {}^*X$, Narens defines $A \sim B$ provided $\mathrm{St}(A) = \mathrm{St}(B)$. He uses this relationship to define a topology which he calls the *compact topology*. In the present paper we will call this same topology the *N-compact topology*. Although the relationship $\sim$ provides a definition of the N-compact topology, it is important for a full understanding of this topology to obtain a description of the actual monads for this topology (see **[14]**, for an elucidation of this point).

The first part of this paper is devoted to the investigation of the N-compact topology and a closely related topology we call the S-compact topology. With either of these topologies Γ is compact and the one point compactification of X is embedded as a closed subset of Γ by the mapping $x \to \{x\}$ with ∞ corresponding to $\emptyset \in \Gamma$. When X is locally compact, the S-compact and N-compact topologies are identical, both are Hausdorff, and the monad, $\mu(F)$, of a point $F \in \Gamma$ is given by $\mu(F) = \{H \in {}^*\Gamma \mid F \sim H\}$. When X is not locally compact the S-compact and N-compact topologies may be different, neither is Hausdorff and neither monad is given by $\{H \in {}^*\Gamma \mid F \sim H\}$. The S-compact topology has a good standard as well as a good nonstandard characterization.

The technique Narens has developed suggests several different topologies on Γ. In the second section we use this technique to obtain a nice description of the Vietoris topology. This description elucidates some of the properties of the Vietoris topology. In the third section of the paper we define a new topology, called the *fine topology* on Γ. This topology has many nice properties and, in particular, may be regarded in some sense as the analog in the topological category of the Hausdorff metric (see Theorem III.8).

Throughout this paper, X, Y and Z will always denote Hausdorff spaces (although Γ may not be Hausdorff). When we are dealing with several spaces X, Y and Z, their extensions will always be taken in a single nonstandard model $^*\mathcal{M}$. That is, we let $\mathcal{M}$ be the complete higher order structure on $X \cup Y \cup Z$ and let $^*\mathcal{M}$ be a higher order elementary extension of $\mathcal{M}$ (**[8]**, **[9]**, **[13]**). If κ is the cardinality of the universe of $\mathcal{M}$ we will assume throughout that $^*\mathcal{M}$ is at least κ^+ saturated (**[1]**, **[3]**, **[4]**, **[5]**, **[11]**). Thus we will assume GCH to insure that such an extension exists.

I. The N-compact and S-compact topologies. Suppose throughout this section that X is a Hausdorff space and that Γ_X (or Γ where confusion is unlikely) is the set of closed subsets of X. We topologize Γ as follows

I.1. DEFINITION. Suppose $F \in \Gamma$ and $H \in {}^*\Gamma$ we write $H \sim F$ whenever $\mathrm{St}(H) = F$. Then the *premonad*, $m(F)$, of F in ${}^*\Gamma$ is defined by $m(F) = \{H \in {}^*\Gamma \mid H \sim F\}$. A subset $A \subseteq \Gamma$ is said to be *N-open* if and only if for each $F \in A$, $m(F) \subseteq {}^*A$. It is easy to verify that the N-open subsets of Γ form a topology which, modifying the terminology of **[12]**, we call the *N-compact topology*.

If $F \in \Gamma$ the *monad*, $\mu_N(F)$, of F in the N-compact topology is defined in the usual way by

$$\mu_N(F) = \bigcap_{F \in A,\, A\ N\text{-open}} {}^*A$$

It is immediate from these definitions that $m(F) \subseteq \mu_N(F)$ although it turns out (See I.9) that frequently $m(F) \neq \mu_N(F)$.

I.2. LEMMA. *Suppose $\mathcal{O}$ is an open subset of X. Define $\hat{\mathcal{O}} \subseteq \Gamma$ by $\hat{\mathcal{O}} = \{F \mid F \cap \mathcal{O} \neq \varnothing\}$. Then $\hat{\mathcal{O}}$ is open in the N-compact topology.*

Proof. Suppose $F \in \hat{\mathcal{O}}$ and $H \sim F$. Since $F \in \hat{\mathcal{O}}$ there is a point $x \in \mathcal{O} \cap F$. Since $\mathcal{O}$ is open $\mu(x) \subseteq {}^*\mathcal{O}$. Since $\mathrm{St}(H) = F$, $x \in \mathrm{St}(H)$. Hence, $\mu(x) \cap H \neq \varnothing$ and thus ${}^*\mathcal{O} \cap H \neq \varnothing$. Therefore $H \in {}^*\hat{\mathcal{O}}$.

I.3. LEMMA. *Suppose K is a compact subset of X. Define $\check{K} \subseteq \Gamma$ by $\check{K} = \{F \in \Gamma \mid F \cap K = \varnothing\}$. Then K is open in the N-compact topology.*

Proof. Suppose $F \in \check{K}$ and $H \sim F$. If $H \notin \check{K}$ then $H \cap {}^*K \neq \varnothing$. Let $x \in H \cap {}^*K$. Since K is compact $y = \mathrm{St}(x)$ exists and is in K. Since $\mathrm{St}(H) = F$, $y \in F$. Thus $K \cap F \neq \varnothing$ and $F \notin \check{K}$. This contradiction completes the proof.

I.4. PROPOSITION. *Γ with the N-compact topology is compact and T_1.*

Proof.

(i) *If $H \in {}^*\Gamma$, $H \in m(\mathrm{St}(H)) \subseteq \mu_N(\mathrm{St}(H))$. Thus every point of ${}^*\Gamma$ is near-standard and Γ is compact.*

(ii) *Suppose $F, G \in \Gamma$, $F \neq G$. Without loss of generality we may assume $F \backslash G \neq \varnothing$. Choose $a \in F \backslash G$. By Lemma* I.3 *$\check{\{a\}}$ is open and*

$G \in \{\check{a}\}$ but $F \notin \{\check{a}\}$. By Lemma I.2 *$\widehat{X\backslash G}$ is open and $G \notin \widehat{X\backslash G}$ while $F \in X\backslash G$.*

It is natural to ask when Γ with the N-compact topology is Hausdorff. The following proposition shows this occurs precisely when the premonads, $m(F)$, are actually the monads, $\mu_N(F)$.

I.5. Proposition. *The following are equivalent.*
(i) *Γ is Hausdorff.*
(ii) *For each $F \in \Gamma$, $m(F) = \mu_N(F)$.*

Proof. (ii)$\rightarrow$(i) is immediate since $F \neq G$ implies $m(F) \cap m(G) = \varnothing$.

(i)$\rightarrow$(ii). We must show $\mu_N(F) \subseteq m(F)$. Suppose $H \notin m(F)$. Hence $\mathrm{St}(H) \neq F$. Since Γ is Hausdorff $\mu_N(\mathrm{St}(H)) \cap \mu_N(F) = \varnothing$. But $H \in m(\mathrm{St}(H)) \subseteq \mu_N(\mathrm{St}(H))$ and, hence, $H \notin \mu_N(F)$. This completes the proof.

I.6. Theorem. *Suppose X is locally compact then Γ is Hausdorff and, hence, by* I.5 *for each $F \in \Gamma$, $m(F) = \mu_N(F)$.*

Proof. Suppose $A, B \in \Gamma$. Without loss of generality we may assume $A \backslash B \neq \varnothing$. Choose $a \in A \backslash B$. Since X is locally compact there is an open set U such that $a \in U$, $\bar{U}$ is compact and $\bar{U} \cap B = \varnothing$. By Lemma I.2 $\hat{U}$ is open. By Lemma I.3 $\check{U}$ is open. Clearly $\hat{U} \cap \check{U} = \varnothing$. But $A \in \hat{U}$ and $B \in \check{U}$ which completes the proof.

The converse of Theorem I.6 is also true as a consequence of the fact (Theorem I.8) that the one-point compactification of X can be embedded in Γ with the point at infinity corresponding to the empty set.

I.7. Definition. Let X^+ denote the one-point compactification of X with ∞ denoting the point at infinity. We define an embedding $e: X^+ \rightarrow \Gamma$ by

$$e(\infty) = \varnothing$$
$$e(x) = \{x\} \quad \text{if} \quad x \neq \infty.$$

I.8. Theorem. *$e: X^+ \rightarrow \Gamma$ is a homeomorphism of X^+ into Γ.*

Proof. (i) Suppose $x \in X$. If $y \in \mu(x)$ then $\{y\} \sim \{x\}$ so ${}^*e(y) \in m(e(x))$. Hence $e(\mu(x)) \subseteq m(e(x)) \subseteq \mu_N(e(x))$. Now, suppose $y \notin \mu(x)$ then there is a standard open set $\mathcal{O}$ such that $x \in \mathcal{O}$ and

$y \notin {}^*\mathcal{O}$. By Lemma I.2 $\hat{\mathcal{O}}$ is open in Γ and $e(x) \in \hat{\mathcal{O}}$ but ${}^*e(y) \notin \hat{\mathcal{O}}$. Hence, ${}^*e(y) \notin \mu_N(e(x))$. Thus $e(\mu(x)) = \mu_N(e(x))$.

(ii) We must show $\mu_N(\varnothing) \cap {}^*e(X^+) = A$ where A is given by $A = \{\varnothing\} \cup \{\{x\} \mid x \in \mu(\infty)\}$. Notice that $\mu(\infty) = \bigcap_{K \text{ compact}} {}^*(X - K) \cup \{\infty\}$. By Lemma I.3 if K is compact $\check{K}$ is open. But $\check{K} \cap e(X^+) = \varnothing \cup e(X - K)$. Hence $\mu_N(\varnothing) \cap {}^*e(X^+) \subseteq A$. Now suppose $\{y\} \notin \mu_N(\varnothing)$. Hence there is an open set $\mathcal{O} \subseteq \Gamma$ s.t. $\varnothing \in \mathcal{O}$ but $\{y\} \notin {}^*\mathcal{O}$. Let $K = \{x \in X \mid \{x\} \notin \mathcal{O}\}$. Notice if $x \in {}^*K$ then $\{x\} \notin {}^*\mathcal{O}$, so $\{x\} \notin m(\varnothing)$ and hence x must be near-standard. Now, if $\{\mathrm{st}(x)\} \in \mathcal{O}$ then $\{x\}$ would also be in ${}^*\mathcal{O}$ so $\{\mathrm{st}(x)\} \notin \mathcal{O}$ and $\mathrm{St}(x) \in K$. Thus, we've shown for each $x \in {}^*K$, $\mathrm{St}(x)$ exists and is in K, so K is compact. Now $y \in {}^*K$, so $y \notin \mu(\infty)$. Hence $A \subseteq \mu_N(\varnothing) \cap e(X^+)$ which completes the proof.

I.9. Corollary. *The following are equivalent*

(i) *X is locally compact*

(ii) *Γ is Hausdorff*

(iii) *For each $F \in \Gamma$, $m(F) = \mu_N(F)$.*

Proof. Immediate from I.5, I.6 and I.8 since X^+ is Hausdorff if and only if X is locally compact.

We would like to obtain a standard description of the compact topology on Γ. Lemmas I.2 and I.3 suggest a topology which is analogous to Vietoris topology. This approach is developed in the following pages. For locally compact spaces the two topologies are identical. However, for more general spaces they may be distinct (see Example I.16).

I.10. Definition. A subset $\mathcal{O}$ of X is said to be *cocompact* whenever $X \backslash \mathcal{O}$ is compact. Suppose that $\mathcal{O}$ is cocompact and that $U_1, U_2, \cdots, U_n$ are open subsets of X. Then let $\langle \mathcal{O}, U_1, U_2, \cdots, U_n \rangle$ denote the set

$$\{F \in \Gamma \mid F \subseteq \mathcal{O} \quad \text{and for} \quad i = 1, 2, \cdots, n \quad F \cap U_i \neq \varnothing\}.$$

By Lemmas I.2 and I.3 the set $\langle \mathcal{O}, U_1, U_2, \cdots, U_n \rangle$ is open in the N-compact topology. Let $\mathcal{B}$ denote the set of all such $\langle \mathcal{O}, U_1, U_2, \cdots, U_n \rangle$. Notice the intersection of two sets in $\mathcal{B}$ is again in $\mathcal{B}$

$$\begin{aligned} &\langle \mathcal{O}, U_1, U_2, \cdots, U_n \rangle \cap \langle \mathcal{O}', V_1, V_2, \cdots, V_k \rangle \\ &\quad = \langle \mathcal{O} \cap \mathcal{O}', U_1, U_2, \cdots, U_n, V_1, V_2, \cdots, V_k \rangle. \end{aligned}$$

So, $\mathscr{B}$ forms a basis for a topology on Γ. This topology is called the *S-compact topology.* By the above remarks every set open in S-compact topology is also open in the N-compact topology. Hence, the S-compact topology is compact. Furthermore, an examination of the proof of Theorem I.8 shows that $e: X^+ \to \Gamma$ is a homeomorphism of X^+ *into* Γ with the S-compact topology.

One of the basic lemmas in Nonstandard Topology is that if X is a topological space, $x \in X$ and $\mu(x)$ its monad then there is a *open set U such that $^*x \in U \subseteq \mu(x)$. For the N-compact topology we have been working primarily with the premonad $m(F)$ rather than the actual monad $\mu_N(F)$ of F in the N-compact topology. It is not true that for this premonad there is always a *open set U in the N-compact topology such that $^*F \in U \subseteq m(F)$. However, when X is locally compact $m(F) = \mu_N(F)$ and such a U can always be found. In particular the following lemma shows such a U can be found in $^*\mathscr{B}$.

I.11. LEMMA. *Suppose X is locally compact and $F \in \Gamma$. Then there is a set $U \in {}^*\mathscr{B}$ such that $F \in U \subseteq m(F)$.*

Proof. (i) For each $x \notin F$ let W_x be in an open set such that $x \in W_x$, $\bar{W}_x$ is compact and $\bar{W}_x \cap F = \varnothing$. By a straightforward enlargements argument there is *compact set K such that $K \cap F = \varnothing$ and for each $x \notin F$, $\bar{W}_k \subseteq K$. Let $\mathscr{O} = X - K$.

(ii) For each $x \in F$ choose a *open set U_x such that $x \in U_x \subseteq \mu(x)$. By a straightforward saturation argument there is an internal *finite collection of *open sets $\{V_1, V_2, \cdots, V_\nu\}$ such that for each $i = 1, 2, \cdots, \nu$, $V_i \cap {}^*F \neq \varnothing$ and for each $x \in F$, $U_x \in \{V_1, V_2, \cdots, V_\nu\}$.

(iii) Let $U = \langle \mathscr{O}, V_1, V_2, \cdots, V_\nu \rangle \in {}^*\mathscr{B}$. It is straightforward to verify that $^*F \in U \subseteq m(F)$.

I.12. COROLLARY. *Suppose X is locally compact. Then $\mathscr{B}$ is a basis for the N-compact topology on Γ. Hence, the S-compact topology and N-compact topology are identical.*

Proof. Immediate from Lemma I.11.

Example I.16 will show that the S-compact and N-compact topologies may be distinct when X is not locally compact. In view of this fact if $F \in \Gamma$ we denote its monad in the S-compact topology by $\mu_S(F)$. Notice, $\mu_N(F) \subseteq \mu_S(F)$. In order to obtain a characterization of $\mu_S(F)$ we need a definition.

I.13. DEFINITION. Suppose X is a topological space and $x \in {}^*X$. x is said to be a *far point* provided for every standard compact subset K of X, $x \notin {}^*K$. Let $\mathrm{FAR}(X) = \{x \in {}^*X \mid x \text{ is far}\}$, when X is locally compact, the far points of ${}^*X$ are precisely the nonnearstandard points.

I.14. PROPOSITION. *Suppose* $F \in \Gamma$ *and* $H \in {}^*\Gamma$ *then* $H \in \mu_S(F)$ *if and only if*

(i) $F \subseteq \mathrm{St}(H)$ *and*

(ii) *For every* $x \in H$ *either* $x \in \mathrm{FAR}(X)$ *or* $\mathrm{St}(x)$ *exists and* $\mathrm{St}(x) \in F$.

Proof. $(\rightarrow)$ (i) Suppose $x \in F$ and U is a standard open set with $x \in U$. Therefore $F \in \langle X, U\rangle$ so $H \in {}^*\langle X, U\rangle$ and $H \cap {}^*U \neq \varnothing$. Hence, by a straighforward saturation argument $H \cap \mu(x) \neq \varnothing$. So $x \in \mathrm{St}(H)$.

(ii) Suppose $x \in H$ and $x \notin \mathrm{FAR}(X)$. Therefore there is a standard compact set K with $x \in {}^*K$. Thus, $\mathrm{St}(x)$ exists and $\mathrm{St}(x) \in K$.

Now suppose $\mathrm{St}(x) \notin F$. Since K is compact there is an open subset U of K such that $\mathrm{St}(x) \in U$ and $\bar{U} \cap F = \varnothing$. $\bar{U}$ is compact since it is a closed subset of K and $x \in {}^*\bar{U}$ since $\mathrm{St}(x) \in U$. Since $\bar{U} \cap F = \varnothing$, $F \in \langle X \setminus \bar{U}\rangle$. But ${}^*\bar{U} \cap H \neq \varnothing$, so $H \notin {}^*\langle X \setminus \bar{U}\rangle$ contradicting $H \in \mu_S(F)$.

$(\leftarrow)$ Suppose $F \in \langle \mathcal{O}, U_1, \cdots, U_n\rangle$. First, suppose $H \not\subseteq {}^*\mathcal{O}$, then $H \cap {}^*(X \setminus \mathcal{O}) \neq \varnothing$. So $t \in H \cap {}^*(X \setminus \mathcal{O})$, $\mathrm{St}(t) \in X \setminus \mathcal{O}$; so $F \not\subseteq {}^*\mathcal{O}$, contradicting $F \in \langle \mathcal{O}, U_1, \cdots, U_n\rangle$. Now, $F \cap U_i \neq \varnothing$. Therefore there is an $x \in F \cap U_i$. But $\mu(x) \subseteq {}^*U_i$ and by (i) there is a $y \in \mu(x) \cap H$. So $H \cap {}^*U_i \neq \varnothing$. This completes the proof.

I.15. COROLLARY.

(a) *Suppose* $F \in \Gamma$ *and* $H \in \mu_N(F)$ *then*

(i) $F \subseteq \mathrm{St}(H)$ *and*

(ii) *For every* $x \in H$ *either* $x \in \mathrm{FAR}(X)$ *or* $\mathrm{St}(x)$ *exists and* $\mathrm{St}(x) \in F$.

(b) *The mapping* $u: \Gamma \times \Gamma \rightarrow \Gamma$ *defined by* $u(H, F) = H \cup F$ *is continuous in the* S*-compact topology.*

Proof.

(a) $\mu_N(F) \subseteq \mu_S(F)$.

(b) For $(H, F) \in \Gamma \times \Gamma$, $\mu(H, F) = \mu_S(H) \times \mu_S(F)$ and clearly $H' \in \mu_S(H)$, $F' \in \mu_S(F)$ implies $H' \cup F' \in \mu_S(H \cup F)$.

By Corollary I.12 if X is locally compact u is continuous in the N-compact topology. However, without this assumption the author does not know whether u is continuous in the N-compact topology. Clearly, if $H' \in m(H)$ and $F' \in m(F)$ then $H' \cup F' \in m(H \cup F)$. However, although this provides some evidence for the continuity of u, it is not by itself sufficient to prove u is continuous.

The following example shows that the N-compact and S-compact topologies are distinct.

I.16. EXAMPLE. We first state carefully two facts necessary for this example.

(i) If $A \subseteq \Gamma$ then A is closed in the N-compact topology if and only if for each $F \in {}^*A$, $\mathrm{St}(F) \in A$. This equivalence is an immediate consequence of Definition I.1.

(ii) If $A \subseteq \Gamma$ then A is closed in the S-compact topology if and only if for each $F \in {}^*A$ and each $H \in \Gamma$, $F \in \mu_S(H)$ implies $H \in A$. Notice that since Γ is not Hausdorff there may be many $H \in A$ such that $F \in \mu_S(H)$. This equivalence is an immediate consequence of Definition I.10.

Now let Q denote the rationals and let G denote the set of closed subgroups of Q. We claim G is closed in the N-compact topology but not in the S-compact topology. The first assertion was proved by Narens in **[12]** by means of (i) above and the observation that if F is a *closed *subgroup of a topological group then so is $\mathrm{St}(F)$. We proceed to the second assertion.

By a straightforward enlargement argument there is an $\alpha \in {}^*Q$ such that

(i) $\alpha \in \mu(1)$

(ii) For each standard integer n

$$n\alpha \in \mathrm{FAR}(Q).$$

Now let $H = \{n\alpha \mid n \in {}^*Z\}$, where Z denotes the set of integers. H is clearly a closed subgroup of ${}^*Q$. By Proposition I.14 $H \in \mu_s(\{1\})$. But $\{1\}$ is not a subgroup of Q. This completes the proof.

Although the N-compact and S-compact topologies on Γ have some very nice properties they also lack some desirable properties. In particular certain constructions on Γ which one might like to be continuous are not continuous with these topologies. We close this section with several such examples before going on to discuss other topologies on Γ in the remainder of this paper.

I.17. EXAMPLES. (i) Suppose $f\colon X \to Y$ is a continuous map. f induces a map $\bar{f}\colon \Gamma_X \to \Gamma_Y$ defined by $\bar{f}(A) = \overline{f(A)}$. One might desire that $\bar{f}$ be continuous. However, this need not be so. In particular, if $\bar{f}$ were continuous this would imply that f had a continuous extension to $g\colon X^+ \to Y^+$ with $g(\infty) = \infty$. No such extension exists, for example, if $f\colon (0,1) \to R$ is the usual inclusion of the open unit interval into the real line.

(ii) Suppose again that $f\colon X \to Y$ is continuous. f induces a map

$\hat{f}: \Gamma_Y \to \Gamma_X$ defined by $f(A) = f^{-1}(A)$. Again one might hope that $\hat{f}$ would be continuous. However, this need not be so. In particular if f is a bijection then the continuity of $\hat{f}$ would imply f is a homeomorphism which is, in general, false.

(iii) Let $X = R$ and define $f: \Gamma \to \Gamma$ by

$$f(A) = \{x \in R \mid a \in A,\ a \leqq x\}$$

f is not continuous since if α is any negative infinite nonstandard real $\{\alpha\} \in \mu_N(\varnothing)$ but $f(\{\alpha\}) = [\alpha, \infty) \notin \mu_S(\varnothing) = \mu_S(f(\varnothing))$.

II. The Vietoris topology. Some of the difficulties noted at the end of §I result from the fact that knowing $H \in \mu_N(F)$ gives us little or no information about the non-nearstandard points in H and $^*F$. In order to obtain a topology on Γ which takes these points into consideration we need a notion of monads for points which are not nearstandard. One such notion is the coarse "monad system" defined in **[14]** and **[15]**. We suppose throughout this section that X is a T_1 space.

II.1. Definition. Suppose $x \in {}^*X$ the *coarse monad* of x, denoted $c(x)$ is defined by

$$c(x) = \bigcap_{\substack{x \in {}^*F \subseteq {}^*U,\ F \text{ standard closed,} \\ U \text{ standard open}}} {}^*U.$$

We collect some results about $c(x)$ in the following proposition.

II.2. Proposition.

(i) *If $x \in X$, $c(x) = \mu(x)$*

(ii) *If $x, y \in {}^*X$, $x \in c(y) \leftrightarrow y \in c(x)$*

(iii) *X is regular $\leftrightarrow$ for every nearstandard x, $c(x) = \mu(\mathrm{St}(x))$.*

(iv) *X is normal $\leftrightarrow$ if $x, y \in {}^*X$ either $c(x) = c(y)$ or $c(x) \cap c(y) = \varnothing$.*

(v) *If $f: X \to Y$ is continuous then for every $x \in {}^*X$ ${}^*f(c(x)) \subseteq c({}^*f(x))$.*

Proof.

(i) Clear.

(ii) If $x \notin c(y)$ then there is a standard closed set F and a standard open set U such that $y \in {}^*F$ but $x \notin {}^*U$. But then $x \in {}^*(X \setminus U)$ and $y \notin {}^*(X \setminus F)$ so $y \notin c(x)$.

(iii) and (iv) see **[14]**.

(v) Clear.

II.3. Definition. Suppose $F \in \Gamma$ and $H \in {}^*\Gamma$. We say $H \underset{c}{\sim} F$ whenever for every $x \in {}^*F$ there is a $y \in H$ such that $y \in \bar{c}(x)$ and for every $y \in H$ there is an $x \in {}^*F$ s.t. $y \in \bar{c}(x)$. The c-monad of F, $\bar{c}(F)$ is defined by $\bar{c}(F) = \{H \in {}^*\Gamma \mid H \underset{c}{\sim} F\}$. Notice, in particular, that $\bar{c}(\emptyset) = \{\emptyset\}$.

II.4. Theorem. *Suppose $F \in \Gamma$, then $\bar{c}(F)$ is the monad of F in the Vietoris topology.*

Proof. First recall that a basis for the Vietoris topology on Γ is given by sets of the form $\langle \mathcal{O}, V_1, \cdots, V_n \rangle = \{F \in \Gamma \mid F \subseteq \mathcal{O}$ and for $i = 1, 2, \cdots, n$ $F \cap V_i \neq \emptyset\}$ where $\mathcal{O}, V_1, \cdots, V_n$ are open subsets of X.

(i) Suppose $F \in \langle \mathcal{O}, V_1, \cdots, V_n \rangle$ and $H \in \bar{c}(F)$. For each $y \in H$ there is an $x \in {}^*F$ such that $y \in c(x)$. Since F is closed and $F \subseteq \mathcal{O}$ $c(x) \subseteq {}^*\mathcal{O}$. Hence $H \subseteq {}^*\mathcal{O}$. Since $F \in \langle \mathcal{O}, V_1, \cdots, V_n \rangle$ there is an $x \in F \cap V_i$ for each i. Since $\{x\}$ is closed $c(x) \subseteq {}^*V_i$ and since $H \underset{c}{\sim} F$ there is a $y \in H \cap c(x)$. Hence $H \cap {}^*V_i \neq \emptyset$. Thus $H \subseteq {}^*\langle \mathcal{O}, V_1, \cdots, V_n \rangle$.

(ii) Suppose $H \notin \bar{c}(F)$. There are two possibilities.

(a) For some $x \in {}^*F$ there is no $y \in H$ such that $y \in c(x)$. Hence, by a straightforward saturation argument there is a standard open U and closed T with $x \in {}^*T \subseteq {}^*U$ and ${}^*U \cap H = \emptyset$. But then $F \in \langle X, U \rangle$ and $H \notin {}^*\langle X, U \rangle$.

(b) For some $y \in H$ there is no $x \in {}^*F$ such that $x \in c(y)$. Hence by a straightforward saturation argument there is a standard open U and closed T with $y \in {}^*T \subseteq {}^*U$ and $U \cap F = \emptyset$. But then $F \in \langle X \backslash T \rangle$ and $H \notin {}^*\langle X \backslash T \rangle$.

Thus, in either case $H \notin \bar{c}(F)$ implies $H \notin$ monad of F in the Vietoris topology, completing the proof.

Notice, that the mapping $X \to \Gamma$ defined by $x \to \{x\}$ is a homeomorphism into using the Vietoris topology but $e: X^+ \to \Gamma$ is not even continuous (unless X is compact) since $\emptyset$ is an isolated point of Γ with the Vietoris topology. In addition it is clear that the mapping $u: \Gamma \times \Gamma \to \Gamma$ defined by $u(F, H) = F \cup H$ is continuous with the Vietoris topology.

With the Vietoris topology the mapping $\bar{f}$ defined in Example I.17 will be continuous if the range is normal. To see this we first observe that Definition II.3 can be extended to the full *power set of X, denoted ${}^*P(X)$.

II.5. Definition. Suppose $A, B \in {}^*P(X)$. We say $A \underset{c}{\sim} B$ whenever for each $a \in A$ there is a $b \in B$ such that $a \in c(b)$ and for each $b \in B$ there is an $a \in A$ such that $a \in c(b)$.

II.6. LEMMA. *Suppose $A \in {}^*P(X)$ and $\bar{A}$ is the *closure of A then $A \underset{c}{\sim} \bar{A}$.*

Proof. Since $A \subseteq \bar{A}$ we need only show that for each $x \in \bar{A}$ there is an $a \in A$ such that $a \in c(x)$. Suppose $x \in {}^*F \subseteq {}^*U$. F standard closed and U standard open. Since ${}^*U$ is *open and $x \in \bar{A}$, $A \cap {}^*U \neq \emptyset$. Hence, by a straightforward saturation argument $A \cap c(x) \neq \emptyset$.

II.7. PROPOSITION. *Suppose f: $X \to Y$ is continuous, Y is normal and $\bar{f}$: $\Gamma_X \to \Gamma_Y$ is defined by $\bar{f}(F) = \overline{f(F)}$. Then $\bar{f}$ is continuous in the Vietoris topology.*

Proof. Since Y is normal it is easy to see using Proposition II.2 (iv) that $\underset{c}{\sim}$ is transitive on ${}^*P(Y)$. From Proposition II.2 (v) it is clear that $F \underset{c}{\sim} H$ implies $f(F) \underset{c}{\sim} f(H)$ but $\bar{f}(F) \underset{c}{\sim} f(F)$ and $\bar{f}(H) \underset{c}{\sim} f(H)$ by the preceding Lemma and, hence by transitivity $\bar{f}(F) \underset{c}{\sim} \bar{f}(H)$.

Although the coarse monad system imposes some control on the non-nearstandard points of X, this is not a very natural monad system, and, in fact, the coarse monads are much too large. As one result the Vietoris Topology has the following interesting property.

II.8. MONOTONE LIMIT THEOREM. Suppose $F_1 \subseteq F_2 \subseteq \cdots$ is an ascending sequence in Γ. Let $F = \overline{\cup F_k}$. Then in the Vietoris topology $\mathrm{Lim}_{k \to \infty} F_k = F$.

Proof. Suppose $F \in \langle \mathcal{O}, U_1, \cdots, U_n \rangle$

(i) $F \subseteq \mathcal{O}$ implies $F_k \subseteq \mathcal{O}$ for each k

(ii) For each i, $F \cap U_i \neq \emptyset$. Hence there is an $x \in F \cap U_i$ and since U_i is open $(\cup F_k) \cap U_i \neq \emptyset$. Therefore for some k_i, $F_{k_i} \cap U_i \neq \emptyset$. Let $K = \max(k_1, \cdots, k_n)$.

Now for each $k \geqq K$, $F_k \in \langle \mathcal{O}, U_1, \cdots, U_n \rangle$.

From our point of view some insight into this theorem can be obtained from the following example.

II.9. EXAMPLE. Suppose $x \in {}^*R$ is an infinite positive nonstandard real and A is a standard set with $x \in {}^*A$. Then $c(x) \cap {}^*A$ contains arbitrarily small and large infinite numbers i.e. for each infinite positive $y \in {}^*R$, $c(x) \cap {}^*A \cap (0, y) \neq \emptyset$ and $c(x) \cap {}^*A \cap (-y, \infty) \neq \emptyset$. The proof of this is a straightforward saturation argument.

In the next section we consider a monad system on X which gives more control over the non-nearstandard points.

III. The fine topology. In **[14]**, **[15]** we obtained a very natural monad system for non-nearstandard points. This monad system enables us to define a topology on Γ which provides a great deal of control around non-nearstandard points. Throughout this section we will restrict our attention to normal spaces.

III.1. DEFINITION.

(i) Suppose U is a *open subset of $^*X$ and F is a *closed subset of $^*X$ with $F \subseteq U$. The pair (U, F) is said to be a *quasi-standard pair, Q.S.P.* provided there is a standard locally finite collection $\mathcal{U} = \{(U_\alpha, F_\alpha)\}_{\alpha \in \mathcal{I}}$ of pairs (U_α, F_α) such that each U_α is open, each F_α is closed, $F_\alpha \subseteq U_\alpha$ and $(U, F) \in {}^*\mathcal{U}$ (See **[14]**, **[15]**).

(ii) If $x \in {}^*X$ we define the *monad of* x, $\mu(x)$ by

$$\mu(x) = \bigcap_{x \in F \subseteq U, (U.F) a Q.S.P.} U.$$

We recall the following facts about this monad system from **[14]**, **[15]**. Our assumption that X is normal is important here.

III.2. PROPOSITIONS.

(i) *If x is standard, $\mu(x)$ is the usual monad* $\mu(x) = \bigcap_{x \in U,\, U \text{ standard open}} {}^*U$

(ii) *If x is nearstandard* $\mu(x) = \mu(\mathrm{St}(x))$

(iii) *For each* $x, y \in {}^*X$ *either* $\mu(x) = \mu(y)$ *or* $\mu(x) \cap \mu(y) = \emptyset$.

(iv) *If* $f: X \to Y$ *is continuous then for every* $x \in {}^*X$, ${}^*f(\mu(x)) \subseteq \mu({}^*f(x))$.

Proof. **[14]**, **[15]**.

III.3. DEFINITION. Suppose $H, F \in {}^*\Gamma$ we say $H \underset{\mu}{\sim} F$ whenever for every $x \in H$, $\mu(x) \cap F \neq \emptyset$ and for every $y \in F$, $\mu(y) \cap H \neq \emptyset$. In view of Proposition III.2 $\underset{\mu}{\sim}$ is an equivalence relation. For each $F \in {}^*\Gamma$, $\bar{\mu}(F)$ is given by $\bar{\mu}(F) = \{H \in {}^*\Gamma \mid H \underset{\mu}{\sim} F\}$. The topology defined by the $\bar{\mu}$ monads is called the *fine topology* on Γ.

We can obtain a standard characterization of the fine topology in the obvious way.

III.4. DEFINITION. Suppose $\mathcal{U} = \{U_\alpha\}_{\alpha \in \mathcal{I}}$ is a locally finite collection of open subsets of X and $\mathcal{O}$ is an open subset of X. Then define $\langle \mathcal{O}, \mathcal{U} \rangle = \{F \in \Gamma \mid F \subseteq \mathcal{O}$ and for each $U_\alpha \in \mathcal{U}$, $U_\alpha \cap F \neq \emptyset\}$. Let $\mathcal{F} = \{\langle \mathcal{O}, \mathcal{U} \rangle \mid \mathcal{O} \subseteq X$ open, $\mathcal{U} = \{U_\alpha\}_{\alpha \in \mathcal{I}}$ a locally finite family of open sets in $X\}$.

III.5. THEOREM. *$\mathcal{F}$ is a basis for the fine topology on* Γ.

Proof.

(i) Suppose $F \in \Gamma$, $F \in \langle \mathcal{O}, \mathcal{U} \rangle$ and $H \in \bar{\nu}(F)$. We must show that $H \in {}^*\langle \mathcal{O}, \mathcal{U} \rangle$.

(a) $F \in \langle \mathcal{O}, \mathcal{U} \rangle$ implies $F \subseteq \mathcal{O}$. Hence for each $x \in F$, $\mu(x) \subseteq {}^*\mathcal{O}$. So, clearly $H \subseteq {}^*\mathcal{O}$.

(b) For each $U_\alpha \in \mathcal{U}$ choose $x_\alpha \in U_\alpha \cap F$. Given $U_\beta \in {}^*\mathcal{U}$ $(U_\beta, \{x_\beta\})$ is a Q.S.P. so $\mu(x_\beta) \subseteq U_\beta$ and since $\mu(x_\beta) \cap H \neq \emptyset$, $U_\beta \cap H \neq \emptyset$.

(ii) Suppose $F \in \Gamma$, and $H \notin \bar{\mu}(F)$. We must find an $\mathcal{O}$, $\mathcal{U}$ such that $F \in \langle \mathcal{O}, \mathcal{U} \rangle$ but $H \notin {}^*\langle \mathcal{O}, \mathcal{U} \rangle$.

(a) Suppose for some $x \in {}^*F$, $\mu(x) \cap H = \emptyset$. Then by a straightforward saturation argument there is a locally finite collection of pairs $\mathcal{U} = \{(U_\alpha, F_\alpha)\}_{\alpha \in \mathcal{I}}$ containing a pair $(U_\beta, F_\beta) \in {}^*\mathcal{U}$ such that $x \in F_\beta$ and $U_\beta \cap H = \emptyset$. Let $\mathcal{V} = \{U_\alpha \mid F \cap U_\alpha \neq \emptyset\}$. Then $F \in \langle X, \mathcal{V} \rangle$ but $H \notin \langle X, \mathcal{V} \rangle$ since $U_\beta \in {}^*\mathcal{V}$.

(b) Suppose for some $y \in H$, $\mu(y) \cap F = \emptyset$. By a straightforward enlargement argument there is a locally finite collection of pairs $\mathcal{U} = \{(U_\alpha, F_\alpha)\}_{\alpha \in \mathcal{I}}$ such that for some β, $y \in F_\beta$ but $U_\beta \cap {}^*F = \emptyset$. Let $T = \bigcup_{\alpha \in \mathcal{I}, F_\alpha \cap F = \emptyset} F_\alpha$. Since $\mathcal{U}$ is locally finite, T is closed. But $F \in \langle X \backslash T, \emptyset \rangle$ and $H \notin {}^*\langle X \backslash T, \emptyset \rangle$.

This completes the proof.

III.6. Proposition. *Suppose $f: X \to Y$ is continuous and $\bar{f}: \Gamma_X \to \Gamma_Y$ is defined by $\bar{f}(F) = \overline{f(F)}$. Then $\bar{f}$ is continuous in the fine topology.*

Proof. Entirely analogous to that of Proposition II.7.

Notice that as with the Vietoris topology the mapping $x \to \{x\}$ is a homeomorphism into of X into Γ but $e: X^+ \to \Gamma$ is not (unless X is compact) since $\emptyset$ is an isolated point of Γ. In addition the mapping $u: \Gamma \times \Gamma \to \Gamma$ given by $u(F, H) = F \cup H$ is easily seen to be continuous in the fine topology. It is easy to see that the Monotone Limit Theorem is false in the fine topology. In fact a counterexample is provided by the sequence $F_n = [-n, n]$ of subsets of $\mathbf{R}$.

In general, the compact topology is coarser than the Vietoris topology which in turn is coarser than the fine topology. Of course, when X is compact all three topologies are identical.

For metric spaces the Hausdorff metric provides a very natural topology on Γ. One difficulty with this topology, however, is that it dependes in an essential way on the metric on X. Recall the definition of the Hausdorff metric.

III.7. Definition. Suppose X is a metric space with metric d.

(i) If $x \in X$ and $F \in \Gamma$ define

$$\rho'(x, F) = \inf_{y \in F} d(x, y) \quad \text{if} \quad F \neq \emptyset$$
$$= \infty \qquad \text{if} \quad F = \emptyset$$

(ii) If $A, B \in \Gamma$ define

$$\rho'(A, B) = \max\left(\sup_{a \in A} \rho'(a, B), \sup_{b \in B} \rho'(b, A)\right).$$

Finally we define the *Hausdorff metric* $\rho_d(A, B)$ by $\rho_d(A, B) = \min(1, \rho'(A, B))$. (See, for example **[2]** or **[6]** and **[7]**).

Since this topology depends in an essential way on the given metric d, it is natural to look for a topology related to the Hausdorff metric which is independent of the metric d.

III.8. DEFINITION. Suppose X is a metrizable space. If d is a metric on X let τ_d denote the set of open subsets of Γ with the metric ρ_d. Let $\mathcal{B} = \bigcup_{d \text{ a metric on } X} \tau_d$. $\mathcal{B}$ is clearly a basis for a topology on Γ.

III.8. THEOREM. *Suppose X is metrizable space. Then $\mathcal{B}$ is a basis for the fine topology on Γ.*

Proof.

(i) Suppose $F \in A \in \tau_d$. We must show $\bar{\mu}(F) \subseteq {}^*A$. Since $A \in \tau_d$ there is a standard $\epsilon > 0$ such that for every $H \in {}^*\Gamma$, ${}^*\rho_d(H, {}^*F) < \epsilon$ implies $H \in {}^*A$. But, now by [**14**, Theorem 2.12] for every $x, y \in {}^*X$, $y \in \mu(x)$ implies ${}^*d(x, y) \sim 0$. Hence, $H \in \bar{\mu}(F)$ implies ${}^*\rho_d(H, {}^*F) \sim 0$; so $H \in {}^*A$. Thus $\bar{\mu}(F) \subseteq \bigcap_{A \in \mathcal{B}} {}^*A$.

(ii) Suppose $H \notin \bar{\mu}(F)$ we must find some $A \in \mathcal{B}$ such that $F \in A$ but $H \notin {}^*A$. Since $H \notin \bar{\mu}(F)$ there is an open set $\mathcal{O}$ and a locally finite collection $\mathcal{U} = \{U_\alpha\}_{\alpha \in \mathcal{I}}$ of open sets such that $F \in \langle \mathcal{O}, \mathcal{U} \rangle$ but $H \notin {}^*\langle \mathcal{O}, \mathcal{U} \rangle$. There are two cases.

(a) $H \not\subseteq {}^*\mathcal{O}$. Since F is closed and $F \subseteq \mathcal{O}$ there is a continuous function $\sigma: X \to [0, 1]$ such that $\sigma(x) = 0$ for $x \in F$ and $\sigma(x) = 1$ for $x \notin \mathcal{O}$. Let d be any metric on X. Define a new metric δ on X by $\delta(x, y) = d(x, y) + |\sigma(x) - \sigma(y)|$. It is straightforward to verify that δ is a metric on X. But if $z \in H \backslash {}^*\mathcal{O}$ and $x \in {}^*F$, $\delta(x, y) \geqq 1$; so $\rho_\delta{}^*(H, {}^*F) \geqq 1$ and, thus, $H \notin {}^*\{T \mid \rho_\delta(T, F) < 1\}$ which is in τ_δ and, hence, $\mathcal{B}$.

(b) For some $\beta \in {}^*\mathcal{I}$, $H \cap U_\beta = \emptyset$. For each α choose $x_\alpha \in U_\alpha \cap F$ and choose a continuous function $\sigma_\alpha: X \to [0, 1]$ so that $\sigma_\alpha(x_\alpha) = 1$ and $\sigma_\alpha(y) = 0$ if $y \notin U_\alpha$. Let d be any metric on X and define a new metric δ on X by $\delta(x, y) = d(x, y) + \max_\alpha |\sigma_\alpha(x) - \sigma_\alpha(y)|$.

Notice for $y \in H$, $^*\delta(x_\alpha, y) \geqq 1$ so $^*\rho_\delta(^*F, H) \geqq 1$. Therefore, we need only verify that δ is a metric on X. But this is a straightforward verification after noticing that for given x and y there are neighborhoods U of x and V of y in which only finitely many σ_α's are nonzero and, thus, $\max_\alpha |\sigma_\alpha(x) - \sigma_\alpha(y)|$ is continuous.

In view of Theorem III.8 the fine topology may be regarded in some sense as the analog of the Hausdorff metric in the topological category.

References

1. J. L. Bell and A. B. Slomson, *Models and Ultraproducts: an Introduction*, North-Holland, Amsterdam, 1969.
2. F. Hausdorff, *Mengenlehre*, Dover, New York. (Translation) *Set Theory*, Chelsea, New York, 1957.
3. H. J. Keisler, *Good ideals in fields of sets*, Ann. of Math., (2) **79** (1964), 338–359.
4. ———, *Ultraproducts and elementary classes*, Nederl. Akad. Wetensch. Proc. Ser. A 64 = Indag. Math., **23** (1961), 477–495.
5. ———, *Ultraproducts and saturated models*, Nederl. Akad. Wetensch. Proc. Ser. A 67 = Indag. Math., **26** (1964), 178–186.
6. K. Kuratowski, *Topology*, Vol. I, Academic Press, New York, 1966.
7. ———, *Topology*, Vol. II, Academic Press, New York, 1968.
8. W. A. J. Luxemberg, *A new approach to the theory of monads*, California Institute of Technology, Pasadena, Calif., 1967.
9. M. Machover and J. Hirschfeld, *Lectures on nonstandard analysis*, Lecture Notes in Math., Vol. 94, Springer-Verlag, New York, 1969.
10. E. Michael, *Topologies on spaces of subsets*, Trans. Amer. Math. Soc., **71** (1951), 152-182.
11. M. Morley and R. Vaught, *Homogeneous universal models*, Math. Scand., **11** (1962), 37–57.
12. L. Narens, *Topologies of closed subsets*, Trans. Amer. Math. Soc., **174** (1972), 55–76.
13. A. Robinson, *Non-standard Analysis*, North-Holland, Amsterdam (1974).
14. F. Wattenberg, *Nonstandard topology and extensions of monad systems to infinite points*, J. Symbolic Logic, **36** (1971), 463–476.
15. ———, *Monads of infinite points and finite product spaces*, Trans. Amer. Math. Soc., **176** (1973), 351–368.

Received May 25, 1976.

UNIVERSITY OF MASSACHUSETTS
AMHERST, MA 01002

PACIFIC JOURNAL OF MATHEMATICS
Vol 68, No 2, 1977

INTEGRAL REPRESENTATION OF TCHEBYCHEFF SYSTEMS

R. A. ZALIK

A representation of Tchebycheff systems in terms of iterated Riemann–Stieltjes integrals, is given.

1. Introduction. A system of real-valued functions $\{u_0, u_1, \cdots, u_n\}$ defined on a totally ordered set is called a Tchebycheff system or T-system (Weak Tchebycheff system or WT-system), provided that for every choice of points $t_0 < t_1 < \cdots < t_n$ of the set,

$$D(u_0, u_1, \cdots, u_n / t_0, t_1, \cdots, t_n) = \det(u_i(t_j);\ i, j = 0, 1, \cdots, n) \tag{1}$$

is strictly positive (nonnegative). A function u is said to be convex with respect to the system $\{u_0, u_1, \cdots, u_n\}$, if $\{u_0, u_1, \cdots, u_n, u\}$ is a WT-system. The set of functions convex with respect to $\{u_0, u_1, \cdots, u_n\}$ is evidently a cone. This cone is referred to as "Generalized Convexity Cone". If $\{u_0, u_1, \cdots, u_i\}$ is a T-system for $i = 0, 1, \cdots, n$, then $\{u_0, u_1, \cdots, u_n\}$ is called a Complete Tchebycheff system or CT-system. Note that no assumptions of continuity have been made in this paragraph.

In 1965 there appeared a paper by M. A. Rutman in which the following proposition is stated (cf. [**4**, Thm. 3]):

THEOREM. *Suppose the system of right-continuous functions $\{1, u_1, u_2, \cdots, u_n\}$ is a CT-system on the open interval (a, b). Then.there is a system $\{1, y_1, y_2, \cdots, y_n\}$ admitting of the following two representations on (a, b):*

$$y_i = u_i + \sum_{j=0}^{i-1} a_{i,j} u_j; \qquad i = 1, 2, \cdots, n, \tag{2}$$

and

$$\begin{aligned} y_1(t) &= \int_c^t dp_1(s) \\ y_2(t) &= \int_c^t \int_c^{s_1} dp_2(s_2) dp_1(s_1) \\ &\text{--------------------} \\ y_n(t) &= \int_c^t \int_c^{s_1} \cdots \int_c^{s_{n-1}} dp_n(s_n) dp_{n-1}(s_{n-1}) \cdots dp_1(s_1), \end{aligned} \tag{3}$$

where $c \in (a, b)$ is arbitrary, and the functions p_i are strictly increasing and right-continuous on (a, b).

Rutman's theorem has apparently nowhere been proved in detail; in fact, as will be shown in the next section, it is not correct with this degree of generality. The purpose of this paper is to investigate the existence of representations of the type (3), in the most general context possible.

If the functions u_i are n times continuously differentiable on an interval I, we can extend the definition of $D(u_0, \cdots, u_n/t_0, \cdots, t_n)$ as given by (1), so as to allow for equalities amongst the t_i: if $t_0 \leqq t_1 \leqq \cdots \leqq t_n$ is any set of points of I, then $D^*(u_0, \cdots, u_n/t_0, \cdots, t_n)$ is defined to be the determinant in the right hand side of (1), where for each set of equal t_i's, the corresponding columns are replaced by the successive derivatives evaluated at the point. With this definition, the system $\{u_0, \cdots, u_n\}$ will be called an Extended Tchebycheff system (ET-system), provided that $D^*(u_0, \cdots, u_n/t_0, \cdots, t_n) > 0$ for every set $t_0 \leqq t_1 \leqq \cdots \leqq t_n$ of points of I. If, moreover, the systems $\{u_0, \cdots, u_i\}$; $i = 0, 1, \cdots, n$ are ET-systems, the system $\{u_0, \cdots, u_n\}$ is called an Extended Complete Tchebycheff system (ECT-system). The validity of Rutman's theorem for ECT-systems has essentially been proved in [**3**, Ch. XI, Theorem 1.2], a fact which will be used further along in our discussion.

We now turn to the statement of our results. Following R. Zielke [**8, 9**], we shall say that a set A has property (D), provided it is totally ordered, it contains no smallest nor greatest element, and for every two distinct elements of A, there is a third element of A in between. The main feature of this paper is the following:

THEOREM 1. *Let $\{u_0, \cdots, u_n\}$ be a CT-system on a nondenumerable set A having property* (D), *and let $c \in A$. Then there is a system of functions $\{y_0, \cdots, y_n\}$ having the following properties:*

(a) *The functions $y_1, \cdots, y_n$ have a representation of the form* (2) *and $y_0 = u_0$ on A.*

(b) *There is a subset B of A, having an at most denumerable complement in A, a real valued strictly increasing function h, defined on A, and a set $\{p_1, \cdots, p_n\}$ of real valued strictly increasing functions, defined on the open interval whose end points are the infimum and the supremum of $h(A)$, such that $p_i[h(c)] = 0$, $i = 1, \cdots, n$ and, for every point t of B,*

$$
\begin{aligned}
&y_1(t) = y_0(t) \int_{h(c)}^{h(t)} dp_1(s) \\
(4) \qquad &y_2(t) = y_0(t) \int_{h(c)}^{h(t)} \int_{h(c)}^{s_1} dp_2(s_2) dp_1(s_1) \\
&\text{---------------------} \\
&y_n(t) = y_0(t) \int_{h(c)}^{h(t)} \int_{h(c)}^{s_1} \cdots \int_{h(c)}^{s_{n-1}} dp_n(s_n) dp_{n-1}(s_{n-1}) \cdots dp_1(s_1).
\end{aligned}
$$

Moreover, if the function y is convex with respect to the system $\{u_0, \cdots, u_n\}$, *it admits of the representation*

$$y(t) = u(t) + y_0(t)\int_{h(c)}^{h(t)}\int_{h(c)}^{s_1}\cdots\int_{h(c)}^{s_n} dp_{n+1}(s_{n+1})\cdots dp_1(s_1) \tag{5}$$

on B, where p_{n+1} *is an increasing function and u is in the linear span of the system* $\{u_0, \cdots, u_n\}$.

(c) *For* $i = 1, \cdots, n$, *the functions* y_i *are uniquely determined, and the functions* p_i *are uniquely determined* (a.e.), *by the functions* $h, u_0, \cdots, u_i$.

(d) *If A is a dense subset of an open interval* (in particular if A is an open interval), *h can be taken to be the identity function*: $h(t) = t$.

(e) *Let* $v^+(t)$ *and* $v^-(t)$ *denote the right and left one-sided limits of the function v at the point t, and let* $v_i = (u_i \circ h^{-1})/(u_0 \circ h^{-1})$. *If t is an element of the set B there is a real number p, contained in the interval* $[0, 1]$, *such that if* $s = h(t)$, *then*

$$z_i(s) = pz_i^+(s) + (1-p)z_i^-(s), \qquad i = 1, \cdots, n+1. \tag{6}$$

Conversely, if for some point t of A the functions z_i *admit of a representation of the form* (6), *with* $s = h(t)$, *then t is contained in B, provided that p be contained in the* open *interval* $(0, 1)$. *If* $p = 0$ *or* $p = 1$, *t may, or may not, belong to B.* (See the first counterexample in the next section.)

REMARKS. (a) Note that if $\{u_0, \cdots, u_n\}$ is a T-system defined on a set having property (D), its linear span contains a basis that constitutes a CT-system, (see **[8]**).

(b) For the case of an ECT-system, the representation (5) is implicit in [**3**, Ch. XI, section 11] (see also **[6]**).

In the following theorem, we have gathered several propositions of independent interest, some of which will be employed in the sequel.

THEOREM 2. *Let* $\{1, y_1, \cdots, y_n\}$ *be a CT-system on a dense subset A of an open interval I, and assume that* $y_{n+1} \in C(1, y_1, \cdots, y_n)$ *thereon. Then*:

(a) *For* $i = 1, \cdots, n+1$, *and every point t of I, the one-sided limits* $y_i^+(t)$ *and* $y_i^-(t)$ *exist and are finite, and the functions* y_i^+ *and* y_i^- *thus defined are of bounded variation in every closed subinterval of I.*

(b) *For every function* $\alpha: I \to [0, 1]$, *if* $z_i = \alpha y_i^+ + (1-\alpha)y_i^-$, *then* $\{1, z_1, \cdots, z_n\}$ *is a CT-system on I, and* $z_{n+1} \in C(1, y_1, \cdots, y_n)$ *thereon, under the additional assumption that A is nondenumerable.*

(c) *If* y_1 *is right* (left) *continuous at a given point of A, all the functions* y_i, $i = 1, \cdots, n+1$ *are right* (left) *continuous at this point.*

(d) *Let* $n > 1$, *and let s be a point of I. If equation* (6) *is satisfied by* *the functions* y_1, y_2, *and* y_3, *where* p *is a point of the interval* $[0, 1]$, *then it is satisfied* (for the same number p) *by all the functions* y_i, $i = 1, \cdots, n + 1$.

REMARK. Note the connection between Theorem 2d. and Theorem 1e.

The proof of Theorem 1 will be divided into two parts, which we shall now outline. In the first part, we shall consider bounded functions u_i defined on an open bounded interval I, such that $\{1, u_1, \cdots, u_n\}$ is a CT-system, $u_{n+1} \in C(1, u_1, \cdots, u_n)$ and, for $i = 1, \cdots, n + 1$, $u_i = (1/2) \cdot (u_i^+ + u_i^-)$. Convolving each function u_i with the Gauss kernel $G_k(s)$ we obtain (for each k), an ECT-system $\{u_0(k, \cdot), \cdots, u_n(k, \cdot)\}$, such that $u_{n+1} \in C(u_0(k, \cdot), \cdots, u_n(k, \cdot))$. As we have already remarked, Rutman's Theorem is valid for ECT-systems. Since $u_i = (1/2)(u_i^+ + u_i^-)$, it can be shown that $\lim_{n\to\infty} u_i(k, \cdot) = u_i$ on I. Thus, the proof of this case will follow (after a number of steps), by letting k tend to infinity. The general case will be considered in the second part of the proof. By a suitable normalization, the original system will be transformed into a CT-system of bounded functions defined on a dense subset D of an open interval I. Redefining these functions so that they will equal the average of their lateral limits everywhere on D, and applying Theorem 2 to extend them to the whole of I, we shall reduce the problem to the one considered in the first part of the proof.

Using Theorems 1 and 2, we shall easily prove the following proposition, which generalizes a result of Bartelt [**1**, Theorem 1]:

THEOREM 3. *Let* $\{u_0, \cdots, u_n\}$ *be a* T*-system on a dense subset* A *of an interval* (a, b). *Then*:

(a) *The system* $\{u_0, \cdots, u_n\}$ *can be extended as a* T*-system to the whole of* (a, b).

(b) *If* $\{u_0^*, \cdots, u_n^*\}$ *is a system of continuous functions on* (a, b), *such that* $u_i^* = u_i$ *on* A, $i = 0, \cdots, n$, *the following propositions are equivalent*:

(i) *The system* $\{u_0^*, \cdots, u_n^*\}$ *is a* T*-system on* (a, b).

(ii) *The linear span of* $\{u_0^*, \cdots, u_n^*\}$ *contains a function that does not vanish on* (a, b).

(iii) *The functions* u_i^* *cannot all vanish at any one point of* (a, b).

Theorem 1 admits of a converse. Indeed, we shall readily prove the following:

LEMMA. *If the functions* y_i *admit of a representation of the form* (4) *on a totally ordered set* A, *where the functions* h, $p_1, \cdots, p_n$ *are strictly increasing, and* y_0 *is strictly positive, then* $\{y_0, \cdots, y_n\}$ *is a* CT*-system on* A.

2. Counterexamples to Rutman's theorem. Let

$$p_1(t) = \begin{cases} t-1, & \text{if} \quad -1 < t < 0 \\ 0, & \text{if} \quad t = 0 \\ t+1, & \text{if} \quad 0 < t < 1 \end{cases}, \qquad p_2(t) = \begin{cases} t-1, & \text{if} \quad -1 < t < 0 \\ \\ t+1, & \text{if} \quad 0 \leqq t < 1 \end{cases},$$

and define

$$y_1(t) = \int_0^t dp_1(s),$$

and

$$y_2(t) = \int_0^t \int_0^{s_1} dp_2(s_2) dp_1(s_1) = \int_0^t [p_2(s) - 1] dp_1(s).$$

Thus,

$$y_1(t) = p_1(t),$$

and

$$y_2(t) = \begin{cases} (1/2)t^2 - 2t & \text{on} \quad (-1, 0) \\ (1/2)t^2 & \text{on} \quad [0, 1). \end{cases}$$

We know from our Lemma, that $\{1, y_1, y_2\}$ is a CT-system on $(-1, 1)$. We thus see from Theorem 2b. or by direct computation, that if

$$u_1(t) = \begin{cases} t-1 & \text{on} \quad (-1, 0) \\ t+1 & \text{on} \quad [0, 1) \end{cases},$$

and $u_2 = y_2$, then also $\{1, u_1, u_2\}$ is a CT-system on $(-1, 1)$. Assume now, that there are two functions $v_1 = u_1 + a$, $v_2 = u_2 + b_1 u_1 + b_2$, and two strictly increasing functions q_1, q_2, such that $v_1(t) = \int_0^t dq_1(s)$, and $v_2(t) = \int_0^t \int_0^{s_1} dq_2(s_2) dq_1(s_1)$. Without any loss of generality, we can assume that $q_1(0) = q_2(0) = 0$; thus $v_1 = q_1 = u_1 + a$. If t is any point of the interval $(-1, 1)$ other than 0, it is clear from their definition that $u_1 = y_1$. Since, as remarked above, $y_1 = p_1$, we conclude that $q_1 = p_1 + a$, identically on $(-1, 1)$, excepting perhaps at zero. Since the function q_2 vanishes at 0, it is readily seen that $v_2 = \int_0^t q_2(s) ds$. On the other hand, from the integral representation of the functions y_i, and bearing in mind that $u_i = y_i$ if $t \neq 0$, we see that, for every point of $(-1, 1)$, with the possible exception of 0, $v_2 = y_2 + b_1 y_1 + b_2 = \int_0^t [p_2(s) + b_1 - 1] dp_1(s) + b_2.$

Let r be the saltus function of p_1, i.e.

$$r(t)=\begin{cases} 0 & \text{on} \quad (-1,0) \\ 1 & \text{if} \quad t=0 \\ 2 & \text{on} \quad (0,1) \end{cases}$$

Then $p_1(t)=(t-1)+r(t)$. Thus, if $t>0$,

$$v_2(t)=\int_0^t [p_2(s)+b_1-1]ds+2[p_2(0)+b_1-1]+b_2.$$

From the integral representation of v_2 in terms of q_2, we see that v_2 vanishes at 0. Passing to the limit in the preceding formula, and bearing in mind that $p_2(0)=1$, we conclude that $b_2+2b_1=0$. Since $u_1(0)=1$, and $u_2(0)=0$, from the representation of v_2 in terms of u_1 and u_2 we conclude, on the other hand, that $b_1+b_2=0$. It is therefore clear that $b_1=b_2=0$. Thus, if $t>0$,

$$v_2(t)=\int_0^t [p_2(s)-1]ds=\int_0^t q_2(s)ds,$$

and we conclude that for positive values of t, $q_2=p_1-1$. Passing to the limit, we readily see that q_2 is right-continuous at 0. But, being that $q_1=u_1+a$, also q_1 is right continuous at 0, and therefore the integral $v_2(t)=\int_0^t q_2(s)dq_1(s)$ cannot exist for negative values of t.

We can also prove that, even if the desired representation exists, the functions p_i in (3) may not be right-continuous. In order to see this, we can consider, for instance, the following example: Let q_1 and q_2 be strictly increasing functions defined in the interval $(-1,1)$. Let q_1 be right-continuous and q_2 left-continuous, and assume that they are bounded, and have the same set of points of discontinuity; assume moreover that $q_1(0)=q_2(0)=0$, and define $u_1(t)=\int_0^t dq_1(s)$, and $u_2(t)=\int_0^t\int_0^{s_1} dq_2(s_2)dq_1(s_1)$. Clearly $u_1=q_1$, and $u_2(t)=\int_0^t q_2(s)dq_1(s)$. Thus also u_2 is right-continuous on $(-1,1)$.

Assume now there are two functions y_1, y_2, $y_1=u_1+a$, $y_2=u_2+b_1u_1+b_2$, and two strictly increasing, right-continuous functions p_1, p_2, such that, for every t in $(-1,1)$,

$$y_1(t)=\int_0^t dp_1(s),$$

and

$$y_2(t) = \int_0^t \int_0^{s_1} dp_2(s_2)dp_1(s_1)$$

Since $u_1(0) = y_1(0) = 0$, we conclude that $a = 0$. Thus $dp_1 = dq_1$. Since also $u_2(0) = y_2(0) = 0$, it is clear that $b_2 = 0$, whence combining the two representations of y_2, we see that

$$y_2(t) = \int_0^t [p_2(s) - p_2(0)]dq_1(s) = \int_0^t [q_2(s) + b_1]dq_1(s).$$

Let $k = b_1 + p_2(0)$. It is clear from the preceding identity, that $q_2 + k = p_2$, on the set of points at which q_1 is continuous. Since q_1 and q_2 have the same set of points of discontinuity, we therefore conclude that also p_2 and q_1 have the same points of discontinuity. But $dp_1 = dq_1$. Thus p_1 and p_2 have the same points of discontinuity. But this is clearly impossible because, being that both p_1 and p_2 are right-continuous, it will suffice that the interval $(0, t)$ contain one point of discontinuity of p_1, for the integral

$$\int_0^t [p_2(s) - p_2(0)]dp_1(s)$$

not to exist.

Rutman's mistake is due to his belief that, if $\{y_0, \cdots, y_n\}$ is a CT-system of right-continuous functions on an interval (a, b) with $y_0 = 1$, the functions $z_i(t) = \lim_{h\to 0^+} [y_i(t+h) - y_i(t)]/[y_1(t+h) - y_1(t)]$ not only exist for every point t of (a, b), (which is true), but that in addition, they are right-continuous thereon (cf. op. cit. Thm. 2). This can be disproved by considering the following counterexample:
Let

$$p_1(t) = \begin{cases} t-1 & \text{on} \quad (-1,0) \\ t & \text{on} \quad [0,1] \end{cases}, \qquad p_2(t) = \begin{cases} t-1 & \text{on} \quad (-1,0] \\ t+1 & \text{on} \quad (0,1) \end{cases}$$

and define: $y_0 = 1$, $y_1 = p_1$, $y_2 = \int_0^t p_2(s)dp_1(s)$.

Referring to our Lemma, we see that $\{y_0, y_1, y_2\}$ is a CT-system of right-continuous functions on (a, b). Moreover, if $t \neq 0$, $z_2(t) = p_2(t)$. However, it is easy to verify that $z_2(0) = 0$, while $\lim_{t\to 0^+} z_2(t) = \lim_{t\to 0^+} p_2(t) = 1$.

3. Proofs.

Proof of Lemma. Taking into consideration [3, p. 10, Example 2b.], it will suffice to consider the case $h(t) = t$. From the linearity of the integral we conclude (as in the proof of [3, Ch. XI, Lemma 2.1]), that if $\{1, y_1, y_2, \cdots, y_i\}$ is a T-system and p is a strictly increasing function, then $\left\{1, \int_c^t 1.dp, \int_c^t y_1 dp, \cdots, \int_c^t y_i dp\right\}$ is a T-system, whence the conclusion follows by an obvious inductive procedure.

Proof of Theorem 2. (a) Let $I = (a, b)$. The assertions are obvious for $n = 0$. Assume they are true if $n = k$, and let $n = k + 1$. It will suffice to carry out the proof for every interval of the form (c, b), $a < c < b$, with $c \in A$. Let $g_i = [y_i - y_i(c)]/[y_1 - y_1(c)]$, and $c < t_1 < \cdots < t_{n+1} < b$. Developing by the first column, we readily see that

$$D(1, y_1, \cdots, y_i/c, t_1, \cdots, t_i) = \left(\prod_{j=1}^{n} [y_1(t_j) - y_1(c)]\right) D(1, g_2, \cdots, g_i/t_1, \cdots, t_i).$$

Since y_1 is strictly increasing, the factors $y_1(t_j) - y_1(c)$ are strictly positive. We thus conclude that $\{1, g_2, \cdots, g_n\}$ is a CT-system on $A \cap (c, d)$, and $g_{n+1} \in C(1, g_2, \cdots, g_n)$ thereon, whence the assertions readily follow from the inductive hypothesis.

(b) Let B denote the set of points of A at which all the functions y_i are continuous. From (a) we know that the set difference $A \sim B$ is at most denumerable; therefore B is dense in I. Clearly $y_i = z_i$ on B; thus $\{1, z_1, \cdots, z_n\}$ is a CT-system on B, and $z_{n+1} \in C(1, \cdots, z_n)$ thereon. We shall first prove that, for $i = 1, 2, \cdots, n + 1$, $\{1, z_1, \cdots, z_i\}$ is a WT-system on I; from this will follow, in particular, that $z_{n+1} \in C(1, z_1, \cdots, z_n)$ on I. The assertion is obvious if the function α can only assume the values 0 and 1; indeed, this simply means that for every point t, either $z_i(t) = y_i^+(t)$, $i = 1, \cdots, n$, or $z_i(t) = y_i^-(t)$, $i = 1, \cdots, n$ whence, since B is dense in I, the proof of our claim follows by an obvious limiting process. In the general case, the assertion follows from the preceding discussion, the linearity of the determinant, and the fact that the functions α and $1 - \alpha$ are nonnegative. For example, let $a < t_0 < t_1 < b$, $\alpha(t_0) = p$, $\alpha(t_1) = q$. Then, setting $z_0 = 1$, we have: $D(1, z_1/t_0, t_1) = \det \| z_i(t_0), z_i(t_1) \| = \det \| p z_i^+(t_0) + (1 - p) z_i^-(t_0), \quad q z_i^+(t_1) + (1 - q) z_i^-(t_1) \| =$
$pq \det \| z_i(t_0), z_i^+(t_1) \| + p(1 - q) \det \| z_i^+(t_0), z_i^-(t_1) \| +$
$(1 - p) q \det \| z_i^-(t_0), z_i^+(t_1) \| + (1 - p)(1 - q) \det \| z_i^-(t_0), z_i^-(t_1) \| \geqq 0.$

To prove that $\{1, z_1, \cdots, z_n\}$ is a CT-system we proceed by induction. The assertion is obvious if $n = 0$. Assume it to be true if $n = k$, and let $n = k + 1$. By inductive hypothesis, $\{1, z_1, \cdots, z_{n-1}\}$ is a CT-system. Thus, we only have to prove that $\{1, z_1, \cdots, z_n\}$ is a

T-system on I. Assume this is not so; then, there is a nontrivial linear combination g of the functions $1, z_1, \cdots, z_n$, that vanishes on a set $t_0 < t_1 < \cdots < t_n$ of points of I. Since by inductive hypothesis $\{1, z_1, \cdots, z_{n-1}\}$ is a T-system on I, the coefficient of z_n in g cannot be zero; thus, without any loss of generality, we can assume it to equal 1. Since B is dense in I, and $\{1, z_1, \cdots, z_n\}$ is a T-system thereon, there is a point t'_n of B, $t_{n-1} < t'_n < t_n$, such that $g(t'_n) \neq 0$. Assume first that $g(t'_n) > 0$; then, since the coefficient of z_n in g is 1, developing by the last row and applying the inductive hypothesis, we see that

$$\begin{aligned} D(1, z_1, \cdots, z_n/t_1, \cdots, t_{n-1}, t'_n, t_n) &= D(1, z_1, \cdots, z_{n-1}, g/t_1, \cdots, t_{n-1}, t'_n, t_n) \\ &= -g(t'_n) \cdot D(1, z_1, \cdots, z_{n-1}/t_1, \cdots, t_n) \\ &< 0. \end{aligned}$$

If $g(t'_n) < 0$, we similarly see that $D(1, z_1, \cdots, z_n/t_0, \cdots, t_{n-1}, t'_n) < 0$. But these conclusions contradict the fact that $\{1, z_1, \cdots, z_n\}$ is a WT-system on I.

(c) The assertion is obviously true for $n = 1$. Assume it to be true for $n = k$, and let $n = k + 1$. Assume for example that y_1 is right continuous at a point s of A. By inductive hypothesis, the functions y_i, $i = 1, \cdots, n$ are right continuous at s. Let $t_0 < t_1 < \cdots < t_{n-1} < s < t_n$, where the t_i's are points of A. Since $D(1, y_1, \cdots, y_{n+1}/t_0, \cdots, t_{n-1}, s, t_n)$ is nonnegative, making t_n tend to s, we see that

$$\begin{aligned} 0 &\leqq D(1, y_1, \cdots, y_{n+1}/t_0, \cdots, t_{n-1}, s, s^+) \\ &= [y^+_{n+1}(s) - y_{n+1}(s)] \cdot D(1, y_1, \cdots, y_n/t_0, \cdots, t_{n-1}, s). \end{aligned}$$

Thus $y^+_{n+1}(s) - y_{n+1}(s) \geqq 0$.

By considering now points t'_i, such that $t'_0 < \cdots < t'_{n-2} < s < t'_{n-1} < t'_n$ (if $n = 2$, such that $s < t'_0 < t'_1$), and making t_{n-1} tend to s, we similarly see that $y^+_{n+1}(s) - y_{n+1}(s) \leqq 0$, whence the conclusion follows.

(d) The assertion is true by hypothesis for $n = 2$. Assume it to be true for $n = k$, and let $n = k + 1$. Let s be any given point of A, and let $t_0 < \cdots < t_n$ be points of the set $A \cap (s, b)$. Let $Q = D(1, y_1, \cdots, y_n/t_0, \cdots, t_n)$ and, for $i = 0, 1, \cdots, n$, let the functions z_i be defined by $z_i(t) = D(1, y_1, \cdots, y_n/t_0, \cdots, t_{i-1}, t, t_{i+1}, \cdots, t_n)$. It is readily seen that $D(z_0, \cdots, z_n/t_0, \cdots, t_n) = Q^{n+1} > 0$, and proceeding as in [7, Lemma 2], we conclude that $\{z_0, \cdots, z_n\}$ is a Tchebycheff system on A. If $z_{n+1}(t) = D(z_0, \cdots, z_n, y_{n+1}/t_0, \cdots, t_n, t)$, it is also clear that $z_{n+1} \in C(z_0, \cdots, z_n)$ thereon. Let $s_0 < \cdots < s_i$ be points of $(a, t_0) \cap A$, $(i < n)$. Then

$$0 < D(z_0, \cdots, z_n/s_0, \cdots, s_i, t_{i+1}, \cdots, t_n) = Q^{i+1} D(z_0, \cdots, z_i/s_0, \cdots, s_i),$$

and it follows that $\{z_0, \cdots, z_n\}$ is a CT-system on $(a, t_0) \cap A$. It is similarly seen that also $\{z_0, \cdots, z_{n-2}, -z_n\}$ is a CT-system on $(a, t_0) \cap A$, and that $z_{n+1} \in C(z_0, \cdots, z_{n-2}, -z_n)$ thereon. Let $v_i = z_i / z_0$, and assume first that v_1 is continuous at s. From (c) we know that all the functions v_i must be continuous at s. It is therefore clear that the functions v_i admit of a representation of the form (6) at the point s, where p can be any number in the interval $[0, 1]$. On the other hand, were v_1 discontinuous at s, if a representation of the form (6) exists at all, p must clearly be unique. Applying therefore the inductive hypothesis to the systems $\{1, v_1, \cdots, v_n\}$ and $\{1, v_1, \cdots, v_{n-2}, -v_n, v_{n+1}\}$ we conclude that the functions $v_1, \cdots, v_{n+1}$ admit of a representation of the form (6) at the point s. In particular, since $1/z_0$ is in the linear span of these functions, it also admits of this representation, and we readily conclude that $p[z_0(s)/z_0^+(s)] + (1-p)[z_0(s)/z_0^-(s)] = 1$. Setting $q = p[z_0(s)/z_0^+(s)]$, and bearing in mind that $z_i = z_0 \cdot v_i$, the conclusion follows.

Proof of Theorem 1. We shall first consider the case in which $\{1, u_1, \cdots, u_n\}$ is a CT-system of bounded functions defined on a bounded open interval (a, b), such that $u_i = (1/2)(u_i^+ + u_i^-)$ on (a, b), for $i = 1, \cdots, n+1$.

Given a real-valued function u defined on (a, b), let $u(k, \cdot)$ be given by

$$u(k, t) = \int_a^b u(s) G_k(t-s) ds, \quad \text{where}$$

$$G_k(s) = (k/\sqrt{2\pi}) \exp[-(1/2)k^2 s^2].$$

Under the conditions imposed on the functions u_i, it can be shown that $\lim_{k \to \infty} u_i(k, \cdot) = u_i$ on (a, b), for $i = 0, 1, \cdots, n$. To see this, extend the functions u_i to the whole real line by stipulating that they should vanish outside of (a, b). Setting $\sigma_i(t) = \int_a^t u_i(s) ds$, and taking into consideration that $G_k(t-s) = G_k(s-t)$, we see that

$$u_i(k, t) = \int_{-\infty}^{\infty} G_k(s-t) d\sigma_i(s) = \int_{-\infty}^{\infty} G_k(s) d\sigma_i(s+t),$$

and the conclusion follows from [**5**, Theorem 4].

From the Basic Composition Formula [**3**, pp. 14, 15], we know that for any fixed integer k, $k = 1, 2, \cdots$, the system $\{u_0(k, \cdot), \cdots, u_n(k, \cdot)\}$ is an ECT-system on $[a, b]$. Let c be any given point of (a, b). From [**3**, Ch. XI, Thm. 1.1] we conclude that by adding to each function $u_i(k, \cdot)$ a suitable linear combination of its predecessors, we obtain a system

$$y_i(k,\cdot) = u_i(k,\cdot) + \sum_{j=0}^{i-1} a_{i,j}(k)u_j(k,\cdot),$$

satisfying the constraints $y_i^{(r)}(k,c) = 0$; $r = 0,1,\cdots,i-1$; $i = 1,2,\cdots,n$.

Proceeding exactly as in the proof of [**3**, Ch. XI, Thm. 1.2], we see that the functions $y_i(k,\cdot)$ admit of a representation of the form $y_0(k,\cdot) = w_0(k,\cdot)$; $y_1(k,t) = w_0(k,t)\int_c^t w_1(k,s)ds$, and in general

$$y_i(k,t) = w_0(k,t)\int_c^t w_1(k,s_1)\int_c^{s_1} w_2(k,s_2)\int_c^{s_2}\cdots\int_c^{s_{i-1}} w_i(k,s_i)ds_ids_{i-1}\cdots ds_1$$

on $[a,b]$, for $i = 1,2,\cdots,n$, where the functions $w_i(k,\cdot)$ are strictly positive. Thus the functions $p_i(k,t) = \int_c^t w_i(k,s)ds$ are strictly increasing. Clearly,

$$(7)\quad y_i(k,t) = y_0(k,t)\int_c^t\int_c^{s_1}\cdots\int_c^{s_{i-1}} dp_i(k,s_i)dp_{i-1}(k,s_{i-1})\cdots dp_1(k,s_1);$$
$$i = 1,2,\cdots,n.$$

Let $a < t_0 < t_1 < \cdots t_{n-1} < c$, where the points t_i are arbitrary but fixed. Let

$$z_i(k,t) = D(u_0(k,\cdot),\cdots,u_i(k,\cdot)/t_0,\cdots,t_{i-1},t)/$$
$$D(u_0(k,\cdot),\cdots,u_{i-1}(k,\cdot)/t_0,t_1,\cdots,t_{i-1}).$$

If $\|f\|_A$ denotes the supremum of the function $|f|$, taken over the set A, it is obvious that the sequences $\{\|u_i(k,\cdot)\|_{[a,b]}\}$; $k = 1,2,\cdots$ are bounded. Thus, also the sequences $\{\|z_i(k,\cdot)\|_{[a,b]}\}$; $k = 1,2,\cdots$ are bounded. Moreover, from [**7**, Lemma 3] we conclude that the functions $z_i(k,\cdot)$; $i = 1,2,\cdots,n$ admit of a representation of the form

$$(8)\quad z_i(k,t) = w_0(k,t)\int_{t_{i-1}}^t w_1(k,s_1)\int_{b_1(k,i)}^{s_1} w_2(k,s_2)\cdots$$
$$\int_{b_{i-1}(k,i)}^{s_{i-1}} w_i(k,s_i)ds_i\cdots ds_1;$$

where $t_0 < b_{i-1}(k,i) < b_{i-2}(k,i) < \cdots < b_1(k,i) < t_{i-1}$. Therefore, by comparing (7) and (8) we conclude that if $t \geqq c$, $0 \leqq y_i(k,t) \leqq z_i(k,t)$. This implies that the sequences $\{\|y_i(k,\cdot)\|_{[c,b]}\}$; $k = 1,2,\cdots$ are bounded, whence we can easily show that the coefficients $a_{i,j}(k)$ are uniformly bounded in i, j and k. Thus, there exist numbers $a_{i,1}, a_{i,2}, \cdots, a_{i,i}$, with $a_{i,i} = 1$, and a sequence $\{k_j\}$; $j = 1, 2, \cdots$ such that, for $i = 1, 2, \cdots, n$,

$$\lim_{j\to\infty} y_i(k_j, \cdot) = \sum_{j=0}^{i} a_{i,j} u_j \equiv y_i^* \tag{9}$$

on (a, b). Clearly $\{1, y_1^*, \cdots, y_n^*\}$ is a CT-system on (a, b). It remains to show that there exist strictly increasing functions $p_1, \cdots, p_n$ such that for $y_1, \cdots, y_n$ as in equations (3), $y_i = y_i^*$, $i = 1, \cdots, n$.

We shall now restrict ourselves to the consideration of any closed subinterval $[\alpha, \beta]$ of (a, b). The proof for the whole of (a, b) will follow from the uniqueness of the functions p_i.

We shall consider two cases, according as to whether the functions $p_i(k_j, \cdot)$ are bounded on $[\alpha, \beta]$ uniformly in $\{k_j\}$, or not. The first case will give us the desired representation, whereas the second will be shown to be impossible.

Since the functions $p_i(k, \cdot)$ are monotone, in the first case we can apply Helly's first theorem to conclude that there are increasing functions p_i and a subsequence $\{k'_j\}$ of k_j, such that $\lim_{j\to\infty} p_i(k'_j, \cdot) = p_i$ on $[\alpha, \beta]$, for $i = 1, \cdots, n$. Passing to the limit under the integral sign in (7), we conclude from (9) that the functions y_i^* have a representation of the form (3) on $[\alpha, \beta]$. (This passage to the limit, in which a Riemann–Stieltjes integral is obtained, can be easily justified by means of Theorems 15.2 and 15.6 of [**2**, Chapter 2]). The functions p_i must be strictly increasing, for assume that some function p_i is not strictly increasing. Then p_i must take a constant value on some subinterval of $[\alpha, \beta]$. By a simple inductive procedure, involving the number of integrations, we see that under this condition y_i^* must be in the linear span of the functions $1, y_1^*, \cdots, y_{i-1}^*$ thereon. But this contradicts the fact that $\{1, y_1^*, \cdots, y_n^*\}$ is a T-system. Setting therefore $y_i = y_i^*$, $i = 1, \cdots, n$, the conclusion follows.

Assume now that for some i the sequence $\{p_i(k, \cdot)\}$ is not uniformly bounded on $[\alpha, \beta]$, and let m be the smallest such i. Since the sequence $\{a_{1,0}(k)\}$; $k = 1, 2, \cdots$ is bounded, we easily conclude from (9) that $m > 1$. Since the functions $p_m(k, \cdot)$ are increasing, we know that one of the sequences $\{p_m(k, \alpha)\}$ or $\{p_m(k, \beta)\}$; $k = 1, 2, \cdots$ is unbounded. Among the several possibilities, let us choose for illustration purposes the case in which $\{p_m(k, \alpha)\}$ is unbounded and $\alpha < c < \beta$. By definition, the functions $p_i(k, \cdot)$ vanish at c and are negative to the left of c. Hence, there exists a sequence $\{k_r\}$ and strictly increasing functions p_i, such that $\lim_{r\to\infty} p_i(k_r, \cdot) = p_i$ on $[\alpha, \beta]$; $i = 1, \cdots, m-1$, and, if $t \leqq \alpha$, $\lim_{r\to\infty} p_m(k_r, t) = -\infty$. Let $I_n(t_1, t_2)$ denote the subset of R^n defined by $t_1 \leqq s_n \leqq s_{n-1} \leqq \cdots \leqq s_1 \leqq t_2$. Then, from (7) we see that, if $t < \alpha$,

$$
\begin{aligned}
|y_m(k,t)/y_0(k,t)| \\
&= \int_{I_m(t,c)} dp_m(k,s_m)\cdots dp_1(k,s_1) \\
&= \int_{I_{m-1}(t,c)} [-p_m(k,s_{m-1})]\,dp_{m-1}(k,s_{m-1})\cdots dp_1(k,s_1) \\
&\geqq \int_{I_{m-1}(t/\alpha)} [-p_m(k,s_{m-1})]\,dp_{m-1}(k,s_{m-1})\cdots dp_1(k,s_1) \\
&\geqq [-p_m(k,\alpha)]\int_{I_{m-1}(t/\alpha)} dp_{m-1}(k,s_{m-1})\cdots dp_1(k,s_1).
\end{aligned}
$$

Thus, $|y^*_m(t)| = \lim_{r\to\infty} |y_m(k_r,t)/y_0(k_r,t)| = \infty$, which is impossible.

Let u be bounded, and convex with respect to $\{u_0,\cdots,u_n\}$. If u is contained in their linear span, in order to obtain a representation of the form (5) it will suffice to take $p_{n+1} = 1$. If u is not contained in the span of the functions u_i, it is easily seen that $\{v_0(k,\cdot),\cdots,v_n(k,\cdot),u(k,\cdot)\}$ is an ECT-system, and we can proceed as above.

In order to establish the uniqueness of the functions p_i, assume that the functions y_i admit of a representation of the form (3) with respect to two sets of functions $\{p_i\}$ and $\{q_i\}$, with $p_i(c) = q_i(c) = 0$; $i = 1,\cdots,n$. Then $\int_c^t dp_i = p_i(t)$ and $\int_c^t dq_i = q_i(t)$. Without loss of generality we can assume that $y_0(t) = 1$. Then $y_1(t) = p_1(t) = q_1(t)$, and

$$
\begin{aligned}
y_i(t) &= \int_c^t \int_c^{s_1}\cdots\int_c^{s_{i-2}} p_i(s_{i-1})\,dp_{i-1}(s_{i-1})\cdots dp_1(s_1) \\
&= \int_c^t \int_c^{s_1}\cdots\int_c^{s_{i-2}} q_i(s_{i-1})\,dq_{i-1}(s_{i-1})\cdots dq_1(s_1); \qquad i = 2,\cdots,n,
\end{aligned}
$$

whence the assertion follows by repeated application of [**2**, Ch. II, Theorem 16.2]. In similar fashion, it can be shown that p_{n+1} is uniquely determined (a.e.), up to a constant, on the set of points at which it is strictly increasing.

We have thus shown that the functions p_i are uniquely determined (a.e.) by the functions $y_1,\cdots,y_i$. In order to establish part (c) of the theorem, assume that the functions $y_1,\cdots,y_k$ admit of the representations (2) and (3), and the functions $z_1,\cdots,z_k$ of a representation of the form (2) with respect to coefficients b_{ij}, and of a representation of the form (3) with respect to functions $q_1,\cdots,q_k$, where $u_0 = y_0 = z_0$, and $p_i(c) = q_i(c) = 0$, $i = 1,\cdots,k$. We have to show that $a_{ij} = b_{ij}$, and $p_i = q_i$, for $j = 0,\cdots,i-1$, and $i = 1,\cdots,k$. Since the systems $\{y_0,\cdots,y_k\}$ and $\{z_0,\cdots,z_k\}$ have both been obtained from the system $\{u_0,\cdots,u_k\}$ by means of a triangular linear transformation, it is clear that there exist coefficients c_{ij}, such that, for $i = 1, \cdots, k$, $y_i = z_i + \sum_{j=0}^{i-1} c_{ij} \cdot z_j$. Since the functions y_i and z_i admit of representations of the form (3), it is clear that they vanish at the point c, whence from the preceding formula we conclude that $0 = c_{i0} \cdot z_0$. Since $z_0 = u_0$, it is strictly positive. We thus conclude that $c_{i0} = 0$, and in particular that $y_1 = z_1$.

The proof of the assertion now easily follows by an inductive argument involving the number of integrations, and the fact that the function p_i is uniquely determined by the functions $y_1, \cdots, y_i$.

Let us now consider the general case. Since u_0 is strictly positive, and u_0, u_1 is a T-system, it is clear that the function u_1/u_0 is strictly increasing. Let Q be the image of A under the function $\operatorname{arctg}(u_1/u_0)$. Since this function is strictly increasing and bounded, it is clear that Q is a bounded set of real numbers, having property (D).

For every element t of Q, let $\mathfrak{l}(t)$ be the least upper bound of the set of points of Q that precede t, and $\mathfrak{u}(t)$ the greatest lower bound of the set of points of Q that follow t. Clearly there is an at most denumerable set of points of Q for which $\mathfrak{l} - \mathfrak{u} > 0$. For every element t of Q, let $q(t) = \mathfrak{l}(t) + \sum_{s<t}[\mathfrak{u}(s) - \mathfrak{l}(s)]$. It is readily seen that q is a strictly increasing function that transforms Q into a dense subset of a bounded open interval. Let $h(t) = q[\arctan(u_1/u_0)(t)]$, and consider the functions defined on $h(A)$ by $z_i(t) = u_i[h^{-1}(t)]/u_0[h^{-1}(t)]$, $i = 0, 1, \cdots, n+1$. Clearly $\{z_0, \cdots, z_n\}$ is a CT-system on $h(A)$, and $z_{n+1} \in C(z_0, \cdots, z_n)$ thereon. Since $z_0 = 1$, and $h(A)$ is a dense subset of an open bounded interval (a, b), we know from Theorem 2(b) that, if $z_i^* = (1/2)[z_i^+ + z_i^-]$, then $\{1, z_1^*, \cdots, z_n^*\}$ is a CT-system on (a, b), and $z_{n+1}^* \in C(1, z_1^*, \cdots, z_n^*)$ thereon. By a procedure similar to the one employed in the proof of Theorem 2(a), we readily see that the functions z_i^* are bounded in every interval of the form (a', b'), with $a < a' < b' < b$. From the uniqueness of the representations (2) and (3), we readily see that the assertions proved in the first part of this proof are also valid without the condition of boundedness. Thus, if c is a point in the set A, and $d = h(c)$, we know there is a system $\{1, y_1^*, \cdots, y_{n+1}^*\}$ admitting of the representation $y_i^* = z_i^* + \sum_{j=0}^{i-1} a_{i,j} z_j^*$, $i = 1, \cdots, n+1$, on (a, b), as well as of a representation of the form (3) thereon (where c is replaced by d), and the validity of statements (a), (b), (c), and (d) readily follows from the fact that $z_i = z_i^*$, except for an at most denumerable set of points. In order to prove (e) first note that, if t is a point of B, it is easily seen, from the basic properties of Riemann–Stieltjes integrals, that the functions y_i^* admit of a representation of the form (6) at the point $s = h(t)$, whence the proof of the necessity readily follows. Conversely, let S denote the set of points of (a, b) at which the functions z_i admit of a representation of the form (6), with p in the interval $[0, 1]$. By virtue of their integral representation, we know that the functions y_i^* can be written in the form

$$(10) \qquad y_i^*(t) = \int_d^t x_i(s)\,dp_1(s),$$

on (a, b), for $i = 1, \cdots, n+1$, where x_i is a function of bounded variation in every closed subinterval of (a, b). In particular, note that, for every point t of (a, b), $y_1^*(t) = p_1(t) - 0 = z_1^*(t) + a_{1,0}$. The function p_1 can be redefined at every point of discontinuity so as to assume any desired value between p_1^+ and p_1^-, without changing the values of the functions y_i^* at points other than these points of discontinuity, with one possible exception: If one of the functions x_i is discontinuous and right (left) continuous at a point t_0 at which also the function p_1 is discontinuous, then p_1 cannot be redefined to equal p_1^+ (or p_1^-) at this point, without affecting the existence of the integral (10) for all points t such that $|t-d| > |t_0 - d|$, as was shown in the first counterexample of §2.

Let therefore p_1^* be defined to equal $z_1 + a_{1,0}$ on the set of points of S for which p is neither 0 nor 1, and to equal p_1 at all other points of (a, b); for $i = 1, \cdots, n+1$, define v_i by means of the formula $v_i(t) = \int_d^t x_i(s)\,dp_1^*(s)$, and let $g_i = z_i + \Sigma_{j=0}^{i-1} a_{i,j} z_j$, where the coefficients $a_{i,j}$ are the same ones that appear in the representation of the functions y_i^* in terms of the functions z_i^*. In the light of the remarks made in the preceding paragraph, and bearing in mind that $v_1 = g_1$ on S, we see that $v_i = g_i$ thereon, for $i = 1, \cdots, n+1$. Thus, $y_i(t) = y_0(t) \cdot v_i[h(t)]$, on S, and the conclusion follows.

Proof of Theorem 3. The proof of part (a) is essentially contained in the preceding discussion, and will therefore be omitted. In order to prove part (b), note that the implication (i) $\Rightarrow$ (ii) is a direct consequence of **[8]**, whereas (ii) $\Rightarrow$ (iii) is trivial. In order to prove that (iii) $\Rightarrow$ (ii), note that in view of the result of **[8]** we can assume, without any loss of generality, that $\{u_0, \cdots, u_n\}$ is a CT-system on A. Thus, it is clear that $\{u_0^*, \cdots, u_k^*\}$ is a WT-system on (a, b), for $k = 0, \cdots, n$. Assume now that $u_0^*(s_0) = 0$, for some point s_0 of (a, b). We claim that $u_i^*(s_0) = 0$, $i = 0, \cdots, n$, in contradiction of (iii). We proceed by induction on i. The assertion is true by hypothesis, for $i = 0$. Assume it to be true for i smaller or equal to m, and let $i = m+1$. Let $a < t_0 < \cdots < t_{m-1} < s_0 < t_m < b$, where the points t_i are contained in A. Clearly $D(u_0^*, \cdots, u_m^*/t_0, \cdots, t_m) > 0$. Thus, since $u_i^*(s_0) = 0$, $i = 0, \cdots, m$, it is clear that

$$\begin{aligned} 0 &\leqq D(u_0^*, \cdots, u_{m+1}^*/t_0, \cdots, t_{m-1}, s_0, t_m) \\ &= -u_{m+1}^*(s_0) D(u_0^*, \cdots, u_m^*/t_0, \cdots, t_m), \end{aligned}$$

whence it is clear that $u_{m+1}^*(s_0) \leqq 0$. Choosing now the points t_i so that $a < t_0 < \cdots < t_m < s_0$, we similarly conclude that $u_{m+1}^*(s_0) \geqq 0$.

In order to prove that (iii) $\Rightarrow$ (i), assume also in this case that $\{u_0, \cdots, u_n\}$ is a CT-system. In view of the proof carried out in the preceding paragraph, it is clear that u_0^* is strictly positive throughout (a, b). By continuity we conclude that $\{u_0^*, \cdots, u_k^*\}$ is a WT-system on (a, b), for $k = 1, \cdots, n$, and the proof is carried out inductively, using (as was done in Theorem 1 to prove that the functions p_i are strictly increasing) the representations (4) and (5), and the Lemma.

ACKNOWLEDGEMENT. The author is greatly indebted to Dr. R. P. Kurshan of Bell Laboratories for his deep and useful comments concerning the organization of this paper, as well as for bringing to his attention a number of mistakes.

REFERENCES

1. M. W. Bartelt, *Weak Chebyshev Sets and Splines*, J. of Approximation Theory, **14** (1975), 30–37.

2. T. H. Hildebrandt, *Introduction to the Theory of Integration*; Academic Press, New York and London, 1963.

3. S. Karlin and W. Studden, *Tchebycheff Systems, with Applications in Analysis and Statistics*; Interscience, New York, 1966.

4. M. A. Rutman, *Integral Representation of Functions Forming a Markov Series*; Dokl. Akad. Nauk. SSSR **164** No. 5 (1965), 989–992 [English translation: Soviet Mathematics Doklady **6**, Amer. Math. Soc. (1965), 1340–1343].

5. H. S. Shapiro, *Smoothing and Approximation of Functions*, Van Nostrand Reinhold, New York, 1969.

6. R. A. Zalik, *Smoothness Properties of Generalized Convex Functions*, Proc. Amer. Math. Soc., **56** (1976), 118–120.

7. ———, *Existence of Tchebycheff extensions*; J. Math. Anal. Appl., **51** (1975), 68–75.

8. R. Zielke, *On Transforming a Tchebyshev-System into a Markov-System*, J. of Approximation Theory, **9** (1973), 357–366.

9. ———, *Alternation Properties of Tchebyshev-Systems and the Existence of Adjoined Functions*, J. of Approximation Theory, **10** (1974), 172–184.

Received January 7, 1975 and in revised form November 22, 1976.

BEN GURION UNIVERSITY OF THE NEGEV
BEERSHEVA 84 120, ISRAEL

Pacific Journal of Mathematics
Vol. 68, No. 2 April, 1977