# Pacific Journal of Mathematics

## IN THIS ISSUE—




**Vol. 48, No. 1** **March, 1973**

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

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# PSEUDO-COMPLETENESS AND THE PRODUCT OF BAIRE SPACES

J. M. AARTS AND D. J. LUTZER

**The class of pseudo-complete spaces defined by Oxtoby is one of the largest known classes $\mathscr{C}$ with the property that any member of $\mathscr{C}$ is a Baire space and $\mathscr{C}$ is closed under arbitrary products. Furthermore, all of the classical examples of Baire spaces belong to $\mathscr{C}$. In this paper it is proved that if $X \in \mathscr{C}$ and if $Y$ is any (quasi-regular) Baire space, then $X \times Y$ is a Baire space. The proof is based on the notion of $A$-embedding which makes it possible to recognize whether a dense subspace of a Baire space is a Baire space in its relative topology. Finally, examples are presented which relate pseudo-completeness to several other types of completeness.**

1. **Introduction.** A space $X$ is a *Baire space* if every nonempty open subset is of second category [2] or, equivalently, if the intersection of countably many dense open subsets of $X$ is dense in $X$. Locally compact Hausdorff spaces and completely metrizable spaces are the classical examples of Baire spaces.

In [10] Oxtoby introduced the notion of a pseudo-complete space (see §2 for precise definitions). Pseudo-complete spaces are Baire spaces and the classical examples of Baire spaces are pseudo-complete. Also, Čech-complete spaces (i.e., $G_\delta$-subsets of compact Hausdorff spaces [3]) as well as subcompact spaces [6] belong to the class of pseudo-complete spaces.

Pseudo-completeness has nice invariance properties. In particular, the topological product of any family of pseudo-complete spaces is pseudo-complete. Thus such a product is a Baire space.

In dealing with pseudo-completeness, assumptions about the usual separation axioms are irrelevant. However, it is often convenient to consider spaces which are *quasi-regular*, i.e., every nonempty open set contains the closure of some nonempty open set (cf. [10]).

Oxtoby [10] has also given an example of a completely regular Baire space whose square is not a Baire space, thus showing that a product theorem for Baire spaces cannot be obtained without some additional condition on (at least one of) the factors.

The main result of our paper is that the product of a quasi-regular Baire space and a pseudo-complete space is a Baire space. The techniques employed here, especially those in §4, are quite different from the usual category type techniques.

This paper is organized as follows. In §2 we discuss some new results on pseudo-completeness and its relation to some other completeness concepts. In §3 the notion of an $A$-embedded subset is introduced. It is shown that the property of being an $A$-embedded subset of a Baire space $Z$ is in a sense complementary to that of being a subset (of $Z$) which is a Baire space. Finally, in §4 the product theorem is proved.

We adopt the following notational convention: given a sequence $S_1, S_2, \cdots$ of subsets of a set $X$ we write $\bigcap \{S_n\}$ instead of $\bigcap \{S_n \mid n = 1, 2, 3, \cdots\}$.

2. **Pseudo-complete spaces.** In this section all spaces are assumed to be *quasi-regular* (see §1).

2.1. We begin with a brief review of results about pseudo-completeness from [10].

DEFINITIONS. A *pseudo-base for* $X$ is a collection $\mathscr{P}$ of nonempty open sets such that each nonempty open $U \subset X$ contains some member of $\mathscr{P}$. A space $X$ is *pseudo-complete* if $X$ has a sequence $\{\mathscr{P}(n)\}$ of pseudo-bases such that if $P_n \in \mathscr{P}(n)$ and $\operatorname{cl} P_{n+1} \subset P_n$ for each $n$, then $\bigcap \{P_n\} \neq \phi$.

THEOREM. (a) *Any dense $G_\delta$-subset of a countably compact space is pseudo-complete.*

(b) *Any product of pseudo-complete spaces is pseudo-complete.*

(c) *Any pseudo-complete space is a Baire space.*

In view of (a) the classical examples of Baire spaces as well as Čech-complete spaces are pseudo-complete. The pseudo-completeness of a subcompact space easily follows from the definitions (if $X$ is subcompact relative to the base $\mathscr{U}$, let $\mathscr{P}(n) = \mathscr{U}$[6]).

2.2. We now mention without proofs some simple facts about pseudo-completeness which are not in [10].

PROPOSITION. (a) *Let $X$ be a dense subspace of $Y$. If $X$ is pseudo-complete, then so is $Y$.*

(b) *Any topological sum (i.e., disjoint union) of pseudo-complete spaces is pseudo-complete.*

(c) *Any open subspace of a pseudo-complete space is pseudo-complete.*

It is easily seen that closed subspaces of pseudo-complete spaces need not be pseudo-complete. For example, Michael's line—the set of real numbers topologized by taking sets of the form $U \cup V$ to be open, where $U$ is open in the usual topology of the reals and $V$ is any set of irrationals—is pseudo-complete (in view of (a)) and yet contains the usual space of rational numbers as a closed subspace.

Observe that the proposition above remains valid if "pseudo-complete space" is replaced by "Baire space". The following corollaries hold for any property of topological spaces for which the corresponding propositions (a), (b), and (c) hold.

COROLLARY 1. *If a space $X$ has an open almost-cover (i.e., a collection of open sets whose union is dense in $X$, cf.* [4]) *by pseudo-complete spaces, then $X$ is pseudo-complete.*

COROLLARY 2. *A space is locally pseudo-complete (i.e., each point has a neighborhood which is pseudo-complete) if and only if it is pseudo-complete.*

2.3. The following result is related to a theorem of McCoy about Baire space extensions [9].

THEOREM. *Any (quasi-regular) space $X$ is a dense subspace of a pseudo-complete space $\tilde{X}$.*

*Proof.* The proof uses standard techniques; we present only an outline. The main difference between our construction and the standard ones is that we do not identify points of $X$ with (open) ultrafilters (since this would require separation axioms of $X$). Let $\mathscr{T}$ denote the topology of $X$. Let $\Delta$ denote the collection of all subfamilies of $\mathscr{T}$ which have the finite intersection property and are maximal with respect to this property. Let $\Omega = \{\xi \mid \xi \in \Delta$ and $\bigcap \{\mathrm{cl}_X F \mid F \in \xi\} = \phi\}$, the free elements of $\Delta$. Let $\tilde{X} = X \cup \Omega$ (observe that the union is disjoint). For each $U \in \mathscr{T}$ let $\tilde{U} = U \cup \{\xi \mid \xi \in \Omega$ and $U \in \xi\}$. $\tilde{X}$ is endowed with the topology for which $\{\tilde{U} \mid U \in \mathscr{T}\} = \tilde{\mathscr{T}}$ serves as a base. It is easily verified that any collection of open sets of $\tilde{X}$ having the finite intersection property has nonempty adherence in $\tilde{X}$. From $\mathrm{cl}_{\tilde{X}}(\tilde{U}) = \tilde{U} \cup \mathrm{cl}_X U$ it follows that $\tilde{X}$ is quasi-regular (since $X$ is assumed to be quasi-regular).
By letting $\mathscr{P}(n) = \tilde{\mathscr{T}}$ for $n \geqq 1$, $\tilde{X}$ is shown to be pseudo-complete. In case $X$ is Hausdorff, $\tilde{X}$ is a Hausdorff-closed extension of $X$ (cf. [7]).

2.4. As is well-known, the open continuous image of a Baire space is a Baire space. It is an open problem whether such mappings also preserve pseudo-completeness. However, if the range space is assumed to be metrizable, there is the following result.

THEOREM. *Suppose $X$ is pseudo-complete and $f: X \to Y$ is continuous, onto and open. If $Y$ is metrizable, then $Y$ contains a dense, completely metrizable, zero-dimensional subspace. In particular, $Y$ is pseudo-complete.*

*Proof*. Let $\{\mathscr{P}(n)\}$ be a sequence of pseudo-bases for $X$ with respect to which $X$ is pseudo-complete. Inductively choose collections $\mathscr{P}'(n) \subset \mathscr{P}(n)$ such that for each $n$:

(a) If $P$ and $Q$ are distinct members of $\mathscr{P}'(n)$, then $f(P) \neq f(Q)$;

(b) The collection $\mathscr{G}(n) = \{f(P) \mid P \in \mathscr{P}'(n)\}$ is a disjoint open collection in $Y$ which is an almost cover[1] of $Y$ with mesh $\leqq 1/n$;

(c) $\mathscr{P}'(n+1)$ refines $\mathscr{P}'(n)$;

(d) If $P_{n+1} \in \mathscr{P}'(n+1)$, $P_n \in \mathscr{P}'(n)$ and $f(P_{n+1}) \subset f(P_n)$, then $\mathrm{cl}_X P_{n+1} \subset P_n$.

Suppose $G_n \in \mathscr{G}(n)$ for $n \geqq 1$ and $G_n \supset G_{n+1}$. Because the collection $\mathscr{G}(n)$ is disjoint, it follows from (a) and (c) that there are unique members $P_n$ and $P_{n+1}$ of $\mathscr{P}'(n)$ and $\mathscr{P}'(n+1)$ respectively having $f(P_n) = G_n$ and $f(P_{n+1}) = G_{n+1}$. According to (d) we have $P_n \supset \mathrm{cl}_X P_{n+1}$. Therefore, $\bigcap\{P_n\} \neq \phi$. For any $x \in \bigcap\{P_n\}$, $f(x) \in \bigcap\{G_n\}$. In view of (b) $\{f(x)\} = \bigcap\{G_n\}$. For each $n$ let $W_n = \bigcup \mathscr{G}(n)$, and let $Z = \bigcap\{W_n\}$. From the preceding observation it follows that $Z$ is a dense subset of $Y$. By virtue of (b), the collection $\mathscr{B} = \{G \cap Z \mid G \in \mathscr{G}(n), n = 1, 2, \cdots\}$ is a base for $Z$ consisting of relatively closed and open sets. The completeness of $Z$ is easily proved (cf. [1], [4], [6]).

COROLLARY. *A metrizable space $Y$ is pseudo-complete if and only if $Y$ contains a dense completely metrizable subspace (which may be taken to be zero-dimensional).*

*Proof*. The "only if" part follows from the preceding theorem by taking $X = Y$. The "if" part follows from 2.1 Theorem (a) and 2.2 Proposition (a).

REMARK. F. G. Slaughter, Jr. pointed out to the authors that in the proofs of the theorem and corollary above it suffices to assume that $Y$ is quasi-regular and developable (instead of metrizable). Thus we have: a pseudo-complete Moore space contains a dense completely metrizable subspace.

The corollary above may be applied to show that pseudo-completeness and the property of being a Baire space are not equivalent, even

[1] Cf. 2.2 Corollary 1.

for metrizable spaces.

EXAMPLE. A space is said to be *totally imperfect* if it contains no uncountable compact subsets. According to a theorem of Bernstein, any separable, metrizable, complete and dense-in-itself space $X$ can be decomposed into two mutually disjoint totally imperfect subsets $Y$ and $Z$. Both $Y$ and $Z$ are known to be Baire spaces, but neither $Y$ nor $Z$ contains a dense completely metrizable subspace. See 2.6, 3.3 Example, and [8] for more details.

2.5. PROBLEM. It is an open problem whether dense $G_\delta$-subsets of pseudo-complete spaces are pseudo-complete. Observe that if dense $G_\delta$-subsets of pseudo-complete spaces and open continuous images of pseudo-complete spaces (see 2.4) are pseudo-complete, a less complicated proof of the Product Theorem 4.2 can be given.

2.6. A part of the theory we are going to present in §§3 and 4 can be developed using other notions of "completeness" instead of pseudo-completeness. For example, we can define a completely regular space $X$ to be *almost Čech-complete* if $X$ contains a dense Čech-complete subspace (cf. [5]). The results of §2.2 also hold for almost Čech-complete spaces and in view of 2.1 Theorem (a) such spaces are pseudo-complete, whence Baire spaces.

We next present an example which shows that the notion of pseudo-completeness is much more general than that of being almost Čech-complete.

EXAMPLE. There is a completely regular, pseudo-complete space $X$ such that each Čech-complete subspace $C$ of $X$ is nowhere dense (i.e., $\operatorname{int}_X \operatorname{cl}_X C = \phi$).

Let $Z = R^c = \Pi\{R_\gamma \mid \gamma \in \Gamma\}$, the continuous product of real lines. Any nonempty $G_\delta$-subset of $Z$ has $\exp c$ points[2] and there are $\exp c$ nonempty $G_\delta$-subsets of $Z$. We shall show that there is a subset $X$ of $Z$ such that neither $X$ nor $Z \backslash X$ contains any nonempty $G_\delta$-subset of $Z$ and that $X$ has the above mentioned properties.

The construction of the set $X$ is very similar to that of a totally imperfect subset of a separable, metrizable, and dense-in-itself space as given in [8]. Well-order the collection of all nonempty $G_\delta$-subsets of $Z$ as $\{H_\alpha \mid \alpha < \eta\}$ where $\eta$ is the first ordinal having cardinality $\exp c$. Using transfinite induction pick two distinct points $x_\alpha$ and $y_\alpha$ from each $H_\alpha$, taking care that at each step only points are picked

[2] We write $\exp c$ for $2^c$.

which have not been selected before. $X$ is the set of all points $x_\alpha$ so obtained.

Now let $\mathscr{B}$ denote the collection of the basic open sets $\bigcap\{\pi_\gamma^{-1}(U_\gamma)\mid\gamma\in\Gamma_0\}$ where $\Gamma_0\subset\Gamma$ is *finite* and each $U_\gamma$ is a *bounded* open interval in $R_\gamma$. Let $\mathscr{P}(n)=\{B\cap X\mid B\in\mathscr{B}\}$ for each $n\geqq 1$. Observe that for each $B\in\mathscr{B}$, $\mathrm{cl}_X(B\cap X)=X\cap\mathrm{cl}_Z B$, because $X$ is dense in $Z$.

In order to prove that $X$ is pseudo-complete, suppose $P_n\in\mathscr{B}$ (so $P_n\cap X\in\mathscr{P}(n)$) and $\mathrm{cl}_X(P_{n+1}\cap X)\subset P_n\cap X$. Then

$$\bigcap\{X\cap P_n\}=\bigcap\{\mathrm{cl}_X(X\cap P_n)\}=X\cap(\bigcap\{\mathrm{cl}_Z P_n\}) .$$

Now, $\bigcap\{\mathrm{cl}_Z P_n\}$ is nonempty because of the boundedness condition imposed on the basic open sets. Moreover, this intersection is a $G_\delta$-subset of $Z$. Hence its intersection with $X$ is nonempty. Thus $X$ is pseudo-complete.

If $D$ is a Čech-complete space which is dense in some open set $U$ of $X$, then $D$ is also dense in some open set $V$ of $Z$. Then $D\cap V$ is a $G_\delta$-subset of $Z$, so $(D\cap V)\cap(Z\backslash X)\neq\phi$. Hence $D$ is not contained in $X$.

## 3. $A$-embedded subsets.

We shall now define the concept of an $A$-embedded subset, which plays a vital role in the proof of the Product Theorem.

3.1. DEFINITION. Let $X$ be a subset of a space $Z$. Then $X$ is said to be *$A$-embedded in $Z$* if each $G_\delta$-subset $H$ of $Z$ which is contained in $X$ is nowhere dense in $X$ (i.e., $\mathrm{int}_X\,\mathrm{cl}_X H=\phi$).

EXAMPLE 1. The set $Q$ of the rational numbers is $A$-embedded in the real line $R$.

To prove this we first observe that in a countable $T_1$-space which is a Baire space, each open set has isolated points (since the complement of a non-isolated point is a dense open set). Now, let $H$ be a $G_\delta$-subset of $R$. If $H\subset Q$, then $H$ is countable, so that $H$ cannot be dense in any interval. It follows that $H$ is nowhere dense in $Q$.

EXAMPLE 2. Let the subspace $S$ of $R$ be defined by $S=\{0\}\cup\{1/n\mid n=1,2,\cdots\}$. The set $\{0\}$ is not $A$-embedded in $S$.

Using the concept of $A$-embedded subsets, we are able to recognize subsets of Baire spaces which, in their relative topology, are Baire spaces.

3.2. THEOREM. *Let $X$ be a dense subspace of a Baire space $Z$.*
(a) *If $Z\backslash X$ is $A$-embedded in $Z$, then $X$ is a Baire space.*

(b) *If* $Z\backslash X$ *is dense in* $Z$, *then* $X$ *is a Baire space if and only if* $Z\backslash X$ *is A-embedded in* $Z$.

*Proof*. To prove (a), observe that if $X$ is not a Baire space, then there is a sequence $H \supset G_1 \supset G_2 \supset \cdots$ of open subsets of $X$ such that each $G_n$ is dense in $H$ and yet $\bigcap \{G_n\} = \phi$. Then there is a sequence $U \supset V_1 \supset V_2 \supset \cdots$ of open subsets of $Z$ such that $H = U \cap X$ and $G_n = V_n \cap X$. Each $V_n$ is dense in $U$ and $U$ is a Baire space. Hence $D = \bigcap \{V_n\}$ is dense in $U$ and therefore in $U \cap (Z\backslash X)$. Since $D \subset Z\backslash X$, $Z\backslash X$ is not $A$-embedded in $Z$.

To prove the "only if" part of (b), we assume that $Z\backslash X$ is not $A$-embedded in $Z$. Let $H$ be a $G_\delta$-subset of $Z$ which is contained in $Z\backslash X$ and which is dense in some relatively open set $U$ of $Z\backslash X$. Let $V$ be an open subset of $Z$ with $V \cap (Z\backslash X) = U$.

Then $F = V \cap H$ is a $G_\delta$-subset of $Z$ which is dense in $V$ and which is contained in $U$. Let $F = \bigcap \{F_n\}$ where each $F_n$ is open in $Z$ and $F_n \subset V$. The sets $F_n \cap X$ are open and dense subsets of $V \cap X$ and yet $\bigcap \{F_n \cap X\} = \phi$. It follows that $V \cap X$ is not a Baire space. Consequently, $X$ is not a Baire space.

REMARK. As is clear from 3.1 Example 2, the hypothesis that $Z\backslash X$ is dense in $Z$ cannot be omitted from part (b) above.

3.3. In connection with Theorem 3.2 we have the following propositions which will be used in the next section.

LEMMA 1. *In every quasi-regular space* $X$ *there are open subspaces* $X_P$ *and* $X_A$ *such that*

(a) $X_P$ *and* $X_A$ *are disjoint and* $X_P \cup X_A$ *is dense in* $X$;

(b) $X_P$ *is pseudo-complete*;

(c) *any pseudo-complete subspace of* $X_A$ *is nowhere dense in* $X_A$.

*Proof*. Let $X_P$ be the union of all open subsets of $X$ which are pseudo-complete in their relative topology.

Let $X_A = X\backslash \mathrm{cl}_X X_P$. The lemma follows from 2.2 Proposition and 2.2 Corollary 1. (Observe that $X_P$ and $X_A$ may be empty.)

LEMMA 2. *Let* $X$, $X_P$ *and* $X_A$ *be as in Lemma* 1. *Then* $X$ *is a Baire space if and only if* $X_A$ *is a Baire space.*

*Proof*. Obvious.

PROPOSITION. *Let* $X$ *be a quasi-regular space such that any pseudo-complete subspace of* $X$ *is nowhere dense in* $X$. *The following properties* *are equivalent*:

(a) *$X$ is a Baire space;*

(b) *For every pseudo-complete space $Y$ such that $X$ is dense in $Y$, the subset $Y \backslash X$ is A-embedded in $Y$;*

(c) *For some pseudo-complete space $Y$ such that $X$ is dense in $Y$, the set $Y \backslash X$ is A-embedded in $Y$.*

*Proof*. For any pseudo-complete space $Y$ such that $X$ is dense in $Y$, the set $Y \backslash X$ is dense in $Y$, because $X$ contains no pseudo-complete open subspaces. That (a) implies (b) now follows from 3.2 Theorem (b). Obviously (c) follows from (b). Finally, 3.2 Theorem (a) shows that (a) follows from (c).

EXAMPLE. According to the above proposition, the totally imperfect subspaces $Y$ and $Z$ of a separable, complete, dense-in-itself metric space $X$ in 2.4 Example are Baire spaces. The proposition can also be applied to spaces which are not totally imperfect. In the special case where $X$ is the Euclidean plane, construct $Y$ and $Z$ as in 2.4 Example and let $Y' = Y \cup L$ where $L$ is a straight line in $X$. Then $Y'$ and $X \backslash Y'$ are each $A$-embedded subspaces of $X$ (so that each is a Baire space) even though $Y'$ is not totally imperfect.

REMARK. It follows directly from Lemma 2 above that a $\sigma$-locally compact space $X$ is a Baire space if and only if $X$ is pseudo-complete: one shows that the $\sigma$-locally compact Baire space $X_A$ must be empty.

## 4. The product theorem.

4.1. PROJECTION LEMMA. *Suppose $X$ is a Baire space and $Y$ is pseudo-complete. If $D$ is a dense $G_\delta$-subset of $X \times Y$, then $\pi_X(D)$ contains a dense $G_\delta$-subset of $X$ (where $\pi_X$ denotes the natural projection of $X \times Y$ onto $X$).*

*Proof*. Write $D = \bigcap \{G_n\}$ where $G_1 \supset G_2 \supset \cdots$ are open subsets of $X \times Y$. Let $\{\mathscr{P}(n)\}$ be a sequence of pseudo-bases for $Y$ with respect to which $Y$ is pseudo-complete.

Let $\mathscr{U}(1) = \{U \mid \phi \neq U$ is open in $X$ and $U \times P \subset G_1$ for some $P \in \mathscr{P}(1)\}$. Let $\mathscr{V}(1)$ be a maximal disjoint subcollection of $\mathscr{U}(1)$. Since $D$ is dense in $X \times Y$, the set $W_1 = \bigcup \mathscr{V}(1)$ is dense in $X$. For each $V \in \mathscr{V}(1)$ choose $P(V, 1) \in \mathscr{P}(1)$ such that $V \times P(V, 1) \subset G_1$.

Inductively define collections $\mathscr{V}(n)$ such that

(a) each $\mathscr{V}(n)$ is a disjoint collection of nonempty open subsets of $X$ and $W_n = \bigcup \mathscr{V}(n)$ is dense in $X$;

(b) $\mathscr{V}(n+1)$ refines $\mathscr{V}(n)$;

(c) for each $V \in \mathscr{V}(n)$ there is a member $P(V, n) \in \mathscr{P}(n)$ having $V \times P(V, n) \subset G_n$;

(d) if $V_{n+1} \in \mathscr{V}(n+1)$ and if $V_n$ is the unique member of $\mathscr{V}(n)$ containing $V_{n+1}$, then $\mathrm{cl}_Y(P(V_{n+1}, n+1)) \subset P(V_n, n)$.

Let $E = \bigcap \{W_n\}$. Since $X$ is a Baire space, $E$ is dense in $X$. To show that $E \subset \pi_X(D)$, let $x \in E$. Then there is a (unique) sequence $\{V_n\}$ such that $V_n \in \mathscr{V}(n)$ and $x \in V_n$. Necessarily $V_{n+1} \subset V_n$ so that $\mathrm{cl}_Y(P(V_{n+1}, n+1)) \subset P(V_n, n)$. Therefore, $\bigcap \{P(V_n, n)\} \neq \phi$ and for any $y \in \bigcap \{P(V_n, n)\}$, $(x, y) \in G_n$ for each $n$ so that $(x, y) \in D$. Hence $x \in \pi_X(D)$.

4.2. THE PRODUCT THEOREM. *If $X$ is a quasi-regular Baire space and $Y$ is pseudo-complete, then $X \times Y$ is a Baire space.*

*Proof.* As in 3.3 Lemma 1, choose open subsets $X_P$ and $X_A$ of $X$. Then $(X_P \times Y) \cup (X_A \times Y)$ is a dense subspace of $X \times Y$. Since $Y_P \times Y$ and $X_A \times Y$ are open subsets of $X \times Y$ and since $X_P \times Y$ is a Baire space in view of 2.1 Theorem (b) and (c), it is enough to show that $X_A \times Y$ is a Baire space. By 3.3 Lemma 2, $X_A$ is a (quasi-regular) Baire space. Thus we need only consider the special case where $X = X_A$.

In view of 2.3 Theorem, there is a pseudo-complete quasi-regular space $\tilde{X}$ which contains $X$ as a dense subspace. Let $X' = \tilde{X} \backslash X$. Then $X'$ is $A$-embedded in $\tilde{X}$ by virtue of 3.3 Proposition. Since $X \times Y$ is a dense subset of the pseudo-complete space $\tilde{X} \times Y$, it will be sufficient to show that $X' \times Y = (\tilde{X} \times Y) \backslash (X \times Y)$ is an $A$-embedded subset of $\tilde{X} \times Y$ (by virtue of 3.2 Theorem).

To this end, suppose $D \subset X' \times Y$ is a $G_\delta$-subset of $\tilde{X} \times Y$ which is dense in some relatively open subset of $X' \times Y$. Then there is a relatively open subset $G$ of $X'$ and an open subset $V$ of $Y$ such that $D \cap (G \times V)$ is dense in $G \times V$. The set $G$ may be written as $G = U \cap X'$ where $U$ is open in $\tilde{X}$. Then $D \cap (U \times V)$ is a dense $G_\delta$-subset of $U \times V$, the product of the Baire space $U$ with the pseudo-complete space $V$. According to the projection lemma, $\pi_U(D \cap (U \times V))$ contains a $G_\delta$-subset $E$ of $U$ which is dense in $U$. Then $E$ is a $G_\delta$-subset of $\tilde{X}$. Furthermore, $E \subset \pi_{\tilde{X}}(D) \subset X'$ which is impossible because $X'$ is known to be $A$-embedded in $\tilde{X}$.

It follows that $X' \times Y$ is $A$-embedded in $\tilde{X} \times Y$ so that $X \times Y$ is a Baire space.

COROLLARY. *If $X$ is a quasi-regular Baire space and if the space $Y$ is (locally) compact Hausdorff or (locally) Čech-complete or (locally) subcompact, then $X \times Y$ is a Baire space.*

REMARK. In the Product Theorem the pseudo-completeness of $Y$ is by no means a necessary condition for $X \times Y$ to be a Baire space. Oxtoby [10] has proved that if $Y$ is any Baire space having a (locally) countable pseudo-base, then $X \times Y$ is a Baire space for any Baire space $X$. This result of Oxtoby and our product theorem overlap, but neither includes the other: (Cf. the examples in 2.4 and 2.6 of which the latter has no (locally) countable pseudo-base. Indeed there are compact Hausdorff spaces having no locally countable pseudo-base.) Furthermore, as has already been mentioned in the Introduction, the techniques employed in the proofs are totally different.

## REFERENCES

1. J. M. Aarts, J. de Groot, and R. H. McDowell, *Cotopology for metrizable spaces*, Duke Math. J., **37** (1970), 291-295.

2. N. Bourbaki, *Elements of Mathematics: General Topology*, part 2, Addison-Wesley, Reading, Mass., 1966.

3. E. Čech, *On bicompact spaces*, Ann. of Math., **38** (1937), 823-844.

4. Z. Frolík, *Generalizations of the $G_\delta$-property of complete metric spaces*, Czech Math. J., **10** (1960), 359-379.

5. ———, *Locally topologically complete spaces*, Sov. Math. Dokl., **2** (1961), 355-357.

6. J. de Groot, *Subcompactness and the Baire category theorem*, Indag. Math., **25** (1963), 761-767.

7. M. Katětov, *Über H-abgeschlossene und bikompakte Räume*, Časopis Pěst. Mat. Fys., **69** (1940), 36-49.

8. C. Kuratowski, *Topology*, Vol. 1, Academic Press, New York, 1966.

9. R. McCoy, *A Baire space extension*, Proc. Amer. Math. Soc., **33** (1972), 199-202.

10. J. C. Oxtoby, *Cartesian products of Baire spaces*, Fund. Math., **49** (1961), 157-166.

Received June 30, 1972 and in revised form August 14, 1972. Partially supported by NSF Grant GP-29401.

DELFT INSTITUTE OF TECHNOLOGY, DELFT, NETHERLANDS
AND
UNIVERSITY OF PITTSBURGH

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# METRIC CHARACTERIZATIONS OF EUCLIDEAN SPACES

G. BERG

**In a metric space an arc which is isometric to a real interval is called a segment. In this paper it is shown that, for $1 \leqq n \leqq 3$, $n$-dimensional Euclidean space ($E^n$) is topologically characterized, among locally compact, $n$-dimensional spaces, by admitting a metric with the following properties: (1) every two points of the space are endpoints of a unique segment, (2) if two segments have an endpoint and one other point in common then one is contained in the other and (3) every segment can be extended, at either end, to a larger segment. This follows from the more general result that, for $1 \leqq n \leqq 3$, a locally compact, $n$-dimensional space which admits a metric with properties (1) and (2) is homeomorphic to an $n$-manifold lying between the closed $n$-ball and its interior.**

**Property (1) suffices to characterize $E^n$, for $n = 1$ or 2, among locally compact, locally homogeneous, $n$-dimensional spaces. For $n > 3$, properties (1), (2), and (3) characterize $E_n$ among locally compact, $n$-dimensional spaces that contain a homeomorph of an $n$-ball.**

1. Introduction. A metric space $(X, d)$ is said to be *convex* provided that every pair of points of $X$ has a midpoint—$m$ is a *midpoint* of $x$ and $y$ if $d(x, m) = d(y, m) = 1/2d(x, y)$. $(X, d)$ is *strongly convex* if every pair of points has a unique midpoint and is *without ramifications* provided that no midpoint of $x$ and $y$ is a midpoint of $z$ and $y$ unless $z = x$. Convex subsets of Euclidean spaces with their inherited metrics are examples of metric spaces with these properties. Lelek and Nitka [8] and Rolfsen [11] have (topologically) characterized the 2-cell and 3-cell, respectively, among compact 2 and 3 dimensional metric spaces by the last two properties. White [12] has shown that a 2-complex is collapsible if and only if it can be given a metric which is strongly convex. Numerous other results have been obtained for metric spaces with the above properties when the underlying space is compact or when the metric is also complete.

In the present paper a number of these results are shown to hold when the underlying space is locally compact. Principally, it is shown that having a strongly convex metric without ramifications (topologically) characterizes $n$-manifolds that lie between the $n$-cell and its interior among locally compact, $n$-dimensional spaces for $n \leqq 3$. This reduces to Lelek and Nitka's or Rolfsen's result when the space is compact and yields a characterization of $E^n$ under various homogeneity conditions.

2. **Existence of segments.** In a metric space $(X, d)$ a set $S$ is said to be a *segment* for the points $x$ and $y$ of $X$ if $x$ and $y$ are elements of $S$ and $S$ is isometric with the real interval $[0, d(x, y)]$. It is well known that in a convex, complete metric space every pair of points has a segment between them. It is shown now that for locally compact spaces the requirement of completeness can be relaxed.

THEOREM 2.1. *Let $(X, d)$ be a locally compact, convex metric space. If, for each pair of points of $X$, the set of midpoints of the pair is compact, then each pair of points is joined by a segment.*

*Proof.* Let $p, q$ be two points of $X$ and let $A_0 = \{p, q\}$. Order $A_0$ by distance from $p$. In general, if $A_n$ has been defined and ordered by distance from $p$, then for each $x \in (A_n - \{q\})$ let $m_x$ be a midpoint for $x$ and the next point of $A_n$. Define $A_{n+1} = A_n \cup \{m_x : x \in A_n - \{q\}\}$. Let $A = \bigcup_{n=0}^{\infty} A_n$, $a = d(p, q)$ and $f = d(p, -) \,|\, \bar{A}$. Clearly $f$ maps $A$ isometrically onto the set of real numbers of the form $r \cdot a$ where $r$ is a dyadic rational in $[0, 1]$, and so $f$ maps $\bar{A}$ isometrically into the interval $[0, a]$. To show that $\bar{A}$ is a segment from $p$ to $q$ it is sufficient to show that $f[\bar{A}] = [0, a]$.

Since the dyadic rationals are dense in $[0, 1]$, the image of $A$ is dense in $[0, a]$. Also, the image of $\bar{A}$ is open in $(0, a)$. To verify this observe that if $x \in (\bar{A} - \{p, q\})$ and $D$ is a compact distance neighborhood of $x$ then $D \cap \bar{A}$ is compact and thus $f[\bar{A} \cap D]$ is closed. Since $f[A \cap D]$ is dense in an interval containing $f(x)$ in its interior, $f(x)$ is an interior point of $f[\bar{A}]$. For the final step let $t$ be any point of $(0, a)$, and let $T$ be a subinterval of $(0, a)$, symmetric about $t$. $T \cap f[\bar{A}]$ and $U$, the reflection of $T \cap f[\bar{A}]$ in $t$, are both open and dense in $T$ and hence their intersection is dense in $T$. Let $\{t_n\}_{n=1}^{\infty}$ be an increasing sequence of points of $T \cap f[\bar{A}] \cap U$ that converges to $t$, so if, for each $n$, $t'_n = 2t - t_n$, then $t'_n \in f[\bar{A}]$ and $t$ is midway between $t_n$ and $t'_n$. Let $x_n = f^{-1}(t_n)$, $x'_n = f^{-1}(t'_n)$ and $M_n = \{y \in X : y$ is midpoint for $x_n$ and $x'_n\}$. If $y \in M_{n+1}$, then $y$ is between $x_n$ and $x'_n$ and also $d(x_n, y) = d(x_n, x_{n+1}) + d(x_{n+1}, y) = d(y, x'_{n+1}) + d(x'_{n+1}, x'_n) = d(y, x'_n)$ so $y \in M_n$. Since each $M_n$ is compact there is a point $x$ in $\bigcap_{n=1}^{\infty} M_n$. Now $d(x_n, x) = 1/2d(x_n, x'_n) = 1/2 \,|\, t_n - t'_n \,|$ and so $\lim_{n\to\infty} d(x_n, x) = 0$. Thus $x \in \bar{A}$ and clearly $f(x) = t$. Evidently $f[\bar{A}] = [0, a]$ and so $\bar{A}$ is a segment from $p$ to $q$.

COROLLARY 2.2. *If $(X, d)$ is a locally compact, strongly convex metric space, then the segment between two points is unique and contains* *all the points between them.*

The proof for the case where $(X, d)$ is also complete carries over without change in light of 2.1.

When segments are unique the segment between $p$ and $q$ is denoted $\overline{pq}$.

## 3. Strongly convex metrics.

DEFINITION 3.1. A topological space $Y$ is said to be *contractible* if there exists a mapping $f: Y \times I \to Y$ such that $f(-, 1)$ is the identity map and $f(-, 0)$ is constant. $Y$ is *locally contractible* if every neighborhood, $U$, of any point contains a neighborhood, $V$, of the point and a map $f: V \times I \to U$ such that $f(-, 1)$ is the identity map and $f(-, 0)$ is constant.

Throughout this section let $(X, d)$ be a locally compact strongly convex metric space. We aim first at showing that $X$ is contractible and locally contractible.

Fix $p \in X$ and define the map $Q: X \times I \to X$ as follows: for $t \in I$, $Q(p, t) = p$ for $t \in I, x \in X - \{p\}$, $Q\{x, t)$ is the point $z$ in $\overline{px}$ such that $d(p, z) = t \cdot d(p, x)$.

$Q$ is the contracting homotopy and in showing $Q$ is continuous the fact that the limit of a sequence of segments is again a segment is used. This fact is the content of 3.3.

NOTATION 3.2. Throughout the paper $D(p, r)$ denotes the set $\{x \in X: d(p, x) \leqq r\}$ and $S(p, r)$ denotes the set $\{x \in X: d(p, x) < r\}$ where $p \in X$ and $r$ is a positive real number.

PROPOSITION 3.3. *Let $x_0, x_1, x_2, \cdots$ be points of $X$ such that $\lim_{i \to \infty} x_i = x_0 \neq p$. Then if*

(1) $y_i \in \overline{px_i}$, $i = 0, 1, 2, \cdots$ *and*

(2) $\lim_{i \to \infty} d(p, y_i) = d(p, y_0)$ *then*

(3) $\lim_{i \to \infty} y_i = y_0$.

*Proof.* Let $A = \{y_0 \in \overline{px_0}$: conditions (1) and (2) imply (3)$\}$. Clearly $x_0 \in A$ because if $\lim_{i \to \infty} d(p, y_i) = d(p, x_0)$, then $\lim_{i \to \infty} d(x_i, y_i) = 0$ since $d(x_i, y_i) = d(p, x_i) - d(p, y_i)$. Let $A'$ be the component of $A$ containing $x_0$ and assume that $q$ is a boundary point of $A'$ relative to $\overline{px_0}$. Choose $r > 0$ so that $D(q, 5r)$ is compact. Let $y_0, y_1, y_2, \cdots$ be a sequence that satisfies conditions (1) and (2) but not (3) and chosen such that $d(y_0, q) < r$. Let $q_0 \in A \cap D(q, r)$. For each $i = 1, 2, \cdots$ let $t_i = \min\{d(p, x_i), d(p, q_0)\}$ and let $q_i$ be a point of $\overline{px_i}$ such that $d(p, q_i) = t_i$, $i = 1, 2, \cdots$. Since $\lim_{i \to \infty} t_i = d(p, q_0)$ it follows that $\lim_{i \to \infty} q_i = q_0$.

Since both of $y_i$, and $q_i$ belong to $\overline{px_i}$, $i = 0, 1, 2, \cdots$ we have $d(q_i, y_i) = |d(p, q_i) - d(p, y_i)|$ and so $\lim_{i\to\infty} d(q_i, y_i) = d(q_0, y_0) \leqq 2r$. Now $d(q, y_i) \leqq d(q, q_0) + d(q_0, q_i) + d(q_i, y_i)$ and so is eventually less than $5r$. Thus the sequence $\{y_i\}_{i=1}^{\infty}$ is eventually in the compact set $D(q, 5r)$. If $z$ is a limit point of the sequence $\{y_i\}_{i=1}^{\infty}$, then it follows from the continuity of the distance function $d$ that $d(p, z) + d(z, x_0) = d(p, x_0)$ and so $z \in \overline{px_0}$. Since $z$ is the same distance from $p$ as $y_0$ it follows that $z = y_0$. Thus $\lim_{i\to\infty} y_i = y_0$ and this contradicts our choice of $y_0$, $y_1, y_2, \cdots$. It follows that $A'$ has no boundary point relative to $\overline{px_0}$ so must be all of $\overline{px_0}$.

PROPOSITION 3.4. *$Q: X \times I \to X$ is continuous.*

*Proof.* Follows from Proposition 3.3 and continuity of distance function.

In contracting $X$ to the point $p$ the map, $Q$, moves every point closer to $p$ so any distance neighborhood of $p$ is contracted in itself. The point $p$ was chosen without restriction so $X$ is also locally contractible.

It follows trivially that $X$ is also connected and locally connected and these conditions for a locally compact metric space imply separability [1].

THEOREM 3.5. *A locally compact, strongly convex metric space is contractible, locally contractible, connected, locally connected and separable.*

THEOREM 3.6. *An $n$-dimensional, locally compact, strongly convex metric space is an $n$-manifold if it is locally homogeneous and contains an $n$-ball.*

*Proof.* Since a locally compact space is second category this follows immediately from a theorem of Bing and Borsuk [3].

THEOREM 3.7. *For $n = 1$ or 2, an $n$-dimensional, locally homogeneous, locally compact metrizable space can be given a strongly convex metric if and only if it is homeomorphic to $E^n$.*

*Proof.* The usual metric for $E^n$ is strongly convex.

It follows from another theorem of Bing and Borsuk [3] that such a space is an $n$-manifold. Since it is contractible as well it must be $E^n$, $n$ being 1 or 2.

**4. Strongly convex metrics without ramifications.** Preliminaries. Throughout this section $(X, d)$ will be a locally compact metric space with $d$ a strongly convex metric without ramifications, briefly an SC-WR metric.

DEFINITION 4.1. For $p$ and $q$ two points of $X$ the set $\{x \in X: x \in \overline{pq}$ or $q \in \overline{px}\}$ is called the *ray* from $p$ through $q$ and is denoted $\overline{pq}\rangle$.

PROPOSITION 4.2. *If* $y \in (\overline{px}\rangle - \{p\})$, *then* $\overline{py}\rangle = \overline{px}\rangle$.

*Proof.* Clearly $x \in \overline{py}\rangle$ so it suffices to show that if $z \in \overline{px}\rangle$ then $z \in \overline{py}\rangle$. We consider four cases:

(1) $y \in \overline{px}$ and $z \in \overline{px}$. In this case it follows immediately from the uniqueness of segments that either $y \in \overline{pz}$ or $z \in \overline{py}$ so $z \in \overline{py}\rangle$.

(2) $y \in \overline{px}$ and $x \in \overline{pz}$. The convexity of the metric yields $y \in \overline{pz}$ and so $z \in \overline{py}\rangle$.

(3) $x \in \overline{py}$ and $z \in \overline{px}$. Same as (2).

(4) $x \in \overline{py}$ and $x \in \overline{pz}$. Unless $\overline{py} \subset \overline{pz}$ or $\overline{pz} \subset \overline{py}$ there would be a ramification point in $\overline{pz} \cap \overline{py}$. Thus $z \in \overline{py}\rangle$.

PROPOSITION 4.3. $\overline{px}\rangle$ *is isometric to a real interval of one of the following forms*: $[0, \infty)$, $[0, a)$, *or* $[0, a]$.

*Proof.* This is evident from the previous proposition and the fact that rays are arc connected.

If $\overline{px}\rangle$ is isometric to the closed interval $[0, a]$, then we say $\overline{px}\rangle$ is a ray with endpoint or a compact ray and the point of $\overline{px}\rangle$ a distance $a$ from $p$ is the endpoint.

DEFINITION 4.4. A metric space $(Y, \rho)$ is said to be *externally convex* if given $p$ and $q$ in $Y$ there is a point $y \in Y$ such that $\rho(p, y) = \rho(p, q) + \rho(q, y)$.

Note that $(X, d)$ is externally convex if and only if no ray has an endpoint.

For rays a result analogous to Proposition 3.3 holds and the proof carries over as well.

PROPOSITION 4.5. *Let* $p, x_0, x_1, x_2, \cdots$ *be points of* $X$ *such that* $\lim_{i\to\infty} x_i = x_0 \neq p$. *Then if*

(1) $y_i \in \overline{px_i}\rangle$, $i = 0, 1, 2, \cdots$ *and*

(2) $\lim_{i\to\infty} d(p, y_i) = d(p, y_0)$, *then*

(3) $\lim_{i\to\infty} y_i = y_0$.

We now define a map which moves points along rays similar to $Q$ in §3. Fix $p$ in $X$ and let

$$D^x = \begin{cases} \{x\} \times [0, \infty) \text{ if } \overline{px}\rangle \text{ is isometric to } [0, \infty) \\ \{x\} \times \left[0, \dfrac{a}{d(p, x)}\right) \text{ if } \overline{px}\rangle \text{ is isometric to } [0, a) \\ \{x\} \times \left[0, \dfrac{a}{d(p, x)}\right] \text{ if } \overline{px}\rangle \text{ is isometric to } [0, a] . \end{cases}$$

Now let $D = \bigcup_{x \in X - \{p\}} D^X$ be the domain of $P$ and define $P: D \to X$ by the following rule:

$$P(p, t) = p$$

$P(x, t) =$ the point $z$ of $\overline{px}\rangle$ such that $d(p, z) = t \cdot d(p, x)$ for $x \neq p$.

PROPOSITION 4.6. *$P$ is continuous.*

*Proof.* Follows from 4.5 and continuity of distance function.

REMARK 4.7. Since the function $P$ depends on the choice of the basepoint $p$, $P$ will be denoted $P_p$ any time confusion might arise. Likewise $D$ is denoted $D_p$.

Our next goal is to show that every open subset of $X$ contains a homeomorphic copy of $X$. The first step is to show that if a sequence of rays converges to a ray with endpoint, then all but finitely many of the sequence of rays have endpoints.

LEMMA 4.8. *Let $p, x_0, x_1, x_2, \cdots$ be points of $X$ such that $\overline{px_0}\rangle = \overline{px_0}$ and $\lim_{i \to \infty} x_i = x_0$. Then, for $i$ sufficiently large there is a point $z_i \in \overline{px_i}\rangle$ such that $\overline{pz_i} = \overline{px_i}\rangle$ and $\lim_{i \to \infty} z_i = x_0$.*

*Proof.* Clearly $\underline{\lim}_{i \to \infty} \operatorname{diam}(\overline{px_i}\rangle) \geqq d(p, x_0)$. On the other hand, if

$$\overline{\lim_{i \to \infty}} \operatorname{diam}(\overline{px_i}\rangle) = d(p, x_0) + \delta > d(p, x_0)$$

then there is an infinite set of integers, $M$, such that for $i \in M$ there is a $y_i \in \overline{px_i}\rangle$ such that $d(p, y_i) = d(p, x_0) + \min\{\delta, r\}$ where $r > 0$ is chosen to make $D(x_0, 2r)$ compact. The set $\{y_i: i \in M\}$ has a limit point, $y$, in $D(x_0, 2r)$. Since $y$ satisfies $d(p, y) = d(p, x_0) + d(x_0, y)$ we have $y \in \overline{px_0}\rangle = \overline{px_0}$. This is a contradiction because $d(p, y) > d(p, x_0)$. Thus $\lim_{i \to \infty} \operatorname{diam}(\overline{px_i}\rangle) = d(p, x_0)$.

Now choose $n_0$ to be an integer such that $x_i \in D(x_0, r)$ and

$$|\operatorname{diam} \overline{px_i}\rangle - d(p, x_0)| \leqq r$$

whenever $i \geqq n_0$. Then the set $((\overline{px_i}\rangle - \overline{px_i}) \cup \{x_i\})$ is in $D(p, 2r)$ when $i \geqq n_0$, and since each of the sets is closed it must be compact. Each of the rays $\overline{px_i}$ with $i \geqq n_0$ is then compact and letting $z_i$ be the endpoint the lemma is proved.

THEOREM 4.9. *Let $p$ be a point of $X$ and $f$ be a map from $X - \{p\}$ into $(0, 1]$. Then the function $G: X \to X$, defined by the formula*

$$g(p) = p$$
$$g(x) = P_p(x, f(x)) \text{ for } x \neq p$$

*is a homeomorphism if it is one-to-one.*

*Proof.* On $X - \{p\}$, $g$ is the composition of continuous functions so is continuous. It is continuous at $p$ as well because

$$d(p, P_p(x, f(x))) = f(x) \cdot d(p, x) \leqq d(p, x) .$$

Assume that $g$ is one-to-one. It remains only to show that $g^{-1}$ is continuous. Note that $g^{-1}$ is continuous at $p$ because if $D(p, r)$ is a compact neighborhood of $p$ and $a = \inf\{f(z): z \in D(p, r)\}$, then $g[D(p, r)]$ contains $D(p, a \cdot r)$ which is a neighborhood of $p$.

The map $g$ restricted to any segment $\overline{px}$ is a homeomorphism of $\overline{px}$ into itself with $p$ remaining fixed so if points are ordered by distance from $p$, $g$ preserves that order. Let $x_0, x_1, x_2, \cdots$ be points of $g[X]$ such that $\lim_{i\to\infty} x_i = x_0$. Let $y_i = g^{-1}(x_i)$ for $i = 0, 1, 2, \cdots$. Consider first the case where $d(p, y_i) \geqq d(p, y_0) + \delta$ for some $\delta > 0$, and all $i = 1, 2, 3, \cdots$. In view of Lemma 4.8, $y_0$ cannot be the endpoint of $\overline{py_0}\rangle$ so we may choose a point $w_0$ in $(\overline{py_0}\rangle - \overline{py_0})$ and within $\delta/2$ of $y_0$. Let $w_i$ be a point of $\overline{px_i}\rangle$ such that $d(p, w_i) = d(p, w_0)$ for $i = 1, 2, 3, \cdots$, and note that $\lim_{i\to\infty} w_i = w_0$. Since $w_0$ is farther than $y_0$ from $p$, $g(w_0)$ is farther than $g(y_0) = x_0$ from $p$. On the other hand, for $i = 1, 2, 3, \cdots$, $w_i$ is closer than $y_i$ to $p$ and so $g(w_i)$ is closer than $x_i = g(y_i)$. From the continuity of $g$, $g(w_0)$ is at least as close to $p$ as $x_0$. This case is ruled out and if no sequence can belong to this case, no infinite subsequence of a sequence can either, so the remaining possibility is that $d(p, y_i) \leqq d(p, y_0)$ for all $i \geqq n_0$, for some integer $n_0$. But then $\bigcup_{i=0}^{\infty} \overline{py_i}$ is compact and $g^{-1}$ is continuous on $g[\bigcup_{i=0}^{\infty} \overline{py_i}] = \bigcup_{i=0}^{\infty} \overline{px_i}$ and so $\lim_{i\to\infty} g^{-1}(x_i) = g^{-1}(x_0)$.

Note that if $f(x)$ is a nondecreasing function of $d(p, x)$, then $g$ is a homeomorphism.

COROLLARY 4.10. *Let $p$ be a point of $X$ and $U$ a neighborhood of $p$. Then there is a homeomorphism, $g$, of $X$ into $U$ leaving $p$ fixed and for $x \in X$, $g(x) \in \overline{px}$.*

*Proof.* Let $0 < r < 1$ be chosen so that $S(p, r) \subset U$, and define the map $h: (0, \infty) \to (0, r)$ by the formula $h(t) = 2r/\pi \tan^{-1} t$. Observe that $h$ is one-to-one and $h(t) \leqq t$. Define $f: X - \{p\} \to (0, 1]$ by the rule $f(x) = h(d(p, x))/d(p, x)$. Let $g$ be defined as in Theorem 4.9 in terms of $f$ and $P_p$. The map $g$ could fail to be one-to-one only by mapping two points of some ray, $\overline{px}\rangle$, to the same point. But $d(p, g(y)) = d(p, y) \cdot f(y) = h(d(p, y))$, so this cannot happen. Moreover, the last expression must be less than $r$ so $g[X] \subset S(p, r) \subset U$.

5. **Endpoints of rays are sparse.** Next we develop some contractibility conditions for $X$ and certain subsets, then show that in an SC-WR metric space of finite dimension the endpoints of rays are contained in a nowhere dense set.

Throughout this section let $(X, d)$ be an SC-WR metric space.

PROPOSITION 5.1. *Let $p$ be a point of $X$ and $r$ be a positive real. Then there is a map $h: X \times [0, 1] \to X$ with the following properties*:

(1) $h(-, 0)$ *is the identity on* $X$;

(2) $h(-, 1)[X] \subset D(p, r)$;

(3) *for* $t \in [0, 1]$ $h(-, t) | D(p, r)$ *is the identity on* $D(p, r)$ *and*

(4) *for* $(x, t) \in (X - D(p, r)) \times [0, 1]$ $h(x, t) \notin S(p, r)$.

*Proof.* Difine the function $m: X \to [0, 1]$ by the formulas

$$m(x) = \begin{cases} 1 \text{ if } x = p \\ \min\left\{1, \dfrac{r}{d(p, x)}\right\} \text{ in } x \neq p \end{cases}$$

and $g: X \times [0, 1] \to [0, 1]$ by $g(x, t) = (1 - t) + t \cdot m(x)$. Let $P$ be the function defined in § 4 with $p$ as its base point. Define $h: X \times [0, 1] \to X$ by $h(x, t) = P(x, g(x, t))$ and it is routine to verify $h$ satisfies conditions (1) through (4).

REMARK 5.2. By virtue of $h$ satisfying conditions (1), (2), and (3), $D(p, r)$ is said to be a *strong deformation retract* of $X$. A subset $A$ of $X$ is a *retract* if there is a map from $X \to A$ which is the identity on $A$.

PROPOSITION 5.3. *Let $p \in X$ and $r$ a positive real. Then for $y \in S(p, r)$, $X - \{y\}$ is contractible (in itself) if and only if $(D(p, r) - \{y\})$ is contractible (in itself).*

*Proof.* Let $y \in S(p, r)$ be given and take $h$ to be the deformation map defined in the previous proposition relative to $D(p, r)$. From properties (3) and (4) of $h$ it is clear that $h[(X - \{y\}) \times [0, 1]] \subset X - \{y\}$. Thus $h$ retracts $X - \{y\}$ onto $(D(p, r) - \{y\})$. A retract of a contractible space is contractible [4, p. 26], so $D(p, r) - \{y\}$ is contractible if $X$-$(y)$ is. The converse is obvious.

PROPOSITION 5.4. *Let $p \in X$ and $r > 0$. Then $D(p, r)$ is contractible and locally contractible.*

*Proof.* In §3 it was shown that $D(p, r)$ is contractible.

Fix $y \in D \equiv D(p, r)$ and $\delta > 0$. Letting $h$ be the deformation map from 5.1 and setting $f = h(-, 1)$ gives us that $f$ is a retraction of $X$ onto $D$. Define the map $g: D \times I \to D$ by the formula $g(x, t) = f(P_y(x, 1 - t))$. Clearly, $g$ is continuous, $g(-, 0)$ is the identity and $g(-, 1)$ is constantly $y$. Choose $r > 0$ so that $D(y, r)$ is compact and it follows that $C_n \equiv D \cap D(y, r/n)$ is compact for each $n = 1, 2, 3, \cdots$. The nested sequence of sets $C_n \times I$ converge to $\{y\} \times I$ and since $g$ is continuous there exists $n_0$ such that $g[C_{n_0} \times I] \subset S(y, \delta) \cap D$. It follows that $g | C_{n_0} \times I$ contracts $C_{n_0}$ to $y$ inside $S(y, \delta)$ and thus $D$ is locally contractible.

DEFINITION 5.5. For a set $A$ in a topological space $Y$ the space $A \times I/A \times \{0\}$, i.e., the upper-semi-continuous decomposition of $A \times I$ whose only nondegenerate element is $A \times \{0\}$, is called the *cone* over $A$.

PROPOSITION 5.6. *Let $(X, d)$ be a locally compact SC-WR metric space. If $A \subset X$ is compact and $p \in (X - A)$ such that, for $x \in A$, $\overline{px} \cap A = \{x\}$, then the set $B = \bigcup_{x \in A} \overline{px}$ is homeomorphic to the cone over $A$.*

*Proof.* The proof of this proposition appears in [8, 6.2] for $X$ compact. The proof carries over for $X$ locally compact in light of the properties shown in the preceding propositions.

The following theorem generalizes a result of D. Rolfsen [11] which was for compact spaces. The proof is identical except that it relies on earlier propositions in this paper for properties of locally compact spaces with SC-WR metrics.

THEOREM 5.7. *Let $(X, d)$ be a locally compact SC-WR metric space with* $\dim X = n$ *and* $0 < n < \infty$. *Then the set $U = \{x: X - \{x\}$ fails to be contractible in itself$\}$ contains a dense, open subset of $X$.*

COROLLARY 5.8. *If $(X, d)$ and $U$ are as in Theorem* 5.7, *then no point of $U$ is the endpoint of a ray.*

*Proof.* Let $x \in X$ and $p \neq x$ such that $\overline{px}\rangle = \overline{px}$. The map $P_p$: $X \times I \to X$ defined earlier when restricted to $(X - \{x\}) \times I$ clearly contracts $X - \{x\}$ to $p$ missing $x$ so $X - \{x\}$ is contractible (in itself). By Theorem 5.7 $x \notin U$.

## 6. Retract properties.

DEFINITION 6.1. Let $Y$ be a topological space and $A$ a subset of $Y$. $A$ is said to be a *neighborhood retract* of $Y$ provided there exists an open set, $O$, of $Y$ such that $A \subset O$ and $A$ is a retract of $O$.

DEFINITION 6.2. A metric space $Y$ is said to be an *absolute retract for metrizable spaces*, or an AR $(M)$-space, if for any metric space $Z$ and a closed subset $A$ of $Z$ with $A$ homeomorphic to $Y$, $A$ is a retract of $Z$. $Y$ is said to be an *absolute neighborhood retract for metrizable spaces*, or an ANR $(M)$-space, if for any metric space $Z$ and closed subset $A$ of $Z$, with $A$ homeomorphic to $Y$, $A$ is a neighborhood retract of $Z$.

DEFINITION 6.3. A metric space $Y$ is said to be an *absolute retract* or AR-space if $Y$ is an AR $(M)$-space and $Y$ is compact, $Y$ is said to be an *absolute neighborhood retract*, or ANR-space, if $Y$ is an ANR $(M)$-space and $Y$ is compact.

PROPOSITION 6.4. *Let $(X, d)$ be a locally compact* SC-WR *metric space of finite dimension. If $D(p, r) \subset X$ is compact, then it is an absolute retract and if no ray from $p$ ends inside $D(p, r)$, then the set $Sh\,(p, r) = \{x \in X: d(x, p) = r\}$ is an absolute neighborhood retract.*

*Proof.* As is evident from Proposition 5.4, $D(p, r)$ is contractible in itself and locally contractible and since it is compact and finite dimensional it is an absolute retract [4, 10.5, p. 122].

To show that Sh $(p, r)$ is an ANR it is sufficient to show that it is a neighborhood retract of the absolute retract $D(p, r)$ [4, 2.4, p. 101]. Since no rays end inside $D(p, r)$ we can retract $D(p, r) - \{p\}$ onto Sh $(p, r)$ by pushing outward along rays from $p$.

THEOREM 6.5. *If $(X, d)$ is a locally compact* SC-WR *metric space of finite dimension, then $X \in$ AR $(M)$.*

*Proof.* For a point $p$ of $X$ there is a positive number $r_p$ so that $D(p, r_p)$ is compact and by Proposition 6.3, $D(p, r_p) \in$ AR $(M)$. As noted in Theorem 3.5, $X$ is separable and since each point of $X$ has a neighborhood which is an ANR $(M)$-space, $X \in$ ANR $(M)$ [4, 10.4,

p. 99]. However, since $X$ is contractible, $X \in \mathrm{AR}(M)$ [4, 9.1, p. 96].

## 7. Existence of cells in low dimension spaces.

LEMMA 7.1. *Let $(X, d)$ be a locally compact,* SC-WR *metric space. Then if $p, x,$ and $y$ are three non-colinear points of $X$, then $\bigcup_{z \in \overline{xy}} \overline{pz}$ is a 2-cell and $\bigcup_{z \in \overline{xy}} \overline{pz}\rangle$ is 2-dimensional and closed.*

*Proof.* Let $A = \bigcup_{z \in \overline{xy}} \overline{pz}$ and $\hat{A} = \bigcup_{z \in \overline{xy}} \overline{pz}\rangle$.

In light of Proposition 5.6 Lelek and Nitka's proof [8] that $A$ is a 2-cell carries over from compact to locally compact spaces.

To establish the second part of the lemma, let $r = \inf\{d(p, z): z \in \overline{xy}\}$. Clearly, $r > 0$, and by Corollary 4.10 there is a homeomorphism of $X$ into $S(p, r)$ that moves points along rays. Under this map, $\hat{A}$ is carried into $A$ and so is 2-dimensional. Moreover, if $q$ is a point of the closure of $\hat{A}$, then the image of $q$ is in the compact set $A$, so $\overline{pq}\rangle$ meets $\overline{xy}$ at a point $z_0$. It follows that $\overline{pq}\rangle = \overline{pz_0}\rangle \subset \hat{A}$ and $q \in \hat{A}$.

THEOREM 7.2. *Let $(X, d)$ be a locally compact* SC-WR *metric space of dimension $n$ with $1 \leqq n \leqq 3$. Then there is a dense, open set $V$ of $X$ such that points of $V$ have closed distance neighborhoods homeomorphic to $I^n$.*

*Proof.* For the case $n = 1$ the theorem follows directly from the lemma. Since dimension $X = 1$, $X$ has two points $p$ and $q$. Since $X$ cannot contain 3 noncolinear points (lemma), $X = \overline{pq}\rangle \cup \overline{qp}\rangle$. Letting $V = X - \{$endpoints of $\overline{pq}\rangle$ and $\overline{qp}\rangle$, if any$\}$ the proof is complete.

The case $n = 2$ or $n = 3$. The argument that Rolfsen [11] gives for a similar theorem with $X$ compact and $\dim X = 3$ carries over to locally compact spaces and, with a small addition, works for $\dim X = 2$ as well. That argument is outlined below with references to results of this paper needed to carry through various of the steps.

Let $U = \{x \in X: X - \{x\}$ fails to be contractible in itself$\}$ and let $V = \operatorname{int} U$. Fix $p \in V$ and choose $\varepsilon > 0$ so that $\bar{N} = D(p, \varepsilon)$ is compact and contained in $V$. Let $S = \{x \in X: d(p, x) = \varepsilon\}$.

(1) $V$ is open and dense in $X$ (Proposition 5.7).

(2) $S$ is compact, $(n-1)$-dimensional and $\bar{N}$ is homeomorphic to the (abstract) cone over $S$ [11, (4), p. 218], (Proposition 6.4).

(3) $S$ is an ANR-space [11, (6), p. 218], (Corollary 4.10).

(4) $S$ does not have the fixed point property [11, (8), p. 218].

(5) For $s \in S$, $S - \{s\}$ is contractible in itself [11, (7), p. 218].

(6) $S$ is connected and if $n = 3$, then no finite set separates $S$ [11, (9), p. 218], (Lemma 7.1).

(7) If $n = 3$, then $S$ is a 2-sphere [11, (10), p. 219].

(8) If $n = 2$, then $S$ is a 1-sphere.

Since $S$ does not have the fixed point property, it follows from a theorem of Lefschetz [5] that for some $k \geqq 0$ the (reduced) singular homology group (integral coefficients), $H_k(S)$ is nontrivial. $S$ is connected so $H_0(S) = 0$ and $\dim(S) = 1$ so $H_k(S) = 0$ for $k \geqq 2$, hence $H_1(S) \neq 0$. Because of (5), $H_k(S - \{s\}) = 0$ for all $k \geqq 0$. It follows from a theorem of McCord's [9] that $S$ is a 1-sphere.

Part (2) along with (7) and (8) yield that $\bar{N}$ is homeomorphic to $I^n$.

## 8. Topological characterizations.

DEFINITION 8.1. A point $y$ in a topological space $Y$ has a *Euclidean neighborhood* if for some neighborhood $V$ of $y$ and some natural number, $n$, $V$ is homeomorphic to $E^n$.

Throughout this section let $(X, d)$ be a locally compact SC-WR metric space.

PROPOSITION 8.2. *If some point of $X$ has a Euclidean neighborhood, then the set, $M \equiv \{x \in X: x$ is the endpoint of some ray$\}$, is closed in $X$ and every point of $(X - M)$ has a Euclidean neighborhood.*

*Proof.* Let $p$ be a point of $X$ with a Euclidean neighborhood $V$, $V$ homeomorphic to $E^n$. There is a homeomorphism of $X$ into $V$ so we may consider $X$, as a topological space, to be imbedded in $E^n$. Let $\operatorname{int} X$ and $\operatorname{Bd} X$ denote the interior and boundary of $X$ as a subset of $E^n$.

For any subset, $Y$, of $E^n$ if $y \in \operatorname{int} Y$, then $Y - \{y\}$ is not contractible (in itself). It follows from proof of Corollary 5.8 that for $x \in M$, $X - \{x\}$ is contractible (in itself), so $M \subset \operatorname{Bd} X$.

Consider a point, $x \in (X - M)$. Since $x$ is not the endpoint of the ray $\overline{px}\rangle$ there is a point $q$ in $\overline{px}\rangle - \overline{px}$. Set $t = d(p, x)/d(p, q)$ and since $0 < t < 1$, the map $P_q(-, t)$ is a homeomorphism of $X$ into itself (Theorem 4.9) that carries $p$ to $x$. By the invariance of domain the image of $V$ under this map is open in $E^n$, hence $x \in \operatorname{int} X$. It follows that $M \subset (X \cap \operatorname{Bd} X)$.

Now $M = (X \cap \operatorname{Bd} X)$ and so $M$ is closed in $X$, and since $(X - M) = \operatorname{int} X$ every point of $(X - M)$ has a Euclidean neighborhood.

REMARK 8.3. Note that the set, $M_p \equiv \{x \in X: \overline{px} = \overline{px}\rangle\}$, where $p$ is a point with Euclidean neighborhood, is contained in $M$. However, in the last part of the above proof it was shown that, in fact, $(X \cap \operatorname{Bd} X) \subset M_p$ so $M_p = M$.

PROPOSITION 8.4. *Let $p \in X$ with $M_p$ closed. The function $r_p: X - \{p\} \to E^1 \cup \{+\infty\}$, defined by $r_p(x) = \operatorname{diam} \overline{px}\rangle$, is lower-semi-continuous.*

*Proof.* Let $x_0, x_1, x_2, \cdots$ be points of $X - \{p\}$ with $\lim_{n\to\infty} x_n = x_0$ and let $t$ be a number less than $r_p(x_0)$. We may as well assume $\lim_{n\to\infty} r_p(x_n)$ exists, and call it $s$. To complete the proof it remains only to rule out the possibility that $s < t$.

If $s < t$, then there is a point $z_0 \in \overline{px_0}\rangle$ such that $d(p, z_0) = s$. We can also assume $r_p(x_n) < \infty$ for $n > 0$, so if we choose $z_n \in \overline{px_n}\rangle$ such that $d(p, z_n) \geqq (\operatorname{diam} \overline{px_n}\rangle - 1/n)$ then, by Proposition 4.5, $\lim z_n = z_0$. Let $D(z_0, r)$ be a compact neighborhood of $z_0$ and, clearly, $T_n \subset D(z_0, r)$ for $n$ sufficiently large where $T_n = \overline{px_n}\rangle - \overline{pz_n}$. Thus $T_n$, and consequently, $\overline{px_n}\rangle$ are compact for $n$ large. Let $y_n$ be the endpoint of the compact $\overline{px_n}\rangle$ and observe that $\lim y_n = z_0$. Since $y_n \in M_p$ and $M_p$ is closed $z_0 \in M_p$, hence $\overline{px_0}\rangle = \overline{pz_0}$, contradicting the choice of $z_0$.

THEOREM 8.5. *Let $p \in (X - M)$ have a Euclidean neighborhood. If $r > 0$ and $D(p, 4r)$ is compact and contained in $(X - M)$, then there is a subset $T$ of $D(p, r)$ such that $S(p, r) \subset T$ and $T$ is homeomorphic to $X$.*

*Proof.* The method of the proof will be to use a sequence of continuous functions approximating $r_p$ to partition $X$ into countably many subsets. These subsets will be mapped homeomorphically onto $S(p, r/2)$ and countably many annuli between $S(p, r/2)$ and $S(p, r)$ along with a subset of $D(p, r) - S(p, r)$.

The map $r_p$ is lower-semi-continuous and has range contained in $[4r, +\infty]$ since no ray ends in $D(p, 4r)$. Let $S = \{x \in X: D(p, x) = r\}$ and $r_p | S$ is lower-semi-continuous. A lower-semi-continuous function on a separable, finite dimensional metric space which is bounded below can be pointwise approximated by a (strictly) increasing sequence of continuous functions [2]. Let $\hat{f}_1, \hat{f}_2, \hat{f}_3, \cdots$ be such a sequence approximating $r_p | S$ and we can assume range of $\hat{f}_n$, all $n$, is contained in $[2r, \infty)$. Extend each $\hat{f}_n$ to all of $X - \{p\}$ by letting $f_n(x) = \hat{f}_n(y)$ where $y$ is the unique point of $S$ in the ray $\overline{px}\rangle$. Clearly the extended functions are continuous on $X - \{p\}$.

Define

$$A_0 \equiv \{x \in X - \{p\}: d(p, x) \leqq f_1(x)\}$$
$$A_n \equiv \{x \in X - \{p\}: f_n(x) \leqq d(p, x) \leqq f_{n+1}(x)\} \text{ for } 0 < n < \infty$$
$$A_\infty \equiv \{x \in X - \{p\}: f_n(x) < d(p, x) \text{ all } n\}\,.$$

The desired homeomorphism $h: X \to D(p, r)$ is defined by the formulas

$$h(p) = p$$
$$h(x) = P_p(x, m(x)) \text{ for } x \neq p$$

where $m: X - \{p\} \to (0, 1]$ is defined as follows:

$$m(x) = \begin{cases} \dfrac{r}{f_1(x)} \cdot \left(\dfrac{2^1 - 1}{2^1}\right) \text{ if } x \in A_0 \\ \left[1 + \dfrac{d(p, x) - f_n(x)}{f_{n+1}(x) - f_n(x)} \cdot \left(\dfrac{1/2}{2^n - 1}\right)\right] \cdot \dfrac{r}{d(p, x)} \cdot \dfrac{2^n - 1}{2^n}, \text{ if } x \in A_n, \\ 0 < n < \infty \\ \dfrac{r}{d(p, x)}, \text{ if } x \in A_\infty . \end{cases}$$

If $x \in A_n \cap A_{n+1}$ for $0 \leqq n < \infty$, then $m(x)$ has two definitions but since $d(p, x) = f_{n+1}(x)$ in that case it is routine to verify that

$$m(x) = \frac{r}{f_{n+1}(x)} \cdot \frac{2^{n+1} - 1}{2^{n+1}}$$

from both definitions. It is also evident that $m$ is continuous on each $A_n$, $0 \leqq n \leqq \infty$ and on $\bigcup_{n<\infty} A_n$ as well.

Observe that for $x \in A_n$, $0 < n < \infty$, then

$$0 \leqq \frac{d(p, x) - f_n(x)}{f_{n+1}(x) - f_n(x)} \leqq 1$$

which yields

$$\frac{r}{d(p, x)}\left(\frac{2^n - 1}{2^n}\right) \leqq m(x) \leqq \frac{r}{d(p, x)}\left(\frac{2^{n+1} - 1}{2^{n+1}}\right).$$

Thus if $x_1, x_2, x_3, \cdots$ is a sequence of points in $\bigcup_{n<\infty} A_n$ with limit $x_0 \in A_\infty$, then

$$\lim_{k \to \infty} m(x_k) = \frac{r}{d(p, x_0)} = m(x_0) .$$

Thus $m$ is continuous on $X$ and the above bounds on $m$ shows that $m$ has range $[1/4, 1/2] \subset (0, 1]$.

In order to show that $h$ is a homeomorphism it only remains to show that $h$ is one-to-one, and because $h$ moves points along rays from $p$, it is sufficient to consider one such ray. Fix $x_0 \in X - \{p\}$ and let $b_1, b_2, b_3, \cdots$ be points of $\overline{px_0}\rangle$ chosen so $d(p, b_n) = f_n(x_0) = f_n(b_n)$. Let $a_n = h(b_n)$ and note

$$d(p, a_n) = m(b_n) \cdot d(p, b_n) = r \cdot \frac{2^n - 1}{2^n} .$$

The function $m$ is constant on $\overline{pb_1} - \{p\}$ so $h$ is one-to-one on $\overline{pb_1}$. In general, $\overline{b_n b_{n+1}} = A_n \cap \overline{px_0}\rangle$, so on $\overline{b_n b_{n+1}}$

$$m(x) = \frac{\alpha_n}{d(p, x)} + \beta_n \text{ where} \begin{cases} \alpha_n = r\cdot\left(\frac{2^n - 1}{2^n}\right)\left[1 + \frac{-f_n(x_0)}{f_{n+1}(x_0) - f_n(x_0)}\cdot\frac{1/2}{2^n - 1}\right] \\ \beta_n = r\left(\frac{2^n - 1}{2^n}\right)\left(\frac{1}{f_{n+1}(x_0) - f_n(x_0)}\right)\left(\frac{1/2}{2^n - 1}\right). \end{cases}$$

Thus $d(p, h(x)) = \alpha_n + \beta_n d(p, x)$ and since $\beta_n > 0$, $h$ is one-to-one on $\overline{b_n b_{n+1}}$, carrying $\overline{b_n b_{n+1}}$ onto $\overline{a_n a_{n+1}}$. $A_\infty \cap \overline{px_0}\rangle$ consists of at most one point whose image lies a distance $r$ from $p$. It follows that $h$ is one-to-one on $\overline{px_0}\rangle$ and also that the image of $\overline{px_0}\rangle$ under $h$ contains $(\overline{px_0}\rangle \cap S(p, r))$.

Let $T = h[X]$ and the theorem is proved.

COROLLARY 8.6. *For $1 \leqq n \leqq 3$ an $n$-dimensional, locally compact metrizable space, $X$, admits an* SC-WR *metric if and only if $X$ is an $n$-manifold (with boundary) and is homeomorphic to a subset of closed unit $n$-ball and which contains the interior of the $n$-ball.*

*Proof.* The necessity is obvious because the usual metric for $E^n$ restricted to such a subset is SC-WR.

To show the sufficiency let $d$ be an SC-WR metric for $X$. By Theorem 7.2 there exists a point $p$ of $X$ and a positive number $t$ such that $D(p, t)$ is homeomorphic to $I^n$. $X$ is homeomorphic to a set $T$ with $S(p, r) \subset T \subset D(p, r)$ where $r = t/4$. $D(p, r)$ is homeomorphic to $I^n$ and thus to $B$, the unit ball in $E^n$. Let $T' \subset B$ be the image of $T$ under the last homeomorphism. Since $T \supset S(p, r)$, $T' \supset \operatorname{int} B$ and $T'$ being locally compact yields that $T'$ contains a relatively open subset of Bd $B$. $T'$ is a $n$-manifold and consequently $X$ is as well.

PROPOSITION 8.7. *Let $(X, d)$ be a locally compact* SC-WR *space. If $X$ is of finite dimension, the following are equivalent:*

(a) *$(X, d)$ is externally convex*
(b) *no ray has an endpoint*
(c) *$X$ is homogeneous*
(d) *$X$ is locally homogeneous.*

*Proof.* The pattern of the proof is (a) $\leftrightarrow$ (b) $\rightarrow$ (c) $\rightarrow$ (d) $\rightarrow$ (b).

(a) $\leftrightarrow$ (b). This equivalence was noted in §4.

(b) $\rightarrow$ (c). Assume (b) holds. We first establish that if $D(p, r)$ is compact and $q \in S(p, r)$, then for $w \in \overline{pq} - \{p\}$ there is a homeomorphism of $X$ onto itself that carries $q$ to $w$.

Let $a = d(p, q)$ and $b = d(p, w)$. Define $A_1 \equiv D(p, a) - \{p\}$, $A_2 \equiv$

$D(p, r) - S(p, a)$ and $A_3 \equiv X - S(p, r)$ and define the map $f: X - \{p\} \to (0, 1]$ by the formula:

$$f(x) = \begin{cases} \dfrac{b}{a} & \text{if } x \in A_1 \\ \dfrac{b}{a} + \left(1 - \dfrac{b}{a}\right)\dfrac{d(p, x) - a}{r - a} & \text{if } x \in A_2 \\ 1 & \text{if } x \in A_3 . \end{cases}$$

Since $f$ is continuous on each of the three closed sets $A_1$, $A_2$, and $A_3$ and uniquely defined on their intersections, it is continuous. Let $h(x) = P_p(x, f(x))$ for $x \neq p$ and $f(p) = p$. $P_p$ moves points along rays and since $f(x)$ is a nondecreasing function of $d(p, x)$, $h$ is one-to-one and therefore a homeomorphism (4.9). Note also that $d(p, h(q)) = d(p, q) \cdot f(q) = a \cdot b/a = b = d(p, w)$, so $h(q) = w$.

On $A_3 \cup \{p\}$, $h$ is the identity and if $x \in X - (A_3 \cup \{p\})$, then the ray $\overline{px}\rangle$ is not contained in $D(p, r)$ so there is a point $y$ in $\overline{px}\rangle$ a distance $r$ from $p$. The segment $\overline{py}$ maps into itself under $h$ and both $y$ and $p$ are fixed so $x$ is the image under $h$ of some point. Thus $h[X] = X$.

Moreover, there is a homeomorphism carrying $q$ to $p$ because there is a point $p'$ in $\overline{qp}\rangle - \overline{qp}$ and close to $p$ which has a compact distance neighborhood contained in $D(p, r)$ and containing $q$ in its interior.

For $x$ and $y$ two points of $X$ there is a finite, simple chain of open distance neighborhoods with the first centered at $x$ and the last at $y$. The above homeomorphisms and their inverses allow us to push $x$ into the second distance neighborhood and then into center of it. Continuing this process a finite number of sets pushes $x$ to $y$.

(c) $\to$ (d). Obvious.

(d) $\to$ (b). Assume that (d) holds and there is a point $q \in X$ such that $q$ is the endpoint of a ray. Since $X$ has finite dimension, there is a point $p \in X$ such that $X - \{p\}$ is not contractible (5.7). Let $U$ and $V$ be neighborhoods of $q$ and $p$, respectively, and $h$ be a homeomorphism of $U$ onto $V$ carrying $q$ to $p$. The point $q$ has arbitrarily small deleted distance neighborhoods that are contractible, so let $D$ be one contained in $U$. $F[D]$ is a neighborhood of $p$, so there exists an $r$ such that $D(p, r) \subset f[D]$, and $D(p, r)$ compact. $D(p, r) - \{p\}$ is a retract of $X - \{p\}$, so is a retract of $f[D] - \{p\} = f[D - \{q\}]$. Since $f[D - \{q\}]$ is contractible and since a retract of a contractible space is contractible [4, p. 26, 13.2], it follows that $D(p, r) - \{p\}$ is contractible. This is a contradiction because $X - \{p\}$ is then contractible (5.3).

THEOREM 8.8. *Let* $(X, d)$ *be a locolly compact* SC-WR *metric space* *of dimension $n$ and let $M \equiv \{x \in X\colon x$ is the endpoint of a ray$\}$. Then if $1 \leqq n \leqq 3$ or some point of $X$ has an $E^n$-like neighborhood, then $(X - M)$ is homeomorphic to $E^n$.*

*Proof.* If $n \leqq 3$, then some points of $X$ have an $E^n$-like neighborhood, so we may, in any case, choose $p \in X$ with an $E^n$-like neighborhood. There exist $r > 0$ and a set $T$ such that $S(p, r) \subset T \subset D(p, r)$ and $T$ homeomorphic to $X$. Under this homeomorphism, if $x$ is not the endpoint of the $\overline{px}\rangle$, then $x$ maps into $S(p, r)$, and if $x$ is the endpoint, it maps into $D(p, r) - S(p, r)$. Thus $(X - M)$ is homeomorphic to $S(p, r)$. By Proposition 5.6, $D(p, r)$ is the cone over $S \equiv \{x \in X\colon d(p, x) = r\}$. M. Brown has shown [5] that if the cone over a set $A$ is $E^n$-like at the vertex, then the (cone over $A$) $- A$ is homeomorphic to $E^n$. Thus, $S(p, r) = D(p, r) - S$ is homeomorphic to $E^n$ and the theorem follows.

COROLLARY 8.9. *Let $(X, d)$ be a locally compact* SC-WR *metric space of dimension $n$. $X$ is homeomorphic to $E^n$ if and only if* (1) *any condition of Proposition* 8.7 *holds, and* (2) $1 \leqq n \leqq 3$ *or some point of $X$ has an $E^n$-like neighborhood.*

Note that if any condition of 8.7 holds, then $X$ is locally homogeneous and by Theorem 3.6, $X$ is an $n$-manifold if it contains an $n$-ball. We can change 8.9 slightly as follows:

COROLLARY 8.10. *A locally compact space of dimension $n$ is homeomorphic to $E^n$ if and only if it admits an* SC-WR, *externally convex metric and, for $n \geqq 4$, contains an $n$-ball.*

9. **Compact spaces.** Rolfsen [10] proved that a compact $n$-manifold (with boundary) which admits an SC-WR metric is homeomorphic to $I^n$ when $n \geqq 6$. In this section it is shown that the result holds for $n = 4$ or 5 whenever there is a terminal point in the space.

DEFINITION 9.1. In a metric space $(X, d)$, a point $p$ is said to be a *terminal* point if for $x, y \in X$, $d(x, y) = d(x, p) + d(p, y)$ implies $p = x$ or $p = y$.

THEOREM 9.2. *Let $(X, d)$ be a compact* SC-WR *metric space. If $X$ is an $n$-manifold and has a terminal point, then $X$ is homeomorphic to $I^n$.*

*Proof.* Let $p$ be a point of $(X - \partial X)$ ($\partial X$ is the boundary of $X$), and let $M = \{x \in X\colon \overline{px}\rangle = \overline{px}\}$. As is evident from the proof of Proposition 8.2, $M$ is the boundary of $X$ in an embedding of $X$ in

$E^n$, so $M = \partial X$. First, note that if $x \in M$ then the segment, $\overline{px}$, meets $M$ only at $x$.

Take $q$ to be a terminal point of $X$ and since $q \in M$, $q$ has a neighborhood, $V$, relative to $M$ which is homeomorphic to $E^{n-1}$. Let $p_1, p_2, p_3, \cdots$ be a sequence of points of $\overline{pq} - \{q\}$ which converges to $q$. For each $i = 1, 2, 3, \cdots$ define the function $h_i: M \rightarrow M$ by the rule: $h_i(x)$ is the endpoint of the ray $\overline{xp_i}\rangle$. If $y = h_i(x)$, then $x \in \overline{yp_i}\rangle$ and by our earlier observation, $h_i(y) = x$. Thus $h_i$ is one-to-one and onto for each $i$, and the continuity of $h_i$ is easily established, so $h_i$ is a homeomorphism of $M$ onto itself.

Suppose that for each integer, $i$, there is a point $x_i \in M - (V \cup h_i[V])$. $M$ is compact, so there is a point $x_0 \in M$ which is a limit of some subsequence of $x_1, x_2, \cdots$. We may assume that $\lim_{i\to\infty} x_i = x_0$. Let $z$ be a limit point of $\{h_i(x_i): i = 1, 2, \cdots\}$ and note $z \notin V$ hence $z \neq q$. But since $\lim p_i = q$, $d(x_0, z) = d(x_0, q) + d(q, z)$ contradicting the choice of $q$. For some $i$ then, $M = V \cup h_i[V]$.

The compact Hausdorff space $M$ being the union of two open $(n-1)$-cells is an $(n-1)$-sphere. The set $\bigcup_{x \in M} \overline{px}$ is homeomorphic to $I^n$ and is all of $X$, so the theorem is proved.

## References


1. P. Alexandroff, *Über die Metrization der im Kleinen kompakten topologischen Räume*, Math. Ann., **92** (1924), 294-301.
2. G. Berg, *Approximating semi-continuous functions*, to appear.
3. R. H. Bing and K. Borsuk, *Some remarks concerning topologically homogeneous spaces*, Annals of Math., **81** (1965), 100-111.
4. K. Borsuk, *Theory of Retracts*, Warsaw, Polish Scientific Publishers, 1967.
5. M. Brown, *The monotone union of open n-cells is an open n-cells*, Proc. Amer. Math. Soc., **12** (1961), 812-814.
6. W. Hurewicz and H. Wallman, *Dimension Theory*, Princeton, Princeton Univ. Press, 1948.
7. S. Lefschetz, *Topics in topology*, Annals of Math. Studies, **10** (1949).
8. A. Lelek and W. Nitka, *On convex metric spaces I*, Fund. Math., **49** (1961), 183-204.
9. M. McCord, *Spaces with acyclic point complements*, Proc. Amer. Math. Soc., **17** (1966), 886-890.
10. D. Rolfsen, *Strongly convex metrics in cells*, Bull. Amer. Math. Soc., **74** (1968), 171-175.
11. ———, *Characterization of the 3-cell by its metric*, Fund. Math., **68** (1970), 315-223.
12. W, White, *A 2-complex is collapsible if and only if it admits a strongly convex metric*, Notices Amer. Math. Soc., **14** (1967), 848.


Received April 7, 1972 and in revised form March 28, 1973. This paper is based on the author's Ph. D. dissertation at Arizona State University under the direction of Prof. E. E. Grace.

ARIZONA STATE UNIVERSITY
AND
NEW MEXICO STATE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# THE SPACE OF BOUNDED SEQUENCES WITH THE MIXED TOPOLOGY

AJIT KAUR CHILANA

**The space of bounded sequences with the mixed topology has some interesting properties and can be used to answer two questions on boundedly generated spaces asked by T. Ito and T. Seidman.**

1. **Introduction.** We consider the locally convex algebra $m$ of bounded complex sequences with pointwise addition and multiplication equipped with the mixed topology [11]. The topology $\tau$ is the same as the strict topology $\beta$ [1] on $C(S)$, when $S$ is taken to be the space of positive integers. This space has a number of interesting properties, some of which can be found in [11], [3] (Example 3), [1], and [4]. In this note we obtain some further results such as: this space is hereditary boundedly generated [6], it has an unconditional Schauder basis, and its Gelfand map is continuous.

The basic ideas are as in [7]. A locally convex space $E$ is said to be *boundedly generated*, in short, $BG$ if it is the closed linear span of a bounded subset; it is said to be *hereditary boundedly generated*, in short, $HBG$, if every closed linear subspace of the space is $BG$ [6]. $E$ is called *sequentially* barrelled if every $\tau_s(E, E')$-null sequence in the topological dual $E'$ of $E$ is equicontinuous [10]. The *sequential dual* $E^+$ of $E$ is the set of sequentially continuous linear functionals $f$ on $E$ (i.e., $f(x_n) \to 0$ whenever $x_n \to 0$ in $E$) [10]. $E$ is called *semi 1-barrelled* if every $\tau_s(E, E')$-bounded sequence in $E'$ is equicontinuous [3]. A barrel in a locally convex algebra $E$ that is also an idempotent set (i.e., $A \subset E$ such that $A.A \subset A$) is called an *m-barrel* and $E$ is said to be *m-barrelled* if every $m$-barrel in $E$ is a neighborhood of 0 [8]. Let $E$ be a complex locally convex algebra. Let $M$ denote the set of nonzero, continuous, multiplicative, linear functionals on $E$, provided with the weak topology induced by $E$. Let $C(M)$ denote the space of complex continuous functions on $M$ with the topology of compact convergence. The Gelfand map $G$ on $E$ to $C(M)$ is given by $G(x)(m) = m(x)$ $(x \in E, m \in M)$. $E$ is called strongly semi-simple if $G$ is an algebraic isomorphism of $E$ into $C(M)$.

Now let $E$ denote the space $m$ with the mixed topology $\tau$ or, equivalently, the strict topology $\beta$. A base of $\tau$-neighborhoods of 0 is given by

$$\mathscr{U} = \{U_a = \{x = (x_n) \colon |x_n| \leqslant a_n,\ n = 1, 2, \cdots\},\ a = (a_n),\ 0 < a_n \to \infty\}\ .$$

Let $B = \{x \in E: |x_n| \leqslant 1, n = 1, 2, \cdots\}$, $||x|| = \sup\{|x_n|: n = 1, 2, \cdots\}$ for $x \in m$. Let $e^i = (x_n)$: $x_i = 1$ and $x_n = 0$ if $i \neq n$. Let

$$l^1 = \left\{x \in m: \sum_{n=1}^{\infty} |x_n| < \infty\right\},$$

and $||x||_1 = \sum_{n=1}^{\infty} |x_n|$ $(x \in l^1)$. The $\alpha$-dual of $m$ is $l^1$ ([7], § 30.1) and $(m, l^1)$ is thus a dual pair.

## 2. Properties of the space $E$.

I. *$E$ is complete and has the Mackey topology $\tau_k(l^1)$.*

It is proved in [1] (also in [3]) that $E$ is complete. The second part follows from [4], Theorems 2 and 4, on taking $S$ to be the set of positive integers with the discrete topology.

II. *No topology on $m$ compatible with duality is barrelled or $m$-barrelled.*

The set $B$ is a barrel [3] and not a neighborhood in $E$, also it is idempotent. As $E$ carries the strongest topology compatible with duality and barrels remain barrels under any topology compatible with duality, the result follows immediately.

III. *$E$ is an HBG space.*

For any subspace $F$ of $E$, $F = \bigcup\{n(B \cap F): n = 1, 2, \cdots\}$ and, therefore, $F$ is the closed linear hull of a bounded subset of itself.

REMARK 1. It was asked in question (2) of [6], if there are any *HBG* spaces that are not Banach or separable Fréchet spaces. II and III give an affirmative answer.

REMARK 2. In [2] the first part of question (3) in [6] was answered in the negative and the following more general question was raised: If $F$ is a *BG* space with dual $F'$, then must there be a barrelled topology compatible with duality $(F, F')$? The example given there to prove that the answer is "No" is artificial in the sense that its completion is a Banach space and thus barrelled. By I, $E$ is a complete space and II and III show that it serves as a better example.

IV. *A sequentially continuous linear functional on $E$ is continuous and $E$ is sequentially barrelled.*

Combining [5], Theorem III (2.8) and I above, we have $E' = E^+$. Thus $\tau = \tau_k(E^+, E)$. Proposition 4.3 in [10] then gives that $E$ is sequentialy barrelled.

V. *E is not semi 1-barrelled.*

Consider $A = \{e^n : n = 1, 2, \cdots\} \subset l^1 = E'$. For each $x \in E$, and for $n = 1, 2, \cdots$, $|e^n(x)| = |x_n| \leqslant \|x\|$. So $A$ is $\tau_s(E', E)$-bounded. Also the polar $A^\circ$ of $A$ in $E$ is $B$ which is not a neighborhood. Therefore, $A$ is not equicontinuous.

REMARK 3. It is known ([3], Proposition 9 (ii), p. 481) that a semi 1-barrelled space is sequentially barrelled. IV and V show that the reverse implication may not be true. We take this opportunity to point out that there are two Proposition 9 in [3] (!) and in Proposition 9 (ii) on p. 481 [3] it should be almost semi-1-barrelled instead of almost semi-barrelled.

VI. *E has a Schauder basis* $(e^n)$, *which is*

( i ) *bounded multiplier,*

( ii ) *boundedly complete,*

(iii) *not of type* $P^*$,

(iv) *unconditional,*

( v ) *shrinking,*

(vi) *not of type* $P$,

(vii) *monotone, and*

(viii) *e-Schauder* [5].

We note that $m$ is perfect and normal ([7], § 30.1), $E' = l' = m^\times$ and $\tau = \tau_k(l^1)$. We can use [5], I (2.5) to obtain (i), (ii), and (iii). To prove (iv) we appeal to [5], I (2.4) and (i) above. The strong dual of $E$ is $(l^1, \|\cdot\|_1)$ and it has $(e^n)$ as a Schauder basis, so (v) is true. For (vi) note that $e^n \to 0$ in $E$. Also $\mathscr{U}$ satisfies the conditions for $(e^n)$ to be monotone and $S_N U_a \subset U_a$ for all $U_a \in \mathscr{U}$ and $N = 1, 2, \cdots$. Thus (vii) and (viii) are true.

REMARK 4. It is well-known that $m$ with the sup-norm topology is not separable and thus cannot have a basis. The above result shows the difference a change in the topology can make.

VII. *$C(M)$ is barrelled.*

Let $0 \neq f \in l^1$ and $f$ be multiplicative on $E$. Because

$$x = \lim_{n\to\infty} \sum_{j=1}^{n} x_j e^j ,$$

$f(x) = \lim_{n\to\infty} \sum_{j=1}^{n} x_j f(e^j)$ for each $x \in E$. So there is an $n$ such that $f(e^n) \neq 0$. Also $f(e^n) = f(e^n e^n) = f(e^n) f(e^n)$, so we must have $f(e^n) = 1$. Now for $j \neq n$ $f(e^n)f(e^j) = f(e^n e^j) = f(0) = 0$, so $f(e^j) = 0$. Thus $f(x) = x_n$ and $f$ can be identified with $e^n \in l^1$. So

$$M = \{e^n \colon n = 1, 2, \cdots\} .$$

Also $\{x\}^\circ \cap M = \{e^n\}$ if $x \in m$ be such that $x_n = 1$ and $x_j = 2$ for $j \neq n$. Hence $M$ can be identified with the set of positive integers with the discrete topology. Therefore, $C(M)$ is the space of all complex sequences with the topology of pointwise convergence and is, thus barrelled.

REMARK 5. We note that $m$-barrelledness of some topology compatible with duality is sufficient in [8], Lemma 3.1 (or [9], Cor. 6.3) and even this condition is not necessary as shown by II and VII.

VIII. *E has jointly continuous multiplication.*

If $a = (a_n)$ be such that $0 < a_n \to \infty$ then for $b = (b_n)$, where $b_n = a_n^{1/2}$, $0 < b_n \to \infty$ and also $U_b U_b \subset U_a$.

IX. *The Gelfand map is continuous but not a homeomorphism.*

It is immediate from the proof of VII.

The next result shows that $E$ does not, however, have a good functional representation.

X. *E cannot be embedded algebraically and topologically in a $C(X)$ for $X$ a locally compact Hausdorff space or for $X$ a completely regular Hausdorff space.*

From the proof of VII we get that $G$ is an isomorphism of $E$ into $C(M)$, and thus $E$ is strongly semi-simple. Also in view of VIII, $E$ is a topological algebra in the sense of [9]. Combining Theorem 4.6 of [9] and IX above we get the required result.

REMARK 6. This space also helps in distinguishing some classes of topological algebras such as $m$-$k$-barrelled algebras, $m$-$k$-infrabarrelled algebras, locally boundedly multiplicatively convex algebras.

I should like to thank the referee for his suggestions regarding the format of the paper.

## REFERENCES

1. R. C. Buck, *Bounded continuous functions on a locally compact space*, Michigan Math. J., **5** (1958), 95–104.

2. A. K. Chilana, *Invariant subspaces for linear operators in locally convex spaces*, J. London Math. Soc., (2), **2** (1970), 493–503.

3. ———, *Some special operators and new classes of locally convex spaces*, Proc. Camb. Phil. Soc., **71** (1972), 475–489.

4. J. B. Conway, *The strict topology and compactness in the space of measures*, Bull. Amer. Math. Soc., **72** (1966), 75–78.

5. Ed Dubinsky and J. R. Retherford, *Schauder bases and Köthe sequence spaces*, Trans. Amer. Math. Soc., Vol. **130**, No. 2, Feb. 1968, 265–280.

6. T. Ito and T. Seidman, *Bounded generators of linear spaces*, Pacific J. Math., (1968), 283-286.

7. G. Köthe, *Topological Vector Spaces I*, (English translation by D. J. H. Garling of Topologische lineare Räume I 1966), Springer-Verlag, 1969.

8. A. Mallios, *On functional representations of topological algebras*, J. Funct. Anal., Vol. **6**, No. 3, (1970), 468-480.

9. P. D. Morris and D. E. Wulbert, *Functional representation of topological algebras*, Pacific J. Math., **22** (1967), 323-337.

10. J. H. Webb, *Sequential convergence in locally convex spaces*, Proc. Camb. Phil. Soc., **64** (1968), 341-364.

11. A. Wiweger, *Linear spaces with mixed topology*, Studia Math., **20** (1961), 47-68.

Received April 25, 1972 and in revised form November 28, 1972.

HINDU COLLEGE, UNIVERSITY OF DELHI, DELHI 7, INDIA.

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# DIFFERENTIABLE OPEN MAPS OF $(p+1)$-MANIFOLD TO $p$-MANIFOLD

P. T. CHURCH AND J. G. TIMOURIAN

**Let $f: M^{p+1} \to N^p$ be a $C^3$ open map with $p \geqq 1$, let $R_{p-1}(f)$ be the critical set of $f$, and let**

$$\dim (R_{p-1}(f) \cap f^{-1}(y)) \leqq 0$$

**for each $y \in N^p$. Then (1.1) there is a closed set $X \subset M^{p+1}$ such that $\dim f(X) \leqq p-2$ and, for every $x \in M^{p+1} - X$, there is a natural number $d(x)$ with $f$ at $x$ locally topologically equivalent to the map**

$$\phi_{d(x)}: C \times R^{p-1} \to R \times R^{p-1}$$

**defined by**

$$\phi_{d(x)}(z, t_1, \cdots, t_{p-1}) = (\mathscr{R}(z^{d(x)}), t_1, \cdots, t_{p-1})$$

**($\mathscr{R}(z^{d(x)})$ is the real part of the complex number $z^{d(x)}$).**

The hypothesis on the critical set is essential [3, (4.11)], but in [4] we show that any real analytic open map satisfies this hypothesis, and thus this conclusion.

COROLLARY 1.2. *If $f: M^{p+1} \to N^p$ is a $C^{p+1}$ open map with* $\dim (R_{p-1}(f)) \leqq 0$, *then at each $x \in M^{p+1}$, $f$ is locally topologically equivalent to one of the following maps:*

(a) *the projection map $\rho: R^{p+1} \to R^p$,*

(b) *$\tau: C \times C \to C \times R$ defined by*
$\tau(z, w) = (2z \cdot \bar{w}, |w|^2 - |z|^2)$, *where $\bar{w}$ is the complex conjugate of $w$.*

(c) *$\psi_d: C \to R$ defined by $\psi_d(z) = \mathscr{R}(z^d)$.*

In order to read the proofs in this paper, the reader will need to have [3] at hand. In particular, the terms *locally topologically equivalent, branch set $B_f$, layer map, extended embedding,* and *0-regular* are defined in [3; (1.3), (1.5), (2.1), (2.3), and (4.1), respectively].

2. Spoke sets. The definition and lemmas of this section are given in somewhat greater generality than needed in this paper (i.e., for open maps), for use in a subsequent paper.

Let $\Gamma^2$ be any 2-manifold (without boundary).

DEFINITION 2.1. Let $\psi_w \times \iota: C \times R^{p-1} \to R \times R^{p-1}$ be defined by $\psi_0 \times \iota(z, t) = (|z|, t)$ and $\psi_w \times \iota(z, t) = (\mathscr{R}(z^w), t)$ $(w = 1, 2, \cdots)$. Thus

$B(\psi_1 \times \iota) = \varnothing$ and $B(\psi_w \times \iota) = \{0\} \times R^{p-1}$ otherwise. For $w = 0$ let $L = D^2 \times D^{p-1}$ and let $J = [-1, 1]$; for $w \geqq 1$ and $\eta > 0$ sufficiently small, let

$$L = (D^2 \times D^{p-1}) \cap (\psi_w \times \iota)^{-1} ([-\eta, \eta] \times D^{p-1})$$

and let $J = [-\eta, \eta]$. These examples motivate the following definition.

Let $f\colon \Gamma^2 \times R^{p-1} \to R \times R^{p-1}$ be a layer map, let $J = [b_0, b_1] \subset R$, and let $W \subset R^{p-1}$ be a closed $q$-cell ($q = 0, 1, \cdots, p-1$). Let $\{\gamma_j\}$ be a (possibly empty) collection of $2w$ disjoint closed arcs in $S^1 (j = 1, 2, \cdots, 2w)$; let $A = \bigcup_j \gamma_j$, and let $\zeta\colon S^1 \times W \to \Gamma^2 \times W$ be a layer embedding such that $B_f \cap \operatorname{imag} \zeta = \varnothing$, $f \circ \zeta\colon \gamma_j \times W \approx J \times W$, and for each component $\Phi$ of $\operatorname{Cl}[S^1 - A]$, $f(\zeta(\Phi \times W)) = \{b_i\} \times W (i = 0$ or $1)$. A *spoke set of $f$ over $J \times W$* is (i) a compact, connected subspace $L \subset f^{-1}(R \times W)$ such that (ii) $L \cap (\Gamma^2 \times \{t\})$ is a 2-cell for each $t \in W$ and (iii) for some $\zeta$ as above, the boundary $\Omega$ of $L$ with respect to $f^{-1}(R \times W)$ is $\operatorname{imag} \zeta$. Thus if $A = \varnothing$, $f(\Omega) = \{b_i\} \times W$ ($i = 0$ or 1). (In case $A \neq \varnothing$ and $q = 1$, $L$ is homeomorphic to the hub and spokes of a wagon wheel, where $\zeta(A \times W)$ corresponds to the ends of the spokes.) The index $\xi(L) = 1 - w$.

LEMMA 2.2. *Let $f\colon \Gamma^2 \times R^{p-1} \to R \times R^{p-1}$ be a layer map with* $\dim (B_f \cap (\Gamma^2 \times \{t\})) = \dim (f(B_f) \cap (R \times \{t\})) \leqq 0$ *for each $t \in R^{p-1}$, let $E \subset B_f$ be compact, let $a \in R^{p-1}$, and let $\varepsilon > 0$. Then there are a closed $(p-1)$-cell neighborhood $W$ of $a$, closed intervals $J_j (j = 1, 2, \cdots, m)$, and spoke sets $L_j$ over $J_j \times W$ such that*

(iv) *$E \cap L_j \neq \varnothing$ and $E \cap (\Gamma^2 \times W) \subset \bigcup_j (L_j - \Omega_j)$,*

(v) *the $L_j - \Omega_j$ are mutually disjoint, and*

(vi) *each* $\operatorname{diam} L_j < \varepsilon$.

*Proof.* Let $F$ be a compact neighborhood of $E$ in $\Gamma^2 \times R^{p-1}$, let $\{U_\alpha\}$ be a cover of $\Gamma^2$ by interiors of closed 2-cells, and let $\delta$ be the Lebesgue number of $\{U_\alpha \times R^{p-1}\}$ as a cover of $F$. We may suppose that $\varepsilon < \min (\delta, d(E, bdy\ F))$. Thus

(1) for each $\Psi \subset F$ with $\operatorname{diam} \Psi < \varepsilon$, there is a closed 2-cell $U$ with $\Psi \subset (\operatorname{int} U) \times R^{p-1}$.

Given $y \in R$ with $(y, a) \in f(E)$ and $X = E \cap f^{-1}(y, a)$, let $Q$ be the finite set and $\nu\colon Q \times D \to \Gamma^2 \times R^{p-1}$ be the extended embedding with $\operatorname{imag} \nu \cap B_f = \varnothing$ given by [3, (2.5)] for $X$ and $\varepsilon$. According to that lemma each component $K$ of $f^{-1}(\operatorname{int} D)$-$\operatorname{imag} \nu$ meeting $X$ has $\operatorname{diam} K < \varepsilon$, and each is open. Since $X = E \cap f^{-1}(y, a)$ and $E$ is compact, one may prove (by contradiction) that it is possible to select the $p$-cell neighborhood $D$ of $(y, a)$ in $R \times R^{p-1}$ sufficiently small that each component $K$ of $f^{-1}(\operatorname{int} D) - \operatorname{imag} \nu$ meeting $E$ has $\operatorname{diam} K < \varepsilon$. Summarizing.

(2) each component $K$ of $f^{-1}(\operatorname{int} D)$-$\operatorname{imag} \nu$ with $K \cap E \neq \varnothing$ has $\operatorname{diam} K < \varepsilon$, so that $\bar{K} \subset \operatorname{int} F$.

Choose a closed interval $J(y) \subset R$ with $y \in \operatorname{int} J(y)$,

$$J(y) \times \{a\} \subset \operatorname{int} D ,$$

and end points $b_0(y)$, $b_1(y)$ with $(b_0(y), a)$, $(b_1(y), a) \notin f(B_f)$. Since $f(F \cap B_f)$ is closed, there is a closed $(p-1)$-cell neighborhood $W(y)$ of $a$ in $R^{p-1}$ such that $(\partial J(y) \times W(y)) \cap f(F \cap B_f) = \varnothing$ and

$$J(y) \times W(y) \subset D .$$

Let $\nu(y)$ be the corresponding extended embedding (restricted) over $J \times W$.

There are $y_1, y_2, \cdots, y_u \in R$ with $(y_j, a) \in f(E)$ and

$$f(E) \cap (R \times \{a\}) \subset \bigcup\nolimits_j \operatorname{int} (J(y_j)) \times \{a\} .$$

The points $\{b_i(y_j)\colon i = 0, 1;\ j = 1, 2, \cdots, u\}$ are the end points of a finite set of closed intervals with mutually disjoint interiors; let $J_h (h = 1, 2, \cdots, r)$ be those intervals with $(J_h \times \{a\}) \cap f(E) \neq \varnothing$. Let $W$ be a closed $(p-1)$-cell neighborhood of $a \in R^{p-1}$ with $W \subset \bigcap_j W(y_j)$. Then $(\partial J_h \times W) \cap f(F \cap B_f) = \varnothing$ and

$$f(E) \cap (R \times W) \subset \bigcup\nolimits_h ((\operatorname{int} J_h) \times W) \quad (h = 1, 2, \cdots, r) .$$

Since each $J_h$ is contained in some $J(y_j)$, restriction of $\nu(y_j)$ yields an extended embedding $\nu_h$ over $J_h \times W$.

Let $J = [b_0, b_1]$ be one of these intervals $J_h$, let

$$\nu\colon (Q \times J) \times W \longrightarrow \Gamma^2 \times R^{p-1}$$

be the layer embedding $\nu_h$, and let $P \subset F$ be a component of

$$f^{-1}(\{b_i\} \times W) - \operatorname{imag} \nu .$$

Since $(\{b_i\} \times W) \cap f(F \cap B_f) = \varnothing$, $f^{-1}(\{b_i\} \times W) \cap \operatorname{int} F$ is a $p$-manifold, $\bar{P}$ is a compact connected $p$-manifold with boundary, and [3, (1.9)] $f \mid \bar{P}\colon \bar{P} \to \{b_i\} \times W$ is a bundle map. Thus [11; p. 53, (11.4)] it is a product bundle map, and since $f$ is a layer map

(3) there is a layer embedding $\lambda\colon \Lambda^1 \times W \to \Gamma^2 \times W$, where $\lambda(\Lambda^1 \times W) = \bar{P}$ and $\Lambda^1 \approx S^1$ or $[0, 1]$.

In particular, $P \cap (\Gamma^2 \times \{s\})$ is a component of $f^{-1}(b_i, s) - \operatorname{imag} \nu$ $(s \in W;\ i = 0, 1)$, and $\operatorname{Cl} [P \cap (\Gamma^2 \times \{s\})] \approx \Lambda^1$. From the compactness of $F$ and the finiteness of $Q$, the number of such components $P$ is finite.

Let $K$ be a component of $f^{-1}(J \times W)$-imag $\nu$ meeting $E$ (thus by (2) $\operatorname{diam} K < \varepsilon$ and $\bar{K} \subset \operatorname{int} F$) and let $T$ be a component of the boundary of $K$ in (i.e., relative to) $\Gamma^2 \times W$. Then

$$T \subset f^{-1}(\{b_0, b_1\} \times W) \cup \operatorname{imag} \nu .$$

Moreover, from (3) there are a finite union (possibly empty) $A$ of disjoint arcs in $S^1$ and a layer embedding $\zeta: S^1 \times W \rightarrow \Gamma^2 \times W$ with $\operatorname{imag} \zeta = T$, $\zeta(A \times W) = T \cap \operatorname{imag} \nu$, and

$$\zeta(\operatorname{Cl}[S^1 - A] \times W) = T \cap f^{-1}(\{b_0, b_1\} \times W) .$$

For each $s \in W$ and component (arc) $\gamma$ of $A$, $f \circ \zeta: \gamma \times s \approx J \times s$, and for each component $\Delta$ of $\operatorname{Cl}[S^1 - A]$, $f(\zeta(\Delta \times \{s\})) = (b_i, s)$ ($i = 0$ or 1). Thus if $A \neq \varnothing$, there are an even number of such components (arcs) $\Delta$, and they alternate in value. Hence there are an even number (possibly zero) of components (arcs) of $A$.

The union of such embeddings $\zeta$ over all $J \in \{J_h: h = 1, 2, \cdots, r\}$ and components $K$ of $f^{-1}(J \times W) - \operatorname{imag} \nu$ is finite: call them

$$\zeta_j \ (j = 1, 2, \cdots, k) .$$

Let $\Omega_j = \operatorname{imag} \zeta_j$ and let $K_j$ be the corresponding component $K$; by (1) there is a closed 2-cell $U_j \subset \Gamma^2$ with $\bar{K}_j \subset (\operatorname{int} U_j) \times W$, and thus each $\bar{K}_j \cap (\Gamma^2 \times \{s\})$ is a 2-cell-with-holes contained in $\operatorname{int} U_j$. Each $\Omega_j$ separates $U_j \times W$ into two components; let $L_j$ be the closure of the component disjoint from $\partial U_j \times W$. Each $L_j \cap (\Gamma^2 \times \{s\})$ is a 2-cell, and since the $K_j$ are mutually disjoint, for $i \neq j$ exactly one of the following is true: $(L_i - \Omega_i) \cap (L_j - \Omega_j) = \varnothing$, $L_i \subset L_j$, or $L_j \subset L_i$. The desired spoke sets are those $L_j$ with $E \cap L_j \neq \varnothing$ and $L_j \not\subset L_i$ for any $i \neq j$. Since each $\operatorname{diam} K_j < \varepsilon$, each $\operatorname{diam} \Omega_j < \varepsilon$, so that $\operatorname{diam} L_j < \varepsilon$. Since $E \cap (\Gamma^2 \times W) \subset \bigcup_j K_j \subset \bigcup_j L_j$, $E \subset B_f$, and $B_f \cap \Omega_j = \varnothing$, $E \cap (\Gamma^2 \times W) \subset \bigcup_j (L_j - \Omega_j)$.

LEMMA 2.3. *Let $f: \Gamma^2 \times R^{p-1} \rightarrow R \times R^{p-1}$ be a layer map, let $L_0$ (resp., $L_j, j = 1, 2, \cdots, q$) be a spoke set over $J \times W$(resp., $J_j \times W'$), and let $s \in W \cap W'$. Suppose that $L_j \cap (\Gamma^2 \times \{s\}) \subset L_0$,*

$$B_f \cap L_0 \cap (\Gamma^2 \times \{s\}) \subset \bigcup_{j>0} (L_j - \Omega_j) ,$$

*and the $L_j - \Omega_j$ are mutually disjoint ($j > 0$). Then*

$$\xi(L_0) = \sum_{j>0} \xi(L_j) .$$

*Proof.* Since $B(f_s) \subset B_f \cap (\Gamma^2 \times \{s\})$ and $\xi(L_j) = \xi(L_j \cap (\Gamma^2 \times \{s\}))$, it suffices to prove the lemma for $f = f_s: \Gamma^2 \rightarrow R$. Thus $L_j \subset L_0$ and $B_f \cap L_0 \subset \bigcup_{j>0} L_j - \Omega_j$. If $A_j$ (see (2.1)) has 2 $w(j)$ components

$(w(j) = 0, 1, \cdots)$, define $g_j$: $L_j \to R$ to agree with $f$ on $\partial L_j = \Omega_j$ and to be topologically equivalent to $\psi_{w(j)}$. Let $h$: $L_0 \to R$ agree with $f$ on $L_0 - \bigcup_{j>0}(L_j - \Omega_j)$ and with $g_j$ on $L_j$ $(j = 1, 2, \cdots, q)$. Then $B(h) = \bigcup_{j>0} B(g_j)$, and so is discrete.

Let $D(L_j)$ be the identification space obtained from

$$(L_j \times \{0\}) \cup (L_j \times \{1\})$$

by identifying $(x, 0)$ with $(x, 1)$ for each $x \in A = A(L_j)$, let $D(g_j)$: $D(L_j) \to R$ be defined by $D(g_j)\ (x, 0) = D(g_j)\ (x, 1) = g_j(x)$, and let $D(h)$ be defined analogously. Define a vector field $u_j$ (resp., $v$) on $D(L_j)$ (resp., $D(L_0)$) which is 0 precisely on the (discrete) branch set $B(D(g_j))$ (resp., $B(D(h))$) and elsewhere is transverse to the level curves of $D(g_j)$ (resp., $D(h)$), i.e., a "gradient vector field" $(j = 0, 1, \cdots, q)$. For any vector field $\alpha$ with isolated zeros, let the sum of the indices of $\alpha$ at its zeros [7, p. 32] be denoted by $\iota(\alpha)$.

Since $L_j \approx D^2$, the Euler characteristic

$$\chi(D(L_j)) = 2 - 2w(j) = 2\xi(L_j)\ .$$

According to the Poincaré-Hopf Theorem [7, p. 35] (differentiability is not really needed in our case) $\chi(D(L_j)) = \iota(u_j)$, so that $2\xi(L_j) = \iota(u_j)$ and $2\xi(L_0) = \iota(u_0) = \iota(v)$. Thus $2\xi(L_0) = \iota(v) = 2\sum_{j>0} \iota(v \mid L_j)$ (by definition of $\iota$) $= \sum_{j>0} \iota(u_j) = 2\sum_{j>0} \xi(L_j)$, so that $\xi(L_0) = \sum_{j>0} \xi(L_j)$ (where $j = 1, 2, \cdots, q$).

Alternatively, we could have used [5, p. 370] or [10, p. 35, (4.3.6)]; in this case we would have removed an open 2-cell with boundary a level circle about each local maximum or minimum point of $g_j$ and $h$, in order to have open maps. Or, we could have used a counting argument based on the Euler characteristics of $L_j$, $L_0$, and $L_0 - \bigcup_j \operatorname{int} L_j$; the first two spaces are 2-cells, and the last one is disjoint from $B_f$, so that information about it can be obtained from [3, (1.9)].

## 3. Spoke sets of open maps.

LEMMA 3.1. *Let $f$: $\Gamma^2 \times R^{p-1} \to R \times R^{p-1}$ be an open layer map, and let $L_0$ be a spoke set over $J \times W$, where $W$ is a closed $(p-1)$-cell. Then*

(a) $f^{-1}(y, t) \cap L_0$ *does not contain a homeomorph of* $S^1$

$$((y, t) \in R \times R^{p-1})$$

(b) $\xi(L_0) \leqq 0$;

(c) $f(L_0) = J \times W$;

(d) $\xi(L_0) \neq 0$ *implies that* $B_f \cap (L_0 - \Omega_0) \cap (\Gamma^2 \times \{t\}) \neq \emptyset$ *for every* $t \in R^{p-1}$;

(e) *if* $\dim (f(B_f) \cap (R \times \{t\})) \leqq 0$ *for every* $t \in R^{p-1}$,

$$\dim (B_f \cap f^{-1}(y, t)) \leqq 0 \quad \textit{for every } (y, t) \in R \times R^{p-1} ,$$

*and* $\xi(L_0) = 0$, *then* $B_f \cap \operatorname{int} L_0 = \emptyset$.

*Proof*. Suppose (a) is false, where $\Lambda$ is the homeomorph of $S^1$. Then $\Lambda$ bounds an open 2-cell $\Delta$ in $L_0 \cap (\Gamma^2 \times \{t\}) \approx D^2$. Since $f_t$: $\Gamma^2 \rightarrow R$ is open, $f_t(\Delta)$ is an open interval, while $f_t(\bar{\Delta})$ is a closed interval with $f_t(\partial \Delta)$ a single point, and a contradiction results.

If $\xi(L_0) > 0$, then $\Omega_0 \cap (\Gamma^2 \times \{t\})$ is a component of $f^{-1}(y, t)$ for some $y \in R$, and a contradiction of (a) results. Thus (b) is true.

From the definition of $L_0$ (2.1), $f(L_0) \subset J \times W$, and from that definition and (b), $f(\Omega_0) = J \times W$, so that (c) $J \times W = f(L_0)$.

If $B_f \cap (L_0 - \Omega_0) \cap (\Gamma^2 \times \{t\}) = \emptyset$ for some $t \in W$, then

$$g \colon L_0 \cap (\Gamma^2 \times \{t\}) \longrightarrow J \times \{t\}$$

defined by restriction of $f$ has $B_g = \emptyset$ [3, (4.10)], and so is a bundle map [3, (1.9)]. Thus [11, p. 53, (11.4)] $L_0 \cap (\Gamma^2 \times \{t\}) \approx J \times F$, where the fiber $F$ is a 1-manifold with boundary. Since $J \times F \approx D^2$ (2.1) (ii), $F$ is connected and $F \not\approx S^1$. Thus $F \approx [0, 1]$, so that $\xi(L_0) = 0$. Conclusion (d) results.

For a spoke set $L$ of $f$ over $I \times U$, let $^*L$ be $L \cap f^{-1}(\operatorname{int}(I \times U))$; thus $^*L - \Omega = \operatorname{int} L$ (interior relative to $\Gamma^2 \times R^{p-1}$). Since the restriction map $\alpha$: $f^{-1} (\operatorname{int} (J \times W)) \rightarrow \operatorname{int} (J \times W)$ is open, $^*L_0 - \Omega_0$ is open in $f^{-1}(\operatorname{int} (J \times W))$, and $B(f \mid L_0) \cap \Omega_0 = \emptyset$, the restriction map $\beta_0$: $^*L_0 \rightarrow \operatorname{int} (J \times W)$ is open. Suppose that $f$ satisfies the hypotheses of (e), i.e., $\xi(L_0) = 0$, while $(x, s) \in B_f \cap \operatorname{int} L_0$. Given $\varepsilon > 0$, which we may assume is less than $d(B_f, \Omega_0)$, let $W'$ and the spoke sets $L_j (j = 1, 2, \cdots, q)$ be as given by (2.2) for $f$, $\varepsilon$, $a = s$, and $E = (B_f \cap L_0)$, where $(x, s) \in \operatorname{int} L_1$. From (b) each $\xi(L_j) \leqq 0$ and from (2.3) $\xi(L_0) = \sum_{j>0} \xi(L_j)$; thus $\xi(L_j) = 0$ for every $j$, so in particular $\xi(L_1) = 0$. Let $\beta_1$: $^*L_1 \rightarrow f(^*L_1)$ be restriction of $f$.

For each $(z, t) \in f(L_i) - f(B_f)$, $(i = 0, 1)$, $(\beta_i)^{-1}(z, t)$ is a 1-manifold with boundary; by (a) each of its components is homeomorphic to $[0, 1]$, and since $\xi(L_i) = 0$, $(\beta_i)^{-1}(z, t) \approx [0, 1]$. By [3, (4.3)(a)] $(\beta_i)^{-1}(y, u)$ is arcwise connected for each $(y, u) \in \operatorname{imag} \beta_i$. Choose $\delta > 0$ such that $S((x, s), \delta) \subset \operatorname{int} L_1$. Then

$$f^{-1}(y, u) \cap S(x, \delta) \subset (\beta_1)^{-1}(y, u) \subset f^{-1}(y, u) \cap S((x, s), \varepsilon) ,$$

so that $f$ is 0-regular at $(x, s)$ [3, (4.1)]. Since $(x, s) \in B_f \cap L_0$ is arbitrary, by [3, (4.2)] $f$ is 0-regular at each point of $L_0$. Thus $\beta_0$ is a bundle map [3, (4.3) (b)], so that $B_f \cap \operatorname{int} L_0 = \varnothing$ .

LEMMA 3.2. *Let $g\colon \Gamma^2 \times R^{p-1} \to R \times R^{p-1}$ be an open layer map, let $L$ be a spoke set over $J \times W$ where $W$ is a $(p-1)$-cell and let $\alpha$; $W \approx B_g \cap L$ with $\pi \circ \alpha$ the identity map. Then $g \mid int\ L$ is topologically equivalent to $\psi_w \times \iota$ ($w = 2, 3, \cdots$; see (2.1)).*

*Proof.* We may as well replace $g$ by its restriction to $g^{-1}(\operatorname{int} J \times \operatorname{int} W)$, and $L$ by $L \cap g^{-1}(\operatorname{int} J \times \operatorname{int} W)$, i.e., we may as well suppose that $\operatorname{int} J = R$ and $\operatorname{int} W = R^{p-1}$. Let $h\colon R \times R^{p-1} \to R \times R^{p-1}$ be the layer homeomorphism defined by $h(y, t) = (y, t) - g(\alpha(t))$, and let $\lambda = h \circ g \mid L$. Then $B_\lambda = B_g \cap L$ and $\lambda(B_\lambda) = \{0\} \times R^{p-1}$.

Let $J_i$ be $(-\infty, 0]$ or $[0, \infty)$ according as $i$ is odd or even. (1) Let $K$ be a component of $\lambda^{-1}((\operatorname{int} J_i) \times R^{p-1})$, and let $\beta\colon K \to \operatorname{int} J_i \times R^{p-1}$ and $\gamma\colon \bar{K} \to J_i \times R^{p-1}$ be the restriction of $\lambda$. Since $B_\beta = \varnothing$, $\beta$ is a bundle map with fiber a 1-manifold $F$ [3, (1.9)], and so $K \approx F \times \operatorname{int} J_i \times R^{p-1}$ [11, p. 53, (11.4)]. Since $K$ is connected, $F$ is also, and by (3.1(a)) $F \approx [0, 1]$. By [3, (4.3)(a)], $\gamma^{-1}(0, t)$ is arcwise connected for each $t \in R^{p-1}$.

Given $(x, s) \in B_\gamma \cap \gamma^{-1}(\{0\} \times R^{p-1})$ and $\varepsilon > 0$ with $S((x, s), \varepsilon) \subset \operatorname{int} L$, let $L'$ be a spoke set over $J' \times W'$ given by (2.2) for $\lambda$, $E = \{(x, s)\}$, $a = s$, and $\varepsilon$. Then $L'$ satisfies the original hypotheses, so that $(r')^{-1}(y, t)$ is arcwise connected for every $(y, t)$. Choose $\delta > 0$ with $S((x, s), \delta) \subset \operatorname{int} L'$. Then

$$S((x, s), \delta) \cap \gamma^{-1}(y, t) \subset (\gamma')^{-1}(y, t) \subset S((x, s), \varepsilon) \cap \gamma^{-1}(y, t)$$

for each $(y, t) \in J' \times W'$, so that $\gamma'$ is 0-regular at $(x, s)$. By [3, (4.2)] $\gamma$ is 0-regular, and (by [3, (4.3)(b)]) (2) $\gamma$ is a (product) bundle map with fiber $[0, 1]$.

For each $t \in R^{p-1}$ and component $K$ (see (1)), $\gamma \mid (\bar{K} \cap (\Gamma^2 \times \{t\}))$ is a product bundle map over $J_i \times (t)$ with fiber $[0, 1]$, so that $\lambda^{-1}(0, t)$ is a deformation retract of $L \cap (\Gamma^2 \times \{t\}) \approx D^2$. Thus $\lambda^{-1}(0, t)$ is connected. Since $\lambda^{-1}(0, t)$ contains no homeomorph of $S^1$ (3.1(a)), and $\lambda^{-1}(0, t) - \{\alpha(t)\}$ is a 1-manifold with boundary points the $2w$ ($\xi(L) = 1 - w$) points of $\lambda^{-1}(0, t) \cap \Omega$ (2.1), it follows that $\lambda^{-1}(0, t)$ is homeomorphic to the union of $2w$ arcs disjoint except for their common endpoint $\alpha(t)$. As a result $\alpha(t) \in \bar{K} \cap (\Gamma^2 \times \{t\})$, so that each $\bar{K}$ contains imag $\alpha$, i.e., $B_\lambda$.

Let $K_i$ ($i = 1, 2, \cdots, 2w$) be the components $K$ enumerated so that for any $t \in R^{p-1}$, $(\operatorname{int} K_i) \cap (\Gamma^2 \times \{t\})$ are the components of

$$(\operatorname{int} L) \cap ((\Gamma^2 \times \{t\}) - \lambda^{-1}(0, t))$$

in counterclockwise order around $\alpha(t)$ with $\lambda(\bar{K}_i) = J_i \times R^{p-1}$. Let

$\Lambda_i = \bar{K}_i \cap \operatorname{int} L$, let $\psi = \psi_w \times \iota$ (see (2.1)), and let $\Delta_i$ be the closures of the components of $\psi^{-1}(\operatorname{int} J_i \times R^{p-1})$ enumerated in analogous fashion.

By (2) there is an orientation-preserving homeomorphism $\mu_i$ of $\Lambda_i$ onto $R \times J_i \times R^{p-1}$ with $\pi \circ \mu_i = \lambda \mid \Lambda_i$. Let $\nu_i$ be the homeomorphism of $R \times J_i \times R^{p-1}$ onto itself defined by

$$\nu_i(x, y, t) = (x, y, t) - \mu_i(\alpha(t)) + (0, 0, t) ,$$

and let $\zeta_i = \nu_i \circ \mu_i$. Then $\zeta_i(\alpha(t)) = (0, 0, t)$, so that

$$\zeta_i(B_\lambda) = \{0\} \times \{0\} \times R^{p-1} .$$

There is an analogous orientation-preserving homeomorphism $\xi_i$ of $\Delta_i$ onto $R \times J_i \times R^{p-1}$ with $\pi \circ \xi_i = \psi \mid \Delta_i$ and $\xi_i(B_\psi) = \{0\} \times \{0\} \times R^{p-1}$.

Let $\Phi = (\operatorname{int} L) \cap \lambda^{-1}(\{0\} \times R^{p-1})$, and let $\Upsilon_i$ (resp., $\Psi_i$) be the closure in $\Phi$ (resp., $\psi^{-1}(\{0\} \times R^{p-1})$) of the component in $\Phi - \beta_\lambda$(resp., $\Psi^{-1}(\{0\} \times R^{p-1}) - B_\psi$) meeting both $\Lambda_i$ and $\Lambda_{i+1}$ (resp., $\Delta_i$ and $\Delta_{i+1}$), where $i$ and $i + 1$ are interpreted mod $2w$. In case $w = 1$ there are two such components, and $\Upsilon_i$ is so chosen that, for each $t \in R^{p-1}$, a counter-clockwise path around $\alpha(t)$ from $\Lambda_i$ to $\Lambda_{i+1}$ passes through $\Upsilon_i$. Then $(\xi_i)^{-1} \circ \zeta_i$ (also $(\xi_{i+1})^{-1} \circ \zeta_{i+1}$) defines a homeomorphism of $\Upsilon_i$ onto $\psi_i$ with $(\xi_i)^{-1} \circ \zeta_i(B_\lambda) = B_\psi$. Let $\rho$: $\Phi \approx \psi^{-1}(\{0\} \times R^{p-1})$ agree with $(\xi_i)^{-1} \circ \zeta_i$ on $\Upsilon_i$.

Let $\sigma_i$ be the layer homeomorphism of $R \times \{0\} \times R^{p-1}$ onto itself which is the restriction of $\xi_i \circ \rho \circ \zeta_i^{-1}$, (on $\zeta_i(\gamma_{i-1})$, $\sigma_i$ agrees with the identity map) and let $\tau_i$ be its first coordinate map. Let $\phi_i$ be the homeomorphism of $R \times J_i \times R^{p-1}$ onto itself defined by $\phi_i(x, y, t) = (\tau_i(x, t), y, t)$, and let $\chi_i = (\xi_i)^{-1} \circ \phi_i \circ \zeta_i$. Then $\chi_i$: $\Lambda_i \approx \Delta_i$, they agree with $\rho$, and they thus define $\chi$: $\operatorname{int} L \approx C \times R^{p-1}$; since $\pi \circ \zeta_i = \lambda \mid \Lambda_i$ and $\pi \circ \xi_i = \psi \mid \Delta_i$, where $\pi$: $R \times J_i \times R^{p-1} \to J_i \times R^{p-1}$ is projection, $\psi \circ \chi = \lambda \mid \operatorname{int} L$. This is the desired conclusion.

## 4. The Proof of the theorem.

REMARK 4.1. According to the Rank Theorem [3, (1.6)] $B_f \subset R_{p-1}(f)$, and we prove (1.1) under the weaker hypothesis that $\dim (B_f \cap f^{-1}(y)) \leqq 0$ for each $y \in N^p$.

*Proof*. Let $X$ be the complement of the set on which $f$ has the desired structure; then $X \subset B_f$ is closed. We suppose that

$$\dim f(X) \geqq p - 1 ,$$

and will obtain a contradiction.

Since $f$ is $C^3$, $\dim (f(R_{p-2}(f))) \leqq p - 2$ [2, p. 1037]. If, for every

$x \in M^{p+1} - f^{-1}(f(R_{p-2}(f)))$, there is an open neighborhood

$$U_x \subset M^{p+1} - f^{-1}(f(R_{p-2}(f)))$$

of $x$ with $\bar{U}_x$ compact and $\dim(f(U_x \cap X)) \leqq p-2$, it follows from the fact that $\{U_x\}$ has a countable subcover that $\dim(f(X)) \leqq p-2$. Thus, there is an $\bar{x} \in M^{p+1} - f^{-1}(f(R_{p-2}(f)))$ such that, (1) for every open neighborhood $U \subset M^{p+1} - f^{-1}(f(R_{p-2}(f)))$ of $\bar{x}$, $\dim(f(U \cap X)) \geqq p-1$.

By [1, p. 87, (1.1)] there are open neighborhoods $U$ of $\bar{x}$ and $V$ of $f(\bar{x})$ and $C^r$ diffeomorphisms $\sigma\colon R^2 \times R^{p-1} \approx U$ and $\rho\colon V \approx R \times R^{p-1}$ such that $\rho \circ f \circ \sigma = g$ is a $C^r$ layer map and $\sigma(0, 0) = \bar{x}$. By hypothesis $\dim(B_g \cap g^{-1}(y, t)) \leqq 0$ for each $(y, t) \in R \times R^{p-1}$.

Since $\sigma^{-1}(X) \subset B_g$, $B_g \subset R_{p-1}(g)$ (by the Rank Theorem [3, (1.6)]), $R_{p-1}(g) \cap (R^2 \times (t)) = R_0(g_t)$, and $\dim(g_t(R_0(g_t))) \leqq 0$ by Sard's Theorem (e.g. [2, p. 1037]), (2) $\dim(g(B_g) \cap (R^2 \times \{t\})) \leqq 0$ and

$$\dim(g(\sigma^{-1}(X)) \cap (R \times \{t\})) \leqq 0 .$$

On the other hand, (by (1)) $\dim(g(\sigma^{-1}(X)) \geqq p-1$, so there is an $r > 0$ such that

$$A = (\mathrm{Cl}[S(0, r)] \times R^{p-1}) \cap \sigma^{-1}(X)$$

has $\dim g(A) \geqq p-1$. If $\pi\colon R \times R^{p-1} \to R^{p-1}$ is projection, then $\dim(\pi(g(A))) \geqq p-1$ (by (2) and [6, p. 91]), and there is an open $(p-1)$-cell $T \subset \pi(g(A))$ [6, p. 44] with $\bar{T}$ compact. Thus (3)

$$A \cap (R^2 \times \{t\}) \neq \varnothing \qquad \text{for each } t \in T .$$

Let $W \subset T$ and the spoke sets $L_j$ $(j = 1, 2, \cdots, q)$ be as given by (2.2) for $g$, any $a \in T$, $E = A \cap (R^2 \times \bar{T})$, and (say) $\varepsilon = 1$. If (4) (i) the cardinality $w(t) \geqq 1$ of $B_g \cap (R^2 \times \{t\}) \cap (\bigcup_j L_j)$ $(t \in \operatorname{int} W)$ is bounded above by $|\sum_j \xi(L_j)|$, choose $s \in \operatorname{int} W$ such that $w(s)$ is maximal and let $(x_i, s)$ $(i = 1, 2, \cdots, w(s))$ be these points. Otherwise, (4) (ii) there are $s \in \operatorname{int} W$ and distinct points $(x_i, s)$ $(i = 1, 2, \cdots, |\sum_j \xi(L_j)| + 1)$ of $B_g \cap (R^2 \times \{t\}) \cap (\bigcup_j L_j)$. Let $w'$ be $w(s)$ in case (4) (i) and $|\sum_j \xi(L_j)| + 1$ in case (4) (ii). Let $\varepsilon > 0$ be less than $d(x_h, x_i)$ for $h \neq i$ and $d(B_g, \bigcup_j \Omega_j)$, and let $W' \subset \operatorname{int} W$ and $\{L'_k\}$ be as given by (2.2) for $g$, $a = s$, $E = \bigcup_j L_j \cap B_g$, and this $\varepsilon$. Thus (5) the $(x_i, s)$, are in distinct spoke sets $L'_k$.

By hypothesis and by (2), the hypothesis of (3.1) (e) is satisfied, so that by (3.1) (d) and (e) $\xi(L_j) = 0$ if and only if $L_j \cap B_g = \varnothing$. We may thus omit those $L_j$ and $L'_k$ with $\xi(L_j) = 0 = \xi(L'_k)$. From (3.1) (b) each $\xi(L_j) < 0$ and $\xi(L'_k) < 0$, and from (5) and (3.1) (d) the cardinality $c$ of $\{L'_k\}$ satisfies $w' \leqq c \leqq |\sum_k \xi(L'_k)|$. Since each $L'_k$ is contained in some $L_j$, $\sum_j \xi(L_j) = \sum_k \xi(L'_k)$ by (2.3), and so $w' \leqq |\sum_j \xi(L_j)|$; this contradicts (4) (ii), and hence (4) (i) must be true.

For $t \in W'$, $w(t) \geqq c$ by (3.1) (d), while $c \geqq w(s)$ by (4) (i), so that $w(t) = w(s)$. Thus (by (3.1) (d)) each $B_g \cap (R^1 \times \{t\}) \cap L'_k$ is a single point for $t \in W'$, and since $B_g$ is closed, there is a homeomorphism $\alpha_i\colon W' \approx L'_k \cap B_g$ with $\pi \circ \alpha_i$ the identity map on $W'$. By (3.2) $\bigcup_k (\sigma^{-1}(X) \cap L'_k) = \emptyset$. But this set contains $\Lambda \cap (R^2 \times W')$, contradicting (3).

REMARK 4.2. In case $p = 1$, $C^3$ may be replaced by $C^2$ and the argument can be shortened considerably. In that case (4.1) results from [12, p. 103, Theorem 1] (cf. [18, pp. 7–8]), and (4.1) in case $B_f$ is discrete is [10, p. 28, (4.3.1)] and [9]. Considerable information relating to open maps $f\colon M^2 \to N^1$ is given in [5], [8], and [10].

4.3. *Proof of* (1.2). The hypotheses of (1.1) are satisfied (with $C^2$ if $p = 1$). In case $p = 1$, $X = \emptyset$, so that at each $x \in M^{p+1}$, $f$ at $x$ is locally topologically equivalent to $\psi_{d(x)}$. In case $p \geqq 2$, for each $x \in M^{p+1} - X$ with $d(x) \neq 1$ (i.e., $x \in B_f$), $\dim B_f = p - 1 \geqq 1$ in a neighborhood of $x$; the assumption that $\dim R_{p-1}(f) \leqq 0$ contradicts the Rank Theorem [3, (1.6)]. Thus $B_f \subset X$, so that

$$\dim f(B_f) \leqq p - 2 .$$

That $f$ is locally topological equivalent to $\rho$ or to $\tau$ is now a consequence of [3, (4.7)].

## REFERENCES


1. P. T. Church, *Differentiable open maps on manifolds*, Trans. Amer. Math. Soc., **109** (1963), 87–100.
2. ———, *On points of Jacobian rank k. II*, Proc. Amer. Math. Soc., **16** (1965), 1035–1038.
3. P. T. Church and J. G. Timourian, *Differentiable maps with O-dimensional critical set, I*, Pacific J. Math., (to appear).
4. ———, *Real analytic open maps*, Pacific J. Math., **41** (1972), 615–630.
5. W. C. Fox, *The critical points of peano interior functions defined on 2-manifolds*, Trans. Amer. Math. Soc., **83** (1956), 338–370.
6. W. Hurewicz and H. Wallman, *Dimension Theory*, 2nd edition, Princeton University Press, Princeton, N. J., 1948.
7. J. Milnor, *Topology from the Differentiable Viewpoint*, The University Press of Virginia, Charlottesville, Virginia, 1965.
8. M. Morse, *Topological Methods in the Theory of Functions of a Complex Variable*, Princeton University Press, Princeton, N. J. 1947.
9. W. D. Nathan, *Open mappings on 2-manifolds*, Pacific J. Math., **41** (1972), 495–501.
10. ———, *Open Mappings on Manifolds*, Ph. D. dissertation, Syracuse University, Syracuse, 1968.
11. N. Steenrod, *The Topology of Fiber Bundles*, Princeton University Press, Princeton, N. J., 1951.

12. Y. K. Toki, *A topological characterization of pseudo-harmonic functions*, Osaka Math. J., **3** (1951), 101-122.

Received June 21, 1972. Work of the first author supported in part by NSF Grant GP-6871, and that of the second author by NSF Grant GP-8888 and NRC Grant A7357.

SYRACUSE UNIVERSITY
UNIVERSITY OF TENNESSEE
AND
UNIVERSITY OF ALBERTA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# ON ADDITIVE FUNCTIONS WHOSE LIMITING DISTRIBUTIONS POSSESS A FINITE MEAN AND VARIANCE

P. D. T. A. ELLIOTT

**In this paper two characterizations are given of those additive arithmetic functions which possess a limiting distribution with a finite mean and variance. It turns out that the study of such functions fits naturally within the framework of the theory of Lambert series.**

1. An arithmetic function $f(n)$ is said to be *additive* if for every pair of coprime positive integers $a$ and $b$ the relation

$$f(ab) = f(a) + f(b)$$

is satisfied. If in addition the relations

$$f(p) = f(p^2) = \cdots$$

hold for each prime power then we say that $f(n)$ is *strongly additive*. For clarity of exposition only we shall confine ourselves to the study of strongly additive functions in this paper.

For each real number $x \geqq 1$ we define the frequency function

$$\nu_x(n; f(n) < z) = x^{-1} \sum_{\substack{n \leq x \\ f(n) < z}} 1 .$$

If as $x \to \infty$ these frequencies converge to a limiting distribution in the usual probabilistic sense then we say that $f(n)$ has a limiting distribution.

2. THEOREM. *For any (real valued) additive function $f(n)$ the following three propositions are equivalent*:

(i) *$f(n)$ has a limiting distribution with finite mean and variance.*

(ii) *The series*

$$\sum f(p)p^{-1} \text{ and } \sum f^2(p)p^{-1}$$

*both converge.*

(iii)

$$\limsup_{x \to \infty} x^{-1} \sum_{n \leq x} f^2(n) < \infty$$

*and*

$$\lim_{x\to\infty} x^{-1} \sum_{n\leq x} f(n)$$

*exist.*

REMARK. The equivalence of Propositions (i) and (ii) is exactly what one should expect from the interpretation of $f(n)$ as the sum of independent random variables which take (respective) values $f(p)$ with probability $p^{-1}$ and zero with probability $1 - p^{-1}$. More surprising, perhaps, is the fact that the hypothesis that $f(n)$ be additive improves the otherwise weak conditions (iii) to equivalence with (i). We shall (perhaps surprisingly) appeal to a result concerning Lambert series.

It will be clear that a form of theorem involving complex-valued additive functions could be proved if we confine our attention to the equivalence of Propositions (ii) and (iii).

3. *Proof that* (i) *implies* (ii).

We define the function

$$f^1(p) = \begin{cases} f(p) & \text{if} \quad |f(p)| < 1 \\ 1 & \text{otherwise} . \end{cases}$$

Then the Erdös-Wintner criterion (see for example Kubilius [3] Theorem 4.5 pp. 74–85) asserts that $f(n)$ possesses a limiting frequency (unrestricted) if and only if both of the series

$$\sum f'(p)p^{-1} \text{ and } \sum (f'(p))^2_{p^{-1}}$$

converge. Let $F(z)$ denote the limiting frequency guaranteed by (i). Then for any positive real number $B$ such that $\pm B$ are continuity points of $F(z)$ we see that

$$x^{-1} \sum_{\substack{n\leq x \\ |f(n)|\leq B}} f^2(n) \longrightarrow \int_{|z|\leq B} z^2 dF(z) \,, \quad (x \longrightarrow \infty) \,.$$

Next, for any real $\varepsilon > 0$ there is a number $A$ such that

$$\liminf_{x\to\infty} \nu_x(n; |f(n)| \leq A) > 1 - \varepsilon \,.$$

From the Erdös-Wintner criterion we see that those primes $q_j$ for which $|f(q_j)| \geq 1$ are such that the series

$$\sum q_j^{-1}$$

converges. Let us denote the set of these primes by $Q$.

A straightforward application of the sieve of Eratosthenes shows that those integers which are prime to every $q_j$ have a natural density. In fact we obtain

$$\nu_x(n;\, q_j \nmid n \forall j) \longrightarrow \prod_{i=1}^{\infty}\left(1 - \frac{1}{q_i}\right), \quad (x \longrightarrow \infty)\,.$$

Set $\alpha$ for this product, and let $A$ be chosen so that the second of our two assertions above holds with $\varepsilon = \alpha/2$. Let the integers $n_i$ run through all those integers $n$ which satisfy both

$$|f(n)| \leqq A \text{ and } q_j \nmid n \forall j\,.$$

From what we have so far said it is clear that

$$\liminf_{x\to\infty} \nu_x(n;\, n = n_i \leqq x) \geqq \alpha/2\,,$$

and in particular we have

$$\nu_x(n;\, n = n_i \leqq x) \geqq \alpha/4$$

for all $x \geqq x_0$, say.

Consider the sum

$$S_x = \sum_{n_i q_j \leqq x}\sum{}' f(n_i q_j)^2$$

where $'$ denotes that the side condition $2A < |f(q_j)| \leqq B - A$ is to be satisfied.

From these restrictions a typical summand satisfies

$$f^2(n_i q_j) \geqq (|f(q_j)| - A)^2 \geqq \frac{1}{4} f^2(q_j)$$

so that

$$\begin{aligned} S_x &\geqq \frac{1}{4}\sum_{q_j^{-1} \leqq x} f^2(q_j) \sum_{n_i \leqq x q^{-1}} 1 \\ &\geqq \frac{1}{4}\sum_{q_j \leqq x x_0^{-1}} f^2(q_j)\frac{1}{4}\alpha\frac{x}{q_j} \end{aligned}$$

and therefore

$$\begin{aligned} \limsup_{x\to\infty} \sum_{q_j \leqq x} \frac{f^2(q_j)}{q_j} &\leqq \limsup_{x\to\infty} x^{-1} S_x \\ &\leqq \int_{|z| \leqq B} z^2 dF(z) \leqq \int_{-\infty}^{\infty} z^2 dF(z)\,. \end{aligned}$$

Since these inequalities hold for any sequence of suitable continuity points $\pm B$ which tend (in absolute value) to infinity, we deduce that for any $B > 0, x \geqq 0$

$$\sum_{q_j \leqq x} f^2(q_j) \leqq \int_{-\infty}^{\infty} z^2 dF(z)$$

where

$$2A < |f(q_j)| \leqq B - A$$

so that letting $B \to \infty$ and then $x$ yields

$$\sum_{|f(q_j)| \geqq 2A} \frac{f^2(q_j)}{q_j} < \infty .$$

Moreover,

$$\sum_{1 \leqq |f(p)| \leqq 2A} \frac{|f(p)|}{p} \leqq 2A \sum_{j=1}^{\infty} \frac{1}{q_j} < \infty ,$$

and

$$\sum_{|f(p)|<1} \frac{f^2(p)}{p} < \infty$$

so that altogether the series

$$\sum f^2(p) p^{-1}$$

converges. The convergence of the second series in (ii) follows immediately.

*Proof that* (ii) *implies* (iii) *and* (i).

We begin with the remark that for any additive function, complex valued or otherwise, the Turan-Kubilius inequality (see for example Kubilius [3] pp. 31–35) asserts that for a suitable positive constant $c$

$$\sum_{n \leqq x} |f(n) - \sum_{p \leqq x} f(p) p^{-1}|^2 \leqq c \sum_{p \leqq x} |f^2(p)| p^{-1} , \quad (x \geqq 1) .$$

In our present circumstances the sums

$$\sum_{p \leqq x} f(p) p^{-1} \text{ and } \sum_{p \leqq x} f^2(p) p^{-1}$$

are uniformly bounded for all real values of $x$, so that

$$\begin{aligned} \sum_{n \leqq x} f^2(n) &\leqq 2 \sum_{n \leqq x} \Big( f(n) - \sum_{p \leqq x} f(p) p^{-1} \Big)^2 + 2x \Big( \sum_{p \leqq x} f(p) p^{-1} \Big)^2 \\ &= O(x) , \end{aligned}$$

and

$$\limsup_{x \to \infty} x^{-1} \sum_{n \leqq x} f^2(n) = D < \infty .$$

From the Erdös-Wintner criterion $f(n)$ possesses a limiting distribution $F(z)$, say. For each real number $B$ such that $\pm B$ are continuity points of this limiting distribution, an application of Fatou's lemma yields

$$\int_{|z| \leqq B} z^2 dF(z) \leqq \liminf_{x \to \infty} x^{-1} \textstyle\sum_{n \leqq x} f^2(n) \leqq D .$$

Since $B$ is otherwise arbitrary $F(z)$ has a finite second moment, and hence a finite mean and variance.

This completes the proof of (i).

Furthermore,

$$x^{-1} \sum_{\substack{n \leq x \\ |f(n)| \leq B}} f(n) \longrightarrow \int_{|z| \leq B} z dF(z) \ , \quad (x \to \infty) \ ,$$

whilst

$$\limsup_{x \to \infty} x^{-1} \sum_{\substack{n \leq x \\ |f(n)| > B}} |f(n)| \leqq B^{-1} \limsup_{x \to \infty} x^{-1} \sum_{n \leqq x} f^2(n) \leqq B^{-1} D$$

from which it follows trivially that as $x \to \infty$

$$x^{-1} \sum_{n \leqq x} f(n)$$

converges to the mean of $F(z)$.

This completes the proof of (iii).

*Proof that* (iii) *implies* (ii) (which will complete the proof of the theorem).

As one would expect this part of the proof takes a little more effort since we have to start, so to speak, from scratch. We recall that an additive function $f(n)$ is said to be *finitely distributed* if and only if there are two positive real numbers $c_1$ and $c_2$ so that for an unbound sequence of real numbers $x \geqq 1$ we can find at least $k \geqq c_2 x$ integers $1 \leqq a_1 < a_2 < \cdots < a_k \leqq x$ so that

$$|f(a_i) - f(a_j)| \leqq c_1$$

holds for every pair $(a_i, a_j)$, $1 \leqq i, j \leqq k$. This concept was introduced by Erdös [1] who proved

LEMMA 1. *A function $f(n)$ is finitely distributed if and only if there is a constant $c_3$ and an additive function $g(n)$ so that*

$$f(n) = c_3 \log n + g(n) \ ,$$

*where*

$$\sum (g'(p))^2 p^{-1} < \infty \ .$$

*There is an alternative proof, on somewhat different lines, given by Ryavec* [4].

In our present circumstances we have

$$x^{-1}\sum_{n\leq x} f^2(n) \leqq E$$

for all $x \geqq 2$ (say). Thus for any positive real number $A > E^{1/2}$,

$$\nu_x(n; |f(n)| \geqq A) \leqq EA^{-2} < 1 , \quad (x \geqq 2) .$$

It follows from Lemma 1 that $f(n)$ is finitely distributed, and has the form

$$c_3 \log n + g(n) .$$

Let $\pi$ denote the set of primes $q_j$ on which $|g(q_j)| > A$. Let $n_i$ run through those squarefree integers which are prime to each $q_j$. Since

$$\sum_{q\in\pi} q^{-1}$$

converges, a straightforward application of the sieve of Eratosthenes shows that

$$\nu_x(n; n = n_i \leqq x) \longrightarrow \prod_{i=1}^{\infty}\left(1 - \frac{1}{q_i + 1}\right) = \beta > 0 , \quad (x \longrightarrow \infty) ,$$

say. For each integer $n$ let $\nu(n)$ denote the number of distinct prime divisors of $n$. We next assume that $c_3 \neq 0$ and obtain a contradiction.

Let $c_4$ be sufficiently large that the inequality $A\nu(n) \leqq c_4 \log n$ holds for all integers $n \geqq 2$. Then for every real number $x \geqq 2$ we have

$$\begin{aligned} Ex &\geqq \sum_{n_i\leqq x} f^2(n_i) \geqq \sum_{n_i\leqq x} (c_3 \log n_i - A\nu(n_i))^2 \\ &= c_3^2 \sum_{n_i\leqq x} \log^2 n_i + O\left(\log x \sum_{n\leqq x} \nu(n)\right) . \end{aligned}$$

For all sufficiently large values of $x$ the first of these two terms is

$$(1 + o(1))\beta c_3^2 x \log^2 x$$

whilst the second is at most $O(x \log x \log\log x)$. This clearly yields a contradiction. Hence $c_3 = 0$ and the additive function $f(n)$ satisfies

$$\sum_p (f'(p))^2 p^{-1} < \infty .$$

We now argue exactly as in the proof that the existence of a limiting distribution for $f(n)$ which has a finite variance implies that the series

$$\sum f^2(p)p^{-1}$$

converges, and deduce the same result.

It remains to secure the convergence of the series

$$\sum f(p)p^{-1}\,.$$

(We do not as yet know that a limiting distribution for $f(n)$ exists, although if we set $\alpha_n = \sum_{p\leqq n} f'(p)p^{-1}$ then we do know that $f(n) - \alpha_n$ has a limiting distribution. See, for example, Kubilius [3] Theorem 4.4 pp. 72–74.)

Consider the generating function

$$G(z) = \sum_{n=1}^{\infty} f(n)z^n\,.$$

If $N$ is any positive integer and $z$ is any complex number then by the Cauchy-Schwarz inequality

$$\begin{aligned}\Big|\sum_{N<n\leqq 2N} f(n)z^n\Big|^2 &\leqq \sum_{n\leqq 2N} f^2(n) \sum_{N<n\leqq 2N} |z|^{2n}\\ &\leqq EN^2|z|^{2N}\,.\end{aligned}$$

It is easily seen that $G(z)$ is defined by an absolutely convergent series if $z$ satisfies $|z| < 1$. By means of the representation

$$f(n) = \sum_{p|n} f(p)$$

we invert the order of summation to obtain:

$$G(z) = \sum_p f(p)\frac{z^p}{1-z^p}\,.$$

Since

$$x^{-1}\sum_{n\leqq x} f(n) \longrightarrow A,\ (x \longrightarrow \infty)\ ,\ \text{say}\ ,$$

it is readily established that for real values of $z$

$$G(z) \sim \frac{A}{1-z} \quad \text{as } z \longrightarrow 1-\,.$$

We now appeal to a Tauberian theorem concerning Lambert series.

LEMMA 2. *Let $a_n$ $n = 1, 2, \cdots$ be a series of real numbers, and define*

$$H(y) = \sum_{n=1}^{\infty} a_n \frac{nye^{-ny}}{1-e^{-ny}}$$

*for positive real values of $y$. Let $H(y) \rightarrow A$ as $y \rightarrow 0+$. Let the sum of the $a_n$ be a slowly decreasing function in the sense of Hardy* [2] §6.2 pp. 124–125, *that is if $x < y$ are real numbers, so that as $x \rightarrow \infty$ and $y \rightarrow \infty$ in such a manner that $y/x \rightarrow 1$, then*

$$\liminf_{x\to\infty} \sum_{x<n\leqq y} a_n \geqq 0 .$$

*Then*

$$\sum_{n\leqq x} a_n \longrightarrow A , \quad (x \longrightarrow \infty) .$$

REMARK. If the $a_n$ are allowed to be complex then provided that we replace the condition of slowly decreasing by a condition of slow oscillation viz:

$$\lim_{x\to\infty} \sum_{x<n\leqq y} a_n = 0 ,$$

the same conclusion may be drawn. A proof of this lemma can be found in Hardy [2], Theorem 261, pp. 373-374.

In our present circumstances we set

$$a_n = \begin{cases} f(p)p^{-1} & \text{if} \quad n = p , \\ 0 & \text{otherwise} \end{cases}$$

and have established that

$$H(y) = yG(e^{-y}) \longrightarrow A , \quad (y \longrightarrow 0 +) .$$

Moreover,

$$\Big(\sum_{x<n\leqq 2x} |a_n|\Big)^2 \leqq \sum_{x<n\leqq 2x} f^2(p)p^{-1} \sum_{x<n\leqq 2x} p^{-1} ,$$

so that since the series $\sum f^2(p)p^{-1}$ converges and

$$\sum_{x<n\leqq 2x} \frac{1}{p} = \log\left(\frac{\log 2x}{\log x}\right) + O((\log x)^{-1}) \leqq c_4 < \infty ,$$

we see that the condition of slow decreasing required for an application of Lemma 2 is satisfied.

We deduce that

$$\lim_{x\to\infty} \sum_{p\leqq x} \frac{f(p)}{p} = A = \lim_{x\to\infty} x^{-1} \sum_{n\leqq x} f(n) .$$

Moreover, by (ii) a limiting distribution exists for $f(n)$, which has the finite mean of value $A$.

This completes the proof of the theorem.

REMARK. The use of the Tauberian theorem in Lemma 2 is very convenient for the study of additive functions. If $f(p)$ assumes complex values the side condition $f(p) = O(\log p)$ will suffice in order for Lemma 2 to be applicable. This is a condition which is satisfied in nearly every case of number theoretical interest.

## REFERENCES

1. P. Erdös, *On the distribution of additive functions*, Ann. Math., **47** (1946), 1-20.
2. G. H. Hardy, *Divergent Series*, Oxford, 1949.
3. J. P. Kubilius, *Probabilistic Methods in the Theory of Numbers*, Amer. Math. Soc. Transl., Vol. 11, 1964.
4. C. A. Ryavec, *A characterization of finitely distributed additive functions*. J. Number Theory, **2** (4) (1970), 393-403.

Received May 22, 1972.

UNIVERSITY OF COLORADO

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# MULTIPLICATIVE AND EXTREME POSITIVE OPERATORS

M. SOLVEIG ESPELIE

**Let $A$ and $B$ denote complex Banach *-algebras and $L(A, B)$ the space of continuous linear operators from $A$ into $B$. Let $P \subset L(A, B)$ be the convex set of positive linear operators of norm $\leqq 1$. If $A$ has an identity, and if $B$ is semi-simple and symmetric, the multiplicative operators of $P$ are shown to be extreme points of $P$. If, on the other hand, it is assumed that, $\|T\| = \|Te\|$ for $T \in P$, then any extreme point $T$ of $P$ satisfies $Te\, Tab = Ta\, Tb$ for all $a, b \in A$. With $A$ as above and $B$ a $B^*$-algebra, the extreme points of $P$ are multiplicative. Thus we characterize the extreme points of $P \subset L(A, C(X))$ as the multiplicative operators. The results are extended to include the case when $A$ has an approximate identity.**

NOTATION. Let $A$ denote a commutative complex Banach algebra with isometric involution *. We call such an algebra a Banach *-algebra. Let $A'$ denote the topological dual space of $A$ and

$$\hat{P}_A = \{f \in A' : f(xx^*) \geqq 0 \text{ for all } x \in A\},$$

the cone of positive functionals on $A$. Define the usual ordering, $\geqq$, on $\hat{P}_A$. Further, let

$$\begin{aligned} P_A &= \{f \in \hat{P}_A : \|f\| \leqq 1\} \\ M'_A &= \{f \in P_A : f(xy) = f(x)f(y) \text{ for all } x, y \in A\} \text{ and} \\ M_A &= \{f \in M'_A : f \neq 0\}. \end{aligned}$$

The sets $M'_A$ and $P_A$ are compact in the weak* topology and if $A$ contains an identity (of norm one), then $M_A$ is compact in this topology. The set $M_A$ is always a weak* closed subset of the maximal ideal space $\Delta_A$ of $A$. The set $M_A$ is the symmetric portion of $\Delta_A$, and if $M_A = \Delta_A$ we call $A$ a symmetric algebra. Equivalently, $A$ is symmetric if and only if $f(x^*) = \overline{f(x)}$ for all $f \in \Delta_A$ and $x \in A$. It is known that the set of extreme points of the convex set $P_A$ is exactly $M'_A$ [4].

We now replace the scalars of the above paragraph by a Banach *-algebra $B$. That is, we consider subsets of $L(A, B)$, the continuous linear operators from $A$ to $B$. An element $T \in L(A, B)$ is called *positive* if for every $a \in A$ there corresponds a finite set $\{b_i : i = 1, \cdots, n\}$ of $B$ such that $T(aa^*) = \sum_{i=1}^{n} b_i b_i^*$. Following the notation of the above paragraph, let $\hat{P}$ be the cone generated by the positive operators and define the usual ordering on $\hat{P}$. Finally, let

$$P = \{T \in \hat{P}\colon \|T\| \leqq 1\}$$

and

$$M = \{T \in P\colon Tab = Ta\, Tb \text{ for all } a, b \in A\} .$$

The question examined in this paper is: When is $M$ exactly the set of extreme points of $P$? We denote this latter set by ext $P$. With $A$ containing an identity and $B$ semi-simple and symmetric we obtain $M \subset \text{ext } P$. An example exists to show that the semi-simplicity of $B$ is necessary. The symmetry of $B$ appears necessary but we fail to give an example. To obtain the inclusion in the other direction we must require that $B$ be a $B^*$-algebra; that is, $\|b\|^2 = \|bb^*\|$ for all $b \in B$. Equivalently, $B$ is linearly isometric and $*$-isomorphic to $C(\varDelta_B)$ under the map $b \rightarrow \hat{b}$ where $\hat{b}(h) = h(b)$ for all $h \in \varDelta_B$. By placing a norm condition on $\hat{P}$ we are able to prove the weaker condition that $Te\, Tab = Ta\, Tb$ for $T \in \text{ext } P$ and $a, b \in A$. We say that $\hat{P} \subset L(A, B)$ satisfies Condition I if $\|T\| = \|Te\|$ for all $T \in \hat{P}$.

THE NORM CONDITION. The restriction that $\|T\| = \|Te\|$ for $T \in \hat{P}$ is not unusual. This is a well known property of positive functionals — that is, if $B$ is the set of scalars the condition is satisfied. Further, if $B$ is a $B^*$-algebra then we have that $\|T\| = \|Te\|$ when $T \in \hat{P}$. This follows from the fact that in this case the unit ball $S$ of $B'$ is the weak* closed absolutely convex hull $\overline{\Gamma(\varDelta_B)}$ of $\varDelta_B$. Hence for $T \in \hat{P}$,

$$\begin{aligned} \|T\| &= \sup_{\|b\|=1} \|Tb\| = \sup_{\|b\|=1} \sup_{\varphi \in \text{ext} S} |\varphi(Tb)| \\ &= \sup_{\|b\|=1} \sup_{\substack{h \in \varDelta_B \\ |\alpha|=1}} |\alpha h(Tb)| = \sup_{\|b\|=1} \sup_{h \in \varDelta_B} |h(Tb)| \\ &\leqq \sup_{h \in \varDelta_B} \|hT\| = \sup_{h \in \varDelta_B} hTe \leqq \|Te\| . \end{aligned}$$

Since the reverse inequality is evident it follows that $\|T\| = \|Te\|$. It is conceivable that if Condition I is valid for $\hat{P} \subset L(A, B)$, for every $A$, then $B$ is a $B^*$-algebra.

To our knowledge the first work in characterizing the extreme points of such sets as $P$ was done by A. and C. Ionescu Tulcea who considered algebras of real valued continuous functions $C(X)$ and $C(Y)$ on compact Hausdorff spaces $X$, $Y$. They showed that the extreme points of

$$P^\# = \{T \in L(C(X), C(Y))\colon T \geqq 0 \text{ and } T(1) = 1\}$$

are exactly the multiplicative elements of $P^\#$ [8, 10]. Various investigators have obtained results related to and extensions of the Ionescu Tulcea result, the work usually being done for algebras of functions (see [2]). In our work the elements of $P$ do not necessarily satisfy $Te = e'$.

In fact when $B$ is a $B^*$-algebra and $P^\# = \{T \in \hat{P}: Te = e'\}$, since $\|T\| = \|Te\|$ for $T \in \hat{P}$, it follows that $P^\# \subset P$ and moreover, $\operatorname{ext} P^\# = P^\# \cap \operatorname{ext} P$. Indeed, we obviously have $(P^\# \cap \operatorname{ext} P) \subset \operatorname{ext} P^\#$. Further, if $T \in \operatorname{ext} P^\#$ and $T \pm S \in P$, then $\|(T \pm S)(e)\| = \|e' \pm Se\| \leqq 1$ and $Se = 0$ since $e'$ is an extreme point of the unit ball of $B$. But then $(T \pm S)(e) = e'$ and $T \pm S \in P^\#$ so that $S = 0$ and $T \in \operatorname{ext} P$. We note that throughout this paper we use the well known characterization of extreme points: If $K$ is a convex set in a linear space $X$ then $x \in \operatorname{ext} K$ if and only if $x \pm y \in K$ for some $y \in X$ implies that $y = 0$.

Recently Watanabe in [11] has dropped the requirement that the algebras be commutative and has placed pseudo-norms on the algebras. His results, applied to the commutative case, show that $M \subset \operatorname{ext} P$ when $A$ is $B^*$, and the algebra $B$ is semi-simple and symmetric, and both algebras have an identity. Thus, the commutative results in [11] are consequences of our paper since all hypotheses of our theorems are satisfied when $A$ or $B$ are $B^*$-algebras with identity.

THEOREM 1. *Suppose $A$ and $B$ are commutative Banach *-algebras, $A$ has an identity, and $B$ is semi-simple and symmetric. If $T \in P$, then $T \in M$ implies $T \in \operatorname{ext} P$.*

*Proof.* Suppose that $T$ is multiplicative and that there exists an element $S \in L(A, B)$ such that $T \pm S \in P$. Let $h \in M_B$, then $hT$ is an extreme point of $P_A$ [4]. Further, $hT \pm hS \geqq 0$ and $\|hT \pm hS\| \leqq 1$. Thus $hSa = 0$ for all $a \in A$. But $h$ was arbitrary and $B$ is semi-simple so that $Sa = 0$ for all $a \in A$, and hence $T \in \operatorname{ext} P$.

The inclusion in the other direction is more difficult to obtain but is valid if we place additional restrictions on the algebras and on the cone of positive operators. We now prove two lemmas which are needed to obtain the implication: If $T$ is an extreme point of $P$ then $T$ is multiplicative.

LEMMA 1. *Let $A$ be a Banach *-algebra and $b$ an element of $A$ such that $b = b^*$ and $\|b\| < 1$. Then for each $a \in A$, the element $aa^* - aba^*$ is of the form $xx^*$ for some $x \in A$.*

*Proof.* This is a known result if the algebra $A$ has an identity, for then $e - b = yy^*$ where $y = (e - b)^{1/2}$ and $y = y^*$ (see [6, p. 245]). Hence imbed $A$ in the algebra $A_1$ with identity adjoined and write $aa^* - aba^*$ as $a(e - b)a^*$.

In $A_1$ the element $y = (e - b)^{1/2} = \sum_{n=1}^{\infty} \binom{1/2}{n}(-b)^n$ exists and in fact, $y = y^*$. But then, since $A$ is a maximal ideal in $A_1$ and $a \in A$, it follows that $x = ay \in A$ and $xx^* = (ay)(ay)^* = ay^2a^* = aa^* - aba^* \in A$.

LEMMA 2. *Let $A$ and $B$ be commutative Banach $*$-algebras and suppose $T \in P$. If $b_1, b_2 \in A$, with $b_i = c_i c_i^*$ for some $c_i \in A$, $i = 1, 2$, define the linear operator $S$ by*

$$S(a) = T(b_1)T(ab_2) - T(b_2)T(ab_1) \text{ for } a \in A .$$

*Then $T \pm S \geqq 0$.*

*Proof.* For any $a \in A$,

$$\begin{aligned}(T + S)(aa^*) &= T(aa^*) + T(b_1)T(b_2aa^*) - T(b_2)T(b_1aa^*) \\ &\geqq T(aa^*) - T(b_2)T(b_1aa^*) \\ &= T(p_1) + T(b_1aa^*) - T(b_2)T(b_1aa^*) \\ &\geqq T(b_1aa^*) - T(b_2)T(b_1aa^*) \geqq 0 .\end{aligned}$$

Since $||b_1|| < 1$ and $b_1 = c_1c_1^*$ from Lemma 1 it follows that the element $p_1 = aa^* - b_1aa^*$ is positive. Similarly, since $||T(b_2)|| < 1$ and $T(b_2)$ is self-adjoint, and since $T(b_1aa^*)$ is positive, with repeated applications of Lemma 1 it follows that $T(b_1aa^*) - T(b_2)T(b_1aa^*)$ is of the form $\sum_{i=1}^{n} c_ic_i^*$ in $B$. In a similar way it can be shown that $(T - S)(aa^*) \geqq 0$ for $a \in A$.

THEOREM 2. *Suppose that $A$ and $B$ are commutative Banach $*$-algebras, that $A$ has an identity and that Condition* I *holds for elements of $\hat{P}$. Then $Te\,Tab = Ta\,Tb$ for all $a, b \in A$ whenever $T$ is an extreme point of $P$.*

*Proof.* Suppose that $T \in \operatorname{ext} P$ and $b \in A$, with $b = cc^*$ for some $c \in A$ and $||b|| < 1$. We let $S(a) = 1/2(T(b)T(a) - T(e)T(ab))$ be the operator defined in Lemma 2 (taking $b_1 = b$ and $b_2 = 1/2e$), so that $T \pm S \geqq 0$.

Since $||T \pm S|| = ||(T \pm S)(e)|| = ||Te|| \leqq 1$ and $T \in \operatorname{ext} P$ it follows that $Te\,Tab = Ta\,Tb$ for all $a \in A$ provided that $b = cc^*$ and $||b|| < 1$. But every element $b \in A$ is a linear combination of at most four elements of the form $cc^*$ with $||cc^*|| < 1$ and hence $Te\,Tab = Ta\,Tb$ for all $a, b \in A$.

COROLLARY 2.1. *With the hypotheses of Theorem 2 it follows that $hT$ lies on an extreme ray of $\hat{P}_A$ for every $h \in M_B$ and $T \in \operatorname{ext} P$.*

*Proof.* Since $Te\,Tab = Ta\,Tb$ for all $a, b \in A$ and $||hT|| = hTe$ for all $h \in M_B$ it follows that if $||hT|| \neq 0$ then $(hT/||hT||)(ab) =$

$(hT/\|hT\|)(a)(hT/\|hT\|)(b)$ for $a, b \in A$ and $(hT/\|hT\|) \in M_A$. But then, either $\|hT\| = 0$ or $hT/\|hT\| \in \operatorname{ext} P_A$; thus $hT$ lies on an extreme ray of $\hat{P}_A$ for each $h \in M_B$.

COROLLARY 2.2. *Suppose that the hypotheses of Theorem* 2 *hold and in addition, that* $B$ *is semi-simple and symmetric. Then* $T$ *is multiplicative if* $\|hT\|$ *is* 0 *or* 1 *for each* $h \in M_B$ *and* $T \in \operatorname{ext} P$.

*Proof.* Since $hT$ lies on an extreme ray of $\hat{P}_A$ and $\|hT\|$ is 0 or 1 then $hT \in M_A$ for each $h \in M_B$. But then, $h(Tab) = h(Ta\,Tb)$ for all $a, b \in A$ and $h \in M_B$, and since $B$ is semi-simple and symmetric, it follows that $T$ is multiplicative.

THEOREM 3. *Let* $A$ *be a commutative Banach *-algebra with identity and let* $B$ *be a* $B^*$*-algebra. Then every extreme point* $T$ *of* $P$ *is multiplicative.*

*Proof.* If $T \in \operatorname{ext} P$ it follows from Theorem 2 that $Te\,Tab = Ta\,Tb$ for all $a, b \in A$. Thus it suffices to prove that $Te\,Ta = Ta$ for all $a \in A$.

Let $S(a) = 1/2(Te\,Ta - Ta)$ for $a \in A$. The method of proof of Lemma 2 shows that $T \pm S \geqq 0$ and so we need only prove that $\|T \pm S\| \leqq 1$. Since $B$ is a $B^*$-algebra and $T \pm S \geqq 0$ it follows from our earlier remarks that $\|T \pm S\| = \|(T \pm S)(e)\|$. Now, $\|T + S\| = \|(T + S)(e)\| = \|1/2Te + 1/2(Te)^2\| \leqq 1/2\|Te\| + 1/2\|Te\|^2 \leqq 1$. Furthermore, since $0 \leqq Te - Se = (3/2)Te - 1/2(Te)^2$ and $\|Te\| \leqq 1$ we see (letting $f = \widehat{Te} \in C(\Delta_B)$) that $f \geqq 0$ and $\|f\| \leqq 1$. But then $0 \leqq f \leqq 1$ so that $(1 - f)(2 - f) \geqq 0$ and therefore $0 \leqq (3/2)f - 1/2f^2 \leqq 1$. Hence $\|Te - Se\| = \|(3/2)f - 1/2f^2\| \leqq 1$ and the proof is complete.

THEOREM 4. *Assume that* $A$, $B$ *are commutative Banach *-algebras, that* $A$ *has an identity and that* $B$ *is semi-simple and symmetric. Let* $T \in L(A, B)$.

(a) *If* $T \in \operatorname{ext} P$ *and* $hT$ *lies on an extreme ray of* $\hat{P}_A$ *for all* $h \in M_B$, *then* $Te\,Tab = Ta\,Tb$ *for* $a, b \in A$.

(b) *If* $T \in \operatorname{ext} P$ *and* $hT = 0$ *or* $hT \in \operatorname{ext} P_A$ *for all* $h \in M_B$, then $Tab = Ta\,Tb$ *for all* $a, b \in A$.

The proofs of (a) and (b) are immediate for if $hT$ lies on an extreme ray of $\hat{P}_A$, then either $\|hT\| = 0$ or $hT/\|hT\| \in M_A$ for each $h \in M_B$. But then $h(Te\,Tab) = h(Ta\,Tb)$ for all $h \in M_B$, and since $B$ is semi-simple and symmetric, it follows that $Te\,Tab = Ta\,Tb$ *for* $a, b \in A$. Finally, if $hTe$ *is* 0 or 1 for all $h \in M_B$, then $T$ is multiplicative.

We state the following result without proof. Using the above methods and Theorem 3 the result follows.

THEOREM 5. *Let $A$ be a commutative Banach *-algebra with identity and $B$ a $B^*$-algebra. If $T \in P$ and $Te\, Tab = Ta\, Tb$ for $a, b \in A$, then $T \in \operatorname{ext} P$ if and only if $Te$ is an extreme point of $K = \{b \in B\colon \|b\| \leqq 1$ and $b = \sum_{i=1}^{n} c_i c_i^*$, $n$ finite$\}$.*

It should be noted that if $Te \in \operatorname{ext} K$ and $Te\, Tab = Ta\, Tb$ for $a$, $b \in A$ then $T \in \operatorname{ext} P$ when $B$ is semi-simple and symmetric. To obtain the converse we employ the hypothesis that $B$ is a $B^*$-algebra so that $K$ can be identified with the set $\{f\colon f \geqq 0$ and $\|f\| \leqq 1\}$ in $C(\Delta_B)$. It is well known that the set of extreme points of this set is $\{f\colon f(x) = 0$ or $1$ for all $x \in \Delta_B\}$.

EXAMPLES. We now display some examples indicating the need for the hypotheses placed on the algebras $A$ and $B$ in the above work. Most of the algebras used in our examples can be found in [10].

Consider the involution algebra $\mathscr{A}$ of functions analytic on the open unit disc and continuous on the closed disc with the usual supremum norm and pointwise multiplication. An involution is defined on $\mathscr{A}$ by $f^*(z) = \overline{f(\bar{z})}$. The algebra $\mathscr{A}$ is semi-simple and not symmetric.

We construct an element of $\operatorname{ext} P \subset L(\mathscr{A}, \mathscr{A})$ which is not in $M$. Denote by $h$ the element of $P$ defined by $h(f) = f(1)$ for $f \in \mathscr{A}$ and denote by $Z^n$ the element of $\mathscr{A}$ defined by $Z^n(w) = w^n$ for $n = 0, 1, 2, \cdots$. Then the operator given by $Tf = Z^2 h(f)$ for $f \in \mathscr{A}$ is an element of $\operatorname{ext} P$ and not in $M$.

It follows that $T \in P$ since $T(ff^*) = f(1)f^*(1)Z^2 = (f(1)Z)(f(1)Z)^*$ and $\|T\| = \sup_{\|f\|=1} \|Z^2 h(f)\| = 1$. Suppose that $T \pm S \in P$ for some $S \in L(\mathscr{A}, \mathscr{A})$ so that $\|T(Z^n) \pm S(Z^n)\| = \|Z^n \pm S(Z^n)\| \leqq 1$. Since $Z^n$, for each $n$, is an extreme point of the unit ball of $\mathscr{A}$ it follows that $S(Z^n) = 0$ and that $S$ is zero on the polynomials, a dense subalgebra of $\mathscr{A}$. Consequently, $S = 0$ and $T \in \operatorname{ext} P$. $T$ is not multiplicative but it should be noted that $T$ satisfies $Te\, Tfg = Tf\, Tg$ for all $f$, $g \in \mathscr{A}$.

We now consider a space $L(A, B)$ such that the algebra $B$ is not semi-simple. Let $\mathscr{A}$ be the Banach space of the above paragraph. However, we now place a multiplication on $\mathscr{A}$ defined in terms of a convolution, $f * g(w) = w\int_0^1 f[(1-t)w]g(tw)dt$ where $|w| \leqq 1$. We denote this algebra by $\mathscr{A}_0$; it is a Banach *-algebra with the involution $f^*(z) = \overline{f(\bar{z})}$ and the supremum norm. With this definition of multiplication we have $\lim_{n\to\infty} \|f^n\|^{1/n} = 0$ for all $f \in \mathscr{A}_0$ since $\|f^n\| = \|f\|^n/(n-1)!$. Thus $\mathscr{A}_0$ is a radical algebra and we consider the algebra $\mathscr{A}_0 + e$, the algebra with the identity adjoined. This algebra has one maximal ideal so that it is symmetric. The element $Z(w) = w$ is positive since it is $I^2$ where $I(w) = 1$ for all $w$ in the disc.

We consider $L(\mathscr{A}, \mathscr{A}_0 + e)$ and define $T \in L(\mathscr{A}, \mathscr{A}_0 + e)$ by $Tf = f(1)Z$, $Z(w) = w$. Using methods similar to these above we can show that $T \in \operatorname{ext} P$ and $T \notin M$. All the hypotheses of Theorem 4.b are satisfied except for the semi-simplicity of the range space. Again we have $Te\, Tfg = Tf\, Tg$ for all $f, g \in \mathscr{A}$.

Finally we display an operator between two algebras which is multiplicative and not an extreme point of $P$ when all hypotheses of Theorem 1 are satisfied except for the semi-simplicity of the range.

Let $\Omega$ be the algebra of all power series $a(z) = \sum_{n=0}^{\infty} a_n z^n$ such that $\sum_{n=0}^{\infty} |a_n|/n! < \infty$; the norm is $\|a(z)\| = \sum_{n=0}^{\infty} (|a_n|/n!)$. Multiplication is defined in the usual way. $\Omega$ is a Banach *-algebra with involution defined by $(a(z))^* = \sum_{n=0}^{\infty} \bar{a}_n z^n$. The identity of $\Omega$ is the series with $a_0 = 1$ and $a_n = 0$ for $n \geqq 1$.

We consider the maximal ideal generated by the series with $a_1 = 1$ and $a_n = 0$ for $n \neq 1$. All elements of this ideal are essentially nilpotent and this is the only maximal ideal in $\Omega$. Thus $\Omega$ is symmetric and not semi-simple. Let $T \in L(\Omega, \Omega)$ be defined by $T(\sum_{n=0}^{\infty} a_n z^n) = \sum_{n=0}^{\infty} a_n z^{2n}$. It can be shown that $T \in P \subset L(\Omega, \Omega)$ and that $T \in M$. Define $S$ by $S(a(z)) = \sum_{n=1}^{\infty} a_{2n-1} z^{4n-2}$. Now $T \pm S \in P$ and $S \neq 0$ so that $T \notin \operatorname{ext} P$.

A GENERALIZATION. If we replace the condition that $A$ has an identity with the condition that $A$ has an approximate identity we obtain analogues of the above statements. The net $\{e_\alpha\}$ is an *approximate identity* in $A$ if $\|e_\alpha\| \leqq 1$ and $e_\alpha > 0$ for all $\alpha$ and $\|e_\alpha x - x\| \xrightarrow[\alpha]{} 0$ for all $x \in A$. (We assume that $e_\alpha > 0$ for all $\alpha$ since $\|xe_\alpha e_\alpha^* - x\| \leqq \|(xe_\alpha - x)e_\alpha^*\| + \|xe_\alpha^* - x\| \leqq \|xe_\alpha - x\| + \|x^* e_\alpha - x^*\|$ for all $x \in A$ so that $\{e_\alpha e_\alpha^*\}$ is an approximate identity whenever $\{e_\alpha\}$ is.) We make use of the fact that for a commutative Banach *-algebra $A$ with approximate identity, if $f \in \hat{P}_A$, then $\|f\| = \lim_\alpha f(e_\alpha)$ and $M_A' = \operatorname{ext} P_A$ [5]. With this result Theorem 1 remains true as stated in this new setting.

To obtain further generalizations we place Condition I′ on $\hat{P}$: $\|T\| = \lim_\alpha \|Te_\alpha\|$ for all $T \in \hat{P}$. We now show that if $T$ is positive in $L(A, B)$ where $A$ is a commutative Banach *-algebra with approximate identity and $B$ is a $B^*$-algebra, then Condition I′ holds. For each $\alpha$ and each $h \in M_B$ we have $hTe_\alpha \leqq \|Te_\alpha\|$ so that $\|hT\| = \lim_\alpha hTe_\alpha = \liminf_\alpha hTe_\alpha \leqq \liminf_\alpha \|Te_\alpha\|$. Since $\|T\| \leqq \sup_{\|h\|=1} \|hT\|$, it follows that $\|T\| \leqq \liminf_\alpha \|Te_\alpha\|$. Moreover, $\|Te_\alpha\| \leqq \|T\|$ for each $\alpha$ yielding $\limsup_\alpha \|Te_\alpha\| \leqq \|T\|$ and hence $\|T\| = \lim_\alpha \|Te_\alpha\|$.

THEOREM 6. *Suppose that $A$, $B$ are commutative Banach *-algebras,* *that $A$ has an approximate identity, and that $\hat{P}$ satisfies Condition I′. Then, for every extreme point $T$ of $P$ it follows that*

$$T(a)T(b) = \lim_\alpha T(e_\alpha)T(ab) \text{ for all } a, b \in A .$$

*Proof.* Let $S(a) = 1/2(T(b)T(e_\alpha a) - T(e_\alpha)T(ba))$ be the operator defined in Lemma 2 with $b_1 = b$ and $b_2 = e_\alpha/2$, so that $T \pm S \geqq 0$.

Since $\hat{P}$ satisfies Condition I′, it follows that $\|T \pm S\| = \lim_\beta \|T(e_\beta) \pm 1/2[T(\mathrm{b})T(e_\alpha e_\beta) - T(e_\alpha)T(be_\beta)]\| = \lim_\beta \|T(e_\beta)\| \leqq 1$ and hence, $T \pm S \in P$. Since $T \in \operatorname{ext} P$ we have $T(b)T(e_\alpha a) = T(e_\alpha)T(ba)$ for $a \in A$ and each $e_\alpha$, so that $\lim_\alpha T(e_\alpha)T(ab) = T(a)T(b)$ for all $a \in A$ and $b \in A$ with $\|b\| < 1$ and $b = cc^*$ for some $c \in A$.

Now, every product, and hence every element of the form $be_\beta$ can be written as the linear combination of four positive elements; that is, $be_\beta = 1/4 \sum_{k=1}^{4} \bar{i}^k(e_\beta^* + i^k b)(e_\beta^* + i^k b)^*$ for all $b \in A$ and all $\beta$. It follows, from the linearity of $T$, that $\lim_\alpha T(e_\alpha)T(abe_\beta) = T(a)T(be_\beta)$ for all $a, b \in A$ and all $\beta$; and, from the continuity of $T$, that $\lim_\alpha T(e_\alpha)T(ab) = T(a)T(b)$ for $a, b \in A$.

COROLLARY 6.1. *With the hypotheses of Theorem* 6 *for $T \in \hat{P}$ and $h \in M_B$, either $hT = 0$ or $hT/\|hT\| \in M_A$.*

*Proof.* From Theorem 6 for $T \in \hat{P}$ we conclude that $\lim_\alpha h(Te_\alpha Tab) = h(Ta\, Tb)$ for $a, b \in A$ and $h \in M_B$. Consequently,

$$\|hT\| hT(ab) = hT(a)hT(b)$$

and either $hT = 0$ or $hT/\|hT\| \in M_A$ for $h \in M_B$. Or equivalently, $hT$ lies on an extreme ray of $\hat{P}_A$ for every $h \in M_B$ and $T \in \hat{P}$.

DEFINITION. As noted earlier, if $A$ is a commutative Banach *-algebra then $A_1$ denotes the algebra with identity adjoined. If $f \in \hat{P}_A$ then $f_1$, defined by $f_1(a + \lambda) = f(a) + \lambda \|f\|$, is the extension of $f$ to $A_1$ and $\hat{P}_{A_1}$ is the cone of positive elements of $A_1'$. Similarly, $P_{A_1}$, denotes the set of positive elements of norm $\leqq 1$. Finally, if $T \in L(A, B)$ and $T \geqq 0$ we define $T_1$ to be that element of $L(A_1, B_1)$ defined by $T_1(a + \lambda) = T(a) + \lambda \|T\|$. (We let $B_1 = B$ if $B$ has an identity.) Furthermore, $\hat{P}_1$ denotes the cone of positive operators and $P_1$ those positive operators of norm $\leqq 1$.

We now prove the statement equivalent to Theorem 3 when $A$ has an approximate identity.

THEOREM 7. *Let $A$ be a commutative Banach *-algebra with approximate identity and $B$ a $B^*$-algebra. Then, every extreme point $T$ of $P$ is multiplicative.*

*Proof.* From the decomposition for products used in Theorem 6 it follows easily that if $T$ is a positive operator, then $T(ab^*) = (T(a^*b))^*$ for $a, b \in A$. Since $A$ has an approximate identity and $T$ is continuous, it follows that $T(a^*) = (Ta)^*$ for $a \in A$. Moreover, since the range of $T$ is contained in $C_0(\Delta_B)$ (or $C(\Delta_B)$ if $B$ contains an identity), letting $h_x$ (or $x$) be that element of $\Delta_B$ defined by $h_x(b) = b(x)$ for $b \in C_0(\Delta_B)$ it follows that $(Ta^*)(x) = \overline{(Ta)(x)}$ for $a \in A$ and $x \in \Delta_B$. Further, for $T$ positive from a Cauchy-Schwarz inequality [9, p. 213] we conclude that $|T(ab^*)|^2(x) \leqq [T(aa^*)T(bb^*)](x)$ for $a, b \in A$ and $x \in \Delta_B$. Letting $b = e_\alpha$ and taking the limit we obtain $|Ta|^2(x) = ||T||\,[T(aa^*)](x)$ for $x \in \Delta_B$ and $a \in A$.

We now show that for any $T \in \hat{P}$ the element $T_1$ defined by $T_1(a + \lambda) = Ta + \lambda\,||T||$ is, in fact, positive. Thus, if $T \in \hat{P}$ and $||T|| \neq 0$, then

$$\begin{aligned}[T_1(a+\lambda)(a^*+\bar{\lambda})](x) &= [Taa^* + \lambda Ta^* + \bar{\lambda}Ta + |\lambda|^2||T||](x_1)\\ &= [Taa^* - 2Re\,(-\bar{\lambda})Ta + |\lambda|^2||T||](x_1)\\ &\geqq [1/||T||\,|Ta|^2 - 2|\lambda|\,|Ta| + |\lambda|^2||T||](x_1)\\ &= 1/||T||\,[|Ta| - ||T||\,|\lambda|]^2(x_1) \geqq 0\end{aligned}$$

for all $a + \lambda \in A_1$ and $x_1 \in \Delta_{B_1}$, where $x_1$ is the extension of the element $x$ in $\Delta_B$ when $B$ does not have an identity. Finally, when $B$ does not have an identity, let $x'$ be that element of $\Delta_{B_1}$ which has $A$ as its corresponding maximal ideal (we note that in this case $\Delta_{B_1}$ is the one point compactification of $\Delta_B$); then $[T_1(a+\lambda)(a^*+\bar{\lambda})](x') = |\lambda|^2(x') \geqq 0$ for $a + \lambda \in A_1$. Thus $T_1 \geqq 0$. Since the range of $T_1$ is contained in a $B^*$-algebra, $\hat{P}_1$ satisfies Condition I and hence $||T_1|| = ||T_1 e||$.

We now show that $T_1$ is an extreme point of $P_1$. Suppose that there exists $S \in L(A_1, B_1)$ such that $T_1 \pm S \in P_1$. Then, $||T_1 e \pm Se|| = ||e' \pm Se|| \leqq 1$ and hence $Se = 0$ since $e'$ is an extreme point of the unit ball of $B_1$. Let $\bar{S}$ denote the restriction of $S$ to $A$, then, since $T \pm \bar{S} \geqq 0$ and $||T \pm \bar{S}|| \leqq 1$, it follows that $\bar{S} = 0$ since $T \in \operatorname{ext} P$. Therefore, $S = 0$ and $T_1$ is an extreme point of $P_1$ so that from Theorem 3 it follows that $T_1$ is multiplicative. Hence, $T$ is multiplicative and the proof is complete.

*Acknowledgements.* I wish to express appreciation to George Maltese for his many helpful conversations and to R. R. Phelps for bringing many improvements in the paper to my attention.

## References

1. R. M. Blumenthal, Joram Lindenstrauss, and R. R. Phelps, *Extreme Operators into C(K)*, Pacific J. Math., **15** (1965), 747-756.

2. F. F. Bonsall, J. Lindenstrauss, and R. R. Phelps, *Extreme Positive Operators on Algebras of Functions*, Math. Scand., **18** (1966), 161-182.

3. N. Bourbaki, *Espaces Vectoriels Topologiques*, Chaps. I–II, 2nd Ed., Hermann, Paris, 1966.
4. R. S. Bucy and G. Maltese, *A representation theorem for positive functionals on involution algebras*, Math. Annalen, **162** (1966), 364–367.
5. M. S. Espelie, *Extreme Functionals*, (submitted).
6. I. Gelfand, D. Raikov, and G. Shilov, *Commutative Normed Rings*, Chelsea, New York, 1964.
7. L. H. Loomis, *An Introduction to Abstract Harmonic Analysis*, D. Van Nostrand Co., Princeton, 1953.
8. R. R. Phelps, *Extreme positive operators and homomorphisms*, Trans. Amer. Math. Soc., **108** (1963), 265–274.
9. C. E. Rickart, *General Theory of Banach Algebras*, D. Van Nostrand Co., Princeton, 1960.
10. A. and C. Ionescu Tulcea, *A note on extremal points*, (unpublished).
11. S. Watanabe, *Positive linear maps of Banach algebras with an involution*, Sci. Rep. Niigata Univ. Ser. A. Nov., **8** (1971), 1–6.

Received August 31, 1971 and in revised form April 18, 1973.

HOWARD UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# DOMAINS OF NEGATIVITY AND APPLICATION TO GENERALIZED CONVEXITY ON A REAL TOPOLOGICAL VECTOR SPACE

JACQUES A. FERLAND

**The purpose of this paper is to derive conditions for the existence of domains of negativity, and then to determine maximal domains of convexity, quasi-convexity, and pseudo-convexity for a quadratic function defined on a real topological vector space.**

1. **Introduction.** Martos, in [14] and [15], and Cottle and the author, in [3], [4], [6], and [7], study quasi-convex and pseudo-convex quadratic functions defined on $E^n$, the $n$-dimensional Euclidean space. Furthermore, in [6] and [7], the author uses the concept of domains of negativity that was introduced, mutatis mutandis, by Koecher in [11]. The purpose of this paper is to derive conditions for the existence of domains of negativity, and then to generalize the results found in [6].

In §2, we briefly review definitions needed in the rest of this paper. We also state relations between the classes of convex, quasi-convex, and pseudo-convex quadratic functions on a convex set. Conditions for the existence of domains of negativity and properties of these are given in §3. In §4, convex quadratic functions are studied. Then, domains of quasi-convexity and pseudo-convexity for quadratic forms are specified in §5, and, in §6, we extend this analysis to quadratic functions.

*Note.* Another approach to this theory have been used by Siegfried Schaible in "Quasi-convex Optimization in General Real Linear Spaces", Zeitschrift für Operations Research, 1972.

2. DEFINITIONS. Let $E^1$ denote the field of real numbers with the natural topology and let $X$ be a vector space over $E^1$. We assume that $X$ admits a *norm*, i.e., there exists a mapping $x \to |x|$ from $X$ into $E^1_+ = \{\alpha \in E^1 | \alpha \geqslant 0\}$ with the following properties:

( i ) $|x| = 0$ if and only if $x = 0$,

(ii) $|\lambda x| = |\lambda|\,|x|$ for all $\lambda \in E^1$ and all $x \in X$,

(iii) $|x + y| \leqslant |x| + |y|$ for all $x$ and $y$ in $X$.

A topology on $X$ is determined by this norm, and $X$, so endowed, is called a *topological vector space over* $E^1$.

Let $X$ and $Y$ be two real vector spaces. The mapping $A: X \to Y$ is a *linear transformation* if and only if for all vectors $x$ and $y$ in

$X$ and for all real numbers $\alpha$ and $\beta$

$$A(\alpha x + \beta y) = \alpha A(x) + \beta A(y).$$

If $Y = E^1$, then $A$ is said to be a *linear form* from $X$ into $E^1$.

The mapping $L: X \times X \rightarrow E^1$ is a *bilinear form* on $X$ if and only if

(i) $L(x, y) = L(y, x)$ for all $x$ and $y$ in $X$,

(ii) $L(x, y)$ is linear and continuous in $y$ for each fixed $x$.

With each bilinear form $L$ is associated a unique *quadratic form* $Q: X \rightarrow E^1$ defined by

$$Q(x) = L(x, x) \text{ for all } x \in X .$$

A *quadratic function* on a real vector space $X$ is a mapping $R$: $X \rightarrow E^1$ defined by

$$R(x) = 1/2Q(x) + P(x) \text{ for all } x \in X ,$$

where $Q$ is a quadratic form and $P$ is a linear form, both defined on $X$.

The *radical* of a bilinear form $L$ is the set

$$X(L) = \{x \in X \mid L(x, y) = 0 \text{ for all } y \in X\} .$$

$L$ is *nondegenerate* on $X$ if $X(L) = 0$. Otherwise, $L$ is degenerate.

If $X_1$ and $X_2$ are subsets of $X$, then the *complement of $X_2$ relative to $X_1$* is the set

$$X_1 \backslash X_2 = \{x \in X_1 \mid x \notin X_2\} .$$

Also, *the sum of $X_1$ and $X_2$* is the set

$$X_1 + X_2 = \{x \in X \mid x = u + v, u \in X_1, \text{ and } v \in X_2\} .$$

If $E_1$ and $E_2$ are subspaces of $X$, then $X = E_1 \oplus E_2$, the *direct sum* of $E_1$ and $E_2$, if and only if for each $x \in X$ there exists a unique pair $u \in E_1$ and $v \in E_2$ such that $x = u + v$.

In [11], Koecher introduces the notion of domains of positivity in a real topological vector space, and mutatis mutandis, we define a *domain of negativity* in $X$ determined by $L$ as a subset $Y$ of $X$ having the following properties:

(i) $Y$ is open and nonempty,

(ii) $L(x, y) < 0$ for all $x$ and $y \in Y$,

(iii) for all $x \notin Y$ there exits a vector $y \in \bar{Y} \backslash X(L)$ such that $L(x, y) \geqslant 0$. (Note that $\bar{Y}$ is the closure of $Y$.)

A subset $S$ of $X$ is said to be *convex* if and only if for all $x, y$ in $S$ and for all $\theta \in [0, 1]$

$$x(\theta) = (1 - \theta)x + \theta y \in S .$$

Furthermore, $S$ is *solid* if and only if it has a nonempty interior, $S^0$.

The quadratic function $R(x) = 1/2Q(x) + P(x)$ is *convex* on a convex set $S$ in $X$ if and only if for all $x$ and $y$ in $S$ and for all $\theta \in [0, 1]$,

$$(1) \qquad R((1-\theta)x + \theta y) \leqslant (1-\theta)R(x) + \theta R(y) .$$

The quadratic function $R(x) = 1/2Q(x) + P(x)$ is *quasi-convex* on a set $S$ in $X$ if and only if for all $x$ and $y$ in $S$

$$(2) \qquad R(y) \leqslant R(x) \text{ implies } L(x, y-x) + P(y-x) \leqslant 0 .$$

The quadratic function $R(x) = 1/2Q(x) + P(x)$ is *pseudo-convex* on a set $S$ in $X$ if and only if for all $x$ and $y$ in $S$

$$(3) \qquad L(x, y-x) + P(y-x) \geqslant 0 \text{ implies } R(y) \geqslant R(x) .$$

Observe that if we take $P(x) = 0$ for all $x \in X$, then (1), (2), and (3) are the conditions for the quadratic form $Q$ to convex, quasi-convex, and pseudo-convex, respectively.

If $S$ is a convex set, then denote by $C(S)$, $QC(S)$, and $PC(S)$ the classes of all quadratic functions $R$ that are convex on $S$, quasi-convex on $S$, and pseudo-convex on $S$, respectively.

Notice that Mangasarian's results in Chapters 6 and 9 of [13] hold for a quadratic function $R(x) = 1/2Q(x) + P(x)$ defined on an arbitrary real topological vector space if we replace the expression $(\nabla R(x), y - x)$ by $L(x, y-x) + P(y-x)$. (Recall that in $E^n$ the gradient of $R$ evaluated at $x$, $\nabla R(x)$, is the column vector of the partial derivatives of $R$ at $x$.) Thus, from [13, Theorem 9.1.4], we have this equivalent definition: a quadratic function $R(x)$ is *quasi-convex* on a *convex* set $S$ in $X$ if and only if for all $x, y \in S$ and for all $\theta \in [0, 1]$

$$(4) \qquad R((1-\theta)x + \theta y) \leqslant \operatorname{Max}\{R(x), R(y)\} .$$

Furthermore the results in [13], [Chapters 6 and 9] imply that if $S$ is a convex set in $X$, then

$$(5) \qquad C(S) \subset PC(S) \subset QC(S) .$$

In [3], Cottle and the author have shown the following.

(6) PROPOSITION. *If the real valued function $h$ is quasi-convex on a nonempty convex set $S$ in $E^n$ and continuous on $\bar{S}$, then $h$ is quasi-convex on $\bar{S}$, the closure of $S$.*

Since this result holds for a quadratic function $R$ defined on an arbitrary real topological vector space, if $S$ is convex, then

$$(7) \qquad QC(S) \subset QC(\bar{S}) .$$

It follows from (5) and (7) that for a convex set $S \subset X$

$$C(S) \subset PC(S) \subset QC(S) \subset QC(\bar{S}) . \tag{8}$$

Observe the similarity with Ponstein's results for $X = E^n$. See [16].

**3. Domains of negativity.** In this section we give necessary and sufficient conditions for a bilinear form to determine a pair of domains of negativity in a real topological vector space. The importance of domains of negativity in the study of quasi-convexity and pseudo-convexity will become apparent in §§5 and 6.

First we introduce the following notation. For each $x \in X$ we denote by $E(x)$ the *subspace generated by* $x$, i.e.,

$$E(x) = \{z \in X \mid z = \alpha x, \alpha \in E^1\} .$$

Given a certain bilinear form $L$ and an arbitrary subspace $E$ of $X$, we denote

$$E_L = \{z \in X \mid L(x, z) = 0 \text{ for all } x \in E\} .$$

Referring to [10, p. 6], the following is true.

(9) PROPOSITION. *If* $x \in X$ *and* $Q(x) \neq 0$, *then* $X = E(x) \oplus E_L(x)$.

Relative to a bilinear form $L$, we say that a nonzero vector $z \in X$ is

$$\begin{aligned} &\textit{positive-valued} \text{ if and only if } Q(z) > 0 , \\ &\textit{negative-valued} \text{ if and only if } Q(z) < 0 , \\ &\textit{zero-valued} \text{ if and only if } Q(x) = 0 . \end{aligned}$$

Suppose that $L$ is a nondegenerate bilinear form, i.e., $X(L) = 0$. Furthermore, suppose there exists a vector $x \in X$ such that $Q(x) = -1$ and $E_L(x)$ is an *inner product space* where $L(u, v)$ is the inner product, i.e.,

$$\begin{aligned} &L(u, v) = L(v, u) \text{ for all } u, v \in E_L(x) \\ &Q(u) \geqslant 0 \text{ for all } u \in E_L(x) \\ &Q(u) = 0 \text{ implies } u = 0. \end{aligned}$$

For details see Schaefer [17, p. 44] or Greub [9, p. 160]. From (9),

$$X = E(x) \oplus E_L(x) .$$

Using the same type of argument as in [9, p. 268], the following can be shown.

(10) PROPOSITION. *If $z$ is a negative-valued vector or if $z$ is a nonzero but zero-valued vector, then* $L(x, z) \neq 0$.

Define the sets

$$Y^+ = \{z \in X \mid Q(z) < 0 \text{ and } L(x, z) < 0\} ,$$
$$Y^- = \{z \in X \mid Q(z) < 0 \text{ and } L(x, z) > 0\} ,$$

Notice that $Y^+$ and $Y^-$ are nonempty since $x \in Y^+$ and $-x \in Y^-$. It is easy to verify that

$$\bar{Y}^+ = \{z \in X \mid Q(z) \leqslant 0 \text{ and } L(x, z) < 0\} \cup \{0\}$$
$$\bar{Y}^- = \{z \in X \mid Q(z) \leqslant 0 \text{ and } L(x, z) > 0\} \cup (0) ,$$

and that $Y^+ \cup \{0\}$, $Y^- \cup \{0\}$, $\bar{Y}^+$, and $\bar{Y}^-$ are solid convex cones. Furthermore, a modified version of arguments [6, (3.22) and (3.32)] shows that $Y^+$ and $Y^-$ are *domains of negativity.*

The definitions of $Y^+$ and $Y^-$ and (10) imply the following result.

(11) THEOREM. *Given the pair of domains of negativity* $Y^+$ and $Y^-$ *in $X$ determined by $L$, then*

(a) $z \in X^- = Y^+ \cup Y^-$ *if and only if* $Q(z) < 0$,

(b) $z \in X^0 = (\bar{Y}^+ \backslash Y^+) \cup (\bar{Y}^- \backslash Y^-)$ *if and only if* $Q(z) = 0$,

(c) $z \in X^+ = X \backslash (\bar{Y}^+ \cup \bar{Y}^-)$ *if and only if* $Q(z) > 0$.

Since $Y^+$ and $Y^-$ are maximal ([11, p. 5]), then it follows from (11) that the pair $Y^+$ and $Y^-$ in $X$ determined by $L$ is unique.

In summary, if the vector $x \in X$ is such that $Q(x) = -1$ and $E_L(x)$ is an inner product space, then there exists a pair of domains of negativity in $X$ determined by $L$. This sufficient condition can be expressed into another form. To see this, we need the following result.

(12) PROPOSITION. *If there exists a vector $x \in X$ such that $Q(x) = -1$ and $E_L(x)$ is an inner product space, then for all $z \in X$ such that $Q(z) <$ 0 the subspace $E_L(z)$ is an inner product space.*

*Proof.* For contradiction, suppose that $Q(z) < 0$ for some $z \in X$ and $E_L(z)$ is not an inner product space. Hence, there exists a nonzero vector $y \in E_L(z)$ such that $Q(y) \leqq 0$. On the other hand, by definition of $x$ there exists a pair $Y^+$ and $Y^-$ of domains of negativity in $X$ determined by $L$.

Suppose $z \in Y^+$. If $Q(y) < 0$, then via (11), either the pair $y$ and $z$ belongs to $Y^+$ or the pair $-y$ and $z$ belongs to $Y^+$. Since $L(y, z) =$

$L(-y, z) = 0$, in either case we have a contradiction to the definition of domains of negativity.

If $Q(y) = 0$, then, via (11), either $y \in \bar{Y}^+ \backslash Y^+$ or $-y \in \bar{Y}^+ \backslash Y^+$. Since $y \neq 0$, either the pair $z$ and $y$ or the pair $z$ and $-y$ contradicts the property that if $u \in Y^+$ and $v \in \bar{Y}^+ \backslash X(L)$, then $L(u, v) < 0$ ([11, Theorem 1 a.]). The proof is complete.

Relying on (12), if the set $\{x \in X \mid Q(x) < 0\}$ is nonempty and for each $x$ in this set the subspace $E_L(x)$ is an inner product space, then there exists a pair of domains of negativity. Other trivial sufficient conditions for the existence of such a pair are $Q(x) < 0$ and $E_L(x)$ empty (i.e., $\dim X = 1$). Now we turn to the necessity of these conditions.

(13) THEOREM. *If there exists a pair $Y^+$ and $Y^-$ of domains of negativity in $X$ determined by $L$, then the set $\{x \in X \mid Q(x) < 0\}$ is nonempty and for all $x \in X$ such that $Q(x) < 0$ the subspace $E_L(x)$ is an inner product space or is empty.*

*Proof.* Since $Y^+$ is nonempty, it follows that $\{x \in X \mid Q(x) < 0\}$ is nonempty. The second condition is shown by a similar argument as in (12), and this completes the proof.

We are left with the problem of studying conditions for the existence of domains of negativity when the bilinear form $L$ is degenerate in $X$, i.e., when $X(L) \neq 0$. Referring to Schaefer [17, p. 20], the vector space $X$ can always be expressed as

$$X = (X/X(L)) \oplus X(L)$$

where $X/X(L)$ is called the *quotient space* of $X$ over $X(L)$. It is well-known that the bilinear form $L$ is nondegenerate on $X/X(L)$.

If there exists a pair $Y_L^+$ and $Y_L^-$ of domains of negativity in $X/X(L)$ determined by $L$, then denote

$$Y^+ = Y_L^+ \oplus X(L)$$
$$Y^- = Y_L^- \oplus X(L) .$$

First, since $Y_L^+$ and $Y_L^-$ are nonempty and open, so are $Y^+$ and $Y^-$. The other conditions for $Y^+$ and $Y^-$ to be domains of negativity in $X$ follow from the fact that if $x, y \in X$, then

$$x = u + t , \qquad u \in X/X(L) \text{ and } t \in X(L) ,$$
$$y = v + z, \qquad v \in X/X(L) \text{ and } z \in X(L) ,$$

and

$$L(x, y) = L(u, v) + L(t, z) = L(u, v) .$$

Hence a pair $Y^+$ and $Y^-$ of domains of negativity in $X$ determined by $L$ exists if and only if such a pair exists when $L$ is restricted to $X/X(L)$.

**4. Domains of convexity for a quadratic function.** In this section, we want to determine the convex sets in $X$ over which a quadratic function is convex. In [2], Cottle has studied this problem for quadratic functions defined on $E^n$, and, as we shall see, these results hold on an arbitrary real topological vector space.

Using definition (1), this result follows immediately.

(14) PROPOSITION. *The quadratic function $R$ is convex on a convex set $S$ in $X$ if and only if the quadratic form $Q$ is convex on $S$.*

The same kind of argument, as when the quadratic form is defined on $E^n$, can be used to show the following result.

(15) PROPOSITION. *The quadratic form $Q$ is convex on a convex set $S$ in $X$ if and only if for all $x$ and $y$ in $S$*

$$Q(x - y) \geqslant 0 .$$

Notice this generalization of Cottle's result [2, (2)].

Recall that a set $K$ in $X$ is said to be a linear manifold if it is of the form

$$K = E + x$$

where $x \in X$ and $E$ is a vector subspace of $X$. ([1]).

With each convex set $S$ in $X$ is associated a *carrying plane* $K(S)$ defined as the linear manifold of least dimension which contains $S$. The same argument as in [2] shows the following property.

(16) PROPOSITION. *If the quadratic form $Q$ is convex on a convex set $S$ in $X$, then $Q$ is convex on $K(S)$.*

It follows that if the quadratic form $Q$ is convex on a solid convex set $S$ in $X$, then $Q$ is convex on $X$.

**5. Domains of quasi-convexity and pseudo-convexity for quadratic forms.** The results found in Chapter 3 of [6] hold even for quadratic forms defined on a real topological vector space. Since only slight modifications of these arguments are needed for the generalization, we will restrict ourselves to the statements of the results.

Suppose that $Y$ is a domain of negativity in $X$ determined by $L$.

(17) THEOREM. *The quadratic form $Q$ is quasi-convex on $\bar{Y}$ and pseudo-convex on $\bar{Y}\backslash X(L)$.*

(18) THEOREM. *If the quadratic form $Q$ is quasi-convex, but not convex, on a solid convex set $S$, then there exists a unique pair of domains of negativity, $Y^+$ and $Y^-$, in $X$ determined by $L$, and $S \subset \bar{Y}^+$ or $S \subset \bar{Y}^-$.*

(19) THEOREM. *If the quadratic form $Q$ is pseudo-convex, but not convex, on a solid convex set $S$, then there exists a unique pair of domains of negativity, $Y^+$ and $Y^-$, in $X$ determined by $L$, and $S \subset \bar{Y}^+\backslash X(L)$ or $S \subset \bar{Y}^-\backslash X(L)$.*

Therefore, if $Y^+$ and $Y^-$ is a pair of domains of negativity in $X$ determined by $L$, then $\bar{Y}^+$ and $\bar{Y}^-$ are *maximal domains of quasi-convexity*, and $\bar{Y}^+\backslash X(L)$ and $\bar{Y}^-\backslash X(L)$ are *maximal domains of pseudo-convexity* for a quadratic form $Q$.

## 6. Domains of quasi-convexity and pseudo-convexity for quadratic functions.

We wish to extend the analysis of Section 5 to quadratic functions.

With each quadratic function $R(x) = 1/2Q(x) + P(x)$, associate the set

$$M = \{a \in X \mid L(a, x) + P(x) = 0 \text{ for all } x \in X\} .$$

A direct generalization of results in Chapter 4 of [6] gives this sufficient condition.

(20) THEOREM. *If $Y \subset X$ is a domain of negativity determined by $L$ and $M$ is nonempty, then the quadratic function $R(x)$ is quasi-convex on $\bar{Y} + M$ and pseudo-convex on $\bar{Y}\backslash X(L) + M$.*

Before we proceed to determine necessary conditions for the quasi-convexity of a quadratic function on a solid convex set, we have to specify under what conditions the set $M$ is nonempty.

It is obvious that the real topological vector space $X$ can be expressed as

$$X = E^+ \oplus E^- \oplus E^0$$

where $E^+$, $E^-$ and $E^0$ are subspaces of $X$ such that

$$Q(x) > 0 \text{ for all } x \in E^+\backslash 0 ,$$
$$Q(x) < 0 \text{ for all } x \in E^-\backslash 0 .$$
$$Q(x) = 0 \text{ for all } x \in E^0 ,$$

This decomposition may not be unique, but for the rest of this section we make the following *assumption*:

(21) There exists at least one decomposition

$$X = E^+ \oplus E^- \oplus E^0$$

where $E^+$ and $E^-$ are *complete* (i.e., each Cauchy sequence in $E^+$ or $E^-$ is convergent).

Under this assumption the following is true:

(22) PROPOSITION. *If $R(x) = 1/2Q(x) + P(x)$, then either the set $M = \{a \in X \mid L(a, x) + P(x) = 0$ for all $x \in X\}$ is nonempty or there exists a vector $t \in X$ such that $P(t) \neq 0$ and $L(x, t) = 0$ for all $x \in X$.*

*Proof.* First we show that both conditions cannot hold simultaneously. Indeed, suppose there is an $a \in M$; i.e., $L(a, x) + P(x) = 0$ for all $x \in X$. On the other hand, if $t$ is such that $L(x, t) = 0$ for all $x \in X$ and $P(t) \neq 0$, then $x = a$ gives a contradiction.

Next, suppose that if $L(x, t) = 0$ for all $x \in X$, then $P(t) = 0$. Hence $X = E^+ \oplus E^- \oplus E^0$ implies that for all $x \in X$

$$L(a, x) + P(x) = (L(a^+, x^+) + P(x^+)) + (L(a^-, x^-) + P(x^-))$$

where $a^+, x^+ \in E^+$ and $a^-, x^- \in E^-$. Relying on [17, p. 44] it follows that there exist at least one $a^+ \in E^+$ and one $a^- \in E^-$ such that for all $x^+ \in E^+$

$$L(a^+, x^+) + P(x^+) = 0$$

and for all $x^- \in E^-$

$$L(a^-, x^-) + P(x^-) = 0 .$$

This shows that $M$ is nonempty and the proof is complete.

Notice this proposition generalizes to an arbitrary real topological vector space $X$, satisfying assumption (21), a well-known result proved in Gale's book [8, Theorem 2.5] for the case $X = E^n$.

This proposition and similar arguments as in [6, (4.4), (4.13), and (4.15)] are combined to show these results.

(23) THEOREM. *If the quadratic function $R(x) = 1/2Q(x) + P(x)$ is quasi-convex, but not convex, on a solid convex set $S$, then*

( i ) *$M$ is not empty,*

(ii) *there exists a unique pair of domains of negativity,* $Y^+$ *and* $Y^-$, *in* $X$ *determined by* $L$,

(iii) $S \subset \bar{Y}^+ + M$ *or* $S \subset \bar{Y}^- + M$.

(24) THEOREM. *If the quadratic function* $R(x) = 1/2Q(x) + P(x)$ *is pseudo-convex, but not convex, on a solid convex set* $S$ *in* $X$, *then*

(i) $M$ *is not empty,*

(ii) *there exists a unique pair of domains of negativity,* $Y^+$ *and* $Y^-$, *in* $X$ *determined by* $L$,

(iii) $S \subset (\bar{Y}^+ \backslash X(L) + M)$ *or* $S \subset (\bar{Y}^- \backslash X(L) + M)$.

Therefore, if $M$ is nonempty and $Y^+$ and $Y^-$ are a pair of domains of negativity in $X$ determined by $L$, then $\bar{Y}^+ + M$ and $\bar{Y}^- + M$ are *maximal domains of quasi-convexity*, and $\bar{Y}^+ \backslash X(L) + M$ and $\bar{Y} - X(L) + M$ are *maximal domains of pseudo-convexity* for a quadratic function $R$.

## REFERENCES


1. C. Berge and A. Ghouila-Houri, *Programming, Games and Transportation Networks*, Methuen, London, 1965.
2. R. W. Cottle, *On the convexity of quadratic forms over convex sets*, Operations Res., **15** (1967), 170-172.
3. R. W. Cottle and J. A. Ferland, *Matrix-theoretic criteria for the quasi-convexity and pseudo-convexity of quadratic functions*, Linear Algebra and Its Appl., **5** (1972), 123-136.
4. ———, *On pseudo-convex functions of nonnegative variables*, Math. Programming, **1** (1971), 95-101.
5. N. Dunford and J. T. Schwartz, *Linear Operators Part I and II*, Interscience Publishers, Inc., New York, 1964.
6. J. A. Ferland, *Quasi-convex and Pseudo-convex Functions on Solid Convex Sets*, Technical Report No. 71-4, Operations Research House, Stanford University, April, 1971.
7. ———, *Maximal domains of quasi-convexity and pseudo-convexity for quadratic functions*, Math. Programming, **3** (1973), 178-192.
8. D. Gale, *The Theory of Linear Economic Models*, McGraw-Hill, New York, 1960.
9. W. G. Greub, *Linear Algebra*, Springer-Verlag, New York, 1963 (second edition).
10. I. Kaplansky, *Linear Alegebra and Geometry*, Allyn and Bacon, Inc., Boston, 1969.
11. M. Koecher, *Positivitätsbereiche im* $R^n$, Amer. J. Math., **79** (1957), 576-596; translated by R. W. Cottle, "*Domains of Positivity in* $R^n$", Technical Report No. 68-8, Operations Research House, Stanford University, May, 1968.
12. O. L. Mangasarian, *Pseudo-convex functions*, SIAM J. Control, **3** (1965), 281-290.
13. ———, *Nonlinear Programming*, McGraw-Hill Book Company, New York, 1969.
14. B. Martos, *Subdefinite matrices and quadratic forms*, SIAM J. Appl. Math., **17** (1969), 1215-1223.
15. ———, *Quadratic programming with a quasi-convex objective function*, Operations Res., **19** (1971), 87-97.
16. J. Ponstein, *Seven kinds of convexity*, SIAM Rev., **9** (1967), 115-119.
17. H. H. Schaefer, *Topological Vector Spaces*, The Macmillan Company, New York, 1966.


Received June 26, 1972. This research was supported by Canadian National Research Council, Grant No. A8312.

UNIVERSITÉ DE MONTRÉAL

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# A COMPACT SET THAT IS LOCALLY HOLOMORPHICALLY CONVEX BUT NOT HOLOMORPHICALLY CONVEX

MICHAEL FREEMAN AND REESE HARVEY

**It is shown that a certain simple imbedding $T$ of the ordinary two-dimensional torus in $C^2$ contains a polynomially convex compact $T$-neighborhood of each of its points, but $T$ is not holomorphically convex in even the weakest presently accepted sense. This example illustrates some of the limitations of a theory of lower dimensional sets in $C^n$. In particular, it shows the difficulty of developing a theory based on local information.**

In the following $K$ will denote a compact set in $C^n$, $\mathscr{C}(K)$ the Banach algebra of continuous functions on $K$, and $\mathscr{O}(K)$ the algebra of functions holomorphic on some $C^n$ neighborhood of $K$. Also, let $A(K)$ denote the Banach subalgebra of $\mathscr{C}(K)$ obtained by taking the closure of the image of $\mathscr{O}(K)$ in $\mathscr{C}(K)$. A compact set $K$ is said to be *holomorphically convex* if $K$ and the spectrum of $A(K)$ are homeomorphic under the natural map. In [5] a notion of the "envelope of holomorphy", for $K$ a compact subset of $C^n$, was introduced; there it was proved, in particular, that $K$ is equal to its envelope if and only if $K$ is holomorphically convex. The Cartan Theorems A and B for open holomorphically convex sets in $C^n$ admit analogues for compact holomorphically convex sets in $C^n$ (see [5]). One might conjecture that the E. E. Levi problem for open sets in $C^n$ admits a compact analogue. That is, one might conjecture that if $K$ is locally holomorphically convex (i.e., for each point $z \in K$ there exists a compact neighborhood $N$ of $z$ in $K$ such that $N$ is holomorphically convex) then $K$ is holomorphically convex. The example presented below shows that this is not the case.

If "holomorphic approximation" holds on a compact set $K \subset C^n$ (i.e., $\mathscr{O}(K)$ is dense in $\mathscr{C}(K)$) then the spectrum of $A(K) = \mathscr{C}(K)$ is of course homeomorphic to $K$ so that $K$ is holomorphically convex according to the above definition. Even if a compact set $K$ has the property that "local holomorphic approximation" holds (i.e., for each point $z \in K$ there exists a compact neighborhood $N$ of $z$ in $K$ such that $\mathscr{O}(N)$ is dense in $\mathscr{C}(N)$) the set $K$ need not be (globally) holomorphically convex because of the example presented below. In particular, this provides an example of a compact set in $C^n$ where local holomorphic approximation holds but global holomorphic approximation does not hold; as distinguished from the well-known fact that if $K$ is a compact subset of the complex line $\boldsymbol{C}$ and local holomorphic approximation holds then it is true that global holomorphic approximation holds (see for example [2]).

In [4] a notion of a "totally real set in $\boldsymbol{C}^n$" was introduced in order to better understand the properties of $\boldsymbol{R}^n$ in $\boldsymbol{C}^n$ which are crucial for the development of Sato's theory of hyperfunctions. Sato's basic theory [12] was shown to hold with $\boldsymbol{R}^n$ replaced by a totally real set. In the definition of a compact totally real subset $K$ of $\boldsymbol{C}^n$ there are two local requirements which heuristically ensure that $K$ has no (locally) "complex structure of dimension $\geqq 1$" (see [4], Definition 3.4 and the Remark 1 afterward). The example presented below shows that the local information contained in the assertion that $K$ is a totally real set (which is more than just local holomorphic convexity but less than local holomorphic approximation) is not sufficient to ensure that $K$ is holomorphically convex. In particular, in the duality result, Corollary 3.10 of [4], the hypothesis that $K$ be holomorphically convex is necessary.

We would like to acknowledge that R. O. Wells has independently verified that the example given here is not holomorphically convex.

The example is very simple. It is just the two-dimensional torus $T$ imbedded in $\boldsymbol{C}^2$ as $T = \{z\colon (|z_1| - 3)^2 + x_2^2 = 1, y_2 = 0\}$. In fact, (a) the envelope of holomorphy of $T$ is the set

$$\tilde{T} = \{z\colon (|z_1| - 3)^2 + x_2^2 \leqq 1, y_2 = 0\}$$

obtained by filling up $T$ in $\boldsymbol{C} \times \boldsymbol{R} \times \{0\}$; but (b) each point $a$ of $T$ has a compact $T$-neighborhood $N$ on which the polynomials $\boldsymbol{C}[z_1, z_2]$ are dense in the Banach space $\mathscr{C}(N)$ of continuous functions on $N$. Of course this implies in particular that each compact subset of $N$ is polynomially and hence holomorphically convex.

The proof of (a) rests on the observation that $T$ has a basis for its neighborhood system consisting of the Hartogs domains

$$U_\varepsilon = \{z\colon |(|z_1| - 3)^2 + x_2^2 - 1| < \varepsilon, |y_2| < \varepsilon\}, \varepsilon > 0$$

(which are clearly circled in $z_1$ for each fixed $z_2$), and on the proposition below, which asserts that the envelope of holomorphy of $U_\varepsilon$ is

$$\tilde{U}_\varepsilon = \{z\colon (|z_1| - 3)^2 + x_2^2 < 1 + \varepsilon, |y_2| < \varepsilon\}, \frac{1}{2} \geqq \varepsilon > 0 .$$

This shows that any function holomorphic in a neighborhood of $T$ has a holomorphic extension to a neighborhood of $\tilde{T}$. Moreover, since each $\tilde{U}_\varepsilon$ is holomorphically convex, so is $\tilde{T} = \bigcap_{\varepsilon>0} \tilde{U}_\varepsilon$. Thus $\tilde{T}$ is the envelope of holomorphy of $T$ (see [5] for the precise definition of envelope of holomorphy of $T$).

PROPOSITION. *$\tilde{U}_\varepsilon$ is the envelope of holomorphy of $U_\varepsilon$ for* $0 < \varepsilon \leqq 1/2$.

*Proof.* The open set $\tilde{U}_\varepsilon$ is a domain of holomorphy because it is pseudoconvex (see [3] or [8]). The fact that the functions $z \to (|z_1| - 3)^2 + x_2^2$ and $z \to |y_2|^2$ are plurisubharmonic on $\tilde{U}_\varepsilon$, $\varepsilon \leqq 1/2$, implies that $\tilde{U}_\varepsilon$ is pseudoconvex by [8] Theorem 2.6.7 (iii).

Each function $f$ holomorphic on $U_\varepsilon$ has a holomorphic extension to $\tilde{U}_\varepsilon$. For this it suffices to see that the Hartogs-Laurent expansion (see [13] page 130) for $f$ on $U_\varepsilon$,

$$f(z) = \sum_{n=-\infty}^{\infty} f_n(z_2) z_1^n \, , \tag{1}$$

is normally convergent on $\tilde{U}_\varepsilon$, for then its sum will extend $f$ as asserted. Here the coefficients $f_n$ are holomorphic on $\{z_2 \colon x_2^2 < 1 + \varepsilon, |y_2| < \varepsilon\}$. From the normal convergence of (1) on $U_\varepsilon$ it follows that

$$\sum_{n=-\infty}^{\infty} \sup \{|f_n(z_2) z_1^n| \colon z \in K_\delta\} < \infty \, , \tag{2}$$

where $0 \leqq \delta < \varepsilon$ and $K_\delta = \{z \colon (|z_1| - 3)^2 + x_2^2 = 1 + \delta, |y_2| \leqq \delta\}$ (a product of a torus in $\boldsymbol{C} \times \boldsymbol{R}$ and a closed interval in $\boldsymbol{R}$). Now the maximum principle applied (for fixed $z_2$) to $z_1 \to f_n(z_2) z_1^n$ shows that the suprema in (2) become no larger if extended over

$$\tilde{K}_\delta = \{z \colon (|z_1| - 3)^2 + x_2^2 \leqq 1 + \delta, |y_2| \leqq \delta\}.$$

Thus (2) holds with $K_\delta$ replaced by $\tilde{K}_\delta$, and since any compact subset of $\tilde{U}_\varepsilon$ is contained in the interior of some $\tilde{K}_\delta$, the normal convergence of (1) on $\tilde{U}_\varepsilon$ is proved. Thus $\tilde{U}_\varepsilon$ is the envelope of holomorphy of $U_\varepsilon$.

PROPOSITION. *Each point $a$ of $T$ has a compact neighborhood $N$ in $T$ such that $\boldsymbol{C}[z_1, z_2]$ is dense in $\mathscr{C}(N)$.*

*Proof.* Two cases will be distinguished.

(1) The point $a$ is not on one of the top or bottom circles $|z_1| = 3$, $z_2 = \pm 1 + 0i$. Then $a$ is a totally real point of $T$ (i.e., the ordinary real-linear tangent space $T_a$ to $T$ at $a$ is not complex-linear). The proposition is known for this case (see [11], [9] or [6]) but a simple direct proof can be based on the real-analyticity of $T$. It will be shown that $a$ has an open neighborhood $U$ such that $U \cap T$ is mapped into $\boldsymbol{R}^2$ by a biholomorphic map $\psi = (\psi_1, \psi_2) \colon U \to \boldsymbol{C}^2$. Then if $N$ is any compact subset of $T \cap U$, the ordinary Weierstrass Theorem implies that $\boldsymbol{C}[w_1, w_2]$ is uniformly dense in $\mathscr{C}(\psi(N))$. Since $\psi$ is invertible, the polynomial combinations of $\psi_1, \psi_2$ are dense in $\mathscr{C}(N)$.

If $U$ is taken from the beginning as a polycylinder, then $\psi_1$ and $\psi_2$ are approximable on $N$ by polynomials in $z_1, z_2$, which proves the proposition in case (1).

The map $\psi$ will be found by constructing its inverse. Note that there is an open neighborhood $V$ of 0 in $\boldsymbol{R}^2$ and a real-analytic map $\phi\colon V \to T$ such that $\phi(0) = a$ and $d_0\phi(\boldsymbol{R}^2) = T_a$. Here $d_0 f$ denotes the Fréchet derivative of $f$ at 0. Then there is an open set $\tilde{V}$ in $\boldsymbol{C}^2$ such that $\tilde{V} \cap \boldsymbol{R}^2 = V$ and a holomorphic map $\tilde{\phi}\colon \tilde{V} \to \boldsymbol{C}^2$ such that $\tilde{\phi} \mid V = \phi$. Clearly, $d_0\tilde{\phi}(\boldsymbol{R}^2) = T_a$. Moreover, $T_a \cap iT_a = \{0\}$, so $\boldsymbol{C}^2 = T_a + iT_a = d_0\tilde{\phi}(\boldsymbol{R}^2) + id_0\tilde{\phi}(\boldsymbol{R}^2) = d_0\tilde{\phi}(\boldsymbol{R}^2 + i\boldsymbol{R}^2) = d_0\tilde{\phi}(\boldsymbol{C}^2)$. Thus $d_0\tilde{\phi}$ is invertible, so $\tilde{\phi}$ has a holomorphic inverse $\psi$ near 0 by the inverse function theorem.

(2) $|a_1| = 3$, $a_2 = \pm 1 + i0$. Then there is a closed disk $D = \{z_1\colon |z_1 - a_1| \leqq \varepsilon\}$ on which the graph of $g(z_1) = (\operatorname{sign} a_2)\sqrt{1 - (|z_1| - 3)^2}$ defines a compact set $N = \{(z_1, g(z_1))\colon |z_1 - a_1| \leqq \varepsilon\} \subset T$. Clearly, $N$ is a $T$-neighborhood of $a$. Moreover, the level curves of $g$, as arcs of radii $> 1$, do not disconnect $\boldsymbol{C}$ and have no interior points. Therefore, by Mergelyan's Theorem [10] (§5, Theorem 1.5), the polynomial combinations of $z_1$ and $g$ are dense in $\mathscr{C}(D)$. The proposition is proved by transporting this property to $N$ via the homeomorphism $z_1 \to (z_1, g(z_1))$.

There is a result (going back to Grauert [3]) of a positive nature which enables one to conclude from local information that a compact subset $K$ of $\boldsymbol{C}^n$ is holomorphically convex. Briefly, the method is as follows (cf. [11], [9] or [6]). Suppose that in some $\boldsymbol{C}^n$ neighborhood $U_a$ of each point $a \in K$ there exists a $\boldsymbol{C}^2$ nonnegative strictly plurisubharmonic function $\varphi$ such that $K \cap U_a$ equals $\{z \in U_a\colon \varphi(z) = 0\}$. By using a partition of unity one can construct a nonnegative strictly plurisubharmonic function $\varphi$ in a neighborhood $U$ of $K$ such that $K = \{z \in U\colon \varphi(z) = 0\}$. Then for sufficiently small $\varepsilon > 0$, each of the sets $W_\varepsilon = \{z \in U\colon \varphi(z) < \varepsilon\}$ is a Stein open neighborhood of $K$ and $K$ is $\mathscr{O}(W_\varepsilon)$-convex. Hence $K$ is holomorphically convex. The use of this result is limited by the fact (see [7]) that sets $K$ which satisfy the local condition described above must be (locally) contained in a $\mathscr{C}^1$ submanifold of $\boldsymbol{C}^n$ all of whose points are totally real.

On the other hand, this technique is extended in [1], where such a function $\varphi$ (which is only required to be plurisubharmonic-not strictly) is constructed in a neighborhood of a point on a two-manifold where its tangent space is complex linear but whose second-order behavior is sufficiently "hyperbolic" (in a precise sense given in [1]). This result is delimited by the above example $T$, which (in the same sense of [1]) exhibits a kind of "parabolic" behavior at such points.

## REFERENCES

1. M. Freeman, *Exceptional points of real submanifolds*, (in preparation).

2. T. Gamelin, *Uniform Algebras*, Prentice Hall, (1969).

3. H. Grauert, *On Levi's problem and the imbedding of real-analytic manifolds*, Ann. of Math., (2) **68** (1958), 460-472.

4. R. Harvey, *The theory of hyperfunctions on totally real subsets of a complex manifold with applications to extension problems*, Amer. J. Math., **91** (1969), 853-873.

5. R. Harvey and R. O. Wells, *Compact holomorphically convex subsets of a Stein manifold*, Trans. Amer. Math. Soc., **136** (1969), 509-516.

6. ———, *Holomorphic approximation on totally real submanifolds of a complex manifold*, Math. Ann., (to appear).

7. ———, *Zero sets of nonnegative strictly plurisubharmonic functions*, Math. Ann.. (to appear).

8. L. Hormander, *An Introduction to Complex Analysis in Several Variables*, D. Van Nostrand, Princeton, N. J., (1966).

9. L. Hormander and J. Wermer, *Uniform approximation on compact sets in $C^n$* Math. Scand., **23** (1968), 5-21.

10. S. N. Mergelyan, *Uniform approximation to functions of a complex variable*, Amer. Math. Soc. Translation, **101** (1954).

11. R. Nirenberg and R. O. Wells, *Approximation theorems on differentiable submanifolds of a complex manifold*, Trans. Amer. Math. Soc., **142** (1969), 15-35.

12. M. Sato, *Theory of hyperfunctions II*, J. Fac. Sci. Univ. Tokyo, Sect. I, **8** (1960), 387-437.

13. V. Vladimirov, *Methods of the Theory of Functions of Several Complex Variables*, The M. I. T. Press, Cambridge, Mass., (1966).

Received June 22, 1972. The work of author Harvey was supported by NSF grant GP-19011.

UNIVERSITY OF KENTUCKY
AND
RICE UNIVERSITY AND THE INSTITUTE FOR ADVANCED STUDY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# POSITIVE-DEFINITE DISTRIBUTIONS AND INTERTWINING OPERATORS

ROE GOODMAN

**An example is given of a positive-definite measure $\mu$ on the group SL$(2, \boldsymbol{R})$ which is extremal in the cone of positive-definite measures, but the corresponding unitary representation $L^\mu$ is *reducible*. By considering positive-definite *distributions* this anomaly disappears, and for an arbitrary Lie group $G$ and positive-definite distribution $\mu$ on $G$ a bijection is established between positive-definite distributions on $G$ bounded by $\mu$ and positive-definite intertwining operators for the representation $L^\mu$. As an application, cyclic vectors for $L^\mu$ are obtained by a simple explicit construction.**

Introduction. The use of positive-definiteness as a tool in abstract harmonic analysis has a long history, the most striking early instance being the Gelfand-Raikov proof via positive-definite functions of the completeness of the set of irreducible unitary representations of a locally compact group [5]. More recently, it was observed by R. J. Blattner [1] that the systematic use of positive-definite *measures* gives very simple proofs of the basic properties of induced representations, and the cone of positive-definite measures on a group was subsequently studied by Effros and Hahn [4].

The purpose of this paper is two-fold. First, we give an example to show that positive-definite measures do not suffice for the study of intertwining operators and irreducibility of induced representations, despite the claim to the contrary in [4]. Specifically, we exhibit a positive-definite measure $\mu$ on $G = \mathrm{SL}(2, \boldsymbol{R})$ such that $\mu$ lies on an extremal ray in the cone of positive-definite measures on $G$, but the associated unitary representation $L^\mu$ is *reducible*, contradicting Lemma 4.16 of [4].

Our second aim is to show that when $G$ is any Lie group, then the correspondence between intertwining operators and positive functionals on $G$ asserted by Effros and Hahn does hold, provided one deals throughout with positive-definite *distributions* instead of just measures. The essential point is the validity of the Schwartz Kernel Theorem for the space $C_0^\infty(G)$, together with a result of Bruhat [3] about distributions on $G \times G$, invariant under the diagonal action of $G$. Using this correspondence, we obtain cyclic vectors for representations defined by positive-definite distributions, using a modification of the construction in [7]. (The proof of cyclicity given in [7] is invalid, since it assumes the existence of a measure on $G$ corresponding to an arbitrary intertwining operator. Cf. [6] for a proof of cyclicity using von Neumann algebra techniques.)

**1. Notation and statement of theorems.** Let $G$ be a Lie group, and denote by $\mathscr{D}(G)$ the space $C_0^\infty(G)$ with the usual inductive limit topology [10]. Fix a left Haar measure $dx$ on $G$; then $d(xy) = \Delta_G(y)dx$, where $\Delta_G$ is the modular function for $G$. If $\phi \in \mathscr{D}(G)$, define $\phi^*(x) = \overline{\phi(x^{-1})\Delta_G(x)^{-1}}$. Denote by $\mathscr{D}'(G)$ the space of Schwartz distributions on $G$. A distribution $\alpha$ is *positive-definite* if $\alpha(\phi^* * \phi) \geqq 0$ for all $\phi \in \mathscr{D}(G)$, where convolution is defined as usual by

$$(\psi * \phi)(x) = \int_G \psi(y)\phi(y^{-1}x)dy .$$

If $\alpha$ and $\beta$ are distributions, say that $\alpha \ll \beta$ if $\beta - \alpha$ is positive-definite.

Given a positive-definite distribution $\mu$, one obtains a unitary representation $L^\mu$ of $G$ by a standard construction: Let $L_y\phi(x) = \phi(y^{-1}x)$ be the left action of $G$ on $\mathscr{D}(G)$. Then $(L_y\phi)^* * (L_y\psi) = \phi^* * \psi$, so the semi-definite inner product $\mu(\phi^* * \psi)$ is invariant under left translations. Define $I_\mu = \{\phi \in \mathscr{D}(G): \mu(\phi^* * \phi) = 0\}$. The quotient space $\mathscr{D}_\mu = \mathscr{D}(G)/I_u$ is then a pre-Hilbert space with inner product $(\tilde{\psi}, \tilde{\phi})_\mu = \mu(\phi^* * \psi)$, where $\phi \rightarrow \tilde{\phi}$ is the natural mapping of $\mathscr{D}(G)$ onto $\mathscr{D}_\mu$. Let $\mathscr{H}_\mu$ be the completion of $\mathscr{D}_\mu$. The operators $L_y$ pass to the quotient to give a strongly continuous unitary representation $y \rightarrow L_y^\mu$ of $G$ on $\mathscr{H}_\mu$.

Suppose now that $\alpha \in \mathscr{D}'(G)$ satisfies $0 \ll \alpha \ll \mu$. Then $I_\alpha \supseteqq I_\mu$, and there exists a unique self-adjoint operator $A$ on $\mathscr{H}_\mu$ such that

$$(A\tilde{\phi}, \tilde{\psi})_\mu = \alpha(\psi^* * \phi) . \tag{1.1}$$

The operator $A$ obviously satisfies

$$0 \leqq A \leqq I \tag{1.2}$$

$$L_x^\mu A = AL_x^\mu , \tag{1.3}$$

since the Hermitian form $\alpha(\phi^* * \phi)$ is nonnegative, bounded by $(\tilde{\phi}, \tilde{\phi})_\mu = ||\tilde{\phi}||_\mu^2$, and invariant under left translations by $G$. It was asserted (without proof) by Effros and Hahn in [4, §4] that when $\mu$ is a *measure*, then every operator $A$ satisfying (1.2) and (1.3) is given by formula (1.1), where $\alpha$ is a positive-definite *measure*. Unfortunately, this is false in general, as shown by the following example:

THEOREM 1. *There is a positive-definite measure $\mu$ on the group $G = \mathrm{SL}(2, \boldsymbol{R})$ such that:*

(i) *The only measures $\alpha$ satisfying $0 \ll \alpha \ll \mu$ are the measures $c\mu$, $c \in [0, 1]$.*

(ii) *The representation $L^\mu$ of $G$ defined by $\mu$ is reducible.*

If we allow positive-definite *distributions* in formula (1.1), however, then we obtain all intertwining operators, as follows:

THEOREM 2. *Let $G$ be a Lie group, and let $\mu$ be a positive-definite distribution on $G$. Suppose $A$ is an operator on $\mathscr{H}_\mu$ satisfying* (1.2) *and* (1.3). *Then there exists a unique positive-definite distribution $\alpha$ on $G$ such that* (1.1) *holds. Furthermore, the local order of $\alpha$ can be bounded in terms of the local order of $\mu$ and the dimension of $G$.*

REMARKS 1. Theorems 1 and 2 show that the cone of positive-definite measures on $\mathrm{SL}(2, \boldsymbol{R})$ is not a *face* of the cone of positive-definite distributions.

2. For a study of *unbounded* intertwining operators, cf. [9].

3. In case $\mu$ is a positive-definite *measure*, then the distribution $\alpha$ in Theorem 2 has finite global order at most $2(\dim G + 1)$.

A sequence $\{\phi_n\} \subset \mathscr{D}(G)$ will be called a *$\delta$-sequence* if $\phi_n(x) \geqq 0$, $\lim_n \int_G \phi_n(x)dx = 1$, and $\mathrm{Supp}\,(\phi_n) \rightarrow \{1\}$ as $n \rightarrow \infty$. Any $\delta$-sequence is an approximate identity under convolution, of course.

COROLLARY. *Let $\{\phi_n\}$ be a delta sequence, and set $w_n = \phi_n^* * \phi_n$. Then the vector $\xi = \Sigma \lambda_n \tilde{w}_n$ will be a cyclic vector for the representation $L^\mu$, provided $\lambda_n > 0$ and $\lambda_n \longrightarrow 0$ sufficiently fast as $n \rightarrow \infty$.*

2. *Proof of Theorem* 1. Let $G = \mathrm{SL}(2, \boldsymbol{R})$ in this section. We distinguish two closed subgroups of $G$: the subgroup $B$ consisting of all matrices $b = \begin{pmatrix} s & t \\ 0 & s^{-1} \end{pmatrix}$, with $s, t$ real, $s \neq 0$, and the subgroup $V$ consisting of all matrices $v = \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}$, $x$ real. One has $B \cap V = \{1\}$, while $V \cdot B$ consists of all unimodular matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that $a \neq 0$. The map $v, b \rightarrow v \cdot b$ is a diffeomorphism from $V \times B$ to the open subset $V \cdot B$ of $G$. Let $dv$ and $db$ be left Haar measures on $V$ and $B$, respectively, and let $\Delta_B$ be the modular function of $B$. Left Haar measure $dx$ on $G$ is then given by the formula

$$(2.1) \qquad \int_G f(x)dx = \int_V \int_B f(vb)\Delta_B(b^{-1})dbdv = \int_B \int_V f(bv)dbdv$$

[2, Chap. VII, §3, Proposition 6].

Suppose that $p$ is a unitary character of $B$. Then $p(b)db$ is a positive-definite measure on $B$, and the measure $\mu$ on $G$ defined by

$$\int_G f(x)d\mu(x) = \int_B f(b)\Delta_B(b)^{-1/2}p(b)db$$

is positive-definite [1]. As in §1, we denote by $L^\mu$ the corresponding representation of $G$ on $\mathscr{H}_\mu$. The representation $L^\mu$ is equivalent to the "principal series" representation of $G$ induced from the one-dimensional representation $p$ of $B$. Using the integration formula (2.1), we can identify the representation space $\mathscr{H}_\mu$ with $L_2(V, dv)$. (This gives the so-called "non-compact picture" for the principal series [8].) Indeed, if $\phi, \psi \in \mathscr{D}(G)$, then an easy calculation using (2.1) shows that

$$(\tilde{\phi}, \tilde{\psi})_\mu = \int_V \varepsilon(\phi)\overline{\varepsilon(\psi)}dv ,$$

where

$$\varepsilon(\phi)(v) = \int_B \phi(vb)\Delta_B(b)^{-1/2}p(b)db .$$

The restriction of $L^\mu$ to the subgroup $V$ becomes simply the left regular representation of $V$ in this picture.

LEMMA 1. *Let $A$ be a bounded operator on $L_2(V)$ which commutes with left translations by $V$, and suppose that there exists a Radon measure $\alpha$ on $G$ such that*

$$(A\varepsilon(\phi), \varepsilon(\psi))_{L_2(V)} = \alpha(\psi^* * \phi) \tag{2.2}$$

*for all $\phi, \psi \in \mathscr{D}(G)$. Then there is a Radon measure $\nu$ on $V$ such that $Af = f*\nu$, for $f \in \mathscr{D}(V)$.*

*Proof.* Since $A$ is translation invariant, it is enough to establish an estimate

$$|(Af)(1)| \leqq C_K \|f\|_\infty , \tag{2.3}$$

for all $f \in \mathscr{D}(V)$ supported on an arbitrary compact set $K \subset V$ ($\|f\|_\infty$ denoting the sup norm). Let $\mathscr{H}^\infty(V)$ be the space of $C^\infty$ vectors for the left regular representation of $V$. By Sobolev's lemma, $\mathscr{H}^\infty(V) \subset C^\infty(V)$, and $A$ leaves the space $\mathscr{H}^\infty(V)$ invariant. Hence, $A\varepsilon(\phi)$ is a $C^\infty$ function for every $\phi \in \mathscr{D}(G)$.

If $f \in \mathscr{D}(V)$ and $g \in \mathscr{D}(B)$, write $f \otimes g$ for the function $f(v)g(b)$. Via the map $v, b \to vb$ we may consider $f \otimes g$ as an element of $\mathscr{D}(G)$. Then $\varepsilon(f \otimes g) = \lambda_g f$, where $\lambda_g = \int_B g(b)\Delta_B(b)^{-1/2}p(b)db$. In particular, if $\{f_n\}$ and $\{g_n\}$ are $\delta$-sequence in $\mathscr{D}(V)$ and $\mathscr{D}(B)$ respectively, then $\lambda_{g_n} \to 1$ as $n \to \infty$ and $f_n \otimes g_n$ is a $\delta$-sequence on $G$ (by the integration formula (2.1)). Hence, we deduce from (2.2) that

$$A\varepsilon(\phi)(1) = \alpha(\phi)$$

for all $\phi \in \mathscr{D}(G)$. Fix $g \in \mathscr{D}(B)$ such that $\lambda_g = 1$. Then for any $f \in \mathscr{D}(V)$ we have $f = \varepsilon(f \otimes g)$, and hence

$$(Af)(1) = \alpha(f \otimes g) \ . \tag{2.4}$$

Since $\alpha$ is a Radon measure, the right side of (2.4) satisfies (2.3), which proves the lemma. (In fact, $\nu$ is the measure $f \to \alpha(f \otimes g)$.)

*Completion of proof of Theorem* 1. Now take for $p$ the character $p(b) = \operatorname{sgn}(s)$, when $b = \begin{pmatrix} s & t \\ 0 & s^{-1} \end{pmatrix}$. Then it is known [8] that the induced representation $L^\mu$ in this case splits into two parts, and when $\mathscr{H}_\mu$ is realized as $L_2(V)$, then any nontrivial intertwining operator is a scalar multiple of the classical Hilbert transform

$$Af(x) = \lim_{\delta \to 0} \frac{1}{\pi} \int_{|y|>\delta} f(x-y)y^{-1}dy \ .$$

(We identify $V$ with $\boldsymbol{R}$ via the map $x \to \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}$.)

The Hilbert transform does not satisfy estimate (2.3). For example, if

$$f_n(x) = \phi(x) \sum_{k=2}^{n} \frac{\sin(kx)}{k \log k} \ ,$$

where $\phi \in \mathscr{D}(\boldsymbol{R})$ is fixed with $\phi(x) = 1$ for $|x| \leqq 1$, then $\operatorname{Supp}(f_n) \subseteqq \operatorname{Supp}(\phi)$ and $\sup_n \|f_n\|_\infty < \infty$ [11, p. 182].
On the other hand,

$$Af_n(0) = \sum_{k=2}^{n} c_k (k \log k)^{-1} + O(1)$$

as $n \to \infty$, where

$$c_k = \frac{1}{\pi} \int_{-1}^{1} x^{-1} \sin(kx) dx \ .$$

Since $c_k \to 1$ as $k \to \infty$, and since $\Sigma(k \log k)^{-1} = +\infty$, it follows that

$$\sup_n |Af_n(0)| = \infty \ .$$

3. *Proof of Theorem* 2 *and Corollary.* Let $G$ be an arbitrary Lie group (assumed countable at infinity), and let $\mu$ be a given positive-

definite distribution on $G$. If we set $\|\phi\|_\mu = \mu(\phi^* * \phi)^{1/2}$, then $\phi \to \|\phi\|_\mu$ is a continuous seminorm on $\mathscr{D}(G)$. Suppose now that $A$ is a bounded operator on the representation space $\mathscr{H}_\mu$. We may associate with $A$ a bilinear form $B_A$ on $\mathscr{D}(G)$ by the formula

$$B_A(\psi, \phi) = (A\tilde{\phi}, \widetilde{J\psi})_\mu . \tag{3.1}$$

Here $\phi \to \tilde{\phi}$ is the canonical map from $\mathscr{D}(G)$ into $\mathscr{H}_\mu$ as in §1, and $J\phi = \bar{\phi}$ (complex conjugate). By the Schwarz inequality and the boundedness of $A$ we see that

$$|B_A(\psi, \phi)| \leqq \|A\| \, \|\phi\|_\mu \|J\psi\|_\mu . \tag{3.2}$$

Clearly, $\psi \to \|J\psi\|_\mu$ is also a continuous seminorm on $\mathscr{D}(G)$. Although $\|J\psi\|_\mu$ need not be bounded in terms of $\|\psi\|_\mu$, nevertheless, the local order of this seminorm is the same as the local order of $\|\cdot\|_\mu$. (If $K \subset G$ is a compact set and $\rho$ is a continuous seminorm on $\mathscr{D}(G)$, we say that $\rho$ has order $\leqq r$ on $K$ if there is a finite set of differential operators $\{D_j\}$ on $G$ each of order $\leqq r$, such that $\rho(\phi) \leqq \max_j \|D_j\phi\|_\infty$ for all $\phi$ with $\operatorname{Supp}(\phi) \subseteqq K$.)

The main analytic fact we need is the following version of the "kernel theorem" for continuous bilinear forms:

LEMMA 2. *Suppose $B$ is a bilinear form on $\mathscr{D}(G)$, and $\rho_1, \rho_2$ are continuous seminorms on $\mathscr{D}(G)$ such that*

$$|B(\phi, \psi)| \leqq \rho_1(\phi)\rho_2(\psi) . \tag{3.3}$$

*Then there is a distribution $T$ on $G \times G$ such that*

$$B(\phi, \psi) = T(\phi \otimes \psi) .$$

*Furthermore, if $K_1$ and $K_2$ are compact subsets of $G$, and if $\rho_j$ has order $\leqq r_j$ on $K_j (j = 1, 2)$, then $T$ has order $\leqq r_1 + r_2 + 2(\dim G + 1)$ on any compact set $M \subset$ Interior $(K_1 \times K_2)$.*

*Proof.* Since multiplication by a $C^\infty$ function is an operator of order zero, we may use a partition of unity and local coordinates to reduce the problem to a local one in $\boldsymbol{R}^d$, $d = \dim G$, such that $K_j = \{|x| \leqq 2\} \subseteqq \boldsymbol{R}^d$ and $M = \{(x, y); |x| \leqq 1, |y| \leqq 1\} \subseteqq \boldsymbol{R}^d \times \boldsymbol{R}^d$.

Let $\phi_0 \in \mathscr{D}(\boldsymbol{R}^d)$ satisfy $\phi_0 = 1$ on $\{|x| \leqq 1\}$ and $\operatorname{Supp}(\phi_0) \subseteqq K_1$. Set $e_n(x) = \phi_0(x)e^{in\cdot x}$, where $n \in \boldsymbol{N}^d$ and $n\cdot x = n_1x_1 + \cdots + n_dx_d$. Then if $D$ is a differential operator of order $r$, one has $\|De_n\|_\infty \leqq C(1 + |n|)^r$. Hence, the a priori estimate (3.3) implies that for some constant $C > 0$,

$$|B(e_m, e_n)| \leqq C(1 + |m|)^{r_1}(1 + |n|)^{r_2} \tag{3.4}$$

for all $m, n \in \boldsymbol{N}^d$.

Suppose now that $f$ is a $C^\infty$ function on $\boldsymbol{R}^d \times \boldsymbol{R}^d$ with $\operatorname{Supp}(f) \subseteqq M$. Then the Fourier series of $f$ can be written as

$$f(x, y) = \sum_{m,n} \hat{f}(m, n) e_m(x) e_n(y) ,$$

where $\{\hat{f}(m, n)\}$ are the Fourier coefficients of $f$. Define

$$T(f) = \sum_{m,n} \hat{f}(m, n) B(e_m, e_n) . \tag{3.5}$$

The series (3.5) is absolutely convergent, and by (3.4) we have the estimate

$$|T(f)| \leqq C_1 \sup_{m,n} \{|\hat{f}(m, n)| (1 + |m|)^{r_1+d+1} (1 + |n|)^{r_2+d+1}\} , \tag{3.6}$$

where $C_1 = C \sum_{m,n} (1 + |m|)^{-d-1} (1 + |n|)^{-d-1} < \infty$. Since the right side of (3.6) is a seminorm of order $r_1 + r_2 + 2d + 2$ on $M$, this proves the lemma.

*Completion of proof of Theorem* 2. Suppose now that the operator $A$ in formula (3.1) commutes with the representation $L^\mu$. Then the distribution $T$ on $G \times G$ such that $B_A(\phi, \psi) = T(\phi \otimes \psi)$, which was constructed in Lemma 2, satisfies for all $z \in G$,

$$T(\delta_z f) = T(f) , \qquad f \in \mathscr{D}(G \times G) , \tag{3.7}$$

where $\delta_z f(x, y) = f(z^{-1}x, z^{-1}y)$.

The structure of distributions satisfying (3.7) was determined by Bruhat [3, Prop. 3.3]. Let $\iota$ denote the distribution on $G$ determined by left Haar measure, and let $\Phi: G \times G \rightarrow G \times G$ be the map $\Phi(x, y) = (x, xy)$. Then (3.7) forces $T$ to have the form

$$T(f) = (\iota \otimes \alpha)(f \circ \Phi) ,$$

where $\alpha$ is a distribution on $G$. Symbolically,

$$T(f) = \iint f(x, xy) dx d\alpha(y) .$$

In particular, if $\phi, \psi \in \mathscr{D}(G)$, then

$$\begin{aligned} (A\tilde{\phi}, \tilde{\psi})_\mu &= T(J\psi \otimes \phi) \\ &= \iint \overline{\psi(x)} \phi(xy) dx d\alpha(y) \\ &= \alpha(\psi^* * \phi) . \end{aligned}$$

Hence, $\alpha$ serves to represent the intertwining operator $A$, and is obviously positive-definite if $A \geqq 0$. Since $\Phi$ is a diffeomorphism, the order of $\iota \otimes \alpha$ on a compact set $M \subset G \times G$ is the same as the order of $T$ on $\Phi^{-1}(M)$. By Lemma 2 and inequality (3.2), the local order of $\iota \otimes \alpha$ (and, hence, the local order of $\alpha$) can, therefore, be bounded in terms of the local order of $\mu$ and the dimension of $G$, as claimed.

*Proof of Corollary.* Using Theorem 2, we are able to rehabilitate the attempted proof of cyclicity in [7]. Given a $\delta$-sequence $\{\psi_n\}$ on $G$, let $K \subset G$ be a compact set such that $K = K^{-1}$ and $\operatorname{Supp}(\psi_n) \subseteqq K$ for all $n$. Since $\|\psi\|_\mu$ is a continuous seminorm on $\mathscr{D}(G)$, there are right-invariant differential operators $D_1, \cdots, D_r$ on $G$ such that

$$\|\psi\|_\mu \leqq \max_j \|D_j\psi\|_\infty \tag{3.8}$$

for all $\psi$ supported on the set $K^2$.

Now set $w_n = \psi_n^* * \psi_n$, and let $\{\lambda_n\}$ be any sequence such that $\lambda_n > 0$ and

$$\sum_n \lambda_n \max_j \|D_j\psi_n\|_\infty^2 < \infty \; . \tag{3.9}$$

The series $\xi = \sum \lambda_n \tilde{w}_n$ then converges absolutely in $\mathscr{H}_\mu$ (since $\|w_n\|_\mu \leqq \|\psi_n\|_\mu^2$). Let $\mathscr{N}$ be the $G$-cyclic subspace generated by $\xi$, and let $A$ be the projection onto $\mathscr{N}^\perp$. Since $A\xi = 0$, we have $\sum \lambda_n (A\tilde{w}_n, \tilde{\phi})_\mu = 0$ for all $\phi \in \mathscr{D}(G)$. But $\widetilde{\phi * \psi} = L_\mu(\phi)\tilde{\psi}$, where $L_\mu(f) = \int f(x) L_\mu(x)dx$ is the integrated form of the representation. Since $A$ commutes with $L_\mu$, this gives $(A\tilde{w}_n, \tilde{\phi})_\mu = (A\tilde{\psi}_n, \widetilde{\psi_n * \phi})_\mu$. Thus taking $\phi = \psi_k$ and letting $k \to \infty$, we see that

$$\lim_{k\to\infty} (A\tilde{w}_n, \tilde{\psi}_k)_\mu = (A\tilde{\psi}_n, \tilde{\psi}_n)_\mu \tag{3.10}$$

(note that $\phi \to \tilde{\phi}$ is continuous from $\mathscr{D}(G)$ to $\mathscr{H}_\mu$). Furthermore, by the Schwartz inequality, the boundedness of $A$, and the calculation just made, we have the estimate

$$\begin{aligned} |(A\tilde{w}_n, \tilde{\psi}_k)_\mu| &\leqq \|\psi_n\|_\mu \, \|\psi_n * \psi_k\|_\mu \\ &\leqq C \max_j \|D_j\psi_n\|_\infty^2 \; . \end{aligned}$$

(Here we have used estimate (3.8), the right-invariance of $D_j$, and the inequality $\|f * g\|_\infty \leqq \|f\|_\infty \|g\|_{L_1}$.) Thus we may apply the dominated convergence theorem to conclude from (3.9) and (3.10) that $\sum \lambda_n (A\tilde{\psi}_n, \tilde{\psi}_n)_\mu = 0$. But $\lambda_n > 0$ and $A \geqq 0$, so in fact $(A\tilde{\psi}_n, \tilde{\psi}_n)_\mu = 0$ for all $n$. (So far we have simply followed the line of proof of [7], replacing uniform convergence of the series $\sum \lambda_n w_n$ by the stronger condition (3.9), in return for allowing $\mu$ which are distributions rather than measures.) Finally let $\alpha$ be the positive-definite distribution on $G$ representing $A$, which exists by Theorem 2. Then $\alpha(\psi_n^* * \psi_n) = 0$ for all $n$. By the Schwarz inequality, this implies that $\alpha(\phi * \psi_n) = 0$ for all $\phi \in \mathscr{D}(G)$ and all $n$. Letting $n \to \infty$, we conclude that $\alpha = 0$.

## REFERENCES

1. R. J. Blattner, *Positive-definite measures*, Proc. Amer. Math. Soc., **14** (1963), 423-428.

2. N. Bourbaki, *Intégration*, Chap. VII-VIII (Act. Sc. Ind., No. 1306, Hermann, Paris, 1963).

3. F. Bruhat, *Sur les représentations induites des groupes de Lie*, Bull. Soc. Math., France, **84** (1956), 97-205.

4. E. G. Effros and F. Hahn, *Locally compact transformation groups and C* algebras*, Mem. Amer. Math. Soc., No. 75, 1967.

5. I. M. Gelfand and D. Raikov, *Irreducible unitary representations of locally compact groups*, Mat. Sbornik, **13** (1943), 301-316.

6. F. Greenleaf and M. Moskowitz, *Cyclic vectors for representations of locally compact groups*, Math. Ann., **190** (1970), 265-288.

7. A. Hulanicki and T. Pytlik, *Cyclic vectors of induced representations*, Proc. Amer. Math. Soc., **31** (1972), 633-634.

8. A. W. Knapp and E. M. Stein, *Intertwining operators for semisimple groups*, Ann. of Math., (2) **93** (1971), 489-578.

9. Neils Skovhus Poulsen, *On $C^\infty$ vectors and intertwining bilinear forms for representations of Lie groups*, J. Functional Analysis, **9** (1972), 87-120.

10. F. Treves, *Topological Vector Spaces, Distributions, and Kernels*, Academic Press, New York, 1967.

11. A. Zygmund, *Trigonometric Series*, Vol. I, 2nd. Ed., Cambridge Univ. Press, Cambridge, 1968.

Received June 6, 1972 and in revised form November 3, 1972. Partially supported by NSF Grant GP 33567.

RUTGERS UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# THE TYPE OF SOME $C^*$ AND $W^*$-ALGEBRAS ASSOCIATED WITH TRANSFORMATION GROUPS

ELLIOT C. GOOTMAN

**Let $(G, Z)$ be a second countable locally compact topological transformation group, $\mathscr{U}(G, Z)$ the associated $C^*$-algebra and $L$ a certain naturally constructed representation of $\mathscr{U}(G, Z)$ on $L^2(G \times Z, dg \times d\alpha)$, $dg$ being left Haar measure on $G$ and $\alpha$ a quasi-invariant ergodic probability measure on $Z$. Representations of $\mathscr{U}(G, Z)$ constructed from positive-definite measures on $G \times Z$ are used to prove that $\mathscr{U}(G, Z)$ is type $I$ if and only if all the isotropy subgroups are type $I$ and $Z/G$ is $T_0$, and, under the assumption of a common central isotropy subgroup, that $L$ has no type $I$ component if $\alpha$ is nontransitive. By means of quasi-unitary algebras, necessary and sufficient conditions are derived for $L$ to be semi-finite under the weaker assumption of a common type $I$ unimodular isotropy subgroup.**

After establishing notation and discussing preliminary material in §2, we prove in §3 that $\mathscr{U}(G, Z)$ is type $I$ if and only if $Z/G$ is $T_0$ and all isotropy subgroups are type $I$. This result, proven by Glimm [9, Theorem 2.2] for the special case in which isotropy subgroups can be chosen "continuously", is not surprising in light of Mackey's Imprimitivity Theorem and the correspondence between representations of $\mathscr{U}(G, Z)$ and systems of imprimitivity based on $(G, Z)$ (see §2). Our general proof, based on the fact that isotropy subgroups can always be chosen "measurably" [1, Proposition 2.3], follows by construction of a direct integral of certain representations which, by being defined in terms of positive-definite measures, are easily specified and shown to form an integrable family.

In §§4 and 5 we consider the type of a $W^*$-algebra $\mathscr{A}$ constructed via an ergodic quasi-invariant probability measure $\alpha$ on $Z$ (see §4 for the construction). This algebra was studied by Murray and von Neumann in [14], [15], and [16] for the case of $G$ discrete (see also [4, pp. 127–137]), by Dixmier in [3, §§10–12] for the case of $G$ acting freely on $Z$ and by Kallman in [10] for the case in which $\alpha$ is transitive. In §4 we first show that $\mathscr{A}$ is the von Neumann algebra generated by the representation of $\mathscr{U}(G, Z)$ determined by the positive-definite measure $\delta_e \times d\alpha$ on $G \times Z$. Then assuming that almost all ($d\alpha$) points in $Z$ have the same isotropy subgroup $H$, we use a direct integral decomposition of $\mathscr{A}$ arising naturally from a consideration of the measure $\delta_e \times d\alpha$ to prove that if $\alpha$ is nontransitive and if $H$ is in addition central in $G$ then $\mathscr{A}$ has no type $I$ component. In §5 we use different methods, namely the theory of quasi-unitary algebras, to derive necessary and sufficient conditions that $\mathscr{A}$ be semi-finite, under the weaker assumption that almost all $(d\alpha)$ points in $Z$ have the same isotropy subgroup $H$ and that $H$ is type $I$ and unimodular.

The results of §3 are contained in the author's Doctoral Dissertation written at the Massachusetts Institute of Technology under the direction of Professor Roe W. Goodman.

**2. Notation and preliminaries.** If $X$ is a second countable locally compact Hausdorff space, we denote by $\mathscr{K}(X)$ the continuous functions on $X$ of compact support, with the inductive limit topology, and by $M(X)$ the dual space of Radon measures on $X$ with the weak *-topology. For $x \in X$, $\delta_x \in M(X)$ is the probability measure on $X$ concentrated at $x$. For a locally compact group $G$, $d_G g$, or simply $dg$, denotes left Haar measure on $G$ and $\Delta_G$ the corresponding modular function. We assume throughout this paper that both $G$ and $Z$ are second countable locally compact Hausdorff spaces, that all Hilbert spaces are separable and that all representations of algebras are nondegenerate.

Although we refer to [6], primarily §§1, 3, and 4, for the construction of and basic results concerning $\mathscr{U}(G, Z)$, we list for convenience some facts, and establish more notation. $\mathscr{K}(G \times Z)$ is a topological *-algebra and is dense in $\mathscr{U}(G, Z)$ [6, pp. 32–35]. The correspondence $L = \langle V, M \rangle$ between representations $L$ of $\mathscr{U}(G, Z)$ on a Hilbert space $\mathscr{H}$ and systems of imprimitivity $\langle V, M \rangle$ based on $(G, Z)$ and acting on $\mathscr{H}$ is completely determined [6, pp. 34–37] by

$$
\begin{aligned}
&\langle L(f)x, y\rangle \\
&= \int_G \langle Mf(g, \cdot)V(g)x, y\rangle dg,\ f \in \mathscr{K}(G \times Z),\ x, y \in \mathscr{H} .
\end{aligned} \tag{2.1}
$$

If there is no possibility of confusion, we shall use the same symbol $M$ for the representation of $\mathscr{K}(Z)$, its extension to the algebra $L^\infty(Z)$ of bounded Borel functions on $Z$, the corresponding projection-valued measure, and the generated $W^*$-algebra in $\mathscr{B}(\mathscr{H})$. We denote by $D(G \times Z)$ the set of positive-definite measures on $G \times Z$, that is, $\{p \in M(G \times Z): p(f^* * f) \geqq 0\ \forall f \in \mathscr{K}(G \times Z)\}$. $p \in D(G \times Z)$ determines a representation $L^p$ of $\mathscr{U}(G, Z)$ on $\mathscr{H}^p$, and there is a canonical continuous map of $\mathscr{K}(G \times Z)$ onto a dense subspace of $\mathscr{H}^p$ [6, §4].

Blattner's results on induced positive-definite measures and their connection with induced representations [2, Theorem 1] can be extended from the group to the transformation group context. Let $H$ be a closed subgroup of $G$ and $L = \langle V, M \rangle$ a representation of $\mathscr{U}(H, Z)$.

As a special case of [18, §3], one can construct an induced system of imprimitivity $\langle \mathrm{ind}\,(V), \mathrm{ind}\,(M)\rangle$ based on $(G, Z)$ and thus by (2-1) an induced representation $\mathrm{ind}\,(L)$ of $\mathscr{U}(G, Z)$. $\mathrm{Ind}\,(V)$ is the usual representation of $G$ induced from the representation $V$ of $H$. If $p \in D(H \times Z)$ define $\tilde{p} \in M(G \times Z)$ by

$$\tilde{p}(f) = p(f \Delta_G^{1/2} \Delta_H^{-1/2}|_{H \times Z}),\ f \in \mathscr{K}(G \times Z)\ . \tag{2.2}$$

LEMMA 2.3. *If $p$ in $D(H \times Z)$ determines a representation $L$ of $\mathscr{U}(H, Z)$, then $\tilde{p} \in D(G \times Z)$ and determines a representation of $\mathscr{U}(G, Z)$ unitarily equivalent to* $\mathrm{ind}\,(L)$.

*Proof.* The proof of Theorem 1 of [2] can be repeated, with obvious modifications, and we omit the details.

LEMMA 2.4. *If $x \to L^x$ is an integrable family of representations of $\mathscr{U}(H, Z)$, then $x \to \mathrm{ind}\,(L^x)$ is an integrable family of representations of $\mathscr{U}(G, Z)$ and $\int \mathrm{ind}\,(L^x)$ is unitarily equivalent to* $\mathrm{ind}\left(\int L^x\right)$.

*Proof.* We sketch the argument. Let $L^x = \langle V^x, M^x\rangle$ on $\mathscr{H}^x$. By using the approximate identity in $\mathscr{K}(H \times Z)$ and the two formulas in [6, Lemma 3.26] one sees that $x \to V^x(s)$ and $x \to M^x(h)$ are measurable operator fields for $s \in H, h \in \mathscr{K}(Z)$. By Theorem 10.1 of [12], $x \to \mathrm{ind}\,(V^x)$ is a measurable field of representations of $G$ on the induced Hilbert spaces $\mathrm{ind}\,(\mathscr{H}^x)$ and $\int \mathrm{ind}\,(V^x)$ on $\int \mathrm{ind}\,(\mathscr{H}^x)$ is unitarily equivalent to $\mathrm{ind}\left(\int V^x\right)$ on $\mathrm{ind}\left(\int \mathscr{H}^x\right)$. A similar argument verifies that $x \to (\mathrm{ind}\,(M^x))(h)$ is measurable for $h \in \mathscr{K}(Z)$ and that the unitary operator implementing the above equivalence for the representations of $G$ transforms $\int (\mathrm{ind}\,(M^x))(h)$ into $\left(\mathrm{ind}\left(\int M^x\right)\right)(h)$. From the fact $x \to \mathrm{ind}\,(V^x)$ and $x \to \mathrm{ind}\,(M^x)$ are measurable it follows that $x \to \mathrm{ind}\,(L^x)$ is measurable. To see this, note that any $u \in \mathscr{U}(G, Z)$ can be approximated in norm by finite sums of the form $\sum f_i \otimes h_i$, $f_i \in \mathscr{K}(G)$ and $h_i \in \mathscr{K}(Z)$, and then apply (2.1). To finish the proof we note that from the 2 formulas in [6, Lemma 3.26] again, it is clear that for any measurable field of representations $x \to L^x = \langle V^x, M^x\rangle$ the system of imprimitivity corresponding to $\int L^x$ is $\left\langle \int V^x, \int M^x \right\rangle$. Thus the representations $\int \mathrm{ind}\,(L^x)$ and $\mathrm{ind}\left(\int L^x\right)$ are unitarily equivalent because their respective systems of imprimitivity $\left\langle \int \mathrm{ind}\,(V^x), \int \mathrm{ind}\,(M^x)\right\rangle$ and $\left\langle \mathrm{ind}\left(\int V^x\right), \mathrm{ind}\left(\int M^x\right)\right\rangle$ are.

For $\nu \in D(H)$ the measure $\tilde{\nu}$, defined by (2.2) with $Z$ ignored, lies in $D(G)$. If $H$ is the isotropy subgroup of $\varphi \in Z$ then $\nu \times \delta\hat{\varphi} \in D(H \times Z)$ and the induced measure on $G \times Z$ is exactly $\tilde{\nu} \times \delta\hat{\varphi}$. If $L = \langle V, M\rangle$ is the representation of $\mathscr{U}(H, Z)$ determined by $\nu \times \delta\hat{\varphi}$, $V$ is unitarily equivalent to the representation of $H$ determined by $\nu$, $M(k) = k(\varphi)I$ for $k \in \mathscr{K}(Z)$, ind $(M)$ is concentrated on the orbit $G\varphi$, and the commutants of ind$(L)$ and $V$ are algebraically isomorphic (see [6, §4] for details).

3. The type of $\mathscr{U}(G, Z)$. For $\varphi \in Z$ let $H_\varphi$ denote the isotropy subgroup of $\varphi$, $d_{H\varphi}$ a left Haar measure on $H_\varphi$ and $\nu_\varphi$ the induced measure $\tilde{d}_{H\varphi}$.

LEMMA 3.1. *There is a choice of left Haar measures on the isotropy subgroups of $G$ so that for each $f \in \mathscr{K}(G \times Z)$, the function $\theta: Z \rightarrow \boldsymbol{C}$ defined by $\theta(\varphi) = (\nu_\varphi \times \delta_\varphi)(f)$ is bounded and Borel.*

*Proof.* Let $\mathscr{S}(G)$ denote the family of all closed subgroups of $G$, endowed with the compact Hausdorff topology described by Fell in [7]. The map $\varphi \rightarrow H_\varphi$ of $Z$ into $\mathscr{S}(G)$ is Borel [1, Proposition 2.3] and left Haar measures $d_H$ can be chosen on the subgroups $H$ of $G$ so that the map $H \rightarrow \tilde{d}_H$ of $\mathscr{S}(G)$ into $M(G)$ is continuous (this follows from [9, appendix] and the proof of Theorem 4.2 of [8]). Thus for $g \in \mathscr{K}(G)$ the composite map $\varphi \rightarrow \nu_\varphi(g)$ is Borel. To show that $\theta$ is bounded and Borel, we need the following estimate (see [9, Lemma 1.1]). Let $K$ be a compact subset of $G$ and $\ell \in \mathscr{K}(G)$ with $\ell \geqq 0$ and $\ell \equiv 1$ on $K$. Since $H \rightarrow \tilde{d}_H(\ell)$ is a continuous function on the compact set $\mathscr{S}(G)$, it is bounded by a positive constant $\alpha$. For any $k \in \mathscr{K}(G \times Z)$ with $\operatorname{supp} k \subseteqq K \times Z$, and for any $H \in \mathscr{S}(G)$, $\varphi \in Z$, we have

$$(*) \qquad |\tilde{d}_H \times \delta_\varphi(k)| = |\tilde{d}_H(k(\cdot, \varphi))| \leqq \|k\|_\infty |\tilde{d}_H(\ell)| \leqq \alpha \|k\|_\infty .$$

Thus $\theta$ is bounded. Let $A$ and $B$ be compact subsets of $G$ and $Z$ contained, respectively, in relatively compact open sets $U$ and $V$. If $f \in \mathscr{K}(G \times Z)$ with $\operatorname{supp} f \subseteqq A \times B$, $f$ can be uniformly approximated by finite sums of the form $\sum g_i \otimes h_i$, $g_i \in \mathscr{K}(G)$, $\operatorname{supp} g_i \subseteqq U$, $h_i \in \mathscr{K}(Z)$, $\operatorname{supp} h_i \subseteqq V$. The estimate $(*)$, applied to the compact set $K = \bar{U}$, implies that $\theta$ is the uniform limit on $Z$ of the Borel functions $\varphi \rightarrow (\nu_\varphi \times \delta_\varphi)(\sum g_i \otimes h_i) = \sum \nu_\varphi(g_i)h_i(\varphi)$, and is thus Borel.

Fix a "measurable" choice of left Haar measures on the isotropy subgroups as allowed by Lemma 3.1 and for $\varphi \in Z$ let $L^\varphi$ denote the representation of $\mathscr{U}(G, Z)$ on the Hilbert space $\mathscr{H}^\varphi$ determined by $\nu_\varphi \times \delta_\varphi$.

LEMMA 3.2. *For every positive Radon measure $\alpha$ on $Z$ the direct* *integral representation* $L = \int_Z L^\varphi d\alpha(\varphi)$ *exists.*

*Proof.* For $f \in \mathscr{K}(G \times Z)$ let $f'(\varphi)$ denote the canonical image of $f$ in $\mathscr{H}^\varphi$. The map $f \to f'(\varphi)$ is continuous with respect to the inductive limit topology on $\mathscr{K}(G \times Z)$ and the norm topology on $\mathscr{H}^\varphi$ (this follows from [6, Lemma 3.7]). Since $G$ and $Z$ are second countable, $\mathscr{K}(G \times Z)$ contains a countable dense set $\{f_i\}$ [6, proof of Corollary 4.12], and by the preceding remarks and Lemma 3.1, it follows that the $f'_i(\varphi)$ are a fundamental sequence of measurable vector fields and thus the direct integral $\mathscr{H} = \int_Z \mathscr{H}^\varphi d\alpha(\varphi)$ exists [4, Chapter II, §1, $n^0$ 4]. Each $u \in \mathscr{U}(G, Z)$ is the limit in norm of a sequence $h_n \in \mathscr{K}(G \times Z)$ and thus $\langle L^\varphi(u) f'_i(\varphi), f'_j(\varphi)\rangle = \lim_n (\nu_\varphi \times \delta_\varphi)(f_j^* * h_n * f_i)$ is a measurable function on $Z$, again by Lemma 3.1, and the direct integral $L = \int_Z L^\varphi d\alpha(\varphi)$ exists.

It follows from [9, Theorem 2.1] and [6, Theorem 4.29 and Lemma 4.30] that each $L^\varphi$ is an irreducible representation of $\mathscr{U}(G, Z)$ and that $L^\varphi \cong L^\eta$ if and only if $\varphi$ and $\eta$ lie in the same $G$-orbit.

THEOREM 3.3. *$\mathscr{U}(G, Z)$ is type I if and only if the orbit space $Z/G$ is $T_0$ and all the isotropy subgroups are type I.*

*Proof.* If $Z/G$ is not $T_0$, there exists an ergodic positive Borel measure $\alpha$ on $Z$ which is not concentrated on any orbit [5, Theorem 2.6]. By Lemma 3.2 and [5, Lemma 4.2], $L = \int_Z L^\varphi d\alpha(\varphi)$ is a factor representation of $\mathscr{U}(G, Z)$ not of type $I$. Also, since a factor representation $W$ of an isotropy subgroup induces a factor representation $L$ of $\mathscr{U}(G, Z)$ of the same type, the commutants of $L$ and $W$ being algebraically isomorphic, $\mathscr{U}(G, Z)$ is not type $I$ if there is a nontype $I$ isotropy subgroup. Conversely if $Z/G$ is $T_0$ every factor representation $L = \langle V, M\rangle$ is induced from an isotropy subgroup by the Imprimitivity Theorem, since the projection-valued measure $M$ is ergodic and thus concentrated on an orbit. If in addition all the isotropy subgroups are type $I$, so therefore is $\mathscr{U}(G, Z)$.

4. On the type of $\mathscr{A}$. Let $\alpha$ be an ergodic quasi-invariant probability measure on $Z$, $g \cdot \alpha$ the measure defined by $g \cdot \alpha(A) = \alpha(g^{-1}A)$, $g \in G$, $A$ Borel $\subseteq Z$, and $\lambda_g(\cdot)$ the Radon-Nikodym derivative $d(g \cdot \alpha)/d\alpha$. Let $\langle W, P\rangle$ be the system of imprimitivity based on $(G, Z)$ and acting on $L^2(Z, d\alpha)$ by

$$(W(g)f)(\varphi) = \lambda_g(\varphi)^{1/2} f(g^{-1}\varphi) , \quad (P(h)f)(\varphi) = h(\varphi)f(\varphi) ,$$

$g \in G$, $\varphi \in Z$, $h \in L^\infty(Z)$ and $f \in L^2(Z, d\alpha)$. Denoting by $U$ the left regular representation of $G$ on $L^2(G)$, we consider the type of the $W^*$-algebra $\mathscr{A}$ on $L^2(G) \otimes L^2(Z, d\alpha)$ generated by the operators $U(g) \otimes W(g)$ and $I \otimes P(h)$, $g \in G$, $h \in L^\infty(Z)$. Our definition of $\mathscr{A}$ is the same as Kallman's [10] except for modifications due to our preference for left rather than right action of $G$ on $Z$.

LEMMA 4.1. *$\mathscr{A}$ is spatially isomorphic to the $W^*$-algebra generated by the representation $L^\alpha$ of $\mathscr{U}(G, Z)$ determined by $\delta_e \times d\alpha$ in $D(G \times Z)$.*

*Proof.* The natural map of the algebraic tensor product $\mathscr{K}(G) \otimes \mathscr{K}(Z)$ onto a dense subspace of $\mathscr{K}(G \times Z)$ clearly extends to an isometry of $L^2(G) \otimes L^2(Z, d\alpha)$ onto $L^2(G \times Z, dg \times d\alpha)$. By the proof of Theorem 5.3 of [13], $\lambda$ can be chosen to be jointly measurable on $G \times Z$ and it is then clear that under the above isometry the system of imprimitivity $\langle U \otimes W, I \otimes P \rangle$ is transformed into the system $\langle V', M' \rangle$ given by

$$(V'(g)f)(t, \varphi) = \lambda_g(\varphi)^{1/2} f(g^{-1}t, g^{-1}\varphi)$$

and

$$(M'(h)f)(t, \varphi) = h(\varphi) f(t, \varphi) ,$$

$g, t \in G$, $\varphi \in Z$, $h \in L^\infty(Z)$ and $f \in L^2(G \times Z, dg \times d\alpha)$. For $f \in \mathscr{K}(G \times Z)$ define $(Rf)(g, \varphi) = f(g, \varphi)\lambda_g(\varphi)^{1/2}$. $Rf$ is measurable on $G \times Z$. Since

$$\int_G \int_Z | f(g, \varphi) |^2 \lambda_g(\varphi) d\alpha(\varphi) dg = \int_G \int_Z | f(g, g\varphi) |^2 d\alpha(\varphi) dg$$

and $k(g, \varphi) = f(g, g\varphi)$ lies in $\mathscr{K}(G \times Z)$, $Rf$ is square-integrable. Routine calculations verify that $R$ extends from $\mathscr{K}(G \times Z)$ to an isometry of $\mathscr{H}^\alpha$, the Hilbert space of $L^\alpha$, onto $L^2(G \times Z, dg \times d\alpha)$ which transforms the system of imprimitivity given by

$$(V(g)f)(t, \varphi) = f(g^{-1}t, g^{-1}\varphi) \quad \text{and} \quad (M(h)f)(t, \varphi) = h(\varphi) f(t, \varphi) ,$$

$t, g \in G$, $\varphi \in Z$, $h \in L^\infty(Z)$ and $f \in \mathscr{K}(G \times Z)$ into $\langle V', M' \rangle$. To check that $V$ transforms into $V'$ requires use of the identity

$$\lambda_{st}(\varphi) = \lambda_s(\varphi)\lambda_t(s^{-1}\varphi) \text{ a.e. } (d\alpha)$$

for each $s, t \in G$. As $\langle V, M \rangle$ is precisely the system of imprimitivity on $\mathscr{H}^\alpha$ determined by $\delta_e \times d\alpha$ (see formulas 4.4 and 4.6 of [6]) and as $\langle V, M \rangle$ generates exactly the same $W^*$-algebra as the corresponding representation $L^\alpha$ of $\mathscr{U}(G, Z)$, we are done.

Now let $\mathscr{A}$ denote the $W^*$-algebra generated by the representation

$L^\alpha$. Henceforth, we assume that $\alpha$ is concentrated on a $G$-invariant Borel set in $Z$ all of whose points have the same isotropy group $H$, which is a priori normal in $G$. The more general case in which it is assumed merely that all isotropy subgroups are conjugate can be reduced to the above case [1, Chapter II, § 2]. If $\pi$ is a representation of $H$, we denote by $g\cdot\pi$ the representation $(g\cdot\pi)(h) = \pi(g^{-1}hg)$. We shall obtain a direct integral decomposition of $\mathscr{A}$ and then use the following lemma to prove that, under additional hypotheses on $H$, $\mathscr{A}$ has no type $I$ component if $\alpha$ is nontransitive. We denote by $[\mathscr{B}, \mathscr{C}]$ the $W^*$-algebra generated by operator algebras $\mathscr{B}$ and $\mathscr{C}$, by $\mathscr{B}'$ the commutant of $\mathscr{B}$ and by $\mathscr{Z}\mathscr{B}$ the center $\mathscr{B} \cap \mathscr{B}'$ of $\mathscr{B}$.

LEMMA 4.2. *Let $\mathscr{B}$ be a $W^*$-algebra on a Hilbert space $\mathscr{H}$ and $\mathscr{C}$ a commutative subalgebra of $\mathscr{B}'$. If $\mathscr{B}$ has a type I component then so does $\mathscr{D} = [\mathscr{B}, \mathscr{C}]$.*

*Proof.* We use the notation of [4, Chapter I, § 2, $n^0 1$] for induced and reduced algebras. $\mathscr{B}$ has a type $I$ component if and only if there is a nonzero projection $F$ in $\mathscr{Z}\mathscr{B}$ and an abelian projection $E$ in $\mathscr{B}_F$ whose central support is the identity (relative to $\mathscr{B}_F$ on the Hilbert space $F\mathscr{H}$) [4, Chapter II, § 8, $n^0 1$, Corollary 1 and $n^0 2$, Theorem 1]. We shall show that the projections $F$ and $E$ satisfy the same properties for $\mathscr{D}$ as they do for $\mathscr{B}$. Since $\mathscr{Z}\mathscr{B} = \mathscr{B} \cap (\mathscr{B} \cap \mathscr{B}') \subseteq \mathscr{D} \cap (\mathscr{C}' \cap \mathscr{B}') = \mathscr{Z}\mathscr{D}$, $F \in \mathscr{Z}\mathscr{D}$. $\mathscr{C}_F$ is clearly a commutative algebra commuting with $\mathscr{B}_F$ and by [4, Chapter I, § 2 $n^0 1$, Proposition 1], $\mathscr{D}_F$ is generated by $\mathscr{C}_F$ and $\mathscr{B}_F$, and $(\mathscr{D}_F)_E$ is generated by elements of the form $EBCE$, $B \in \mathscr{B}_F$ and $C \in \mathscr{C}_F$. This is because products of the form $BC$, $B \in \mathscr{B}_F$, $C \in \mathscr{C}_F$, form a generating subset of $\mathscr{D}_F$ closed under involution and multiplication. That $(\mathscr{D}_F)_E$ is abelian follows easily now from the hypothesis that $(\mathscr{B}_F)_E$ is abelian and from the fact that $E$ lies in $\mathscr{B}_F$ and thus commutes with $\mathscr{C}_F$. Since $E \in \mathscr{B}_F \subseteq \mathscr{D}_F$, $E$ clearly has central support equal to the identity with respect to the larger algebra $\mathscr{D}_F$, and we are done.

THEOREM 4.3. *Let $\alpha$ be a nontransitive ergodic quasi-invariant probability measure on $Z$, and assume that almost all $(d\alpha)$ points of $Z$ have the same isotropy subgroup $H$. If the left regular representation $T$ of $H$ can be decomposed as a direct integral $T = \int_\Gamma T^\gamma d\gamma$ of irreducibles $T^\gamma$ on $\mathscr{H}^\gamma$, so that* a.e. $(d\gamma)$, *$g\cdot T^\gamma$ is unitarily equivalent to $T^\gamma$ for all $g \in G$, then $\mathscr{A}$ has no type $I$ component.*

*Proof.* We note first that the hypotheses on $T$ are certainly satisfied if $H$ is central in $G$. In any case, since $H$ is normal in $G$, $\Delta_G|_H = \Delta_H$ and $\delta_e \times d\alpha \in D(G \times Z)$ is induced from the measure $\delta_e \times d\alpha \in D(H \times Z)$ (see formula (2.2)). Thus $L^\alpha$ is induced from the representation $R^\alpha$ of $\mathscr{U}(H, Z)$ determined by $\delta_e \times d\alpha$ in $D(H \times Z)$. By applying Lemma 4.1 to $R^\alpha$ one obtains a unitary equivalence between $R^\alpha$ and $\pi^\alpha = \langle T \otimes I, I \otimes Q \rangle$ on $L^2(H) \otimes L^2(Z, d\alpha)$, where $Q$ is the natural projection-valued measure from $Z$ to $L^2(Z, d\alpha)$. As $H$ leaves almost all $(d\alpha)$ points of $Z$ fixed, each $\langle T^\gamma \otimes I, I \otimes Q \rangle$ is a system of imprimitivity, based on $(H, Z)$ and acting on $\mathscr{H}^\gamma \otimes L^2(Z, d\alpha)$. Denoting by $\sigma^\gamma$ the corresponding representation of $\mathscr{U}(H, Z)$ and by $\operatorname{ind} \sigma^\gamma$ the induced representation of $\mathscr{U}(G, Z)$, we have by Lemma 2.4 and its proof a unitary equivalence $L^\alpha \cong \int_\Gamma \operatorname{ind} \sigma^\gamma d\gamma$. If $\mathscr{A}$ had a type $I$ component so would $[\mathscr{A}, L^\infty(\Gamma, d\gamma)]$ by Lemma 4.2, and therefore [4, Chapter II, §3, Exercise 1] so would the representations $\operatorname{ind} \sigma^\gamma$ for $\gamma$ in a set of positive measure on $\Gamma$. We shall use Lemma 4.2 of [5] to verify that in fact $\operatorname{ind} \sigma^\gamma$ is a.e. $(d\gamma)$ a nontype $I$ factor representation, and the theorem will be proven. $Q$ has a natural direct integral decomposition $Q(h) = \int_Z Q^\varphi(h) d\alpha(\varphi)$, where $Q^\varphi(h)$ is multiplication by $h(\varphi)$ on $\mathbf{C}$, $h \in L^\infty(Z)$. Fix $\gamma \in \Gamma$. The system of imprimitivity $\langle T^\gamma \otimes I, I \otimes Q^\varphi \rangle$, or simply $\langle T^\gamma, Q^\varphi \rangle$, on $\mathscr{H}^\gamma \otimes \mathbf{C} = \mathscr{H}^\gamma$ determines a representation $\tau^\varphi$ of $\mathscr{U}(H, Z)$ and again by Lemma 2.4, $\operatorname{ind} \sigma^\gamma \cong \int_Z \operatorname{ind} \tau^\varphi d\alpha(\varphi)$. It follows from Theorem 2.1 of [9] and the discussion preceding that theorem that each $\operatorname{ind} \tau^\varphi$ is an irreducible representation of $\mathscr{U}(G, Z)$, since $T^\gamma$ is an irreducible representation of $H$, and furthermore that $\operatorname{ind} \tau^\varphi$ is unitarily equivalent to $\operatorname{ind} \tau^\eta$ if and only if $\varphi = g \cdot \eta$ and $T^\gamma \cong g \cdot T^\gamma$ for some $g \in G$. If $g \cdot T^\gamma \cong T^\gamma$ for all $g \in G$, $\operatorname{ind} \tau^\varphi \cong \operatorname{ind} \tau^\eta$ if and only if $\varphi$ and $\eta$ lie in the same $G$-orbit. $\alpha$ is thus ergodic with respect to the relation of unitary equivalence among the components $\operatorname{ind} \tau^\varphi$ of $\operatorname{ind} \sigma^\gamma$, and by Lemma 4.2 of [5] $\operatorname{ind} \sigma^\gamma$ is a nontype $I$ factor representation. By hypothesis, this is true a.e. $(d\gamma)$ and we are done.

5. On the type of $\mathscr{A}$ (continued). We derive necessary and sufficient conditions for $\mathscr{A}$ to be semi-finite, under the assumption of a common isotropy group $H$ which is type $I$ and unimodular. Our proof is modelled on Dixmier's in [3, §§10–12], where the case of free action is considered. As there, we assume that the Radon-Nikodym derivative $d(g \cdot \alpha)/d\alpha = \lambda_g(\cdot)$, considered as a function on $G \times Z$, is continuous and strictly positive. With no loss of generality, we also assume that support $\alpha = Z$. We start with the realization of $\mathscr{A}$ as the $W^*$-algebra on $L^2(G \times Z, dg \times d\alpha)$ generated by $\{V(g), M(h): g \in G, h \in L^\infty(Z)\}$, where

$$(5.1)\qquad \begin{aligned} (V(g)f)(t, \varphi) &= \lambda_g(\varphi)^{1/2} f(g^{-1}t, g^{-1}\varphi) \quad \text{and} \\ (M(h)f)(t, \varphi) &= h(\varphi) f(t, \varphi)\,, \quad f \in L^2(G \times Z, dg \times d\alpha)\,. \end{aligned}$$

(See the proof of Lemma 4.1, where $V$ and $M$ are denoted by $V'$ and $M'$.)

For $f \in \mathscr{K}(G \times Z)$, define

$$(5.2)\qquad \begin{aligned} f^j(g, \varphi) &= \Delta(g)^{-1/2} \lambda_g(\varphi)^{1/2} f(g, \varphi) \quad \text{and} \\ f^s(g, \varphi) &= \Delta(g)^{-1/2} \lambda_g(\varphi)^{1/2} \bar{f}(g^{-1}, g^{-1}\varphi)\,. \end{aligned}$$

Also, let $\langle W, N \rangle$ be the system of imprimitivity on $L^2(G \times Z, dg \times d\alpha)$ given by

$$(5.3)\qquad \begin{aligned} (W(g)f)(t, \varphi) &= \Delta(g)^{1/2} f(tg, \varphi) \quad \text{and} \\ (N(h)f)(t, \varphi) &= h(t^{-1}\varphi) f(t, \varphi)\,. \end{aligned}$$

Our definitions differ from Dixmier's due essentially to our preference for left action of $G$ on $Z$. Denote by $L$ and $R$ the representations of $\mathscr{U}(G, Z)$ corresponding, respectively, to $\langle V, M \rangle$ and $\langle W, N \rangle$.

LEMMA 5.4. *$\mathscr{K}(G \times Z)$, with $f^j$ and $f^s$ as in* (5.2), *convolution as multiplication and inner product as in $L^2(G \times Z, dg \times d\alpha)$, is a quasi-unitary algebra with underlying Hilbert space $L^2(G \times Z, dg \times d\alpha)$. Its left algebra $\mathscr{R}^l$ is $\mathscr{A}$ and its right algebra $\mathscr{R}^r = (\mathscr{R}^l)'$ is the algebra generated by $\langle W, N \rangle$.*

*Proof.* That the conditions on [3, p. 277] are satisfied can be verified as in [3, Proposition 9] and we omit the computations. For $f \in \mathscr{K}(G \times Z)$, denote by $\pi^l(f)$ and $\pi^r(f)$, respectively, the bounded operators on $L^2(G \times Z, dg \times d\alpha)$ of left and right convolution by $f$. $\mathscr{R}^l$ and $\mathscr{R}^r$ are, respectively, the $W^*$-algebras generated by all $\pi^l(f)$, $\pi^r(f)$, $f \in \mathscr{K}(G \times Z)$ (see [3, p. 278]). The remainder of the lemma follows by use of (2.1) to verify that $L(f) = \pi^l(f \cdot \lambda^{1/2})$ and $R(f) = \pi^r(\bar{f}^s \cdot \lambda^{-1/2})$ for all $f \in \mathscr{K}(G \times Z)$.

We denote by $J$ the positive self-adjoint extension of $f \to f^j$, by $S$ the isometric extension of $f \to f^s$ [3, p. 278], by $P^l(P^r)$ the set of operators in $\mathscr{R}^l(\mathscr{R}^r)$ commuting with $J$, and by $Q^l(Q^r)$ the operators in $\mathscr{R}^l(\mathscr{R}^r)$ commuting with all of $P^l(P^r)$. Theorem 2 of [3] and Theorem 1 of [17] yield the following: $\mathscr{R}^l$ is semi-finite if and only if there exist (unbounded) positive invertible self-adjoint operators $A$ and $A'$ belonging to $\mathscr{R}^l$ and $\mathscr{R}^r$, respectively, so that $A' = SAS$ and $J$ is the minimal closed extension of $A(A')^{-1}$, and if this is the case, then $A$ and $A'$ belong to $Q^l$ and $Q^r$, respectively, and $Q^l \subseteqq P^l$, $Q^r \subseteqq P^r$. As in [3], we derive necessary and sufficient conditions for $A, A'$ as above to exist in terms of the action of $G$ on some measure space $(B, db)$ by investigating how $A$ and $A'$ correspond to the operators of multiplication by certain elements of $L^\infty(B, db)$. In the case of a nontrivial isotropy subgroup $H$, this necessitates an examination of various direct integral decompositions. We assume familiarity with the notation and results of [4, Chapter II, §§1–3]. If $\pi$ is a representation of a group $K$, we denote by $\pi(K)$ the $W^*$-algebra generated by $\{\pi(k): k \in K\}$.

$L^2(G \times Z, dg \times d\alpha)$ is naturally isometric with the direct integral over $(Z, d\alpha)$ of the constant field of Hilbert spaces $\varphi \to \mathscr{H}(\varphi) = L^2(G)$, with the algebra $M$ corresponding naturally to the algebra of diagonalizable operators. Denote by $\mathscr{L}$ the algebra on $L^2(G \times Z, dg \times d\alpha)$ generated by multiplication by bounded Borel functions on $G \times Z$ and by $\mathscr{L}_1$ the subalgebra generated by the bounded Borel functions on $G/H \times Z$, considered as functions on $G \times Z$. Let $\bar{V}$ and $\bar{W}$ denote, respectively, the left and right regular representations of $G$ on $L^2(G)((\bar{W}(g)f)(t) = \Delta(g)^{1/2}f(tg))$, and $M(G)(M(G/H))$ the algebra on $L^2(G)$ generated by multiplication by bounded Borel functions on $G(G/H)$. Then clearly (see (5.1) and (5.3)) the $W^*$-algebras $\mathscr{L}, \mathscr{L}_1, [\mathscr{L}, V(H)]$, $[\mathscr{L}, \mathscr{Z} V(H)]$ and $[M, \mathscr{Z} V(H)]$ are all the direct integrals, respectively, of the constant fields of $W^*$-algebras $\varphi \to M(G)$, $M(G/H)$, $[M(G), \bar{V}(H)]$, $[M(G), \mathscr{Z} \bar{V}(H)]$ and $\mathscr{Z} \bar{V}(H)$ on $L^2(G)$. Also, each operator $W(g)$ decomposes as $\int_Z \bar{W}(g) d\alpha(\varphi)$.

LEMMA 5.5. *If $H$ is unimodular, then* $Q^l \subseteqq [M, \mathscr{Z} V(H)]$.

*Proof*. It follows from (5.1) and (5.2) that $M \subseteqq P^l$ and that $V(H) \subseteqq P^l$ for $H$ unimodular. If $A \in Q^l \subseteqq \mathscr{R}^l = (\mathscr{R}^r)'$, then $A \in [W(G), N]'$ by Lemma 5.4, and $A \in [M, V(H)]'$ by the preceding remark. By modifying the proof of [3, Lemme 26] so that instead of dealing with compact subsets $K, K'$ of $G$ one deals with subsets of the form $KH$, $K'H$, $K$ and $K'$ compact, it follows that $[M, N] = \mathscr{L} \cap V(H)' = \mathscr{L}_1$. As $(\mathscr{L} \cap V(H)')' = [\mathscr{L}', V(H)] = [\mathscr{L}, V(H)]$ we have $A \in [M, N]' = [\mathscr{L}, V(H)]$. Thus $A = \int_Z A(\varphi) d\alpha(\varphi)$, $A(\varphi) \in [M(G), \bar{V}(H)]$ a.e. $(d\alpha)$, and we must show $A(\varphi) \in \mathscr{Z} \bar{V}(H)$ a.e. $(d\alpha)$. From the fact that $A \in V(H)'$ it follows that $A(\varphi) \in \bar{V}(H)'$ a.e. $(d\alpha)$, and from the fact that $A \in (W(G))' \cap \mathscr{L}_1'$ it follows that $A(\varphi) \in (\bar{W}(G))' \cap (M(G/H))'$ a.e. $(d\alpha)$. By a commutation theorem of Takesaki [19, Theorem 3] the latter algebra is exactly $\bar{V}(H)$ (note that the left and right coset spaces $G/H$, $H\backslash G$ are identical) and we are done.

We now decompose $L^2(G)$ explicitly with respect to the abelian $W^*$-algebra $\mathscr{Z} \bar{V}(H)$. Choose left Haar measure $dh$ and $d\bar{g}$ on $H$ and $G/H$, respectively, so that $\int_G f(g) dg = \int_{G/H} \int_H f(gh) dh\, d\bar{g}$, for all $f \in$

$\mathscr{K}(G)$. Let $\sigma$ denote a Borel cross-section from $G/H$ to $G$ with $\sigma(\bar{e}) = e$, and let $\eta(g) = \sigma(\bar{g})^{-1}g$, so that every $g \in G$ may be written uniquely $g = \sigma(\bar{g})\eta(g)$, $\eta(g) \in H$. Define $\theta(g)$ by

$$\int_H f(ghg^{-1})dh = \theta(g)\int_H f(h)dh,\ f \in \mathscr{K}(H)\ ,$$

and denote by $U_g$ the isometry of $L^2(H)$ into itself given by $(U_g f)(h) = \theta(g)^{1/2}f(g^{-1}hg)$. Let $\tilde{V}$ be the left regular representation of $H$ on $L^2(H)$ and $\int_{\hat{H}} n(\gamma)R^\gamma d\gamma$ its canonical central decomposition. $(G, \hat{H})$ is a Borel transformation group [18, Theorem 2.4].

LEMMA 5.6. *$L^2(G)$ is isometric with the direct integral over $(G/H, d\bar{g})$ of the constant field of Hilbert spaces $\bar{g} \rightarrow L^2(H)$. The operator $U$ implementing the isometry is $(Uf)(\bar{g}, h) = f(\sigma(\bar{g})h)$, $f \in L^2(G)$. For $\mathscr{l} \in L^2(G/H, d\bar{g}, L^2(H))$, $(U^{-1}\mathscr{l})(g) = \mathscr{l}(\bar{g}, \eta(g))$. Furthermore,*

$$U\tilde{V}(h)U^{-1} = \int_{G/H} (\sigma(\bar{g})\cdot \tilde{V})(h)d\bar{g}$$

*and $(\sigma(\bar{g})\cdot \tilde{V})(h) = U^{-1}_{\sigma(\bar{g})}\tilde{V}(h)U_{\sigma(\bar{g})}$, so that $\mathscr{Z}\tilde{V}(H)$ is transformed by $U$ into*

$$\left\{\int_{G/H} U^{-1}_{\sigma(\bar{g})}AU_{\sigma(\bar{g})}d\bar{g}\colon A \in \mathscr{Z}\tilde{V}(H)\right\}.$$

*$\mathscr{Z}\tilde{V}(H)$ is invariant under $A \rightarrow U^{-1}_{\sigma(\bar{g})}AU_{\sigma(\bar{g})}$, and if $A \in \mathscr{Z}\tilde{V}(H)$ corresponds to $f \in L^\infty(\hat{H}, d\gamma)$, $U^{-1}_{\sigma(\bar{g})}AU_{\sigma(\bar{g})}$ corresponds to the function $g^{-1}\cdot f$, given by $(g^{-1}\cdot f)(\gamma) = f(g\cdot\gamma)$.*

*Proof.* All of the statements except the last are either standard results or can be verified easily by direct computation. We note that $\mathscr{l} \in L^2(G/H, d\bar{g}, L^2(H))$ can indeed be considered as a jointly measurable function on $(G/H \times H, d\bar{g} \times dh)$ by [11, Lemma 3.1]. For the last statement of the lemma see, for example, [1, Introduction, Proposition 10.2].

REMARK 1. The automorphisms $A \rightarrow U^{-1}_{\sigma(\bar{g})}AU_{\sigma(\bar{g})}$ of $\mathscr{Z}\tilde{V}(H)$ into itself define an action of $G/H$ on $\mathscr{Z}\tilde{V}(H)$, for if $h \in H$, $U_h$ is the product of a left and a right translation by elements of $H$ and thus commutes with $\mathscr{Z}\tilde{V}(H)$ [3, Theoreme 1]. Thus $G/H$ is an automorphism group on $(\hat{H}, d\gamma)$, as indeed it is on $(Z, d\alpha)$, but we shall continue to regard $G$ as the group acting on these spaces. Since $H$ acts trivially and is unimodular, however, the following equalities, which we shall use shortly, hold: $\sigma(\bar{g})\varphi = g\varphi$, $\Delta(\sigma(\bar{g})) = \Delta(g)$, $\lambda_{\sigma(\bar{g})} = \lambda_g$ and $\theta(\sigma(\bar{g})) = \theta(g)$, $g \in G$, $\varphi \in Z$.

REMARK 2. We shall use Lemma 3.1 of [11], without further explicit mention, to identify $L^2(X, dx, L^2(Y, dy))$ with $L^2(X \times Y, dx \times dy)$ and the space of essentially bounded measurable functions from $(X, dx)$ to $L^\infty(Y, dy)$ with $L^\infty(X \times Y, dx \times dy)$, where $(X, dx)$ and $(Y, dy)$ are each one of the spaces $(Z, d\alpha)$, $(G/H, d\bar{g})$ or $(\hat{H}, d\gamma)$.

By Lemma 5.6 and the discussion preceding Lemma 5.5, an operator $A \in [M, \mathscr{X} V(H)]$ corresponds, after direct integral decomposition of $L^2(G \times Z, dg \times d\alpha)$ over $(Z, d\alpha)$ and $(G/H, d\bar{g})$, to

$$(*) \qquad \int_Z \int_{G/H} U^{-1}_{\sigma(\bar{g})} A(\varphi) U_{\sigma(\bar{g})} d\bar{g} d\alpha(\varphi) \; , \quad A(\varphi) \in \mathscr{X} \tilde{V}(H) \; .$$

But after decomposition over $(\hat{H}, d\gamma)$, $A(\varphi)$ corresponds to multiplication by $f^\varphi \in L^\infty(\hat{H}, d\gamma)$ and $U^{-1}_{\sigma(\bar{g})} A(\varphi) U_{\sigma(\bar{g})}$ corresponds to multiplication by $g^{-1} \cdot f^\varphi$. Regarding $f(\varphi, \gamma) = f^\varphi(\gamma)$ as an element of $L^\infty(Z \times \hat{H}, d\alpha \times d\gamma)$, which may involve changing values of $f$ on a $(d\alpha \times d\gamma)$ null set, we have $A$ corresponding to multiplication by $m(\varphi, \bar{g}, \gamma) = f(\varphi, g \cdot \gamma)$. We now examine what $SAS$ and $J$ correspond to, and we shall obtain our final result.

LEMMA 5.7. *Let $A$ and $f$ be as above. After decomposing over $Z, G/H$ and $\hat{H}$, $SAS$ corresponds to multiplication by $k(\varphi, \bar{g}, \gamma) = \bar{f}(g^{-1}\varphi, \gamma)$ and $J$ corresponds to multiplication by*

$$\ell(\varphi, \bar{g}, \gamma) = \Delta(g)^{-1/2} \lambda_g(\varphi)^{1/2} \; .$$

*Proof.* The result for $J$ follows directly from (5.2). Let $U_1$ be the isometry implementing the decomposition over $Z$ and $G/H$. For $r \in L^2(Z \times G, d\alpha \times dg)$, $(U_1 r)(\varphi, \bar{g}, h) = r(\varphi, \sigma(\bar{g})h)$, and for $r \in L^2(Z \times G/H, d\alpha \times d\bar{g}, L^2(H))$, $(U_1^{-1} r)(\varphi, g) = r(\varphi, \bar{g}, \eta(g))$. $U_1 A U_1^{-1}$ is given by $(*)$. We shall compute $U_1 S U_1^{-1}$ and then

$$(**) \qquad U_1(SAS)U_1^{-1} = (U_1 S U_1^{-1})(U_1 A U_1^{-1})(U_1 S U_1^{-1}) \; .$$

Although the computation of $U_1 S U_1^{-1}$ and other operators by pointwise evaluation yields (pointwise) formulas valid only a.e., these formulas still uniquely determine the element of $L^\infty(Z \times G/H \times \hat{H}, d\alpha \times d\bar{g} \times d\gamma)$ to which $SAS$ corresponds. Thus we may for simplicity ignore a.e. considerations. For $r \in L^2(Z \times G/H, d\alpha \times d\bar{g}, L^2(H))$, it can be verified directly that

$$(U_1 S U_1^{-1} r)(\varphi, \bar{g}, h) = \Delta(g)^{-1/2} \lambda_g(\varphi)^{1/2} \bar{r}(g^{-1}\varphi, \bar{g}^{-1}, \eta(h^{-1}\sigma(\bar{g})^{-1})) \; .$$

Now

$$\begin{aligned} \eta(h^{-1}\sigma(\bar{g})^{-1}) &= \sigma(\bar{g}^{-1})^{-1} h^{-1} \sigma(\bar{g})^{-1} \\ &= (\sigma(\bar{g}^{-1})^{-1} h^{-1} \sigma(\bar{g}^{-1}))(\sigma(\bar{g}^{-1})^{-1} \sigma(\bar{g})^{-1}) \; . \end{aligned}$$

Defining $\Phi(g) = \sigma(\bar{g}^{-1})^{-1}\sigma(\bar{g})^{-1}$ and an operator $\tilde{S}$ on $L^2(H)$ by $(\tilde{S}a)(h) = \bar{a}(h^{-1})$, one can compute directly from the above formulae that

$$\begin{aligned}&(U_1SU_1^{-1}r)(\varphi, \bar{g})\\ &= (\Delta(g)^{-1/2}\theta(g)^{1/2}\lambda_g(\varphi)^{1/2}U_{\sigma(\bar{g}^{-1})}\tilde{V}(\Phi(g))\tilde{S})(r(g^{-1}\varphi, \bar{g}^{-1}))\end{aligned}$$

as elements of $L^2(H)$, and again by direct computation and $(**)$ it follows that

$$\begin{aligned}&(U_1SASU_1^{-1}r)(\varphi, \bar{g})\\ &= (U_{\sigma(\bar{g}^{-1})}\tilde{V}(\Phi(g))\tilde{S}U^{-1}_{\sigma(\bar{g}^{-1})}A(g^{-1}\varphi)U_{\sigma(\bar{g}^{-1})}U_{\sigma(\bar{g})}\tilde{V}(\Phi(g^{-1}))\tilde{S})(r(\varphi, \bar{g}))\ .\end{aligned}$$

It is clear that $\tilde{S}\tilde{S} = I$ on $L^2(H)$, and therefore the operator on $L^2(H)$ given by the right-hand side of the above equation equals the product $T_1T_2T_3$, where

$$\begin{aligned}T_1 &= U_{\sigma(\bar{g}^{-1}))}\tilde{V}(\Phi(g))\tilde{S}U^{-1}_{\sigma(\bar{g}^{-1})}\tilde{S}\ ,\\ T_2 &= \tilde{S}A(g^{-1}\varphi)\tilde{S} \quad \text{and}\\ T_3 &= \tilde{S}U_{\sigma(\bar{g}^{-1})}U_{\sigma(\bar{g})}\tilde{V}(\Phi(g^{-1}))\tilde{S}\ .\end{aligned}$$

Now $A(g^{-1}\varphi) \in \mathscr{L}\tilde{V}(H)$ and thus $\tilde{S}A(g^{-1}\varphi)\tilde{S} = A^*(g^{-1}\varphi) \in \mathscr{L}\tilde{V}(H)$ by [3, Corollaire, p. 283]. By a tedious but straightforward computation, one checks that $T_1 = \tilde{V}(\Phi(g^{-1}))$ and thus

$$\begin{aligned}T_1T_2T_3 &= T_2(T_1T_3)\\ &= A^*(g^{-1}\varphi)(U_{\sigma(\bar{g}^{-1})}\tilde{V}(\Phi(g))\tilde{S}U_{\sigma(\bar{g})}\tilde{V}(\Phi(g^{-1}))\tilde{S})\ .\end{aligned}$$

But $T_1T_3$ equals the identity (again a straightforward computation) and we have finally that

$$(U_1SASU_1^{-1}r)(\varphi, \bar{g}) = A^*(g^{-1}\varphi)(r(\varphi, \bar{g}))\ .$$

Thus $SAS$ corresponds to $k(\varphi, \bar{g}, \gamma) = \bar{f}(g^{-1}\varphi, \gamma)$ and we are done.

THEOREM 5.8. *$\mathscr{A}$ is semi-finite if and only if there exists a positive measurable function $\psi$ on $(Z \times \hat{H}, d\alpha \times d\gamma)$ such that*

$$\frac{\psi(\varphi, g\gamma)}{\psi(g^{-1}\varphi, \gamma)} = \Delta(g^{-1})\lambda_g(\varphi) \text{ a.e. } (dg\times d\alpha\times d\gamma) \quad \text{on} \quad G \times Z \times \hat{H}\ .$$

*Proof.* See [3, Theoreme 7 and Proposition 12] for the proof. Also see [3, Remarque 1, p. 318] for a slight strengthening of the theorem and [3, Remarque 2, p. 319] for the measure-theoretic significance of the hypothesis on $\psi$.

## REFERENCES

1. L. Auslander and C. C. Moore, *Unitary representations of solvable Lie groups*, Mem. Amer. Math. Soc., No. **62** (1966).

2. R. J. Blattner, *Positive definite measures*, Proc. Amer. Math. Soc., **14** (1963), 423-428.
3. J. Dixmier, *Algebres quasi-unitaires*, Comment. Math. Helv., **26** (1952), 275-322.
4. ———, *Les algebres d'operateurs dans l'espace hilbertien* (*Algebres de von Neumann*), Cahiers Scientifiques, fasc. XXV, Gauthier-Villars, Paris, 1957.
5. E. Effros, *Transformation groups and C*-algebras*, Ann. of Math., (2) **81** (1965), 38-55.
6. E. Effros and F. Hahn, *Locally compact transformation groups and C*-algebras*, Mem. Amer. Math. Soc., No. **75** (1967).
7. J. M. G. Fell, *A Hausdorff topology for the closed subsets of a locally compact non-Hausdorff space*, Proc. Amer. Math. Soc., **13** (1962), 472-476.
8. ———, *Weak containment and induced representations of groups. II*, Trans. Amer. Math. Soc., **110** (1964), 424-447.
9. J. Glimm, *Families of induced representations*, Pacific J. Math., **12** (1962), 885-911.
10. R. R. Kallman, *A problem of Gelfand on rings of operators and dynamical systems*, Canad. J. Math., **22** (1970), 514-517.
11. G. W. Mackey, *A theorem of Stone and von Neumann*, Duke Math. J., **16** (1949), 313-326.
12. ———, *Induced representations of locally compact groups. I*, Ann. of Math., (2) **55** (1952), 101-139.
13. ———, *Unitary representations of group extensions. I*, Acta Math., **99** (1958), 265-311.
14. F. J. Murray and J. von Neumann, *On rings of operators*, Ann. of Math., (2) **37** (1936), 116-229.
15. ———, *On rings of operators. IV*, Ann. of Math., (2) **44** (1943), 716-808.
16. J. von Neumann, *On rings of operators. III*, Ann. of Math., (2) **41** (1940), 94-161.
17. L. Pukanszky, *On the theory of quasi-unitary algebras*, Acta Sci. Math. (Szeged), **16** (1955), 103-121.
18. M. Takesaki, *Covariant representations of C*-algebras and their locally compact automorphism groups*, Acta Math., **119** (1967), 273-303.
19. ———, *A generalized commutation theorem for the regular representation*, Bull. Soc. Math. France, **97** (1969), 289-297.

Received June 26, 1972.

UNIVERSITY OF GEORGIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# ANGULAR LIMITS OF LOCALLY FINITELY VALENT HOLOMORPHIC FUNCTIONS

DAVID C. HADDAD

**A function $f$ defined in a domain $D$ is *$n$-valent* in $D$ if $f(z) - w_0$ has at most $n$ zeros in $D$ for each complex number $w_0$. The purpose of this paper is to show that a sufficient condition for a holomorphic function $f$ in $|z| < 1$ to have angular limits almost everywhere on $|z| = 1$ is that there exist a positive integer $n$ and a positive number $r_0$ such that $f$ is $n$-valent in each component of the set $\{z: |f(z)| > r_0\}$.**

We have previously shown that the same conditions on $f$ imply that $f$ is a quasi-normal function of order at most $n - 1$ [3, Theorem 2], and $f$ has angular limits at a dense subset of $|z| = 1$ [3, Corollary 1]. Note that the bound $n$ on the valence of $f$ is the same for each component of $\{z: |f(z)| > r_0\}$. This uniformity on $n$ is essential to the conclusion that $f$ has angular limits almost everywhere on $|z| = 1$; for we have shown in the example in [2] that if the uniformity is dropped, then $f$ need not even have asymptotic values at a dense subset of $|z| = 1$.

If $w = f(z)$ is a nonconstant, holomorphic function in $|z| < 1$, we denote by $F$ the Riemann surface of $f^{-1}$ (as a covering surface over the $w$-plane). If $S$ is a subset of $|z| = 1$, then $m(S)$ denotes the Lebesgue measure of $S$.

A Jordan arc $T = \{z = h(t): 0 < t < 1\}$ lying in a domain $D$ is a *crosscut* of $D$ if $h(t) \to z_0 \in \partial D$ as $t \downarrow 0$, $h(t) \to z_1 \in \partial D$ as $t \uparrow 1$, and $z_0 \neq z_1$. If $z_0 = z_1$, then $T$ is a *loopcut* of $D$.

If a holomorphic function $f$ in $|z| < 1$ is $n$-valent in a component $D(r)$ of the set $\{z: |f(z)| > r\}$ then the connectivity of $D(r)$ is as most $n + 1$ [3, Lemma 3]. We denote by $D^*(r)$ the simply connected domain obtained by adding to $D(r)$ those (at most $n$) components of $\{z: |f(z)| \leqq r\}$ that punch holes in $D(r)$.

LEMMA 1. *Let $f$ be a nonconstant, holomorphic function in $|z| < 1$ that is $n$-valent in each component of the set $\{z: |f(z)| > r_0\}$. For each $r > r_0$, let $\{D_k(r)\}$ denote the at most countable collection of components of $\{z: |f(z)| > r\}$. Then there exists a countable subset $E$ of $(r_0, \infty)$ such that $\partial D_k^*(r)$ is a Jordan curve for all $k$ and all $r \in (r_0, \infty) - E$.*

*Proof.* Define a set $R = \{r: r > r_0$, and $F$ has no branch points lying over the circle $|w| = r\}$. Then the set $(r_0, \infty) - R$ is at most countable. If $r \in R$, then for each $k$, $\partial D_k^*(r) \cap \{|z| < 1\}$ consists of at most countably many crosscuts and loopcuts $T_j^k$ of $|z| < 1$ by [2, Corollary 1].

We show that if for a fixed $k$ there are infinitely many curves $T_j^k$, then their diameters tend to zero as $j \to \infty$. If the diameters did not tend to zero, then the sequence $\{T_j^k\}$ would have an accumulation continuum in $|z| \leqq 1$. Since $f$ is a nonconstant, holomorphic function, $\{T_j^k\}$ cannot have an accumulation continuum in $|z| < 1$. By [2, Theorem 3], $f$ has asymptotic values at a dense subset of $|z| = 1$, and hence, by a theorem of MacLane [4, Theorem 1], the sequence $\{T_j^k\}$ of level curves cannot have an arc of $|z| = 1$ for an accumulation continuum. Hence, the diameters of the curves $T_j^k$ tend to zero as $j \to \infty$.

We still must show that there exists a countable subset $E$ of $(r_0, \infty)$ such that $\partial D_k^*(r)$ has no double points for all $k$ and all $r \in (r_0, \infty) - E$. Suppose to the contrary that $S$ is an uncountable subset of $R$ and that for each $r \in S$ there exists a component $D(r)$ of the set $\{z: |f(z)| > r_0\}$ such that $\partial D^*(r)$ has double points. This implies that for each $r \in S$, $\partial D^*(r)$ contains a loopcut $T_r$, since the curves comprising $D^*(r) \cap \{|z| < 1\}$ are Jordan arcs for all $r \in R$. The domain $D^*(r)$ cannot be interior to a loopcut; for if it were, $f$ would be unbounded in $D^*(r)$ by the extended maximum principle, and, consequently, the loopcut would determine two distinct asymptotic tracts ending at one point contradicting [2, Theorem 2]. (See [4] or [2] for the definition of an asymptotic tract.) Let $G_r$ denote the domain interior to the loopcut $T_r$. The uncountable collection of open sets $G_r$ must contain a pair that intersect, say $G_q$ and $G_s$ where $q < s$. Since the loopcuts $T_q$ and $T_s$ cannot intersect inside $|z| < 1$, then $G_q \subset G_s$, and $T_q$ and $T_s$ end at the same point of $|z| = 1$. By [2, Corollary 1], $T_q$ and $T_s$ determine at least two (since $q \neq s$) asymptotic tracts ending at one point contradicting [2, Theorem 2]. Thus, there must exist a countable subset $E$ of $(r, \infty)$ such that $\partial D_k^*(r)$ is a Jordan curve for all $k$ and all $r \in (r_0, \infty) - E$.

LEMMA 2. *Let $f$ be a nonconstant, holomorphic function in $|z| < 1$ that is n-valent in each component of $\{z: |f(z)| > r_0\}$. If $r_1 > r_0$ and $D(r_1)$ is a component of $\{z: |f(z)| > r_1\}$, then $f$ has angular limits almost everywhere on $E(r_1) = \overline{D(r_1)} \cap \{|z| = 1\}$.*

*Proof.* We assume $m(E(r_1)) > 0$, for, otherwise, there is nothing to prove. For $r < r_1$, we denote by $D(r)$ the component of $\{z: |f(z)| > r\}$ containing $D(r_1)$, and we write $E(r) = \overline{D(r)} \cap \{|z| = 1\}$. We first show that there exists $s \in (r_0, r_1)$ such that $\partial D^*(s)$ is a rectifiable Jordan curve.

By Lemma 1, the set $R = \{r \in (r_0, r_1)$ such that $F$ has no branch points over $|w| = r$ and $\partial D^*(r)$ is a Jordan curve} is the whole interval $(r_0, r_1)$ minus possibly a set of measure zero. Let $C(r) = \partial D(r) \cap \{|z| < 1\}$, and let $\Gamma$ be the family $\{C(r): r \in R\}$. By passing to the Riemann surface $F$, it is not hard to show that the extremal length of the family $\Gamma$ is bounded by $2n\pi \log r_1/r_0$, and this implies $\partial D(r)$ is rectifiable for infinitely many values $r \in R$ (for example, see [2, Theorem 1]). Thus, we can choose $s \in (r_0, r_1)$ such that $\partial D^*(s)$ is a rectifiable Jordan curve.

By the Riemann mapping theorem and Carathéodory's theorem on boundary correspondence there exists a homeomorphism $g$ of $\overline{D^*(s)}$ onto $|\zeta| \leqq 1$ that is a conformal mapping of $D^*(s)$ onto $|\zeta| < 1$. Since the connectivity of $D(s)$ is finite, $|f(g^{-1}(\zeta))| > s$ in some annulus $t < |\zeta| < 1$. Hence, $f \circ g^{-1}$ has angular limits almost everywhere on $|\zeta| = 1$ by a simple extension of theorems of Fatou [1, p. 19] and F. and M. Riesz [1, p. 22] on angular limits. Since $\partial D^*(s)$ is a rectifiable Jordan curve, $g^{-1}$ maps a set of measure zero on $|\zeta| = 1$ onto a set of measure zero on $\partial D^*(s)$ by a theorem of F. and M. Riesz [1, p. 50]. Thus $f$ has asymptotic values almost everywhere on $E(s)$ and hence angular limits almost everywhere on $E(s)$ by [3, Theorem 3]. This completes the proof of the lemma since $E(s) \supset E(r_1)$.

LEMMA 3. *Let $\{I_j\}$ be a sequence of mutually disjoint open arcs on $|z| = 1$, and let $C = \bigcup_j I_j$. Let $f$ be a continuous function on $\{|z| < 1\} \cup C$ that is holomorphic in $|z| < 1$. Let $|f(z)| = r_0$ for $z \in C$, $|f(0)| > r_0$, and the set $D = \{z: |z| < 1, |f(z)| > r_0\}$ be a connected set whose boundary contains the circle $|z| = 1$. If $f$ is $n$-valent in $D$, then $|f(0)| \leqq r_0 \exp [2\pi^4 n/m(C)^2]$.*

*Proof.* Let $\gamma(r)$ be the level set $\{z: |f(z)| = r\}$. The proof consists of finding bounds on the extremal length $\lambda(\Gamma)$ of the family $\Gamma = \{\gamma(r): r_0 < r < |f(0)|$, and $F$ has no branch points lying over $|w| = r\}$. By passing to the Riemann surface $F$, it can be shown that $\lambda(\Gamma) \leqq 2\pi n/\log |f(0)|/r_0$ (for example, see [2, Theorem 1]).

By our hypotheses on $f$, each arc $I_j$ must be separated from the point $z = 0$ by a level curve of $\{z: |f(z)| = r\}$ for each $r$ in the interval $(r_0, |f(0)|)$. None of these curves can be relatively compact curves encircling the point $z = 0$ by the maximum principle. Thus, the Euclidean length of a level curve separating $I_j$ from $z = 0$ is bounded below by $\min (2, (2/\pi)m(I_j))$. Hence, the Euclidean length of each $\gamma(r) \in \Gamma$ is bounded below by $(1/\pi)m(C)$. By considering the linear density $\rho(z)$ defined to be 1 on $D$ and 0 elsewhere, we can easily obtain the inequality $\lambda(\Gamma) \geqq (1/\pi^3)m(C)^2$. Combining the two bounds on $\lambda(\Gamma)$ we have $|f(0)| < r_0 \exp [2\pi^4 n/m(C)^2]$, which completes the proof of the lemma.

A point $e^{i\theta}$ is a *Plessner point* for a function $f$ defined in $|z| < 1$ if for every Stolz angle $S$ at $e^{i\theta}$, the cluster set of $f$ at $e^{i\theta}$ with respect to the domain $S$ is total.

THEOREM. *A sufficient condition for a holomorphic function $f$ in $|z| < 1$ to have finite angular limits almost everywhere on $|z| = 1$ is that there exist a positive number $r_0$ and a positive integer $n$ such that $f$ is $n$-valent in each component of the set $\{z: |f(z)| > r_0\}$.*

*Proof.* Suppose to the contrary that the set of points of $|z| = 1$ at which $f$ does not have finite angular limits has positive measure. Then, by a theorem of Plessner [1, p. 147] and a theorem of Priwalow [1, p. 146], $f$ must be a nonconstant function whose set of Plessner points $P$ has positive measure.

For each $r > 0$, let $\{D_j(r)\}$ denote the at most countable collection of components of the set $\{z: |f(z)| > r\}$. By Lemma 1, there exists $r_1 > r_0$ such that $\partial D_j^*(r_1)$ is a Jordan curve for each $j$ and $F$ has no branch points over the circle $|w| = r_1$. Thus, $\partial D_j^*(r_1) \cap \{|z| < 1\}$ consists of at most countably many level curves which are crosscuts of $|z| < 1$. Write $D_j = D_j(r_1)$, $E_j = \overline{D_j^*} \cap \{|z| = 1\}$, and $E_j' = \{|z| = 1\} - E_j$. Since by Lemma 2, $f$ has angular limits almost everywhere on $\bigcup_j E_j$, we can assume $P \subset \bigcap_j E_j'$. Let $\omega_j$ be the harmonic measure in $D_j^*$ of the set $\partial D_j^* \cap \{|z| < 1\}$. We need the following lemma whose proof we postpone.

LEMMA 4. *There exists a harmonic function $v$ in $|z| < 1$ having angular limit 0 almost everywhere on $\bigcap_j E_j'$, and $\omega_j(z) \geqq 1 - v(z)$ for $z \in D_j^* (j = 1, 2, \cdots)$.*

Thus, there exists a point $z_0 \in P$ at which $v$ has angular limit 0. Then, by the definition of $P$, there exists a sequence $\{z_k\}$ of points lying inside a Stolz angle at $z_0$ and converging to $z_0$ such that $|f(z_k)| > r_1$ for each $k$ and $f(z_k) \to \infty$ as $k \to \infty$. At most finitely many $z_k$ can lie in the same component $D_j$ since $z_0 \in P \subset \bigcap_j E_j'$. Hence, we can assume (by taking subsequences if necessary) that $z_j \in D_j$ $(j = 1, 2, \cdots)$ and $D_j \cap D_k = \phi$ for j $\neq k$. By Lemma 4, $\omega_j(z_j) \to 1$ as $j \to \infty$.

By the Riemann mapping theorem and Carathéodory's theorem on boundary correspondence, there exists a homeomorphism $g_j$ of $\overline{D_j^*}$ onto $|\zeta| < 1$ that is a holomorphic map of $D_j^*$ onto $|\zeta| < 1$ sending $z_j$ into 0. Applying Lemma 3 to the function $h_j = f \circ g^{-1}$ and the set $C_j = g_j(\partial D_j^* \cap |z| < 1)$ we have $|h_j(0)| \leqq r_1 \exp[2\pi^4 n/m(C_j)^2]$. On the one hand, $h_j(0) = f(z_j) \to \infty$ as $j \to \infty$. On the other hand, $m(C_j) = 2\pi\omega_j(g_j^{-1}(0)) = 2\pi\omega_j(z_j) \to 2\pi$ as $j \to \infty$, and this implies $h_j(0) \nrightarrow \infty$ as $j \to \infty$. Thus to complete the proof of the theorem, we need only prove Lemma 4.

Let $u_j$ be the harmonic measure in $|z| < 1$ of the set $E'_j$, and let $v_k(z) = \sum_{j=1}^{k}(1 - u_j(z))$. Clearly, $\{v_k\}$ is an increasing sequence of nonnegative harmonic functions, and $v_k$ has angular limit 0 at each point of the set $\bigcap_{j=1}^{k} E'_j$. Since the set $\overline{D_j^*} \cap \overline{D_q^*}$ can contain at most two points for $j \neq q$, each point $e^{i\theta}$ lies in at most one of the sets $E_1, E_2, \cdots, E_k$ for all but finitely many values of $\theta$ in the interval $[0, 2\pi)$. Hence, $\overline{\lim}_{z\to e^{i\theta}} v_k(z) \leqq 1$ $(k = 1, 2, \cdots)$ for all but finitely many values $\theta \in [0, 2\pi)$. It follows from the extended maximum principle that $v_k(z) \leqq 1$ for $|z| \leqq 1 (k = 1, 2, \cdots)$. By Harnack's theorem, the sequence $\{v_k\}$ converges in $|z| < 1$ to a bounded harmonic function $v(z)$. Let $I = \{\theta\colon 0 \leqq \theta < 2\pi, e^{i\theta} \varepsilon \bigcap_j E'_j$, and $v, v_1, v_2, \cdots$ have angular limits at $e^{i\theta}\}$. Then, writing $v(e^{i\theta})$ for the angular limit of $v$ at $e^{i\theta}$, we have

$$\begin{aligned}\int_I v(e^{i\theta})d\theta &= \int_I v(e^{i\theta}) - v_k(e^{i\theta})d\theta \\ &\leqq \int_0^{2\pi} v(e^{i\theta}) - v_k(e^{i\theta})d\theta \\ &= v(0) - v_k(0)\ .\end{aligned}$$

Thus, $v$ has angular limit 0 at $e^{i\theta}$ for almost all $\theta \in I$, since $v(0) - v_k(0) \to 0$ as $k \to \infty$. Since the set $\bigcap_j E'_j - \{e^{i\theta}\colon \theta \in I\}$ has measure zero by Fatou's theorem, $v$ has angular limit 0 almost everywhere on $\bigcap_j E'_j$. Clearly, $v(z) \geqq v_j(z) \geqq 1 - u_j(z)$ for all $j$ and $|z| < 1$, and by Carleman's principle of domain extension, $\omega_j(z) \geqq u_j(z)$ for $z \in D_j^*$ $(j = 1, 2, \cdots)$. This completes the proof of Lemma 4 and hence of the theorem.

REMARK. The conclusion of the theorem raises the following question. Are all functions that satisfy the hypotheses of the theorem of bounded characteristic? This seems to be a difficult question to answer. The best we can presently show is that $T(r) = o(1/1 - r)$, where $T$ is the Nevanlinna characteristic of $f$.

## REFERENCES

1. E. F. Collingwood and A. J. Lohwater, *The Theory of Cluster Sets*, Cambridge University Press, 1966.

2. D. C. Haddad, *Asymptotic values of finitely valent functions*, Duke Math. J., **39** (1972), 391-367.

3. ———, *Boundary behavior and quasi-normality of finitely valent holomorphic functions*, to appear, Canad. J. Math., (1973).

4. G. R. MacLane, *Asymptotic Values of Holomorphic Functions*, Rice University Studies, **49** No. 1 (1963).

Received June 22, 1972. The author is indebted to the late Professor G. R. MacLane for his considerable assistance.

WEST VIRGINIA COLLEGE OF GRADUATE STUDIES

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# ON QUASI-COMPLEMENTS

WILLIAM B. JOHNSON

**Results of H. P. Rosenthal and the author on $w^*$-basic sequences are combined with known techniques and applied to quasi-complementation problems in Banach spaces.**

1. **Introduction.** Recall that (closed, linear) subspaces $Y, Z$ of the Banach space $X$ are *quasi-complements* (respectively *complements*) provided $Y \cap Z = \{0\}$ and $Y + Z$ is dense in $X$ (respectively, $Y + Z = X$).

Suppose that $Y, Z$ are quasi-complements, but not complements, for the separable space $X$. We show that there exist closed subspaces $Y_1$ and $Y_2$ of $X$ with $Y_1 \subset Y \subset Y_2$, $\dim Y/Y_1 = \infty = \dim Y_2/Y$, such that $Y_1, Z$ are quasi-complements and $Y_2, Z$ are quasi-complements. This generalizes a theorem of James [5], who proved the existence of $Y_1$ for the case of general separable $X$ and the existence of $Y_2$ for separable, reflexive $X$. Our proof uses James' method (and $w^*$-basic sequences), but seems simpler than James' construction. Also, our argument provides information for some nonseparable spaces.

We show also the following.

THEOREM 2. *Suppose $Y$ is a subspace of $X$ and $Y^*$ is weak*-separable. If $X/Y$ has a separable, infinite dimensional quotient space, then $Y$ is quasi-complemented in $X$.*

Theorem 2 was discovered by J. Lindenstrauss and H. P. Rosenthal [unpublished], both of whom apparently use an idea from [3]. Our argument uses $w^*$-basic sequences and Rosenthal's proof of Theorem 2 in the case where $X/Y$ has a reflexive, infinite dimensional quotient (cf. [12]).

The final result of the paper is that every subspace of a separable conjugate space admits a weak*-closed quasi-complement which is spanned by a boundedly complete $w^*$-basic sequence.

The notation and terminology agree with [6]. In particular, subspaces and quotients are assumed to be infinite dimensional and complete. For $A \subset X$, $A^\perp$ is the annihilator of $A$ in $X^*$, while for $B \subset X^*$, $B^\tau$ is the annihilator of $B$ in $X$ and $\widetilde{B}$ is the weak*-closure of $B$ in $X^*$.

II. THE THEOREMS. We recall the definition of $w^*$-basic sequence

[6]: A sequence $(y_n) \subset X^*$ is called $w^*$-basic provided that there exists $(x_n) \subset X$ biorthogonal to $(y_n)$ and, for each $y$ in the weak*-closure $\widetilde{[y_n]}$ of the closed linear span $[y_n]$ of $(y_n)$, $y = w^*\text{-}\lim_n \sum_{i=1}^n y(x_i)y_i$.

In [6] it was proved that, when $X$ is separable, if $(y_n) \subset X^*$, $y_n \xrightarrow{w^*} 0$, but $\liminf \|y_n\| > 0$, then $(y_n)$ contains a $w^*$-basic subsequence. Let us note that the same result is true when $X$ admits a weakly compact fundamental set. Indeed, in this case there exists by [1] a norm one projection $P$ on $X$ with $PX$ separable and $(y_n) \subset P^*X^*$. $P^*X^*$ is isometric to $(PX)^*$ and the relative weak* topology on $P^*X^*$ from $X^*$ agrees with the weak* topology on $P^*X^*$ considered as the conjugate of $PX$. Therefore, the above mentioned result from [6] applies to show that $(y_n)$ has a $w^*$-basic subsequence.

First we prove the extension of James' theorem:

THEOREM 1. *Suppose that $Y, Z$ are quasi-complements, but not complements, for $X$.*

(a) *If $Y$ has a weakly compact fundamental subset, then there exists a subspace $Y_1$ of $Y$ with dim $Y/Y_1 = \infty$ and $Y_1, Z$ are quasi-complements.*

(b) *If $X/Y$ has a weakly compact fundamental subset (in particular, if $X$ does), then there exists a subspace $Y_2$ of $X$ with $Y_2 \supset Y$, dim $Y_2/Y = \infty$, and $Y_2, Z$ are quasi-complements.*

*Proof.* Pick positive numbers $(a_n)$ less than 1 so that $a_1 + a_1a_2 + a_1a_2a_3 + \cdots < \infty$. Let $p$ be a bijection of $N \times N$ onto $N$ ($N$ is the set of natural numbers) so that for each $n$ and $j$, $p(n, j) \geqq j$.

To prove (a), we use the fact that $Y + Z$ is not closed to select unit vectors $(y_n)$ in $Y$ with $d(y_n, Z) \equiv \inf\{\|y_n + z\|: z \in Z\} \to 0$. Since $Y \cap Z = \{0\}$, 0 is the only possible weak cluster point of $(y_n)$, and hence either $y_n \xrightarrow{w} 0$ or the weak closure of $(y_n)$ is not weakly compact. Thus, by either [2] or [11], $(y_n)$ has a basic subsequence, which we also denote by $(y_n)$.

Let $(y_n^*)$ be a bounded sequence of functionals in $Y^*$ biorthogonal to $(y_n)$. Since $Y$ admits a weakly compact fundamental set, the unit ball of $Y^*$ is weak* sequentially compact (cf. [1]), so we may assume, by passing to a subsequence, that $y_n^* \xrightarrow{w^*} y^*$. $(y_n^* - y^*)$ converges $w^*$ to 0 and is bounded away from zero, so it has a $w^*$-basic subsequence. Thus by passing to a subsequence of $(y_n, y_n^* - y^*)$, we have that there exists a biorthogonal sequence $(x_n, x_n^*)$ in $Y$ with $\|x_n\| = 1$, $(\|x_n^*\|)$ bounded, $d(x_n, Z) \leqq n^{-1}a_1a_2a_3 \cdots a_n$, $(x_n)$ is basic, and $(x_n^*)$ is $w^*$-basic.

Let $Y_1 = [(x_i^*)^\tau \cup (a_i x_{p(n,i)} - x_{p(n,i+1)})_{i,n=1}^\infty]$. (The annihilator of $(x_i^*)$ is of course taken in $Y$.) We claim that $Y_1 \cap [x_{p(n,1)}] = \{0\}$. To see this, first note that $w_n^* = x_{p(n,1)}^* + a_1 x_{p(n,2)}^* + a_1 a_2 x_{p(n,3)}^* + \cdots$ is absolutely convergent, $w_n^*(x_{p(n,1)}) = 1$, while $w_n^*(x_{p(m,1)}) = 0$ when $n \neq m$. By construction, $Y_1 \subset (w_n^*)^\tau$, and $(w_n^*)^\tau \cap [(x_{p(n,1)})] = \{0\}$ because $(x_{p(n,1)})$ is basic under some ordering and $(x_{p(n,1)}, w_n^*)$ is biorthogonal. Hence, $Y_1 \cap [x_{p(n,1)}] = \{0\}$, whence dim $Y/Y_1 = \infty$.

We complete the proof by showing that $Y_1 + Z$ is dense in $X$. Now $(x_n^*)^\tau + [x_n]$ is dense in $Y$ because $(x_n^*)$ is $w^*$-basic, so we need show only that $(x_{p(n,1)}) \subset \overline{Y_1 + Z}$. But

$$\begin{aligned} x_{p(n,1)} &- a_1^{-1}(a_1 x_{p(n,1)} - x_{p(n,2)}) - (a_1 a_2)^{-1}(a_2 x_{p(n,2)} - x_{p(n,3)}) \\ &- \cdots - (a_1 a_2 \cdots a_j)^{-1}(a_j x_{p(n,j)} - x_{p(n,j+1)}) \\ &= (a_1 a_2 \cdots a_j)^{-1} x_{p(n,j+1)} \,. \end{aligned}$$

Since $d(x_{p(n,j+1)}, Z) \leqq p(n, j+1)^{-1} a_1 a_2 \cdots a_{p(n,j+1)} \leqq (j+1)^{-1} a_1 a_2 \cdots a_j$, it follows that $d(x_{p(n,1)}, Y_1 + Z) \leqq (j+1)^{-1}$. Since $j$ is arbitrary, this completes the proof of (a).

The proof of (b) is very similar to the above: Since $Y, Z$ are not complements, $Y^\perp + Z^\perp$ is not closed in $X^*$. Thus there exists a sequence $(y_n^*)$ of unit vectors in $Y^\perp$ with $d(y_n^*, Z^\perp) \to 0$. Of necessity, $y_n^* \xrightarrow{w^*} 0$. Now $Y^\perp = (X/Y)^*$ in the canonical way, so $(y_n^*)$ has a $w^*$-basic subsequence. Hence for an appropriate subsequence $(x_n^*)$ of $(y_n^*)$, we have that there exists a biorthogonal sequence $(x_n, x_n^*)$ in $X$ with $(\|x_n\|)$ bounded, $\|x_n^*\| = 1$, $(x_n^*) \subset Y^\perp$, $(x_n^*)$ $w^*$-basic, and $d(x_n^*, Z^\perp) \leqq n^{-1} a_1 a_2 \cdots a_n$.

We define $Y_2^\perp$ to be the weak*-closure of $[Y^\perp \cap (x_n)^\perp \cup (a_i x_{p(n,i)}^* - x_{p(n,i+1)}^*)_{n,i=1}^\infty]$. Since $Y_2^\perp \subset Y^\perp$, we have $Y_2 \supset Y$. To show that dim $Y_2/Y = \infty$, it clearly suffices to prove that $Y_2^\perp \cap \widetilde{[x_{p(n,1)}^*]} = \{0\}$. But note that $y_n \equiv x_{p(n,1)} + a_1 a_2 x_{p(n,2)} + a_1 a_2 a_3 x_{p(n,3)} + \cdots$ is absolutely convergent, $x_{p(n,1)}^*(y_n) = 1$, while $x_{p(m,1)}^*(y_n) = 0$ when $m \neq n$. By construction, $(y_n)^\perp \supset (a_i x_{p(n,i)}^* - x_{p(n,i+1)}^*)_{n,i=1}^\infty$ and $(y_n)^\perp \supset (x_n)^\perp$, hence $(y_n)^\perp \supset Y_2^\perp$. But $(y_n)^\perp \cap \widetilde{[x_{p(n,1)}^*]} = \{0\}$ because $(x_{p(n,1)}^*)$ is $w^*$-basic in some ordering and $(y_n, x_{p(n,1)}^*)$ is biorthogonal.

Since $Y_2^\perp \cap Z^\perp \subset Y^\perp \cap Z^\perp = \{0\}$, we have that $Y_2 + Z$ is dense in $X$. To show that $Y_2 \cap Z = \{0\}$, we prove the equivalent fact that $Y_2^\perp + Z^\perp$ is $w^*$ dense in $X^*$. But $Y^\perp \cap (x_n)^\perp + [x_n^*]$ is $w^*$ dense in $Y^\perp$ because $(x_n^*)$ is $w^*$-basic, so we need only show that each $x_{p(n,1)}^*$ is in the closure of $Y_2^\perp + Z$. To see that this last statement is true, write

$$\begin{aligned} x^*_{p(n,1)} &- a_1^{-1}[a_1x^*_{p(n,1)} - x^*_{p(n,2)}] - (a_1a_2)^{-1}[a_2x^*_{p(n,2)} - x^*_{p(n,3)}] - \cdots \\ &- (a_1a_2\cdots a_j)^{-1}[a_jx^*_{p(n,j)} - x^*_{p(n,j+1)}] \\ &= (a_1a_2\cdots a_j)^{-1}x^*_{p(n,j+1)} . \end{aligned}$$

Since $d(x^*_{p(n,j+1)}, Z) \leqq p(n, j+1)^{-1}a_1a_2\cdots a_{p(n,j+1)} \leqq (j+1)^{-1}a_1\cdots a_j$, we have $d(x^*_{p(n,1)}, Y_2^\perp + Z) \leqq (j+1)^{-1}$ for arbitrary $j$.

Next we prove the result of Lindenstrauss and Rosenthal.

*Proof of Theorem* 2. Since $X/Y$ has a separable quotient, there exists a biorthogonal sequence $(x_n, x_n^*)$ in $X$ with $(x_n^*) \subset Y^\perp$, $(x_n^*)$ $w^*$-basic, and normalized so that $||x_n|| = 1$. Since $Y^*$ is $w^*$-separable, a biorthogonalization argument (cf., e.g., [8] or [7]) shows that there exists a biorthogonal sequence $(y_n, y_n^*)$ for $Y$ with $(y_n^*) \subset X^*$, $Y \cap (y_n^*)^\tau = \{0\}$, and normalized so that $||y_n^*|| = 1$.

Define $T: X \to X$ by $Tx = \sum_{n=1}^\infty 2^{-n-1}y_n^*(x)x_n$. Then $||T|| \leqq 1/2$, so $I + T$ is an isomorphism. Hence $(I+T)^*$ is a weak*-isomorphism on $X^*$, whence $(x_n^* + T^*x_n^*)$ is a $w^*$-basic sequence $w^*$-equivalent to $(x_n^*)$.

Computing $T^*x_n^*$, we have $T^*x_n^*(x) = x_n^*Tx = x_n^*\sum_{m=1}^\infty 2^{-m-1}y_m^*(x)x_m = 2^{-n-1}y_n^*(x)$; i.e., $T^*x_n^* = 2^{-n-1}y_n^*$.

We claim that $(x_n^* + 2^{-n-1}y_n^*)^\tau$ is a quasi-complement to $Y$. First we show that $Y^\perp \cap \widetilde{[x_n^* + 2^{-n-1}y_n^*]} = \{0\}$ (so that $Y + (x_n^* + 2^{-n-1}y_n^*)^\tau$ is dense). But if $x^* \in \widetilde{[x_n^* + 2^{-n-1}y_n^*]}$, then, since $(x_n^* + 2^{-n-1}y_n^*)$ is $w^*$-equivalent to $(x_n^*)$, we can write $x^* = w^*\text{-}\lim_{n\to\infty}\sum_{i=1}^n \alpha_i x_i^* + \sum_{i=1}^\infty 2^{-i-1}\alpha_i y_i^*$ for some sequence $(\alpha_i)$ of scalars. Thus for each $n$, $x^*(y_n) = 2^{-n-1}\alpha_n$, hence, since $x^* \in Y^\perp$, $\alpha_n = 0$.

We complete the proof by showing that $Y \cap (x_n^* + 2^{-n-1}y_n^*)^\tau = \{0\}$. For suppose $y$ is in this intersection. Since $y \in Y$, $x_n^*(y) = 0$ for each $n$. Hence $y_n^*(y) = 0$ for each $n$, whence $y \in (y_n^*)^\tau \cap Y = \{0\}$.

THEOREM 3. *Suppose $X^*$ is separable and $Y$ is a subspace of $X^*$ with $\dim X^*/Y = \infty$. Then there exists a weak*-closed subspace $Z$ of $X^*$ with $Y, Z$ quasi-complements and $Z = [z_n]$ for some boundedly complete, $w^*$-basic sequence $(z_n)$.*

*Proof.* Mackey [8] showed that $Y$ has a quasi-complement, say, $W$. Let $(w_n, w_n^*)$ be a biorthogonal sequence in $W$ with $||w_n|| = 1$ and $[w_n] = W$ (cf. [9]). By Theorem III. 2 of [6], there exists a biorthogonal sequence $(x_n, x_n^*)$ in $X$ with $(x_n^*) \subset Y$, $(x_n^*)$ boundedly complete and $w^*$-basic, normalized so that $\|x_n\| = 1$.

Define $T: X \to X$ by $Tx = \sum_{n=1}^\infty 2^{-n-1} w_n(x) x_n$. Then $\|T\| \leqq 1/2$, so $I + T$ is an isomorphism and hence $(I + T)^*$ is a weak*-isomorphism. One checks that $T^* x_n^* = 2^{-n-1} w_n$, so that $(x_n^* + 2^{-n-1} w_n)$ is a $w^*$-basic sequence $w^*$-equivalent to $(x_n^*)$. Letting $Z = [x_n^* + 2^{-n-1} w_n]$, we have by Proposition 1 of [6] that $Z$ is weak*-closed.

Certainly $Z + Y \supset (w_n)$, so $Z + Y \supset Y + W$ and thus is dense. Suppose that $z \in Z \cap Y$. Then $z = \sum_{n=1}^\infty \alpha_n (x_n^* + 2^{-n-1} w_n)$ for some scalars $(\alpha_n)$ because $(x_n^* + 2^{-n-1} w_n)$ is basic. Hence also $\sum_{n=1}^\infty \alpha_n x_n^*$ converges, whence $z - \sum_{n=1}^\infty \alpha_n x_n^* = \sum_{n=1}^\infty \alpha_n 2^{-n-1} w_n$ is again in $Y$. Certainly $\sum_{n=1}^\infty \alpha_n 2^{-n-1} w_n$ is also in $W$ so that $\sum_{n=1}^\infty \alpha_n 2^{-n-1} w_n = 0$. Thus $\alpha_n 2^{-n-1} = w_n^*(\sum_{m=1}^\infty \alpha_m 2^{-m-1} w_m) = 0$, so that $z = 0$.

REMARK. Separability of $X^*$ in Theorem 3 is essential to get that $Z$ is weak*-closed. Indeed, regard $m = l_1^*$. Rosenthal [12] showed that $c_0$ is quasi-complemented in $m$. However, if $Z$ is a quasi-complement for $c_0$ in $m$, then $Z$ cannot be weak*-closed. For if $Z$ were $w^*$-closed, then $m/Z$ would be isomorphic to $(Z^\tau)^*$. But $m/Z$ is separable, hence reflexive (cf. [4]). Thus $Z^\tau$ would be a reflexive subspace of $l_1$, a contradiction (cf., e.g., [10]).

## REFERENCES


1. D. Amir and J. Lindenstrauss, *The structure of weakly compact sets in Banach spaces*, Ann. Math., (2), **88** (1968), 35-46.
2. C. Bessaga and A. Pelczynski, *On bases and unconditional convergence of series in Banach spaces*, Bull. Acad. Polon. Sci., **4** (1959), 379-383.
3. W. J. Davis and J. Lindenstrauss, *On total non-norming subspaces*, Proc. Amer. Math. Soc., **31** (1972), 109-111.
4. A. Grothendieck, *Sur les applications lineaires faiblement compactes d'espaces du-type C(K)*, Canad. J. Math., **5** (1953), 129-173.
5. R. C. James, *Quasi-complements*, J. Approximation Theory, **6** (1972), 147-160.
6. W. B. Johnson and H. P. Rosenthal, *On w*-basic sequences and their applications to the study of Banach spaces*, Studia Math., **43** (1972), 77-92.
7. V. Klee, *On the Borelian and projective types of linear subspaces*, Math. Scand., **6** (1958), 189-199.
8. G. Mackey, *Note on a theorem of Murray*, Bull. Amer. Math. Soc., **52** (1946), 322-325.
9. A. Markuschevich, *Sur les bases (au sens large) dans les espaces linéaires*, C. R. (Doklady) Acad. Sci. URSS, **41** (1943), 227-229.
10. A. Pelczynski, *Projections in certain Banach spaces*, Studia Math., **19** (1960), 209-228.
11. ———, *A note on the paper of I. Singer, 'Basic sequences and reflexivity of Banach spaces'*, Studia Math., **21** (1962), 371-374.
12. H. P. Rosenthal, *On quasi-complemented subspaces of Banach spaces, with an appendix on compactness of operators from $L^p(\mu)$ to $L^r(\nu)$*, J. Functional Analysis, **4** (1969), 176-214.

Received May 16, 1972 and in revised form April 10, 1973. The author was supported in part by NSF GP 28719. Most of the research for this paper was done at the University of Houston.

OHIO STATE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# ON 2-TRANSITIVE COLLINEATION GROUPS OF FINITE PROJECTIVE SPACES

WILLIAM M. KANTOR

**In 1961, A. Wagner proposed the problem of determining all the subgroups of $P\Gamma L(n, q)$ which are 2-transitive on the points of the projective space $PG(n-1, q)$, where $n \geqq 3$. The only known groups with this property are: those containing $PSL(n, q)$, and subgroups of $PSL(4, 2)$ isomorphic to $A_7$. It seems unlikely that there are others. Wagner proved that this is the case when $n \leqq 5$. In unpublished work, D. G. Higman handled the cases $n = 6, 7$. We will inch up to $n \leqq 9$. Our result is that nothing surprising happens. The same is true if $n = r^\alpha + 1$ for a prime divisor $r$ of $q - 1$.**

**One of Wagner's results is that it suffices to only consider subgroups of $PGL(n, q)$. Once this is done, it becomes simpler to view the problem as one concerning linear groups: find all those subgroups $G$ of $GL(n, q)$ which are 2-transitive on the 1-spaces of the underlying vector space $V$. Our approach is based primarily on three facts. (1) Wagner showed that the global stabilizer in $G$ of any 3-space of $V$ induces at least $SL(3, q)$ on that 3-space. (2) Unless $G \geqq SL(n, q)$ or $n = 4$, $q = 2$, and $G \approx A_7$, no nontrivial element of $G$ can fix every 1-space of some $n$-2-space of $V$. (3) $G \leqq SL(n, q)$ if $|G|$ is divisible by a prime which is a primitive divisor of $q^m - 1$ for a suitable $m \leqq n - 2$.**

Wagner's results are in [10]. Higman's result, and the case $n = 2^\alpha + 1$ and $q$ odd, are mentioned by Dembowski [1], p. 39. The result mentioned above in (2) is an easy consequence of results of Wagner. The idea used in (3) is due to Perin [8] and, independently, to G. Hare and E. Shult.

I am indebted to G. Seitz for several helpful remarks.

2. Notation and preliminaries. As already mentioned, we will be dealing with linear groups. Let $V$ be an $n$-dimensional vector space over $GF(q)$. We write $GL(V) = GL(n, q)$ and $SL(V) = SL(n, q)$. It will be convenient to regard everything as taking place in the relative holomorphic $V \cdot GL(V)$. For any subgroups $K, L$ of this semidirect product we can then consider the normalizer $N_L(K)$ and centralizer $C_L(K)$. If $L \leqq GL(V)$ and $W$ is an $L$-invariant subspace of $V$, we write $L^W = L/C_L(W)$ for the subgroup of $GL(W)$ induced by $L$. $C_L(V/W)$ and $L^{V/W}$ are defined similarly. For any group $G$, as usual $G'$ is its commutator subgroup, $Z(G)$ its center, and $\Phi(G)$ its Frattini subgroup.

A group $A$ is said to be *involved* in a group $B$ if $A \approx C/D$ with $B \geqq C \trianglerighteq D$.

(2.1) If $R \leqq GL(V)$ has prime power order and $(|R|, q) = 1$, then $V = C_V(R) \oplus [V, R]$, where $[V, R] = \langle v - vr \mid v \in V, r \in R\rangle$ is $N_{GL(V)}(R)$-invariant.

*Proof.* [3], p. 177.

(2.2) Let $R \leqq GL(V)$ have prime power order with $(|R|, q) = 1$. Let $W$ be an $R$-invariant subspace. Then $\dim C_V(R) = \dim C_W(R) + \dim C_{V/W}(R)$.

*Proof.* [3], p. 187, or (2.1).

Both (2.1) and (2.2) will be used frequently, generally without reference.

A *primitive divisor* of $q^k - 1$ is a prime $r$ satisfying $r \mid q^k - 1$ but $r \nmid q^i - 1$ for $1 \leqq i < k$; clearly $k \mid r - 1$.

(2.3) (i) If $q$ is a prime power and $k \geqq 2$, then $q^k - 1$ has a primitive divisor unless $k = 6$, $q = 2$, or $k = 2$ and $q$ is a Mersenne prime.

(ii) Let $r$ be a primitive divisor of $q^k - 1$, and let $R$ be an $r$-subgroup of $GL(V)$ for a $GF(q)$-space $V$. If $C_V(R) = 0$, then $k$ divides $\dim V$.

*Proof.* (i) [12].

(ii) This is clear if $|R| \leqq r$. Let $|R| > r$, and let $R_1 \leqq Z(R)$ have order $r$. Then $V = W \oplus [V, R_1]$, where $W = C_V(R_1)$ is $R$-invariant and $C_W(R) = 0$. By induction, $k$ divides $\dim W$ and $\dim [V, R_1]$.

(2.4) Suppose $\dim V = \alpha m$, $r$ is a primitive divisor of $q^m - 1$, and $R \leqq GL(V)$ is an $r$-group such that $C_V(R) = 0$. Then:

(i) Each noncyclic composition factor of $N = N_{GL(V)}(R)$ is involved in $PSL(\alpha, q^m)$; and

(ii) If $R$ is abelian, each noncyclic composition factor of $N/C_N(R)$ is involved in the symmetric group $S_\alpha$.

*Proof.* Write $V = W_1 \oplus \cdots \oplus W_\beta$, with each $W_i$ a sum of $R$-isomorphic irreducible $R$-spaces and no two $W_i$ having isomorphic irreducible $R$-subspaces. Set $R_i = C_R(W_i)$. Then $Z(R/R_i)$ is cyclic and nontrivial; let $Z_i$ be its subgroup of order $r$. By (2.3 ii), $\dim W_i = me_i$ for some $e_i$. Consequently, $\beta \leqq \alpha$ and $e_i \leqq \alpha$.

$N$ permutes the $W_i$. Let $K$ be the kernel of this permutation representation. Then $N/K$ is involved in $S_\beta \leqq S_\alpha$, and hence in $GL(\alpha, q^m)$.

Set $K_i = N_{GL(W_i)}(Z_i)$. Then $K$ is contained in $K_1 \times \cdots \times K_\beta$. Moreover, $K_i$ is contained in $\Gamma L(e_i, q^m)$. This proves (i).

Now assume that $R$ is abelian. Then $R/R_i$ is a cyclic group normalized by $K$. Since $\cap R_i = 1$, it follows that $K/C_K(R)$ is abelian. Since $N/K$ is involved in $S_{\alpha'}$ this proves (ii).

(2.5) Let $q$ be odd, and let $H \leqq GL(V)$. Suppose that $H \trianglerighteq A \neq 1$, where $A$ is an elementary abelian 2-group. Set

$$m = \min \{|H: N_H(B)| \mid B < A, |A: B| = 2\} .$$

Then $m \leqq \dim V$.

*Proof.* (G. Seitz.) Let $\bar{V}$ be an $H$-irreducible section of $V$ on which $A$ acts nontrivially. Let $\bar{H}$ and $\bar{A}$ be the groups induced by $H$ and $A$. Then $\bar{A} \neq 1$, and the corresponding $\bar{m} \geqq m$. We may thus assume that $V = \bar{V}$ is $H$-irreducible. By Clifford's Theorem ([3], p. 70), $V = V_1 \oplus \cdots \oplus V_t$ with the $V_i$ direct sums of $A$-isomorphic irreducible $A$-spaces, no two $V_i$ having a common irreducible constituent. Here $A$ induces a group of order 2 on each $V_i$, while $H$ is transitive on $\{V_1, \cdots, V_t\}$. Thus, $\{C_A(V_i) \mid i = 1, \cdots, t\}$ is an orbit of $H$ of subgroups of $A$ of index 2. Consequently, $t \geqq m$, so $\dim V \geqq m$.

(2.6) Let $L$ be a finite group and $K \lhd L$ with $L/K$ simple. Suppose $L$ has no proper subgroup $L_0$ for which $L_0/L_0 \cap K \approx L/K$. Then:

(i) $K$ is nilpotent; and

(ii) Each proper normal subgroup of $L$ is contained in $K$.

*Proof.* (i) Let $S$ be a Sylow subgroup of $K$. By the Frattini argument, $L = KN_L(S)$, so our conditions on $L$ imply that $L = N_L(S)$.

(ii) Let $M \trianglelefteq L$ and $M \nleqq K$. Since $1 \neq MK/K \trianglelefteq L/K$, $MK = L$ and hence $M = L$.

(2.7) Let $d > e \geqq 2$ and $t \geqq 1$. Then $PSL(d, q)$ is not involved in $PSL(e, q^t)$.

*Proof.* If $p$ is the prime dividing $q$, then $p$-Sylow subgroups of $PSL(d, q)$ and $PSL(e, q^t)$ have nilpotence class $d - 1$ and $e - 1$, respectively.

We now come to our main technical lemma.

(2.8) Let $q = p^e$, where $p$ is a prime, and $m = \dim V$. Suppose either $m = 3, 4$, or $5$, or $m = 6$ and $p = 2$. Let $L \leqq GL(V)$ and $H, K \lhd L$, where $H \leqq K$, $L/K \approx PSL(3, q)$, and $L/H \approx PSL(3, q)$ or $SL(3, q)$. Assume that $L$ has no proper subgroup $L_0$ for which $L_0/L_0 \cap K \approx PSL(3, q)$. Finally, assume: ($\sharp$) If $1 \neq h \in H$ and $p \nmid |h|$, then $\dim C_V(h) \leqq m - 3$.

Then there are $L$-invariant subspaces $X$, $Y$ with $X > Y$ such that the following hold.

(a) $K = P \times C$ with $P$ a $p$-group, $|C| = (3, q-1)$, and $H = P$ or $K$.

(b) $L/P \approx SL(3, q)$.

(c) $P^{V/X}$, $P^{X/Y}$ and $P^Y$ are all 1.

(d) $\dim X/Y = 3$ and $L^{X/Y} = SL(X/Y)$.

(e) If $m \leqq 5$ and $q \neq 2$, then $L^{V/X}$ and $L^Y$ are 1. Moreover, some element $g$ of order $p$ in the center of a $p$-Sylow subgroup of $L$ satisfies $\dim C_V(g) \geqq m-2$, and even $\dim C_V(g) = m-1$ if $P = 1$.

*Proof.* Everything is obvious if $m = 3$, so assume $m > 3$. We will proceed by a series of steps.

(i) Clearly $L = L'$. We can apply (2.6) to $L$. In particular, $K$ is nilpotent.

(ii) Suppose that there are $L$-invariant subspaces $V_1$, $V_2$ with $V_1 \geqq V_2$ and $\dim V_1/V_2 \leqq 2$. We claim that $L$ centralizes $V_1/V_2$. For, $C_L(V_1/V_2) \trianglelefteq L$, and since $L^{V_1/V_2}$ does not have $PSL(3, q)$ as a homomorphic image, (2.6) implies that $C_L(V_1/V_2) = L$.

(iii) Next, suppose that there are $L$-invariant subspaces $X$, $Y$ with $X > Y$, $\dim X/Y = 3$ and $L^{X/Y} \neq 1$. We claim that (a)—(e) hold.

Arguing as in (ii) we find that $L^{X/Y} = SL(X/Y)$, while $L^{V/X}$ and $L^Y$ are both 1 or $SL(3, q)$. Write $K = P \times C$ with $P$ a $p$-group and $C$ a $p'$-group. $C$ induces a group of order 1 or $(3, q-1)$ on $V/X$, $X/Y$, and $Y$. By (2.2), (a) holds unless $|C| = 9$ and $m = 6$. However, in this case $C \leqq Z(L)$, so $L/P = (L/P)'$ is a central extension of $SL(3, q)$ by a group of order 9, and this is impossible [2].

Thus, (a), (b), (c), and (d) hold.

Now let $m \leqq 5$. Then $\dim V/X$ and $\dim Y$ are $\leqq 2$, so $L^{V/X}$ and $L^Y$ are 1 by (ii). If $P \neq 1$ then, by (c), each $g \neq 1$ in $P$ satisfies $\dim C_V(g) \geqq m-2$.

Suppose $P = 1$, so $L \approx SL(3, q)$. By results of Higman [4], §5, if $q \neq 2$ then there is an $L$-invariant 3-space $T$, and each element of $L$ inducing a transvection on $T$ is a transvection of $V$. This proves (e).

(iv) From now on we assume that $m$ and $L$ are chosen with $m$ minimal such that (2.8) is false. Then $m > 3$.

$L$ is irreducible on $V$. For otherwise, there is an $L$-invariant subspace $W$ with $V > W > 0$.

Then $L^W \neq 1$ and $L^{V/W} \neq 1$. For suppose, say, that $L^{V/W} = 1$. Consider $L^W$, $K^W$, and $H^W$. By (2.2), (#) is inherited by $L^W$. Also, if $L_0 \leqq L$ and $L_0^W/L_0^W \cap K^W \approx PSL(3, q)$ then $L_0K/K \approx L_0/L_0 \cap K$ has $PSL(3, q)$ as a homomorphic image, so that $L_0K = L$ and hence $L_0 = L$.

Consequently, $L^W$ satisfies the hypotheses of (2.8). Then we can find subspaces $X$ and $Y$ of $W$ such that (iii) applies, whereas (2.8) is assumed false. Thus, $L^W \neq 1$ and $L^{V/W} \neq 1$.

By (ii) we must have $m = 6$ and $\dim W = 3$. Then (iii) again applies, and this is again impossible.

(v) By (iv) and the nilpotence of $K$, $(|K|, q) = 1$.

$K$ is not central in $L$. For suppose $K \leqq Z(L)$. Since $L = L'$, $L$ is a homomorphic image of the covering group of $PSL(3, q)$. Then $L$ is $PSL(3, q)$ or $SL(3, q)$ (see, e.g., [2]).

On the other hand, $L$ has an irreducible $GF(q)$-representation of degree $m$, where $4 \leqq m \leqq 6$ and $q$ is even if $m = 6$. No such representation exists by [7] and [9].

(vi) Let $r$ be a prime and $R_1$ an $r$-Sylow subgroup of $K$ such that $R_1 \nleqq Z(L)$. Set $R = R_1 \cap H$. Then $R \nleqq Z(L)$ and $R \lhd L$.

Let $A$ be a characteristic elementary abelian subgroup of $R$. By (#), $|A| \leqq r^{m-3}$.

We claim that $A \leqq Z(L)$. For otherwise, $L$ has a nontrivial $GF(r)$-representation of degree $\leqq m - 3 \leqq 3$. By (2.6 ii), $PSL(3, q)$ is involved in $GL(3, r)$. Thus, $q = 2$ and $r \neq 3$. Since $A$ is a noncyclic elementary abelian subgroup of $GL(6, 2)$, $|A| = 7^2$. Then $L$ acts transitively on $A - \{1\}$. However, not all elements of $A - \{1\}$ are conjugate in $GL(6.2)$.

Thus, $A \leqq Z(L)$. In (iv), $|A| = r$. In particular, $Z(R)$ is cyclic.

(vii) Suppose $r \nmid q - 1$. By (vi), $R \leqq GL(6, q)$ is nonabelian, so $r = 3 | q + 1$ and $m = 6$. Moreover, $R \rhd B$ with $|R: B| = 3$ and $B$ abelian. By (vi) we can find $B_1 \neq B$ with $R \rhd B_1$, $|R: B_1| = 3$, and $B_1$ abelian. Then $B \cap B_1 \leqq Z(R)$ and $|R/Z(R)| \leqq 9$. Consequently, $L$ centralizes $Z(R)$, $R/Z(R)$, and hence also $R$, which is not the case.

Thus, $r | q-1$. In (iv), $A \leqq L \cap Z(GL(V)) \leqq Z(SL(V))$, so $r | (q-1, m)$.

There are now just three possibilities: $m = 4, r = 2$; $m = 5, r = 5$; and $m = 6, r = 3$.

(viii) Let $m = 4, r = 2$. By (vii), $-1 \in R$. There is an involution $t \neq -1$ in $R$. Either $\dim C_V(t) \geqq 2$ or $\dim C_V(-t) \geqq 2$. This contradicts (#).

(ix) Let $m = 5, r = 5$. A 5-Sylow subgroup of $GL(5, q)$ has a normal abelian subgroup of index 5 (the "diagonal subgroup"). Thus, we can find $B \leqq R$ with $B$ abelian and $|R: B| = 1$ or 5. By (vi), $|R: B|$ is 5 and $B$ is not characteristic in $R$. Let $B_1 < R$, $B_1 \neq B$, satisfy the same conditions as $B$. Then $B_1 \cap B \leqq Z(R)$ and $|R: Z(R)| \leqq 5^2$. By (vi), $Z(R)$ is cyclic, so $L$ centralizes $Z(R)$, $R/Z(R)$, and hence also $R$, which is not the case.

(x) Finally, let $m = 6, r = 3$, and $q = 2^i$. Here $3 | q - 1$. On the one hand, $L/C_L(R/\Phi(R))$ can be regarded as a subgroup of $GL(e, 3)$ for some $e$; on the other hand, using (2.6) and $(|K|, q) = 1$, we find that this group has an elementary abelian 2-subgroup of order $q^2$ whose normalizer is transitive on the nontrivial elements. By (2.5), $e \geqq q^2 - 1$. However, a 3-Sylow subgroup of $SL(6, q)$ has order $\leqq 3(q-1)^6$. Thus, $3^{q^2-1} \leqq 3^e \leqq |R| < 3q^6$, and since $q \geqq 4$ this is ridiculous.

This contradiction completes the proof of (2.8).

3. **Wagner's results and some corollaries.** Let $V$ be $n$-dimensional over $GF(q)$, $n \geqq 3$, and let $G \leqq GL(V)$ be 2-transitive on 1-spaces.

(3.1) For each 3-space $T$, $N_G(T)^T \geqq SL(T)$.

*Proof.* Wagner [10], p. 417.

(3.2) If $n \leqq 5$ then $G \geqq SL(V)$, unless $n = 4$, $q = 2$, and $G \approx A_7$.

*Poof.* Wagner [10], p.422.

(3.3) For each $n$-1-space $W$, $N_G(W)$ is 2-transitive on the 1-spaces of $V$ not in $W$.

*Proof.* [6], p. 6.

(3.4) If $G$ has an element $g \neq 1$ such that $\dim C_V(g) \geqq n - 2$, then $G \geqq SL(V)$ or $n = 4$, $q = 2$, and $G \approx A_7$.

*Proof.* We may assume that $|g|$ is prime and $n > 5$. Since $\dim [V, g] \leqq 2$ and $g$ centralizes $V/[V, g]$, there is a 3-space $T > [V, g]$ such that $g^T \neq 1$. Then $1 \neq C_G(V/T)^T \trianglelefteq N_G(T)^T$. By (3.1), $C_G(V/T)^T \geqq SL(T)$. Choose $g' \in C_G(V/T)$ with $|g'| \mid q+1$ and $\dim C_T(g') = 1$. Then $\dim C_V(g') = n - 2$.

We may thus assume that $(|g|, q) = 1$. Since $g^{[V,g]} \neq 1$, as before $C_G(V/T)^T \geqq SL(T)$ for each 3-space $T > [V, g]$. By the 2-transitivity of $G$, this holds for every 3-space of $V$.

Choose $m \leqq n$ maximal with repect to $C_G(V/U)^U \geqq SL(U)$ for all $m$-spaces $U$. Suppose $m < n$. By Wagner [10], p. 420, $m \leqq n - 2$. Take any subspace $W$ of dimension $m + 1$ or $m + 2$. For each $m$-space $U < W$, $C_G(V/U)$ fixes $W$ and centralizes $V/W$, while $C_G(V/U)^U \geqq SL(U)$. By Wagner [10], p. 420, and (3.2), $C_G(V/W)^W \geqq SL(W)$ for each $m + 1$-space $W$. This contradicts the maximality of $m$.

(3.5) Let $s$ be a prime and $S$ an $s$-group maximal with respect to $\dim C_V(S) \geqq 3$. Then $N_G(S)$ is 2-transitive on the 1-spaces of $C_V(S)$.

*Proof.* Take any 3-space $T \leqq C_V(S)$. Then $S$ is Sylow in $C_G(T)$. By the Frattini argument and (3.1), $(N_G(S) \cap N_G(T))^T = N_G(T)^T \geqq SL(T)$. Our assertion follows immediately.

**4. The case $n = r^\alpha + 1$.** There is one very easy case of our problem.

(4.1) THEOREM. *Let $r$ be a prime divisor of $q - 1$, and let $\alpha \geqq 1$. Then every collineation group of $PG(r^\alpha, q)$ which is 2-transitive on points contains $PSL(r^\alpha + 1, q)$.*

We first prove:

(4.2) Let $r$ be a prime divisor of $q - 1$, and let $\alpha \geqq 1$. Let $V$ be an $r^\alpha$-dimensional vector space over $GF(q)$. If $G \leqq \Gamma L(V)$ is transitive on $V - \{0\}$, then $r \mid |G \cap Z(GL(V))|$.

*Proof.* Let $r^\beta$ be the largest power of $r$ dividing $q^d - 1$, where $d = r^\alpha$. Then $q$ is not an $r^\beta$th power, so $r \mid |G \cap GL(V)|$.

Let $R$ be an $r$-Sylow subgroup of $G$. By [11], p. 6, each orbit of $R$ on $V - \{0\}$ has length divisible by $r^\beta$.

$R$ fixes no nontrivial proper subspace of $V$. For, if it did we would have $r^\beta \mid q^m - 1$ with $1 \leqq m < d$. Set $e = (d, m)$. Then $r^\beta \mid q^e - 1$. However, as $d/e$ is a power of $r, (q^d - 1)/(q^e - 1)$ is divisible by $r$, and this contradicts the definition of $r^\beta$.

Let $x \in Z(R) \cap GL(V)$ have order $r$. Since $r \mid q - 1$, $x$ can be diagonalized. By the preceding paragraph, $x$ is a scalar transformation, that is, $x \in Z(GL(V))$.

(4.3) Let $r$ be a prime divisor of $q - 1$, and let $\alpha \geqq 1$. Then a collineation group of the affine space $AG(r^\alpha, q)$ which is 2-transitive on points contains the translation group.

*Proof.* (4.2).

Now (4.1) follows immediately from (3.3) and (4.3).

**5. Primes dividing $|G|$.** We will consider the following situation in the remainder of this paper.

*Let $V$ be an $n$-dimensional $GF(q)$-space, $n \geqq 6$, and $G$ be a subgroup of $GL(V)$, 2-transitive on 1-spaces, such that $G \ngeqq SL(V)$.* We may clearly assume that $G > Z = Z(GL(V))$.

In this section let $s$ be a prime dividing $(|G|, q^m - 1)$, $1 < m \leqq n - 2$, such that $s$ is a primitive divisor of $q^m - 1$. (5.1) is essentially due to Perin [8] and, independently, to E. Shult and G. Hare.

(5.1) If $m = n - 2$ then $q = 2$ and $n$ is even.

(5.2) Suppose that $n = \alpha m + \beta$, $\alpha < \beta \leqq m + 2$, and an element of order $s$ centralizes some 3-space $X$. Then, for some $n'$ satisfying $5 < n' < n$ and $n' \equiv n \pmod m$, there is a subgroup of $GL(n', q)$, not containing $SL(n', q)$, which is 2-transitive on the points of $PG(n' - 1, q)$.

Clearly (5.2) has an inductive flavor. Since the proofs are similar, we will only prove the second of the above results.

*Proof of* (5.2). Choose $S \leqq C_G(X)$ as in (3.5). Set $W = C_V(S)$, $W^* = [V, S]$, and $N = N_G(S)$. Then $V = W \oplus W^*$, $C_{W^*}(S) = 0$, and $N^W$ is 2-transitive on 1-spaces.

Set $n' = \dim W$, so $n' \geqq 3$. By (2.3 ii), since $\beta \leqq m + 2$ we have $\dim W^* = \gamma m$ with $\gamma \leqq \alpha$. Then $n' = n - \gamma m \geqq n - \alpha m = \beta > \alpha \geqq \gamma$.

We must show that $n' > 5$ and $N^W \ngeqq SL(W)$. Deny this. Then either $N^W \geqq SL(W)$ or $n' = 4$, $q = 2$, and $N^W \approx A_7$. In particular, the commutator subgroup $N'^W$ contains a nontrivial element centralizing an $n'$-2-space.

In this situation, $C_{N'}(W^*)^W \leqq Z(GL(W))$. For otherwise, $C_{N'}(W^*)^W \trianglelefteq N'^W$ implies that $C_{N'}(W^*)^W = N'^W$. Then $C_{N'}(W^*)$ has a nontrivial element $g$ centralizing an $n'$-2-space of $W$. Hence, $\dim C_V(g) \geqq n - 2$, which contradicts (3.4).

It follows that $N'^{W^*}$ has $PSL(n', q)$ as a homomorphic image, unless $n' = 4$ and $q = 2$, in which case $A_7$ may be a homomorphic image.

Since $C_{W^*}(S) = 0$, we can apply (2.4): each noncyclic composition factor of $N^{W^*}$ is involved in $PSL(\gamma, q^m)$. Since $n' > \gamma$, by (2.7) $PSL(n', q)$ cannot be such a composition factor. Thus, $n' = 4$, $q = 2$, $\gamma \leqq 3$, and $A_7$ is a composition factor of $N'^{W^*}$. However, $A_7$ is not involved in $PSL(3, 2^m)$. This is a contradiction.

REMARK. It is useful to note that the above proof holds under slightly weaker hypotheses: $s$ is a primitive divisor of $q^m - 1$, $S \neq 1$ is an $s$-subgroup of $G$ with $W = C_V(S)$ of dimension $n' \geqq 3$, $(n - n')/m < n'$, and $N_G(S)^W$ is 2-transitive on 1-spaces.

We conclude this section with two miscellaneous results.

(5.3) Assume that $G$ has a cyclic subgroup $H$ of order $q^n - 1$ containing an $r$-Sylow subgroup of $G$ for some prime $r$ dividing $q^2 + q + 1$. Then $q = 2$ and $n$ is even.

*Proof.* Suppose $q \neq 2$ or $q = 2$ and $n$ is odd. By (2.3), $H$ is transitive on $V - \{0\}$. Thus, $H$ is transitive on the 3-spaces fixed by its subgroup $R$ of order $r$.

On the other hand, by (3.1) each 3-space is fixed by a conjugate of $R$. Thus, $G$ is transitive on 3-spaces, and this contradicts Perin [8] or (5.1) since $n \geqq 6$.

(5.4) Assume that $G$ has a cyclic subgroup of order $q^{n-1} - 1$ fixing some $n - 1$-space $W$ and transitive on $W - \{0\}$. Then $N_G(W)$ is 2-transitive on the 1-spaces of $W$, $q = 2$, and $n$ is even.

*Proof.* We may assume that $G - Z$ has no element fixing all 1- spaces in $W$. By [6], Lemma 7.3, $N_G(W)$ is 2-transitive on the 1-spaces of $W$. The result now follows from (2.3) and (5.1).

6. The case $n \leqq 9$. Let $n$, $V$, $G$, and $Z$ be as in §5, so $G \ngeqq SL(V)$. Let $p$ be the prime dividing $q$.

Assume that $6 \leqq n \leqq 9$.

(6.1) $n \neq 6$.

*Proof.* Suppose $n = 6$. If $q = 2$ then $q^5 - 1$ is a prime. By (5.4), the stabilizer of a 5-space $W$ is 2-transitive on $W - \{0\}$. By (3.2) and (3.4), $G \geqq SL(V)$, which is not the case.

Thus, $q > 2$. Let $r$ be a prime dividing $q - 1$.

Suppose that there is 3-space $T$ for which $N_G(T) - Z$ contains an element inducing a scalar transformation of order $r$ on $T$. Using $Z$, we find that $r \mid |C_G(T)|$. Let $R$ be an $r$-Sylow subgroup of $C_G(T)$. By (3.4), $T = C_V(R)$. By (3.5), $N_G(R)^T \geqq SL(T)$. Also, $N_G(R)$ normalizes the 3-space $[V, R]$. An element of order $p$ in the center of a $p$-Sylow subgroup of $N_G(R)$ centralizes 2-spaces of both $C_V(R)$ and $[V, R]$, and hence centralizes a 4-space of $V = C_V(R) \oplus [V, R]$. This contradicts (3.4). Thus, no element of $G - Z$ of order $r$ has an eigenspace of dimension $> 2$.

Now take any 3-space $T$, and write $T = X \oplus Y$ with $\dim X = 2$ and $\dim Y = 1$. Set $F = N_G(X) \cap N_G(Y)$, so $F^X = GL(X)$. Take $R \leqq F$ of order $r$ with $R \nleqq Z$ and $R^T \leqq Z(F^T)$. By the Frattini argument, $N_F(R)^X = GL(X)$. Let $E \leqq N_F(R)$ be minimal with respect to $E^X = SL(X)$.

Since $R$ is diagonalizable and each of its eigenspaces has dimension 1 or 2, we can write $V = X \oplus W_1 \oplus W_2$ with $W_1 > Y$, $\dim W_i = 2$, and $W_i$ invariant under $N_G(R)$. If $q \neq 3$, $E = E'$ centralizes $W_1$, so an element of $E$ of order $p$ centralizes a 4-space, which contradicts (3.4). If $q = 3$, $R$ cannot have more than two eigenspaces as $|R| = 2$, which is again a contradiction.

(6.2) $q$ is even.

*Proof.* Assume that $q$ is odd. There is an involution $t \in G - Z$. Since $n \geqq 6$, $\dim C_V(t)$ or $\dim C_V(-t)$ is $\geqq 3$. Let $S$ be a 2-group in $G$ maximal with respect to $\dim C_V(S) \geqq 3$. Set $W = C_V(S)$ and $W^* = [V, S]$, so $V = W \oplus W^*$. Set $M = N_G(S)$. By (3.5), $M^W$ is 2-transitive on 1-spaces. Since $M > Z$ and all involutions in $M^W$ centralize at most a 2-space (by the maximality of $S$), $\dim W \leqq 4$. Consequently, by (3.2), $M^W \geqq SL(W)$.

By (4.1) and (6.1), $n = 7$ or 8, so $\dim W^* \leqq 5$.

We claim that $C_M(W^*)^W \leqq Z(GL(W))$. For otherwise, $C_M(W^*)^W \trianglelefteq M^W$ yields $C_M(W^*)^W \geqq SL(W)$. Then $C_M(W^*)$ contains a nontrivial transvection of $V$, which contradicts (3.4).

Thus, $C_M(W^*)$ is cyclic and $M'^{W^*}$ has $PSL(W)$ as a homomorphic image.

Suppose that $\dim W = 4$. Then $\dim W^* = 3$ or $4$. Use of $M'^{W^*}$ yields $\dim W^* = 4$ and $M'^{W^*} \geqq SL(W^*)$. If $g \neq 1$ is in the center of a $p$-Sylow subgroup of $M'$ then $g^W$ and $g^{W^*}$ are transvections, and this contradicts (3.4).

Thus, $\dim W = 3$. Let $L \leqq M$ be minimal with respect to having $PSL(3, q)$ as a homomorphic image. Let $H = C_L(W) \leqq K \lhd L$ with $L/K \approx PSL(3, q)$. Then (2.8) applies to $W^*$, $L^{W^*}$, $K^{W^*}$, and $H^{W^*}$.

Choose $g \in L$ so that $g^{W^*}$ is as in (2.8 e). If $g \in H = C_L(W)$, then $\dim C_V(g) \geqq n - 2$. If $H^{W^*} = 1$ then $H = 1$, and both $g^W$ and $g^{W^*}$ are transvections, so once again $\dim C_V(g) \geqq n - 2$. In either case we have contradicted (3.4).

(6.3) $n \neq 7, 8$.

*Proof*. Let $n = 7$ or $8$. Fix a prime $r \mid q + 1$.

Take any 3-space $T$. By (3.1), $N_G(T)^T \geqq SL(T)$. Also, $N_G(T)$ acts on $V/T$. By (3.4), $C_G(V/T)^T \leqq Z(GL(T))$ (since otherwise, $C_G(V/T)$ would have an element of order $r$), so $C_G(V/T)$ is solvable. Thus, $N_G(T)^{V/T}$ has $PSL(3, q)$ as a composition factor. By (2.8), there is an $r$-group $R \neq 1$ in $N_G(T)$ such that $\dim C_{V/T}(R) \geqq 2$, and then $\dim C_V(R) \geqq 3$.

This contradicts (5.2) with $n = 2 \cdot 2 + 3$ or $2 \cdot 2 + 4$.

(6.4) If $n = 9$ then $q = 2$ or $4$.

*Proof*. Suppose $n = 9$ and $q > 4$ is even.

(i) By (5.2) with $n = 2 \cdot 3 + 3$, no nontrivial element of order dividing $(q^2 + q + 1)/(q + 1, 3)$ can centralize a 1-space.

(ii) Let $T$ be any 3-space. Let $L \leqq N_G(T)$ be minimal with respect to having $PSL(3, q)$ as a homomorphic image. By (3.4), $C_G(V/T)^T \leqq Z(GL(T))$, so (2.8) applies to $L^{V/T}$. Consequently, by (i) there is a 6-space $Y > T$ such that $L^{Y/T} = SL(Y/T)$ and $L^{V/Y} = SL(V/Y)$.

(iii) Let $s$ be a prime dividing $q + 1$. By (ii), there is an element of order $s$ centralizing a 3-space.

Let $S$ be an $s$-group maximal with respect to $\dim C_V(S) \geqq 3$. By (3.5), $N_G(S)$ is 2-transitive on the 1-spaces of $C_V(S)$. In view of (i), it follows from (3.2), (6.1), and (6.3) that $\dim C_V(S) = 3$.

Let $T = C_V(S)$ in (ii), and choose $L \leqq N_G(S)$ there. By (i) and the proof of (2.4), $(LS)^{[V,S]}$ acts as a subgroup of $\Gamma L(3, q^2)$, with $S$ inducing scalar transformations.

(iv) Since $q > 4$, by (2.3 i) there is a prime $r \neq 3$ dividing $q - 1$. Moreover, if $q \neq 16$ we can choose $r \neq 5$.

We claim that some element of order $r$ centralizes a 4-space. For, since $r \neq 3$, in (iii) we can find $g \in L - Z$ of order $r$ such that $g^{[V,S]}$ has an eigenspace of dimension $\geqq 4$. Consequently, some element of $\langle g, Z \rangle$ of order $r$ centralizes a 4-space.

(v) Let $R$ be an $r$-group maximal with respect to $\dim C_V(R) \geqq 3$; by (iv), $R \neq 1$. Set $T = C_V(R)$ and $T^* = [V, R]$. By (3.5), $N_G(R)^T$ is 2-transitive on 1-spaces, so $\dim T = 3$ by (i). We can thus choose $L \leqq N_G(R)$ in (ii).

We claim that $LR$ centralizes $R$ and that $R$ is diagonalizable. Certainly $(LR)^{T^*} \leqq GL(T^*)$. Suppose $r > 5$. Then an $r$-Sylow subgroup of $GL(6, q)$ is diagonalizable, and hence abelian. By (2.4 ii) (with $m = 1, \alpha = 6$), each composition factor of $L/C_L(R)$ is involved in $S_6$. By (2.6 ii), $L = C_L(R)$, so $R \leqq Z(LR)$.

Consider the case $r = 5, q = 16$. Suppose $L > C_L(R)$. Then $L$ acts nontrivially on $R/\Phi(R)$, where $|R/\Phi(R)| \leqq 5^7$. By (2.6 ii), $16 + 1$ divides $|GL(7, 5)|$, which is not the case.

Thus, $L$ centralizes $R$. There is an $s$-group $S_0 < L$ such that $\dim C_{T^*}(S_0) = 2$. Since $R$ normalizes $C_{T^*}(S_0)$ and $[T^*, S_0]$, it follows that $R$ is again diagonalizable. Thus, $R \leqq Z(LR)$.

(vi) $T^*$ is the direct sum of $R$-invariant subspaces, each invariant under $LR$. By (ii) and (v), there are 3-spaces $X$ and $X'$ such that $T^* = X \oplus X'$, $R^X$ and $R^{X'}$ consist of scalar transformations, $L^X = SL(X')$, and $L^{X'} = SL(X')$.

Consequently, for each $h \in R$, $\dim C_V(h) = 3, 6$, or $9$.

(vii) By (iv), there is an $r$-group $R_1 \neq 1$ maximal with respect to $\dim C_V(R_1) \geqq 4$. By (vi), $W = C_V(R_1)$ has dimension 6. Set $M = N_G(R_1)$.

Take any 3-space $T < W$. Let $R \geqq R_1$ be an $r$-Sylow subgroup of $C_G(T)$. If $R = R_1$ then $N_M(T)^T \geqq SL(T)$ by the Frattini argument. If $R > R_1$ then the choice of $R_1$ implies that $C_V(R) = T$, and hence that $R$ is an $r$-group maximal with respect to $\dim C_V(R) \geqq 3$; by (v), $C_G(R)^T \geqq SL(T)$, so again $N_M(T)^T \geqq SL(T)$.

Consequently, $M^W$ is 2-transitive on 1-spaces. Then $(q^6-1)/(q-1)$ divides $|G|$, and this contradicts (5.2).

(6.5) If $n = 9$ then $q \neq 4$.

*Proof*. Suppose $n = 9$ and $q = 4$. We will try to imitate the proof of (6.4) using $r = 3$. Steps (i) and (ii) of that proof still hold.

We begin by showing the existence of $x \in G$ of order 3 such that $x^y = x^{-1}$ for some 2-element $y$. Take $T$ and $L$ as in (ii). Then we can find $x, y \in L$ with $|x| = 3$, $y$ a 2-element, and $x^y = x^{-1}a, a \in C_L(T)$.

By (2.8), $C_L(T) = P \times C$ with $P$ a 2-group and $|C| = 1$ or 3. Then $\langle x \rangle$ is Sylow in $\langle x, y \rangle P$. By the Frattini argument, some element of $\langle y \rangle P$ inverts $\langle x \rangle$, and we may assume this is $y$.

We next claim that some element of order 3 centralizes a 4-space. For, assume that this is false, and choose $x, y$ as above. Since $q = 4$, $x$ is diagonalizable and has at most 3 eigenspaces. However, no element of $\langle x, Z \rangle - \{1\}$ centralizes a 4-space, so $C_V(x) = T$ is a 3-space and $x$ has two other 3-dimensional eigenspaces $T_1$, $T_2$. Moreover, by our assumption, $C_G(T)$ has a cyclic 3-Sylow subgroup. Thus, by the Frattini argument, $N_G(\langle x \rangle)^T \geqq SL(T)$, so $C_G(x)^T \geqq SL(T)$. Since $|GL(T): SL(T)| = 3$, $y^T \in SL(T)$, so we can find $c \in C_G(X)$ such that $c^{-1}y \in C_G(T)$. Clearly $c^{-1}y$ inverts $x$, so there is an involution $t \in \langle c^{-1}y \rangle$. Here, $t$ centralizes $T$ and centralizes 2-spaces of each $T_i$, so $\dim C_V(t) \geqq 7$. This contradicts (3.4), and proves our claim.

Now define $R$, $T$, $T^*$, and $L$ as in (v). We will be able to obtain a contradiction precisely as in (vi) and (vii) if we can show that $R \leqq Z(LR)$ and $R$ is diagonalizable.

By (2.6), $L \rhd K$ with $L/K \approx PSL(3, 4)$ and $K$ nilpotent. By (2.2) and (2.8), $K = P \times C$ with $|C| = 3$ or 9 and $P$ a 2-group; moreover, there is an $L$-invariant 3-space $X < T^*$ such that $L^X = SL(X)$, $L^{T^*/X} = SL(T^*/X)$, and $P$ centralizes $T$, $X$, and $T^*/X$. By (3.4), no nontrivial element of $P$ centralizes a 4-space of $T^*$. Consequently, $P$ is elementary abelian of order $\leqq 4^3$. Thus, if $P \nleqq Z(L)$ then $PSL(3, 4)$ is isomorphic to a subgroup of $GL(6, 2)$, which is not the case ([7], [9]). Thus, $K \leqq Z(L)$.

Now suppose that $L$ acts nontrivially on $R$, and hence on $R/\Phi(R)$. Since $R \leqq GL(6, 4)$, $|R/\Phi(R)| \leqq 3^6 \cdot 3^2$. Thus, $PSL(3, 4)$ or $SL(3, 4)$ is isomorphic to a subgroup of $GL(8, 3)$. Then $GL(8, 3)$ has an elementary abelian subgroup of order $4^2$ whose normalizer is transitive on the nontrivial elements. By (2.5), this is impossible.

Consequently, $L \leqq C_G(R)$. An element of $L$ of order 5 centralizes 1-spaces of $X$ and $T^*/X$. It follows that $T^*$ is the sum of $R$-invariant 2-spaces. Thus, $R$ is diagonalizable and $R \leqq Z(LR)$. This completes the proof of (6.5).

Last, and least:

(6.6) If $n = 9$ then $q \neq 2$.

*Proof*. Suppose $n = 9$ and $q = 2$. Using (5.1) and (5.2) we find that $|G| = 2^\alpha \cdot 3^\beta \cdot 5 \cdot 7 \cdot 17 \cdot 73$ for some $\alpha$, $\beta$.

Let $S$ be a 73-Sylow subgroup of $G$. By (5.3), $|C_G(S)| = 73$. Thus, $|N_G(S)| = 3^\gamma \cdot 73$ with $\gamma \leqq 2$.

By Sylow's theorem, $2^\alpha \cdot 3^{\beta-\gamma} \cdot 5 \cdot 7 \cdot 17 \equiv 1 \pmod{73}$. A little arithmetic shows that this is impossible.

In view of (3.2) and the results of this section, we can now state:

THEOREM 6.7. *Let $H$ be a subgroup of $P\Gamma L(n, q)$ which is 2-transitive on the points of $PG(n-1, q)$. If $3 \leqq n \leqq 9$, then $H \geqq PSL(n, q)$ or $n = 4, q = 2$, and $H \approx A_7$.*

## REFERENCES

1. P. Dembowski, *Finite Geometries*, Springer, Berlin-Heidelberg-New York, 1968.
2. W. Feit, *The current situation in the theory of finite simple groups*, Proc. Int. Cong. Math. Nice, 1970, Vol. 1, 55-93.
3. D. Gorenstein, *Finite Groups*, Harper and Row, New York, 1968.
4. G. Hare, *Thesis*, Southern Illinois University, 1970.
5. D. G. Higman, *Flag-transitive collineation groups of finite projective spaces*, Illinois J. Math., **6** (1962), 434-446.
6. W. M. Kantor, 2-*Transitive symmetric designs*, Trans. Amer. Math. Soc., **146** (1969), 1-28.
7. C. Mark, *Thesis*, University of Toronto, 1939.
8. D. Perin, *On collineation groups of finite projective spaces*, Math. Z., **126** (1972), 135-142.
9. R. Steinberg, *Representations of algebraic groups*, Nagoya Math. J., **22** (1963), 33-56.
10. A. Wagner, *On collineation groups of finite projective spaces*, I. Math. Z., **76** (1961), 411-426.
11. H. Wielandt, *Finite Permutation Groups*, Academic Press, New York, 1964.
12. K. Zsigmondy, *Zur Theorie der Potenzreste*, Monatsh. f. Math. u. Phys., **3** (1892), 265-284.

Received November 6, 1971 and in revised form October 13, 1972. This research was supported in part by NSF Grant GP28420.

UNIVERSITY OF OREGON

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# COMPLETIONS AND CLASSICAL LOCALIZATIONS OF RIGHT NOETHERIAN RINGS

J. LAMBEK AND G. MICHLER

**Given a right Noetherian ring $R$ and a prime ideal $P$ of $R$, the injective hull of the right $R$-module $R/P$ is a finite power of a uniquely determined indecomposable injective $I_P$. One forms the ring of right quotients $R_P$ of $R$ relative to $I_P$ and the right ideal $M = PR_P$ of $R_P$ generated by $P$. The $M$-adic and $I_P$-adic topologies are compared; they turn out to coincide on every finitely generated $R_P$-module when $R_P$ is a classical quasi-local ring with maximal ideal $M$. This condition also implies that $R$ satisfies the right Ore condition with respect to the multiplicative set $\mathscr{C}(P)$ introduced by Goldie, that the $M$-adic completion $\hat{R}_P$ of $R_P$ is the bicommutator of $I_P$, and that $\hat{R}_P$ is an $n$ by $n$ matrix ring over a complete local ring.**

**Introduction.** If $P$ is a prime ideal of the commutative Noetherian ring $R$, then, by a theorem of Matlis [8], the completion $\hat{R}_P$ of the ring of quotients of $R$ at $P$ is the bicommutator of the injective hull of the $R$-module $R/P$. Recently Kuzmanovich [5] proved an analogous result for Noetherian Dedekind prime rings. Both these results are special cases of Theorem 6 below: Let $P$ be a two-sided prime ideal of the right Noetherian ring $R$, and assume that the ring of right quotients $R_P$ at $P$ is a *classical quasi-local* ring with maximal ideal $M = PR_P$, that is, $R_P/M$ is a simple Artinian ring and, for every right ideal $E$ of $R_P$, $\bigcap_{n=1}^{\infty} E + M^n = E$. Then the bicommutator of the $R$-injective hull of $R/P$ is the $M$-adic completion of $R_P$. The hypothesis of Theorem 6 is satisfied by the prime ideals of the enveloping algebra of a finitely generated nilpotent Lie algebra, by the augmentation ideal of a group ring of a finite group over a right Noetherian prime ring of characteristic zero, and by the nonidempotent prime ideals of a right and left Noetherian hereditary prime ring.

These results are consequences of Theorem 5, which states that $R_P$ is a classical quasi-local ring with maximal ideal $M$ if and only if $R_P/M$ is a simple Artinian ring, and, on any finitely generated $R_P$-module, the $M$-adic topology coincides with the $I_P$-adic topology. Here $I_P$ denotes the unique (up to isomorphism) $P$-torsionfree indecomposable injective $R$-module with associated prime $P$. By [7], Theorem 3.9, the injective hull $I_R(R/P)$ of the $R$-module $R/P$ is isomorphic to a direct sum of $g$ copies of $I_P$, where $g$ is the Goldie dimension of the prime ring $R/P$. Thus the bicommutators of $I_P$ and $I_R(R/P)$ are isomorphic.

Concerning terminology, we refer to [6], [7], and [8]. All rings are associative and have a unity element. Modules are right $R$-modules and unitary. We put

$$\mathscr{C}(P) = \{c \in R \mid \forall_{r \in R} cr \in P \Longrightarrow r \in P\} .$$

We begin by comparing topologies, generalizing the known result when $R$ is commutative [6].

PROPOSITION 1. *If $R$ satisfies the right Ore condition with respect to $\mathscr{C}(P)$, then on any finitely generated $R_P$-module the $I_P$-adic topology contains the $M$-adic topology, where $M = PR_P$.*

*Proof.* Let $G$ be any finitely generated $R_P$-module. Take any fundamental open neighborhood $GM^n$ of zero in the $M$-adic topology. We claim that $GM^n$ is also open in the $I_P$-adic topology, in fact, $G/GM^n \in \mathscr{F}$, the class of all $R_P$-modules isomorphic to submodules of finite powers of $I_P$.

Since $\mathscr{F}$ is closed under module extensions, and since

$$GM^n \subseteqq GM^{n-1} \subseteqq \cdots \subseteqq GM \subseteqq G ,$$

it suffices to show that $GM^k/GM^{k-1} \in \mathscr{F}$. Put $H = GM^k$, then $H$ is a finitely generated $R_P$-module. Now $R_P/M$ is a simple Artinian ring, by [7], Theorem 5.6. Hence $H/HM$ is isomorphic to a finite direct sum of minimal right ideals of $R_P/M$.

It remains to show that $R_P/M \in \mathscr{F}$. Indeed, in view of [7], Lemma 5.4, the mapping $R \xrightarrow{h} R_P \rightarrow R_P/M$ has kernel $P$, and so $R_P/M$ may be regarded as an $R$-module extension of $R/P$. Actually, it is an essential extension; for, if $0 \neq [q] \in R_P/M$, then $q \notin M$, but $qc \in h(R)$, for some $c \in \mathscr{C}(P)$, and $qc \notin h(P)$, since otherwise $q = qcc^{-1} \in h(P)R_P = PR_P = M$. Thus $R_P/M$ is isomorphic to an $R$-submodule of $I_R(R/P) = I_P^g$. By [7], Theorem 5.6, $R_P/M$ is torsionfree and divisible, hence $R_P/M$ is also isomorphic to an $R_P$-submodule of $I_P^g$, and so $R_P/M \in \mathscr{F}$.

This completes the proof. In the converse direction we have the following result. We remark that condition (1) plays an important rôle in [4], Theorem 5.3.

PROPOSITION 2. *Suppose $M = PR_P$ is a two-sided ideal of $R_P$. Then* (1) $\Rightarrow$ (2) $\Rightarrow$ (3):

(1) *For each right ideal $E$ of $R_P$ there exists a natural number $n$ such that $E \cap M^n \subseteqq EM$.*

(2) *For each element $i \in I_P$ there exists a natural number $n$ such that $iM^n = 0$.*

(3) *On any finitely generated $R_P$-module the $I_P$-adic topology is contained in the $M$-adic topology.*

*Proof.* Assume (1). Let $0 \neq i \in I_P$, and put $F = \{q \in R_P \mid iq = 0\}$, $E = \{q \in R_P \mid qM \subseteqq F\}$. Note that $EM \subseteqq F \subseteqq E$. Pick $n$ so that $E \cap M^n \subseteqq EM$, then $E \cap (M^n + F) = (E \cap M^n) + F = F$. Since $I_P$ is indecomposable, $F$ is meet-irreducible, hence $F = E$ or $M^n + F = F$. We shall prove that $F \neq E$, hence $M^n \subseteqq F$, and so (1) implies (2).

As $P$ is the associated prime ideal of the $R$-module $I_P$, $P$ is the right annihilator of some nonzero $R$-submodule $U$ of $I_P$. Putting $V = UR_P$, we see that $VM = 0$ and $V \neq 0$. Now $iR_P \cong R_P/F$, hence $0 \neq V \cap iR_P \cong G/F$, say, where $F \subseteqq G \subseteqq E, F \neq G$, hence $F \neq E$, as remained to be shown.

Assume (2), and let $G$ be a finitely generated $R_P$-module. Take any fundamental open neighborhood of zero in the $I_P$-adic topology. By definition, this has the form $\operatorname{Ker} f$, where $f: G \to I_p^n$ for some positive integer $n$. Let $p_k: I_p^n \to I_P$ be the canonical projections, with $k = 1, 2, \cdots, n$, and put $G_k = p_k(f(G))$. Then $G_k$ is a finitely generated $R_P$-submodule of $I_P$.

By assumption, there is a natural number $u(k)$ such that $G_k M^{u(k)} = 0$. Let $u = \operatorname{Max}\{u(1), \cdots, u(n)\}$, then $f(G)M^u = 0$, hence $\operatorname{Ker} f$ contains $GM^u$, a fundamental open neighborhood of 0 in the $M$-adic topology. It follows that every open set in the $I_P$-adic topology is also open in the $M$-adic topology. Thus $(2) \Rightarrow (3)$, and the proof is complete.

We know from [7], Lemma 5.2, that for each element $q \in R_P$ there exists an element $c \in \mathscr{C}(P)$ such that $qc \in h(R)$, where $h: R \to R_P$ canonically. This does not imply that $R_P$ is the classical ring of quotients of $h(R)$ with denominators in $h(\mathscr{C}(P))$, unless $R$ satisfies the right Ore condition with respect to $\mathscr{C}(P)$. (See [7], Proposition 5.5.) However, we have the following.

PROPOSITION 3. *Let $P$ be a two-sided prime ideal of the right Noetherian ring $R$, and assume that $M = PR_P$ is a maximal two-sided ideal of $R_P$ such that $R_P/M$ is Artinian. Then, for every integer $n \geqq 1$, $R_P/M^n$ is the classical ring of right quotients of $h(R)/(M^n \cap h(R))$, and its elements have the form $[h(r)][h(c)]^{-1}$, with $r \in R$ and $c \in \mathscr{C}(P)$.*

We could deduce this from [9], Theorem 2.4, by first proving that the ideals $h^{-1}(M^n)$ are the $n$th symbolic powers $P^{(n)}$ defined there in a different fashion. However, it is a bit quicker to deduce this directly from the following result by Small. (See [10], Theorem 1.)

Suppose $P$ is the prime radical of the right Noetherian ring $S$, and $\mathscr{C}$ is a multiplicatively closed subset of $S$ consisting of elements with zero right annihilators. Suppose the classical ring of right quotients of $S/P$ has elements of the form $[s][c]^{-1}$, with $s \in S$ and $c \in$

$\mathscr{C}$. Then $S$ satisfies the right Ore condition with respect to $\mathscr{C}$ and has a classical ring of right quotients with elements of the form $sc^{-1}$.

*Proof.* In [7], Theorem 5.6, in the proof of the implication (1) $\Rightarrow$ (2), it is shown that $R_P/M$ is the classical ring of right quotients of $R/P$, and that its elements have the form $[r][c]^{-1}$, where $r \in R$, and $c \in \mathscr{C}(P)$. Since $h^{-1}(M) = P$, by [7], Lemma 5.4, the result holds for $n = 1$.

To obtain the result for $n = 2$, we shall apply Small's Theorem to the ring $S = h(R)/(M^2 \cap h(R))$. To this purpose we must show that the elements of $h(\mathscr{C}(P))$ modulo $M^2$ have zero right annihilators. In fact, we shall see that they have left inverse in $R_P/M^2$.

Take any $c \in \mathscr{C}(P)$. In view of the case $n = 1$, we have $R_P = R_Pc + M$. Hence $M = MR_P = Mc + M^2$, and so $R_P = R_Pc + Mc + M^2 = R_Pc + M^2$. By Small's Theorem, $R_P/M^2$ is the classical ring of right quotients of $h(R)/(M^2 \cap h(R))$, and denominators may be taken in $\mathscr{C}(P)$ modulo $M^2$.

Repeating the same argument, we see that $R_P = R_Pc + M^3$, and that $R_P/M^3$ is the classical ring of right quotients of $h(R)/(M^3 \cap h(R))$, with denominators in $\mathscr{C}(P)$ modulo $M^3$. Etc, etc.

In accordance with [7], we call the ring $S$ a *classical quasi-local* ring if it is right Noetherian, it has a maximal ideal $M$ such that $S/M$ is Artinian, and every right ideal of $S$ is closed in the $M$-adic topology. In view of the following lemma, this implies that $M$ is the Jacobson radical of $S$.

LEMMA 4. *Suppose $M$ is a primitive ideal of the ring $S$, and every finitely generated right ideal of $S$ is closed in the $M$-adic topology. Then $M$ is the Jacobson radical of $S$.*

*Proof.* The first assumption assures that $M$ contains the radical. We claim the second assumption implies the converse. We shall prove that if $E$ is any right ideal of $S$ and $M + E = S$ then $E = S$.

Suppose $M + E = S$. Without loss in generality, we may take $E$ to be finitely generated. Now $M = SM = M^2 + EM$, hence $M^2 + E = M^2 + EM + E = M + E = S$. Similarly $M^3 + E = S$, and so on. Hence the $M$-adic closure $\bigcap_{n=1}^{\infty} (E + M^n)$ of $E$ is also $S$. By assumption, $E$ is closed, hence $E = S$.

THEOREM 5. *Let $P$ be a two-sided prime ideal of the right Noetherian ring $R$, and put $M = PR_P$, where $R_P$ is the ring of right quotients of $R$ at $P$. Then the following conditions are equivalent:*

( * ) *$R$ satisfies the right Ore condition with respect to $\mathscr{C}(P)$ and, for each right ideal $E$ of $R_P$, there exists a natural number $n$*

*such that* $E \cap M^n \subseteqq EM$.

( ** ) $R_P$ *is a classical quasi-local ring with maximal ideal* $M$.

( *** ) $M$ *is a two-sided ideal of* $R_P$, $R_P/M$ *is a simple Artinian ring, and on any finitely generated right* $R_P$*-module the* $I_P$*-adic and* $M$*-adic topologies coincide.*

(****) $M$ *is a two-sided ideal of* $R_P$, $R_P/M$ *is simple Artinian, and for each finitely generated right ideal* $E$ *of* $R_P$ *there exists a natural number* $n$ *such that* $E \cap M^n \subseteqq EM$.

*Proof.* We shall show that (*) $\Rightarrow$ (**) $\Rightarrow$ (***) $\Rightarrow$ (****) $\Rightarrow$ (*).

Assume (*). In view of [7], Theorem 5.6, (**) will follow if we show that every right ideal $F$ of $R_P$ is closed in the $M$-adic topology. Now its closure is given by $E = \bigcap_{n=1}^{\infty} (F + M^n)$. Pick $n$ so that $E \cap M^n \subseteqq EM$, then

$$E \subseteqq (F + M^n) \cap E = F + (M^n \cap E) \subseteqq F + EM .$$

Take any $e \in E$, then $e = f + \sum_{i=1}^{k} e_i m_i$, where $f \in F$, $e_i \in E$, and $m_i \in M$. Then $[e] = \sum_{i=1}^{k} [e_i] m_i$, modulo $F$, hence $E/F \subseteqq (E/F)M$.

It was pointed out in the discussion preceding [7], Theorem 5.6, that $R_P$ is right Noetherian. Thus $E$ and $E/F$ are finitely generated $R_P$-modules. We may therefore invoke Nakayama's Lemma and deduce that $E/F = 0$. Thus $F = E$, and so (**) holds.

Assume (**). By Lemma 4, $M$ is the Jacobson radical of $R_P$. By [7], Theorem 5.6, $R$ satisfies the right Ore condition with respect to $\mathscr{C}(P)$. Let $G$ be any finitely generated right $R_P$-module. Then, by Proposition 1, the $I_P$-adic topology on $G$ contains the $M$-adic topology. By Proposition 2 and [4] Theorem 5.3, the converse is true. Thus (***) holds.

Assume (***). Suppse $E$ is any finitely generated right ideal of $R_P$. Then $EM$ is an open subset of $E$ in the $M$-adic, hence in the $I_P$-adic topology. Now the $I_P$-adic topology on any module induces the $I_P$-adic topology on any submodule. Therefore, $EM = E \cap V$, where $V$ is an open subset of $R_P$ in the $I_P$-adic topology. Since $R_P$ is a finitely generated $R_P$-module, $V$ is an open set in the $M$-adic topology, hence $M^n \subseteqq V$ for some $n$, and so $E \cap M^n \subseteqq E \cap V = EM$. Thus (****) holds.

Assume (****). It remains to prove the right Ore condition. Given $a \in R$ and $c \in \mathscr{C}(P)$, we see from Proposition 3 that, for each positive integer $n$, there exist $a_n \in R$ and $c_n \in \mathscr{C}(P)$ such that $h(ac_n - ca_n) = h(u_n) \in M^n \cap h(R)$.

Let $F$ be the right ideal generated by the $u_n$, then $F = u_1 R + \cdots + u_k R$, since $R$ is right Noetherian. Taking $E = FR_P$ in the above, we see that $FR_P \cap M^n \subseteqq FM$, for some $n$. Hence $h(u_n) = u_1 m_1 + \cdots + u_k m_k$, where the $m_i \in M$.

Pick $d \in \mathscr{C}(P)$ so that all $m_i d \in h(R)$, then $m_i d \in M \cap h(R) = h(P)$, and we may write $m_i d = h(p_i)$, where $p_i \in P$.

Put $c' = c_n d - \sum_{i=1}^k c_i p_i$ and $a' = a_n d - \sum_{i=1}^k a_i p_i$, then an easy calculation shows that $h(ac') = h(ca')$. Moreover $C' \in \mathscr{C}(P)$, since $c_n d \in \mathscr{C}(P)$ and $\sum c_i p_i \in P$. Since $ac' - ca' \in \operatorname{Ker} h$, we can find $d' \in \mathscr{C}(P)$ so that $(ac' - ca')' = 0$, hence $a(c'd') = c(a'd')$. This establishes the right Ore condition for $R$, and our proof is complete.

THEOREM 6. *Let $P$ be a two-sided prime ideal of the right Noetherian ring $R$ such that $R_P$ is a classical quasi-local ring with maximal ideal $M = PR_P$. Then*

(a) *the $M$-adic completion $\hat{R}_P$ of $R_P$ is the bicommutator of the $P$-torsionfree indecomposable injective $R$-module $I_P$ with associated prime $P$,*

(b) *$\hat{R}_P$ is a ring of $n \times n$ matrices over a complete local ring $\hat{D}$ whose Jacobson radical $J$ is finitely generated.*

*Proof.* (a) By Theorem 5, $R$ satisfies the right Ore condition with respect to $\mathscr{C}(P)$. By [7], Theorem 5.6, every torsionfree $R$-module is $P$-divisible. In view of [6], Proposition 2, $R_P$ is therefore a dense subring of the bicommutator $S$ of $I_P$ with respect to the finite topology, as the $P$-torsion theory coincides with that determined by $I_P$, by [7], Corollary 3.10. By [6], Corollary 1, the finite topology coincides with the $I_P$-adic topology on $R_P$, and $S$ is the completion of $R_P$. By Theorem 5, the $I_P$-adic topology on $R_P$ coincides with the $M$-adic one. Therefore $S$ is the $M$-adic completion of $R_P$.

(b) follows immediately from (a) and [9], Corollary 2.7.

REMARK 7. By [9], Remark 3, there exists a right Noetherian ring $R$ with a two-sided prime ideal $P$ such that $R$ satisfies the right Ore condition with respect to $\mathscr{C}(P)$, even though $R_P$ is not a classical quasi-local ring with maximal ideal $M$. In that example $R_P$ is not Hausdorff with respect to the $M$-adic topology, hence the bicommutator of $I_P$ is not the $M$-adic completion of $R_P$.

Thus the right Ore condition does not imply the second part of (*) in Theorem 5. Conversely, Example 5.9 of [7] shows that the second part of (*) does not imply the right Ore condition.

We conclude by giving some classes of examples satisfying the condition of Theorem 5. But first we note that, in view of Theorem 3.3 of [9], each of these is also equivalent to the following, which involves only the ring $R$ itself:

(+) For every right ideal $F$ of $R$ there exists a positive integer $n$ such that $F \cap P^{(n)} \subseteqq \mathrm{cl}_P(FP)$, where $P^{(n)}$ is the $n$th right symbolic power of $P$.

For notation see [9].

COROLLARY 8. *Let $R$ be the enveloping algebra of a finitely generated nilpotent Lie algebra, and assume that $P$ is a nonzero prime ideal of $R$. Then the conclusions* (a) *and* (b) *of Theorem* 6 *hold.*

*Proof.* In Theorem 2.6 of [9], it is shown that, if $R$ is right and left Noetherian, $P^{(n)}$ coincides with the symbolic $n$th power defined by Goldie in [4]. To deduce (+), we therefore refer to [2], namely to Theorem 6, Corollary 7 and Remark I.

COROLLARY 9. *Let $R = AG$ be the group ring of a finite group $G$ over a right Noetherian prime ring $A$ of characteristic zero, and let $P$ be the augmentation ideal of $R$. Then the conclusions* (a) *and* (b) *of Theorem* 6 *hold.*

*Proof.* Condition (+) holds by Corollary 3.7 of [9].

Actually, in this example $R_P$ is the classical ring of right quotients of $R$, and $M = PR_P = 0$, because $P$ is the $P$-torsion ideal of $R$.

COROLLARY 10. *Let $R$ be a right and left Noetherian hereditary prime ring, and assume that $P$ is not idempotent. Then the conclusions* (a) *and* (b) *of Theorem* 6 *hold. Furthermore, $\hat{D}$ is a complete discrete rank one valuation ring.*

*Proof.* By [11], $R/P$ is a simple Artinian ring. It is known that $P$ is an invertible ideal. By Lemma 1.1 of [3], it then follows that $P$ has the Artin-Rees property. Now, by Corollary 2.8 of [9], $P^n = P^{(n)}$, hence condition (+) holds.

It remains to show that $\hat{D}$ is a rank one valuation ring. By the remark preceding Theorem 5.6 in [7], $R_P$ is hereditary Noetherian and quasi-local. As is well-known, this implies that $\hat{R}_P$ is hereditary Noetherian. By Morita equivalence, $\hat{D}$ is hereditary Noetherian. But it is local, hence a discrete rank one valuation ring.

For the sake of completeness, we shall show that $P$ is invertible. Let $Q$ be the maximal ring of right and left quotients of $R$ and put $R.\cdot P = \{q \in Q \mid qP \subseteqq R\}$. It is known [3] that $R \subseteqq P(R.\cdot P)$ provided $P$ is finitely generated and projective as a right $R$-module and "dense" in a technical sense, which means that $P$ has zero left annihilator in $R$ when $P$ is a two-sided ideal. Since $R$ is right Noetherian, right hereditary, and prime, $P$ satisfies all three conditions.

Now $P \subseteqq (R.\cdot P)P \subseteqq R$, and $P$ is maximal. Therefore $(R.\cdot P)P = P$ or $R$. Suppose the former, then $P \subseteqq P(R.\cdot P)P = P^2$, which would lead to the contradiction that $P$ is idempotent. Therefore, $(R.\cdot P)P = R$.

Finally, consider $P\cdot .R = \{q \in Q \mid Pq \subseteqq R\}$. Then

$$P\cdot .R = (R.\cdot P)P(P\cdot .R) \subseteqq R.\cdot P\,.$$

By symmetry we obtain $P(P\cdot .R) = R$ and $R.\cdot P \subseteqq P\cdot .R$, and so $P$ is invertible in $Q$.

For the sake of completeness, we shall also include the argument of [1] to show that $P$ has the Artin-Rees property. Let $E$ be any right ideal of $R$ and put $E_k = (E \cap P^k)P^{-k}$. Since $R$ is right Noetherian, there exists a positive integer $k$ such that $E_k \subseteqq E_1 + \cdots + E_{k-1}$. Then $E \cap P^k = E_k P^k \subseteqq \sum_{i<k} (E \cap P^i)P^{k-i} \subseteqq EP$, and this is the Artin-Rees property.

## References

1. A. W. Chatters and S. M. Ginn, *Localization in hereditary rings*, J. Algebra, **22** (1972), 82–88.

2. J. C. Connell, *Localization in enveloping rings*, J. London Math. Soc., **43** (1968), 421–428.

3. G. D. Findlay, *Multiplicative ideal theory and rings of quotients, I*, Proc. Roy. Soc. Edinburgh, **68** (1968), 30–53.

4. A. W. Goldie, *Localization in non-commutative Noetherian rings*, J. Algebra, **5** (1967), 89–105.

5. J. Kuzmanovich, *Completions of Dedekind prime rings and second endomorphism rings*, Pacific J. Math., **36** (1971), 721–729.

6. J. Lambek, *Bicommutators of nice injectives*, J. Algebra, **21** (1972), 60–73.

7. J. Lambek and G. Michler, *The torsion theory at a prime ideal of a right Noetherian ring*, J. Algebra, **25** (1973), 364–389.

8. E. Matlis, *Injective modules over Noetherian rings*, Pacific J. Math., **8** (1958), 511–528.

9. G. Michler, *Right symbolic powers and classical localization in right Noetherian rings*, Math. Z., **127** (1972), 57–69.

10. L. Small, *Orders in Artinian rings II*, J. Algebra, **9** (1968), 266–273.

11. D. B. Webber, *Ideals and modules of simple Noetherian hereditary rings*, J. Algebra, **16** (1970), 239–242.

Received June 7, 1972.

McGill University
and
Universität Giessen

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# BOREL SETS OF PROBABILITY MEASURES

KENNETH LANGE

**Let $M(X)$ be the collection of probability measures on the Borel sets of a Polish space $X$. The Borel structure of $M(X)$ generated by the weak* topology is investigated. Various collections of probability measures arising in nonparametric statistics are shown to Borel sets of $M(X)$. Attention is particularly focused on collections arising from restrictions on distribution functions, density functions, and supports of the underlying probability measures.**

1. Introduction. Baynesian statisticians assume prior distributions on certain families of probability measures. This amounts to putting a probability measure on a family of probability measures. Now families of probability measures typically arising in statistics are parametrized by some Borel set of Euclidean $n$-space. In such cases, one has a natural Borel structure or $\sigma$-algebra of subsets with which to deal. In nonparametric situations the natural Borel structure is not so obvious. Ideally, one might desire each commonly occurring family of probability measures to be a Borel set of some properly chosen complete separable metric space. Then a prior distribution could be viewed as a probability measure on the entire space which is concentrated on the given Borel set. Our aim is to show that many, if not most, nonparametric families of probability measures are indeed Borel sets of complete separable metric spaces. This advances slightly the cause of nonparametric Baynesian statistics, but does not overcome the more difficult barrier of finding reasonable prior distributions in nonparametric situations.

In our probabilistic model we suppose $X$ to be a complete separable metric space. Let $C(X)$ be the bounded real-valued continuous functions on $X$ under the sup norm topology. Then the collection of probability measures $M(X)$ on the Borel sets of $X$ can be viewed as a subset of the dual of $C(X)$ under the weak* topology. It is well known that $M(X)$ is metrizable as a complete separable metric space with this topology [16]. Our investigations will center on the Borel structure of $M(X)$.

Dubins and Freedman have done the spadework for the subsequent discussion in their basic paper [8]. Section 3 generalizes their analysis of the relationship of distribution and density functions to probability measures. Section 4 explores the connection between a probability measure and its support when the underlying space is no longer compact. Section 5 collects some further examples not considered in

[8], and §6 adds to the fund of counterexamples.

Finally, before moving on to some preliminary definitions, let us cite two other areas where the Borel structure of collections of measures can be fruitfully pursued. Much work has been done on so called ergodic decompositions of invariant measures. The original stimulus for this research came from classical statistical mechanics. The reader may consult [20] for a detailed theoretical discussion and further references. Another area of potential applications is the analysis of Poisson and point processes. See [14] for steps in this direction.

## 2. Preliminary definitions.

2.1. *Borel spaces.* First, let us give a compressed account of Borel spaces. The reader is advised to consult §§1–3 of Chap. 1 of [2] for a fuller treatment. A *Borel space* $(X, \mathscr{A})$ consists of a set $X$ together with a distinguished $\sigma$-algebra of subsets $\mathscr{A}$. Quite often $X$ itself is said to be the Borel space and $\mathscr{A}$ is tacitly understood. For example, if $X$ is a topological space, then $\mathscr{A}$ is always taken to be the smallest $\sigma$-algebra containing the open sets. A function $f\colon X \to Y$ between two Borel spaces $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ is called *Borel* if $f^{-1}(B) \in \mathscr{A}$ whenever $B \in \mathscr{B}$.

The sets in the $\sigma$-algebra $\mathscr{A}$ of a Borel space $(X, \mathscr{A})$ are also termed *Borel sets.* Every subset $Z$ of $X$ inherits a *relative Borel structure* $\{Z \cap A\colon A \in \mathscr{A}\}$. If $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ are two Borel spaces, the *product Borel structure* $\mathscr{A} \times \mathscr{B}$ is defined to be the smallest $\sigma$-algebra of subsets of $X \times Y$ containing the Borel rectangles $A \times B$, $A \in \mathscr{A}$, and $B \in \mathscr{B}$. Suppose $\sim$ is an equivalence relation on a Borel space $(X, \mathscr{A})$ and $\pi\colon X \to X/\!\sim$ is the projection taking each point into its equivalence class. The *quotient Borel structure* on $X/\!\sim$ is the largest $\sigma$-algebra making $\pi$ a Borel map.

Certain Hausdorff topological spaces have very well behaved Borel structures. Among these are *Polish spaces.* A Polish space is a topological space which is metrizable by a complete separable metric. It is well known that any locally compact space with a countable neighborhood basis is Polish. Such spaces will be referred to as locally compact and separable.

One property of Polish spaces will be particularly useful later on: Suppose $X$ and $Y$ are Polish spaces. Let $B$ be a Borel set of $X$ equipped with the relative Borel structure. Also let $f\colon B \to Y$ be a one-to-one Borel map. Then $f(B)$ is a Borel set of $Y$. (See Cor. 3.3 of Chap. 1 of [16].) This fact will be applied to show that certain Borel maps are *Borel isomorphisms.* A map $f\colon X \to Y$ between two Borel spaces $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ is said to be a Borel isomorphism if it is one-to-one, onto, Borel and its inverse is Borel.

Yet another class of Borel spaces appearing in the sequel is the class of *analytic Borel spaces*. To define this notion it is necessary to mention a second notion. A Borel space $(X, \mathscr{A})$ is called *countably separated* if there exists a countable collection of Borel sets $\{A_n\}_{n=1}^{\infty}$ such that for any two points $u, v \in X$ there is some $A_n$ with either $u \in A_n$, $v \notin A_n$ or $u \notin A_n$, $v \in A_n$. A Borel space $(X, \mathscr{A})$ is analytic if it is countably separated and the image of a Polish space under a Borel map. A subset of a Borel space is called analytic if it is an analytic Borel space with its relative Borel structure; it is called *complementary analytic* if its complement is analytic. Every Borel set of an analytic Borel space is analytic, but not every analytic set is Borel. If $f: X \to Y$ is a Borel map between two analytic Borel spaces, then the image and inverse image of every analytic set under $f$ is analytic.

2.2. *Notation for topological spaces.* Suppose $T$ is a topological space and $Y$ is a subset of $T$. $Y^-$ will denote the closure of $Y$, $Y^0$ the interior, and $Y'$ the complement. For two subsets $Y$ and $Z$ put $Y \Delta Z = (Y \backslash Z) \cup (Z \backslash Y)$, the symmetric difference of $Y$ and $Z$. $R^n$ will mean Euclidean $n$-space and $C^n$ the space of $n$-tuples of complex numbers. If $u = (u_1, \cdots, u_n)$ and $v = (v_1, \cdots, v_n)$ are in $R^n$, then $\langle u, v \rangle = \sum_{i=1}^n u_i v_i$ will denote the usual inner product and $\|u\| = (\sum_{i=1}^n u_i^2)^{1/2}$ the usual norm.

2.3. *Multi-index notation.* A multi-index $j = (j_1, \cdots, j_n)$ is a finite sequence of nonnegative integers. $|j|$ denotes the sum $\sum_{i=1}^n j_i$. If $x = (x_1, \cdots, x_n) \in R^n$, put $x^j = x_1^{j_1} \cdots x_n^{j_n}$. For a function $f: R^n \to C$ continuously differentiable of order $|j|$, set $D_i = \partial/\partial x_i$ and $D^j f = D_1^{j_1} \cdots D_n^{j_n} f$.

2.4. *Comments on probability measures.* Suppose $X$ is a Polish space. The *support* of a probability measure $\mu \in M(X)$ can be characterized as the complement of the largest open set on which $\mu$ vanishes. If $\nu$ is a $\sigma$-finite measure on the Borel sets of $X$ and $\mu$ is absolutely continuous with respect to $\nu$, then the *density* of $\mu$ will mean the Radon-Nikodym derivative $d\mu/d\nu$. On $R^n$ densities will always refer to Radon-Nikodym derivatives with respect to Lebesgue measure. The *distribution function* $F_\mu: R^n \to [0, 1]$ of $\mu \in M(R^n)$ is defined by $F_\mu(x_1, \cdots, x_n) = \mu((-\infty, x_1] \times \cdots \times (-\infty, x_n])$.

Finally, for $X$ Polish we should mention an alternate description of the Borel structure of the collection of probability measures $M(X)$. Varadrajan proved that the Borel structure on $M(X)$ generated by the weak* topology is precisely the smallest Borel structure making each of the maps $\mu \to \mu(A)$ Borel, where $\mu \in M(X)$ and $A$ is a Borel set of $X$. (See Lemma 2.3 of [20].)

**3. Borel properties of $M(R^n)$ in terms of distribution and density functions.** Let us treat the case of distribution functions first. Our opening lemma is a generalization of a well known result in the theory of stochastic processes. (See Thm. T 47 of Chap. 4 of [12].)

LEMMA 3.1. *Suppose $M_1$ is a Borel space and $M_2$ a metric space. Consider a function $f(x_1, \cdots, x_n, y)$ from $R^n \times M_1$ into $M_2$. If $f$ is Borel in $y$ for $x_1, \cdots, x_n$ fixed and right continuous in each $x_i$ for all other variables fixed, then $f$ is Borel in all variables jointly. The same conclusion holds if $R^n$ is replaced by a product of intervals.*

*Proof.* The general case follows from the one dimensional case since one can replace $M_1$ by $(\prod_{i=1}^{n-1} U_i) \times M_1$, where the $U_i$'s are intervals. So consider $f: U \times M_1 \to M_2$ satisfying the condition of the lemma, where $U$ is an interval. For each $k = 1, 2, \cdots$, choose pairwise disjoint intervals $F_k^n (n = 1, \cdots, n_k, n_k$ finite or $\infty)$ which are closed on the right, have union $U$ and satisfy length $(F_k^n) < 1/k$. For each $k$ and $n$ let $x_k^n$ be the right endpoint of $F_k^n$. Now define $f_k(x, y) = f(x_k^n, y)$ whenever $x \in F_k^n$. Since $f$ is continuous on the right in its first variable, $f_k \to f$ pointwise on $U \times M_1$. Since $f$ is Borel in its second variable, each $f_k$ is Borel on $U \times M_1$. Finally, because $M_2$ is a metric space, the limit $f$ of the Borel functions $f_k$ is Borel.

DEFINITIONS. For $0 \leqq k \leqq \infty$ let $C^k(R^n)$ be the topological vector space of all complex-valued functions on $R^n$ having continuous derivatives of order $k$ under the topology of uniform convergence on compact sets. Note that $f_m \to f$ in $C^k(R^n)$ means whenever $j$ is a multi-index with $|j| \leqq k$, $D^j f_m \to D^j f$ uniformly on each compact set of $R^n$. $C^k(R^n)$ is a separable Frechet space since the polynomials with rational coefficients form a countable dense subset [19]. Lip $(R^n)$ will denote the collection of complex-valued functions $f$ on $R^n$ satisfying $|f(x) - f(y)| \leqq c_f \|x - y\|$ for all $x$ and $y$ and some constant $c_f$ depending only on $f$. Finally, let $H(C^n)$ be the topological vector space of holomorphic functions in $n$ complex variables equipped with the topology of uniform convergence on compact sets. $H(C^n)$ is also a separable Frechet space.

LEMMA 3.2. *Each of the spaces $C^k(R^n)$, $1 \leqq k \leqq \infty$, and* Lip $(R^n)$ *is a Borel subset of $C^0(R^n)$. If $H(R^n)$ denotes the holomorphic functions in $n$ real variables, then $H(R^n)$ is a Borel set of $C^0(R^n)$ too.*

*Proof.* The natural injection of $C^k(R^n)$, $1 \leqq k \leqq \infty$, into $C^0(R^n)$ is continuous. Hence its image in $C^0(R^n)$ is a Borel set. Lip $(R^n)$ is a Borel set of $C^0(R^n)$ because the map

$$f \longrightarrow \sup_{x \neq y} \frac{|f(x) - f(y)|}{||x - y||}$$

is lower semicontinuous in the classical sense (not to be confused with the notion of lower semicontinuity for set valued maps discussed in the next section), hence Borel. Last of all, the map $H(C^n) \to C^0(R^n)$ given by restricting a holomorphic function on $C^n$ to $R^n$ is certainly continuous. It is one-to-one because a holomorphic function on $C^n$ is completely determined by its values on $R^n$. (See 9.4.4 of [6].)

THEOREM 3.3. *For each probability measure $\mu$ on $R^n$ let $F_\mu$ be the distribution function of $\mu$. Then the set of probability measure $\mu$ with $F_\mu$ satisfying any of the conditions below forms a Borel set of $M(R^n)$:*

1. *$F_\mu$ is continuous.*
2. *$F_\mu$ is Lipschitz, i.e., $|F_\mu(x) - F_\mu(y)| \leqq c||x - y||$ for some constant $c$ and all $x$ and $y$.*
3. *$F_\mu$ is continuously differentiable of order $k, 1 \leqq k \leqq \infty$.*
4. *$F_\mu$ is holomorphic.*

*Proof.* Let us prove the assertion in part 1. first. Designate by $\{r_k\}_{k=1}^{\infty}$ the points in $R^n$ having all components rational. Since $\mu \to \mu(A)$ is Borel from $M(R^n)$ to $R$ for every Borel set $A$ in $R^n$, the function

$$\mu \longrightarrow \inf_{\substack{n \in Z \\ n>0}} \sup_{||r_k - r_l|| < 1/n} |F_\mu(r_k) - F_\mu(r_l)| = g(\mu)$$

is Borel. Our claim is that $\{\mu: g(\mu) = 0\}$ is the collection of all probability measures with continuous distribution functions. In fact, if $x = (x_1, \cdots, x_n)$ and $y = (y_1, \cdots, y_n)$ are two points in $R^n$, the expansion

$$F_\mu(x) - F_\mu(y) = \sum_{k=1}^{n} (F_\mu(z^{k-1}) - F_\mu(z^k)),\ z_i^k = \begin{cases} x_i \text{ if } k < i \\ y_i \text{ if } k \geqq i \end{cases},$$

leads immediately to the conclusion that $g(\mu) = 0$ implies $F_\mu$ is continuous. On the other hand, if $F_\mu$ is continuous, it is actually uniformly continuous. This follows from the fact that for every $\varepsilon > 0$ there is a compact set $K$ with $\mu(K) > 1 - \varepsilon$. The uniform continuity of $F_\mu$ then clearly entails $g(\mu) = 0$.

To prove the assertion of the theorem for parts 2.—4. it is sufficient by Lemma 3.2 to prove that the map $\mu \to F_\mu$ is Borel from the collection of probability measures having continuous distribution functions into $C^0(R^n)$. To do this it is enough according to Thm. 2 of the appendix of [13] to show $\mu \to l(F_\mu)$ Borel whenever $l$ is a continuous linear functional on $C^0(R^n)$. Since the dual space of $C^0(R^n)$ consists of the complex measures $\nu$ with compact support (13.19.3 of [7]), the problem reduces to showing $\mu \to \int F_\mu d\nu$ Borel for every complex measure $\nu$ with compact support. But this follows from Fubini's Theorem once one notes Lemma 3.1 says $F_\mu(x)$ is jointly Borel in $\mu$ and $x$.

COROLLARY 3.4. *For $|j| \leqq k$ the map $(x, \mu) \to D^j F_\mu(x)$ is jointly Borel in $x$ and $\mu$ on the Cartesian product of $R^n$ and the set of all $\mu \in M(R^n)$ with distribution functions in $C^k(R^n)$.*

*Proof.* Since the natural injection of $C^k(R^n)$ into $C^0(R^n)$ is Borel, its inverse is also. Hence the map $\mu \to F_\mu$ is Borel from $\{\mu \in M(R^n)$: $F_\mu \in C^k(R^n)\}$ into $C^k(R^n)$. Also the map $f \to D^j f$ of $C^k(R^n)$ into $C^0(R^n)$ is continuous, and the map $(x, g) \to g(x)$ of $R^n \times C^0(R^n)$ into the complex plane is Borel by Lemma 3.1. Now combine these facts and the statement of the corollary follows.

Our next aim is to prove analogues of the preceding for probability measures having densities. Some of the assertions for densities follow by noting that the density of a probability measure on $R^n$ can be recovered by differentiating the distribution function a sufficient number of times. For the sake of completeness though, it seems preferable to proceed directly.

THEOREM 3.5. *Let $S$ be a Polish space and $\mu$ a $\sigma$-finite measure on the Borel sets of $S$. The collection of probability measures absolutely continuous with respect to $\mu$ and the collection of probability measures equivalent to $\mu$ both form Borel sets of $M(S)$. Furthermore, the map taking a probability measure $\nu$ absolutely continuous with respect to $\mu$ into its density $f_\nu \in L^1(\mu)$ is a Borel isomorphism.*

*Proof.* Consider the separable Banach space $L^1(\mu)$ of $\mu$-integrable functions. Take a countable collection of Borel sets $\{A_k\}_{k=1}^\infty$ in $S$ which are dense in the measure algebra derived from $\mu$. The set

$$B = \bigcap_{k=1}^{\infty} \left\{ f \in L^1(\mu) \colon \int f \chi_{A_k} d\mu \geqq 0 \right\} \cap \left\{ f \in L^1(\mu) \colon \int f d\mu = 1 \right\}$$

is a Borel set of $L^1(\mu)$. We claim that $B$ consists of the function in $L^1(\mu)$ which are nonnegative $\mu$-almost everywhere and have integral 1. Suppose $f \in B$. Then it is necessary to show $\int_A f d\mu \geqq 0$ for every Borel set $A$. But

$$\int_A f d\mu = \int_{A_k} f d\mu - \int_{A_k \setminus A} f d\mu + \int_{A \setminus A_k} f d\mu \ * .$$

By taking $\mu(A_k \Delta A)$ small enough, we can make the second and third contributions on the right of * as small as desired.

Now map $f$ in $B$ into the measure $\mu_f(A) = \int_A f d\mu$. This map is one-to-one and continuous for the weak* topology on the set of probability measures. The image of $B$ under the map is a Borel set and reduces simply to the probability measures absolutely continuous with respect to $\mu$. Furthermore, the inverse of $f \to \mu_f$ is also a Borel map.

To prove the assertion about probability measures equivalent to $\mu$ it suffices to prove that the set of $f \in B$ with $f > 0$ $\mu$-almost everywhere forms a Borel subset of $B$. Now this set is

$$\Big(\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{\substack{A_k \\ \mu(A_k)>1/m}}\Big\{f \in L^1(\mu): \int f\chi_{A_k} d\mu \geqq \frac{1}{n}\Big\}\Big) \cap \Big\{f \in L^1(\mu): \int f d\mu = 1\Big\} = C .$$

For, if $f \in B \backslash C$, then there is an $m$ and a sequence $\{A_{k_i}\}_{i=1}^{\infty}$ with

$$\mu(A_{k_i}) > \frac{1}{m} \text{ and } \int_{A_{k_i}} f d\mu \to 0 .$$

Assuming $f > 0$ $\mu$-almost everywhere, this contradicts the absolute continuity of $\mu$ with respect to $\mu_f$. On the other hand, if $f \in C$ and $\mu(A) > 0$, then take $A_k$ so that $\mu(A_k \Delta A)$ is small and $\mu(A_k) > 1/m$ for some $m$. Then our earlier representation * of $\int_A f d\mu$ shows that $\int_A f d\mu > 0$.

DEFINITIONS. $D(R^n)$ will denote the space of infinitely differentiable functions with compact support. It is well known that $D(R^n)$ is a countable union of separable Frechet spaces. [19] $S(R^n)$ will denote the space of infinitely differentiable functions rapidly decreasing at infinity together with all derivatives. With its usual topology $S(R^n)$ is a separable Frechet space. Finally, for $1 \leqq p < \infty$, $L^p(R^n)$ means the Banach space of equivalence classes of Borel functions whose $p$th powers are integrable.

LEMMA 3.6. *For $1 \leqq p < \infty$ each of the following spaces is a Borel set of $L^p(R^n)$:*

1. $C^k(R^n) \cap L^p(R^n)$ $\qquad k = 0, 1, 2, \cdots, \infty$
2. $\mathrm{Lip}\,(R^n) \cap L^p(R^n)$
3. $H(R^n) \cap L^p(R^n)$
4. $S(R^n)$
5. $D(R^n)$.

*Proof.* Since each of the spaces $C^k(R^n)$, $1 \leqq k \leqq \infty$, $\mathrm{Lip}\,(R^n)$, $H(R^n)$, $S(R^n)$, and $D(R^n)$ reside in $C^0(R^n)$ as Borel sets, it suffices to show that $C^0(R^n) \cap L^p(R^n)$ is a Borel set of $C^0(R^n)$ and that the injection $C^0(R^n) \cap L^p(R^n) \to L^p(R^n)$ is Borel. The first ingredient of the proof follows from the fact that

$$f \longrightarrow \sup_{\substack{n \in Z \\ n>0}} \int_{\|x\| \leqq n} |f(x)|^p dx = \int |f(x)|^p dx$$

is a Borel map of $C^0(R^n)$ into the extended real line. As for the second, note that

$$f \longrightarrow \lim_{n \to \infty} \int_{\|x\| \leqq n} f(x)g(x)dx = \int f(x)g(x)dx$$

is Borel from $C^0(R^n) \cap L^p(R^n)$ into $C$ for every $g \in L^q(R^n)$, where $1/p + 1/q = 1$ and $L^q(R^n)$ is identified with the dual of $L^p(R^n)$. According to the criterion already cited in Theorem 3.3, this implies that the injection $C^0(R^n) \cap L^p(R^n) \to L^p(R^n)$ is Borel.

THEOREM 3.7. *The collection of probability measures on $R^n$ whose densities lie in any of the spaces $C^k(R^n)$, $\mathrm{Lip}\,(R^n)$, $H(R^n)$, $S(R^n)$ or $D(R^n)$ form a Borel set of $M(R^n)$.*

*Proof.* The map taking a probability measure to its density with respect to Lebesgue measure is Borel into $L^1(R^n)$. Now apply the last lemma.

**4. Supports of probability measures.** In order to analyze the relation of a probability measure to its support, we find it convenient to introduce the Fell topology [10]. Let $X$ be a locally compact separable space. Denote the collection of closed subsets of $X$ by $\mathscr{C}(X)$. The Fell topology on $\mathscr{C}(X)$ can be given by specifying a basis of open sets of the form $U(C, \{V_1, \cdots, V_n\})$. $C$ is compact in $X$ and $\{V_1, \cdots, V_n\}$ is a finite, but possibly empty, family of open sets. $U(C; \{V_1, \cdots, V_n\}) = \{Y \in \mathscr{C}(X) : Y \cap C = \phi$ and $Y \cap V_i \neq \phi,\ i = 1, \cdots, n\}$. Fell demonstrates that $\mathscr{C}(X)$ is a compact Hausdorff space with this topology. Furthermore, as Fell observes, $\mathscr{C}(X)$ is separable and thus a compact metric space if $X$ is separable. Indeed, if $\mathscr{B}$ is a basis for the topology of $X$ and each element of $\mathscr{B}$ has compact closure, then the sets $U(C; \mathscr{F})$, where $C$ is the closure of the union of a finite subset of $\mathscr{B}$ and $\mathscr{F} \subset \mathscr{B}$ is finite, form a countable basis for $\mathscr{C}(X)$.

Next we wish to define notions of semicontinuity for maps into $\mathscr{C}(X)$. Since our definitions differ slightly from Kuratowski's [11] and Berge's [3], we feel it prudent to give a detailed discussion.

(Especially with reference to the next lemma see Chap. 6, Sec. 2 of [3].)

DEFINITION. Let $f\colon T \to \mathscr{C}(X)$ be a map from a topological space $T$ into $\mathscr{C}(X)$. $f$ is said to be *lower semicontinuous* if $\{t \in T\colon f(t) \cap V \neq \phi\}$ is open whenever $V$ is open in $X$. $f$ is *upper semicontinuous* if $\{t \in T\colon f(t) \cap C = \phi\}$ is open whenever $C$ is compact in $X$. Obviously, a map $f\colon T \to \mathscr{C}(X)$ will be continuous if and only if it is both lower and upper semicontinuous.

LEMMA 4.1. *Suppose $\{f_i\colon T \to \mathscr{C}(X)\}_{i \in \Lambda}$ is a nonempty collection of maps from a topological space $T$ into $\mathscr{C}(X)$.*

1. *If each $f_i$ is lower semicontinuous, then $t \to (\bigcup_{i\in\Lambda} f_i(t))^-$ is lower semicontinuous.*

2. *If each $f_i$ is upper semicontinuous, then $t \to \bigcap_{i\in\Lambda} f_i(t)$ is upper semicontinuous.*

3. *If the index set $\Lambda$ is finite and each $f_i$ is upper semicontinuous, then $t \to \bigcup_{i\in\Lambda} f_i(t)$ is upper semicontinuous.*

*Proof.* 1. Take $V$ open in $X$. Then $\{t \in T\colon (\bigcup_{i\in\Lambda} f_i(t))^- \cap V \neq \phi\} = \{t \in T\colon (\bigcup_{i\in\Lambda} f_i(t)) \cap V \neq \phi\} = \bigcup_{i\in\Lambda}\{t \in T\colon f_i(t) \cap V \neq \phi\}$ is open in $T$.

2. Take $C$ compact in $X$ and suppose $t_0 \in T$ satisfies $\bigcap_{i\in\Lambda} f_i(t_0) \cap C = \phi$. It is enough to show the existence of some neighborhood of $t_0$ where $\bigcap_{i\in\Lambda} f_i(t) \cap C = \phi$ continues to hold. To this end select one function $g$ from the collection $\{f_i\colon T \to \mathscr{C}(X)\}_{i\in\Lambda}$ and reduce the collection to $\{f_i\colon T \to \mathscr{C}(X)\}_{i\in\Sigma}$ by eliminating $g$. If $g(t_0) \cap C = \phi$, then our choice for the neighborhood of $t_0$ is obvious. Otherwise, put $K = g(t_0) \cap C$. For each $x \in K$ there is at least one index $i \in \Sigma$ with $x \notin f_i(t_0)$. Choose a neighborhood $V_x$ of $x$ with compact closure and satisfying $f_i(t_0) \cap V_x^- = \phi$. Since $f_i$ is upper semicontinuous, it is possible to select a neighborhood $W_x$ of $t_0$ such that $f_i(t) \cap V_x^- = \phi$ for all $t \in W_x$. Let $V_{x_1}, \cdots, V_{x_n}$ cover $K$. Applying the upper semicontinuity of $g$ there also exists a neighborhood $W$ of $t_0$ where $g(t) \cap (C\backslash(\bigcup_{m=1}^n V_{x_m})) = \phi$. It now follows that on $W \cap (\bigcap_{m=1}^n W_{x_m})$, $(\bigcap_{i\in\Lambda} f_i(t)) \cap C = \phi$.

3. For $C$ compact in $X_1$ $\{t \in T\colon (\bigcup_{i\in\Lambda} f_i(t)) \cap C = \phi\} = \bigcap_{i\in\Lambda}\{t \in T\colon f_i(t) \cap C = \phi\}$.

The next lemma shows that upper and lower semicontinuous mappings are Borel. (Compare Lemma 9.4 of [17].)

LEMMA 4.2. *The following are necessary and sufficient conditions for a map $f\colon T \to \mathscr{C}(X)$ from a Borel space $T$ into $\mathscr{C}(X)$ to be Borel.*

1. $\{t \in T: f(t) \cap V \neq \phi\}$ *is a Borel set for each open set* $V$.
2. $\{t \in T: f(t) \cap C = \phi\}$ *is a Borel set for each compact set* $C$.

*Proof.* Both conditions are clearly necessary for $f$ to be Borel. To prove their sufficiency consider a basic open set $U(C; \{V_1, \cdots, V_n\})$ in $\mathscr{C}(X)$. $U(C; \{V_1, \cdots, V_n\}) = U(C; \phi) \cap U(\phi; \{V_1\}) \cap \cdots \cap U(\phi; \{V_n\})$. Let $d$ be a metric for $X$ and define $W_m$ by $\{x \in X: d(x, C) = \inf_{y \in C} d(x, y) < 1/m\}$. Then $U(C; \phi) = \bigcup_{m=1}^{\infty} U(\phi; \{W_m\})'$. On the other hand, if $\{K_l\}_{l=1}^{\infty}$ is a sequence of compact sets whose union is some $V_i$, then $U(\phi; \{V_i\}) = \bigcup_{l=1}^{\infty} U(K_l; \phi)'$. Thus any basic open set can be expressed in terms of a countable number of open sets of type $U(K; \phi)$ or $U(\phi; \{V\})$.

LEMMA 4.3. *The map* $\mu \rightarrow$ *support* $(\mu)$ *is lower semicontinuous from* $M(X)$ *into* $\mathscr{C}(X)$.

*Proof.* For $V$ open in $X$, $\{\mu: \text{support}(\mu) \cap V \neq \phi\} = \{\mu: \mu(V) > 0\}$ is open in $M(X)$. In fact, $\{\mu: \mu(V) = 0\}$ is closed because $\mu(V) \leqq \lim_n \inf \mu_n(V)$ holds for every converging sequence $\mu_n \rightarrow \mu$ in $M(X)$. (See Thm. 6.1 of Chap. 2 of [16]).

LEMMA 4.4. *Each of the following maps is Borel*:

1. $F: \prod_{n=1}^{\infty} \mathscr{C}(X) \rightarrow \mathscr{C}(X)$ *given by* $F(\prod_{n=1}^{\infty} Y_n) = (\bigcup_{n=1}^{\infty} Y_n)^-$. $\prod_{n=1}^{\infty} \mathscr{C}(X)$ *has the product topology.*
2. $G: \prod_{n=1}^{\infty} \mathscr{C}(X) \rightarrow \mathscr{C}(X)$ *given by* $G(\prod_{n=1}^{\infty} Y_n) = \bigcap_{n=1}^{\infty} Y_n$.
3. $H: \mathscr{C}(X) \rightarrow \mathscr{C}(X)$ *given by* $H(Y) = (Y^0)'$.
4. $J: \mathscr{C}(X) \rightarrow \mathscr{C}(X)$ *given by* $J(Y) = (Y^0)^-$.

*Proof.* 1. The projection $F_k: \prod_{n=1}^{\infty} \mathscr{C}(X) \rightarrow \mathscr{C}(X)$ taking $\prod_{n=1}^{\infty} Y_n$ into $Y_k$ is continuous. Hence $F(\prod_{n=1}^{\infty} Y_n) = (\bigcup_{k=1}^{\infty} F_k(\prod_{n=1}^{\infty} Y_n))^-$ is lower semicontinuous.

2. $G$ is upper semicontinuous.

3. It suffices to show that the set of $Y \in \mathscr{C}(X)$ satisfying $(Y^0)' \cap V = \phi$ is Borel for each open set $V$. Now $(Y^0)' \cap V = \phi$ iff $Y^0 \supset V$ iff $Y \supset V$ iff $Y \supset V^-$ iff $Y \cap V^- = V^-$. The collection of $Y$ satisfying $Y \cap V^- = V^-$ is a Borel set because it is precisely the set where the Borel maps $Y \rightarrow Y \cap V^-$ and $Y \rightarrow V^-$ agree.

4. Again it is enough to prove that the collection of $Y \in \mathscr{C}(X)$ satisfying $(Y^0)^- \cap V = \phi$ is Borel for each open set $V$. Let $W = (V')^0$. Then $(Y^0)^- \cap V = \phi$ iff $V' \supset (Y^0)^-$ iff $V' \supset Y^0$ iff $W \supset Y^0$ iff $W' \subset (Y^0)'$ iff $W' \cap (Y^0)' = W'$. But the collection of $Y$ satisfying $W' \cap (Y^0)' = W'$ is precisely where the Borel maps $Y \rightarrow W' \cap H(Y)$ and $Y \rightarrow W'$ agree.

THEOREM 4.5. *Each of the following collection of sets in* $\mathscr{C}(X)$

*is Borel*:

1. *For every positive integer $k$, the collection of $Z \in \mathscr{C}(X)$ having $k$ or fewer points.*
2. *The collection of compact sets $Z \in \mathscr{C}(X)$.*
3. *The collection of compact connected sets $Z \in \mathscr{C}(X)$.*
4. *The single closed set $Y \in \mathscr{C}(X)$.*
5. *The collection of $Z \in \mathscr{C}(X)$ contained within a given closed set $Y$.*
6. *The collection of $Z \in \mathscr{C}(X)$ containing a given closed set $Y$.*
7. *The collection of $Z \in \mathscr{C}(X)$ with empty interior.*
8. *The collection on $Z \in \mathscr{C}(X)$ with no isolated points.*
9. *The collection of $Z \in \mathscr{C}(X)$ which are open as well as closed.*

*Proof*. 1. $\mathscr{D}^k$ is the collection of $Z \in \mathscr{C}(X)$ having $k$ or fewer points. According to Fell [10], $\mathscr{D}^1$ is closed and therefore compact in $\mathscr{C}(X)$. Since the union map $\prod_{n=1}^{k} \mathscr{C}(X) \to \mathscr{C}(X)$ is both lower and upper semicontinuous, the image of $\prod_{n=1}^{k} \mathscr{D}^1$, which is $\mathscr{D}^k$, is compact in $\mathscr{C}(X)$.

2. Let $\{C_n\}_{n=1}^{\infty}$ be an increasing sequence of compact sets having $\bigcup_{n=1}^{\infty} C_n^0 = X$. The collection of compact sets $Z \in \mathscr{C}(X)$ is the countable union of closed sets $\bigcup_{n=1}^{\infty} U(\phi; \{C_n'\})'$.

3. The compact disconnected sets coincide with the intersection of the collection of compact sets with the union of the basic open sets $U((V_1 \cup V_2)'; \{V_1, V_2\})$, where $V_1 \cap V_2 = \phi$ and $(V_1 \cup V_2)'$ is compact.

5. $Z \in \mathscr{C}(X)$ is contained within $Y$ iff $Z \notin U(\phi; \{Y'\})$.

6. $Z \in \mathscr{C}(X)$ contains $Y$ iff $Z \cap Y = Y$. Hence the collection of $Z \in \mathscr{C}(X)$ containing $Y$ is where the Borel maps $Z \to Z \cap Y$ and $Z \to Y$ agree.

7. Let $H$ be the map of last lemma. Then $Z$ has empty interior iff $H(Z) = X$.

8. Let $J$ be the map of the last lemma. Then $Z$ has no isolated points iff $J(Z) = Z$.

9. $Z \in \mathscr{C}(X)$ is open iff $H(Z) \cap Z = \phi$.

COROLLARY 4.6. *The collection of probability measures in $M(X)$ having support in any one of the families* 1.—9. *listed above is Borel.*

We now wish to introduce notions of congruence and symmetry for the space of closed sets $\mathscr{C}(X)$. To be specific, suppose $G$ is a locally compact separable topological group which acts on the right of $X$. If for each fixed $g \in G$, the map $x \to xg$ is continuous, then $G$ also acts on $\mathscr{C}(X)$. For $Y \in \mathscr{C}(X)$ and $g \in G$ define $Yg$ to be $\{yg: y \in Y\}$.

LEMMA 4.7. *If the action $X \times G \to X$ is jointly continuous, then* *so is the action* $\mathscr{C}(X) \times G \to \mathscr{C}(X)$. (Compare Prop. 2.2 of Chap. 2 of [2].)

*Proof.* For $Y \in \mathscr{C}(X)$ and $g \in G$ let $U(C; \{V_1, \cdots, V_n\})$ be a basic neighborhood of $Yg$. Since $Yg \cap V_i \neq \phi$, there exists $x_i \in Y$, a neighborhood $W_i$ of $x_i$, and a neighborhood $U_i$ of $g$ with $W_i U_i \subset V_i$. Let $V = \bigcap_{i=1}^n U_i$ and reduce $V$ if necessary so that $V^-$ is compact. Also define $K = \{xh^{-1}: x \in C, h \in V^-\}$. $K$ is compact, and by a further reduction of $V$ it is possible to assume $Y \cap K = \phi$. Now the neighborhood $U(K; \{W_1, \cdots, W_n\}) \times V$ of $(Y, g)$ maps into $U(C; \{V_1, \cdots, V_n\})$.

The next theorem provides some more instances of Borel sets of probability measures.

THEOREM 4.8. *If $G$ is a locally compact separable topological group, then the collection of closed subgroups is closed in $\mathscr{C}(G)$. The collection of closed normal subgroups is also closed in $\mathscr{C}(G)$. If $G$ acts continuously on a locally compact separable space $X$, the fixed points of $\mathscr{C}(X)$ under the action of $G$ form a closed set of $\mathscr{C}(X)$. Finally, the orbit of any $Y \in \mathscr{C}(X)$ under the action of $G$ is a Borel s et o $\mathscr{C}(X)$.*

*Proof.* The first statement is just Fell's observation. Indeed, the collection of closed subgroups is the complement of the union of $U(\{e\}; \phi)$ with the basic open neighborhoods of the form $U(V_1^-(V_2^-)^{-1}; \{V_1, V_2\})$, $e$ being the identity of $G$ and $V_1^-$ and $V_2^-$ being compact. The second statement is a special case of the third statement. Simply note that $G$ acts on itself by $(x, g) \to g^{-1}xg$ and hence on $\mathscr{C}(G)$ by $(Y, g) \to Yg = \{g^{-1}xg: x \in Y\}$. The fixed points of this action form a closed subset of $\mathscr{C}(G)$. Intersecting the collection of fixed points with the collection of closed subgroups gives the closed normal subgroups. Finally, the third statement is obvious, and the fourth statement follows because an orbit in $\mathscr{C}(X)$, $\{Yg: g \in G\}$, can be written as a countable union of compact sets, $\bigcup_{n=1}^\infty \{Yg: g \in K_n\}$, if $\{K_n\}_{n=1}^\infty$ is a sequence of compact subsets of $G$ whose union is $G$.

EXAMPLES. The sets in $\mathscr{C}(R^n)$ spherically symmetric about the origin are just the fixed points of $\mathscr{C}(R^n)$ under the action of the orthogonal group. Furthermore, under the action of the orthogonal group, the orbit of a subspace of dimension $m$, $m \leqq n$, is the collection of all subspaces of dimension $m$. The collection of all sets in $\mathscr{C}(R^n)$ geometrically congruent to a given closed set $Y$ lies on the orbit of $Y$ under the affine orthogonal group, i.e., the group generated by all orthogonal transformations and translations. As a consequence, the collection of all affine subspaces of a given dimension $m$ forms a Borel set of $\mathscr{C}(R^n)$. Similarly, considering the group of dilations (resp. homothetic transformation with the origin as center), it is clear that the collection of closed spheres (resp. closed spheres centered at the origin) forms a Borel set of $\mathscr{C}(R^n)$. (Consult [4] for the geometric terminology.)

The remainder of this section deals with convexity and subspaces and is inspired by [18]. Since the present proofs are different from those in [18] where results overlap, and require perhaps less background of the reader, we have furnished complete arguments.

THEOREM 4.9. *On $R^n$ the map $Y \to \mathrm{clconv}\,(Y)$ which takes a closed set into its closed convex hull is Borel from $\mathscr{C}(R^n)$ into $\mathscr{C}(R^n)$. Hence the collection of closed convex sets forms a Borel set of $\mathscr{C}(R^n)$, and the collection of probability measures with convex support forms a Borel set of $M(R^n)$.*

*Proof.* Let $\{l_i\}_{i \in A}$ be the collection of linear functionals on $R^n$. For each positive integer $k$ and $i \in A$ define $L_i^k \colon \mathscr{C}(R^n) \to \mathscr{C}(R^n)$ as follows: $L_i^k(Y) = \{x \in R^n \colon l_i(x) \leqq \sup_{w \in Y \cap S_k} l_i(w)\}$, where $S_k$ is the sphere $\{w \in R^n \colon \|w\| \leqq k\}$ and $L_i^k(Y)$ is taken to be $\phi$ if $Y \cap S_k = \phi$. Let us check that $L_i^k$ is upper semicontinuous. For $C$ compact $\{Y \in \mathscr{C}(R^n)\colon L_i^k(Y) \cap C = \phi\} = \{Y \in (R^n)\colon Y \cap S_k \cap \{t \in R^n \colon l_i(t) \geqq \inf_{s \in C} l_i(s)\} = \phi\}$, so $L_i^k$ is upper semicontinuous because $S_k \cap \{t \in R^n \colon l_i(t) \geqq \inf_{s \in C} l_i(s)\}$ is compact. Now note that $Y \to \bigcap_{i \in A} L_i^k(Y)$ is upper semicontinuous and $\bigcap_{i \in A} L_i^k(Y)$ is the closed convex hull of $Y \cap S_k$. It follows that $Y \to \bigcup_{k=1}^{\infty} \bigcap_{i \in A} L_i^k(Y)$ is a Borel map taking $Y$ into its closed convex hull.

The second claim of the theorem is true because the closed convex sets $Y$ of $\mathscr{C}(R^n)$ are simply where the Borel maps $Y \to \bigcup_{k=1}^{\infty} \bigcap_{i \in A} L_i^k(Y)$ and $Y \to Y$ agree.

DEFINITION. Suppose $Y$ is a nonempty closed convex set of $R^n$. For any point $x \in R^n$, $\mathrm{prox}\,(x, Y)$ is defined to be the unique point of $Y$ closest to $x$. In other words, if $d$ is the metric on $R^n$ derived from the Euclidean norm, then $\mathrm{prox}\,(x, Y)$ is the unique point of $Y$ where $d(x, Y) = \inf_{w \in Y} d(x, w)$ is attained.

LEMMA 4.10. *Suppose $f\colon T \to \mathscr{C}(R^n)$ is a Borel map with $f(t)$ a nonempty closed convex set for every $t \in T$. Then for every $x \in R^n$, $t \to$* prox $(x, f(t))$ *is a Borel map of $T$ into $R^n$.*

*Proof.* First, let us show that $t \to d(x, f(t))$ is a Borel map of $T$ into $R^+ = \{z \in R \colon z \geqq 0\}$. It suffices to show that $\{t \in T\colon d(x, f(t)) < \varepsilon\}$

is a Borel set of $T$ for each $\varepsilon > 0$. But $\{t \in T: d(x, f(t)) < \varepsilon\} = \{t \in T: \{y \in R^n: d(x, y) < \varepsilon\} \cap f(t) \neq \phi\}$. Next the map $s \to \{y \in R^n: d(x, y) \leqq s\}$ is Borel from $R^+$ into $\mathscr{C}(R^n)$ because for $K$ compact

$$\{s \in R^+: \{y \in R^n: d(x, y) \leqq s\} \cap K = \phi\} = \begin{cases} \phi & \text{if } \; x \in K \\ 0 \leqq s < d(x, K) \,. \end{cases}$$

Finally, note that $\operatorname{prox}(x, f(t)) = f(t) \cap \{y \in R^n: d(x, y) \leqq d(x, f(t))\}$. Since $t \to \operatorname{prox}(x, f(t))$ is a Borel map into $\mathscr{C}(R^n)$, it follows at once that it is a Borel map into $R^n$.

LEMMA 4.11. *Suppose $f: T \to \mathscr{C}(R^n)$ has the same properties as in the last lemma. Then the map $t \to f(t)^{\perp} = \{y \in R^n : \forall x \in f(t) \langle x, y\rangle = 0\}$ is also Borel from $T$ into $\mathscr{C}(R^n)$.*

*Proof.* Let $\{x_m\}_{m=1}^{\infty}$ be a dense collection of points in $R^n$. Since $f(t)$ equals the closure of $\bigcup_{m=1}^{\infty} \operatorname{prox}(x_m, f(t))$, $\{y \in R^n: \forall x \in f(t)\langle x, y\rangle \leqq 0\} = \bigcap_{m=1}^{\infty} \{y \in R^n: \langle \operatorname{prox}(x_m, f(t)), y\rangle \leqq 0\}$. Now $t \to \{y \in R^n: \langle \operatorname{prox}(x_m, f(t)), y\rangle \leqq 0\}$ is a Borel map into $\mathscr{C}(R^n)$ because $s \to \{y \in R^n: \langle s, y\rangle \leqq 0\}$ is Borel from $R^n$ into $\mathscr{C}(R^n)$. In fact, for a compact set $K$, $s \to \inf_{y \in K} \langle s, y\rangle$ is Borel and $\{s \in R^n: \{y \in R^n: \langle s, y\rangle \leqq 0\} \cap K \neq \phi\} = \{s \in R^n: \inf_{y \in K} \langle s, y\rangle \leqq 0\}$. Similarly, $t \to \{y \in R^n: \forall x \in f(t)\langle x, y\rangle \geqq 0\}$ is Borel. To finish the lemma observe that $t \to \{y \in R^n: x \in f(t)\langle x, y\rangle = 0\}$ is the intersection of two Borel maps.

THEOREM 4.12. *The map $\mathscr{C}(R^n)\backslash\{\phi\} \to \mathscr{C}(R^n)$ taking a closed set $Y$ into the smallest subspace containing $Y$ is Borel. Likewise, the map $\mathscr{C}(R^n)\backslash\{\phi\} \to \mathscr{C}(R^n)$ taking $Y$ into the smallest affine subspace containing $Y$ is Borel.*

*Proof.* The smallest subspace containing $Y$ is $[\operatorname{clconv}(Y)]^{\perp\perp}$. The smallest affine subspace containing $Y$ is

$$[\operatorname{clconv}(Y) - \operatorname{prox}(0, \operatorname{clconv}(Y))]^{\perp\perp} + \operatorname{prox}(0, \operatorname{clconv}(Y)) \,.$$

The second map is Borel because the action $\mathscr{C}(R^n) \times R^n \to \mathscr{C}(R^n)$ defined by translation is jointly continuous.

COROLLARY 4.13. *For each $m \leqq n$, the collection of probability measures on $R^n$ whose supports lie within a subspace (affine subspace) of dimension $m$ forms a Borel set of $M(R^n)$.*

## 5. Further examples of Borel sets of probability measures.

EXAMPLE 5.1. Let $G$ be a metric group acting continuously on a Polish space $X$. Then $G$ acts by translation on the set of probability measures $M(X)$. Indeed, define for $\mu \in M(X)$ and $g \in G$ $\mu g$ to be the probability measure assigning measure $\mu(Ag^{-1})$ to each Borel set $A$ of $X$.

THEOREM 5.2. *$G$ acts continuously on $M(X)$.*

*Proof.* Suppose $\mu_n \to \mu$ in $M(X)$ and $g_n \to g$ in $G$ as $n \to \infty$. It is sufficient to prove that

$$\int f(sg_n)d\mu_n(s) \longrightarrow \int f(sg)d\mu(s)$$

for every bounded continuous real-valued function $f$. For $\varepsilon > 0$ let $K$ be a compact subset of $X$ with $\mu_n(K) \geqq 1 - \varepsilon$ for all $n$. Now estimate as follows:

$$\left|\int f(sg_n)d\mu_n(s) - \int f(sg)d\mu(s)\right| \leqq \int_K |f(sg_n) - f(sg)|d\mu_n(s)$$
$$+ \int_{X\setminus K} |f(sg_n) - f(sg)|d\mu_n(s) + \left|\int f(sg)d\mu_n(s) - \int f(sg)d\mu(s)\right| . \quad **$$

The last term in ** can be made small since $\mu_n \to \mu$. The middle term on the right of ** is bounded by $2 \sup_{s \in X} |f(s)|\varepsilon$. The first term on the right can be made small because $sg_n \to sg$ uniformly on $K$, as we prove momentarily, and because $f$ is uniformly continuous on the compact set $\{kh: k \in K, h \in \{g_n\}_{n=1}^{\infty} \text{ or } h = g\}$. To show that $sg_n \to sg$ uniformly on $K$ let $d$ be the metric on $X$ and suppose $d(s_m g_{n_m}, s_m g) > \delta$ for some subsequence $\{g_{n_m}\}$ of $\{g_n\}$, a sequence $\{s_m\}$ of $K$ and some $\delta > 0$. Since $K$ is compact we may assume $s_m \to s \in K$. Then $s_m g_{n_m} \to sg$ and $s_m g \to sg$ by the joint continuity of the action of $G$ on $S$. This contradicts the assumption that $d(s_m g_{n_m}, s_m g) > \delta$ for all $m$. Hence $sg_n \to sg$ uniformly and this completes the proof of the theorem.

COROLLARY 5.3. *The invariant measures form a closed subset of $M(X)$ since they are the fixed points for the action of $G$ on $M(X)$. If $G$ is a Polish group, then the orbit of any probability measure under $G$ is Borel in $M(X)$.*

*Proof.* For the second assertion see Lemma 3.4 of [9].

*Applications.* On $R^n$, the orbit of any nondegenerate normal distribution under the group of invertible affine transformations is the whole collection of nondegenerate normal distributions. Also the collection of translates of any probability measure is Borel in $M(R^n)$.

Other commonly occurring groups acting on $R^n$ are the orthogonal group and the group permuting the coordinates of any point. The latter group arises in the theory of order statistics.

REMARK. A Borel set $B$ of a space $X$ is called invariant under a group action if $Bg = B$ for every group element $g$. If $\mu$ is a probability measure on $X$ and $\mu(B) = 0$ or $1$ for every invariant Borel set $B$, then $\mu$ is said to be ergodic. Varadarajan has shown that the collection of invariant ergodic probability measures is Borel in $M(X)$ if $X$ is a Polish space and the underlying group is locally compact and separable. (See Thm. 4.2 of [20].)

EXAMPLE 5.4. Suppose $\{X_n\}_{n=1}^{\infty}$ is a sequence of separable metric spaces. Consider the probability measures $M(\prod_{n=1}^{\infty} X_n)$ on the product space $\prod_{n=1}^{\infty} X_n$. We claim that the set $P = \{\mu \in M(\prod_{n=1}^{\infty} X_n) : \mu = \prod_{n=1}^{\infty} \mu_n,\ \mu_n \in M(X_n)\}$ is closed in $M(\prod_{n=1}^{\infty} X_n)$. Our reasoning goes as follows: The map $M(\prod_{n=1}^{\infty} X_n) \to \prod_{n=1}^{\infty} M(X_n)$ taking a probability measure into its sequence of marginal probability measures is continuous. Also the map $\prod_{n=1}^{\infty} M(X_n) \to M(\prod_{n=1}^{\infty} X_n)$ taking a sequence of probability measures into their product is continuous. (Modify slightly the proof of Lemma 1.1 of Chap. 3 of [16].) $P$ is the set where the composition of these two maps agrees with the identity map on $M(\prod_{n=1}^{\infty} X_n)$.

If each $X_n$ is the same, then the set of probability measures on $\prod_{n=1}^{\infty} X_n$ having all marginals the same is certainly closed too. Hence the set of probability measures on $\prod_{n=1}^{\infty} X_n$ which are product measures with equal components is closed.

EXAMPLE 5.5. Suppose $X$ is locally compact and separable. According to Corollary 4.6, the collection of probability measures concentrated at $k$ or fewer points is Borel in $M(X)$. A stronger assertion is possible.

LEMMA 5.6. *Let $X$ be a Polish space. For each $1 \geqq \delta > 0$ the collection of probability measures having $k$ or fewer atoms with total mass $\geqq \delta$ is closed in $M(X)$.*

*Proof.* An easy induction using Prohorov's Theorem. (See Thm. 6.7 of Chap. 2 of [16].)

For another application of Lemma 5.6 put $A_{k,n} = \{\mu \in M(X) : \mu$ has $k$ or less atoms with total mass $\geqq 1 - 1/n\}$. Then the Borel set $\bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} A_{k,n}$ consists of those probability measures concentrated on a finite or countable set of points.

EXAMPLE 5.7. If one is more interested in the number or geometry of the atoms rather their total weight, one can proceed as follows in the locally the compact case: For each $\mu \in M(X)$ and $\varepsilon > 0$ let $d_\varepsilon(\mu)$ be the set of atoms of $\mu$ having individual mass of at least $\varepsilon$. Since $\{\mu \in M(X): d_\varepsilon(\mu) \cap C \neq \phi\}$ is closed in $M(X)$ for each compact set $C$ of $X$, $d_\varepsilon: M(X) \to \mathscr{C}(X)$ is upper semicontinuous. (See Prop. I. 2. 8 of [1].) Apply part 1. of Lemma 4.4 to conclude that $\mu \to (\bigcup_{n=1}^\infty d_{1/n}(\mu))^-$ is a Borel map into $\mathscr{C}(X)$. $\mu \to (\bigcup_{n=1}^\infty d_{1/n}(\mu))^-$ can be used to keep track of the cardinality of the atoms and their positions.

EXAMPLE 5.8. The collection of probability measures on $R^n$ with some moment (all moments) existing is a Borel set of $M(R^n)$. To prove this let $\{g_i: R^n \to [0, 1]\}_{i=1}^\infty$ be a collection of continuous functions satisfying

$$g_i(x) = \begin{cases} 1 & \text{if } \|x\| \leqq i \\ 0 & \text{if } \|x\| \geqq i+1 . \end{cases}$$

Now note that

$$\mu \longrightarrow \sup_i \int |x^k| g_i(x) d\mu(x) = \int |x^k| d\mu(x)$$

is lower semicontinuous in the classical sense for each multi-index $k$. Hence the collection of probability measures having finite $k$th moment is a countable union of closed sets. In general, this collection is neither open nor closed. For instance, on $R$ it is possible to show that the collection of probability measures having finite first moment is neither open nor closed. Furthermore, the collection of probability measures lacking a first moment is dense in $M(R)$.

EXAMPLE 5.9. Occasionally it is convenient to deal only with those probability measures on $R^n$ having continuous, strictly increasing distribution functions. To characterize this family of probability measures consider for each pair of positive integers $n$ and $m$ the map $G_{n,m}: C^0(R^n) \to R$ defined by

$$G_{n,m}(g) = \inf_{\substack{\|z-w\|>1/n \\ \|z\| \leqq m \\ \|w\| \leqq m \\ z \geqq W}} g(z) - g(w) ,$$

where $z \geqq w$ means $z_i - w_i \geqq 0$ for each component of $z - w$. It is easy to check that $G_{n,m}$ is upper semicontinuous in the classical sence. Hence $\mu \to G_{n,m}(F_\mu)$ is a Borel map into $R$ from the collection of probability measures having continuous distribution functions. $\mu$ has strictly increasing distribution function iff $G_{n,m}(F_\mu) > 0$ for every $n$

and $m$. Similar arguments can be used to show that $\{(\mu, \nu): F_\mu, F_\nu$ continuous, $F_\mu(x) > F_\nu(x)\ \forall x\}$ is Borel in $M(R^n) \times M(R^n)$. However, removing the continuity assumptions makes both problems much more difficult.

EXAMPLE 5.10. A probability measure $\mu$ on $R$ is said to be symmetric if the Fourier transform $\hat{\mu}(\theta)$ of $\mu$ can be written as $e^{i\theta t}r(\theta)$, where $r(\theta)$ is a real-valued function of $\theta$ and $t$ is some real constant. Now for each positive integer $m$, $\{e^{i\theta t}r(\theta): t \in [-m, m]\ r$ real-valued and in $C^0(R)\}$ is closed in $C^0(R)$. Since $\mu_n \to \mu$ in $M(R)$ iff $\hat{\mu}_n \to \hat{\mu}$ in $C^0(R)$, the collection of symmetric probability measures on $R$ is a countable union of closed sets of $M(R)$.

EXAMPLE 5.11. For $\mu$ a probability measure on $R$ and $p \in (0, 1)$, $t \in R$ is called a $p$th percentile of $\mu$ if $\mu(-\infty, t) \leqq p$ and $\mu(-\infty, t] \geqq p$. When $p = 1/2$ the term median is used instead of percentile. It is easy to show that the set of $p$th percentiles for $\mu$ is a compact interval whose right endpoint is $\{^+(p, \mu) = \sup\{r \in Q: \mu(-\infty, r) \leqq p\}$ and whose left endpoint is $\{^-(p, \mu) = \inf\{r \in Q: \mu(-\infty, r] \geqq p\}$, where $Q$ is the set of rationals. Moreover, $\{^+(p, \mu)$(resp. $\{^-(p, \mu)$) is right (resp. left) continuous in $p$ for fixed $\mu$ and Borel in $\mu$ for fixed $p$. Hence Lemma 3.1 implies $\{^+(p, \mu)$ and $\{^-(p, \mu)$ are jointly Borel in $p$ and $\mu$. Using this fact one can show various hypotheses in nonparametric statistics involving the set of $p$th percentiles to be Borel. Perhaps it is worth pointing out that $(p, \mu) \to [\{^-(p, \mu), \{^+(p, \mu)]$ is a Borel map into $\mathscr{C}(R)$.

EXAMPLE 5.12. Let us indicate briefly now the Borel structure on $M(X)$ furnishes a natural framework for the description of several ideas in probability and statistics. For instance, in the theory of Markov processes one can define transition functions as Borel maps from $X$ into $M(X)$. If $\mu$ is a probability measure on a Polish space $X$ and $\pi$ is a Borel map onto another Polish space $Y$, define $\bar{\mu}$ on $Y$ by $\bar{\mu}(A) = \mu(\pi^{-1}(A))$ for every Borel set $A$ of $Y$. Then $\mu$ has a regular conditional probability distribution given $\pi$. From our perspective this means a Borel map $y \to \mu_y$ from $Y$ into $M(X)$ such that $\bar{\mu}$-almost all $\mu_y$ are concentrated on $\pi^{-1}(y)$ and $\mu(B) = \int \mu_y(B) d\bar{\mu}(y)$ for each Borel set $B$ of $X$. Finally, we should cite empirical distribution functions. Let $\{f_i: S \to R\}_{i=1}^\infty$ be a sequence of independent and identically distributed random variables on a Borel space $S$ with probability measure $\mu$. For each positive integer $n$ define a Borel map $\mu_n: S \to M(R)$ by taking $\mu_n(s)$ to be the probability measure giving equal weight to $f_1(s), \cdots, f_n(s)$. $\mu_n$ is Borel because for every Borel set

$$A \subset R,\ \mu_n(s)(A) = \frac{1}{n} \sum_{i=1}^{n} \chi_A(f_i(s)) ,$$

where $\chi_A$ is the indicator function of $A$.

**6. Counterexamples.** Obviously not all subsets of $M(X)$ are Borel. Here are some counter-examples.

EXAMPLE 6.1. For $X$ a Polish space it is well known that $X$ is homeomorphic to the collection of unit point masses, $\{\delta_w \in M(X): \delta_w(\{w\}) = 1, w \in X\}$ [16]. If $X$ is uncountable, then there exists $Y \subset X$ which is not Borel. But then $\{\delta_w \in M(X): w \in Y\}$ cannot be Borel in $M(X)$ either.

EXAMPLE 6.2. Our second counterexample involves the notion of equivalence between probability measures. It is transparent that mutual absolute continuity, denoted $\sim$, is an equivalence relation on $M(X)$. By the axiom of choice it is possible to choose one representative probability measure from each equivalence class. The next theorem shows when this "transversal" can also be taken to be a Borel set of $X$.

THEOREM 6.3. *Suppose $X$ is a Polish space. Then a Borel transversal exists for $\sim$ on $M(X)$ iff $X$ is countable or finite.*

*Proof.* Suppose $X$ is the set of positive integers. Give $Z_n = \{0, 1\}$ the discrete topology and consider the product space $\prod_{n=1}^{\infty} Z_n$. Subtract off from $\prod_{n=1}^{\infty} Z_n$ the countable number of sequences in which 1 appears only finitely often and call the remainder $Z$. Map $Z$ into $M(X)$ by taking the sequence $\{w_n\}_{n=1}^{\infty}$ into the probability measure giving mass $w_n(1/2)^{w_1+\cdots+w_n}$ to the integer $n$. This map is one-to-one, Borel, and provides the desired Borel transversal. The case of $X$ finite is even simpler.

Now assume $X$ is uncountable. Since any two uncountable Polish spaces $X$ and $Y$ are Borel isomorphic, (Thm. 2.12 of Chap. 1 of [16]), it is easy to see that $M(X)$ and $M(Y)$ will be Borel isomorphic too. Hence it is enough to establish the necessary part of the theorem for the space $\prod_{n=1}^{\infty} Z_n$ above. But this is the content of Lemma 5.1 of [15]. Here it is proved that $M(\prod_{n=1}^{\infty} Z_n)/\sim$ is not countably separated. If a Borel (even analytic) transversal existed in this case, then Prop. 2.12 of Chap. 1 of [2] would be contradicted, since $M(\prod_{n=1}^{\infty} Z_n)/\sim$ cannot be analytic if it fails to be countably separated. Note that $\sim$ is Borel as a subset of $M(\prod_{n=1}^{\infty} Z_n) \times M(\prod_{n=1}^{\infty} Z_n)$ because of 2.11 of [8].

EXAMPLE 6.4. Our next two counterexamples partially justify sticking to locally compact spaces when discussing the relation of a probability measure to its support. Suppose $X$ is a Polish space. Define a Borel structure on the space $\mathscr{C}(X)$ of closed subsets of $X$ by requiring every collection, $\{A \in \mathscr{C}(X): A \subset B\}$ to be Borel whenever $B \in \mathscr{C}(X)$. Christensen shows in Thm. 1 of [5] that this Borel structure is analytic and on the subspace of nonempty closed sets coincides with the Borel structure generated by the Hausdorff metric associated with any precompact metric on $X$. Furthermore, if $X$ is locally compact, this is the Borel structure generated by the Fell topology.

Now it is evident that $\mu \rightarrow$ support $(\mu)$ is Borel from $M(X)$ into $\mathscr{C}(X)$. If $X$ is a real infinite dimensional separable Hilbert space, Christensen proves that the collection $W$ of $Z \in \mathscr{C}(X)$ contained in the open unit sphere is complementary analytic but not analytic. (See Thm. 8 of [5].) Since every $Z \in W$ is the support of some $\mu \in M(X)$, the inverse image of $W$ under $\mu \rightarrow$ support $(\mu)$ fails to be Borel or even analytic.

This counterexample also illustrates that the intersection map $\mathscr{C}(X) \times \mathscr{C}(X) \rightarrow \mathscr{C}(X)$ need not be Borel when $X$ is not locally compact. Indeed, let $Y$ be the complement of the open unit sphere. Then $W = \{Z \in \mathscr{C}(X): Z \cap Y = \phi\}$.

EXAMPLE 6.5. The same phenomenon of Example 6.4 occurs if $X$ is a countable Cartesian product of the positive integers. Then the collection of $Z \in \mathscr{C}(X)$ which are open as well as closed is complementary analytic but not analytic. (See Thm. 5 of [5].)

ACKNOWLEDGEMENTS. I would like to thank Richard Dudley, Eric Nordgren, Don Cohn, and M. Takesaki for assistance with some of the proofs of this paper and for pointing out valuable references.

## REFERENCES

1. Alfsen, *Compact Convex Sets and Boundary Integrals*, Springer, Berlin, 1971.
2. Auslander and C. Moore, *Unitary representations of solvable Lie groups*, Memoirs Amer. Math. Soc., **62**, Amer. Math. Soc., Providence, 1966.
3. Berge, *Espaces Topologiques et Fonctions Multivoques*, Dunod, Paris, 1959.
4. Choquet, *Geometry in a Modern Setting*, Houghton Mifflin, Boston, 1964.
5. Christensen, *On some properties of Effros Borel structure on spaces of closed subsets*, Math. Ann., **195** (1971), 17-24.
6. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York, 1960.
7. ———, *Treatise on Analysis Vol. II*, Academic Press, New York, 1970.
8. Dubins and Freedman, *Measurable sets of measures*, Pacific J. Math., **14** (1964), 1211-1222.
9. Effros, *Transformation Groups and C*-Algebras*, Ann. of Math., **81** (1965), 38-55.
10. Fell, *A Hausdorff topology for the closed subsets of a locally compact non-Hausdorff space*, Proc. Amer. Math. Soc., **13** (1962), 472-476.
11. Kuratowski, *Topology Vol. I*, Academic Press, New York, 1966.
12. Meyer, *Probability and Potentials*, Ginn (Blaisdell), Boston, 1966.

13. R. Moore, *Measurable, continuous, and smooth vectors for semigroups and group representation*, Memoirs Amer. Math. Soc., **78**, Amer. Math. Soc., Providence, 1968.
14. Mori, *On random translations of point processes*, Yokoama Math. J., **19** (1971), 119-139.
15. Nielsen, *Borel Sets of Von Neumann Algebras*, Amer. J. Math., **95** (1973), 145-164.
16. Parthasarathy, *Probability Measures on Metric Spaces*, Academic Press, New York, 1967.
17. Ramsay, *Virtual groups and group actions*, Advances in Math., **6** (1971), 253-328.
18. Rockafellar, *Measurable dependence of convex sets and functions on parameters*, J. Math. Anal. Appl., **28** (1969), 4-25.
19. Treves, *Topological Vector Spaces, Distributions and Kernels*, Academic Press, New York, 1967.
20. Varadarajan, *Groups of automorphisms of Borel spaces*, Trans. Amer. Math. Soc., **109** (1963), 191-220.

Received May 11, 1972.

UNIVERSITY OF NEW HAMPSHIRE

*Current address*:

Department of Biomathematics
School of Medicine
University of California
Los Angeles, CA 90024.

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# PRODUCT INTEGRALS FOR AN ORDINARY DIFFERENTIAL EQUATION IN A BANACH SPACE

DAVID LOWELL LOVELADY

**Let $Y$ be a Banach space with norm $|\ |$, and let $R^+$ be the interval $[0, \infty)$. Let $A$ be a function on $R^+$ having the properties that if $t$ is in $R^+$ then $A(t)$ is a function from $Y$ to $Y$ and that the function from $R^+ \times Y$ to $Y$ described by $(t, x) \rightarrow A(t)[x]$ is continuous. Suppose there is a continuous real-valued function $\alpha$ on $R^+$ such that if $t$ is in $R^+$ then $A(t) - \alpha(t)I$ is dissipative. Now it is known that if $z$ is in $Y$, the differential equation $u'(t) = A(t)[u(t)]$; $u(0) = z$ has exactly one solution on $R^+$. It is shown in this paper that if $t$ is in $R^+$ then $u(t) = {}_0\prod^t \exp[(ds)A(s)][z] = {}_0\prod^t[I - (ds)A(s)]^{-1}[z]$, where the exponentials are defined by the solutions of the associated family of autonomous equations.**

The dissipitavity condition on $A$ is simply that if $(t, x, y)$ is in $R^+ \times Y \times Y$ and $c$ is a positive number then

$$(1)\qquad |[I - cA(t)][x] - [I - cA(t)][y]| \geqq [1 - c\alpha(t)]|x - y| .$$

The author and R. H. Martin, Jr. [5] have shown that if (1) holds, and $z$ is in $Y$, then there is exactly one continuously differentiable function $u$ from $R^+$ to $Y$ such that

$$(2)\qquad u(0) = z$$

and

$$(3)\qquad u'(t) = A(t)[u(t)]$$

whenever $t$ is in $(0, \infty)$. In the present article we shall show that $u$ can be expressed as a product integral in each of two forms:

$$(4)\qquad u(t) = \prod_0^t \exp[(ds)A(s)][z]$$

and

$$(5)\qquad u(t) = \prod_0^t [I - (ds)A(s)]^{-1}[z] .$$

Our work is related to results of J. V. Herod [2, §6] and G. F. Webb [7], [8]. Herod showed that representation (5) is valid if the mapping $(t, x) \rightarrow A(t)[x]$ is bounded on bounded subsets of $R^+ \times Y$. Webb obtained in [7] a representation similar to (4) under a set of hypotheses different from, and independent of, those used here. In

[8], Webb showed that (5) is valid if $A$ is independent of $t$. (Actually Webb in [8] restricted his attention to the case $\alpha = 0$, but his proofs adapt easily to the general time-independent case.)

II. **Product integrals.** We shall assume throughout that $A$ and $\alpha$ are as in our introduction, and that (1) is true whenever $(t, x, y)$ is in $R^+ \times Y \times Y$ and $c$ is a positive number. Now it follows from either of [5] and [6] that if $(t, x)$ is in $R^+ \times Y$ then there is exactly one solution $v$ of the problem

$$v'(s) = A(t)[v(s)];\ v(0) = x\,. \tag{6}$$

Furthermore, this problem generates an operator semigroup, which we shall denote $\{\exp[sA(t)]: s \text{ is in } R^+\}$, i.e., if $s$ is in $R^+$ then $\exp[sA(t)]$ is a function from $Y$ to $Y$ such that if $x$ is in $Y$ then $\exp[sA(t)][x] = v(s)$, where $v$ solves (6).

It is clear from (1) that there is no loss in assuming $\alpha$ to be $R^+$-valued, and we shall. It follows from [6] that if $(c, t)$ is in $R^+ \times R^+$ and $c\alpha(t) < 1$ then $I - cA(t)$ is a bijection on $Y$, and

$$|[I - cA(t)]^{-1}[x] - [I - cA(t)]^{-1}[y]| \leqq [1 - c\alpha(t)]^{-1}|x - y|$$

whenever $(x, y)$ is in $Y \times Y$. If $\{B_1, \cdots, B_n\}$ is a set of functions from $Y$ to $Y$, and $x$ is in $Y$, then $\prod_{j=1}^{0} B_j[x] = x$ and $\prod_{j=1}^{k} B_j[x] = B_k[\prod_{j=1}^{k-1} B_j[x]]$ whenever $k$ is an integer in $[1, n]$. If $(t, x, y)$ is in $R^+ \times Y \times Y$ then the statement

$$y = \prod_0^t [I - (ds)A(s)]^{-1}[x]$$

means that if $\varepsilon$ is a positive number then there is a chain $\{r_j\}_{j=0}^m$ from 0 to $t$ such that if $\{s_k\}_{k=0}^n$ is a refinement of $\{r_j\}_{j=0}^m$, and $\{\tilde{s}_k\}_{k=1}^n$ is a $[0, t]$-valued sequence such that if $k$ is an integer in $[1, n]$ then $\tilde{s}_k$ is in $[s_{k-1}, s_k]$, then

$$\left| y - \prod_{k=1}^{n} [I - (s_k - s_{k-1})A(\tilde{s}_k)]^{-1}[x] \right| < \varepsilon\,.$$

The statement

$$y = \prod_0^t \exp[(ds)A(s)][x]$$

is defined analogously.

THEOREM. *Let $z$ be in $Y$, and let $u$ solve* (2) *and* (3). *Then each of* (4) *and* (5) *is true whenever $t$ is in $R^+$.*

Let $m_-$ be that function from $Y \times Y$ to the real numbers given by

$$m_-[x, y] = \lim_{\delta \to 0-} (1/\delta)(|x + \delta y| - |x|) .$$

Now (1) is equivalent to requiring that

$$m_-[x - y, A(t)[x] - A(t)[y]] \leqq \alpha(t)\,|x - y|$$

whenever $(t, x, y)$ is in $R^+ \times Y \times Y$ (compare [1, p. 3]). Also, if $f$ is a function from a subset of $R^+$ to $Y$, if $c$ is in the domain of $f$, if $f'_-(c)$ (the left derivative of $f$ at $c$) exists, and if $P$ is given on the domain of $f$ by $P(t) = |f(t)|$, then $P'_-(c)$ exists and $P'_-(c) = m_-[f(c), f'_-(c)]$ (compare [1, p. 3]). If $(x, y, z)$ is in $Y \times Y \times Y$ then $m_-[x, y + z] \leqq m_-[x, y] + |z|$ (see [4, Lemma 6]). We are now prepared to prove our theorem.

*Proof of the theorem.* Let $b$ be a positive number, and let $\beta$ be a positive upper bound for the set $\{\alpha(t)\colon t \text{ is in } [0, b]\}$. Let $\varepsilon$ be a positive number, and let $\delta$ be a positive number such that $(\delta/\beta)(e^{\beta b} - 1) < \varepsilon$. Now $\{u(t)\colon t \text{ is in } [0, b]\}$ is a compact subset of $Y$, so the function described by $(t, x) \to A(t)[x]$ is uniformly continuous on $[0, b] \times \{u(t)\colon t$ is in $[0, b]\}$. In particular, there is a positive number $\eta$ such that if $(r, s, t)$ is in $[0, b] \times [0, b] \times [0, b]$ and $|r - s| < \eta$ then $|A(r)[u(t)] - A(s)[u(t)]| < \delta$. Let $\{t_k\}_{k=0}^n$ be a chain from 0 to $b$ such that $t_k - t_{k-1} < \eta$ whenever $k$ is an integer in $[1, n]$, and let $\{\tilde{t}_k\}_{k=1}^n$ be a $[0, b]$-valued sequence such that if $k$ is an integer in $[1, n]$ then $\tilde{t}_k$ is in $[t_{k-1}, t_k]$. Let $v$ be that function from $[0, b]$ to $Y$ having the property that if $k$ is an integer in $[1, n]$ and $t$ is in $[t_{k-1}, t_k]$ then

$$v(t) = \exp\,[(t - t_{k-1})A(\tilde{t}_{k-1})] \prod_{j=1}^{k-1} \exp\,[(t_j - t_{j-1})A(\tilde{t}_j)][z] .$$

Clearly now $v$ is continuous. Also, $v$ is left differentiable on $(0, b]$: if $k$ is an integer in $[1, n]$ and $t$ is in $(t_{t-1}, t_k]$ then

$$v'_-(t) = A(\tilde{t}_{k-1})[v(t)] .$$

Let $P$ be given on $[0, b]$ by $P(t) = |v(t) - u(t)|$. Now $P(0) = 0$. Suppose that $t$ is in $(0, b]$ and $k$ is an integer in $[1, n]$ and $t$ is in $(t_{k-1}, t_k]$. Now

$$\begin{aligned} P'_-(t) &= m_-[v(t) - u(t), v'_-(t) - u'(t)] \\ &= m_-[v(t) - u(t), A(\tilde{t}_{k-1})[v(t)] - A(t)[u(t)]] \\ &= m_-[v(t) - u(t), A(\tilde{t}_{k-1})[v(t)] - A(\tilde{t}_{k-1})[u(t)] \\ &\quad + A(\tilde{t}_{k-1})[u(t)] - A(t)[u(t)] \end{aligned}$$

$$
\begin{aligned}
&\leqq m_-[v(t) - u(t), A(\tilde{t}_{k-1})[v(t)] - A(\tilde{t}_{k-1})[u(t)]] \\
&\quad + |A(\tilde{t}_{k-1})[u(t)] - A(t)[u(t)]| \\
&\leqq \beta P(t) + \delta .
\end{aligned}
$$

Hence [3, Theorem 1.4.1, p. 15],

$$P(t) \leqq \int_0^t \delta e^{\beta(t-s)} ds = (\delta/\beta)(e^{\beta t} - 1)$$

whenever $t$ is in $[0, b]$. In particular,

$$
\begin{aligned}
&\left| u(b) - \prod_{k=1}^{n} \exp [(t_k - t_{k-1}) A(\tilde{t}_k)][z] \right| \\
&\quad = |u(b) - v(b)| \\
&\quad = P(b) \\
&\quad \leqq (\delta/\beta)(e^{\beta b} - 1) < \varepsilon .
\end{aligned}
$$

Thus we have proved that representation (4) is valid.

Now let $b$ and $\beta$ be as before. Let $c$ be a positive number such that $c\beta < 1/2$. Now if $t$ is in $[0, b]$ and $r$ is in $[0, c]$ then

$$
\begin{aligned}
&|[I - rA(t)]^{-1}[x] - [I - rA(t)]^{-1}[y]| \\
&\quad \leqq [1 - r\beta]^{-1} |x - y| \\
&\quad \leqq (1 + 2r\beta) |x - y| \\
&\quad \leqq e^{2r\beta} |x - y|
\end{aligned}
$$

whenever $(x, y)$ is in $Y \times Y$.

Now let $K = \{u(t): t \text{ is in } [0, b]\}$, and recall that $K$ is compact. Let $\varepsilon$ be a positive number. By the aforementioned uniform continuity, there is a positive number $\eta_1$ such that if $(s, t, x, y)$ is in $[0, b] \times [0, b] \times K \times K$ and $|s - t| < \eta_1$ and $|x - y| < \eta_1$ then $|A(s)[x] - A(t)[y]| < (\varepsilon/b)e^{-2\beta b}$. Let $\eta_2$ be a positive number such that if $(s, t)$ is in $[0, b] \times [0, b]$ and $|s - t| < \eta_2$ then $|u(s) - u(t)| < \eta_1$. Let $\delta = \min \{\eta_1, \eta_2, c\}$. Suppose that $0 \leqq r \leqq s \leqq t \leqq b$ and $t - r < \delta$. Let $\{\xi_k\}_{k=0}^n$ be a chain from $r$ to $t$, and let $\{\tilde{\xi}_k\}_{k=1}^n$ be a $[r, t]$-valued sequence such that if $k$ is an integer in $[1, n]$ then $\tilde{\xi}_k$ is in $[\xi_{k-1}, \xi_k]$. Now

$$
\begin{aligned}
&\left| \sum_{k=1}^{n} (\xi_k - \xi_{k-1}) A(\tilde{\xi}_k)[u(\tilde{\xi}_k)] - (t - r)A(s)[u(t)] \right| \\
&\leqq \sum_{k=1}^{n} (\xi_k - \xi_{k-1}) |A(\tilde{\xi}_k)[u(\tilde{\xi}_k)] - A(s)[u(t)]| \\
&\leqq \sum_{k=1}^{n} (\xi_k - \xi_{k-1})(\varepsilon/b)e^{-2\beta b} = (t - r)(\varepsilon/b)e^{-2\beta b} .
\end{aligned}
$$

It is now clear that

$$\left| \int_r^t A(\xi)[u(\xi)]d\xi - (t-r)A(s)[u(t)] \right| \leqq (t-r)(\varepsilon/b)e^{-2\beta b} .$$

Let $\{t_k\}_{k=0}^n$ be a chain from 0 to $b$, and suppose that $t_k - t_{k-1} < \delta$ whenever $k$ is an integer in $[1, n]$. Let $\{\tilde{t}_k\}_{k=1}^n$ be a $[0, b]$-valued sequence such that if $k$ is an integer in $[1, n]$ then $\tilde{t}_k$ is in $[t_{k-1}, t_k]$. Now

$$\begin{aligned}
&\left| \prod_{k=1}^n [I - (t_k - t_{k-1})A(\tilde{t}_k)]^{-1}[z] - u(b) \right| \\
&\leqq \sum_{k=1}^n \left| \prod_{j=k+1}^n [I - (t_j - t_{j-1})A(\tilde{t}_j)]^{-1}[u(t_k)] \right. \\
&\qquad \left. - \prod_{j=k}^n [I - (t_j - t_{j-1})A(\tilde{t}_j)]^{-1}[u(t_{k-1})] \right| \\
&\leqq \sum_{k=1}^n e^{2\beta(b-t_k)} |u(t_k) - [I - (t_k - t_{k-1})A(\tilde{t}_k)]^{-1}[u(t_{k-1})]| \\
&\leqq e^{2\beta b} \sum_{k=1}^n |[I - (t_k - t_{k-1})A(\tilde{t}_k)][u(t_k)] - u(t_{k-1})| \\
&= e^{2\beta b} \sum_{k=1}^n |u(t_k) - u(t_{k-1}) - (t_k - t_{k-1})A(\tilde{t}_k)[u(t_k)]| \\
&= e^{2\beta b} \sum_{k=1}^n \left| \int_{t_{k-1}}^{t_k} u'(\xi)d\xi - (t_k - t_{k-1})A(\tilde{t}_k)[u(t_k)] \right| \\
&= e^{2\beta b} \sum_{k=1}^n \left| \int_{t_{k-1}}^{t_k} A(\xi)[u(\xi)]d\xi - (t_k - t_{k-1})A(\tilde{t}_k)[u(t_k)] \right| \\
&\leqq e^{2\beta b} \sum_{k=1}^n (t_k - t_{k-1})(\varepsilon/b)e^{-2\beta b} = \varepsilon .
\end{aligned}$$

The proof of the theorem is complete.

## REFERENCES

1. W. A. Coppel, *Stability and Asymptotic Behavior of Differential Equations*, D. C. Heath & Co., Boston, 1965.

2. J. V. Herod, *A pairing of a class of evolution systems with a class of generators*, Trans. Amer. Math. Soc., **157** (1971), 247-260.

3. V. Lakshmikantham and S. Leela, *Differential and Integral Inequalities*, vol. 1, Academic Press, New York, 1969.

4. D. L. Lovelady, *A functional differential equation in a Banach space*, Funkcialaj Ekvacioj, **14** (1971), 111-122.

5. D. L. Lovelady and R. H. Martin, Jr., *A global existence theorem for a nonautonomous differential equation in a Banach space*, Proc. Amer. Math. Soc., **35** (1972), 445-449.

6. R. H. Martin, Jr., *A global existence theorem for autonomous differential equations in a Banach space*, Proc. Amer. Math. Soc., **26** (1970), 307-314.

7. G. F. Webb, *Product integral representation of time dependent nonlinear evolution equations in a Banach space*, Pacific J. Math., **32** (1970), 269-281.

8. G. F. Webb, *Nonlinear evolution equations and product integration in a Banach space,* Trans. Amer. Math. Soc., **148** (1970), 273-282.

Received June 7, 1972.

FLORIDA STATE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# A HOM-FUNCTOR FOR LATTICE-ORDERED GROUPS

JORGE MARTINEZ

**Results are presented that characterize subdirect products of reals (respectively, integers) functorially.**

By defining a quasi-order on the lattice-homomorphisms (henceforth: $l$-homomorphisms) of one abelian lattice-ordered group (henceforth: $l$-group) to another, one can set up a co-compatible system of partially ordered groups (henceforth: p. o. groups). Their co-limit $L(A, B)$, where $A$ and $B$ are the $l$-groups in question, is a directed, semi-closed p. o. group. If $A$ is a totally ordered group (henceforth: o-group) then $L(A, B)$ is simply the subgroup of Hom $(A, B)$ generated by the o-homomorphisms. On the other hand, if $B = \boldsymbol{R}$, the additive group of real numbers with the usual order, then $L(A, B)$ is a cardinal sum of copies of $\boldsymbol{R}$, one for each maximal $l$-ideal of $A$. In general the co-compatible system mentioned above is far from being directed.

$L(\cdot, B)$ is a contravariant functor; not much happens functorially in the second variable. It transforms $l$-epimorphisms (onto maps) into o-embeddings. The functor also preserves *finite* cardinal sums.

If the sequence $0 \to A \to B \to C \to 0$ is exact, i.e., $C \simeq B\backslash A$, then $0 \to L(C, X) \to L(B, X) \to L(A, X)$ is exact for any o-group $X$, provided $B \to C$ is a retraction. This happens in all of the following nontrivial cases: (1) $C$ is a projective $l$-group relative to all $l$-epimorphisms; (2) $B$ is divisible and $A$ is a prime subgroup of $B$; (3) $B$ is a direct, lexicographic extension of $A$ by $C$.

1. **Preliminaries.** Suppose $\{G_i \mid i \in I\}$ is a family of p. o. groups. If $G$ is the direct sum of the $G_i$ we call $G$ the *cardinal sum* of the $G_i$ if we define $0 \leqq g \in G$ if and only if $0 \leqq g_i \in G_i$ for all $i \in I$; notation: $G = \boxplus \{G_i \mid i \in I\}$. If each $G_i$ is an $l$-group and $G$ is the cardinal sum, then $G$ is also an $l$-group. $\boldsymbol{Z}$ (resp. $\boldsymbol{R}$) denotes the additive group of integers (resp. real numbers), with the usual ordering. We observe that an Archimedean o-group is o-isomorphic to a subgroup of $\boldsymbol{R}$ in its usual order; (Hölder's theorem, [3]). A *prime subgroup* $N$ of the $l$-group $G$ is a convex $l$-subgroup such that $G/N$ is an o-group. A p. o. group $G$ is *semi-closed* if given $g \in G$ and $ng \geqq 0$, with $n$ a positive integer, it follows that $g \geqq 0$.

We use $(\subset) \subseteqq$ for (proper) containment of sets; the symbol $\backslash$ for complementation in sets.

All groups in this discussion shall be abelian. If $A$ and $B$ are $l$-groups $\mathscr{L}(A, B)$ will denote the set of $l$-homomorphisms of $A$ into

$B$. We would like to construct a group $L(A, B)$ which "comes close" to behaving like a group of homomorphisms; the problem is of course that the sum of two $l$-homomorphisms need not be an $l$-homomorphism. Conrad and Diem have come up with a rather large set of $l$-endomorphisms of an $l$-group, which does turn out to be a semigroup under the usual addition of homomorphisms; they are the so-called $p$-endomorphisms, or polar-preserving endomorphisms (see [2]). We shall mention them in the sequel.

Suppose $A$ and $B$ are $l$-groups and $\theta, \phi \in \mathscr{L}(A, B)$. We say that $\phi$ *dominates* $\theta$ if $a\phi \wedge b = 0$ implies $a\theta \wedge b = 0$, for all $0 \leqq a \in A$ and $0 \leqq b \in B$; our notation for this is $\theta \prec \phi$. It is immediate that $\prec$ is a quasi-ordering of $\mathscr{L}(A, B)$. In fact, if $\theta \prec \phi$ and also $\phi \prec \theta$ we write $\phi \sim \theta$ and call $\phi$ and $\theta$ *polar equivalent*; $\sim$ is indeed an equivalence relation. Moreover, it induces a partial order on the equivalence classes, which we shall index $\{\mathscr{L}_i(A, B) \mid i \in I\}$: $\mathscr{L}_i(A, B) \leqq \mathscr{L}_j(A, B)$ if and only if some $\phi \in \mathscr{L}_j(A, B)$ dominates a $\theta \in \mathscr{L}_i(A, B)$. Now for each $i \in I$ let $L_i(A, B)^+ = \{\theta \in \mathscr{L}(A, B) \mid \theta \prec \phi,$ with $\phi \in \mathscr{L}_i(A, B)\} = \bigcup_{j \leqq i} \mathscr{L}_i(A, B)$. (We think of $I$ as being partially ordered so as to be compatible with the order induced on the equivalence classes.)

We are almost ready to state our first lemma; Hom $(A, B)$ is of course the full homomorphism group, $\mathscr{B}(A, B)$ the subgroup of Hom $(A, B)$ generated by $\mathscr{L}(A, B)$. Thus $\mathscr{B}(A, B) = \{\theta_1 - \theta_2 \mid \theta_1, \theta_2$ are sums of $l$-homomorphisms of $A$ into $B\}$.

LEMMA 1.1. (a) *For each* $i \in I$ $L_i(A, B)^+$ *is a subsemigroup of Hom* $(A, B)$; *that is, if* $\theta_1, \theta_2 \prec \phi$ then $\theta_1 + \theta_2 \in \mathscr{L}(A, B)$ and $\theta_1 + \theta_2 \prec \phi$.
(b) *For each* $i \in I$ $\mathscr{L}_i(A, B)$ *is a subsemigroup of* $L_i(A, B)^+$.

*Proof*. (a) Suppose $x \wedge y = 0$ in $A$; then $x\phi \wedge y\phi = 0$ for $\phi \in \mathscr{L}(A, B)$. If $\theta_1, \theta_2 \prec \phi$ we get $x\theta_1 \wedge y\phi = 0$, and in turn $x\theta_1 \wedge y\theta_2 = 0$. Likewise $x\theta_2 \wedge y\theta_1 = 0$, and of course $x\theta_i \wedge y\theta_i = 0$ for $i = 1, 2$, so that $(x\theta_1 + x\theta_2) \wedge (y\theta_1 + y\theta_2) = 0$, and so $\theta_1 + \theta_2$ is an $l$-homomorphism. If $a\phi \wedge b = 0$ then $a\theta_i \wedge b = 0$ for both $i = 1, 2$, so $a\theta_1 + a\theta_2 \wedge b = 0$, which means that $\theta_1 + \theta_2 \prec \phi$.

(b) We check that if $\theta_1$ and $\theta_2$ are polar equivalent to $\phi$ then so is $\theta_1 + \theta_2$. We already know that $\phi$ dominates $\theta_1 + \theta_2$. Yet if $a\theta_1 + a\theta_2 \wedge b = 0$ then since $0 \leqq a\theta_1 \leqq a\theta_1 + a\theta_2$ it follows that $a\theta_1 \wedge b = 0$, whence $a\phi \wedge b = 0$. The conclusion is that $\phi \prec \theta_1 + \theta_2$, and hence $\theta_1 + \theta_2 \sim \phi$.

For each $i \in I$ let $L_i(A, B)$ be the subgroup of $\mathscr{B}(A, B)$ generated by $L_i(A, B)^+$. If we declare an element $\phi \in L_i(A, B)$ positive when it is an $l$-homomorphism, one easily sees that $L_i(A, B)$ becomes a (directed) p.o. group whose cone is $L_i(A, B)^+$. If $i \leqq j$, let $f_{ij}$ stand for the inclusion map of $L_i(A, B)$ into $L_j(A, B)$. We define $L(A, B)$ to be the *co-limit* in the category of abelian groups of the system $\{L_i(A, B) \mid \{f_{ij}\}\}$. (It is easily verifiable that $f_{ii}$ is the identity on $L_i(A, B)$, and that if $i \leqq j \leqq k$ then $f_{ij}f_{jk} = f_{ik}$.)

PROPOSITION 1.2. *$L(A, B)$ is obtained as a quotient group of the direct sum of the $L_i(A, B)$ by factoring out the subgroup generated by all elements of the form*
$(\cdots, 0, \cdots, 0, \phi, 0, \cdots, 0, -\phi, 0, \cdots)$ *(with the two nonzero entries in the $i$th and $j$th position respectively, and $\phi \in L_i(A, B)$ while $i \leqq j$).*

*Proof.* The statement of the proposition merely sets out in detail the definition of a co-limit in the category of abelian groups.

Thus a typical element of $L(A, B)$ is a vector $(\cdots, \phi_i, \cdots)$ which is finitely nonzero, while addition and equality of vectors is subject to the identification imposed by Proposition 1.2; the entry $\phi_i \in L_i(A, B)$. The direct sum of the $L_i(A, B)$ may be ordered cardinally using the partial orders on the $L_i(A, B)$; it is clear also that the subgroup being factored out is trivially ordered in this partial order. We therefore have a partial order on $L(A, B)$ defined by $0 \leqq \phi \in L(A, B)$ if $\phi$ has a representation $(\cdots, \phi_i, \cdots)$ where each $\phi_i$ is an $l$-homomorphism.

A representation $\phi = (\cdots, \phi_i, \cdots)$ is said to be in *reduced form* if (1) for all $i \neq j$ in the support of $(\cdots, \phi_i, \cdots)$ $i$ and $j$ have no common upper bound in $I$, and (2) the cardinality of the support is minimal with respect to satisfying (1). The following lemma is obvious.

LEMMA 1.3. (a) *Each $\phi \in L(A, B)$ can be put in reduced form. If $(\cdots, \phi_i, \cdots)$ and $(\cdots, \theta_i, \cdots)$ are reduced forms of $\phi$, then their supports have the same cardinality, and there is a bijection $\pi$ of the supports such that $\phi_i = \theta_{\pi(i)}$.*

(b) *$0 \leqq \phi \in L(A, B)$ if and only if it has a reduced form $(\cdots, \phi_i, \cdots)$ such that $\phi_i \in \mathscr{L}(A, B)$ for each $i \in I$. If so then any reduced representation is by $l$-homomorphisms.*

PROPOSITION 1.4. *$L(A, B)$ is a directed, semi-closed p. o. group.*

*Proof.* $L(A, B)$ is obviously directed, so we need only verify it is semi-closed. Let $\phi \in L(A, B)$ and suppose $(\cdots, \phi_i, \cdots)$ is a reduced form of $\phi$. Suppose $n\phi \geqq 0$ for a positive integer $n$; the representation $(\cdots, n\phi_i, \cdots)$ of $n\phi$ is clearly again in reduced form. Hence by Lemma 1.3 each $n\phi_i$ is an $l$-homomorphism; one can easily check that each $\phi_i$ is in fact an $l$-homomorphism.

PROPOSITION 1.5. *If $B$ is an Archimedean l-group then $L(A, B)$ is an Archimedean p.o. group, in the sense that if $0 \leqq \phi \in L(A, B)$ and $n\theta \leqq \phi$ for each positive integer $n$, then $\theta \leqq 0$.*

*Proof.* Suppose $n\theta \leqq \phi, \phi \geqq 0$ and $\theta = (\cdots, \theta_i, \cdots)$ and $\phi = (\cdots, \phi_i, \cdots)$ are both given in reduced form. After reducing $\phi - n\theta$ we have three possibilities for an index $i$ of the support of $\phi - n\theta$:

$$(\phi - n\theta)_i = \begin{cases} \phi_j & \text{for some } j \in I \\ -n\theta_k & \text{for some } k \in I \\ \text{a sum of the above} . \end{cases}$$

Again invoking Lemma 1.3 it follows that $-n\theta_k$ is an $l$-homomorphism, and $\phi_j - n\theta_k \geqq 0$ for all $n = 1, 2, \cdots$ whenever the third choice occurs. In either case, (in the latter using the archimedeaneity of $B$) it follows that $\theta_k \leqq 0$. This shows $\theta \leqq 0$, and we are done.

For some information concerning the structure of $L(A, B)$ we look in the remainder of this section at some special cases.

*$A$ is an o-group* 1.6. In this situation the $l$-homomorphisms of $A$ into $B$ are simply the o-homomorphisms. The index set $I$ is then directed, since the sum of two o-homomorphisms is an o-homomorphism. $\mathscr{B}(A, B)$ then reduces to $\{\phi_1 - \phi_2 \mid \phi_i$ are o-homomorphisms of $A$ into $B\}$. Since each $L_i(A, B)$ is a subgroup of $\mathscr{B}(A, B)$ we may take their union over $I$; it is easily seen that this union is *precisely* $\mathscr{B}(A, B)$. Moreover, $L(A, B)$ is now the direct limit of the $L_i(A, B)$; it is well known that the direct limit of subgroups of an abelian group is the union of the subgroups. Hence $\mathscr{B}(A, B) \simeq L(A, B)$.

We have a converse of sorts:

PROPOSITION 1.6(a). *Suppose $A$ is not an o-group; then there is an o-group $B$ so that the index set $I$ in the construction of $L(A, B)$ is not directed.*

*Proof.* Suppose $A$ is not an o-group, and select $0 < x, y \in A$ such that $x \wedge y = 0$. Let $M$ (resp. $N$) be a prime subgroup that fails to contain $x$ (resp. $y$); then $y \in M$ and $x \in N$. Note that $A/M$ and $A/N$ are o-groups; we form $B$, the direct lexicographic extension of $A/M$ by $A/N$. We consider two $l$-homomorphisms $\phi$ and $\theta$ from $A$ into $B$: $\phi$ is the canonical map from $A$ onto $A/M$, followed by the (convex) inclusion of $A/M$ in $B$; $\theta$ is the canonical map from $A$ onto $A/N$ followed by the inclusion of that in $B$. Now observe that

$$(\phi + \theta)(x - y) \vee 0 = [(\phi + \theta)x - (\phi + \theta)y] \vee 0$$
$$= (M + x, N - y) \vee 0 = 0$$

whereas

$$(\phi + \theta)[(x - y) \vee 0] = (\phi + \theta)x = (M + x, 0) > 0 .$$

We conclude therefore that $\phi + \theta$ is not an $l$-homomorphism. The index set $I$ that arises in the construction of $L(A, B)$ is then not directed.

*B is an o-group* 1.7. One can verify with little trouble that $\phi \in \mathscr{L}(A, B)$ dominates $\theta \in \mathscr{L}(A, B)$ if and only if $\mathrm{Ker}(\phi) \subseteqq \mathrm{Ker}(\theta)$. Hence $\phi$ and $\theta$ are polar equivalent if and only if they have the same kernel. The kernels are all prime subgroups of $A$, and so $I$ is anti-isomorphic to a subset of the root system of primes (see [1], Theorem 1.7, the $l$-ideals containing a prime subgroup lie on a chain). $I$ is therefore a *tree-system*: no two incomparable elements of $I$ have a common upper bound; plainly, $I$ is far from being directed.

Now if $\phi \in L(A, B)$ then any vector representing $\phi$ is "almost" in reduced form; that is, it satisfies the first defining condition, except the support may be too large.

$B = \boldsymbol{R}$ 1.8. From the discussion in 1.7 it is clear that the index set $I$ is trivially ordered. We will show there is in fact an index $i \in I$ for each maximal $l$-ideal of $A$, and that $L(A, B)$ is a cardinal sum of copies of $\boldsymbol{R}$, one for each maximal $l$-ideal of $A$.

If $\phi: A \to B$ is an $l$-homomorphism, then $M = \mathrm{Ker}(\phi)$ is a maximal $l$-ideal. Using the fundamental theorem of $l$-homomorphisms there is an o-isomorphism $\bar{\phi}: A/M \to B$, which, by a well known corollary to Hölder's theorem, is a left multiplication by a positive real number. Thus the $l$-homomorphisms of $A$ into $B$ with kernel $M$ form a semigroup which is o-isomorphic to the additive semigroup of positive real numbers. This proves that each $L_i(A, B)$ is a copy of $\boldsymbol{R}$. It is clear that one such copy appears for each maximal $l$-ideal of $A$, since the corresponding quotient groups are all o-isomorphic to subgroups of $\boldsymbol{R}$.

Finally, the subgroup one factors out of the direct sum of these copies of $\boldsymbol{R}$ to get $L(A, B)$ is trivial here, and we conclude that $L(A, B)$ is a cardinal sum of copies of $\boldsymbol{R}$.

A similar argument can be made for $B = \boldsymbol{Z}$; one then obtains that $L(A, B)$ is a cardinal sum of copies of $\boldsymbol{Z}$, one for each maximal $l$-ideal of $A$ with cyclic factor in $A$.

A *polar preserving endomorphism* of an $l$-group $A$ is an $l$-endomorphism $\phi$ with the property that $x \wedge y = 0$ in $A$ implies that $x\phi \wedge y = 0$. (For an in-depth discussion of these endomorphisms the reader is referred to [2].) In our notation the semigroup of polar preserving endomorphisms is precisely the set of $l$-endomorphisms which are dominated by the identity on $A$. The subgroup they generate is one of the $L_i(A, A)$.

If $\phi$ is an $l$-homomorphism of $A$ *onto* $B$ and $\theta$ is a polar preserving endomorphism ($p$-endomorphism) of $B$, then $\phi\theta \prec \phi$, for if $x\phi \wedge y = 0$ then $x\phi\theta \wedge y = 0$. Conversely, if $\phi' \in \mathscr{L}(A, B)$ and $\phi' \prec \phi$ one easily sees that $\operatorname{Ker}(\phi) \subseteqq \operatorname{Ker}(\phi')$. This implies the existence of an endomorphism $\theta$ of $B$ satisfying $b\theta = a\phi'$ if $b = a\phi$. $\theta$ is certainly well defined, and it is a $p$-endomorphism since $\phi$ dominates $\phi'$. It follows then that if $i$ is the index in $I$ determined by $\phi$, $L_i(A, B)$ is o-isomorphic to the group generated by the $p$-endomorphisms of $B$.

We close this section with a rather general comment: for arbitrary $l$-groups $A$ and $B$ the groups $L_i(A, B)$ are subgroups of $\mathscr{B}(A, B)$; the inclusion mappings are compatible with the $f_{ij}$, so by the definition of co-limits we have a "natural" homomorphism of $L(A, B)$ into $\mathscr{B}(A, B)$. It assigns to $\phi = (\cdots, \phi_i, \cdots)$ the sum of the $\phi_i$ in $\mathscr{B}(A, B)$. About all that is on the surface concerning this mapping is that it is onto and an o-homomorphism. As a major unanswered question we might pose the following: when is this mapping an o-isomorphism? In most of the examples one can dream up it is, but as the $l$-groups get more complex, our knowledge of the structure of $L(A, B)$ decreases rapidly.

2. The functor $L(\cdot, B)$. We will show that $L(\cdot, B)$ is a contravariant functor from the category of abelian $l$-groups and $l$-homomorphisms into the category of directed, semi-closed p. o. groups with o-homomorphisms. ($L(A, \cdot)$ does not seem to be a functor at all.)

Suppose $\phi: A \rightarrow A'$ is an $l$-homomorphism; if $\theta_1, \theta_2: A' \rightarrow B$ are $l$-homomorphisms and $\theta_1 \prec \theta_2$ then $\phi\theta_1 \prec \phi\theta_2$. Thus $\phi$ induces an o-homomorphism $\phi^i$ of each $L_i(A', B)$ into some $L_{\phi(i)}(A, B)$; the map $i \rightarrow \phi(i)$ is an order preserving map of $I(A', B)$ into $I(A, B)$. We have canonical embeddings $\mu_i: L_i(A, B) \rightarrow L(A, B)(i \in I(A, B))$ and

$$\bar{\mu}_j: L_j(A', B) \longrightarrow L(A', B)(j \in I(A', B)) .$$

We also have the connecting embeddings $\{f_{ij}\}$, for $i \leqq j \in I(A, B)$, and $\{\bar{f}_{ij}\}$, for $i \leqq j \in I(A', B)$. Consider now for each $i \in I(A', B)$ the map $\phi^i\mu_{\phi(i)}: L_i(A', B) \rightarrow L(A, B)$. We show that if $i \leqq j$ in $I(A', B)$ then

$$\bar{f}_{ij}\phi^j\mu_{\phi(j)} = \phi^i\mu_{\phi(i)} ;$$

for if $0 \leqq \alpha_1, \alpha_2 \in L_i(A', B)$

$$\begin{aligned}(\alpha_1 - \alpha_2)\bar{f}_{ij}\phi^j\mu_{\phi(j)} = (\phi\alpha_1 - \phi\alpha_2)\mu_{\phi(j)} &= (\phi\alpha_1 - \phi\alpha_2)f_{\phi(i)\phi(j)}\mu_{\phi(j)} \\ &= (\phi\alpha_1 - \phi\alpha_2)\mu_{\phi(i)} \\ &= (\alpha_1 - \alpha_2)\phi^i\mu_{\phi(i)} .\end{aligned}$$

By the definition of the co-limit there is a unique homomorphism $L(\phi, B)$: $L(A', B) \to L(A, B)$ such that $\bar{\mu}_i L(\phi, B) = \phi^i \mu_{\phi(i)}$, for each $i \in I(A', B)$. Thus if

$$\alpha = (\cdots, \alpha_i, \cdots) \in L(A', B), \quad \alpha L(\phi, B) = (\cdots, (\phi\alpha_i)_{\phi(i)}, \cdots) ;$$

it is clear then that $L(\phi, B)$ is order preserving.

The next two lemmas are easy to prove; consequently we shall not bore the reader with their proofs.

LEMMA 2.1. *$L(\cdot, B)$ is a contravariant functor; that is if $\phi: A_1 \to A_2$ and $\theta: A_2 \to A_3$ are l-homomorphisms then*

$$L(\phi\theta, B) = L(\theta, B)\cdot L(\phi, B) ,$$

*and $L(1_A, B) = 1_{L(A,B)}$. ($1_G$ denotes the identity mapping on $G$.)*

LEMMA 2.2. *If $\phi, \phi': A \to A'$ are l-homomorphisms and $\phi + \phi'$ is too, then $L(\phi + \phi', B) = L(\phi, B) + L(\phi', B)$.*

In a category $\mathscr{C}$ with zero the *co-kernel* of a morphism $f: A \to B$ is a morphism $\gamma: B \to C$ such that $f\gamma = 0$, and having the property that if $\delta: B \to D$ is any morphism with $f\delta = 0$, then there is a unique morphism $\delta': C \to D$ such that $\gamma\delta' = \delta$. In the category of abelian $l$-groups the co-kernel of an $l$-homomorphism $\phi: A \to B$ is the canonical mapping $\eta: B \to B/J$ where $J$ is the convex hull of the image of $\phi$. All epimorphisms of this category have zero co-kernel, but not conversely. For instance, the embedding $j: Z \to Z \boxplus Z$ onto the diagonal has zero co-kernel, but if $\phi$ denotes the $l$-automorphism of $Z \boxplus Z$ given by $(a, b)\phi = (b, a)$ then $j\phi = j\cdot 1_{Z\boxplus Z} = j$, so $j$ is not epic.

THEOREM 2.3. *If $\alpha: A \to B$ is an l-homomorphism with zero co-kernel then $L(\alpha, X)$ has a trivially ordered kernel. This holds in particular if $\alpha$ is epic. If $\alpha$ is onto $B$ then $L(\alpha, X)$ is one-to-one.*

*Proof*. Suppose $\phi = (\cdots, \phi_i, \cdots) \in L(B, X)$ with each $\phi_i \geqq 0$, and assume $\phi L(\alpha, X) = 0$. Thus $(\cdots, (\alpha\phi_i)_{\alpha(i)}, \cdots) = 0$; this means that the vector $(\cdots, (\alpha\phi_i)_{\alpha(i)}, \cdots)$ of $\boxplus \{L_i(A, X) \mid i \in I(A, X)\}$ is in a trivially ordered subgroup. Thus each $\alpha\phi_i = 0$, and since $\alpha$ has zero co-kernel, each $\phi_i = 0$.

Now suppose $\alpha$ is onto and $\theta \in L(B, X)$. If $\theta = (\cdots, \theta_i, \cdots)$ is in reduced form then we seek to show $(\cdots, (\alpha\theta_i)_{\alpha(i)}, \cdots)$ is too. Clearly $i \in I(B, X)$ is in the support of $(\cdots, \theta_i, \cdots)$ if and only if $\alpha(i)$ is in the support of $(\cdots, (\alpha\theta_i)_{\alpha(i)}, \cdots)$ since $\alpha$ is onto. Suppose now that $\alpha(i)$, $\alpha(j)$ are both in the support of $(\cdots, (\alpha\theta_i)_{\alpha(i)}, \cdots)$ and $k \in I(A, X)$ exceeds both of them. Then whenever $0 \leqq \gamma_i \in L_i(B, X)$ and $0 \leqq \gamma_j \in L_j(B, X)$, $\alpha\gamma_i + \alpha\gamma_j = \alpha(\gamma_i + \gamma_j)$ is an $l$-homomorphism of $A$ into $X$. Again using the fact that $\alpha$ is onto one can then readily show that $\gamma_i + \gamma_j$ is an $l$-homomorphism. But then some index of $I(B, X)$ exceeds $i$ and $j$, contradicting the hypothesis that $(\cdots, \theta_i, \cdots)$ is reduced. A similar argument shows that the size of the support of $(\cdots, (\alpha\theta_i)_{\alpha(i)}, \cdots)$ is minimal; it now follows that $(\cdots, (\alpha\theta_i)_{\alpha(i)}, \cdots)$ is reduced.

Thus if $0 = \theta L(\alpha, X) = (\cdots, (\alpha\theta_i)_{\alpha(i)}, \cdots)$ then each $\alpha\theta_i = 0$ and so $\theta_i = 0$ for all $i \in I(B, X)$; hence $\theta = 0$ and so $L(\alpha, X)$ is one-to-one. (We shall see later that $L(\alpha, X)$ is in fact an o-embedding.)

The natural question here is: what does $L(\cdot, X)$ do to short exact sequences of $l$-groups? (We call a sequence $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ of $l$-homomorphisms *short exact* if $\alpha$ is one to one, $\beta$ is onto and $\operatorname{Ker}(\beta) = \operatorname{Im}(\alpha)$.) We will show presently that $L(\beta, X)$ is an o-monomorphism. Certainly $L(\beta, X)\cdot L(\alpha, X) = L(\alpha\beta, ) = 0$, but do we get exactness at $L(B, X)$? We shall give some partial answers, and then make some (hopefully) educated guesses.

PROPOSITION 2.4. *If* $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ *is a short exact sequence of l-groups, and* if $0 \leqq \phi \in \operatorname{Ker}(L(\alpha, X))$ *then* $\phi \in (L(C, X)^+)L(\beta, X)$. *In particular* $L(\beta, X)$ *is an o-embedding.*

*Proof.* If $\phi = (\cdots, \phi_i, \cdots) \geqq 0$ and $\phi L(\alpha, X) = 0$ then

$$(\cdots, (\alpha\phi_i)_{\alpha(i)}, \cdots) = 0 .$$

This means that in $\boxplus \{L_i(A, X) \mid i \in I(A, X)\{(\cdots, (\alpha\phi_i)_{\alpha(i)}, \cdots)$ is a vector whose components add to zero. But each entry $\alpha\phi_i$ is 0 or an $l$-homomorphism; if the sum of $l$-homomorphisms is zero each of them is zero. Thus $\alpha\phi_i = 0$ for each $i \in I(B, X)$; since $\beta$ is the co-kernel of $\alpha$, there is an $l$-homomorphism $\gamma^i: C \to X$ such that $\beta\gamma^i = \phi_i$. This determines a $\gamma \in L(C, X)$ whose image under $L(\beta, X)$ is $\phi$; clearly $0 \leqq \gamma$ and our proposition is proved.

PROPOSITION 2.5. *If* $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ *splits cardinally, i.e.* $B \cong A \boxplus C$, *then* $L(B, X) \cong L(C, X) \boxplus L(A, X)$.

*Proof.* If $B \cong A \boxplus C$ we have $l$-homomorphisms $\rho: C \to B$ and $\sigma: B \to A$ such that $\alpha\sigma = 1_A$, $\rho\beta = 1_C$, $\rho\sigma = 0$ and $1_B = \sigma\alpha + \beta\rho$. For each $l$-group $X$ we have

$$L(\sigma, X)L(\alpha, X) = 1_{L(A,X)}\,, \qquad L(\beta, X)L(\rho, X) = 1_{L(C,X)},$$
$$L(\sigma, X)L(\rho, X) = 0\,, \qquad L(\beta, X)L(\alpha, X) = 0\,,$$

and finally by Lemma 2.2

$$L(\alpha, X)L(\sigma, X) + L(\rho, X)L(\beta, X) = 1_{L(B,X)}\,.$$

This proves $L(B, X) \cong L(C, X) \boxplus L(A, X)$.

PROPOSITION 2.6. *Let $j\colon G \rightarrow \bar{G}$ be the natural embedding of the $l$-group $G$ in its divisible hull. For each $l$-group $X$ $L(j, X)$ is an o-embedding. If $X$ is divisible $L(j, X)$ is onto.*

*Proof.* If $\phi_1$ and $\phi_2$ are *any two homomorphisms* of $\bar{G}$ into the $l$-group $X$ which agree on $G$, then since each $x \in \bar{G}$ is of the form $x = (1/n)g$, for a suitable positive integer $n$, we have

$$n(x\phi_1) = n((1/n)g)\phi_1 = g\phi_1 = g\phi_2 = n((1/n)g)\phi_2 = n(x\phi_2)\,,$$

which implies that $x\phi_1 = x\phi_2$, since $X$ is torsion free. Clearly then $L(j, X)$ is one-to-one. Moreover, if $\phi\colon \bar{G} \rightarrow X$ is a homomorphism whose restriction to $G$ is an $l$-homomorphism then $\phi$ is an $l$-homomorphism; for if $x = (1/n)g \in \bar{G}$ with $g \in G$ then

$$\begin{aligned} n(x \vee 0)\phi = n((1/n)g \vee 0)\phi &= (g \vee 0)\phi = g\phi \vee 0 = [n(1/n)g\phi] \vee 0 \\ &= n[(1/n)g\phi \vee 0] = n(x\phi \vee 0)\,. \end{aligned}$$

This says that $L(j, X)$ is an o-embedding. Finally, if $X$ is divisible then each $l$-homomorphism of $G \rightarrow X$ extends (uniquely) to an $l$-homomorphism of $\bar{G} \rightarrow X$; in other words, $L(j, X)$ is onto.

We shall for the remainder of the section study the question of exactness of $L(\cdot, X)$ for $o$-groups $X$; according to 1.7 the picture we get of $L(A, X)$ is somewhat less cluttered. The preceding result tells us that if $X$ is divisible we might as well assume that $A$ is. So we ask: given an o-group $X$, which exact sequences $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$ go to exact sequences

$$0 \rightarrow L(C, X) \rightarrow L(B, X) \rightarrow L(A, X)\ ?$$

Prior to going into these questions more deeply we record some interesting properties of $L(\cdot, X)$.

PROPOSITION 2.7. *Let $\phi\colon A \rightarrow B$ be an $l$-homomorphism onto $B$. If $L(\phi, X)$ is an o-isomorphism for each o-group $X$ then $\phi$ itself is an isomorphism.*

REMARK. An analogous statement holds for o-groups $X$ with a minimal nonzero convex subgroup.

*Proof.* If $\phi$ is not one-to-one pick $0 < x \in \operatorname{Ker}(\phi)$ and let $N$ be a prime subgroup that fails to contain $x$. Set $X = A/N$ and $\eta: A \to X$ to be the canonical $l$-homomorphism. Then $(\cdots, 0, \cdots, \eta, \cdots, 0, \cdots) \in L(A, X)$ is not an image under $L(\phi, X)$.

THEOREM 2.8. *Let $A$ be an $l$-group; $A$ is a subdirect product of reals if and only if whenever $\phi: A \to B$ is an $l$-homomorphism onto $B$ then $L(\phi, \boldsymbol{R}): L(B, \boldsymbol{R}) \to L(A, \boldsymbol{R})$ is an $l$-isomorphism if and only if $\phi$ is an $l$-isomorphism.*

*Proof.* Suppose $\phi: A \to B$ is an $l$-homomorphism onto $B$. Let us examine what $L(\phi, \boldsymbol{R})$ does. There is a one-to-one correspondence between the maximal $l$-ideals of $B$ and the maximal $l$-ideals of $A$ that contain $K = \operatorname{Ker}(\phi)$. Now $L(B, \boldsymbol{R})$ and $L(A, \boldsymbol{R})$ are both cardinal sums of copies of $\boldsymbol{R}$, one for each maximal $l$-ideal of $B$ and $A$ respectively. So $L(\phi, \boldsymbol{R})$ is nothing more than the injection of $L(B, \boldsymbol{R})$ onto that portion of $L(A, \boldsymbol{R})$ corresponding to maximal $l$-ideals of $A$ that contain $K$.

If $L(\phi, \boldsymbol{R})$ is then onto for some $\phi$ with nonzero kernel, then every maximal $l$-ideal of $A$ contains $K$ and so $A$ is not a subdirect product of reals. Conversely, if $A$ is not a subdirect product of reals let $D$ be the intersection of all the maximal $l$-ideals of $A$; $D \neq 0$. Let $B = A/D$ and $\phi$ be the canonical mapping of $A$ onto $B$. By our arguments in the previous paragraph $L(\phi, \boldsymbol{R})$ is an $l$-isomorphism.

REMARK. A similar theorem holds for subdirect products of integers.

THEOREM 2.9. *Let $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ be a short exact sequence of $l$-groups. If $X$ is any Archimedean o-group then the induced sequence*

$$0 \to (C, X) \to L(B, X) \to L(A, X) \qquad \text{is exact.}$$

*If $X = \boldsymbol{R}$ then $L(\alpha, X)$ is onto if and only if every maximal $l$-ideal of $A$ is the meet of a maximal $l$-ideal of $B$ with $A$. If this is the case $L(B, X) \cong L(C, X) \boxplus L(A, X)$. If $X = \boldsymbol{Z}$ then $L(\alpha, X)$ is onto if and only if every maximal $l$-ideal of $A$ with cyclic factor is the meet with $A$ of a maximal $l$-ideal of $B$ with cyclic factor.*

*Proof.* As in 1.8 we have that if $\phi: B \to X$ is an $l$-homomorphism its kernel $M$ is a maximal $l$-ideal and $\phi$ determines an o-isomorphism from $B/M \to X$ which is a right multiplication by a suitable positive real number. The difference here is that not all maximal $l$-ideals appear as indices for $L_i(B, X)$, and the $L_i(B, X)$ themselves need not be full copies of $\boldsymbol{R}$. Still $L(B, X)$ is a cardinal sum of subgroups of $\boldsymbol{R}$ one for each "admissible" maximal $l$-ideal. Now $L(\beta, X)$ acts as in the proof of 2.8: there still is a one-to-one correspondence between maximal $l$-ideals of $C$ that appear as kernels of $l$-homomorphisms into $X$ and the same type of maximal $l$-ideals of $B$ that contain $A$. So $L(\beta, X)$ is the injection of $L(C, X)$ onto that portion of $L(B, X)$ corresponding to those maximal $l$-ideals of $B$ that contain $A$.

As for $L(\alpha, X)$ we have the following: if $\phi\colon B \to X$ is once again an $l$-homomorphism, and $M = \operatorname{Ker}(\phi) \not\supseteq A$ then $M \cap A$ is a maximal $l$-ideal of $A$ and it is the kernel of $\alpha\phi$. Thus $L(\alpha, X)$ has the effect of annihilating all the components of $L(B, X)$ corresponding to maximal $l$-ideals of $B$ that contain $A$, and being the identity on the remaining components.

It is now clear that $0 \to L(C, X) \to L(B, X) \to L(A, X)$ is exact, and also that the last part of the theorem holds, in the special cases when $X = \boldsymbol{R}$ or $X = \boldsymbol{Z}$.

In fact, after we record the following definition we have a better theorem.

Let $X$ be an o-group and $G$ be any $l$-group; a prime subgroup $N$ of $G$ is an *$X$-entry* of $G$ if it appears as the kernel of some $l$-homomorphism of $G$ into $X$. Thus:

THEOREM 2.9a. *If $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ is exact then $L(B, X) \cong L(C, X) \boxplus L(A, X)$ for an Archimedean o-group $X$ if and only if every $X$-entry of $A$ is the meet of an $X$-entry of $B$ with $A$.*

We have the following sufficient condition for the exactness of $0 \to L(C, X) \to L(B, X) \to L(A, X)$, when $X$ is an arbitrary o-group.

THEOREM 2.10. *If $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ is exact, then $0 \to L(C, X) \to L(B, X) \to L(A, X)$ is exact if $A + N = B$ for every $X$-entry of $B$ which does not contain $A$.*

The proof of this theorem depends upon the following lemma, which is known and quite easy to prove. (See [1], Theorem 1.14.)

LEMMA 2.11. *Let $G$ be an $l$-group, $A$ be a nonzero $l$-ideal of $G$. There is an o-isomorphism between the set of prime subgroups of $G$ that do not contain $A$ and the proper prime subgroups of $A$ via the mapping $N \mapsto N \cap A$.*

*Proof of* 2.10. Since the index sets $I(\cdot, X)$ are inversely o-isomorphic to a subset of prime subgroups we shall use the prime subgroups themselves to index the groups that make up the $L(\cdot, X)$'s.

Suppose then that $\phi = (\cdots, \phi_N, \cdots)$ is in reduced form and $\phi L(\alpha, X) = 0$, that is $(\cdots, (\alpha\phi_N)_{N\cap A}, \cdots) = 0$. If $N \supseteq A$ then $\alpha\phi_N$ is identically zero; to see this write $\phi_N = \phi_N^+ - \phi_N^-$, with $\phi_N^+, \phi_N^- \in \mathscr{L}(B, X)$; the kernels of $\phi_N^+$ and $\phi_N^-$ contain $N$ and hence $A$. In this case we need not worry about $\phi_N$; pick $\theta, \psi \in L(C, X)$ such that $\beta\theta = \phi_N^+$ and $\beta\psi = \phi_N^-$; then $\beta(\theta - \psi) = \phi_N$.

We are therefore left to consider those prime subgroups $N$ of $B$ which do not contain $A$. By Lemma 2.11 the support of $(\cdots, (\alpha\phi_N)_{N\cap A}, \cdots)$ is determined by precisely those prime subgroups; the lemma also guarantees that the representation is reduced. We have then that $\alpha\phi_N = 0$ for each prime subgroup $N \not\supseteq A$. Once again, writing $\phi_N = \phi_N^+ - \phi_N^-$ as a difference of $l$-homomorphisms (whose kernels contain $N$ *but not* $A$, for otherwise they would also vanish when restricted to $A$) we have $\alpha\phi_N^+ = \alpha\phi_N^-$.

Our assumption is though that $A + N = B$ for each such prime subgroup $N$, and this implies that $\phi_N^+ = \phi_N^-$. The conclusion here is that the support of $\phi = (\cdots, \phi_N, \cdots)$ consists of those $X$-entries which contain $A$. Our first paragraph in this proof then makes it clear that $\phi$ is the image of some element of $L(C, X)$ under $L(\beta, X)$. This completes the proof of the theorem.

COROLLARY 2.10.1. *Suppose $A$ is a maximal $l$-ideal of $B$, let $C = B/A$ and $0 \to A \to B \to C \to 0$ be the induced exact sequence. If $A$ is also a minimal prime subgroup then $0 \to L(C, X) \to L(B, X) \to L(A, X)$ is exact for all o-groups $X$.*

An $l$-group $G$ is *hyper-archimedean* if it is Archimedean and every $l$-homomorphic image of $G$ is Archimedean. It is well known (see for instance [1], Theorem 2.4) that $G$ is hyper-archimedean if and only if every prime subgroup is maximal (and hence minimal).

COROLLARY 2.10.2. *If $B$ is a hyper-archimedean $l$-group and $0 \to A \to B \to C \to 0$ is exact then $0 \to L(C, X) \to L(B, X) \to L(A, X)$ is exact for every o-group $X$.*

*Proof.* Every prime subgroup of $B$ is both maximal and minimal; consequently, if $N$ is an $X$-entry of $B$ that does not contain $A$ we have $B = A + N$. Theorem 2.10 now applies.

Another sufficient condition for the exactness of $0 \to L(C, X) \to$

$L(B, X) \to L(A, X)$ is obtained by requiring that $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ be "right splitting", i.e., that $\beta$ be a retract.

THEOREM 2.11. *Let* $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ *be an exact sequence of l-groups, and suppose* $\rho: C \to B$ *is an l-homomorphism such that* $\rho\beta = 1_C$. *Then for each o-group* $X$ $0 \to L(C, X) \to L(B, X) \to L(A, X)$ *is exact.*

*Proof.* We use the notation of the proof of Theorem 2.10. Let $\phi = (\cdots, \phi_N, \cdots)$ be an element of $L(B, X)$ in reduced form and consider $\phi L(\alpha, X) = (\cdots, (\alpha\phi_N)_{N \cap A}, \cdots)$; as shown in 2.10 this is once again reduced. So if $\phi L(\alpha, X) = 0$ we have $\alpha\phi_N = 0$ for all $X$-entries $N$ of $B$. As before, write $\phi_N = \phi_N^+ - \phi_N^-$ as the difference of $l$-homomorphisms of $B$ into $X$. For each $X$-entry define $\theta_N^+, \theta_N^-: C \to X$ by $\theta_N^+ = \rho\phi_N^+$ and $\theta_N^- = \rho\phi_N^-$. We claim that $\theta L(\beta, X) = \phi$, where $\theta = (\cdots, \theta_N, \cdots)$ and $\theta_N = \theta_N^+ - \theta_N^-$.

Note that $\rho$ induces a *group* direct sum $B \cong A \oplus C$; more precisely, each $b \in B$ can be expressed uniquely as $b = a\alpha + c\rho$, where $c = b\beta$. Thus $b\beta\theta_N^+ = b\beta\rho\phi_N^+$ and $b\beta\theta_N^- = b\beta\rho\phi_N^-$, while $b\phi_N^+ = a\alpha\phi_N^+ + c\rho\phi_N^+ = a\alpha\phi_N^+ + b\beta\rho\phi_N^+ = a\alpha\phi_N^+ + b\beta\theta_N^+$; likewise $b\phi_N^- = a\alpha\phi_N^- + b\beta\theta_N^-$, which implies that $b\phi_N = b\beta\theta_N$, for all $b \in B$.

This suffices to prove that $\theta L(\beta, X) = \phi$, and our theorem is proved.

COROLLARY 2.11.1. *Let* $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ *be an exact sequence; in all of the cases below* $0 \to L(C, X) \to L(B, X) \to L(A, X)$ *is exact for each o-group* $X$.

(a) *$C$ is a projective l-group.*

(b) *$B$ is divisible and $A$ is a prime subgroup of $B$.*

(c) *$B$ is a direct lexicographic extension of $A$ by $C$.*

*Proof.* In each of the above cases $\beta$ is a retract and the theorem applies.

COROLLARY 2.11.2. *If* $0 \to A \to B \to C \to 0$ *is exact where $A$ is a prime subgroup of $B$ then* $0 \to L(C, X) \to L(B, X) \to L(A, X)$ *is exact for each divisible o-group $X$.*

*Proof.* Apply Proposition 2.6 and Corollary 2.11.1 (b).

The following example may serve to illustrate a bit the difficulty in deciding which conjectures ought to be made in connection with this functor. Let $X = Z \overrightarrow{\times} Z$ with the lexicographic order: that is,

$(m, n) \geqq 0$ if $m > 0$ or $m = 0$ and then $n \geqq 0$. We will show that if $0 \to A \to B \to C \to 0$ is a short exact sequence then $0 \to L(C, X) \to L(B, X) \to L(A, X)$ is exact. So consider an exact sequence $0 \to A \to B \to C \to 0$, suppose $\phi = (\cdots, \phi_N, \cdots)$ is in reduced form and $\phi L(\alpha, X) = 0$ ($\phi \in L(B, X)$). As in the proof of 2.10 it suffices to consider those $X$-entries $N$ such that $N \not\supseteq A$. As before write $\phi_N = \phi_N^+ - \phi_N^-$ as a difference of $l$-homomorphisms whose kernels do not contain $A$. By our assumption $\alpha\phi_N^+ = \alpha\phi_N^-$; $\phi_N^+$ and $\phi_N^-$ have a common kernel, and after factoring out this kernel we have two o-embeddings of $X$ into itself, say $\theta_1$ and $\theta_2$, which agree on the nonzero proper convex subgroup of $X$. The o-homomorphisms of $X$ into itself are given by triangular integral matrices

$$\begin{pmatrix} m & p \\ 0 & n \end{pmatrix} \text{ with } m > 0, n \geqq 0 \text{ or } m = n = 0 \text{ and } p \geqq 0 .$$

If $\theta_i = \begin{pmatrix} m_i & p_i \\ 0 & n_i \end{pmatrix}$ $(i = 1, 2)$ and $\theta_1$ agrees with $\theta_2$ as specified, then $n_1 = n_2$, so clearly $\theta_1 - \theta_2$ is either order preserving or order inverting.

Lifting back to $B$ $\phi_N^+ - \phi_N^-$ is either an $l$-homomorphism or the additive inverse of one. Since $\alpha(\phi_N^+ - \phi_N^-) = 0$ there is a unique $l$-homomorphism $\psi\colon C \to X$ such that $\beta\psi = \pm(\phi_N^+ - \phi_N^-)$. This suffices to prove the exactness of the sequence.

The reader will appreciate the special nature of the above example.

3. **Comments and questions.** It appears that our functor will be of little use as the classical Hom-functor is in extension theory of abelian groups and modules. One might try to define an Ext-like functor using projective resolutions; in that case the question of independence of the resolution used appears to be an impossible problem. Or one could choose some "standard" free resolution; here it is obvious that computations could become nightmarish.

In view of some of our results, particularly Theorems 2.8 and 2.9, one can expect $L(\cdot, X)$ to be useful in characterizing certain lattice-group theoretical concepts. In any case, one large disadvantage of our construction is that there is no functoriality in the second variable.

Another possibility is that $L(\cdot, \boldsymbol{R})$ might serve as a "duality" functor between $l$-groups and abelian groups. Then one practically has to restrict oneself to subdirect products of reals, ($L(A, \boldsymbol{R}) = 0$ if $A$ has no maximal $l$-ideals), and then two such subdirect products of reals might very well have the same dual, (if they have the same number of maximal $l$-ideals.) A true duality can be realized, at least for subdirect products of reals, if one computes $L(A, X)$ for every Archimedean o-group, and then associates for each $A$ the whole "spectrum" $\{L(A, X) \mid X \text{ is a subgroup of } R\}$. Such a duality is evidently too cumbersome.

## REFERENCES

1. P. Conrad, *Lattice-Ordered Groups*, Tulane University, 1970.
2. P. Conrad and J. Diem, *The ring of polar-preserving endomorphisms of an abelian lattice-ordered group*; Illinois J. Math., **15**, No. 2, (1971), 222-240.
3. L. Fuchs, *Partially Ordered Algebraic Systems*, Pergamon Press, 1963.
4. B. Mitchell, *Theory of Categories*, Academic Press, 1965.

Received June 19, 1972.

UNIVERSITY OF FLORIDA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# WEAKLY ALMOST PERIODIC HOMEOMORPHISMS OF THE TWO SPHERE

W. K. MASON

**A self-homeomorphism $f$ of the 2-sphere $S^2$ is weakly almost periodic (w.a.p.) if the collection of orbit closures forms a continuous decomposition of $S^2$. It is shown that if $f$ is orientation-preserving, w.a.p. and nonperiodic, then $f$ has exactly two fixed points, and every nondegenerate orbit closure is an homology 1-sphere. There is an example with an orbit closure which is an homology 1-sphere but not a real 1-sphere. If $f$ is orientation-reversing, w.a.p. and has a fixed point, then $f$ is shown to be periodic. The orbit structure of orientation-reversing, w.a.p., nonperiodic homeomorphisms on $S^2$ is studied.**

1. Introduction. Let $f$ be a periodic mapping of the 2-sphere $S^2$ to itself. Kerékjártó [8] and Eilenberg [3] showed that $f$ is topologically equivalent either to the identity (every point fixed), to a rotation (two fixed points), a reflection (a simple closed curve of fixed points), or to a rotation followed by a reflection (no fixed points). If $f$ satisfies the weaker condition of being almost periodic (equivalent to having equicontinuous iterates), then the fixed point set of $f$ again is either empty or an $i$-sphere, $0 \leqq i \leqq 2$, [9]. (For related results on almost periodic mappings of subsets of $S^2$, see Hemmingsen [7].)

In the present paper we study the weakly almost periodic homeomorphisms on $S^2$, (the collection of orbit closures forms a continuous decomposition of $S^2$), and show that the set of fixed points is still either empty or an $i$-sphere, $0 \leqq i \leqq 2$, (Theorem 3 and Corollary 5). Some other results are: if $f$: $S^2 \rightarrow S^2$ is weakly almost periodic (w.a.p.), orientation-reversing, and has a fixed point, then $f$ is periodic (Theorem 4); if $f$: $S^2 \rightarrow S^2$ is w.a.p., orientation-preserving, and not periodic, then every nondegenerate orbit closure is an homology 1-sphere (Theorem 5).

A homeomorphism of $S^2$ to itself which is w.a.p. but not almost periodic is given in [12, Example 1]. This example is not almost periodic since it has an orbit closure which is not locally connected, (see [7, Section 5]). The collection of orbit closures, however, is easily seen to be continuous.

Our main theorems are given in §§ 6 and 7. Section 3 gives a summary of results in the theory of prime ends which we need. Section 4 discusses the fixed point theory used in §§ 5, 6, and 7. (Those familiar with prime ends and local fixed point index may skip

§§ 3 and 4.) Many of our techniques are based on those of Cartwright and Littlewood in [2].

**2. Definitions and notation.** If $f: X \to X$ is a homeomorphism and $x \in X$, then the *orbit closure* of $x$ is the closure of the set of iterates $\{f^n(x)\}$, $n = 0, \pm 1, \pm 2, \cdots$, $(f^0 = Id)$.

The original definition of weakly almost periodic was given by Gottschalk in [5]. For compact spaces the original definition is equivalent to requiring that the orbit closures form a continuous decomposition [5, Theorem 5]. The equivalent definition which we shall use in our proofs is: $f: S^2 \to S^2$ is *weakly almost periodic* if (a) the collection of orbit closures is a decomposition of $S^2$, (if two orbit closures meet, they are equal), and (b) for any closed set $B$, the union of all orbit closures which intersect $B$ is a closed set, [6, Theorem 4. 24, p. 34].

A point $x \in X$ is a *nonwandering point* if for every neighborhood $U$ of $x$, there is a nonzero integer $n$ such that $f^n(U) \cap U \neq \phi$. If $x$ is not a nonwandering point it is a *wandering point*. It is easily seen that if $f: S^2 \to S^2$ is w.a.p. then every point is a nonwandering point.

A *domain* is a connected open set. If $A$ is a set $\mathrm{Cl}(A)$ and $\mathrm{Bd}(A)$ denote the closure and boundary, respectively, of $A$. If $U$ is a domain of $S^2$ and $x$ is a point in a component $R$ of $S^2 - \mathrm{Cl}(U)$, then $\mathrm{Bd}(R)$ is the *outer boundary of $U$ with respect to $x$*.

An *homology 1-sphere* $K$ in $S^2$ is a continuum (closed, connected set) such that $S^2 - K$ has exactly two components.

An open triod is a set homeomorphic to the set of all points $(x, y)$ in the plane such that either $-1 < x < 1$ and $y = 0$, or $x = 0$ and $0 \leqq y < 1$. The points $(-1, 0)$, $(1, 0)$, $(0, 1)$ are called the *feet* of the triod.

If $U$ is a domain then a *crosscut* of $U$ is an open arc in $U$ whose closure is an arc which intersects $\mathrm{Bd}(U)$ in two points. An *endcut* of $U$ is a half-open arc in $U$ whose closure is an arc which intersects $\mathrm{Bd}(U)$ in one point.

**3. Prime ends.** In this section we state the results and definitions concerning prime ends which we shall use in §§ 5 and 6. The material in the present section is taken from [2], [11], and [15].

Let $U$ be a simply-connected domain in $S^2$ with a nondegenerate boundary. A *C-transformation* of $U$ onto the open unit disk $D$ is a homeomorphism $T: U \to D$ such that the image of any crosscut in $U$ is a crosscut in $D$, and the endpoints of such images of crosscuts of $U$ are dense in the boundary of $D$. The conformal mapping of $U$ onto $D$ given by the Riemann mapping theorem shows that $C$-transformations always exist. However, $C$-transformations may be constructed by topological methods, without using conformal mapping theory, [15, Appendix 2].

Given a homeomorphism $f$ of the closure of $U$ onto itself, and a $C$-transformation $T$ of $U$ onto $D$, we have that $TfT^{-1}$: $D \rightarrow D$ is a $C$-transformation which may be extended to a homeomorphism of the closed unit disk onto itself, [15, (4.10) on page 6, and (A1.7) on page 27].

A collection of crosscuts $Q_1, Q_2, \cdots$ of the simply connected domain $U$ is a *chain* if (a) the arcs $\mathrm{Cl}(Q_1)$, $\mathrm{Cl}(Q_2), \cdots$ are pairwise disjoint, (b) $Q_n$ separates $Q_{n-1}$ from $Q_{n+1}$ in $U$, (c) there is a point on $\mathrm{Bd}(U)$ whose greatest distance from $Q_n$ approaches 0 as $n \rightarrow \infty$. Corresponding to each $Q_n$ there is a domain $G_n$ of $U - Q_n$ containing $Q_{n+1}$. Note that $G_1 \supset G_2 \supset \cdots$.

If $\{Q_i\}$, $\{R_i\}$ are chains of crosscuts, and $\{G_i\}$, $\{H_i\}$ are their respective corresponding domains, then $\{Q_i\}$, $\{R_i\}$ are *equivalent* chains if for every $n$ there is an $m$ such that $H_m \subset G_n$ and $G_m \subset H_n$. Equivalent chains are said to define the same *prime end*. Thus, a prime end of $U$ is an equivalence class of chains of $U$.

If $Q_1, Q_2, \cdots$ is a chain of crosscuts in $U$, then their images $T(Q_1)$, $T(Q_2), \cdots$ under the $C$-transformation $T$: $U \rightarrow D$ is a chain in $D$, [15, Appendix 2]. If $\{Q_i\}$ and $\{R_i\}$ are equivalent chains in $U$, then $\{T(Q_i)\}$, and $\{T(R_i)\}$ are equivalent chains in $D$, and in fact converge to the same point on the boundary of $D$, ($\{Q_i\}$ and $\{R_i\}$ may not converge to the same point on $\mathrm{Bd}(U)$). Thus, $T$ sets up a $1-1$ correspondence between prime ends of $U$ and points of the unit circle [11, p. 621].

If $f$: $\mathrm{Cl}(U) \rightarrow \mathrm{Cl}(U)$ is a homeomorphism and $E$ is a prime end of $U$, then $E$ is *fixed* by $f$ if for some chain $\{Q_i\}$ defining $E$, we have that $\{Q_i\}$ and $\{f(Q_i)\}$ are equivalent chains. This definition is easily seen to be independent of which defining chain is used. If $T$: $U \rightarrow D$ is a $C$-transformation, $h$: $\mathrm{Cl}(D) \rightarrow \mathrm{Cl}(D)$ is the extension of $TfT^{-1}$, and $e$ is the point on $\mathrm{Bd}(D)$ corresponding to the fixed prime end $E$, then $h(e) = e$. Conversely, every fixed point of $h$ on $\mathrm{Bd}(D)$ corresponds to a fixed prime end of $f$.

If $E$ is a prime end of $U$, $\{Q_i\}$ is a defining chain for $E$, and $p$ is the point on $\mathrm{Bd}(U)$ to which the crosscuts $\{Q_i\}$ converge, then $p$ is a *principal point* of $E$. (We remark that there exists a $U$ with a prime end $E$ such that every point of $\mathrm{Bd}(U)$ is a principal point of $E$, [13].)

If $A$ is an endcut in $U$ with an endpoint $s \in \mathrm{Bd}(U)$, then there is a chain $\{Q_i\}$ defining a prime end $E$ such that $s$ is a principal point of $E$ and each crosscut $Q_i$ separates the endpoint of $A$ in $U$ from some (open) subarc of $A$ having $s$ as an endpoint. $E$ is the *prime end determined by* $A$. If $T: U \to D$ is a $C$-transformation, and $e$ is the point on $\mathrm{Bd}(D)$ corresponding to $E$, then $T(A)$ is an endcut in $D$ having $e$ as an endpoint, [15, page 5].

**4. Lefschetz number and local fixed point index.** In this section we state the results concerning fixed points which we shall use in §§ 5, 6, and 7.

If $X$ is a compact polyhedron and $f: X \to X$ is a map (continuous function), then there is a certain rational number $L(f)$, called the *Lefschetz number* of $f$, associated with $f$ and $X$, [14, p. 195]. We shall use the following two facts about $L(f)$.

*Fact* 1. If $X$ is a two cell, then $L(f) = 1$.

*Fact* 2. If $X$ is a 2-sphere and $f$ is an orientation-preserving homeomorphism, then $L(f) = 2$.

For proofs of Facts 1 and 2, see [14, p. 196].

If $e$ is the category of compact polyhedra and maps, let $A(e)$ denote the set of pairs $(f, U)$, where $f: X \to X$ is a map in $e$ and $U$ is an open subset of $X$ such that $f$ has no fixed points on the boundary of $U$. Then there is a function $i$, the *local fixed point index*, from $A(e)$ into the rationals which satisfies the following axioms:

A1. If $(f, U)$, $(g, U)$ belong to $A(e)$, and $f = g$ on the closure of $U$, then $i(f, U) = i(g, U)$.

A2. If $f_t$ is a homotopy such that $(f_t, U) \in A(e)$ for each $t$, $0 \leqq t \leqq 1$, then $i(f_0, U) = i(f_1, U)$.

A3. If $(f, U) \in A(e)$ and $U$ contains mutually disjoint open sets $V_j, j = 1, \cdots, k$, such that $f$ has no fixed points on $U - \bigcup_{j=1}^{k} V_j$, then

$$i(f, U) = \sum_{j=1}^{k} i(f, V_j) .$$

In particular, if $f$ has no fixed points on $U$, $i(f, U) = 0$.

A4. If $f: X \to X$ belongs to $e$, then $i(f, X) = L(f)$.

A5. If the maps $f: X \to Y$, $g: Y \to X$ belong to $e$, and

$$(gf, U) \in A(e) ,$$

then $i(gf, U) = i(fg, g^{-1}(U))$.

For further discussion of the local fixed point index see [4] or [1].

REMARK. If $D$ is the open unit disk, and $h$ is a map of the closure of $D$ to itself with no fixed points on $\mathrm{Bd}(D)$, then $i(h, D) = 1$. For, by Fact 1 and Axiom A4, $1 = \mathrm{L}(h) = i(h, \mathrm{Cl}(D))$. Then, by Axiom A3, $i(h, \mathrm{Cl}(D)) = i(h, D)$.

## 5. Preliminary lemmas.

Our first lemma is based on Lemma 11 of [2].

LEMMA 1. *Suppose $f\colon S^2 \to S^2$ is a homeomorphism, $U$ is a simply connected domain with nondegenerate boundary, $f(U) = U$, and every point of $U$ is a nonwandering point. Suppose also that $E$ is a prime end of $U$ which is fixed by $f$. Then every principal point of $E$ is a fixed point of $f$.*

*Proof.* Let $Q_1, Q_2, \cdots$ be a chain of crosscuts defining $E$ which converge to the principal point $p$ of $E$.

*Case* 1. $f(Q_i) \cap Q_i = \phi$ for some $i$. Let $V$ be the component of $U - Q_i$ containing $Q_{i+1}, Q_{i+2}, \cdots$. $E$ is fixed by $f$, so $\{Q_j\}$ and $\{f(Q_j)\}$ are equivalent chains, hence $f(V) \cap V \neq \phi$. But then $f(V)$ either contains or is contained in $V$. Assume $f(V) \subset V$. Let $W$ be the nonempty open set $V - \mathrm{Cl}(f(V))$. Then $f^n(W) \cap W = \phi$ if $n \neq 0$. Thus no point of $W$ is a nonwandering point. This contradiction shows that Case 1 cannot occur.

*Case* 2. $f(Q_i) \cap Q_i \neq \phi$ for all $i$, $i = 1, 2, \cdots$. For each $i$, select a point $x_i \in Q_i$ such that $f(x_i) \in Q_i$. The crosscuts $Q_1, Q_2, \cdots$ converge to the principal point $p$, hence $\{x_i\} \to p$, hence $\{f(x_i)\} \to f(p)$. But $f(x_i) \in Q_i$, hence $\{f(x_i)\} \to p$. Hence $f(p) = p$ and the proof of Lemma 1 is complete.

LEMMA 2. *Suppose $f\colon S^2 \to S^2$ is a homeomorphism, $M$ is an invariant continuum in $S^2$ which contains no fixed point of $f$, and every point of $S^2$ is a nonwandering point. Then $i(f, U) = 1$ for every component $U$ of $S^2 - M$ which is invariant under $f$. (See § 4 for discussion of the fixed point index $i(f, U)$.)*

*Proof.* Let $U$ be a component of $S^2 - M$ such that $f(U) = U$. $M$ is connected, hence $U$ is simply connected. Also, $\mathrm{Bd}(U)$ is nondegenerate, since $M$ contains no fixed point of $f$. Let $T$ be a $C$-transformation of $U$ onto the open unit disk $D$. Extend $TfT^{-1}$ to a homeomorphism $h$ of $\mathrm{Cl}(D)$ onto itself. Since $\mathrm{Bd}(U)$ contains no fixed point of $f$, we see by Lemma 1 that $U$ has no fixed prime ends. Hence $h$ has no fixed points on $\mathrm{Bd}(D)$. Hence $i(h, D) = 1$ by the Remark, § 4.

We would like to conclude from Axiom A5 of § 4 that $i(f, U) = 1$. However, $D$ and $U$ are not compact polyhedra. We overcome this difficulty as follows: let $X$ be an open 2-cell which contains the fixed points of $f$ in $U$ and whose closure is contained in $U$. Let $Y$ be a closed 2-cell in $U$ containing $\mathrm{Cl}(X) \cup f(\mathrm{Cl}(X))$. Let $r_1$: $\mathrm{Cl}(D) \to T(Y)$, and $r_2$: $S^2 \to Y$ be retractions. Since $T(X)$ contains all fixed points of $h$, we have:

$$\begin{aligned}
1 = i(h, D) &= i(h, T(X)) && \text{by Axiom A3} \\
&= i(Tr_2 f T^{-1} r_1, T(X)) && \text{by A1} \\
&= i(f T^{-1} r_1 T r_2, X) && \text{by A5} \\
&= i(f, X) && \text{by A1} \\
&= i(f, U) && \text{by A3}\,.
\end{aligned}$$

The proof of Lemma 2 is complete.

## 6. Fixed point sets of weakly almost periodic homeomorphisms on $S^2$.

THEOREM 3. *Suppose $f$: $S^2 \to S^2$ is a w.a.p. orientation-preserving homeomorphism. Then either $f$ is the identity or $f$ has exactly two fixed points.*

*Proof.* Let $\mathrm{Fix}(f)$ denote the set of fixed points of $f$. Assume $\mathrm{Fix}(f) \neq S^2$. Since $f$ is orientation-preserving it is easily seen that $f$ leaves every component of $S^2 - \mathrm{Fix}(f)$ invariant, and so we may select an arc $A$ in one of these components such that $f(A) \cap A \neq \phi$. Denote by $M$ the union of all orbit closures which meet $A$. $M$ is closed, since $f$ is w.a.p.; $M$ contains no fixed point of $f$; and $M$ is connected since $M$ is the union of the connected set

$$\bigcup_{n=-\infty}^{\infty} f^n(A)$$

and limit points of this set.

Since $M$ and $\mathrm{Fix}(f)$ are disjoint closed sets, we see that $\mathrm{Fix}(f)$ is contained in a finite number $U_1, \cdots, U_s$ of components of $S^2 - M$. By Axioms A3, A4, and Fact 2 of § 4, we have

$$2 = L(f) = i(f, S^2) = \sum_{j=1}^{s} i(f, U_j)\,.$$

But by Lemma 2, $i(f, U_j) = 1$, $1 \leqq j \leqq s$. Hence $s = 2$.

It remains to show that $\mathrm{Fix}(f) \cap U_j$, $j = 1, 2$, is a single point.

Let $U$ be the component of $U_1 - \mathrm{Fix}(f)$ with $\mathrm{Bd}(U_1) \subset \mathrm{Bd}(U)$. Since $\mathrm{Bd}(U_1)$ and $\mathrm{Fix}(f)$ are disjoint closed sets, we see that $\mathrm{Bd}(U) - \mathrm{Bd}(U_1)$ is a closed nonempty subset of $\mathrm{Fix}(f)$.

*Case* 1. $\mathrm{Bd}(U) - \mathrm{Bd}(U_1)$ has more than one component. Then by [16, Corollary 3.11, p. 109], there is a simple closed curve $J$ in $U$ which separates $\mathrm{Bd}(U) - \mathrm{Bd}(U_1)$. Let $B$ be an arc with one endpoint on $\mathrm{Bd}(U_1)$, the other on $J$, and contained in $U$ except for one endpoint. Then $\mathrm{Bd}(U_1) \cup J \cup B$ is connected, and

$$f(\mathrm{Bd}(U_1) \cup J \cup B) \cap (\mathrm{Bd}(U_1) \cup J \cup B) \neq \phi .$$

Thus if we denote by $N$ the union of all orbit closures which intersect $\mathrm{Bd}(U_1) \cup J \cup B$, we see that $N$ is an invariant continuum which contains no fixed point of $f$ (this follows similarly to the case of $M$ above). Let $V_1, \cdots, V_t$ be the (finite) number of components of $S^2 - N$ such that $\mathrm{Fix}(f) \cap V_j \neq \phi$ and $V_j \subset U_1$, $1 \leqq j \leqq t$. By Lemma 2, $i(f, V_j) = 1$, $1 \leqq j \leqq t$. By Axiom A3,

$$1 = i(f, U_1) = \sum_{j=1}^{t} i(f, V_j) = t .$$

But $J$ separates two points of $\mathrm{Fix}(f) \cap U_1$, hence $t > 1$. This contradiction shows that Case 1 cannot occur.

*Case* 2. $\mathrm{Bd}(U) - \mathrm{Bd}(U_1)$ is connected. The proof will be complete if we show that $\mathrm{Bd}(U) - \mathrm{Bd}(U_1)$ is a single point. We assume that $\mathrm{Bd}(U) - \mathrm{Bd}(U_1)$ is a nondegenerate continuum and derive a contradiction.

Assuming $\mathrm{Bd}(U) - \mathrm{Bd}(U_1)$ is a nondegenerate continuum we establish

*Claim* 1. There is a simply connected invariant domain $C_v$ containing two endcuts $A$ and $B$ such that the endpoint of $B$ on $\mathrm{Bd}(C_v)$ is not a fixed point of $f$, and the endpoint of $A$ on $\mathrm{Bd}(C_v)$ is a fixed point of $f$ which is not a limit point of $\mathrm{Bd}(C_v) - \mathrm{Fix}(f)$.

Let $Q$ be a crosscut in $U$ both of whose endpoints lie on

$$\mathrm{Bd}(U) - \mathrm{Bd}(U_1) .$$

Let $V$ be the component of $U - Q$ whose boundary does not intersect $\mathrm{Bd}(U_1)$, [15, (5.3), p. 6]. $V$ is a component of

$$S^2 - ((\mathrm{Bd}(U_1) - \mathrm{Bd}(U)) \cup Q) .$$

Let $p$ be a point of $\mathrm{Bd}(V) - \mathrm{Cl}(Q)$. Note that $p$ is a fixed point of $f$.

Denote by $L$ the union of all orbit closures which intersect $\mathrm{Cl}(Q)$. $L$ is a continuum. $p$ is not a limit point of $L$ so there is a connected neighborhood 0 of $p$ which misses $L$. Let $A$ be an endcut of $V$ which is contained in 0. Let $C_v$ be the component of

$$S^2 - ((\mathrm{Bd}(U) - \mathrm{Bd}(U_1)) \cup L)$$

which contains the endcut $A$. The endpoint of $A$ in $\mathrm{Bd}(C_v)$ has a neighborhood 0 which misses $L$, hence $0 \cap \mathrm{Bd}(C_v) \subset \mathrm{Fix}(f)$.

Let $B'$ be an endcut of $V$ with one endpoint $b$ in $C_v$ and the other in the crosscut $Q$. Then the component of $B' \cap C_v$ containing $b$ is the required endcut $B$.

$C_v$ is simply connected because $(\mathrm{Bd}(U) - \mathrm{Bd}(U_1)) \cup L$ is connected, (see [15, (5.3), p. 6] and [10, Theorem 74, p. 217]).

$C_v$ is invariant because (a) $(\mathrm{Bd}(U)) - \mathrm{Bd}(U_1)) \cup L$ is invariant, (b) $\mathrm{Bd}(C_v)$ contains a continuum of fixed points of $f$, and (c) $f$ is orientation-preserving, (for further details see proof of Claim 2 below). The proof of Claim 1 is complete.

*Claim* 2. The prime end $E$ of $C_v$ determined by the endcut $A$ is a fixed prime end of $f$.

Let $S_1, S_2, \cdots$ be a chain of crosscuts converging to the endpoint $s$ of $A$ and defining the prime end $E$. Since $s$ is not a limit point of $\mathrm{Bd}(C_v) - \mathrm{Fix}(f)$, we may assume that the endpoints of $S_i$ are fixed points of $f$ for every $i$, $i = 1, 2, \cdots$. We also may assume that every crosscut $S_i$ intersects $A$. From the crosscut $S_i$ and the endcut $A$ we may construct an open triod $T_i$ (see § 2 for definition) whose feet are fixed points of $f$. Since $f$ is orientation-preserving, we see easily that $f(T_i) \cap T_i \neq \phi$. (Hence $f(C_v) \cap C_v \neq \phi$, and since $(\mathrm{Bd}(U) - \mathrm{Bd}(U_1)) \cup L$ is invariant, we have $f(C_v) = C_v$.)

Since $f(T_i) \cap T_i \neq \phi$ for $i = 1, 2, \cdots$, we see that $\{S_i\}$ and $\{f(S_i)\}$ are equivalent chains, hence $E$ is a fixed prime end of $f$. The proof of Claim 2 is complete.

Let $T$ be a $C$-transformation of $C_v$ onto the open unit disk $D$. Extend the homeomorphism $TfT^{-1}\colon D \to D$ to a homeomorphism $h$ of the closed unit disk onto itself. $h$ is orientation-preserving, since $f$ is.

By Claim 2, there is a fixed prime end of $C_v$; hence $h$ has a fixed point on $\mathrm{Bd}(D)$. But then, since $h$ is orientation-preserving, every point of $\mathrm{Bd}(D)$ is either a fixed point of $h$ or converges to a fixed point under positive iterates of $h$ [2, Lemma 14].

Consider the endcut $B$ of Claim 1. The endpoint of $B$ on $\mathrm{Bd}(C_v)$

is not fixed by $f$, but this endpoint is a principal point of the prime end $F$ determined by $B$. Hence, by Lemma 1, $F$ is not a fixed prime end. Hence, if $e$ is the endpoint of $T(B)$ on $\mathrm{Bd}(D)$, $e$ is not a fixed point of $h$. But then, there is a fixed point $m$ of $h$ on $\mathrm{Bd}(D)$ such that $\{h^n(e)\}_{n=0}^{\infty} \to m$. If $M$ is the prime end of $C_v$ corresponding to the point $m$, then by Lemma 1, every principal point of $M$ is a fixed point of $f$.

Let $X_1, X_2, \cdots$ be a chain of crosscuts of $C_v$ defining the prime end $M$. We claim that for large $j$, $T(X_j)$ intersects the orbit under $h$ of $T(B)$. To see this we proceed as follows. Let $b$ be the endpoint of $B$ in $C_v$. Then the orbit closure of $b$ is contained in $C_v$; therefore, the orbit closure of $T(b)$ under $h$ is contained in $D$. In particular, $m$ is not a limit point of the orbit of $T(b)$. Hence, for large $j$, the closure of the crosscut $T(X_j)$ separates $m$ and the orbit of $T(b)$ in $\mathrm{Cl}(D)$. But the other endpoint $e$ of $T(B)$ converges to $m$ under positive iterates of $h$, so for large $j$, there is a positive integer $n$ such that $h^n(\mathrm{Cl}(T(B)))$ intersects both components of

$$\mathrm{Cl}(D) - \mathrm{Cl}(T(X_j)) .$$

Hence $h^n(T(B))$ intersects $T(X_j)$, and our claim is established.

Hence, for large $j$, $X_j$ intersects the orbit under $f$ of $\mathrm{Cl}(B)$.

But the chain $X_1, X_2, \cdots$ of crosscuts converges to a principal point $q$ of the prime end $M$. But then $q$ is a fixed point of $f$ which is a limit point of the orbit of $\mathrm{Cl}(B)$. Therefore, the union of all orbit closures which intersect $\mathrm{Cl}(B)$ is not a closed set. This contradicts the fact that $f$ is w.a.p.

This final contradiction establishes that $\mathrm{Bd}(U) - \mathrm{Bd}(U_1)$ is a single point. Similarly, $\mathrm{Fix}(f) \cap U_2$ is a single point, and so $f$ has exactly two fixed points. The proof of Theorem 3 is complete.

THEOREM 4. *Suppose $f\colon S^2 \to S^2$ is a w.a.p. orientation-reversing homeomorphism. Then either $f$ has no fixed points, or $f$ is periodic with period* 2.

*Proof.* Suppose $f$ has a fixed point.

*Claim.* $f$ has more than two fixed points.

Suppose the claim is not true. Let $A$ be an arc intersecting no fixed point, such that $A \cap f(A) \neq \phi$. Denote by $M$ the union of all orbit closures which intersect $A$. $M$ is an invariant continuum containing no fixed points of $f$. Let $U$ be a component of $S^2 - M$ containing a fixed point of $f$. Then $f(U) = U$ and $U$ is simply connected with a nondegenerate boundary. Let $T$ be a $C$-transformation of $U$

onto the open unit disk $D$. Extend $TfT^{-1}$ to a homeomorphism $h$ of the closed unit disk onto itself. $h$ is orientation-reversing, since $f$ is. But then $h$ must have two fixed points on $\mathrm{Bd}(D)$, [16, Theorem 7.3, p. 264]. These fixed points correspond to fixed prime ends of $U$. By Lemma 1, the principal points of these prime ends are fixed points of $f$. This contradicts the assumption that $M$ contains no fixed points of $f$. The proof of our claim is complete.

But now consider the homeomorphism $f^2\colon S^2 \to S^2$. $f^2$ is orientation-preserving, w.a.p. [6, Theorem 4.24, p. 34 and Theorem 2.33, p. 17], and by our claim, has more than two fixed points. Hence, by Theorem 3, $f^2 = Id$. The proof of Theorem 4 is complete.

COROLLARY 5. *Suppose $f\colon S^2 \to S^2$ is a w. a. p. orientation-reversing homeomorphism. Then the set of fixed points of $f$ is either empty or is a simple closed curve.*

*Proof.* Follows from Theorem 4 and [3].

## 7. Orbit closures of weakly almost periodic homeomorphisms on $S^2$.

THEOREM 6. *Suppose $f\colon S^2 \to S^2$ is a w.a.p. orientation-preserving homeomorphism which is not periodic. Then every nondegenerate orbit closure is a 1-dimensional homology 1-sphere.*

*Proof.* $f \neq Id$ so by Theorem 3, $f$ has exactly two fixed points. Let $K$ be a nondegenerate orbit closure. We show that $K$ separates the fixed points of $f$. Suppose not. Then there is a simple closed curve $J$ which separates $K$ and the fixed points of $f$, (connect the fixed points by an arc missing $K$, then "enlarge" the arc slightly to obtain a disk whose boundary is $J$). We must have $f(J) \cap J \neq \phi$, since otherwise every point of $J$ would be a wandering point. Denote by $M$ the union of all orbit closures which intersect $J$. Then $M$ is an invariant continuum which separates $K$ and the fixed points of $f$. Let $U$ be a component of $S^2 - M$ which intersects $K$. Since every point of $U$ is a nonwandering point, there is an integer $n$ such that $f^n(U) \cap U \neq \phi$. Since $M$ is invariant, $f^n(U) = U$.

$f^n$ is a w.a.p. orientation-preserving homeomorphism [6, p. 34 and p. 17]. $f$ is not periodic, hence $f^n \neq Id$, hence by Theorem 3, $f^n$ has exactly two fixed points. These fixed points are the original fixed points of $f$, and so the domain $U$ contains no fixed points of $f^n$. But by Lemma 2, $i(f^n, U) = 1$. This contradiction shows that the orbit closure $K$ must separate the fixed points of $f$.

We now show that $K$ is connected. Let $V$ be a component of $S^2 - K$ containing a fixed point of $f$. Let $B$ be the outer boundary of $V$ with respect to the fixed point of $f$ not in $V$, (see § 2 for definitions). $B$ is connected, [10, Theorem 25, p. 176]. And $V$ and the fixed points of $f$ are invariant, hence $B$ is invariant. But $K$ is a minimal invariant set, and $B \subset K$, hence $B = K$.

$K$ is one dimensional, since outer boundaries contain no interior points.

Finally, $S^2 - K$ has exactly two components. For, if there were more than two components, then some component $U$ would contain no fixed point of $f$, and we would arrive at the same contradiction as in proving that $K$ separates the fixed points of $f$.

Thus $K$ is a 1-dimensional homology 1-sphere and the proof of Theorem 6 is complete.

REMARK. [12, Example 1] is an example of a w.a.p. orientation-preserving homeomorphism with an orbit closure which is an homology 1-sphere but not a real 1-sphere.

THEOREM 7. *Suppose $f$: $S^2 \rightarrow S^2$ is a w.a.p. orientation-reversing homeomorphism which is not periodic. Then, with two exceptions, every orbit closure is the union of two disjoint homology* 1-*spheres. The exceptions are* (a) *a period* 2 *orbit, and* (b) *one orbit closure which is an homology* 1-*sphere* (*the "axis of reflection"*).

*Proof.* $f^2$ is a w.a.p., orientation-preserving, nonperiodic homeomorphism. Hence, by Theorems 3 and 6, $f^2$ has two fixed points, and every nondegenerate orbit closure is an homology 1-sphere. The orbit closure under $f$ of a point $x$ is the union of the orbit closure of $x$ under $f^2$ and the orbit closure of $f(x)$ under $f^2$. Thus, the two fixed points of $f^2$ correspond to a period 2 orbit under $f$, and every other orbit closure under $f$ is the union of two homology 1-spheres which are either disjoint or equal. Let $H$ denote the collection of orbit closures under $f$ which are homology 1-spheres. We show that $H$ has exactly one element.

Let $G$ be the decomposition space whose points are orbit closures under $f^2$. Let $w$: $S^2 \rightarrow G$ be the natural decomposition map [16, p. 125]. If $K$ is any nondegenerate orbit closure under $f^2$, then $w(K)$ is a cut point of $G$, since $K$ separates $S^2$, $w$ is an open map, [16, p. 130], and orbit closures are connected. Hence $G$ has exactly two noncut points, (the fixed points of $f^2$), hence $G$ is an arc, [16, p. 54]. Define a map $g$: $G \rightarrow G$ by $g(w(K)) = w(f(K))$ for all orbit closures $K$ of $f^2$. It is easily seen that $g$ is a nontrivial period 2 map of the arc $G$ onto itself. Fixed points of $G$ correspond to elements of the set $H$ defined above. But $g$ has exactly one fixed point [16, p. 264]. The proof of Theorem 7 is complete.

## References

1. R. Brown, *The Lefschetz Fixed Point Theorem*, Scott, Foresman and Co., Glenview, III., 1971.
2. M. Cartwright and J. Littlewood, *Some fixed point theorems*, Annals of Math., **54** (1951), 1-37.
3. S. Eilenberg, *Sur les transformations périodiques de la surface de sphere*, Fund. Math., **22** (1934), 28-41.
4. E. Fadell, *Recent results in the fixed point theory of continuous maps*, Bull. Amer. Math. Soc., **76** (1970), 10-29.
5. W. Gottschalk, *Almost periodic points with respect to transformation semi-groups*, Annals of Math., **47** (1946), 762-766.
6. W. Gottschalk and G. Hedlund, *Topological Dynamics*, Amer. Math. Soc. Coll. Pub., Vol. **36**, Providence, 1955.
7. E. Hemmingsen, *Plane continua admitting nonperiodic autohomeomorphisms with equicontinuous iterates*, Math. Scand., **2** (1954), 119-141.
8. B. de Kerékjártó, *Über die periodischen transformationen der kreisscheibe und der Kugelfläche*, Math. Ann., **80** (1919-1920), 36-38.
9. ———, *Topologische Charakterisierung der linearen abbildungen*, Acta Sci. Math. (Szeged), **6** (1934), 235-262.
10. R. L. Moore, *Foundations of Point Set Topology*, Amer. Math. Soc. Coll. Pub., Vol. **13**, Providence, 1962.
11. M. Ohtsuka, *Dirichlet Problem, Extremal Length, and Prime Ends*, Van Nostrand, New York, 1970.
12. R. Remage, *On minimal sets in the plane*, Proc. Amer. Math. Soc., **13** (1962), 41-47, and "Correction", p. 1000.
13. N. Rutt, *Prime ends and indecomposability*, Bull. Amer, Math. Soc., **41** (1935), 265-273.
14. E. Spanier, *Algebraic Topology*, McGraw-Hill, New York, 1966.
15. H. Ursell and L. Young, *Remarks on the Theory of Prime Ends*, Memoirs of Amer. Math. Soc., No. **3**, Providence, 1951.
16. G. Whyburn, *Analytic Topology*, Amer. Math. Soc., Coll. Pub., Vol. **28**, Providence, 1963.

Received June 12, 1972 and in revised form August 24, 1972.

RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# LIMITS FOR MARTINGALE-LIKE SEQUENCES

ANTHONY G. MUCCI

**The concept of a martingale is generalized in two ways. The first generalization is shown to be equivalent to convergence in probability under certain uniform integrability restrictions. The second generalization yields a martingale convergence theorem.**

## 1. Introduction.

In what follows $\{X_n, \mathfrak{B}_n\}$ is a sequence of integrable random variables and sub-sigma fields on the probability space $(\Omega, \mathfrak{B}, P)$ such that

$$X_n \text{ is } \mathfrak{B}_n\text{-measurable}$$
$$\mathfrak{B}_n \subset \mathfrak{B}_{n+1}$$
$$\mathfrak{B} = \sigma\left(\bigcup_1^\infty \mathfrak{B}_n\right).$$

We call the sequence $\{X_n, \mathfrak{B}_n\}$ an adapted sequence. In [2] Blake defines $\{X_n, \mathfrak{B}_n\}$ as a game which becomes fairer with time provided

$$E(X_n | \mathfrak{B}_m) - X_m \xrightarrow{P} 0 \quad \text{as} \quad n \geqq m \longrightarrow \infty ,$$

i.e., provided, for all $\varepsilon > 0$:

$$\lim_{n>m} P(| E(X_n | \mathfrak{B}_m) - X_m | > \varepsilon) = 0 \quad \text{as} \quad m \longrightarrow \infty .$$

It is proven in [1] that if $\{X_n, \mathfrak{B}_n\}$ becomes fairer with time, and if there exists $Z \in L_1$ with $|X_n| \leqq Z$ for all $n$, then $X_n \xrightarrow{\mathscr{L}_1} X$, some $X \in \mathscr{L}_1$.

In the present paper we will show that $X_n \xrightarrow{\mathscr{L}_1} X$ under the less restrictive assumption that $\{X_n\}$ is uniformly integrable. We will further show that in the presence of uniform integrability $\{X_n, \mathfrak{B}_n\}$ becomes fairer with time if and only if $\{X_n\}$ converges in probability, i.e.,

$$E(X_n | \mathfrak{B}_m) - X_m \xrightarrow{P} 0 \Longleftrightarrow X_n - X_m \xrightarrow{P} 0 .$$

Finally, by using the more restrictive concept that $\{X_n, \mathfrak{B}_n\}$ is a martingale in the limit, namely,

$$\lim_{n \geqq m \to \infty} (E(X_n | \mathfrak{B}_m) - X_m) = 0 \quad \text{a.e.},$$

we will prove (Theorem (2)) a generalization of a standard martingale convergence theorem.

2. PROPOSITION 1. *Let the sequence $\{X_n\}$ be uniformly integrable and assume*

$$\lim_{n\to\infty}\int_A X_n \textit{ exists, all } A\in\bigcup_1^\infty \mathfrak{B}_n .$$

*Then there exists $X\in\mathscr{L}_1$ such that*

$$\lim_{n\to\infty}\int_A X_n=\int_A X\,,\quad \textit{all}\quad A\in\mathfrak{B}\,.$$

*Proof.* Let $A\in\mathfrak{B}$, $\delta>0$. There exists $A_0\in\bigcup_1^\infty\mathfrak{B}_n$ with $P(A\varDelta A_0)\leqq\delta$. This, together with the augument in Neveu [3] (page 117) proves the desired result.

REMARKS. Let $\Omega=[0,1)$ with Lebesgue measure. Let $\mathfrak{B}_n$ be the $\sigma$-field generated by the subintervals $A_{k,n}\equiv[k/2^n,(k+1)/2^n)$, $k=0,1,\cdots,2^n-1$. Set $X_n=\sum_{k=0}^{2^n-1}(-1)^k I_{A_{k,n}}$ where $I_A$ is the indicator function of $A$. Then for any $A\in\bigcup\mathfrak{B}_n$ we have $\lim_{n\to\infty}\int_A X_n=0$. Further, $\{X_n\}$ is uniformly integrable. However, $\{X_n\}$ does not converge in the $\mathscr{L}_1$-norm.

PROPOSITION 2. *Let $\{X_n\}$ be uniformly integrable and assume $\{X_n\}$ becomes fairer with time:*

$$(*)\qquad \lim_{n\geqq m\to\infty}P(|E(X_n|\mathfrak{B}_m)-X_m|>\varepsilon)=0\,.$$

*Then there exists $X\in\mathscr{L}_1$ such that $X_n\xrightarrow{\mathscr{L}_1}X$.*

*Proof.* Let $A\in\mathfrak{B}_m$, $p\geqq q\geqq m$. Then

$$\begin{aligned}\left|\int_A X_p-\int_A X_q\right| &= \left|\int_A E(X_p|\mathfrak{B}_q)-X_q\right|\\ &\leqq \int_{A(|E(X_p|\mathfrak{B}_q)-X_q|>\varepsilon)}|E(X_p|\mathfrak{B}_q)-X_q|+\varepsilon\\ &\leqq 2\sup_k\int_{A(|E(X_p|\mathfrak{B}_q)-X_q|>\varepsilon)}|X_k|+\varepsilon\,.\end{aligned}$$

By uniform integrability and the assumption (*) we see that

$$\lim_{n\to\infty}\int_A X_n\quad\text{converges, all}\quad A\in\bigcup_1^\infty\mathfrak{B}_n\,.$$

By Proposition 1, there exists $X\in\mathscr{L}_1$ with

$$\lim_{n\to\infty}\int_A X_n=\int_A X\,,\quad\text{all}\quad A\in\mathfrak{B}\,.$$

Note that $\{E(X|\mathfrak{B}_n), \mathfrak{B}_n\}$ is a martingale and $E(X|\mathfrak{B}_n) \to X$ both in the $\mathscr{L}_1$ and the almost sure sense (Levy's Theorem). Since

$$\int |X_n - X| \leqq \int |X_n - E(X|\mathfrak{B}_n)| + \int |E(X|\mathfrak{B}_n) - X| ,$$

it will be enough to show $\int |X_n - E(X|\mathfrak{B}_n)| \to 0$. Now

$$\int |X_n - E(X|\mathfrak{B}_n)| = \int_{(X_n \geqq E(X|\mathfrak{B}_n))} (X_n - E(X|\mathfrak{B}_n)) + \int_{(X_n < E(X|\mathfrak{B}_n))} (E(X|\mathfrak{B}_n) - X_n) .$$

Letting $n' \geqq n$ and setting $A = (|E(X_{n'}|\mathfrak{B}_n) - X_n| > \varepsilon)$, we have

$$\begin{aligned}\int_{(X_n \geqq E(X|\mathfrak{B}_n))} (X_n - E(X|\mathfrak{B}_n)) &\leqq \int_A |X_n| + \int_A |X_{n'}| \\ &\quad + \left|\int_{(X_n \geqq E(X_n|\mathfrak{B}_n))} (X_{n'} - X)\right| + \varepsilon \\ &\leqq 2 \sup_k \int_A |X_k| \\ &\quad + \left|\int_{(X_n \geqq E(X|\mathfrak{B}_n)} (X_{n'} - X)\right| + \varepsilon .\end{aligned}$$

By uniform integrability and condition (*), the first integral is small. Letting $n' \to \infty$, the difference in the remaining integral tends to zero. An identical analysis shows

$$\int_{(X_n < E(X|\mathfrak{B}))} (E(X|\mathfrak{B}_n) - X_n) \longrightarrow 0 .$$

REMARKS. Suppose $X_n \xrightarrow{\mathscr{L}_1} X$. Then since

$$\int_A |X_n| \leqq \int |X_n - X| + \int_A |X| ,$$

we see that $\{X_n\}$ is uniformly integrable. Further

$$\begin{aligned}P(|E(X_n|\mathfrak{B}_m) - X_m| > \varepsilon) &\leqq \frac{1}{\varepsilon}\int |E(X_n|\mathfrak{B}_m) - X_m| \\ &\leqq \frac{1}{\varepsilon}\int |X_n - X_m| ,\end{aligned}$$

so $\{X_n, \mathfrak{B}_n\}$ becomes fairer with time. It is shown (Neveu [3], page 52):

$\{X_n\}$ is Cauchy in the $\mathscr{L}_1$ norm $\Longleftrightarrow \{X_n\}$ is uniformly integrable and $\{X_n\}$ is Cauchy in probability.

We tie these results together with Proposition 2 to get

THEOREM 1. *Let* $\{X_n, \mathfrak{B}_n\}$ *be an adapted sequence. Then the following three statements are equivalent:*

(a) *There exists* $X \in \mathscr{L}_1$ *and* $X_n \xrightarrow{\mathscr{L}_1} X$.

(b) $\{X_n\}$ *is uniformly integrable and* $E(X_n \mid \mathfrak{B}_m) - X_m \xrightarrow{P} 0$.

(c) $\{X_n\}$ *is uniformly integrable and* $X_n - X_m \xrightarrow{P} 0$.

COROLLARY 1. *Let the adapted sequence* $\{X_n, \mathfrak{B}_n\}$ *be uniformly integrable. Then*

$$E(X_n \mid \mathfrak{B}_m) - X_m \xrightarrow{p} 0 \Longleftrightarrow X_n - X_m \xrightarrow{p} 0 .$$

REMARKS. In the absence of uniform integrability we have neither implication. Consider the following two examples:

(1) Set $X_n = \sum_1^n y_k$ where $\{y_k\}$ is a sequence of independent identically distributed random variables with zero means. Set $\mathfrak{B}_n = \sigma(y_1, y_2, \cdots, y_n)$. Clearly $\{X_n, \mathfrak{B}_n\}$ is a martingale, so $E(X_n \mid \mathfrak{B}_m) - X_m \xrightarrow{P} 0$. But, if, for instance

$$y_k = \begin{cases} 1 & \text{with probability } \dfrac{1}{2} \\ -1 & \text{with probability } \dfrac{1}{2} \end{cases},$$

then

$$\begin{aligned} P(|X_n - X_m| \geqq 1) &= P\left(\left|\sum_1^{n-m} y_k\right| \geqq 1\right) \\ &= 1 - P\left(\sum_1^{n-m} y_k = 0\right) \sim 1 - \frac{c}{\sqrt{n-m}} \not\rightarrow 0 , \end{aligned}$$

so $X_n - X_m \overset{P}{\not\rightarrow} 0$.

(2) Let $\{y_k\}$ independent where $P(y_k = k^2) = 1/k^2$ and $P(y_k = 0) = 1 - 1/k^2$.

Then, setting $X_n = \sum_1^n y_k$ we have

$$|E(X_n \mid \mathfrak{B}_m) - X_m| = E \sum_{m+1}^{n} y_k \geqq 1$$

while

$$\begin{aligned} P(|X_n - X_m| \geqq \varepsilon) &= P\left(\sum_{m+1}^{n} y_k \geqq \varepsilon\right) = P\left(\bigcup_{m+1}^{n} (y_k \geqq \varepsilon)\right) \\ &\leqq \sum_{m+1}^{n} P(y_k \geqq \varepsilon) = \sum_{m+1}^{n} \frac{1}{k^2} \longrightarrow 0 , \end{aligned}$$

so in this case $X_n - X_m \xrightarrow{p} 0$ while $E(X_n \mid \mathfrak{B}_m) - X_m \overset{p}{\not\rightarrow} 0$.

Recall now the definition that $\{X_n, \mathfrak{B}_n\}$ be a martingale in the limit, namely:

$$(**) \qquad E(X_n \mid \mathfrak{B}_m) - X_m \longrightarrow 0 \text{ almost everywhere.}$$

THEOREM 2. *Let the adapted sequence $\{X_n, \mathfrak{B}_n\}$ be uniformly integrable and a martingale in the limit. Then there exists $X \in \mathscr{L}_1$ such that*

$$X_n \longrightarrow X \textit{ almost everywhere and in the } \mathscr{L}_1\textit{-norm.}$$

*Proof.* Clearly, $\{X_n, \mathfrak{B}_n\}$ becomes fairer with time, so from Theorem 1 there exists $X \in \mathscr{L}_1$ with $X_n \xrightarrow{\mathscr{L}_1} X$. Now, for an arbitrary subsequence $\{n'\}$,

$$|X_m - X| \leqq |X_m - E(X_{n'} \mid \mathfrak{B}_m)| + |E(X_{n'} - X \mid \mathfrak{B}_m)| + |E(X \mid \mathfrak{B}_m) - X| .$$

By Levy's theorem, the third term is less than $\varepsilon/3$ for large enough $m$. The first term is also bounded by $\varepsilon/3$ for large $m, n'$ since $\{X_n, \mathfrak{B}_n\}$ is a martingale in the limit. We must now show that the second term is small. Note first that for an arbitrary $\sigma$-field $\mathscr{A}$ we have

$$E(X_n \mid \mathscr{A}) \xrightarrow{\mathscr{L}_1} E(X \mid \mathscr{A} .$$

Now start with the $\sigma$-field $\mathfrak{B}_1$ and note that the convergence $E(X_n \mid \mathfrak{B}_1) \xrightarrow{\mathscr{L}_1} E(X \mid \mathfrak{B}_1)$ implies the existence of subsequence $\{n_1\} \subset \{n\}$ with $E(X_{n_1} \mid \mathfrak{B}_1) \to E(X \mid \mathfrak{B}_1)$ almost everywhere. Continuing, we have $E(X_{n_1} \mid \mathfrak{B}_2) \xrightarrow{\mathscr{L}_1} E(X \mid \mathfrak{B}_2)$, and we can extract $\{n_2\} \subset \{n_1\}$ with $E(X_{n_2} \mid \mathfrak{B}_2) \to E(X \mid \mathfrak{B}_2)$ almost everywhere. Thus, there exists a subsequence $\{\bar{n}\} \subset \{n\}$ with $E(X_{\bar{n}} \mid \mathfrak{B}_m) \to E(X \mid \mathfrak{B}_m)$ a.e. for all $m$, namely the diagonal subsequence. Now choose $\{n'\}$ as a subsequence of $\{\bar{n}\}$, and we can bound the second term above by $\varepsilon/3$.

*Applications.* 1. Let $\{y_k\}$ be a sequence of independent random variables such that

$$\lim_{\substack{m \to \infty \\ n \to \infty}} \int \left| \sum_m^n y_k \right| = 0 .$$

Then $\sum_1^\infty y_k$ exists a.e. and in the $\mathscr{L}_1$-norm.

*Proof.* Set $S_n = \sum_1^n y_k$. Then

$$\int_A |S_n| \leqq \int_A |S_m| + \int \left| \sum_{m+1}^n y_k \right| ,$$

so it is clear that $\{S_n\}$ is uniformly integrable. Further, setting $\mathfrak{B}_n = \sigma(y_1, y_2, \cdots, y_n)$, we have

$$|E(S_n|\mathfrak{B}_m) - S_m| = \left|\int \sum_{m+1}^{n} y_k\right| \leqq \int \left|\sum_{m+1}^{n} y_k\right| ,$$

so $\{S_n, \mathfrak{B}_n\}$ is a uniformly integrable martingale in the limit.

2. Let $\{X_n, \mathfrak{B}_n\}$ be an adapted uniformly integrable sequence with $|E(X_{n+1}|\mathfrak{B}_n) - X_n| \leqq c_n$ where $\{c_n\}$ is a sequence of constants with $\sum_1^\infty c_n < \infty$. Then there exists $x \in \mathscr{L}_1$ with $X_n \rightarrow X$ almost everywhere and in the $\mathscr{L}_1$-norm.

*Proof.* We have

$$\begin{aligned} E(X_n|\mathfrak{B}_m) - X_m &= \sum_{m}^{n-1} E(X_{k+1} - X_k|\mathfrak{B}_m) \\ &= \sum_{m}^{n-1} E(E_{k+1} - X_k|\mathfrak{B}_k)|\mathfrak{B}_m). \end{aligned}$$

Thus

$$|E(X_n|\mathfrak{B}_m) - X_m| \leqq \sum_{m}^{n-1} c_k .$$

*Editorial note.* See also R. Subramanian, "On a generalization of Martingales due to Blake," Pacific J. Math., 48, No. 1, (1973), 275–278.

## REFERENCES

1. L. Blake, *A note concerning a class of games which become fairer with time,* to appear, Glasgow Math. J.
2. ———, *A generalization of martingales and two consequent convergence theorems,* Pacific J. Math., **35**, No. 2, (1970).
3. J. Neveu, *Mathematical Foundations of the Calculus of Probability,* Holden-Day, (1965).

Received June 23, 1972 and in revised form October 2, 1972. This research was supported by a University of Maryland General Research Board Grant, Summer 1972.

UNIVERSITY OF MARYLAND

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# RELATIONALLY INDUCED SEMIGROUPS

EUGENE M. NORRIS

**This paper gives sufficient conditions, of a relation-theoretic nature, in order that a quotient of the state space of a recursion (or topological machine) be a topological semigroup iseomorphic to the endomorphism semigroup of the recursion, generalizing recent function-theoretic results.**

Relations. By a *relation* $R$ from a set $A$ to a set $B$, we mean that $R$ is a subset of $A \times B$. If $A$ and $B$ are topological spaces, we say that $R$ is *closed* to mean that it is a closed subset of the product space. If $R$ is a relation from $A$ to $B$ and $S$ is a relation from $B$ to $C$, their *composition* is the relation $S \circ R$ from $A$ to $C$ defined by $(a, c) \in S \circ R$ if and only if there is some $b \in B$ with $(a, b) \in R$ and $(b, c) \in S$. This is contrary to the notation in [1], but agrees with the usual (non-algebraist's) notation for the composition of functions. The *inverse* of a relation $R$ is the relation $R^{-1}$ defined by $(b, a) \in R^{-1}$ if and only if $(a, b) \in R$. A relation from $A$ to $A$ is reflexive if $R$ contains $\Delta_A = \{(a, a): a \in A\}$, *symmetric* if $R^{-1} \subseteqq R$ (whence follows $R^{-1} = R$), and *transitive* if $R \circ R \subseteqq R$. $R$ is an *equivalence relation* if it is reflexive, symmetric, and transitive. For any relation $R$ from $A$ to $B$ and any subsets $A' \subseteqq A$, $B' \subseteqq B$, $A'R$ denotes the set $\{b \in B: (a, b) \in R$ for some $a \in A'\}$; $RB'$ is then defined to be the set $B'R^{-1}$. We write $aR$ rather than $\{a\}R$ and $Rb$ for $R\{b\}$, for simplicity's sake. It is known that if $A'$ is compact and $R$ is closed then $A'R$ is closed; if $A$, $B$, and $C$ are all compact Hausdorff spaces and $R$ and $S$ are closed relations from $A$ to $B$ and $B$ to $C$ respectively then $S \circ R$ is also closed. It is also known that if $A$ is compact and $R$ is a closed equivalence on $A$ then the quotient space $A/R = \{aR: a \in A\}$ is compact Hausdorff. See Kelley [3] for topological details.

After Riguet [5, 6], a relation $R$ from $A$ to $B$ is called *difunctional* if $R \circ R^{-1} \circ R \subseteqq R$; we observe that any function is difunctional and any symmetric, transitive relation is difunctional; in elementary geometry, the relation of orthogonality is difunctional, as Riguet noted. We use Riguet's 1950 results freely [6] and note in particular that if $R$ is a difunctional relation from $A$ to $B$ satisfying $A = RB$ and $B = AR$, then $R^{-1} \circ R$ and $R \circ R^{-1}$ are equivalence relations on $A$ and $B$, respectively, closed if $R$ is closed and $A$ and $B$ are compact Hausdorff. Furthermore, $A/(R^{-1} \circ R) = \{Rb: b \in B\}$ and $B/(R \circ R^{-1}) = \{aR: a \in A\}$. For any difunctional relation $R$, the *slices* $aR$ and $a'R$ either coincide or are disjoint, a property well-known for equivalence relations; the same property holds for slices $Rb, Rb'$, since $R^{-1}$ is difunctional if and only if $R$ is difunctional. In fact, this property of slices characterizes difunctional relations. Unfortunately, the composition of difunctional relations need not be difunctional.

**Recursions.** A *recursion* is a triple $(T, X, \cdot)$, where $T$ and $X$ are spaces and $T \times X \xrightarrow{\cdot} X$ is a continuous binary operation, the value $t \cdot x$ of which at the point $(t, x)$ is usually denoted by juxtaposition, unless emphasis seems wise. For $T' \subseteqq T$ and $X' \subseteqq X$, we write $T'X'$ (or occasionally $T' \cdot X'$) to denote the set $\{tx: t \in T' \text{ and } x \in X'\}$. We frequently avoid the use of curly brackets, writing $Tx$ for $T\{x\}$ and so forth. In particular, if $R \subseteqq T \times X$ and $t, z$ are elements of $T$, then $t(zR) = \{t\} \cdot (zR)$, the translate of the slice $zR$. A recursion is *c.o.d.* if both spaces are compact Hausdorff or both are discrete.

For the sake of completeness we state below an easily established folkloric lemma that A. D. Wallace attribute to G. E. Schweigert [7], and a generalization, the Induced Function Theorem (IFT for short), proved in [1]. The lemma is frequently used in what follows.

LEMMA 0. *If $A$, $B$, and $C$ are all compact or all discrete spaces, if $f: A \rightarrow B$ and $g: A \rightarrow C$ are continuous functions with $f$ surjective and if the condition $f(a) = f(a')$ implies $g(a) = g(a')$ for all $a, a'$ holds then there is a unique continuous function $h: B \rightarrow C$, satisfying $h(f(a)) = g(a)$ for all $a$ in $A$.*

*Induced Function Theorem. Let $A$ and $B$ be both compact Hausdorff or both discrete spaces, $R \subseteqq A \times B$ a closed relation from $A$ to $B$, and $E$ and $F$ closed equivalence relations on $A$ and $B$, respectively. If $A = RB$ and $R \circ E \circ R^{-1} \subseteqq F$ then there is a unique continuous function $h$ making the following diagram of projection and quotient functions analytic:*

$$\begin{array}{ccccc} A & \longleftarrow & R & \longrightarrow & B \\ \downarrow & & & & \downarrow \\ A/E & & \xrightarrow{\quad h \quad} & & B/F \end{array} .$$

*Furthermore, if in addition to the previous hypothesis $B = AR$ and $R^{-1} \circ F \circ R \subseteqq E$, then $h$ is a homeomorphism.*

**Results.**

THEOREM 1. *Suppose $(T, X, \cdot)$ is a c.o.d. recursion and $R \subseteqq T \times X$ is a closed difunctional relation satisfying, for all $t', t'', t$ and $s \in T$,*

(1) $t'R = t''R \Rightarrow t'(tR) = t''(tR)$

(2) $tR = t'(t''R) \Rightarrow t(sR) = t'(t''(sR))$

(3) $T = RX$ *and* $X = TR$

(4) *for each* $t, t'$ *in* $T$ *there is some* $t''$ *in* $T$ *with* $t(t'R) = t''R$. *Then* $X/(R \circ R^{-1})$ *is a topological semigroup with multiplication* $*$ *satisfying* $tR^*t'R = t(t'R)$ identically.

*Proof.* From difunctionality and hypothesis (3), $R^{-1} \circ R$ and $R \circ R^{-1}$ are equivalence relations on $T$ and on $X$, respectively, and are closed if $T$ and $X$ are compact. The Induced Function Theorem implies that there is a unique homeomorphism $h$ making the following diagram of projection and quotient maps analytic.

$$\begin{array}{ccc} T & \longleftarrow R \longrightarrow & X \\ \downarrow & & \downarrow \\ T/(R^{-1} \circ R) & \longrightarrow & X/(R \circ R^{-1}) \end{array}$$

In the following diagram,

$$\begin{array}{ccc} T \times X & \xrightarrow{\cdot} & X \\ {\scriptstyle hp \times q}\downarrow & & \downarrow{\scriptstyle q} \\ X/(R \circ R^{-1}) \times X/(R \circ R^{-1}) & \longrightarrow & X/(R \circ R^{-1}) \end{array}$$

we note that $(t, x) \in R$ iff $tR = q(x)$ and $Rx = p(t)$ iff $h(Rx) = tR$. If $(t, x)$ and $(t', x')$ satisfy $[hp \times q]\ (t, x) = [hp \times q]\ (t', x')$ then $h(p(t)) = h(p(t'))$ and $q(x) = q(x')$, so that $tR = t'R$. If $t'' \in Rx$ then $x \in t''R$, hence $tx \in t(t''R) = t'(t''R)$ by hypothesis (1). We also have $t'x' \in t'(t''R)$ since $Rx = Rx'$. Hypothesis (4) allows us to conclude that $(tx, t'x') \in R \circ R^{-1}$, i.e., $q(tx) = q(t'x')$. Hence Lemma 0 applies to give a unique continuous function$^*$ making the diagram analytic. We observe that $tR^*q(x) = tq(x)$ for all $t \in T$ and all $x \in X$. Now$^*$ is associative, for if $t, t't'' \in T$, then there is some $s \in T$ such that $t(t'R) = sR$, and hence $(tR^*t'R)^*t''R = t(t'R)^*t''R = sR^*t''R = s(t''R) = t(t'(t''R)) = tR^*t'(t''R) = tR^*(t'R^*t''R)$, using hypothesis (2).

THEOREM 2. *Suppose* $(T, X, \cdot)$ *is a c.o.d. recursion and* $R \subseteq T \times X$ *is a closed difunctional relation satisfying*

(1) $T = RX$ and $X = TR$

(2) *the set* $Z = \{z \in T\colon tR = t'R \Rightarrow t(zR) = t'(zR)\}$ *is not empty*

(3) *for each* $t, t' \in T$ *there is some* $t'' \in T$ *with* $t(t'R) = t''R$

(4) *if* $tR = t'(t''R)$ *then for any* $z \in Z$, $t(zR) = t'(t''(zR))$.

*Then* $\{zR\colon z \in Z\}$ *is a topological semigroup in the quotient topology with multiplication$^*$ satisfying* $zR^*z'R = z(z'R)$ *for all* $z, z' \in Z$.

*Proof.* For simplicity, let $\bar{Z} = \{zR\colon z \in Z\}$ be the subspace of the quotient space $A/(R \circ R^{-1})$. We dispose topological considerations first.

One verifies easily that if $T$ and $X$ are compact, then $Z$ is closed, and it follows by standard results that $ZR$ and finally $\bar{Z}$ are compact. Of course, if $T$ and $X$ are discrete, so is $\bar{Z}$.

On the algebraic side, we observe that $Z \cdot ZR \subseteqq ZR$, for if $z, z' \in Z$ and $tR = z(z'R)$, then it will be seen that $t \in Z$ (such $t$ exists by hypothesis (3)). To this end, suppose that $t'R = t''R$, and let $t'(zR) = sR$ to infer that $t'(tR) = t'(z(z'R)) = s(z'R)$, by hypothesis (4). Since $z \in Z$, then $t'(zR) = t''(zR)$ and hence $sR = t''(zR)$; it then follows from hypothesis (4) that $s(z'R) = t''(z(z'R))$, so that $t'(tR) = t''(tR)$, implying that $t \in Z$. Hence $Z \cdot ZR \subseteqq ZR$.

If $x$ and $x'$ are points in $ZR$ satisfying $(x, x') \in R \circ R^{-1}$, then $(zx, zx') \in R \circ R^{-1}$ also, and hence we may infer from Lemma 0 that the function $Z \times \bar{Z} \xrightarrow{*} \bar{Z}$ given by $z{*}z'R = z \cdot (z'R)$ is continuous.

Finally, if $R'$ is the relation from $Z$ to $\bar{Z}$ defined by $(z, z'R) \in R'$ if $\{z\} \times z'R \subseteqq R$, then we can easily see that $R'$ is closed and difunctional, so that $R'$ and the compact or discrete recursion $(Z, \bar{Z}, *)$ satisfy the hypothesis of Theorem 1. Theorem 2 now follows.

Representation. Assuming the hypothesis of Theorem 2, let $S$ be the semigroup (with compact open topology) of all continuous functions from the quotient space $A/(R \circ R^{-1})$ into itself, and let *end* denote the subsemigroup of $S$ defined by $f \in$ *end* if $t \cdot f(\bar{x}) = f(t \cdot \bar{x})$ for all $t \in T$ and all $\bar{x}$ in $X/(R \circ R^{-1})$. The function $F \colon T \to S$, given by $F_t(t'R) = t' \cdot (tR)$, is easily seen to be continuous and maps $Z$ into *end*; let $F'$ denote the restriction of $F$ to $Z$. In a similar way, the map $G \colon Z \to ZR/(R \circ R^{-1})$ given by $G(z) = zR$ is a continuous surjection. Lemma 0 is seen easily to apply, giving a continuous function $H \colon ZR/(R \circ R^{-1}) \to$ *end* satisfying $H \circ G = F$, from which we see that for any $z \in Z$ and any $t \in T$, $[H(zR)]\,(tR) = t(zR)$. Routine computation, using hypothesis (3) and (4), shows that $H(zR{*}z'R) = H(z'R) \circ H(zR)$, so that $H$ is an anti-homorphism.

THEOREM 3. *If, in addition to the hypothesis of Theorem 2, for some $z_0 \in Z$ and all $t \in T$ it is the case that $t(z_0R) = tR = z_0(tR)$, then $H$ is an anti-iseomorphism and $ZR/(R \circ R^{-1})$ is a monoid with $z_0R$ its identity; furthermore, the set $z_0R$ is a set of generators for $X$, i.e., $T(z_0R) = X$.*

*Proof.* That $z_0R$ generates $X$ is clear from the equations $t(z_0R) = tR$ and $TR = X$. That $z_0R$ is the identity follows from the fact that for any $z \in Z$, $zR{*}z_0R = z(z_0R) = zR = z_0(zR) = z_0R{*}zR$. If $H(zR) = H(z'R)$ then $zR = z_0(zR) = z_0(z'R) = z'R$, so that $H$ is injective. To see that $H$ is also surjective, let $f \in$ *end*, and suppose $f(z_0R) = t_0R$;

we will see that $t_0 \in Z$. To see this suppose $t'R = t''R$ and compute: $t'(t_0R) = t'f(z_0R) = f(t'(z_0R)) = f(t'R)$; similarly, $t''(t_0R) = f(t''R)$; it follows that $t_0 \in Z$. Now for any $t \in T$, we see that $f(tR) = f(t(z_0R)) = tf(z_0R) = t(t_0R) = [H(t_0R)](tR)$, implying that $H$ is surjective.

REMARKS. Theorem 2 obviously generalizes Theorem 1 and also contains a previous result of the author [4]. When $R$ is a continuous function from $T$ onto $X$ it is a closed, difunctional relation and $R \circ R^{-1} = \Delta_X$, so that $X/(R \circ R^{-1})$ is homeomorphic to $X$, and the set $ZR$ is just the image of $Z$; $R$ is surjective just in case $X = TR$. Hence Theorem 1 generalizes the theorem of [7] and Theorem 2, the theorem of [8], which in turn elegantly generalize theorems of Aczel-Wallace, Hosszu, Barnes, Fleck, Weeg, Oehmke *et. al.* (see [8] for references). Other applications will be announced elsewhere. In view of recent results of Fay [2], the present work allows one to induce semigroups "in" the objects of many categories. The details of this extension will be left for another time.

## REFERENCES

1. A. R. Bednarek and A. D. Wallace, *A relation-theoretic result in topological algebra*, Mathematical Systems Theory, 1 No. 3 (1967), 217-224.

2. T. H. Fay, *The Induced Morphism Theorem: A Result in Categorical Relation Theory*, to appear.

3. J. L. Kelley, *General Topology*, D. Van Nostrand and Company, Princeton, New Jersey, 1955.

4. Eugene M. Norris, *More on externally induced operations*, Notices Amer. Math. Soc., # 689-A15, October, 1971.

5. J. Riguet, *Relations binaires, Fermetures, correspondances de Galois*, Bull. Soc. Math. France, **76** (1948), 114-155.

6. ———, *Quelque propriêtês des relations difonctionelles*, C. R. Acad. Sci. Paris, **230** (1950), 1999-2000.

7. A. D. Wallace, *Recursions with semigroup state spaces*, Revue Roumaine de Mathématiques Pures et Appliquées, Tome XII, No. **9** (1967), 1411-1415.

8. ———, *Externally Induced Operations*, Jahresbericht der Deutschen Mathematiker-Vereinigung, Bd. 73, Heft 1 (1971), 48-52.

Received June 6, 1972. This research was supported initially by the Center for Informatics Research of the University of Florida. The author is grateful to A. R. Bednarek for alerting him to Riguet's work.

UNIVERSITY OF SOUTH CAROLINA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# A COMPARISON OF $c$-DENSITY AND $k$-DENSITY

ARTHUR E. OLSON, JR.

**In this paper a comparison is made between $c$-density and $k$-density in the general setting of Freedman density spaces in additive number theory. The comparison is motivated by the following question of Freedman: Does there exist a density space and a set such that the $c$-density of that set is positive and the $k$-density is zero? The answer is yes. More generally, there exists a density space such that for any two real numbers $\rho_1$ and $\rho_2$ with $0 \leqq \rho_1 \leqq \rho_2 < 1$, a set can be constructed such that the $k$-density of the set is $\rho_1$ while the $c$-density is $\rho_2$.**

Let $S$ be any nonempty subset of an abelian group $G$ with binary operation $+$ and identity element 0. We define a relation $<$ on $S$ by saying $y < x$ whenever $x - y \in S\backslash\{0\}$. The set $S$ is called an *s-set* whenever the following conditions hold:

(s.1) $0 \in S$

(s.2) $S\backslash\{0\} \neq \phi$

(s.3) $S\backslash\{0\}$ is closed with respect to $+$.

(s.4) $L(x) = \{y \mid y \in S, y < x \text{ or } y = x\}$ is finite for each $x \in S\backslash\{0\}$.

Corresponding to each $x \in S\backslash\{0\}$, let $H(x)$ be a subset of $S$ satisfying the following three conditions:

(c.1) $\{0, x\} \subseteqq H(x)$

(c.2) $H(x) \subseteqq L(x)$

(c.3) if $y \in H(x)\backslash\{0\}$, then $H(y) \subseteqq H(x)$.

Let $\mathscr{F}(H) = \{F \mid F \subseteqq S, F$ finite, $0 \in F, F\backslash\{0\} \neq \phi, x \in F\backslash\{0\}$ implies $H(x) \subseteqq F\}$.

Then $\mathscr{F}(H)$ is said to be a *fundamental family on S*.

Freedman [1] calls the ordered pair $(S, \mathscr{F}(H))$ a *density space* whenever $S$ is an *s-set* and $\mathscr{F}(H)$ is a fundamental family on $S$.

For any two sets $X, D \subseteqq S$ with $D$ finite, let $X(D)$ denote the number of nonzero elements in $X \cap D$.

DEFINITION. The *k-density* of a set $A \subseteqq S$ with respect to $\mathscr{F}(H)$, written $\alpha_k$, is

$$\alpha_k = \text{glb}\left\{\frac{A(F)}{S(F)}\middle| F \in \mathscr{F}(H)\right\} .$$

DEFINITION. The *c-density* of a set $A \subseteqq S$ with respect to $\mathscr{F}(H)$, written $\alpha_c$, is

$$\alpha_c = \operatorname{glb}\left\{\frac{A(H(x))}{S(H(x))} \,\middle|\, x \in S\backslash\{0\}\right\}.$$

We begin our comparison of $k$-density and $c$-density by stating without proof the following two results of Freedman.

THEOREM 1. *Let $(S, \mathscr{F}(H))$ be a density space. For any set $A \subseteqq S$ we have $0 \leqq \alpha_k \leqq \alpha_c \leqq 1$.*

THEOREM 2. *Let $(S, \mathscr{F}(H))$ be a density space and $A$ be a subset of $S$ with $0 \in A$. The following three conditions are equivalent:* (i) $\alpha_k = 1$, (ii) $\alpha_c = 1$, *and* (iii) $A = S$.

For the remainder of this paper we suppose that $\alpha_c < 1$. Freedman posed the following question [1]: Does there exist a density space $(S, \mathscr{F}(H))$ and a subset $A$ of $S$ such that $\alpha_c > 0$ and $\alpha_k = 0$? The answer is yes.

EXAMPLE 1. Let $I$ be the set of nonnegative integers with the usual addition and let $d$ be any positive integer. Let $H(x)$ be defined by

$$H(x) = \begin{cases} \{0, 1, 2, \cdots, d\} \cup \{x\} & \text{if } \ x \geqq d + 1\,, \\ \{0, x\} & \text{otherwise}\,, \end{cases}$$

where $x \in I\backslash\{0\}$. Then $(I, \mathscr{F}(H))$ is a density space. Let $A = \{0, 1, 2, \cdots, d\}$. Then $\alpha_k = 0$, but $\alpha_c = d/(d + 1) > 0$.

Example 1 shows that there are density spaces for which $\alpha_k = 0$ and $\alpha_c$ is arbitrarily close to (but not equal) 1. Example 1 also answers a second question of Freedman [1]: Does $0 \in A$ and $\alpha_c > 0$ imply that $A$ is a basis for $I$? The answer is, of course, no. The set $A$ has finite cardinality and hence cannot be a basis for $I$.

EXAMPLE 2. Let $I^\infty$ denote the set of all zero terminating sequences of nonnegative integers. Then $(I^\infty, \mathscr{F}(L))$ is a density space. For any positive integer $N \geqq 2$, let

$$I^\infty\backslash A = \{(x_1, x_2, \cdots) \mid x_i \leqq N \text{ for all } i \text{ and } x_i = N \text{ for exactly one } i\}\,.$$

Then $\alpha_k = 0$ and $\alpha_c = (N - 1)/N$ which again answers Freedman's first question. This density space is less artificial than the space in Example 1. However, it does not serve as an answer to Freedman's second question.

In the final theorem of this paper we show that it is possible to create a density space for which there exist sets having any $k$-density and $c$-density we want as long as Theorems 1 and 2 are not violated.

THEOREM 3. *There exists a density space* $(I, \mathscr{F}(H))$ *such that if* $0 \leqq \rho_1 \leqq \rho_2 < 1$, *then there is a set* $A \subseteq I$ *such that* $\alpha_k = \rho_1$ *and* $\alpha_c = \rho_2$.

*Proof.* Let $\{(d_i, b_i)\}$ be a sequence of ordered pairs of positive integers satisfying $1 < d_i \leqq b_i$ where all possible such pairs occur and occur infinitely often.

For all $x \in I \backslash \{0\}$, define

$$(1)\quad H(x) = \begin{cases} \{0, 1, 2, \cdots, x\} & \text{if } 1 \leqq x \leqq d_1 - 1 \,, \\ \{0, 1, 2, \cdots, d_1 - 1, x\} & \text{if } d_1 \leqq x \leqq b_1 \,, \\ \{0, b_1 + 1, b_1 + 2, \cdots, x\} & \text{if } b_1 + 1 \leqq x \leqq b_1 + d_2 - 1 \,, \\ \{0, b_1 + 1, b_1 + 2, \cdots, b_1 + d_2 - 1, x\} & \text{if } b_1 + d_2 \leqq x \leqq b_1 + b_2 \,, \\ \vdots & \\ \{0, \sum_{i=1}^{j} b_i + 1, \sum_{i=1}^{j} b_i + 2, \cdots, x\} & \text{if } \sum_{i=1}^{j} b_i + 1 \leqq x \leqq \sum_{i=1}^{j} b_i + d_{j+1} - 1 \,, \\ \{0, \sum_{i=1}^{j} b_i + 1, \sum_{i=1}^{j} b_i + 2, \cdots, \sum_{i=1}^{j} b_i + d_{j+1} - 1, x\} & \text{if } \sum_{i=1}^{j} b_i + d_{j+1} \leqq x \leqq \sum_{i=1}^{j+1} b_i \,, \\ \vdots & . \end{cases}$$

The space $(I, \mathscr{F}(H))$ is a density space.

Let $\rho_1$ and $\rho_2$ satisfy the hypothesis of the theorem. Let $\{u_i\}$ and $\{v_i\}$ be strictly decreasing sequences of positive rational numbers less than 1 such that

$$\rho_1 = \operatorname{glb} \{u_i \mid i = 1, 2, \cdots\},$$
$$\rho_2 = \operatorname{glb} \{v_i \mid i = 1, 2, \cdots\}, \text{ and } u_i \leqq v_i \text{ for each } i \,.$$

Since $u_i$ and $v_i$ are positive rationals, there exist positive integers $a_i$, $b'_i$, and $d'_i$ such that $u_i = a_i/b'_i$ and $v_i = a_i/d'_i$. Since $0 < u_i \leqq v_i < 1$, we have $1 \leqq a_i < d'_i \leqq b'_i$. Furthermore, there is a strictly increasing function $r(s)$ such that $b'_s = b_{r(s)}$ and $d'_s = d_{r(s)}$ for $s = 1, 2, \cdots$. Let

$$A = \{0\} \cup \left\{x \,\middle|\, x \in I, \sum_{i=1}^{j} b_i + 1 \leqq x \leqq \sum_{i=1}^{j+1} b_i, \text{ where } j \geqq 0 \text{ and } j + 1 \neq r(s) \text{ for all } s\right\}$$
$$\cup \bigcup_{s=1}^{\infty} \left\{x \,\middle|\, x \in I, \sum_{i=1}^{r(s)-1} b_i + 1 \leqq x \leqq \sum_{i=1}^{r(s)-1} b_i + a_s\right\}.$$

We now show that $\alpha_c = \rho_2$ and $\alpha_k = \rho_1$.

For each positive integer $s$, we have

$$\frac{A\left(H\left(\sum_{i=1}^{r(s)-1} b_i + d_{r(s)}\right)\right)}{I\left(H\left(\sum_{i=1}^{r(s)-1} b_i + d_{r(s)}\right)\right)} = \frac{a_s}{d_{r(s)}} = v_s .$$

Since $\rho_2 = \mathrm{glb}\,\{v_s \mid s = 1, 2, \cdots\}$, we have $\alpha_c \leqq \rho_2$. Also for any positive integer $m$ there is an integer $j \geqq 0$ such that $\sum_{i=1}^{j} b_i + 1 \leqq m \leqq \sum_{i=1}^{j+1} b_i$. If $j + 1 = r(s)$ for some $s$, then

$$\frac{A(H(m))}{I(H(m))} \geqq \frac{A\left(H\left(\sum_{i=1}^{r(s)-1} b_i + d_{r(s)}\right)\right)}{I\left(H\left(\sum_{i=1}^{r(s)-1} b_i + d_{r(s)}\right)\right)} = \frac{a_s}{d_{r(s)}} = v_s .$$

Otherwise $A(H(m)) = I(H(m))$.
Therefore,

$$\alpha_c = \mathrm{glb}\left\{\frac{A(H(m))}{I(H(m))} \middle| m = 1, 2, \cdots\right\} \geqq \mathrm{glb}\,\{v_s \mid s = 1, 2, \cdots\} = \rho_2 .$$

Hence we have $\alpha_c = \rho_2$.

It is more difficult to show that $\alpha_k = \rho_1$. For each integer $j \geqq 0$, define that set $F_j$ by

$$F_j = \left\{0, \sum_{i=1}^{j} b_i + 1, \sum_{i=1}^{j} b_i + 2, \cdots, \sum_{i=1}^{j+1} b_i\right\} .$$

By formula (1) we have

$$F_j = \bigcup_{m=m_1(j)}^{m_2(j)} H(m) ,$$

where $m_1(j) = \sum_{i=1}^{j} b_i + 1$ and $m_2(j) = \sum_{i=1}^{j+1} b_i$, and hence $F_j \in \mathscr{F}(H)$. If $j + 1 = r(s)$ for some $s$, then

$$\frac{A(F_j)}{I(F_j)} = \frac{a_s}{b_{j+1}} = \frac{a_s}{b_{r(s)}} = u_s . \tag{2}$$

Since $\rho_1 = \mathrm{glb}\,\{u_s \mid s = 1, 2, \cdots\}$ we have $\alpha_k \leqq \rho_1$. Now consider any $F \in \mathscr{F}(H)$. For each integer $j \geqq 0$, let $G_j = F \cap F_j$. Now since $F \in \mathscr{F}(H)$ and $F_j \in \mathscr{F}(H)$, we have $G_j \in \mathscr{F}(H) \cup \{\{0\}\}$. Now $i \neq j$ implies $F_i \cap F_j = \{0\}$ and hence $G_i \cap G_j = \{0\}$. Also $F$ is finite. Hence there is a finite integer

$$J = \max\,\{j \mid G_j \backslash \{0\} \neq \phi\} .$$

Now if $G_j \backslash \{0\} \neq \phi$, then $G_j \in \mathscr{F}(H)$. If $G_j \backslash A \neq \phi$, then $j + 1 = r(s)$ for some $s$ and

$$\left\{\sum_{i=1}^{r(s)-1} b_i + 1, \sum_{i=1}^{r(s)-1} b_i + 2, \cdots, \sum_{i=1}^{r(s)-1} b_i + a_s\right\} \subseteqq G_j \subseteqq F_j ,$$

and so

$$\frac{A(G_j)}{I(G_j)} \geqq \frac{A(F_j)}{I(F_j)} . \tag{3}$$

If $G_j \backslash A = \phi$, then $A(G_j) = I(G_j)$ and inequality (3) still holds. Therefore, by statement (3) and since $G_i \cap G_j = \{0\}$ for $i \neq j$, we have

$$\frac{A(F)}{I(F)} = \frac{A(\bigcup_{j=1}^{J} G_j)}{I(\bigcup_{j=1}^{J} G_j)} = \frac{\sum_{j=1}^{J} A(G_j)}{\sum_{j=1}^{J} I(G_j)} \geqq \frac{A(G_i)}{I(G_i)} \geqq \frac{A(F_i)}{I(F_i)}$$

for some $i$ $(1 \leqq i \leqq J)$. If $i + 1 = r(s)$ for some $s$, then by statement (2), we have

$$\frac{A(F_i)}{I(F_i)} = u_s .$$

If $i + 1 \neq r(s)$ for all $s$, then $A(F_i) = I(F_i)$. In either case, $A(F)/I(F) \geqq u_s$ for some $s$. Therefore,

$$\alpha_k = \operatorname{glb}\left\{\frac{A(F)}{I(F)} \middle| F \in \mathscr{F}(H)\right\} \geqq \operatorname{glb}\{u_s \mid s = 1, 2, \cdots\} = \rho_1 .$$

Hence we have $\alpha_k = \rho_1$.

## REFERENCE

1. Allen R. Freedman, *Some generalizations in additive number theory*, Reine Angew. Math., **235**: (1969), 1-19.

Received June 22, 1972. The author is indebted to the referee who provided Example 2 and made suggestions which strengthened Theorem 3.

U. S. AIR FORCE ACADEMY, COLORADO

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# ON THE SEMISIMPLICITY OF GROUP RINGS OF LINEAR GROUPS II

D. S. PASSMAN

**In this paper we continue the study of the semisimplicity problem for group rings of linear groups. We consider the case in which the characteristics of the two fields involved are both equal to $p > 0$ and we obtain appropriate necessary and sufficient conditions in terms of the abstract structure of the group.**

Let $K[G]$ denote the group ring of $G$ over the field $K$. In this paper we study the semisimplicity problem for $K[G]$ with $G$ a linear group. If $\operatorname{char} K = 0$ and if $G$ is a linear group over any field, then it is trivial to see that $JK[G] = 0$. Thus the only case of interest occurs when $\operatorname{char} K = p > 0$. A study of this situation was initiated by A. E. Zalesskii in [4] and continued somewhat in [3]. Here we solve the problem in case $G$ is a linear group over a field $L$ and $\operatorname{char} L = \operatorname{char} K$. Before we can properly state the result it is necessary to describe a certain characteristic subgroup $\mathscr{L}(G)$ of $G$. Therefore, we postpone the statement until the next section. We follow the notation of [2] and [3].

## 1. Normal $p$-subgroups.

Let $G$ be a linear group over a field $L$ of characteristic $p > 0$. That is, $G$ is a subgroup of the group of units of $L_u$, the ring of $u \times u$ matrices over $L$. Of course $G$ is also contained in $\tilde{L}_u$, where $\tilde{L}$ is the algebraic closure of $L$ and thus without loss of generality we may assume that $L$ is algebraically closed. Thus for the remainder of this work $L$ will denote a fixed algebraically closed field of $\operatorname{char} p > 0$ and any subgroup of $L_u$ for any $u$ will be called an $L$-linear group.

It is apparent from [4] that a necessary ingredient here must be a consideration of the normal $p$-subgroups of $G$. We start with a few elementary observations. If $G$ is any group let $\boldsymbol{O}_p(G)$ denote its maximal normal $p$-subgroup. It is clear that $\boldsymbol{O}_p(G)$ always exists. If $G \subseteq L_u$ we let $LG$ donote its $L$-linear span. Thus certainly $LG$ is an $L$-subalgebra of $L_u$.

LEMMA 1.1. *Let $G$ be an $L$-linear group. Then*

(i) *$\boldsymbol{O}_p(G)$ is a nilpotent group.*

(ii) *$G/\boldsymbol{O}_p(G)$ is an $L$-linear group.*

(iii) *If $\boldsymbol{O}_p(G) = \langle 1 \rangle$, then $G$ can be represented as an $L$-linear group in such a way that $LG$ is semisimple.*

(iv) *If $LG$ is semisimple and $H \lhd G$ then $LH$ is semisimple.*

*Proof.* Observe that $LG$ is a finite dimensional $L$-algebra so $JLG$, its Jacobson radical, is nilpotent. We start by proving (iv). If $x \in G$ then since $H \lhd G$, $x$ normalizes $H$ and hence clearly $x$ acts as an algebra automorphism on $LH$. Since $JLH$ is characteristic in $LH$ we have $x^{-1}(JLH)x = JLH$ so $(JLH)x = x(JLH)$. Thus since $LG$ is spanned by all such $x$ we obtain easily $(JLH)(LG) = (LG)(JLH)$. Now $JLH$ is nilpotent and therefore by the above so is the ideal $(JLH)(LG)$. Thus $(JLH)(LG) \subseteq JLG = 0$ and $JLH = 0$. This yields (iv).

Now let $\mu: LG \to LG/JLG$ be the natural map and let $P = \{g \in G \mid \mu(g) = 1\}$. Since $G \subseteq LG$, $P$ is a subgroup of $U = \{1 + \alpha \mid \alpha \in JLG\} \subseteq LG$. Now $JLG$ is nilpotent and $\operatorname{char} L = p > 0$ so we see easily that $U$ is a nilpotent $p$-group. Thus $P$ is a nilpotent $p$-group and $P \subseteq \boldsymbol{O}_p(G)$.

Now $\mu(LG) = LG/JLG$ is a finite dimensional $L$-algebra so it is contained in $L_w$ for some integer $w$. Furthermore, $LG/JLG$ contains the group $\bar{G} = G/P$ and is clearly spanned by it. This shows that $\bar{G}$ is an $L$-linear group with $L\bar{G}$ semisimple. If $\boldsymbol{O}_p(G) = \langle 1 \rangle$ then certainly $P = \langle 1 \rangle$ so $G = \bar{G}$ and (iii) is proved.

Observe that if we show that $P = \boldsymbol{O}_p(G)$ then (i) and (ii) will follow and to do this we need only show that $\bar{Q} = \boldsymbol{O}_p(\bar{G}) = \langle 1 \rangle$. Since $L\bar{G}$ is semisimple, part (iv) and $\bar{Q} \lhd \bar{G}$ implies that $L\bar{Q}$ is also semisimple. Let $I$ be the subalgebra of $L\bar{Q}$ spanned by all $1 - x$ with $x \in \bar{Q}$. Then $I$ is an ideal of $L\bar{Q}$ and $I$ is a finite dimensional algebra (without 1) spanned by the nilpotent elements $1 - x$. As is well known (see for example the proof of Lemma 10.1 (ii) of [2]) this implies that $I$ is nilpotent so $I \subseteq JL\bar{Q} = 0$. If $x \in \bar{Q}$ then $1 - x \in I = 0$ so $x = 1$. Thus $\bar{Q} = \langle 1 \rangle$ and the lemma is proved.

Let $G$ be any group and let $H$ be a subgroup of $G$. We set

$$\boldsymbol{D}_G(H) = \{x \in G \mid [H: \boldsymbol{C}_H(x)] < \infty\} .$$

Clearly $\boldsymbol{D}_G(H)$ is a subgroup of $G$ and if $H$ is normal or characteristic in $G$ then so is $\boldsymbol{D}_G(H)$. Furthermore,

$$\boldsymbol{D}_G(G) = \Delta(G) = \{x \in G \mid [G: \boldsymbol{C}_G(x)] < \infty\}$$

is the $F.C.$ subgroup of $G$. Finally $\Delta^p(G)$ is defined to be the subgroup of $\Delta(G)$ generated by all $p$-elements, that is elements whose order is a power of $p$. We say that $G$ is a $\Delta$-group if $G = \Delta(G)$.

LEMMA 1.2. *Let $G$ be an $L$-linear group.*

(i) *If $H \lhd G$ and $G = \boldsymbol{D}_G(H)$ then $[H: H \cap \boldsymbol{Z}(G)] < \infty$.*

(ii) *If $\boldsymbol{O}_p(G) = \langle 1 \rangle$ then $\Delta^p(G)$ is finite.*

*Proof.* Since $LG$ is finite dimensional we can choose some finite number of group elements $x_1, x_2, \cdots, x_n$ which span $LG$. By assumption for each $i$, $[H: C_H(x_i)] < \infty$ and thus by Lemma 1.1 of [2], $[H: Z] < \infty$ where $Z = \bigcap_1^n C_H(x_i)$. Now $Z \subseteq LG$ is centralized by a spanning set so it is, therefore, centralized by all of $LG$ and hence by all of $G$. This shows that $Z \subseteq Z(G)$ and thus (i) follows.

Suppose $O_p(G) = \langle 1 \rangle$ and set $H = \Delta^p(G)$. Then $H = \Delta(H)$ so by part (i) applied to $H$ we conclude that $[H: Z(H)] < \infty$. Now $O_p(G) = \langle 1 \rangle$ and $Z(H) \lhd G$ so $O_p(Z(H)) = \langle 1 \rangle$ and since $Z(H)$ is abelian this says that $Z(H)$ has no elements of order $p$. Thus $\Delta^p(Z(H)) = \langle 1 \rangle$. On the other hand, since $[\Delta^p(G): Z(H)] < \infty$, Lemma 19.3 (v) of [2] implies that $[Z(H): \Delta^p(Z(H))] < \infty$. Thus $\Delta^p(Z(H)) = \langle 1 \rangle$ yields $|Z(H)| < \infty$ and hence $|H| < \infty$. This completes the proof.

Let $G$ be any group. We define a characteristic subgroup $\mathscr{L}(G)$ of $G$ as follows. Let $P = O_p(G)$ and set $G^* = D_G(P)$ so that $G^* \cap P = D_P(P) = \Delta(P)$. Then $\mathscr{L}(G)$ is the subgroup of $G^*$ given by

$$G^* \supseteq \mathscr{L}(G) \supseteq \Delta(P), \qquad \mathscr{L}(G)/\Delta(P) = \Delta^p(G^*/\Delta(P)) .$$

LEMMA 1.3. *Let $G$ be an $L$-linear group. Then with the above notation $[\mathscr{L}(G): \Delta(P)]$ is finite and $\mathscr{L}(G)$ is a characteristic $\Delta$-subgroup of $G$.*

*Proof.* $\mathscr{L}(G)$ is clearly characteristic by its construction. Now $G^* \lhd G$ so $O_p(G^*) \subseteq O_p(G) = P$ and thus $O_p(G^*) = \Delta(P)$. Therefore, by Lemma 1.1 (ii), $G^*/\Delta(P)$ is an $L$-linear group and certainly $O_p(G^*/\Delta(P)) = \langle 1 \rangle$. Thus Lemma 1.2 (ii) implies that $\Delta^p(G^*/\Delta(P))$ is finite and we see that $[\mathscr{L}(G): \Delta(P)]$ is finite. Furthermore, since clearly $G^* = D_{G^*}(\Delta(P))$, Lemma 1.2 (ii) yields $[\Delta(P): \Delta(P) \cap Z(G^*)] < \infty$ and this and the above show that $\mathscr{L}(G)$ has a center of finite index. Therefore, $\mathscr{L}(G)$ is a $\Delta$-group.

We can now state our main result. If $H$ is a subgroup of $G$ we say that $H$ has locally finite index in $G$ and write $[G: H] = \text{l.f.}$ if for all finitely generated subgroups $S$ of $G$ we have $[S: S \cap H] < \infty$.

THEOREM. *Let $K$ be a field of characteristic $p > 0$ and let $G$ be a linear group over a field of the same characteristic $p$. Then $JK[G] \neq 0$ if and only if there exists an element $h \in \mathscr{L}(G)$ of order $p$ with $[G: C_G(h)] = l.f.$*

Observe that the above necessary and sufficient conditions concern the abstract structure of $G$ and not how $G$ is written as a linear group.

2. The case: $O_p(G) = \langle 1 \rangle$. The linear groups with $O_p(G) = \langle 1 \rangle$ were studied in [4] under the additional assumption that $K = L$, that is the two fields are the same, and the semisimplicity problem was solved in that case. Here we modify the original argument slightly to handle the case in which $K$ and $L$ are different.

If $S$ is a subset of any group $G$ we say that $S$ has finite index in $G$ and write $[G: S] < \infty$ if $G$ can be written as a finite union, $G = \bigcup_1^n Sx_i$, of right translates of $S$.

LEMMA 2.1. *Let $G$ be an $L$-linear group and let $T_1, T_2, \cdots, T_j$ be a finite number of $L$-subspaces of $LG$ properly smaller than $LG$. Let $S$ be a subset of $G$ and suppose that*

$$G = S \cup \bigcup_1^j (G \cap T_i) .$$

*Then either $[G: S] < \infty$ or $G = \bigcup_1^j (G \cap T_i)$ and $G$ has a subgroup $H$ of finite index with $LH \neq LG$.*

*Proof.* We assume that $[G: S]$ is infinite and we consider all ways of writing $G$ as a finite union

$$G = \bigcup_1^s Sx_i \cup \bigcup_1^t (G \cap M_i)$$

where $x_i \in G$ and the $M_i$ are $L$-subspaces of $LG$ each contained in some $T_{i'}$. By assumption such a decomposition exists. For each such union we associate an ordered pair $(d, r)$ where $d = \max \dim M_i$ and $r$ is the number of $M_i$ of dimension $d$. We say $(d_1, r_1) < (d_2, r_2)$ if $d_1 < d_2$ or $d_1 = d_2$ and $r_1 < r_2$. This then is a well ordering and assume the above union is so chosen that $(d, r)$ is minimal. By definition $d < \dim LG$. We may assume that $\dim M_i = d$ for $i = 1, 2, \cdots, r$. Note that the $M_i$ terms must occur since $[G: S] = \infty$.

Fix $k \leqq r$ and $g \in G$. Then

$$(G \cap M_k)g \subseteqq G = \bigcup_1^s Sx_i \cup \bigcup_1^t (G \cap M_i)$$

so

$$\begin{aligned} G \cap M_k &\subseteqq \bigcup_1^s Sx_ig^{-1} \cup \bigcup_1^t (G \cap M_i)g^{-1} \\ &= \bigcup_1^s Sx_ig^{-1} \cup \bigcup_1^t (G \cap M_ig^{-1}) \end{aligned}$$

and thus

$$G \cap M_k \subseteqq \bigcup_1^s Sx_ig^{-1} \cup \bigcup_1^t (G \cap (M_ig^{-1} \cap M_k)) .$$

Thus replacing the term $G \cap M_k$ in the original union by the above yields a new such union with the subspace $M_k$ replaced by the finitely many subspaces $M_i g^{-1} \cap M_k$ for $i = 1, 2, \cdots, t$. If $\dim(M_i g^{-1} \cap M_k) < \dim M_k$ for all $i$, we then get a new decomposition with some smaller parameter $(d', r')$. Since this cannot happen we conclude that for some $i$, $M_i g^{-1} \cap M_k = M_k$ or $M_i \supseteqq M_k g$. Since $M_k$ has the largest dimension of all the subspaces we therefore have $M_i = M_k g$ for some $i \leqq r$.

We have therefore shown that $G$ permutes by right multiplication the subspaces $M_1, M_2, \cdots, M_r$ and hence if $H$ is the stabilizer of $M_1$ then $[G: H] < \infty$. If $LH = LG$ then $M_1 H = M_1$ implies that $M_1(LH) = M_1$ and then $M_1 G = M_1$. Again by the minimality of $(d, r)$ and $[G: S] = \infty$ we have $G \cap M_1 \neq \emptyset$ so let $y \in G \cap M_1$. Then $M_1 G = M_1$ yields $M_1 \supseteqq yG = G$. Thus $M_1 \supseteqq LG$, a contradiction. This shows that $LH \neq LG$ and therefore $LH$ is a proper subalgebra of $LG$.

Finally let $1 = g_1, g_2, \cdots, g_m$ be a set of right coset representatives for $H$ in $G$. By renumbering the $M_i$'s if necessary we may assume that $M_1 g_i = M_i$. Let $T_{i'}$ be chosen with $M_i \subseteqq T_{i'}$. Now $M_1 H = M_1$ yields $yH \subseteqq M_1$ so $yHg_i \subseteqq M_1 g_i = M_i$. Thus

$$G = yG = \bigcup_1^m yHg_i \subseteqq \bigcup_1^m M_i \subseteqq \bigcup_1^m T_{i'}$$

so clearly $G = \bigcup_1^j (G \cap T_i)$ and the lemma is proved.

For the remainder of this work we let $K$ denote a fixed field of characteristic $p$. If $G$ is a group and if $x, y \in G$ we use $x \sim_G y$ to indicate that $x$ and $y$ are conjugate in $G$.

LEMMA 2.2. *Let* $\alpha = \sum_{i=1}^k a_i g_i \in K[G], \alpha \neq 0$ *and suppose that* $\alpha$ *is nilpotent. Then for some* $i \neq j$ *and some integer* $n$ *we have* $g_i^{p^n} \sim_G g_j^{p^n}$.

*Proof.* Let $S$ denote the subspace of $K[G]$ spanned by all Lie products $[\beta, \gamma] = \beta\gamma - \gamma\beta$ with $\beta, \gamma \in K[G]$. Then $S$ is spanned by all Lie products $[x, y] = xy - yx$ with $x, y \in G$. Now $yx = x^{-1}(xy)x$ so $yx \sim_G xy$ and, therefore, we see that if $\delta \in S$ then the sum of the coefficients of $\delta$ over any conjugacy class of $G$ is zero.

By assumption $\alpha$ is nilpotent so we can choose $n \geqq 0$ with $\alpha^{p^n} = 0$. Then Lemma 3.4 of [2] yields

$$0 = \alpha^{p^n} = \sum_{i=1}^k a_i^{p^n} g_i^{p^n} + \delta$$

for some $\delta \in S$. If $a_i \neq 0$ then since the sum of the coefficients in the conjugacy class of $g_i^{p^n}$ must be zero in the above and since the contribution of $\delta$ to this sum is zero, we conclude that some $j \neq i$ must exist with $g_j^{p^n} \sim_G g_i^{p^n}$.

LEMMA 2.3. *Let $G$ be an $L$-linear group with $LG$ semisimple. Since $L$ is algebraically closed, $LG$ is a finite direct sum of full matrix rings over $L$ and we embed $LG$ in $L_u$ for some $u$ by placing the matrix rings of $LG$ in blocks along the diagonal of $L_u$. Then* tr, *the matrix trace map on $L_u$, yields a nondegenerate symmetric bilinear form $(\alpha, \beta) = \operatorname{tr} \alpha\beta$ on $LG$.*

*Proof.* The form $(\alpha, \beta) = \operatorname{tr} \alpha\beta$ is certainly bilinear and symmetric. We need only show that it is nondegenerate on $LG$. Let $\alpha \in LG$ with $(\alpha, LG) = 0$. Then

$$\operatorname{tr}(LG)\alpha(LG) = \operatorname{tr} \alpha(LG)(LG) = \operatorname{tr} \alpha(LG) = 0$$

so every element of the ideal $(LG)\alpha(LG)$ has trace zero. But any nonzero ideal of $LG$ contains one of the full matrix ring and certainly all its elements cannot have trace 0. Thus $\alpha$ must be zero and the lemma is proved.

We now obtain our generalization of Zalesskii's result by modifying the proof of [4]. It is apparent that the proof could be greatly simplified if we only knew that the radical was a nil ideal.

LEMMA 2.4. *Let $G$ be an $L$-linear group with $\boldsymbol{O}_p(G) = \langle 1 \rangle$. Then $G$ has a normal subgroup $G_0$ of finite index and a representation of $G_0$ as an $L$-linear group in such a way that $LG_0$ is semisimple and if $[G_0 \colon H] < \infty$ then $LG_0 = LH$.*

*Proof.* Since $\boldsymbol{O}_p(G) = \langle 1 \rangle$, Lemma 1.1 (iii) implies that $G$ can be represented as an $L$-linear group with $LG$ semisimple. We now consider all normal subgroups $H$ of $G$ of finite index and all ways in which $H$ can be represented as an $L$-linear group with $LH$ semisimple and we choose $G_0$ to give the minimum possible dimension of $LG_0$.

Thus we have $G_0 \lhd G$, $[G \colon G_0] < \infty$ and $G_0$ is an $L$-linear group with $LG_0$ semisimple. Furthermore, let $H$ be a subgroup of $G_0$ of finite index. Then $[G \colon H] < \infty$ so $H_0$, the intersection of the finitely many $G$-conjugates of $H$, is a normal subgroup of $G$ of finite index. Since $H_0 \lhd G_0$ we have $LH_0$ semisimple by Lemma 1.1 (iv) and thus by the minimality of the dimension of $LG_0$ we have $LG_0 = LH_0$ and hence $LG_0 = LH$.

PROPOSITION 2.5. *Let $G$ be an $L$-linear group with $\boldsymbol{O}_p(G) = \langle 1 \rangle$. Then $JK[G]$ is nilpotent.*

*Proof.* Let $G_0$ be the normal subgroup of $G$ of finite index given in the preceding lemma and let us write $LG_0$ as described in Lemma 2.3. Thus $LG_0 \subseteq L_u$ and tr yields a nondegenerate bilinear form on $LG_0$. We show now that $K[G_0]$ is semisimple.

Suppose by way of contradiction that $\alpha = \sum_{i=1}^k a_i g_i \in JK[G_0]$ with $\alpha \neq 0$ and with the group elements $g_i$ distinct. If $x \in G_0$ then also $\alpha x = \sum_{i=1}^k a_i g_i x \in JK[G_0]$. Thus if $G_1$ is the finitely generated subgroup of $G_0$ given by $G_1 = \langle g_1, g_2, \cdots, g_k, x\rangle$ then $\alpha x \in JK[G_0] \cap K[G_1] \subseteq JK[G_1]$ by Lemma 16.9 of [2]. We show now that for some $i \neq j$, $\operatorname{tr}(g_i x) = \operatorname{tr}(g_j x)$.

Suppose this is not the case and let $G\tilde{F}(p)$ denote the algebraic closure of $GF(p)$. Since $G_1$ is a finitely generated subgroup of $L_u$ we can find, by the Extension Theorem for Places, a place $\varphi: L \to G\tilde{F}(p) \cup \{\infty\}$ such that $\varphi$ is finite on all the matrix entries of the generators of $G_1$ and their inverses and furthermore for all $i \neq j$, $\varphi(\operatorname{tr}(g_i x)) \neq \varphi(\operatorname{tr}(g_j x))$. If $\mathcal{O}$ denotes the corresponding valuation ring in $L$ then clearly $G_1 \subseteq \mathcal{O}_u$ and $\varphi$ can be extended to a homomorphism $\varphi: \mathcal{O}_u \to G\tilde{F}(p)_u$ and therefore $\varphi(G_1)$ is finite.

Consider the natural map $\eta: K[G_1] \to K[\varphi(G_1)]$. Since $\eta$ is an epimorphism, $\eta(JK[G_1]) \subseteq JK[\varphi(G_1)]$ and thus

$$\eta(\alpha x) = \sum_{i=1}^k a_i \varphi(g_i x) \in JK[\varphi(G_1)] .$$

Now $\varphi(G_1)$ is finite so $JK[\varphi(G_1)]$ is nilpotent and therefore $\sum_{i=1}^k a_i \varphi(g_i x)$ is nilpotent. Thus Lemma 2.2 implies that for some $i \neq j$ and some integer $n$, $\varphi(g_i x)^{p^n} \sim_{\varphi(G_1)} \varphi(g_j x)^{p^n}$. Let $\widetilde{\operatorname{tr}}$ denote the trace map in $G\tilde{F}(p)_u$. Since similar matrices have the same trace and since the fields have characteristic $p > 0$ we conclude that

$$[\widetilde{\operatorname{tr}}\, \varphi(g_i x)]^{p^n} = \widetilde{\operatorname{tr}}\, [\varphi(g_i x)^{p^n}] = \widetilde{\operatorname{tr}}\, [\varphi(g_j x)^{p^n}] = [\widetilde{\operatorname{tr}}\, \varphi(g_j x)]^{p^n}$$

and thus $\widetilde{\operatorname{tr}}\, \varphi(g_i x) = \widetilde{\operatorname{tr}}\, \varphi(g_j x)$. But certainly $\widetilde{\operatorname{tr}} \circ \varphi = \varphi \circ \operatorname{tr}$ so we obtain

$$\varphi(\operatorname{tr}(g_i x)) = \widetilde{\operatorname{tr}}\, \varphi(g_i x) = \widetilde{\operatorname{tr}}\, \varphi(g_j x) = \varphi(\operatorname{tr}(g_j x))$$

a contradiction.

We have, therefore, shown that for each $x \in G_0$ there exists some $i \neq j$ with $\operatorname{tr} g_i x = \operatorname{tr} g_j x$. For each $i \neq j$ let $T_{ij}$ be the $L$-subspace of $LG_0$ given by

$$T_{ij} = \{\delta \in LG_0 \mid \operatorname{tr}(g_i - g_j)\delta = 0\} .$$

Since tr yields a nondegenerate bilinear form we see that $T_{ij} \neq LG_0$ and by the above we have

$$G = \bigcup_{i \neq j} G \cap T_{ij} .$$

But then Lemma 2.1 with $S = \varnothing$ implies that $G_0$ has a subgroup $H$ of finite index with $LH \neq LG_0$, a contradiction. This shows that $K[G_0]$ is semisimple. Since $[G: G_0] < \infty$, Lemma 16.8 of [2] implies that $JK[G]$ is nilpotent and result follows.

**3. A local situation.** We now study a group $G$ with a rather special structure. We say $G$ has property (*) if $G$ has a normal series $G \supseteqq W \supseteqq P \supseteqq Z$ satisfying

1. $G/W$ is infinite cyclic.
2. $\bar{G} = G/P$ is an $L$-linear group with $O_p(\bar{G}) = \langle 1 \rangle$.
3. $P$ is an abelian $p$-group.
4. $[P: Z] < \infty$ and W centralizes $Z$.

We say that $G$ has property (**) if $G$ satisfies all of the above and in addition

5. $P \cap \Delta(G) = \langle 1 \rangle$.

Our aim is essentially to completely determine $JK[G]$ if $G$ satisfies (*). We start by assuming that $G$ satisfies (**) and prove that $JK[G]$ is nilpotent. For the remainder of this section we assume that $G$ satisfies (**) and is given as above. We start by introducing some more notation.

LEMMA 3.1. *There exists a subgroup $G_0$ of $G$ of finite index with $G \supseteqq G_0 \supseteqq P$ and such that*

( i ) *$\bar{G}_0 = G_0/P$ has a representation as an L-linear group with $L\bar{G}_0$ semisimple and with $L\bar{G}_0 = L\bar{H}$ for all subgroups $\bar{H} \subseteqq \bar{G}_0$ of finite index.*

(ii) *$G_0$ centralizes the quotient $P/Z$.*

(iii) *If $W_0 = G_0 \cap W$ then $G_0/W_0$ is infinite cyclic.*

*Proof.* The existence of a group $G_0$ satisfying (i) is an immediate consequence of Lemma 2.4. Furthermore, it is clear that this same property holds for any subgroup of $G_0$ of finite index which contains $P$. Now $G_0$ acts on finite group $P/Z$ and $P$ centralizes this quotient. Thus we may certainly replace $G_0$ by $C_{G_0}(P/Z)$ if necessary and then this new $G_0$ also satisfies (ii). Finally

$$G_0/W_0 = G_0/(W \cap G_0) \cong G_0W/W$$

is a subgroup of finite index in the infinite cyclic group $G/W$ and the result follows.

We will show that $K[G_0]$ is semisimple. Thus by way of contradiction we assume now that $G_0$ is given as above and $JK[G_0] \neq 0$.

LEMMA 3.2. *There exists a nonzero element* $\gamma = \alpha\beta \in JK[G_0] \cap K[W_0]$ *satisfying*

(i) $\alpha = \hat{Q}$, *the sum of all the elements of* $Q$, *where* $Q$ *is a finite subgroup of* $P$.

(ii) $\beta = \sum_{i=1}^{n} a_i g_i$ *where the* $g_i$ *are in distinct cosets of* $P$ *in* $W_0$.

(iii) $\gamma$ centralizes $K[P]$.

*Proof.* By assumption $JK[G_0] \neq 0$ and since $G_0/W_0$ is infinite cyclic Theorem 17.7 of [2] implies that

$$I = JK[G_0] \cap K[W_0]$$

is a nonzero ideal of $K[W_0]$. Choose $\gamma \in I, \gamma \neq 0$ such that Supp $\gamma$ is contained in the smallest number $n$ of cosets of $P$. By multiplying $\gamma$ by a group element if necessary we may assume that one of these cosets is the identity coset. Thus

$$\gamma = \sum_{i=1}^{n} \alpha_i g_i$$

where $\alpha_i \in K[P]$ and $g_1 = 1, g_2, \cdots, g_n$ are in distinct cosets of $P$.

Let $Q$ be the subgroup of $P$ generated by the support of all the $\alpha_i$. Then $Q$ is a finitely generated and hence finite subgroup of the abelian $p$-group $P$. Therefore, as is well known, the unique minimal ideal of $K[Q]$ consists of all $K$-multiples of $\hat{Q}$ and thus $\hat{Q}$ is a multiple of $\alpha_1$ in $K[Q]$. By multiplying $\gamma$ on the left by this suitable factor we may clearly assume that $\alpha_1 = \hat{Q}$. Let $h \in Q$. Then $(1-h)\alpha_1 = 0$ so $(1-h)\gamma \in I$ has support contained in a smaller number of cosets. This implies that $(1-h)\gamma = 0$ for all $h \in Q$ and thus we have for all $i$, $\alpha_i = a_i\hat{Q}$ for some $a_i \in K$. This yields

$$\gamma = \alpha\beta = \hat{Q}\Big(\sum_{i=1}^{n} a_i g_i\Big)$$

and (i) and (ii) are proved.

Finally let $h \in P$. Since $P$ is abelian and $g_1 = 1$ we see $h^{-1}\gamma h - \gamma \in I$ has support in fewer cosets of $P$. By the minimality of $n$ we conclude that $h^{-1}\gamma h - \gamma = 0$ for all $h \in P$ and (iii) follows.

We now define an even smaller subgroup of $G$. Again we fix the above notation for the remainder of this section. Let

$$T = \{h \in Q \mid h \neq 1, \boldsymbol{C}_{G_0}(h) \not\subseteq W_0\} .$$

Now define the subgroup $G_1$ by

$$G_1 = \bigcap_{h \in T} W_0 C_{G_0}(h)$$

with the understanding that $G_1 = G_0$ if $T = \varnothing$.

LEMMA 3.3. *Let $G_1$ be as above. Then*
(i) $G_0 \supseteqq G_1 \supseteqq W_0$, $[G_0 : G_1] < \infty$ *and $G_1/W_0$ is infinite cyclic.*
(ii) *If $h \in T$ then $G_1 = W_0 C_{G_1}(h)$.*

*Proof.* By definition we have $G_0 \supseteqq G_1 \supseteqq W_0$. Moreover, $G_1/W_0$ is the intersection of finitely many nonidentity subgroups of the infinite cyclic group $G_0/W_0$. Thus $G_1/W_0$ is infinite cyclic and $[G_0 : G_1] < \infty$.
Finally let $h \in T$. Then $W_0 \subseteqq G_1 \subseteqq W_0 C_{G_0}(h)$ so

$$G_1 = W_0(G_1 \cap C_{G_0}(h)) = W_0 C_{G_1}(h)$$

and the lemma is proved.

The reason for working with $G_1$ rather than $G_0$ will be apparent in the following result.

LEMMA 3.4. *Let $x \in G_1 - W_0$ and let $\alpha$ be as above. Suppose that for infinitely many integers $s$ (positive or negative) there exists an integer $r = r(s) \geqq 1$ with*

$$\alpha\alpha^{x^{-s}}\alpha^{x^{-2s}} \cdots \alpha^{x^{-rs}} = 0 .$$

*Then for some $h \in T$ we have $x \in C_{G_1}(h)$.*

*Proof.* The assumption on $\alpha$ clearly implies that for each such $s$ the group $QQ^{x^{-s}}Q^{x^{-2s}} \cdots Q^{x^{-rs}}$ is not a direct product of the indicated factors. Since there are infinitely many such $s$ there are certainly infinitely many positive or infinitely many negative ones. Therefore, by Lemmas 3 and 4 of [1], there exists $h \in Q$, $h \neq 1$ and a positive integer $m$ with $x^{-m}$ or $x^m$ in $C_{G_1}(h)$ and hence $x^m \in C_{G_1}(h)$. Now $x \in G_1 - W_0$ and $G_1/W_0$ is infinite cyclic so $x^m \notin W_0$ and by definition of $T$ we must have $h \in T$.

Since $G_1 = W_0 C_{G_1}(h)$ by Lemma 3.3 (ii) we can write $x = wy$ with $w \in W_0$ and $y \in C_{G_1}(h)$. Therefore, $y \in C_{G_1}(h)$, $x^m = (wy)^m \in C_{G_1}(h)$ and since $W_0$ centralizes $P/Z$ we have $h^w = hz$ for some $z \in Z$. It then follows easily by induction on $i$ that

$$h^{(wy)^i} = hz^y z^{y^2} \cdots z^{y^i}$$

and therefore

$$h = h^{(wy)^m} = hz^y z^{y^2} \cdots z^{y^m}$$

so we have $z^y z^{y^2} \cdots z^{y^m} = 1$. We now conjugate this last expression by $y^{-1}$ and obtain

$$zz^y \cdots z^{y^{m-1}} = 1 = z^y z^{y^2} \cdots z^{y^m} .$$

Thus since $P$ is abelian we have $z = z^{y^m}$.

Since $G$ satisfies (**) we know that $W$ centralizes $Z$ and thus we have $C_{G_1}(z) \supseteqq \langle W_0, y^m \rangle$. Furthermore, $G_1/W_0$ is infinite cyclic and $y \notin W_0$ since $x \notin W_0$ so clearly $[G_1: C_{G_1}(z)] < \infty$. Hence $[G: C_G(z)] < \infty$ and we have $z \in P \cap \Delta(G)$. Again by assumption (**), $P \cap \Delta(G) = \langle 1 \rangle$ so $z = 1$. Finally $h^x = h^{wy} = hz^y = h$ so $x \in C_{G_1}(h)$ and the result follows.

Let $^{-}$ denote the natural map $G_0 \to \bar{G}_0 = G_0/P$ and we extend this to the map $K[G_0] \to K[\bar{G}_0]$. Thus for $\beta = \sum_{i=1}^n a_i g_i$ as given before we have $\bar{\beta} = \sum_{i=1}^n a_i \bar{g}_i$. We now represent $\bar{G}_0$ as an $L$-linear group as in Lemma 3.1 (i) so that $L\bar{G}_0$ is semisimple.

LEMMA 3.5. *We can embed $L\bar{G}_0$ in the matrix ring $L_u$ in such a way that* tr, *the matrix trace map on $L_u$, yields a nondegenerate symmetric bilinear form $L\bar{G}_0$. Futhermore, if for each $i \neq j$ we define $T_{ij}$ by*

$$T_{ij} = \{\bar{\delta} \in L\bar{G}_0 \mid \operatorname{tr}(\bar{g}_i - \bar{g}_j)\bar{\delta} = 0\}$$

*then $T_{ij}$ is a proper $L$-subspace of $L\bar{G}_0 = L\bar{G}_1$.*

*Proof.* The first part follows immediately from Lemma 3.1 (i) and Lemma 2.3. The second part about $T_{ij}$ follows from the nondegeneracy of the bilinear form and the fact that $\bar{g}_i \neq \bar{g}_j$ by Lemma 3.2 (ii). Finally $L\bar{G}_0 = L\bar{G}_1$ by Lemma 3.1 (i).

LEMMA 3.6. *Let $x \in G_1 - W_0$ and let $\beta$ be as above. Suppose that $\bar{\beta}\bar{x}^s \in K[\bar{G}_1]$ is nilpotent for all integers $s$ (positive or negative) with possibly finitely many exceptions. Then for some $i \neq j$ we have $\bar{x} \in T_{ij}$.*

*Proof.* Since $x \in G_1 - W_0$ and $\bar{G}_1/\bar{W}_0$ is infinite cyclic we see that $\langle \bar{x} \rangle$ is infinite. We consider $\langle \bar{x} \rangle$ as an $L$-linear subgroup of $\bar{G}_0$. Let $V$ denote the finite set of exceptional integers in the above and let

$$S = \{\bar{x}^v \mid v \in V\} .$$

Then $S$ is a finite subset of $\langle \bar{x} \rangle$ so clearly $[\langle \bar{x} \rangle : S] = \infty$. Now let $s$ be an integer not in $V$. Since

$$\bar{\beta}\bar{x}^s = \sum_{i=1}^n a_i \bar{g}_i \bar{x}_i^s$$

is nilpotent we conclude from Lemma 2.2 that for some $i \neq j$ and some integer $t \geqq 0$

$$(\bar{g}_i\bar{x}^s)^{p^t} \sim_{\bar{G}_1} (\bar{g}_j\bar{x}^s)^{p^t} .$$

Thus since similar matrices have the same trace and since char $L = p > 0$ we have

$$(\operatorname{tr} \bar{g}_i\bar{x}^s)^{p^t} = \operatorname{tr} (\bar{g}_i\bar{x}^s)^{p^t} = \operatorname{tr} (\bar{g}_j\bar{x}^s)^{p^t} = (\operatorname{tr} \bar{g}_j\bar{x}^s)^{p^t} .$$

Hence $\operatorname{tr} \bar{g}_i\bar{x}^s = \operatorname{tr} \bar{g}_j\bar{x}^s$ and $\bar{x}^s \in T_{ij}$.

We have therefore shown that

$$\langle \bar{x} \rangle = S \cup \bigcup_{i \neq j} (\langle \bar{x} \rangle \cap T_{ij})$$

and since $[\langle \bar{x} \rangle : S] = \infty$, Lemma 2.1 implies that

$$\langle \bar{x} \rangle = \bigcup_{i \neq j} (\langle \bar{x} \rangle \cap T_{ij}) .$$

This shows that $\bar{x} \in T_{ij}$ for some $i \neq j$ and the lemma is proved.

We now come to the main result of this section.

PROPOSITION 3.7. *Let $G$ be a group satisfying* (**). *Then $JK[G]$ is nilpotent.*

*Proof.* We use all the above notation and show first that $JK[G_0] = 0$. If this is not the case then all of the above lemmas and notation apply.

Let $x \in G_1 - W_0$ and let $s \neq 0$ be an integer (positive or negative). Since $G_1/W_0$ is infinite cyclic, the element $x^{-s}$ has infinite order modulo $W_0$. Since $\gamma \in JK[G_0] \cap K[W_0]$, Lemma 21.3 of [2] implies that for some integer $r = r(s) \geqq 1$ we have

$$\gamma\gamma^{x^{-s}}\gamma^{x^{-2s}} \cdots \gamma^{x^{-rs}} = 0 .$$

Now $\gamma = \alpha\beta$ so this yields

$$\alpha\beta\alpha^{x^{-s}}\beta^{x^{-s}}\alpha^{x^{-2s}} \cdots \alpha^{x^{-rs}}\beta^{x^{-rs}} = 0 .$$

By Lemma 3.2 (iii) $\gamma$ centralizes $K[P]$ and hence since $P \lhd G$, $\gamma^{x^{-is}}$ also centralizes $K[P]$.

We use this latter fact to rearrange the terms in the above product. First since the product is

$$\gamma\gamma^{x^{-s}} \cdots \gamma^{x^{-(r-1)s}}\alpha^{x^{-rs}}\beta^{x^{-rs}}$$

we can shift the $\alpha^{x^{-rs}}$ factor past all the $\gamma^{x^{-is}}$ and obtain

$$\alpha^{x^{-rs}}\gamma\gamma^{x^{-s}} \cdots \gamma^{x^{-(r-2)s}}\alpha^{x^{-(r-1)s}}\beta^{x^{-(r-1)s}}\beta^{x^{-rs}} .$$

We next shift the $\alpha^{x^{-(r-1)s}}$ term all the way to the left and continuing this process we clearly obtain

$$(\alpha\alpha^{x^{-s}}\alpha^{x^{-2s}}\cdots\alpha^{x^{-rs}})(\beta\beta^{x^{-s}}\beta^{x^{-2s}}\cdots\beta^{x^{-rs}}) = 0\,.$$

Let $\sigma$ denote the above first factor and $\tau$ the second. Suppose that $\sigma \neq 0$. Now $P$ is an abelian $p$-group and char $K = p$ so $JK[P]$ is the unique maximal ideal of $K[P]$. This implies that every element of $K[P] - JK[P]$ is a unit in $K[P]$. If we now write $\tau$ as $\tau = \Sigma\tau_i y_i$ with $\tau_i \in K[P]$ and the $y_i$ in distinct cosets of $P$, $\sigma\tau = 0$ and $\sigma \neq 0$ therefore implies that $\tau_i \in JK[P]$ and hence $\tau \in (JK[P])K[G_0]$. But this ideal is precisely the kernel of the homomorphism $K[G_0] \to K[\bar{G}_0]$ and therefore $\bar{\tau} = 0$. Thus

$$0 = \bar{\tau} = \bar{\beta}\bar{\beta}^{\bar{x}^{-s}}\bar{\beta}^{\bar{x}^{-2s}}\cdots\bar{\beta}^{\bar{x}^{-rs}} = (\bar{\beta}\bar{x}^s)^{r+1}\bar{x}^{-rs}$$

and $(\bar{\beta}\bar{x}^s)^{r+1} = 0$.

We have therefore shown that for each $s \neq 0$ either

$$\alpha\alpha^{x^{-s}}\alpha^{x^{-2s}}\cdots\alpha^{x^{-rs}} = 0$$

for some $r = r(s) \geqq 1$ or $\bar{\beta}\bar{x}^s$ is nilpotent. If the first fact occurs for infinitely many $s$ then by Lemma 3.4, $x \in C_{G_1}(h)$ for some $h \in T$. If this first fact occurs for only finitely many $s$, then $\bar{\beta}\bar{x}^s$ is nilpotent for all but finitely many $s$ and Lemma 3.6 yields $\bar{x} \in T_{ij}$ for some $i \neq j$.

Observe that the above holds for any $x \in G_1 - W_0$. Thus we see that

$$\bar{G}_1 = S \cup \bigcup_{i \neq j} (\bar{G}_1 \cap T_{ij})$$

where

$$S = \bar{W}_0 \cup \bigcup_{h \in T} \overline{C_{G_1}(h)}\,.$$

We apply Lemma 2.1 and there are two possible conclusions. First there exists a subgroup $\bar{H}$ of $\bar{G}_1$ of finite index with $L\bar{H} \neq L\bar{G}_1$. But $[\bar{G}_0 : \bar{G}_1] < \infty$ so $[\bar{G}_0 : \bar{H}] < \infty$ and Lemma 3.1 (i) then yields $L\bar{G}_1 = L\bar{G}_0 = L\bar{H}$, and contradiction. Secondly we have $[\bar{G}_1 : S] < \infty$ and this says that $\bar{G}_1$ is a finite union of cosets of the subgroups $\bar{W}_0$ and $\overline{C_{G_1}(h)}$ for all $h \in T$. Then by Lemma 1.2 of [2] we see that one of these subgroups must have finite index in $\bar{G}_1$. Since $\bar{G}_1/\bar{W}_0$ is infinite cyclic we, therefore, have for some $h \in T$, $[\bar{G}_1 : \overline{C_{G_1}(h)}] < \infty$. Moreover, $[G : G_1] < \infty$ and $C_{G_1}(h) \supseteq P$ so this yields $[G : C_{G_1}(h)] < \infty$. Thus $h \neq 1$ and $h \in P \cap \Delta(G)$, a contradiction since $G$ satisfies (**).

We have therefore shown that $JK[G_0] = 0$. Since $[G : G_0] < \infty$, Lemma 16.8 of [2] implies that $JK[G]$ is nilpotent and the proposition is proved.

4. The main theorem. In this section we prove our result. However, we first need a few additional facts about groups satisfying condition (*).

LEMMA 4.1. *Let $G$ satisfy (*) and suppose that $P \cap \Delta(G)$ is finite. Then $JK[G]$ is nilpotent.*

*Proof.* Let $Q = P \cap \Delta(G) \lhd G$ and consider $G/Q$. Then $G/Q$ has a normal series

$$G/Q \supseteqq W/Q \supseteqq P/Q \supseteqq ZQ/Q$$

and it is trivial to see that $G/Q$ has property (*). In addition $G/Q$ satisfies (**) as follows. Let $h \in P$ with $hQ/Q \in \Delta(G/Q)$. Then the $G$ conjugates of $h$ are contained in only finitely many cosets of $Q$. Since $Q$ is finite this implies that $h \in P \cap \Delta(G) = Q$ and $hQ/Q = 1$. Thus $P/Q \cap \Delta(G/Q) = \langle 1 \rangle$ and Proposition 3.7 implies that $JK[G/Q]$ is nilpotent.

Consider the natural map $K[G] \to K[G/Q]$. Since $Q$ is a finite $p$-group the kernel of this map is the nilpotent ideal $(JK[Q])K[G]$. Moreover, we have

$$JK[G]/(JK[Q])K[G] \cong JK[G/Q]$$

and since both $JK[G/Q]$ and $(JK[Q])K[G]$ are nilpotent, the lemma is proved.

LEMMA 4.2. *Let $Q$ be a periodic normal subgroup of a group $G$ with $Q \subseteqq \Delta(G)$. Let $g, y \in G$ and suppose that $gQ/Q \in \Delta(G/Q)$. Then there exists an integer $m \geqq 1$ such that $y^m$ centralizes $g$.*

*Proof.* Since $hQ/Q \in \Delta(G/Q)$ it follows that some power $y^{m'}$ of $y$ with $m' \geqq 1$ centralizes $gQ/Q$ and thus $(y^{m'}, g) \in Q$. Moreover, since $Q$ is a periodic normal subgroup of $G$ contained in $\Delta(G)$, there exists a finite normal subgroup $H$ of $G$ with $(y^{m'}, g) \in H$. This implies that $y^{m'}$ normalizes the finite coset $Hg$ and therefore some possibly bigger power $y^m$ of $y$ centralizes $g$.

At this point we could completely determine the structure of $JK[G]$ if $G$ satisfies (*). However, we will content ourselves with observing the following key fact. If $\alpha \in K[G]$ we let

$$p\text{-Supp } \alpha = \{h \in \text{Supp } \alpha \mid h \neq 1 \text{ has order a power of } p\} .$$

PROPOSITION 4.3. *Let $G$ satisfy* (*) *and let $x \in G$. Suppose that $\alpha \in JK[G]$ with $1 \in \operatorname{Supp} \alpha$. Then there exists $h \in p\text{-Supp}\, \alpha$ and an integer $n \geqq 1$ such that $x^n$ centralizes $h$ and $hP/P \in \Delta^p(W/P)$.*

*Proof.* Let $Q = P \cap \Delta(G) \lhd G$ and consider $G/Q$. Then $G/Q$ has a normal series

$$G/Q \supseteqq W/Q \supseteqq P/Q \supseteqq ZQ/Q$$

and it is trivial to see that $G/Q$ also satisfies (*). Suppose $z \in Z$ with $zQ/Q \in \Delta(G/Q)$ and choose $y \in G$ with $G = \langle W, y\rangle$. Then Lemma 4.2 applies and we conclude that $y^m$ centralizes $z$ for some $m \geqq 1$. Since $z \in Z$ we therefore have $C_G(z) \supseteqq \langle W, y^m\rangle$ and hence

$$[G: C_G(z)] < \infty, z \in P \cap \Delta(G) = Q \quad \text{and} \quad zQ/Q = 1 .$$

We have shown that the group $G/Q$ satisfies (*) and in addition $ZQ/Q \cap \Delta(G/Q) = \langle 1\rangle$. Since $[P/Q: ZQ/Q] < \infty$ we therefore conclude that $P/Q \cap \Delta(G/Q)$ is finite and hence by Lemma 4.1, $JK[G/Q]$ is nilpotent.

Write $\alpha$ as

$$\alpha = \sum_{i=1}^{t} \alpha_i g_i$$

with $\alpha_i \in K[Q]$ and with $g_1 = 1, g_2, \cdots, g_t$ in distinct cosets of $Q$ in $G$. Since $1 \in \operatorname{Supp} \alpha$ we can assume that $1 \in \operatorname{Supp} \alpha_i$ for all $i$ and hence $g_i \in \operatorname{Supp} \alpha$.

Suppose first that $\alpha_1 \in JK[Q]$. Since $1 \in \operatorname{Supp} \alpha_1$ it follows that there exists $h \in \operatorname{Supp} \alpha_1 \subseteqq \operatorname{Supp} \alpha$ with $h \neq 1$. Then $h$ has order a power of $p$ and $h \in \Delta(G)$ so certainly $x^n$ centralizes $h$ for some $n$. Finally $hP/P = 1 \in \Delta^p(W/P)$.

Now assume that $\alpha_1 \notin JK[Q]$ and let $\sim$ denote the natural map $K[G] \to K[G/Q]$. Since $Q$ is an abelian $p$-group we see that the kernel of $\sim$ is $(JK[Q])K[G]$ and therefore for each $i$, $\tilde{\alpha}_i = a_i\tilde{1}$ for some $a_i \in K$ and by assumption $a_1 \neq 0$. Then

$$\tilde{\alpha} = \sum_{i=1}^{t} a_i \tilde{g}_i \in JK[G/Q]$$

has $\tilde{1}$ in its support. Furthermore, $JK[G/Q]$ is nilpotent so Theorem 20.2 (i) and Lemma 3.5 of [2] imply that for some $i \neq 1$, $\tilde{g}_i \in \Delta^p(G/Q)$ and $\tilde{g}_i$ has order a power of $p$. Since $Q$ is a $p$-group we see that $g_i$ has order a power of $p$ and by Lemma 4.2, $x^n$ centralizes $g_i$ for some

$n \geqq 1$. Now $g_i$ has finite order a power of $p$ and $G/W$ is infinite cyclic so $g_i \in W$. Moreover, $g_iQ/Q$ has only finitely many conjugates in $G/Q$ so certainly $g_iP/P$ has only finitely many conjugates in $W/P$. Thus $g_iP/P \in \Delta^p(W/P)$ and the proposition is proved.

The following is well known.

LEMMA 4.4. *Let $G$ be a group and let $H$ be a normal $\Delta$-subgroup of $G$. Suppose that there exists an element $h \in H$ of order $p$ with $[G: C_G(h)] = l.f.$ Then $JK[G] \cap K[H] \neq 0$.*

*Proof.* Let $h$ and $H$ be given as above and let $H^* = \langle h \rangle^H$ be the normal closure of $\langle h \rangle$ in $H$. Then $H^*$ is a finite normal subgroup of $H$ whose order is divisible by $p$. We show that $JK[H^*] \subseteqq JK[G]$. Since $JK[H^*] \neq 0$ and $JK[H^*] \subseteqq K[H]$ this will yield the result.

Since $H^*$ is finite, it clearly suffices by Lemma 17.6 of [2] to show that if $S$ is a finitely generated subgroup of $G$ with $S \supseteqq H^*$ then $JK[H^*] \subseteqq JK[S]$. Now by definition $[S: C_S(h)] < \infty$ so since $C_S(h)$ clearly normalizes $H^*$ we have $[S: N_S(H^*)] < \infty$. Let $N$ denote the core of $N_S(H^*)$ in $S$, that is the intersection of all conjugates of $N_S(H^*)$. Then $[S: N] < \infty$ and $N \lhd S$. Since $H^* \subseteqq S \cap H \lhd S$ and $S \cap H \subseteqq N_S(H^*)$ we have $H^* \subseteqq S \cap H \subseteqq N$ and clearly $H^* \lhd N$. By Lemma 19.4 of [2], $JK[H^*] \subseteqq JK[N]$ and by Theorem 16.6 of [2], $JK[N] \subseteqq JK[S]$. Thus $JK[H^*] \subseteqq JK[S]$ and the result follows.

We can now prove our main theorem.

*Proof of the Theorem.* Let $G$ be an $L$-linear group. Suppose first that there exists an element $h \in \mathscr{L}(G)$ of order $p$ with $[G: C_G(h)] = \text{l.f.}$ Then by Lemmas 1.3 and 4.4 we have $JK[G] \cap K[\mathscr{L}(G)] \neq 0$ and hence $JK[G] \neq 0$.

Conversely let us assume that $JK[G] \neq 0$. There are three cases to consider.

*Case* 1. $O_p(G) = \langle 1 \rangle$.

By definition, $\mathscr{L}(G) = \Delta^p(G)$ here and by Proposition 2.3, $JK[G]$ is nilpotent. Thus by Theorem 20.2 there exists an element $h \in \Delta^p(G)$ of order $p$. Since $h \in \Delta^p(G)$ we have $[G: C_G(h)] < \infty$ and hence $[G: C_G(h)] = \text{l.f.}$

*Case* 2. $G$ has a finite normal nonidentity $p$-subgroup.

Let this subgroup be $Q$. Then $Q \subseteqq O_p(G)$ so $Q \subseteqq \Delta(O_p(G)) \subseteqq \mathscr{L}(G)$. Let $h$ be an element of order $p$ in $Q$. Then again $h \in \Delta(G)$ implies that $[G: C_G(h)] < \infty$ and hence $[G: C_G(h)] = \text{l.f.}$

*Case* 3. $P = \boldsymbol{O}_p(G) \neq \langle 1 \rangle$ and $G$ has no finite normal nonidentity $p$-subgroups.

Set $G^* = \boldsymbol{D}_G(P)$. Since $JK[G] \neq 0$ and $P$ is nilpotent by Lemma 1.1 (i), it follows from results of [5], that $JK[G] \cap K[G^*] \neq 0$. Thus we may choose $\alpha \in JK[G] \cap K[G^*]$ with $1 \in \operatorname{Supp} \alpha$. We set $T = p\text{-}\operatorname{Supp} \alpha \cap \mathscr{L}(G)$.

Since $P$ is nilpotent and $P \neq \langle 1 \rangle$ we have $\Delta(P) \neq \langle 1 \rangle$ and hence by assumption $\Delta(P)$ is infinite. On the other hand, Lemma 1.2 (i) implies that $[\Delta(P) \colon (\Delta(P) \cap \boldsymbol{Z}(G^*))] < \infty$. Thus we can choose $h_0 \in \Delta(P) \cap \boldsymbol{Z}(G^*)$ to be an element of order $p$. We show now that in the notation of [3]

$$G = \sqrt{\boldsymbol{C}_G(h_0)} \cup \bigcup_{h \in T} \sqrt{\boldsymbol{C}_G(h)} \, .$$

Let $x \in G$ and suppose first that $xG^*/G^*$ has infinite order. We consider the group $\tilde{G} = \langle G^*, x \rangle$ and show that it satisfies condition (*). First we have the normal series

$$\tilde{G} \supseteqq G^* \supseteqq \Delta(P) \supseteqq Z$$

where $Z = \Delta(P) \cap \boldsymbol{Z}(G^*)$. By assumption $\tilde{G}/G^*$ is generated by $xG^*/G^*$ and is therefore infinite cyclic. This yields condition (1). Now $G^* \cap P = \Delta(P)$, and since $\tilde{G}/G^*$ is infinite cyclic we have $\tilde{G} \cap P = G^* \cap P = \Delta(P)$. Thus since $G/P$ is an $L$-linear group by Lemma 1.1 (ii) so is $\tilde{G}/\Delta(P) \cong \tilde{G}P/P \subseteqq G/P$. Again since $\tilde{G}/G^*$ is infinite cyclic, $\boldsymbol{O}_p(\tilde{G}) = \boldsymbol{O}_p(G^*) \lhd G$ so $\boldsymbol{O}_p(\tilde{G}) \subseteqq P \cap \tilde{G} = \Delta(P)$ and therefore $\boldsymbol{O}_p(\tilde{G}/\Delta(P)) = \langle 1 \rangle$ so condition (2) is satisfied. Moreover, Lemma 1.2 (i) clearly yields (4). Finally $\Delta(P)$ has a center of finite index so by Lemma 2.1 of [2], $\Delta(P)'$ is finite. Then this is a finite normal $p$-subgroup of $G$ so by assumption $\Delta(P)' = \langle 1 \rangle$, $\Delta(P)$ is abelian and condition (3) holds.

Thus $\tilde{G}$ satisfies (*). Now $\alpha \in JK[G] \cap K[\tilde{G}] \subseteqq JK[\tilde{G}]$ by Lemma 16.9 of [2] so Proposition 4.3 implies that there exists $h \in p\text{-}\operatorname{Supp} \alpha$ and an integer $n \geqq 1$ such that $x^n$ centralizes $h$ and $h\Delta(P)/\Delta(P) \in \Delta^p(G^*/\Delta(P))$. Note that the latter condition really says that $h \in \mathscr{L}(G)$. Thus $h \in T$ and

$$x \in \bigcup_{h \in T} \sqrt{\boldsymbol{C}_G(h)} \subseteqq \sqrt{\boldsymbol{C}_G(h_0)} \cup \bigcup_{h \in T} \sqrt{\boldsymbol{C}_G(h)} \, .$$

Now let $x \in G$ with $xG^*/G^*$ of finite order. Then $x^n \in G^*$ for some $n \geqq 1$ and hence by the choice of $h_0$, $x^n \in \boldsymbol{C}_G(h_0)$. Therefore, in this case also we have

$$x \in \sqrt{\boldsymbol{C}_G(h_0)} \subseteqq \sqrt{\boldsymbol{C}_G(h_0)} \cup \bigcup_{h \in T} \sqrt{\boldsymbol{C}_G(h)} \, .$$

Thus we have show that

$$G = \sqrt{C_G(h_0)} \cup \bigcup_{h \in T} \sqrt{C_G(h)} \,.$$

Therefore, since $G$ is a linear group, Proposition 7 of [3] implies that for some $g \in \{h_0\} \cup T$ we have $[G\colon C_G(g)] = \text{l.f.}$ Now by definition $\{h_0\} \cup T \subseteqq \mathscr{L}(G)$ and hence $g \neq 1$ is an element of $\mathscr{L}(G)$ of order a power of $p$. Finally if $h$ is an element of order $p$ in $\langle g \rangle$, then $h \in \mathscr{L}(G)$ and $C_G(h) \supseteqq C_G(g)$ so $[G\colon C_G(h)] = \text{l.f.}$ and the theorem is proved.

5. **Comments.** The preceding proof is complicated by having to handle a number of small details. In each case if our knowledge of the situation was only a little more complete, a simplification of the proof would occur. For example, the unpleasantness of the place argument in Proposition 2.5 could be avoided if we knew that $JK[G]$ was a nil ideal. In addition much of the work in § 3 would be simpler if we could assume that $P \subseteqq \Delta(W)$ or in other words if we knew that for an $L$-linear group $G$, $\Delta(P) \subseteqq \Delta(G^*)$ where $P = \boldsymbol{O}_p(G)$ and $G^* = \boldsymbol{D}_G(P)$.

Actually even a greater simplification would occur if only we could handle the equation

$$G = \bigcup_{i=1}^{n} \sqrt{H_i} \cup \bigcup_{j=1}^{m} (G \cap T_j)$$

where the $H_i$ are centralizer subgroups of $G$ and the $T_j$ are proper $L$-subspaces of $LG$ where $G$ is an $L$-linear group. We would of course want to conclude from the above that either $[G\colon H_i] = \text{l.f.}$ for some $i$ or else that some subgroup of finite index has smaller linear span than $G$. However, this does not appear to be true at least in this generality. For example we have

EXAMPLE 5.1. Consider the $2 \times 2$ linear group over the complex numbers $C$ given by

$$G = \left\{ \begin{bmatrix} 1 & 0 \\ a & b \end{bmatrix} \middle| \, a, b \in C \text{ and } b \text{ is a root of unity} \right\} .$$

Then $G$ has a normal subgroup $H$

$$H = \left\{ \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix} \middle| \, a \in C \right\}$$

isomorphic to $C^+$, the additive group of $C$. Note that $C^+$ has no proper subgroups of finite index and thus if $\tilde{G}$ is a subgroup of $G$ of finite index then $\tilde{G} \supseteqq H$ and it follows easily that $C\tilde{G} = CG$.

Let

$$T = \left\{ \begin{bmatrix} d & 0 \\ a & d \end{bmatrix} \middle| a, d \in C \right\} .$$

Then $T$ is a proper $C$-subspace of $CG$ and $H \subseteqq T$. Now suppose $x \in G - H$. Then $x = \begin{bmatrix} 1 & 0 \\ a & b \end{bmatrix}$ for some $b \neq 1$ and thus clearly the matrix $x$ is similar to $\begin{bmatrix} 1 & 0 \\ 0 & b \end{bmatrix}$. Since $b$ is a root of unity, this implies that $x$ has finite order and hence certainly $x \in \sqrt{C_G(g)}$ where $g = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

We have therefore shown that

$$G = \sqrt{C_G(g)} \cup (G \cap T)$$

and certainly $[G:C_G(g)]$ is not locally finite since $C_G(g) \cap H = \langle 1 \rangle$. Thus we see that we cannot conclude from such a decomposition of $G$ what we would like to.

Finally it would appear from the main result here and also the result for solvable groups given in [5] (or see [3] for a description of this fact) that $JK[G] \neq 0$ must imply in general that $G$ has a nonidentity normal $\Delta$-subgroup. However, this is unfortunately not the case as we see below.

Let $p$ be a prime and let $A = Z_p$ be the cyclic group of order $p$ if $p > 2$ and $A = Z_4$ if $p = 2$.

LEMMA 5.2. *Let $H$ be an infinite $p$-group and let $G$ be the Wreath product $G = A \wr H$. If $N$ is a normal $\Delta$-subgroup of $G$ then $N$ is contained in the normal abelian subgroup of $G$ which in $\Sigma A$.*

*Proof.* Write $G = WH$ where $W = \Sigma A$ is the direct sum of copies of $A$, one for each element of $H$. If $N \not\subseteq W$ choose $x \in N - W$ with $x^p \in W$. Then $N \supseteqq (x, W)$ but we see easily since $H$ is infinite that $[(x, W): C_{(x,W)}(x)] = \infty$, a contradiction.

EXAMPLE 5.3. Let $G_1$ be an infinite locally finite $p$-group and define $G_1 \subseteqq G_2 \subseteqq G_3 \subseteqq \cdots$ inductively by $G_{n+1} = A \wr G_n$. Then $G = \bigcup_{n=1}^{\infty} G_n$ is a locally finite $p$-group. If $N \neq \langle 1 \rangle$ is a normal $\Delta$-subgroup of $G$ choose $n$ so that $N \cap G_n \neq \langle 1 \rangle$. Then $N \cap G_{n+1}$ is a normal $\Delta$-subgroup of $G_{n+1} = A \wr G_n$ not contained in $\Sigma A$, a contradiction by the above lemma.

Thus $G$ has no nonidentity normal $\Delta$-subgroup. On the other hand, if $K$ is a field of characteristic $p$ then $JK[G]$ is the augmentation ideal of $K[G]$, since $G$ is a locally finite $p$-group. Therefore, $JK[G] \neq 0$.

## REFERENCES

1. D. S. Passman, *On the semisimplicity of modular group algebras II*, Canad. J. Math., **21** (1969), 1137-1145.
2. ———, *Infinite Group Rings*, Marcel Dekker, New York, 1971.
3. ———, *On the semisimplicity of group rings of linear groups*, Pacific J. Math., **47** (1973), 221-228.
4. A. E. Zalesskii, *On group rings of linear groups*, Siberian Math. J., **12** (1971), 246-250.
5. ———, *On the semisimplicity of a modular group algebra of a solvable group*, Soviet Math., **14** (1973), 101-105.

Received May 16, 1972. Research supported in part by NSF contract GP-29432.

UNIVERSITY OF WISCONSIN, MADISON

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1. 1973

# ERGODICITY IN VON NEUMANN ALGEBRAS

CHARLES RADIN

**We investigate the ergodicity of elements of a von Neumann algebra $\mathfrak{A}$ under the action of an arbitrary cyclic group of inner *-automorphisms of $\mathfrak{A}$. A simple corollary of our results is the following characterization: A von Neumann algebra $\mathfrak{A}$ is finite if and only if for each $A \in \mathfrak{A}$ and inner *-automorphism $\alpha$ of $\mathfrak{A}$, there exists $\bar{A} \in \mathfrak{A}$ such that $1/N \sum_{n=0}^{N-1} \alpha^n(A) \xrightarrow[N\to\infty]{} \bar{A}$ in the weak operator topology.**

**1. Introduction.** Our purpose is to explore in a new direction the ergodic theory of von Neumann algebras presented by Kovács and Szücs [2]. In [2] the essential contribution was the introduction of a certain restriction (called $G$-finiteness) on a group of *-automorphisms of a von Neumann algebra, fashioned so that all elements of the algebra behave ergodicly with respect to the group. Instead we consider the action of a natural class of (cyclic) groups of *-automorphisms, namely the inner ones, and investigate which elements of the algebra behave ergodicly with respect to all such groups.

**2. Behavior of infinite projections.** From the ergodic theory developed in [2], we note the following simple consequence.

THEOREM 0. (Kovács and Szücs). *Let $\mathfrak{A}$ be a finite von Neumann algebra. For each $A \in \mathfrak{A}$ and each inner *-automorphism $\alpha$ of $\mathfrak{A}$, there exists $\bar{A} \in \mathfrak{A}$ such that $1/N \sum_{n=0}^{N-1} \alpha^n(A) \xrightarrow[N\to\infty]{} \bar{A}$ in the strong operator topology.*

Our first result is a complement to this and provides a new characterization of finiteness for von Neumann algebras.

THEOREM 1. *Let $\mathfrak{A}$ be a von Neumann algebra. For each nonzero infinite projection $P \in \mathfrak{A}$ there exists an infinite projection $\theta \in \mathfrak{A}$, $\theta \leqq P$, and a unitary $U \in \mathfrak{A}$, such that $1/N \sum_{n=0}^{N-1} U^n \theta U^{-n}$ does not converge in the weak operator topology.*

First we need the following lemma.

LEMMA. *There exists a nonzero properly infinite projection $P' \leqq P$.*

*Proof.* Let $S$ be the set of all central projections $E$ of $\mathfrak{A}$ such that $EP$ is finite. $0 \in S$ so $S$ is not empty. Let $\{E_\alpha\}$ be an orthogonal family of elements of $S$. If $\sum_\alpha E_\alpha P \sim Q \leqq \sum_\alpha E_\alpha P$ (where $\sim$ is the usual equivalence relation for projections in $\mathfrak{A}$), then $E_\alpha P \sim E_\alpha Q \leqq E_\alpha P$ so that $E_\alpha Q = E_\alpha P$ and therefore $Q \geqq \sum_\alpha E_\alpha Q = \sum_\alpha E_\alpha P$. Therefore, $Q = \sum_\alpha E_\alpha P$ and $\sum_\alpha E_\alpha P$ is finite. It follows easily that there exists a (unique) maximal element $F$ in $S$. From [1, III.2.3.5] it follows that $(I-F)P$ is nonzero and infinite. Assume it is not properly infinite. Then from [1, III.2.5.9] there exists a central projection $G$ such that $0 \neq G(I-F)P$ is finite. But then from [1, III.2.3.5] $F < F + G(I-F) \in S$, which contradiction proves our lemma with $P' \equiv (I-F)P$.

*Proof of Theorem* 1. From [1, III.8.6.2] there exists a set $\{P_n \mid n \in \boldsymbol{Z}\}$ of nonzero projections $P_n \in \mathfrak{A}$ such that $P_nP_m = \delta_{n,m}P_n$ and $P_n \sim P_m$ for all $m, n \in \boldsymbol{Z}$, and such that $\sum_{|n| \leq m} P_n \xrightarrow[m \to \infty]{} P'$ in the strong operator topology. Therefore, there exist $V_n \in \mathfrak{A}$ such that $V_n^* V_n = P_n$ and $V_n V_n^* = P_{n+1}$ for all $n \in \boldsymbol{Z}$, so that $P_{n+1}V_n = V_nP_n$ and $P_n V_n^* = V_n^* P_{n+1}$ for all $n \in \boldsymbol{Z}$. Define for each $f \in \mathscr{H}$ (the Hilbert space of definition of $\mathfrak{A}$),

$$Uf = (\text{norm} \lim_{m \to \infty} \sum_{|n| \leqq m} V_nP_nf) + (I - P')f \,,$$

where the limit exists since $\| V_nP_nf \| = \| P_nf \|$ and $V_nP_nf = P_{n+1}V_nf$ so that $\{V_nP_nf \mid n \in \boldsymbol{Z}\}$ are pairwise orthogonal and

$$\sum_{|n| \leqq m} \| V_nP_nf \|^2 = \sum_{|n| \leqq m} \| P_nf \|^2 \leqq \|P'f\|^2 \,.$$

In fact $U$ is clearly a linear and norm preserving surjection, and therefore unitary. Now since

$$\Big(\sum_{|k| \leqq l} V_kP_k\Big) \text{norm} \lim_{m \to \infty} \sum_{|n| \leqq m} P_nf = \sum_{|n| \leqq l} V_nP_nf$$

it follows that $U_l \equiv I - P' + \sum_{|k| \leqq l} V_kP_k$ has $U$ as a strong operator limit as $l \to \infty$. Therefore, $U \in \mathfrak{A}$. It also follows that $UP_nU^{-1} = P_{n+1}$ for all $n \in \boldsymbol{Z}$, and so by induction $U^mP_nU^{-m} = P_{n+m}$ for all $m, n \in \boldsymbol{Z}$. Now define $g\colon \boldsymbol{N} \to \{0, 1\}$ by

$$g(n) = \begin{cases} 1 & \text{if} \quad 3^{2m} \leqq n < 3^{2m+1} \quad \text{for some} \quad m \in \boldsymbol{N} \\ 0 & \text{if} \quad 3^{2m+1} \leqq n < 3^{2m+2} \quad \text{for some} \quad m \in \boldsymbol{N} \,. \end{cases}$$

Then define $\theta$ as the strong operator limit as

$$K \to -\infty \quad \text{of} \quad \textstyle\sum_{m=K}^{0} g(-m)P_m \,,$$

and let $\psi$ be a unit vector in $P_0\mathscr{H}$. Now consider

$$\left\langle \psi, 1/N \sum_{n=0}^{N-1} U^n \theta U^{-n} \psi \right\rangle = 1/N \sum_{n=0}^{N-1} \left\langle \psi, U^n \theta U^{-n} P_0 \psi \right\rangle$$
$$= 1/N \sum_{n=0}^{N-1} \sum_{m=-\infty}^{0} g(-m) \left\langle \psi, P_{n+m} P_0 \psi \right\rangle$$
$$= 1/N \sum_{n=0}^{N-1} g(n) .$$

It is easy to see that for all $M \in N$, $1/3^{2M+1} \sum_{n=0}^{3^{2M+1}-1} g(n) \geqq 2/3$ yet $1/3^{2M+2} \sum_{n=0}^{3^{2m+2}-1} g(n) \leqq 1/3$, and the theorem is proven.

Using Theorem 0, we have immediately,

COROLLARY 1 (resp.2). *A von Neumann algebra $\mathfrak{A}$ is finite if and only if for each $A \in \mathfrak{A}$ and inner *-automorphism $\alpha$ of $\mathfrak{A}$, there exists $\bar{A} \in \mathfrak{A}$ such that $1/N \sum_{n=0}^{N-1} \alpha^n(A) \xrightarrow[N\to\infty]{} \bar{A}$ in the weak (resp. strong) operator topology.*

**3. Finite elements.** Theorem 1 raises the question of the ergodic behavior, under arbitrary inner *-automorphisms, of "finite elements" of infinite von Neumann algebras. The following theorem gives some information in this direction.

THEOREM 2. *Let $\mathfrak{A}$ be a von Neumann algebra and $\tau$ a faithful normal semi-finite trace on $\mathfrak{A}^+$ invariant under the *-automorphism $\alpha$ of $\mathfrak{A}$. Then for each $A \in \mathfrak{A}$ such that $\tau(A^*A) < \infty$, there exists $\bar{A} \in \mathfrak{A}$ such that $1/N \sum_{n=0}^{N-1} \alpha^n(A) \xrightarrow[N\to\infty]{} \bar{A}$ in the strong operator topology.*

*Proof.* First we define the following (standard) objects: see e.g. [1, I.6.2.2]

$$\| \; \|_2 : A \in \mathfrak{A} \longrightarrow [\tau(A^*A)]^{1/2}$$
$$\mathcal{N} = \{A \in \mathfrak{A} \mid \| A \|_2 < \infty\} .$$

Let $L_2$ be the abstract completion of $\mathcal{N}$ in the norm $\| \; \|_2$, and extend $\| \; \|_2$ to $L_2$ in the usual way. Let $i$ be the isometric embedding of $\mathcal{N}$ into $L_2$. $L_2$ is a Hilbert space with the obvious addition and scalar multiplication, and inner product $<, >$ defined as the extension to $L_2 \times L_2$ of

$$\tau: A \times B \in \mathcal{N} \times \mathcal{N} \longrightarrow \tau(A^*B) .$$

We note the simple inequalities

$$\| AB \|_2 \leqq \| A \| \, \| B \|_2 \qquad \text{for all } B \in \mathcal{N},\ A \in \mathfrak{A}$$
$$\| AB \|_2 \leqq \| A \|_2 \, \| B \| \qquad \text{for all } B \in \mathcal{N},\ B \in \mathfrak{A} .$$

We then define the $C^*$-representation $\pi$ of $\mathfrak{A}$ on $L_2$ by

$$\pi(A)i(B) \equiv i(AB)$$

and noting that $\| \pi(A)i(B) \|_2 = \| AB \|_2 \leqq \| A \| \| B \|_2$ so that $\pi(A)$ extends uniquely to $L_2$ by continuity. It is easy to see that $\pi$ is faithful and normal and that

$$U: i(B) \longrightarrow i(\alpha[B]) \qquad \text{for } B \in \mathcal{N}$$

extends to a unitary operator on $L_2$. Defining, for $B \in \mathfrak{A}$,

$$B_N = \frac{1}{N} \sum_{n=0}^{N-1} \alpha^n(B), \quad \text{we know by von Neumann's}$$

mean ergodic theorem that for each $A \in \mathcal{N}$, $i(A_N)$ is $\| \ \|_2$-Cauchy. Define for each $B \in \mathcal{N}$,

$$D_A: i(B) \longrightarrow \text{norm} \lim_{N \to \infty} \pi(A_N)i(B)$$

which limit exists since

$$\| \pi(A_N - A_M)i(B) \|_2 \leqq \| A_N - A_M \|_2 \| B \| .$$

$D_A$ is obviously linear. Furthermore,

$$\| D_A i(B) \|_2 = \lim_{N \to \infty} \| \pi(A_N)i(B) \|_2 \leqq \| A \| \| B \|_2$$

so $D_A$ extends uniquely to a bounded operator on $L_2$ by continuity. It is easy to see that $\pi(A_N)$ converges to $D_A$ in the strong operator topology. Since $\pi$ is normal, $\pi(\mathfrak{A})$ is strong operator closed [1, I.4.3.2] so there exists $\bar{A} \in \mathfrak{A}$ such that $D_A = \pi(\bar{A})$. Since $\pi$ is faithful, $A_N \xrightarrow[N \to \infty]{} \bar{A}$ in the strong operator topology [1, I.4.3.1].

COROLLARY 1. *Let $\mathfrak{A}$ be a countably decomposable von Neumann algebra. For each finite projection $P \in \mathfrak{A}$ and inner *-automorphism $\alpha$ of $\mathfrak{A}$, there exists $\bar{P} \in \mathfrak{A}$ such that*

$$\frac{1}{N} \sum_{M=0}^{N-1} \alpha^n(P) \xrightarrow[N \to \infty]{} \bar{P} \quad \textit{in the strong operator topology} .$$

*Proof.* Let

$$A \in \mathfrak{A} \longrightarrow A_1 \oplus A_2 \in \mathfrak{A}_1 \oplus \mathfrak{A}_2$$

be the canonical decomposition of $\mathfrak{A}$ into its countably decomposable semi-finite and purely infinite components. From [1, I.6.7.9] we know that any finite countably decomposable von Neumann algebra has a faithful, normal, tracial state. Inserting this fact into the proof of

[3, 2.5.3], we see that there exists a countable faithful family $\{\tau_n \mid n \in N\}$ of normal semi-finite traces on $\mathfrak{A}_1^+$ with pairwise orthogonal supports such that $\tau_n(P_1) < \infty$ for all $n \in N$. Define

$$\tau' = \sum_{n=0}^{\infty} \tau_n / [\tau_n(P_1) + 2]^n$$

on $\mathfrak{A}_1^+$; it is faithful, normal and semi-finite. Since $\alpha$ is also inner for $\mathfrak{A}_1$ and therefore leaves $\tau'$ invariant, we may apply Theorem 2 to $\mathfrak{A}_1$. Since $P_2 = 0$ from [1, III.2.4.8], we are finished.

In the countably decomposable case, Theorem 2 gives us an essentially different proof of Theorem 0, namely

COROLLARY 2. *Let $\mathfrak{A}$ be a finite countably decomposable von Neumann algebra. For each $A \in \mathfrak{A}$ and inner $*$-automorphism $\alpha$ of $\mathfrak{A}$, there exists $\bar{A} \in \mathfrak{A}$ such that*

$$\frac{1}{N} \sum_{n=0}^{N-1} \alpha^n(A) \xrightarrow[N\to\infty]{} \bar{A} \quad \textit{in the strong operator topology}\,.$$

*Proof.* Just combine the existence of a faithful finite normal trace on $\mathfrak{A}^+$ [1, I.6.7.9] with Theorem 2.

## REFERENCES

1. J. Dixmier, *Les algèbres d'opérateurs dans l'espace Hilbertien*, 2nd Ed., Paris, 1969.
2. I. Kovács and J. Szücs, *Ergodic type theorems in von Neumann algebras*, Acta Sci. Math., **27** (1966), 233-246.
3. S. Sakai, *C*-Algebras and W*-Algebras*, New York, 1971.

Received June 6, 1972. Research supported by AFOSR under Contract F44620-71-C-0108.

PRINCETON UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# ON THE SINGULARITIES OF THE FUNCTION GENERATED BY THE BERGMAN OPERATOR OF THE SECOND KIND

P. ROSENTHAL

**Let $\psi(\lambda, y) = P_2(f)$ be Bergman's integral operator of the second kind with domain of definition**

$$W = \{(\lambda, y) \mid 3^{1/2} \mid \lambda \mid < y,\ \lambda \leqq 0,\ y > 0\}\ .$$

**Let $f(q) = (q - A)^{-1}$, $A \in W$. In this paper it is shown that $\psi(\lambda, y)$ has singular points $z = 2A$, $2A(1 - w)$, where $w = A^{-1}\lambda$ and $z = \lambda + iy$.**

Let

$$\psi(z, z^*) = P_2(f) = \int_l E(z, z^*, t) f\left(\frac{z}{2}(1 - t^2)\right) \frac{dt}{\sqrt{1 - t^2}}$$

be Bergman's integral operator of the second kind. $P_2(f)$ maps functions $f$ analytic in one variable in the neighborhood of the origin into solutions of the linear partial differential equation

$$\psi_{zz^*} + N\left(\frac{z + z^*}{2}\right)(\psi_z + \psi_{z^*}) = 0\ , \quad z = \lambda + iy\ , \quad z^* = \lambda - iy\ ,$$

$N(\lambda) = -(1/12\lambda)(1 + b_1(-\lambda)^{2/3} + \cdots)$ is analytic for $-\infty < \lambda < 0$ and singular at $\lambda = 0$. $E(z, z^*, t)$, called the generating function of the operator, is analytic in the three variables $z$, $z^*$, and $t$ providing $\mid z + z^* \mid < \mid t^2 z \mid$, $l$ is some rectifiable Jordan curve in the upper complex $t$-plane connecting the points $-1$ and 1, [1], [3].

In a previous paper [7] we obtained some results on the singularities of $P_2(f)$ where $f$ is meromorphic and $z$, $z^*$ were treated as independent complx variables. In this paper we let $z^* = \bar{z}$ (conjugate of $z$) and $N(\lambda) = -1/12\lambda$ (Tricomi case). With these assumptions,

$$\psi(\lambda, y) = \int_{-1}^{1} E(u) \frac{f(q)}{\sqrt{1 - t^2}} dt\ , \quad \text{where} \quad u = \frac{t^2 z}{2\lambda}\ ,$$

$$(1) \qquad q = \frac{1}{2} z(1 - t^2)\ ,$$

$$z = \lambda + iy\ ,$$

$E(u) = H(\lambda)(F^{(1)}(u) + F^{(2)}(u))$, $F^{(1)}(u) = C_1 u^{-1/6} F_1(1/6, 2/3, 1/3, 1/u)$, $F^{(2)}(u) = C_2 u^{-5/6} F_2(5/6, 4/3, 5/3; 1/u)$, $F_j$ is the hypergeometric function $j = 1, 2$, $H(\lambda) = C_3 \lambda^{-1/6}$, $C_j$ are constants, $j = 1, 2, 3$, $(\lambda, y) \in W = \{(\lambda, y) \mid 3^{1/2} \mid \lambda \mid < y,\ \lambda \leqq 0,\ y > 0\}$,

$l$ (the path of integration $= \{t \mid t = e^{i\theta}, 0 \leqq \theta \leqq \pi\}$, [4, p. 107] .

THEOREM. *Let* $f(q) = (A - q)^{-1}$, $A = \lambda_0 + iy_0 \in W$, $\lambda/A = w = s + i\sigma$, $z = \lambda + iy$, $S_1 = \{(w, z) \mid z = 2A, \pi/2 \geqq \arg w \geqq \alpha_1, \pi/2 > \alpha_1 > \pi/3, 0 < \delta_1 \leqq |w| \leqq 1/4 - \delta_2, 1/4 > \delta_1, \delta_2 > 0, 1/4 > \delta_1 + \delta_2\}$,

$$S_2 = \{(w, z) \mid z = 2A(1 - w), \textit{ same conditions on } w \textit{ as in } S_1\} ,$$

$S_3 = \{(0, 2y_0)\}$. *Let* $T = S_1 \cup S_2 \cup S_3$. *Then* $T$ *is a singular set for at least one of the branches of* $\psi(w, z)$ *defined in* (1).

*Proof.* We consider first the case where $E(u) = H(\lambda)F^{(1)}(u)$.

*Domain considerations.* (3), (4) imply $\psi(w, z)$ is analytic function of the two complex variables $w$, $z$ for disc neighborhoods satisfying $0 < |w| < 1/4$, $|A/2| < |z| < |A|$, where we have extended $\lambda$ to the complex variable $Aw$. Note (1) implies we must specify branch cuts in our definition of $\psi(w, z)$. Since $z = \lambda + iy$ (see (1)), we must also consider the extension of $\lambda$, $y$ to complex values subject to the above inequalities. Thus we can also obtain nonempty neighborhoods $N_\delta(\lambda)$, $N_\delta(y)$ such that $\psi(\lambda, y)$ is an analytic function in $\lambda$, $y$, where $\lambda$, $y$ now have been extended to complex values.

In what follows we treat $\psi(w, z)$ as an analytic function in $z$ for fixed $w$.

Consider the function obtained from (1) where we have used the series definition for $F_1(u)$,

$$f(\lambda, y) = \int_{-1}^{1} \sum_{p=0}^{\infty} \frac{a_p t^{-2p}(2\lambda)^p}{z^p} \cdot \sum_{p=0}^{\infty} \left(\frac{z}{2A}\right)^p (1 - t^2)^p t^{-1/3} \frac{dt}{\sqrt{1 - t^2}} , \tag{2}$$

$a_p = (\Gamma(p + 1/6)\Gamma(p + 2/3)/\Gamma(p + 1/3)\Gamma(p + 1))$, $\Gamma$ is the Gamma function, $|z| < A$, $|2\lambda| < |z|$. From (2) we obtain two series,

$$\begin{aligned} &\sum_{p=0}^{\infty} \left(\sum_{k=0}^{\infty} a_k \left(\frac{\lambda}{A}\right)^k t^{-2k-1/3}(1 - t^2)^{p+k-1/2}\right) \left(\frac{z}{2A}\right)^p \\ &+ \sum_{p=1}^{\infty} \left(\sum_{k=0}^{\infty} a_{p+k} \left(\frac{\lambda}{A}\right)^k t^{-(2(p+k+1/6))}(1 - t^2)^{k-1/2} \left(\frac{2\lambda}{z}\right)^p\right) \\ &|z| < |A| , \qquad |2\lambda| < |z| . \end{aligned} \tag{3}$$

We will limit ourselves to the first series in (3) for our analysis of the singularities of $P_2(f)$. When $|\lambda| \leqq |A/2| - \delta$, $|z| \leqq |A| - \delta$, $|A/2| > \delta > 0$, the operations of summation and integration (with respect to $t$) can be interchanged in the first series of (3), our integrals are in the improper Riemann sense. Integrating the first part of (3) by parts, then using the formula,

$$\int_{-1}^{1} t^{-1/3}(1-t^2)^{\nu+1/6}\frac{dt}{\sqrt{1-t^2}} = -\frac{1}{2}(1-e^{2\pi i/3})\frac{\Gamma(1/3)\Gamma(\nu+2/3)}{\Gamma(\nu+1)},$$

[2, p. 33], we obtain the function

$$f_1(w, z) = \sum_{p=0}^{\infty} \beta_p(w)\left(\frac{z}{2A}\right)^p, \tag{4}$$

where

$$\beta_p(w) = \sum_{k=0}^{\infty} A_{pk}w^k, \quad A_{pk} = \frac{\Gamma(k+1/6)\Gamma(k+p+1/2)}{\Gamma(k+1/3)\Gamma(k+1)\Gamma(p+5/6)}, \quad w = \frac{\lambda}{A}. \tag{5}$$

(5) can be rewritten as

$$\beta_p(w) = \alpha_p c_p(w), \qquad \alpha_p = \frac{\Gamma(1/6)\Gamma(1/2+p)}{\Gamma(1/3)\Gamma(5/6+p)}, \tag{6}$$

$c_p(w) = F_3(1/6, 1/2+p, 1/3;\ w)$, $F_3$ a hypergeometric function. Using the asymptotic formula for $F_3$ for large $p$ [6, pp. 235, 241 (23)], we can write $c_p(w)$ as

$$\begin{aligned} c_p(w) &= a_p(w)e^{i\pi/6}R_1 + a_p(w)(1-w)^{-1/3-p}R_2, \\ a_p(w) &= \left(\Gamma\left(\frac{1}{6}\right)\right)^{-1}\Gamma\left(\frac{1}{3}\right)(pw)^{-1/6}, \end{aligned} \tag{7}$$

$p$ sufficiently large, $w \in T_1 = \{w \mid 0 < \delta_1 \leqq |w| \leqq 1/4 - \delta_2,\ 1/4 > \delta_2,\ \delta_1 > 0,\ 1/4 > \delta_1 + \delta_2,\ \pi/2 \geqq \arg w \geqq \alpha_1,\ \pi/2 > \alpha_1 > \pi/3\}$,

$$R_j(p, w) = 1 + R_0^{(j)}(p, w),$$

$\lim_{p\to\infty} pR_0^{(j)}(p, w) = h_j(w) \neq 0$ uniformly for $w \in T_1$, $j = 1, 2$. Using (6), (7), we can rewrite (4) as

$$\begin{aligned} f_1(w, z) &= \sum_{p=0}^{p=p_0\geqq 1} \alpha_p c_p(w)\left(\frac{z}{2A}\right)^p + \sum_{p=p_0+1}^{\infty} c_1(p, w)z_1^p + \sum_{p=p_0+1}^{\infty} c_2(p, w)z_2^p, \\ z_1 &= \frac{z}{2A}, \qquad z_2 = (1-w)^{-1}\frac{z}{2A}, \end{aligned} \tag{8}$$

and

$$c_1(p, w) = \alpha_p a_p(w)e^{i\pi/6}R_1, \qquad c_2(p, w) = \alpha_p a_p(w)(1-w)^{-1/3}R_2, \tag{9}$$

see (6) for the definition of $\alpha_p$, (7) for $a_p(w)$.

From (9) we obtain

$$\rho = \lim_{p\to\infty} |c_j(p, w)|^{-1/p} = 1, \tag{10}$$

the radius of convergence of the second and third series in (8), and $-\varepsilon < \arg c_j(p, w) < \varepsilon$, $0 < \varepsilon < \pi/2$, $p$ sufficiently large, $w \in T_1$, $j = 1, 2$.

*Proof of* (10). From (7) we obtain

$$1 + \varepsilon \geqq |R_j(p, w)| \geqq 1 - \varepsilon > 0 ,$$

$1 > \varepsilon > 0$, $p$ sufficiently large, $w \in T_1$. So we can take the $p$th root (say principle branch) of $c_j(p, w)$, $j = 1, 2$, cf (9).

Using the asymptotic formula $(\Gamma(p + A)/\Gamma(p + B)) \sim p^{A-B}$, we conclude the first part of (10). Since $\lim_{p\to\infty} (1 + R_0^{(j)}(p, w) = 1$, $w \in T_1$, see (7), the second part of (10) follows.

$$\text{(11)} \qquad z = 2A \quad \text{and} \quad z = 2A(1 - w) , \quad w \in T_1 ,$$

are singular points of (8).

*Proof of* (11). (10) satisfies the hypotheses of a theorem of Dienes [5, p. 227]. From this theorem we conclude $z = 2A$ and $z = 2A(1 - w)$ are singular points respectively of the second and third series in (8). Further, $c_j(p = \xi = \rho e^{i\psi}, w)$ (see (9)) is an analytic function in $\xi$ in the half-plane $x_1 \geqq 1$, $\xi = x_1 + iy_1$, and

$$|c_j(1 + \rho e^{i\psi}, w)| < e^{\varepsilon\rho} , \ \varepsilon > 0 ,$$

and arbitrarily small, $\rho > 0$ and sufficiently large, and $-\pi/2 \leqq \psi \leqq \pi/2$, $w \in T_1$, $j = 1, 2$. This follows from a definition of the remainder term $R_0^{(j)}(p, w)$ of (7), see [6, p. 235]. Hence by a theorem of Le Roy and Lindelöf [5, p. 340], we conclude the only possible singular points of the second series in (8) are the points on the ray $\varphi = \varphi_0$, $\varphi_0 = \arg 2A$, joining $2A$ to infinity and the only possible singular points of the third series in (8) are the points on the ray $\varphi = \theta_0$, $\theta_0 = \arg 2A(1 - w)$, $w \in T_1$, joining $2A(1 - w)$ to infinity. Further, $\arg 2A \neq \arg(2A(1 - w))$, $w \in T_1$. Hence the singular points $z = 2A$, $z = 2A(1 - w)$, $w \in T_1$, of the second and third series respectively are not removed upon addition of these two series in (8). This completes the proof of (11).

$$\text{(12)} \qquad (0, 2y_0) \text{ is a singular point of } \psi(w, z) .$$

*Proof*. Let $w = \lambda/A = \lambda_0 = 0$. (3) then reduces to the first series, and (4) reduces to the hypergeometric function $F_4(1, 1/2, 5/6; (y/2y_0))$ times a constant. $F_4$ is singular at the point $y = 2y_0$, so (12) holds.

From (11), (12) we conclude $T$ is a singular set (see Theorem for the definition of $T$) of $\psi(w, z)$ for the case $F_1$.

*Proof.* We note the second series in (3) when integrated with respect to $t$ gives rise to a function $f_2(w, z)$ which is regular at the points in $T$.

For the case $F_2$ (see (1)) we use the formula

$$\int_{-1}^{1} t^{-5/3}(1-t^2)^{\nu+5/6}\frac{dt}{\sqrt{1-t^2}} = \frac{1}{2}(1-e^{-(2\pi i/3)})\frac{\Gamma(-1/3)\Gamma(\nu+4/3)}{\Gamma(\nu+1)}$$

[2, p. 33].

Proceeding as above, we then conclude $T - \{(0, 2y_0)\}$ is a singular set for the case $F_2$. (1) thus can be written as the sum of two functions,

$$\psi(w, z) = \frac{1}{z^{5/6}}(g(w, z) = z^{2/3}P_1(w, z) + P_2(w, z)) , \tag{13}$$

where $P_j$ is singular at the points in $T - \{(0, 2y_0)\}$, $j = 1, 2$. This follows from the linearity of the operator $P_2(f)$.

(14) At least one of the branches of $g(w, z)$ of (15) is singular for points in $T - \{(0, 2y_0)\}$ .

*Proof of* (14). $z^{2/3}$ can be one of the three branches,

$$\alpha_1 = R^{2/3}e^{i2/3\theta} , \quad \alpha_2 = R^{2/3}e^{i(2/3\theta+2/3\pi)} , \quad \alpha_3 = R^{2/3}e^{i(2/3\theta+4/3\pi)}, \ \pi > \theta > -\pi .$$

We form the sum

$$\sum_{i=1}^{3} g_i(w, z) = \sum_{i=1}^{3} \alpha_i P_1(w, z) + 3P_2(w, z) .$$

We note $\sum_{i=1}^{3} \alpha_i P_1(w, z) = 0$, $|w| < 1/4$; $|A/2| < |z| < |A|$ (see (3)). So if all the branches of $\psi(w, z)$ in (13) were regular at the points in $T - \{(0, 2y_0)\}$, then $P_2(w, z)$ would be regular at the same points, a contradiction. For $w = \lambda/A = \lambda_0 = 0$, $P_2(0, z) = 0$, hence $(0, 2y_0)$ is a singular point for all branches (13) (see (12)). This completes the proof of our Theorem.

## References


1. S. Bergman, *Two-dimensional transonic flow patterns*, Amer. J. Math., **70** (1948), 856-891.
2. ———, *On an initial value problem in the theory of two-dimensional transonic flow patterns*, Pacific J. Math., **32** (1970), 29-46.
3. S. Bergman and R. Bojanic, *Application of integral operators to the theory of partial differential equations with singular coefficients*, Arch. Rational Mech. Anal., **4** (1962), 323-340.
4. ———, *Integral Operators in the Theory of Linear Partial Differential Equations*, Springer-Verlag, Band 23, 3rd ed., 1971.

5. P. Dienes, *The Taylor Series*, Oxford University Press, 1931.
6. Luke Yudell, *The Special Functions and Their Approximation*, Vol. 1, Academic Press, 1969.
7. P. Rosenthal, *On the Bergman integral operator for an elliptic partial differential equation with a singular coefficient*, Pacific J. Math., **35** (1970), 493-497.

Received June 16, 1972. This work was supported in part by NSF Grant DP-33779 at Stanford University.

STANFORD UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# MULTIPLICATIONS ON HOMOGENEOUS SPACES, NONASSOCIATIVE ALGEBRAS AND CONNECTIONS

A. A. SAGLE AND J. R. SCHUMI

**In this paper we show how nonassociative algebras over the real numbers arise from multiplications on certain homogeneous spaces; that is, an analytic function $\mu: M \times M \rightarrow M$. Then these algebras are used to obtain an invariant connection $\nabla$ on the homogeneous space and we give some applications of nonassociative algebras to these topics. Conversely every finite dimensional nonassociative algebra over the real numbers arises from an invariant connection and a local multiplication on a homogeneous space. Thus, analogous to the theory of Lie groups and Lie algebras, much of the basic theory of nonassociative algebras can be formulated in terms of multiplications and connections and conversely.**

**1. Introduction.** Let $G$ be a connected Lie group with Lie algebra $g$ and let $H$ be a closed subgroup with Lie algebra $h$. Then the pair $(G, H)$ or $(g, h)$ is called a *reductive pair* if there exists a subspace $m$ of $g$ so that $g = m + h$ (subspace direct sum) and $(\operatorname{Ad} H)(m) \subset m$. The corresponding homogeneous space $G/H$ is called a *reductive homogeneous space* which is an analytic manifold. An analytic function

$$\mu: G/H \times G/H \longrightarrow G/H$$

such that $\mu(\bar{e}, \bar{e}) = \bar{e} = eH$ is called a *multiplication* on $G/H$; for example, Lie groups, Moufang loops and certain $H$-spaces are reductive homogeneous spaces with a multiplication.

The nonassociative algebras arise from studying the local behavior of a multiplication $\mu$ on $G/H$ which we now consider. Thus let $\pi$: $G \rightarrow G/H$ be the natural projection and let $g = m + h$ be a fixed (reductive) decomposition. From [1, p. 113] we know that for the map $\psi = \exp | m$ there exists a neighborhood $U$ of 0 in $m$ which is mapped homeomorphically into $G$ under $\psi$ and such that $\pi$ maps $\psi(U)$ homeomorphically onto a neighborhood $N^*$ of $\bar{e}$ in $G/H$. Thus by the analyticity of $\mu$ and $\pi \circ \psi$ there exists a neighborhood $D$ of 0 in $m$ contained in $U$ so that for all $X, Y \in D$

$$\mu(\pi \exp X, \pi \exp Y) = \pi \exp F(X, Y)$$

is in $N^*$ where $F: D \times D \rightarrow U$ is a function which is analytic at $\theta = (0, 0) \in m \times m$. Thus $\mu$ is determined locally by $F$ which has the Taylor's series expansion [5] $F(X, Y) = F(\theta) + F^1(\theta)(X, Y) + 1/2F^2(\theta)(X, Y)^2 +$

$\cdots$ for $X, Y \in D$. We will show that $F(\theta) = 0$ and the function

$$\alpha(X, Y) = F^2(\theta)[(X, 0), (0, Y)]$$

is bilinear. Therefore the multiplication $\mu$ on $G/H$ determines a nonassociative algebra with linear space $m$ and composition $\alpha: m \times m \to m$ and we denote this algebra by $(m, \alpha)$.

Next we require that for all $u \in H$ the mappings

$$\tau(u): G/H \longrightarrow G/H: yH \longrightarrow uyH$$

are automorphisms of the multiplication $\mu$ on $G/H$, and call the pair $(G/H, \mu)$ with $\tau(H) \subset \mathrm{Aut}\,(G/H, \mu)$ a *multiplicative system*. We show that $\tau(H) \subset \mathrm{Aut}\,(G/H, \mu)$ implies $\mathrm{Ad}\,H$ is in the automorphism group of the algebra $(m, \alpha)$ and this allows us to use the result [6] which gives a bijective correspondence between $G$-invariant connections $\nabla$ on $G/H$ and nonassociative algebras $(m, \alpha)$ with $\mathrm{Ad}\,H \subset \mathrm{Aut}\,(m, \alpha)$. Thus a multiplicative system $(G/H, \mu)$ induces a $G$-invariant connection via the algebra $(m, \alpha)$. Conversely, we show that every such algebra comes from a local multiplicative system. In particular, any finite dimensional nonassociative algebra $A$ over $R$ can be regarded as an algebra $(g, \alpha)$ for a Lie algebra $g$ of suitable dimension and consequently $A$ arises from a local multiplicative system defined on $G$ and also from a $G$-invariant connection defined on $G$.

The above multiplicative systems $(G/H, \mu)$ are too general and not particularly related to the action of $G$ on $G/H$. We will now describe an "invariance" restriction for the multiplicative system $(G/H, \mu)$. The multiplication $\mu$ defines a function $\measuredangle(\mu, X): G/H \to T(G/H)$ from $G/H$ to the tangent bundle of $G/H$ as follows; (see [9] for the case of Lie groups). Let $T(G/H, \bar{a})$ denote the tangent space at $\bar{a} \in G/H$; thus $T(G/H, \bar{e}) = m$. Let $T$ denote the differential of a function, then for each $X \in m$ we set

$$\measuredangle(\mu, X)(\bar{a}) = [T\mu(\bar{a}, \bar{e})](0, X) \text{ for } (0, X) \in T(G/H, \bar{a}) \times m\ .$$

That is, $\measuredangle(\mu, X)(\bar{a})$ is the differential of $\mu$ evaluated at $(\bar{a}, \bar{e})$ on $(0, X)$. The function $\measuredangle(\mu, X)$ is a vector field if and only if $(G/H, \mu)$ has $\bar{e}$ as a right identity element; also the vector field is analytic and depends linearly on $X$. For $\Gamma$ a subset of $\tau(G)$ containing $\tau(e)$, whose precise definition will be given in §4, we say that $\mu$ is $\Gamma$-invariant if for all $X \in m$, $\measuredangle(\mu, X)$ is a vector field invariant under all the maps $\tau(a): G/H \to G/H: \bar{x} \to a\bar{x}$ for all $\tau(a) \in \Gamma$. That is, for all $\tau(a) \in \Gamma$, and all $X \in m$, $\measuredangle(\mu, X)$ satisfies

$$T\tau(a)(\bar{e})\measuredangle(\mu, X)(\bar{e}) = \measuredangle(\mu, X)(\tau(a)\bar{e})\ .$$

In particular if $(G, \mu)$ is a Lie group, then $\Gamma$ can be taken to be

$L(G) = \{L(a): a \in G\}$ and $\ell(\mu, X)$ is the usual left invariant vector field generated by $X$, so our results are consistent with Lie theory. In the case of a $\Gamma$ invariant multiplication $\mu$, we obtain the connection induced by $\mu$ is given by the algebra $(m, \alpha)$ with $\alpha(X, Y) = F^2(\theta)[(X, 0), (0, Y)] = 1/2[XQY]$ where $Q$ is the endomorphism of $m$ given by $Q: m \to m: Y \to F^1(\theta)(0, Y)$. Although we have used a global multiplication in the above discussion, most of the results concern the algebra $(m, \alpha)$, thus it suffices to consider local multiplications on $G/H$. However the globalization of these local results present many topological problems. For example, every sphere is a reductive homogeneous space $G/H$ and consequently has in a suitable neighborhood of $\bar{e} = eH$ a local multiplication with $\bar{e}$ as an identity element. Thus any sphere is a local $H$-space but only $S^1$, $S^3$, and $S^7$ are global $H$-spaces.

**2. Multiplications.** Using the notations of §1 we have for $X$, $Y$ in a suitable neighborhood $D$ of 0 in $m$ that

$$\mu(\pi \exp X, \pi \exp Y) = \pi \exp F(X, Y)$$

where $F: D \times D \to U$ is analytic at $\theta = (0, 0) \in m \times m$ and $U \subset D$ is a neighborhood 0 in $m$ for which $\pi \circ \exp = \pi \circ \psi$ is a diffeomorphism. Thus analogous to local Lie groups we have a local multiplication system $(U, F)$.

We now consider the Taylor's series for $F$ near the origin $\theta = (0, 0) \in m \times m$. Thus for $Z = (X, Y) \in m \times m$ and $Z^k = (Z, \cdots, Z)$ $k$-times we have for $t$ in a suitable interval $(-\delta, \delta) \subset R$ that

$$F(tX, tY) = F(\theta) + tF^1(\theta)Z^1 + t^2/2F^2(\theta)Z^2 + \cdots$$

where $F^k(\theta) = D^kF(\theta)$ is the $k$th derivative of $F$ at $\theta$ and is regarded as a symmetric $k$-linear function on $(m \times m)^k$ into $m$ [5]. In particular since $\mu(\bar{e}, \bar{e}) = \bar{e}$, $F(\theta) = F(0, 0) = 0$. Next writing $Z = (X, 0) + (0, Y)$ we see that

$$\begin{aligned} DF(\theta)Z &= DF(\theta)[(X, 0) + (0, Y)] \\ &= [DF(\theta)](X, 0) + [DF(\theta)](0, Y) \\ &= PX + QY \end{aligned}$$

where $[DF(\theta)](X, 0)$ is regarded as a linear function of $X$ and denoted by $PX$ with $P$ an endomorphism of $m$ and similarly $[DF(\theta)](0, Y) = QY$.

Using the symmetric bilinearity of $D^2F(\theta)$ we obtain for $Z = (X, Y)$

$$\begin{aligned} 1/2[F^2(\theta)](Z, Z) &= 1/2F^2(\theta)[(X, 0) + (0, Y), Z] \\ &= 1/2F^2(\theta)[(X, 0), (X, 0)] + 1/2F^2(\theta)[(0, Y), (0, Y)] \\ &\quad + F^2(\theta)[(X, 0), (0, Y)] \,. \end{aligned} \tag{2.1}$$

Now we note that for $X, Y \in m$

$$\alpha(X, Y) = F^2(\theta)[(X, 0), (0, Y)]$$

defines a bilinear function $\alpha$ on $m \times m$ into $m$ as follows. For $a \in R$ and $X, Y, Z \in m$

$$\begin{aligned} \alpha(aX + Z, Y) &= F^2(\theta)[(aX + Z, 0), (0, Y)] \\ &= F^2(\theta)[(aX, 0) + (Z, 0), (0, Y)] \\ &= F^2(\theta)[(aX, 0), (0, Y)] + F^2(\theta)[(Z, 0), (0, Y)] \\ &= a\alpha(X, Y) + \alpha(Z, Y) \end{aligned}$$

and similarly $\alpha$ is right-linear. Thus $m$ with the bilinear function $\alpha$ becomes a nonassociative algebra denoted by $(m, \alpha)$.

Note that the converse is true locally. Thus given a nonassociative algebra $(m, \alpha)$ we can find a neighborhood $D$ of 0 in $m$ so that $\mu(\pi \exp X, \pi \exp Y) = \pi \exp (X + Y + \alpha(X, Y))$ defines a local multiplicative system on some neighborhood $N^*$ of $\bar{e}$; this is analogous to formal Lie groups. Furthermore note that this multiplicative system has $\bar{e}$ as a two-sided identity and for $F(X, Y) = X + Y + \alpha(X, Y)$ we have $1/2F^2(\theta)(X, Y)^2 = \alpha(X, Y)$.

If the multiplicative system $(G/H, \mu)$ has $\bar{e} = eH$ as a right identity $(\mu(\bar{a}, \bar{e}) = \bar{a})$, then in the above notation

$$P = I \text{ and } F^k(\theta)(X, 0)^k = 0 \,.$$

For if $t$ is in a suitable interval $(-\delta, \delta)$ of $R$ we have for $X \in m$ that

$$\pi \exp tX = \mu(\pi \exp tX, \pi \exp 0) = \pi \exp F(tX, 0)$$

and since $\pi \circ \exp$ suitably restricted to $m$ is a diffeomorphism as previously discussed we have

$$t\,X = F(tX, 0) = t\,PX + \frac{t^2}{2}F^2(\theta)(X, 0)^2 + \cdots$$

for $t$ in a suitable interval about 0 in $R$. Differentiating this formula at $t = 0$ gives the results. A similar result holds if $(G/H, \mu)$ has $\bar{e}$ as a left identity. Thus if $(G/H, \mu)$ is an $H$-space; that is, $\bar{e}$ is a two-sided identity, then $F(X, Y) = X + Y + \alpha(X, Y) + \cdots$.

3. **Automorphisms.** From the bijective correspondence between $G$-invariant affine connections on $G/H$ and nonassociative algebras $(m, \alpha)$ [6, 8] we see that $\operatorname{Ad} H$ must be in the automorphism group, $\operatorname{Aut}(m, \alpha)$, of the algebra $(m, \alpha)$. We shall show in this section that this condition is implied by $\tau(H) \subset \operatorname{Aut}(G/H, \mu)$; thus we want to consider multiplicative systems $(G/H, \mu)$ with $\tau(H) \subset \operatorname{Aut}(G/H, \mu)$.

DEFINITION. An analytic diffeomorphism $\eta: G/H \rightarrow G/H$ is an *automorphism* of $(G/H, \mu)$ if $\eta(\bar{e}) = \bar{e}$ and $\eta\mu(\bar{x}, \bar{y}) = \mu(\eta\bar{x}, \eta\bar{y})$ for all $\bar{x}, \bar{y} \in G/H$. We denote the set of such automorphisms by $\mathrm{Aut}\,(G/H, \mu)$. An endomorphisms $S \in GL(m)$ is an *automorphism* of the algebra $(m, \alpha)$ if $S\alpha(X, Y) = \alpha(SX, SY)$ for all $X, Y \in m$. We denote the set of such automorphisms by $\mathrm{Aut}\,(m, \alpha)$.

For $\eta \in \mathrm{Aut}\,(G/H, \mu)$ and for $X \in m$ sufficiently near 0 in $m$ we can write

$$\eta(\pi \exp X) = \pi \exp(\varphi(X)) \tag{3.1}$$

where $\varphi: m \rightarrow m$ is analytic at $0 \in m$ and $\varphi(0) = 0$. Thus for $X, Y$ sufficiently near 0 in $m$ we can also write

$$\eta\mu(\pi \exp X, \pi \exp Y) = \eta(\pi \exp F(X, Y)) = \pi \exp(\varphi F(X, Y))$$

and

$$\begin{aligned} \mu(\eta(\pi \exp X), \eta(\pi \exp Y)) &= \mu(\pi \exp(\varphi X), \pi \exp(\varphi Y)) \\ &= \pi \exp F(\varphi X, \varphi Y)\ . \end{aligned}$$

But since $\eta \in \mathrm{Aut}\,(G/H, \mu)$ we can conclude for $X, Y$ sufficiently near 0 in $m$

$$\varphi F(X, Y) = F(\varphi X, \varphi Y)$$

that is, $\varphi$ is an automorphism of a suitable local multiplicative system $(U, F)$.

We shall now expand $\varphi$ and $F$ in their Taylor's series to find conditions on $\eta \in \mathrm{Aut}\,(G/H, \mu)$ so that the differential $(T\eta)(\bar{e})$ is in $\mathrm{Aut}\,(m, \alpha)$. First we note from (3.1) and the chain rule we have for $X \in m$

$$\begin{aligned} T\eta(\bar{e})(X) &= [T(\pi \circ \exp \circ \varphi)(0)](X) \\ &= [T\pi(e) \circ T\exp(0) \circ T\varphi(0)](X) \\ &= T\varphi(0)(X) \\ &= \varphi^1(0)(X) \end{aligned} \tag{$**$}$$

because $T\exp(0)$ is the identity on $g$ and $T\pi(e)$ is the identity on $m$. From the Taylor's series

$$F(X, Y) = PX + QY + \frac{F^2(\theta)}{2}(X, Y)^2 + \cdots$$

and

$$\varphi(X) = \varphi^1(0)X + \frac{\varphi^2(0)}{2}X^2 + \cdots$$

we have for $X, Y$ sufficiently near 0 in $m$

$$\begin{aligned}
\varphi(F(X, Y)) &= \varphi^1(0)F(X, Y) + 1/2\varphi^2(0)(F(X, Y), F(X, Y)) + \cdots \\
&= \varphi^1(0)(PX + QY + 1/2F^2(\theta)(X, Y)^2 + \cdots) \\
&\quad + 1/2\varphi^2(0)(PX + QY + \cdots, PX + QY + \cdots) + \cdots \\
&= \varphi^1(0)PX + \varphi^1(0)QY \\
&\quad + 1/2\varphi^1(0)F^2(\theta)(X, Y)^2 + 1/2\varphi^2(0)(PX + QY, PX + QY) \\
&\quad + \varepsilon_3
\end{aligned}$$

where $\varepsilon_3$ is of order three. Also

$$\begin{aligned}
F(\varphi X, \varphi Y) &= P\varphi(X) + Q\varphi(Y) + 1/2F^2(\theta)(\varphi X, \varphi Y)^2 + \cdots \\
&= P(\varphi^1(0)X + 1/2\varphi^2(0)X^2 + \cdots) + Q(\varphi^1(0)Y + 1/2\varphi^2(0)Y^2 \\
&\quad + \cdots) \\
&\quad + 1/2F^2(\theta)(\varphi^1(0)X + \cdots, \varphi^1(0)Y + \cdots)^2 + \cdots \\
&= P\varphi^1(0)X + Q\varphi^1(0)Y + 1/2P\varphi^2(0)X^2 + 1/2Q\varphi^2(0)Y^2 \\
&\quad + 1/2F^2(\theta)(\varphi^1(0)X, \varphi^1(0)Y)^2 + \varepsilon_3 .
\end{aligned}$$

Since $\varphi F(X, Y) = F(\varphi X, \varphi Y)$ we compare terms of the same degree to obtain

$$[\varphi^1(0), P] = [\varphi^1(0), Q] = 0$$

and

$$\begin{aligned}
1/2\varphi^1(0)F^2(\theta)(X, Y)^2 + 1/2\varphi^2(0)(PX + QY, PX + QY) \\
= 1/2P\varphi^2(0)X^2 + 1/2Q\varphi^2 Y^2 \\
+ 1/2F^2(\theta)(\varphi^1(0)X, \varphi^1(0)Y)^2 .
\end{aligned}$$

From (2.1) and this last equation we obtain by considering the expressions in both $X$ and $Y$ (i.e. replacing $X$ by $sX$ and $Y$ by $tY$):

$$\begin{aligned}
\varphi^1(0)F^2(\theta)[(X, 0), (0, Y)] - F^2(\theta)[(\varphi^1(0)X, 0), (0, \varphi^1(0)Y)] \\
= - \varphi^2(0)(PX, QY) .
\end{aligned}$$

Recalling $\alpha(X, Y) = F^2(\theta)[(X, 0)(0, Y)]$ and equation (**) and by definition $T\eta(\bar{e})$ is nonsingular we obtain the following.

LEMMA 3.1. *Let $(G/H, \mu)$ be a multiplicative system given locally by $\mu(\pi \exp X, \pi \exp Y) = \pi \exp F(X, Y)$ where $F(X, Y) = PX + QY + 1/2F^2(\theta)(X, Y)^2 + \cdots$ and let $\eta \in \mathrm{Aut}(G/H, \mu)$ given locally by $\eta(\pi \exp X) = \pi \exp(\varphi(X))$. Then*

(1) $[P, T\eta(\bar{e})] = [Q, T\eta(\bar{e})] = 0$

(2) $T\eta(\bar{e}) \in \mathrm{Aut}(m, \alpha)$ *if and only if* $\varphi^2(0)(PX, QY) = 0$ *for all* $X, Y \in m$.

Now for $u \in H$ and for $x \in G$ with $\pi x = xH \in G/H$ we have

$$\tau(u)\pi(x) = uxH = uxu^{-1}(uH) = \pi\sigma(u)(x)$$

where $\sigma(u)\colon G \to G\colon x \to uxu^{-1}$ is the inner automorphism of the group $G$ defined by $u$. Next recall [1] that $[T\sigma(u)](e) = \operatorname{Ad} u$ and for an automorphism $\sigma$ of $G$ that $\sigma(\exp X) = \exp(T\sigma(e)X)$. So we assume for $u \in H$ that $\eta = \tau(u) \in \operatorname{Aut}(G/H, \mu)$. Then the local representation gives

$$\begin{aligned}\pi \exp(\varphi(X)) &= \eta(\pi \exp X)\\ &= (\tau(u)\circ\pi)(\exp X)\\ &= \pi\circ\sigma(u)(\exp X)\\ &= \exp \pi\,([T\,\sigma(u)(e)](X))\\ &= \pi \exp(\operatorname{Ad} u(X))\ .\end{aligned}$$

Since $(G, H)$ is a reductive pair we have $\operatorname{Ad} H(m) \subset m$ and consequently for all $X$ in a suitable neighborhood of 0 in $m$ we have

$$\varphi(X) = \operatorname{Ad} u(X)\ .$$

Thus since $\varphi = \operatorname{Ad} u$ is linear we have from this equation, and (**) applied to $\eta = \tau(u)$ that

$$T\tau(u)(\bar{e}) = T\varphi(0) = \operatorname{Ad} u\ . \tag{3.2}$$

Since $\varphi = \operatorname{Ad} u$ is linear, its second derivative is zero; that is, $\varphi^2(0) = 0$. This and Lemma 3.1 yield the following.

PROPOSITION 3.2. *Let $(G/H, \mu)$ be a multiplicative system so that $\tau(H) \subset \operatorname{Aut}(G/H, \mu)$. Let $\mu$ be given locally by $\mu(\pi \exp X, \pi \exp Y) = \pi \exp F(X, Y)$ where $F(X, Y) = PX + QY + 1/2F^2(\theta)(X, Y)^2 + \cdots$ and let $(m, \alpha)$ be the algebra determined by $F^2(\theta)$. Then*

(1) *$[P, \operatorname{Ad} u] = [Q, \operatorname{Ad} u] = 0$ all $u \in H$.*

(2) *$\operatorname{Ad} H \subset \operatorname{Aut}(m, \alpha)$.*

(3) *The algebra $(m, \alpha)$ defines a $G$-invariant affine connection on $G/H$.*

4. **Invariant multiplications.** Let $(G/H, \mu)$ be a multiplicative system defined on the reductive space $G/H$ and let $g = m + h$ be the corresponding fixed decomposition. For $X \in m$ and for $T(G/H)$ the tangent bundle of $G/H$ define functions

$$\ell(\mu, X)\colon G/H \longrightarrow T(G/H)\colon \bar{a} \longrightarrow [(T\mu)(\bar{a}, \bar{e})](0, X)$$

and

$$r(\mu, X)\colon G/H \longrightarrow T(G/H)\colon \bar{a} \longrightarrow [(T\mu)(\bar{e}, \bar{a})](X, 0)$$

where $T\mu$ is the differential of the function $\mu: G/H \times G/H \rightarrow G/H$ which is evaluated, for example, at $(\bar{a}, \bar{e})$ and acting on tangent vectors $(0, X) \in T(G/H, \bar{a}) \times T(G/H, \bar{e})$.

Next note $\ell(\mu X)$(resp. $r(\mu, X)$) is a vector field if and only if $\mu(\bar{a}, \bar{e}) = \bar{a}$(resp. $\mu(\bar{e}, \bar{a}) = \bar{a}$). For if $\ell(\mu, X)$ is a vector field and $\mu(\bar{a}, \bar{e}) = \bar{b}$, then $\ell(\mu, X)(\bar{a}) \in T(G/H, \bar{b})$ which is the tangent space of $G/H$ at $\bar{b}$. But $\ell(\mu, X)$ being a vector field means

$$\bar{a} = idy\,(\bar{a}) = p \circ \ell(\mu, X)(\bar{a}) = \bar{b}$$

where $p: T(G/H) \rightarrow G/H$ is the corresponding projection map. Conversely $\mu(\bar{a}, \bar{e}) = \bar{a}$ easily implies $p \circ \ell(\mu, X) = idy$; that is, $\ell(\mu, X)$ is a vector field. Similarly for $r(\mu, X)$.

Also it is not difficult to see that in this case $\ell(\mu, X)$ and $r(\mu, X)$ are analytic vector fields which depend linearly on the parameter $X$. Since the results for $r(\mu, X)$ are similar to those for $\ell(\mu, X)$, we restrict ourselves to $\ell(\mu, X)$.

We now define the concept of an invariant multiplication which reduces to the familiar notion in Lie groups. Recall [4] if $\tilde{X}: M \rightarrow T(M)$ is an analytic vector field on a manifold $M$ where $T(M)$ is the tangent bundle over $M$ and if $f: M \rightarrow M$ is an analytic diffeomorphism, then $\tilde{X}$ is $f$-invariant if

$$Tf(p)(\tilde{X}(p)) = \tilde{X}(f(p))$$

for all $p \in M$ where $Tf(p): T(M, p) \rightarrow T(M, f(p))$ is the differential of $f$ at $p$.

DEFINITION. Let $(G, H)$ be a reductive pair with $g = h + m$ the corresponding decomposition for $g$ and let $G/H$ be the homogeneous space of left cosets. Let $(G/H, \mu)$ be a multiplicative system with $\tau(H) \subset \mathrm{Aut}\,(G/H, \mu)$ and let $\Gamma = \{\tau(\exp A): A \in O_m\}$ where $O_m$ is a neighborhood of 0 in $m$ on which exp is one-to-one and $(\exp O_m) \cap H = \{e\}$; [1, p. 113]. Let $\bar{e} = eH$ be a right identity for $(G/H, \mu)$; that is, $\mu(\bar{a}, \bar{e}) = \bar{a}$, then $\mu$ is called *Γ-invariant* if for all $X \in m$ the vector fields $\ell(\mu, X)$ are invariant relative to the functions in $\Gamma$ as follows:

$$T\tau(\exp A)(\bar{e}) \cdot \ell(\mu, X)(\bar{e}) = \ell(\mu, X)(\tau(\exp A)\bar{e})$$

for all $A$ in $O_m$.

REMARK. (1) Before considering the general case we first consider the system $(G, \mu)$. In this case $\Gamma$ can be replaced by all of $L(G) = \{L(a): a \in G\}$ where $L(a): G \rightarrow G: x \rightarrow ax$, the multiplication in the group $G$. In particular we see that if $e$ is a right identity of $(G, \mu)$, then the vector field $\ell(X)$ is $\Gamma$-invariant if and only if $TL(a)(p) \cdot \ell(\mu, X)(p) = \ell(\mu, X)(L(a)p)$; that is, the $\Gamma$-invariance at $e$ is actually global. Also it should be noted that when $\mu$ is the Lie group multiplication in $G$, then the $L(G)$-invariant vector field $\measuredangle(X)$ equals the usual left $G$-invariant $\tilde{X}$; note remark (3) below.

EXAMPLE 1. Let $f: G \to G$ be an analytic function on the Lie group $G$ so that $f(e) = e$, then the multiplication

$$\mu(x, y) = m(x, f(y)) = xf(y)$$

is $L(G)$-invariant where $m$ is the Lie group multiplication in $G$. First $\mu(x, e) = xf(e) = x$ for all $x \in G$ so that $\measuredangle(\mu, X)$ is a vector field on $G$. From remark (1) we have $m$ is a $L(G)$-invariant multiplication so that

$$[(Tm)(a, e)](0, U) = \measuredangle(m, U)(a) = TL(a)(e)[(Tm)(e, e)\cdot(0, U)] \,.$$

Thus noting $\mu(x, y) = [m\circ(idy \times f)](x, y)$ we have using the chain rule

$$\begin{aligned}
\measuredangle(\mu, X)(a) &= [(T\mu)(a, e)](0, X) \\
&= [T(m\circ(idy \times f))(a, e)](0, X) \\
&= [T\,m(a, e)\circ(T\,idy(a) \times Tf(e))](0, X) \\
&= [T\,m(a, e)](T\,idy(a)\cdot 0,\ Tf(e)\cdot X) \\
&= [T\,m(a, e)](0,\ Tf(e)\cdot X) \\
&= T\,L(a)(e)[T\,m(e, e)\cdot(T\,idy(e)\cdot 0,\ Tf(e)\cdot X)] \\
&= T\,L(a)(e)[T(m\circ(idy \times f))(e, e)\cdot(0, X)] \\
&= T\,L(a)(e)[(T\mu)(e, e)\cdot(0, X)] \\
&= T\,L(a)(e)\measuredangle(\mu, X)(e)
\end{aligned}$$

so that $\mu$ is left $L(G)$-invariant. Other examples can easily be constructed where the multiplication need not have the "separation of variable property". Thus locally an $L(\exp g)$-invariant multiplication $\mu$ can be given by $\mu(\exp X, \exp Y) = \exp F(X, Y)$ where

$$F(X, Y) = C(X, f(X, Y))$$

with $f: g \times g \to g$ analytic at $(0, 0)$, and $f(X, 0) = 0$, and $C(X, Y) = X + Y + 1/2[XY] + \cdots$ is the Campbell-Hausdorff formula. We leave it as an open problem to see if this is the most general way of obtaining the local expression for an $L(G)$-invariant multiplication on $G$.

REMARKS. We shall soon give the local formula for a $\Gamma$-invariant multiplication on $G/H$. But first we give a few remarks and formulas.

(2) If $(G/H, \mu)$ is a Lie group; i.e. $H$ normal, then since $\tau(a)\bar{x} = axH$ we have $\mu(\bar{a}, \bar{x}) = L(\bar{a})\bar{x}$. Thus locally $\mu$ is $L(G/H)$-invariant (as a group) if and only if $\mu$ is $\Gamma$-invariant (as a homogeneous space).

Also when $H$ is a normal subgroup with $\mu(\bar{a}, \bar{b}) = abH$, then $\tau(H) \subset \operatorname{Aut}(G/H, \mu)$ because for $u \in H$ we have

$$\begin{aligned}\tau(u)\mu(\bar{a}, \bar{b}) &= uabH \\ &= (uau^{-1}ubu^{-1})H \\ &= uau^{-1}H \cdot ubu^{-1}H \\ &= \mu(\tau(u)\bar{a}, \tau(u)\bar{b}) .\end{aligned}$$

(3) Let $(G/H, \mu)$ be a multiplicative system with $\mu$ given locally by

$$F(X, Y) = X + QY + \alpha(X, Y) + 1/2F^2(\theta)(0, Y)^2 + \cdots$$

as in §2, where $\theta = (0, 0)$ and $\alpha(X, Y) = F^2(\theta)[(X, 0), (0, Y)]$. Then the vector field $\swarrow(\mu, X)$ satisfies

$$\swarrow(\mu, X)(\bar{e}) = QX .$$

For from $\mu(\pi \circ \exp X, \pi \circ \exp Y) = \pi \circ \exp F(X, Y)$ we obtain $\mu \circ (\pi \circ \exp \times \pi \circ \exp) = \pi \circ \exp \circ F$. Using $T(\pi \circ \exp)(0) = idy$ on $m$ and the chain rule we obtain

$$\begin{aligned}\swarrow(\mu, X)(\bar{e}) &= [(T\mu)(\bar{e}, \bar{e})](0, X) \\ &= T[(\pi \circ \exp \circ F)(0, 0)](0, X) \\ &= T(\pi \circ \exp)(0)(0 + QX) \\ &= QX\end{aligned}$$

recalling $[(TF)(0, 0)](U, V) = U + QV$.

(4) We must restrict ourselves to the *set* $\Gamma = \{\tau(\exp A): A \in O_m\}$ and not the group generated by $\tau(\exp O_m)$ because this group is frequently $G$ since $m$ frequently generates $g$. For example, if $g$ is simple, then $m + [m\, m]$ is an ideal of $g$ and therefore equals $g$.

However if for the system $(G/H, \mu)$ we have $\mu$ is $\tau(G)$-invariant, then it is $\tau(H)$-invariant. This with $\tau(H) \subset \operatorname{Aut}(G/H, \mu)$ yield the following computations which indicate that $\tau(G)$-invariance is too strong of a condition. With $\mu$ given by $F(X, Y) = X + QY + \cdots$ as in remark (3) we see from Proposition 3.2 that $\tau(H) \subset (G/H, \mu)$ implies $[\operatorname{Ad} u, Q] = 0$ for all $u \in H$. From $\mu$ being $\tau(H)$ invariant we have

$$T\tau(u)(\bar{e}) \cdot \swarrow(\mu, X)(\bar{e}) = \swarrow(\mu, X)(\tau(u)\bar{e}) = \swarrow(\mu, X)(\bar{e}) .$$

But from formula (3.2) we have $T\tau(u)(\bar{e}) = \operatorname{Ad} u$ and from remark (3) we have $\swarrow(\mu, X)(\bar{e}) = QX$; thus

$$(\operatorname{Ad} u)(QX) = QX .$$

This gives, since $[\operatorname{Ad} u, Q] = 0$ that $Q(\operatorname{Ad} u - I)X = 0$ for all $u \in H$ and $X \in m$. Thus for $u = \exp U$ with $U \in h$ and using $\operatorname{Ad}(\exp U) = e^{\operatorname{ad} U}$

we obtain

$$0 = (Q \circ \operatorname{ad} U)(X) = (\operatorname{ad} U \circ Q)(X) .$$

If $Q$ is nonsingular we obtain $(\operatorname{ad} h)(m) = 0$ so that $h$ is an ideal in $g$; this is usually not the case.

We now obtain sufficient information concerning the multiplication $\mu$ from the Taylor's series for $F(X, Y)$; note the converse statement in Proposition 4.7.

THEOREM 4.6. *Let $(G, H)$ be a reductive pair with $(G/H, \mu)$ a multiplicative system with $\tau(H) \subset \operatorname{Aut}(G/H, \mu)$. Let $\mu$ be $\Gamma$-invariant and for $X, Y \in m$ in a suitable neighborhood of* 0 *in $m$ let $\mu$ be given locally by $\mu(\pi \exp X, \pi \exp Y) = \pi \exp F(X, Y)$ where $F$ is given by the Taylor's series*

$$F(X, Y) = X + QY + \alpha(X, Y) + 1/2F^2(\theta)(0, Y)^2 + \sum_{n=3}^{\infty} 1/n!\, F^n(\theta)(X, Y)^n .$$

*Then*

(1) $\alpha(X, Y) = 1/2X \cdot QY (\equiv 1/2[X\, QY]_m)$

(2) If $\pi' = T\pi(e)$ *and if*

$$F_n \equiv F_n(U, V) = F^n(\theta)[(U, 0), \cdots, (U, 0), (0, V)]$$

*where $(U, 0)$ occurs $(n-1)$-times and $n > 2$, then for $F_1 \equiv F_1(X, Y) = QY$ we have*

$$0 = \pi'[\mu(n, 0)F_n + \mu(n, 1)\operatorname{ad} X(F_{n-1}) + \cdots + \mu(n, k)(\operatorname{ad} X)^k(F_{n-k}) + \cdots + \mu(n, n-1)(\operatorname{ad} X)^{n-1}F_1]$$

*where $\mu(n, k) = (-1)^k/(k+1)!\,(n-k-1)!$. Thus we have an iterative formula for part of the Taylor's series for $F$ which is the best possible obtainable from the $\Gamma$-invariance condition.*

(3) *For each $u \in H$ and $n = 2, 3, \cdots$*

$$(\operatorname{Ad} u) \cdot F^n(\theta)(X, Y)^n = F^n(\theta)(\operatorname{Ad} u \cdot X, \operatorname{Ad} u \cdot Y)^n .$$

*In particular* $(\operatorname{Ad} u) \cdot F_n(X, Y) = F_n(\operatorname{Ad} u \cdot X, \operatorname{Ad} u, Y)$.

*Proof.* We have $\tau(a) \circ \pi = \pi \circ L(a)$ and from remark (3)

$$[\mu \circ (\pi \times \pi) \circ (\exp \times \exp)](X, Y) = [\pi \circ \exp \circ F](X, Y) .$$

Using this equation and the chain rule we obtain for $\bar{a} = \pi \exp X$ and

$\bar{e} = \pi \exp 0$ that

$$[(T\mu)(\bar{a}, \bar{e})](0, Y) = [T(\pi \circ \exp \circ F)(X, 0)](0, Y)$$

which is used in the fifth equality below.

From [1, p. 95] we have for $A$ in a suitable neighborhood of 0 in $g$ that

$$T[L(\exp(-A)) \circ \exp](A) = \frac{I - e^{-\operatorname{ad} A}}{\operatorname{ad} A}$$

where $(I - e^{-P})/P = \sum_{k=0}^{\infty}(-P)^k/(k+1)!$. Also from $\tau(a^{-1}) \circ \tau(a) = idy$ on $G/H$ we obtain

$$T\tau(a^{-1})(\bar{a}) \circ T\tau(a)(\bar{e}) = I$$

which gives the inverse for $T\tau(a)(\bar{e})$.

We now use the above formulas and the chain rule to obtain for $X$, $Y$ in a suitable neighborhood of 0 in $m$ and $a = \exp X$,

$$\begin{aligned}
QY &= \not\iota(\mu, Y)(\bar{e}) \\
&= [T\tau(a)(\bar{e})]^{-1} \not\iota(\mu, Y)(\bar{a}) \\
&= T\tau(a^{-1})(\bar{a}) \cdot [(T\mu)(\bar{a}, \bar{e})](0, Y) \\
&= [T(\tau(a^{-1}) \circ \mu)(\bar{a}, \bar{e})](0, Y) \\
&= [T(\tau(a^{-1}) \circ \pi \circ \exp \circ F)(X, 0)](0, Y) \\
&= T\pi(e) \cdot T(L(\exp(-X)) \circ \exp)(X) \cdot [(TF)(X, 0)](0, Y) \\
&= T\pi(e) \cdot \frac{I - e^{-\operatorname{ad} X}}{\operatorname{ad} X} \cdot [(TF)(X, 0)](0, Y)
\end{aligned}$$

where

$$\begin{aligned}
F' &\equiv [(TF)(X, 0)](0Y) \\
&= \lim_{t \to 0} \frac{1}{t}[F(X, tY) - F(X, 0)] \\
&= \lim_{t \to 0} \frac{1}{t}[X + tQY + t\alpha(X, Y) + t^2/2F^2(\theta)(0, Y)^2 + \cdots - X] \\
&= QY + \alpha(X, Y) + \sum_{n=3}^{\infty} \frac{1}{(n-1)!} F_n(X, Y) .
\end{aligned}$$

To see this last equality just note that by induction we have

$$\begin{aligned}
F^n(\theta)(X, tY)^n &= F^n(\theta)(X, 0)^n + ntF_n(X, Y) + o(t^2) \\
&= ntF_n(X, Y) + o(t^2)
\end{aligned}$$

using $F^n(\theta)(X, 0)^n = 0$ since $(G/H, \mu)$ has $\bar{e}$ as a right identity.

From the series for $(I - e^{-\operatorname{ad} X})/\operatorname{ad} X$ we obtain

$$\begin{aligned}
QY &= \pi' \sum_{k=0}^{\infty}\left(\frac{(-\operatorname{ad} X)^k}{(k+1)!}\right)(F') \\
&= \pi'\left[IF' - \frac{\operatorname{ad} X}{2!}F' + \cdots + \frac{(-\operatorname{ad} X)^k}{(k+1)!}F' + \cdots\right] \\
&= \pi'\Big[QY + \alpha(X, Y) + \sum_{n=3}^{\infty} F_n/(n-1)! \\
&\quad - 1/2(\operatorname{ad} X)\Big(QY + \alpha(X, Y) + \sum_{n=3}^{\infty} F_n/(n-1)!\Big) \\
&\quad + \cdots + \frac{(-\operatorname{ad} X)^k}{(k+1)!}\Big(QY + \alpha(X, Y) + \sum_{n=3}^{\infty} F_n/(n-1)!\Big) + \cdots\Big] \\
&= \pi'[QY + \alpha(X, Y) - 1/2(\operatorname{ad} X)(QY) + \cdots] \\
&= \pi' QY + \pi'\alpha(X, Y) - \frac{\pi'}{2}(\operatorname{ad} X)(QY) + \cdots .
\end{aligned}$$

Thus since $QY$ and $\alpha(X, Y)$ are in $m$ and $\pi' \mid m$ is the identity, we obtain

$$\begin{aligned}
\alpha(X, Y) &= \frac{\pi'}{2}(\operatorname{ad} X)(QY) \\
&= \frac{\pi'}{2}(X \cdot QY + [XQY]_h) \\
&= \frac{1}{2} X \cdot QY .
\end{aligned}$$

(Recall $[UV]_h$ is the component of $[UV]$ which is in $h$). Similarly by noting $F_k(X, Y)$ is homogeneous in $X$ of degree $k-1$ we combine those terms of degree $n-1$ in $X$ to obtain

$$\begin{aligned}
0 = \pi'\Big[&\frac{F_n}{(n-1)!} - \frac{1}{2}(\operatorname{ad} X)\frac{F_{n-1}}{(n-2!)} + \cdots + \frac{(-1)^k(\operatorname{ad} X)^k F_{n-k}}{(k+1)!(n-k-1)!} \\
&+ \cdots + \frac{(-1)^{n-1}}{n!}(\operatorname{ad} X)^{n-1} F_1\Big]
\end{aligned}$$

where $F_1 = QY$.

Equation (3) in the theorem follows from Proposition 3.2 and the remarks preceding it. Thus for $\eta \in \operatorname{Aut}(G/H, \mu)$ we wrote locally $\eta(\pi \exp X) = \pi \exp(\phi(X))$ and showed $\phi F(X, Y) = F(\phi X, \phi Y)$. In particular for $\eta = \tau(u)$ we showed $\phi = \operatorname{Ad} u$ for $u \in H$ so that from the Taylor's series for $F$ and the linearity of $\phi = \operatorname{Ad} u$ we obtain (3).

EXAMPLE 2. These formulas can also be used to construct examples locally. Thus let $G$ be nilpotent so that $\operatorname{ad} X$ is nilpotent for all $X \in g$; that is, there exists $n$ so that for all $X \in g$, $(\operatorname{ad} X)^n = 0$. Let the function $F_k$ be given by the iteration formulas: $F_0 =$

$X$, $F_1(X, Y) = QY$, $0 = \pi'[\mu(2, 0)F_2 + \mu(2, 1)\text{ad}\, XF_1]$, etc. Set $K(X, Y) = \sum_{k=0}^{\infty} F_k(X, Y)/k!$ which is a finite sum by nilpotency of $G$. Note that $K$ is analytic and for the $k$-th derivative $K^k(\theta)(X, Y)^k = F_k(X, Y) \equiv K_k(X, Y)$. Thus $K$ has the above Taylor's series and we can define locally $\mu(\pi \exp X, \pi \exp Y) = \pi \exp K(X, Y)$. This is (locally) $\Gamma$-invariant because $K$, in terms of its Taylor's series, satisfies the iteration equations of the theorem and the process of the proof is reversable.

Next by induction using the iteration equation we also obtain $\phi F_k(X, Y) = F_k(\phi X, \phi Y)$ for $\phi = \text{Ad}\, u$ with $u \in H$. This uses Proposition 3.2 as follows: $\phi F_1(X, Y) = \text{Ad}\, u\, QY = Q(\text{Ad}\, u Y) = F_1(X, \phi Y) = F(\phi X, \phi Y)$ since "$X$" does not occur in the formula for $F_1$. Also reductivity of the pair $(G, H)$ is used to commute $\pi'$ and $\text{Ad}\, u$: for $Y \in m$, $(\pi' \circ \text{Ad}\, u)(Y) = \text{Ad}\, u(Y) = (\text{Ad}\, u \circ \pi')(Y)$. Thus since $K(X, Y) = \sum F_k(X, Y)/k!$ we have $\phi K(X, Y) = K(\phi X, \phi Y)$ so that using the results of §3 and the definition of $\phi$ by $\tau(u)(\pi \exp X) = \pi \exp(\phi X)$ we have locally $\tau(H) \subset \text{Aut}(G/H, \mu)$ as follows:

$$\begin{aligned} \tau(u)\mu(\pi \exp X, \pi \exp Y) &= \tau(u)(\pi \exp K(X, Y)) \\ &= \pi \exp(\phi K(X, Y)) \\ &= \pi \exp K(\phi X, \phi Y) \end{aligned}$$

and

$$\begin{aligned} \mu(\tau(u)(\pi \exp X), \tau(u)(\pi \exp Y)) &= \mu(\pi \exp(\phi Y), \pi \exp(\phi Y)) \\ &= \pi \exp K(\phi X, \phi Y)\,. \end{aligned}$$

We extend the above notions in the following result to obtain a converse to Theorem 4.6.

PROPOSITION 4.7. *Let $(G, H)$ be a reductive pair with fixed decomposition $g = m + h$. Let $F_k$: $m \times m \to m$ for $k = 1, 2, \cdots$ be a sequence of multilinear functions which satisfy the iterative equation* (2) *and equation* (3) *of Theorem* 4.6; *that is, for all $u \in H$, $(\text{Ad}\, u) \cdot F_k(X, Y) = F_k(\text{Ad}\, u \cdot X, \text{Ad}\, u \cdot Y)$. Then for all $X$, $Y$ in a suitable neighborhood of* 0 *in $m$ the series $X + \sum_{k=1}^{\infty} 1/k!\, F_k(X, Y)$ converges absolutely and uniformly to a function $K(X, Y)$ which is analytic at $\theta = (0, 0) \in m \times m$ and the multiplication $\mu(\pi \exp X, \pi \exp Y) = \pi \exp K(X, Y)$ defines a local multiplicative system $(G/H, \mu)$ so that $\mu$ is locally $\Gamma$-invariant.*

*Proof*. Using the obvious extension of the results in example (2) above, it suffices to prove the series converges to an analytic function $K$ so that the derivatives $K^k(\theta)(X, Y)^k = F_k(X, Y)$. To show that the series converges absolutely and uniformally for $X$, $Y$ in a suitable neighborhood of 0 in $m$, we let $B_k = F_{k+1}/k!$ for $k = 1, 2, \cdots$ and let $S = \pi'$ and $T = \text{ad}\, X$. Then from the iteration formula we obtain $F_1 = SQY$ and

$$B_1 = \frac{ST}{2!}F_1$$

$$B_2 = \frac{ST}{2!}B_1 - \frac{ST^2}{3!}F_1$$

$$B_3 = \frac{ST}{2!}B_2 - \frac{ST^2}{3!}B_1 + \frac{ST^3}{4!}F_1$$

$$\vdots$$

$$B_n = \frac{ST}{2!}B_{n-1} - \frac{ST^2}{3!}B_{n-2} + \cdots + \frac{(-1)^{n+1}}{(n+1)!}ST^nF_1 .$$

Now let $\|\ \ \|$ denote either the operator or the Euclidean norm, then we have $\|S\| = 1$. Let $r = 1$ and let $c$ be a fixed number with $c > 5\,e$ where $e = 2.71 \cdots$. Then for $F_1 = SQY$ with $X \in \mathscr{A} \equiv \{X \in m\colon \|\mathrm{ad}\, X\| < r\}$ and $Y \in \mathscr{B} \equiv \{Y \in m\colon \|QY\| < 1\}$ we have $\|T\| = \|\mathrm{ad}\, X\| < r = 1$ and $\|B_1\| = 1/2\|STF_1\| \leqq 1/2\|S\|\cdot\|T\|\cdot\|F_1\| < 1/2 < c$. Assume for all $k < n$ that $\sum_{i=1}^{k}\|B_i\| < c$, then for $k = n$ we have from the equations for the $B$'s

$$\begin{aligned}
\sum_{i=1}^{n}\|B_i\| = {} & \left\|\frac{ST}{2!}F_1\right\| \\
& + \left\|\frac{ST}{2!}B_1 - \frac{ST^2}{3!}F_1\right\| \\
& + \left\|\frac{ST}{2!}B_2 - \frac{ST^2}{3!}B_1 + \frac{ST^3}{4!}F_1\right\| + \cdots \\
& + \left\|\frac{ST}{2!}B_{n-1} - \frac{ST^2}{3!}B_{n-2} + \cdots + \frac{(-1)^{n+1}}{(n+1)!}ST^nF_1\right\| \\
\leqq {} & \|S\|\left(\frac{\|T\|}{2!} + \frac{\|T^2\|}{3!} + \cdots + \frac{\|T^n\|}{(n+1)!}\right)\|F_1\| \\
& + \frac{\|ST\|}{2!}\left(\sum_{i=1}^{n-1}\|B_i\|\right) + \frac{\|ST^2\|}{3!}\left(\sum_{i=1}^{n-2}\|B_i\|\right) \\
& + \cdots + \frac{\|ST^{n-1}\|}{n!}\|B_1\| \\
\leqq {} & \left(\frac{r}{2!} + \frac{r^2}{3!} + \cdots + \frac{r^n}{(n+1)!}\right) \\
& + \frac{r}{2!}c + \frac{r^2}{3!}c + \cdots + \frac{r^{n-1}}{n!}c
\end{aligned}$$

using the induction hypothesis and $\|T\| < r$. But if $d(k, r) = \sum_{i=1}^{k} r^i/(i+1)!$, then

$$rd(k, r) + 1 + r = 1 + r + \frac{r^2}{2!} + \cdots + \frac{r^{k+1}}{(k+1)!} < e^r = e$$

since $r = 1$. Thus $d(k, r) < (2.8 - 2)/1 = .8$. Therefore $\sum_{i=1}^{n} \|B_i\| < .8 + .8c = .8(1+c) < c$ because $c > 5e$. Thus the series $\sum B_k(X, Y)$ converges absolutely for all $(X, Y)$ with $X \in \mathscr{A}$ and $Y \in \mathscr{B}$ to a function $\bar{k}(X, Y)$. Since $\|B_k\|/k + 1 < \|B_k\|$ we see that $\sum B_k/(k+1)$ converges absolutely to a function $k(X, Y)$.

This series converges uniformly on $\mathscr{A} \times \mathscr{B}$ to $k(X, Y)$ as follows. On $\mathscr{A} \times \mathscr{B}$ the partial sums $\sum_{k=1}^{n} B_k$ are bounded by $c > 5e$. Thus since $1/k + 1 \to 0$ we have from a standard result that the series $\sum B_k/k + 1$ converges uniformly on $\mathscr{A} \times \mathscr{B}$. But $B_k/k + 1 = F_{k+1}/(k+1)!$ so that $X + \sum F_k(X, Y)/k!$ converges uniformly on $\mathscr{A} \times \mathscr{B}$ to $K(X, Y)$. Using this we see from [2, §3] that $K$ is analytic at $\theta = (0, 0)$ and $K^k(\theta)(X, Y)^k = F_k(X, Y)$ as desired.

5. **Connections and holonomy.** From [6] there is a bijective correspondence between $G$-invariant connections on the reductive space $G/H$ and non-associative algebras $(m, \alpha)$ with $\mathrm{Ad}\, H \subset \mathrm{Aut}\,(m, \alpha)$. Thus if this algebra $(m, \alpha)$ is induced by a multiplicative system $(G/H, \mu)$ we obtain a connection "induced by $\mu$" and we discuss such connections and the corresponding holonomy algebra (Lie algebra of the holonomy group). Thus for $X, Y, Z \in m$ let

$$a(X)\colon m \to m\colon Y \to \alpha(X, Y) \quad \text{and} \quad R(X, Y)\colon m \to m\colon Z \to R(Y, Y)Z$$

where $R(X, Y)Z = \alpha(X, \alpha(Y, Z)) - \alpha(Y, \alpha(X, Z)) - \alpha(XY, Z) - [h(X, Y)Z]$ is the curvature evaluated at $\bar{e} = eH$ in $G/H$ [6]; recall that $XY = [XY]_m$ (resp. $h(X, Y) = [XY]_h$) is the projection of $[XY]$ in $g$ into $m$(resp. $h$). From [7] the holonomy algebra, denoted by $\mathrm{hol}\,(\alpha)$, is the smallest Lie algebra $\mathrm{hol}\,(\alpha)$ of endomorphisms of $m$ so that $R(X, Y) \in \mathrm{hol}\,(\alpha)$ and $[a(X), \mathrm{hol}\,(\alpha)] \subset \mathrm{hol}\,(\alpha)$ for all $X, Y \in m$.

We shall say that the holonomy group acts irreducibly on $G/H$ in case $\mathrm{hol}\,(\alpha)$ act irreducibly on $m$. This can be stated in terms of the algebra $(m, \alpha)$ as follows. A left ideal of the algebra $(m, \alpha)$ is a subspace $n$ of $m$ such that $\alpha(m, n) \subset n$. Thus from the formula for $R(X, Y)Z$ we see that a left ideal $n$ which is invariant under $\mathrm{ad}\, h(m, m) = \{\mathrm{ad}\, h(X, Y)\colon X, Y \in m\}$ is $\mathrm{hol}\,(\alpha)$-invariant. Therefore the holonomy irreducibility of $G/H$ implies $(m, \alpha)$ has no left ideals which are $\mathrm{ad}\, h(m, m)$-invariant.

We now consider the connection of the first kind which is a well behaved, easy to construct connection. From [6] we see that on the reductive space $G/H$ there exists a unique $G$-invariant connection which has zero torsion tensor and such that a 1-parameter subgroup $x(t) = \exp tX$ of $G$ generated by $X \in m$ projects by $\pi\colon G \to G/H\colon x(t) \to$

$\bar{x}(t)$ into a geodesic $\bar{x}(t)$ in $G/H$. In this case $\alpha(X, Y) = 1/2XY$ and the connection is called the connection of the first kind relative to a fixed decomposition $g = m + h$.

Thus since this multiplication $\alpha(X, Y) = 1/2XY$ is anti-commutative, a left ideal is a two sided ideal; therefore holonomy irreducibility implies $(m, \alpha)$ has no ideals invariant under ad $h(m, m)$. But using the Jacobi identity, ad $h(m, m)$ is contained in the derivation algebra of $(m, \alpha)$ so that the algebra $(m, \alpha)$ must contain no proper ideals or $mm = 0$; that is, the holonomy irreducibility implies $(m, \alpha)$ is the zero algebra or simple. This uses a result in [8] which states if a finite dimensional nonassociative algebra over $R$ which is not the zero algebra has a proper ideal, then it has a proper ideal invariant under its derivation algebra.

If the connection on $G/H$ induced by $(m, \alpha)$ is pseudo-Riemannian, then from [6] there exists a nondegenerate symmetric bilinear form $C: m \times m \rightarrow R$ satisfying

$$C(\operatorname{ad} U \cdot X, Y) + C(X, \operatorname{ad} U \cdot Y) = 0$$

and

$$C(a(Z) \cdot X, Y) + C(X, a(Z) \cdot Y) = 0$$

for all $X, Y, Z$ in $m$ and $U$ in $h$; that is, the endomorphisms ad $U$ and $a(Z)$ are $C$-skew symmetric. Also for this connection we have [6], $0 = \operatorname{Tor}(X, Y) = \alpha(X, Y) - \alpha(X, Y) - XY$ and the multiplication function $\alpha$ is determined by

$$2C(Z, \alpha(X, Y)) = C(Z, XY) + C(ZX, Y) + C(X, ZY) . \tag{5.1}$$

We shall denote the algebra $m$ with multiplication $\alpha(X, Y) = 1/2XY$ by $(m, 1/2XY)$ and we shall denote the algebra $(m, \alpha)$ with a nondegenerate from $C$ inducing a pseudo-Riemannian connection (i.e. satisfying the above equations) by $(m, \alpha, C)$. In particular, if $C$ is positive definite so that it induces a Riemannian connection, then from the deRham decomposition [4] the original connection is built up from its irreducible components.

We next use the algebra $(m, \alpha)$ obtained from a multiplication to obtain a connection. Thus let $(G/H, \mu)$ be a multiplicative system as before and let

$$\mu(\pi \exp X, \pi \exp Y) = \pi \exp F(X, Y)$$

where we have

$$F(X, Y) = PX + QY + \alpha(X, Y) + \cdots$$

with $\alpha(X, Y) = F^2(\theta)[(X, 0), (0, Y)]$ a bilinear multiplication on $m$ so that $\operatorname{Ad} H \subset \operatorname{Aut}(m, \alpha)$. For a $\Gamma$-invariant multiplication we obtained in §4 (using the notation $XY = X\cdot Y$)

$$\alpha(X, Y) = 1/2X\cdot QY$$

thus if $L(X)\colon m \to m\colon Y \to XY$ we have for all $X \in m$ and $U \in h$ that

$$a(X) = 1/2L(X)\circ Q \text{ and } [\operatorname{ad} U, Q] = 0$$

using the results of §3.

LEMMA 5.1. *Let $(G/H, \mu)$ be a multiplicative system as above and let $\mu$ be $\Gamma$-invariant. Then the kernel of $Q$ is an* ad *$h$-invariant left ideal of $(m, \alpha)$.*

*Proof.* Let $n = \ker Q$, then since $[\operatorname{ad} h, Q] = 0$ we see that $n$ is ad $h$-invariant. Also $\alpha(m, n) = 1/2m\cdot Qn = 0$ so that $\alpha(m, n) \subset n$; that is, $n$ is a left ideal.

LEMMA 5.2. *Let $(G/H, \mu)$ be a multiplicative system as before which induces a nonzero algebra $(m, \alpha)$ and a corresponding connection on $G/H$. Let $\mu$ be $\Gamma$-invariant and let* hol $(\alpha)$ *be irreducible, then $Q$ is nonsingular.*

*Proof.* Suppose $\mu$ is $\Gamma$-invariant and hol $(\alpha)$ is irreducible. Then from the remarks at the beginning of this section, the algebra $(m, \alpha)$ has no left ideals which are ad $h(m, m)$-invariant. But from Lemma 5.1, the kernel of $Q$ is such an ideal. Thus the kernel of $Q$ is zero since we are assuming $\alpha(X, Y) = 1/2X\cdot QY$ is not identically zero.

We use these lemmas in the next two results where we compare an irreducible connection induced by a multiplication with the irreducible connection of the first kind.

THEOREM 5.3. *Let $(G, H)$ be a reductive pair so that for the decomposition $g = m + h$ we have $[m, m]_m \neq 0$. Let $(G/H, \mu)$ be a multiplicative system as before so that $\mu$ is $\Gamma$-invariant and let the connection induced by $\mu$ via the algebra $(m, \alpha)$ be a holonomy irreducible pseudo-Riemannian connection. If the algebra $(m, 1/2XY)$ is simple, then $\alpha(X, Y) = 1/2XY$; thus the connection by $\mu$ is of the first kind.*

*Proof.* First assume $\mu$ is $\Gamma$-invariant, then from $\alpha(X, Y) = 1/2X\cdot QY$ and

$$\begin{aligned} 0 &= \operatorname{Tor}(X, Y) = \alpha(X, Y) - \alpha(Y, X) - XY \\ &= 1/2X\cdot QY - 1/2\cdot QX - XY \end{aligned}$$

we obtain

$$L(X)\circ Q = L((2I - Q)X) .$$

Thus from $a(X)Y = \alpha(X, Y)$,

$$2a(X) = L(X)\circ Q = L((2I - Q)X) . \tag{5.3}$$

Next by Lemma 5.2 and the hypothesis of irreducibility we see that $Q$ is nonsingular. From this we obtain $2I - Q$ is nonsingular as follows. If $(2I - Q)A = 0$, then using $(5.3) L(A)\circ Q = 0$. But since $Q$ is nonsingular, $L(A) = 0$. Thus if $A \neq 0$, this means that the one dimensional subspace $RA$ is an ideal in the simple algebra $(m, XY)$; consequently $A = 0$. Since $2I - Q$ is nonsingular, $(2I - Q)m = m$ and from formula (5.3) we obtain $a(m) = L(m)$. Next recall that the elements of $a(m)$ are $C$-skew so that the elements of $L(m)$ are also $C$-skew; thus

$$C(ZX, Y) + C(X, ZY) = 0 .$$

But from formula (5.1) which uniquely determines $\alpha$ in terms of $C$ and $(m, XY)$ we obtain

$$2C(Z, \alpha(X, Y)) = C(Z, XY)$$

that is, $\alpha(X, Y) = 1/2XY$. Since $(m, XY)$ is simple, this also implies $Q = I$.

COROLLARY 5.4. *Let the reductive pair $(G, H)$ and the multiplicative system $(G/H, \mu)$ be as in Theorem 5.3. If the corresponding Lie algebra $g$ is simple and $h$ is semi-simple and $g = m + h$ where $m = h^{\perp}$ the orthogonal complement relative to the Killing form, then the connection induced by $\mu$ is of the first kind.*

*Proof.* This uses the result from [8] that if $g = m + h$ as above and $[m\, m]_m \neq 0$, then the algebra $(m, 1/2XY)$ is simple.

## REFERENCES


1. S. Helgason, *Differential Geometry and Symmetric Spaces*, Academic Press, New York, 1962.
2. E. Hille and R. Phillips, *Functional Analysis and Semi-Groups*, Amer. Math. Soc. Providence, R. I., 1957.
3. N. Jacobson, *Lie Algebras*, Interscience, New York, 1962.
4. S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry I and II*, Interscience, New York, 1963.
5. S. Lang, *Analysis* I, Addison-Wesley, Reading, Mass., 1968.
6. K. Nomizu, *Invariant affine connections on homogeneous spaces*, Amer. J. Math., **76** (1954), 33-65.
7. ———, *Recent development in the theory of connections and holonomy groups*, Advances in Mathematics, Academic Press, New York, 1961.
8. A. Sagle and D. Winter, *On homogeneous spaces and reductive subalgebras of simple Lie algebras*, Trans. Amer. Math. Soc., **128** (1967), 142-147.

9. J. Tits, *Liesche Gruppen und Algebren*, Der Universitat Bonn, Bonn, 1965.

Received February 17, 1971 and in revised form June 6, 1972. This research was supported in part by NSF Grants GP-13328 and GP-29224.

UNIVERSITY OF MINNESOTA
AND
UNIVERSITY OF HAWAII, HILO COLLEGE

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# EXISTENCE OF DIRICHLET FINITE BIHARMONIC FUNCTIONS ON THE POINCARÉ 3-BALL

LEO SARIO AND CECILIA WANG

**In an earlier study we discussed the existence of quasiharmonic functions, i.e., solutions of $\Delta u = 1$. We showed, in particular, that there exist Dirichlet finite quasiharmonic functions on the Poincaré 3-ball**

$$B_\alpha: \{|x| < 1, ds = (1 - |x|^2)^\alpha |dx|\}$$

**if and only if $\alpha \in (-3/5, 1)$. We now ask: Is the existence of these functions entailed by that of Dirichlet biharmonic functions? This is known to be the case for dimension 2. We shall show that, perhaps somewhat unexpectedly, it is no longer true for dimension 3.**

For preparation we first solve the problem, of significance in its own right, of the existence of Dirichlet finite biharmonic functions. In the notation of No. 1 below, we give the complete characterization

$$B_\alpha \notin O_{H^2D} \Longleftrightarrow \alpha > -\frac{3}{5}\,.$$

The problem also offers considerable technical interest, as the generating harmonic functions can not be presented in a closed form, but only by means of expansions at the regular singular point of the related differential equation. This makes the estimates somewhat delicate. Also, the four cases $\alpha \geqq 1$, $\alpha \in (-3/5, 1)$, $\alpha < -3/5$, and $\alpha = -3/5$ must be treated separately, each with its own approach.

To deduce the above result (Theorem 1), we first expand a harmonic function on $B_\alpha$ in terms of spherical harmonics with respect to our non-Euclidean metric (Theorem 2). As important applications of Theorem 1 to the classification theory we obtain a decomposition of the totality of Riemannian 3-manifolds into three disjoint nonempty subclasses induced by $O_{QD}$ and $O_{H^2D}$ (Theorem 3), and establish the existence of parabolic 3-manifolds which carry $H^2D$-functions and of hyperbolic 3-manifolds which do not carry $H^2D$-functions (Theorem 4).

An interesting open problem is whether $B_\alpha \notin O_{H^2D}$ if and only if $\alpha > -3/(N+2)$.

1. A function $u$ is harmonic or biharmonic according as it satisfies $\Delta_\lambda u = 0$ or $\Delta_\lambda^2 u = 0$, where $\Delta_\lambda$ is the Laplace-Beltrami operator $\Delta_\lambda = d\delta + \delta d$ with respect to the metric $ds = \lambda(x)|dx|$. Denote by $H^2$ the family of nonharmonic biharmonic functions, by $D$ the family of

functions $f$ with finite Dirichlet integrals $D(f) = \int df \wedge *df < \infty$, and set $H^2D = H^2 \cap D$. Let $O_{H^2D}$ be the class of Riemannian manifolds which do not carry $H^2D$-functions. We assert:

THEOREM 1. $B_\alpha \notin O_{H^2D} \Leftrightarrow \alpha > -3/5$.

The proof will be given in Nos. 2–7.

2. We start by expanding a harmonic function on $B_\alpha$ in spherical harmonics. We recall that a function $S_n(\theta^1, \theta^2)$, in polar coordinates $(r, \theta^1, \theta^2)$, is called a spherical harmonic of degree $n$ if $r^n S_n(\theta^1, \theta^2)$ is harmonic with respect to the Euclidean metric. Every such function is a unique linear combination of $2n+1$ linearly independent fundamental spherical harmonics $S_{nm}$ of degree $n$. The class $\{S_{nm}; n = 0, 1, 2, \cdots; m = 1, 2, \cdots, 2n+1\}$ is not only an orthogonal system with respect to the inner product $(f, g) = \int_\omega fg\, dS$, with $\omega$ the 2-sphere and $dS$ the surface element, but also a complete system with respect to the family of $L^2$-functions. For every harmonic function $h$ in $B_\alpha$, we have a Fourier expansion

$$h(r, \theta) = \sum_{n=0}^{\infty} \sum_{m=1}^{2n+1} d_{nm}(r) S_{nm}(\theta) \tag{1}$$

with $\theta = (\theta^1, \theta^2)$.

By virtue of

$$\Delta_\lambda(f(r)S_{nm}(\theta)) = -\lambda^{-2}[f''(r) + \left(\frac{2}{r} + \frac{\lambda'}{\lambda}\right)f'(r) - n(n+1)r^{-2}f(r)]S_{nm}(\theta) ,$$

and $\lambda'\lambda^{-1} = -2\alpha r(1-r^2)^{-1}$, the function $f(r)S_{nm}(\theta)$ is harmonic on $B_\alpha$ if and only if $f(r)$ satisfies the differential equation

$$\begin{aligned} r^2(1-r^2)f''(r) + r[2(1-r^2) - 2\alpha r^2]f'(r) \\ - n(n+1)(1-r^2)f(r) = 0 . \end{aligned} \tag{2}$$

We shall denote the solution of equation (2) for each $n$ by $f_n(r)$. Since all coefficients in (2) can be expanded into power series of $r$, the point 0 is a regular singular point of the equation. Thus there exists at least one solution of (2) in the form

$$f_n(r) = r^{p_n} \sum_{i=0}^{\infty} c_{ni} r^i , \tag{3}$$

$c_{n0} \neq 0$. On substituting in (2) we have

$$(4)\qquad \sum_{i=0}^{\infty}[(p_n+i-1)(p_n+i)+2(p_n+i)-n(n+1)]c_{ni}r^{p_n+i}$$
$$-\sum_{i=2}^{\infty}\{(p_n+i-3)(p_n+i-2)+(2+2\alpha)(p_n+i-2)$$
$$-n(n+1)\}c_{n,i-2}r^{p_n+i}=0\,.$$

To determine $p_n$ we equate to 0 the coefficient of $r^{p_n}$ and obtain the indicial equation

$$(p_n-1)p_n+2p_n-n(n+1)=0$$

which gives $p_n=n$ or $p_n=-(n+1)$. Since $0\in B_\alpha$, $p_n$ can not be negative, and therefore $p_n=n$.

We then equate to 0 the coefficient $2(n+1)c_{n1}$ of $r^{p_n+1}$ and obtain $c_{n1}=0$.

To find $c_{ni}$, $i\geqq 2$, we equate to 0 the coefficient of $r^{p_n+i}$:

$$[(p_n+i)(p_n+i+1)-n(n+1)]c_{ni}$$
$$=[(p_n+i-2)(p_n+i-1+2\alpha)-n(n+1)]c_{n,i-2}\,.$$

On letting $p_n=n$ and $c_{n0}=1$ we have

$$(5)\qquad c_{n,2i}=\prod_{j=1}^{i}\frac{(n+2j-2)(n+2j-1+2\alpha)-n(n+1)}{(n+2j)(n+2j+1)-n(n+1)}$$

for $i\geqq 1$, and $c_{n,2i+1}=0$ for $i\geqq 0$.

The limit of $f_n(r)=\sum_{i=0}^{\infty}c_{n,2i}r^{n+2i}$ as $r\rightarrow 1$ exists since the $c_{n,2i}$ are of constant sign as soon as $i$ is sufficiently large. Furthermore, this limit can not be zero, for otherwise $\lim_{r\rightarrow 1}f_nS_{nm}\equiv 0$, and consequently $f_n\equiv 0$, contrary to $c_{n0}=1$. In a similar fashion we see that $f_n(r)\neq 0$ for $0<r<1$. Hence for arbitrary but fixed $r_0$, $0<r_0<1$, there exist constants $a_{nm}$ such that $a_{nm}f_n(r_0)S_{nm}=d_{nm}(r_0)S_{nm}$, and

$$(6)\qquad \sum_{n=0}^{\infty}\sum_{m=1}^{2n+1}a_{nm}f_n(r)S_{nm}(\theta)$$

is a series of functions harmonic on $B_\alpha$ which converges absolutely and uniformly to $h(r_0,\theta)$ on the 2-sphere of radius $r_0$. Now let $r_0<r'<1$; then by the same argument there exist constants $a'_{nm}$ such that

$$(7)\qquad \sum_{n=0}^{\infty}\sum_{m=1}^{2n+1}a'_{nm}f_n(r)S_{nm}(\theta)$$

converges to $h$ on the ball of radius $r'$. Hence (6) and (7) are identical on the ball of radius $r_0$, so that $a_{nm}=a'_{nm}$ for all $(n,m)$.

We have proved:

THEOREM. *Every harmonic function $h(r, \theta^1, \theta^2)$ on the Poincaré ball $B_\alpha$ has the expansion in terms of the fundamental spherical harmonics $S_{nm}$,*

$$h(r, \theta^1, \theta^2) = \sum_{n=0}^{\infty} \sum_{m=1}^{2n+1} a_{nm} \sum_{i=0}^{\infty} c_{n,2i} r^{n+2i} S_{nm}(\theta^1, \theta^2) , \tag{8}$$

*where the $c_{n,2i}$ are given by* (5).

3. After this preparation, we proceed with the proof of Theorem 1. An essential aspect of the proof is that the cases $\alpha \geqq 1, \alpha \in (-3/5, 1), \alpha < -3/5$, and $\alpha = -3/5$ all require a different treatment. We first establish the following crucial estimate:

LEMMA 1. *If $\alpha \geqq 1$, then*

$$f_1(r) = \sum_{i=0}^{\infty} c_{1,2i} r^{1+2i} = O((1-r)^{-2\alpha}) \text{ as } r \longrightarrow 1 .$$

*Proof.* By (5),

$$\begin{aligned} c_{1,2i} &= \prod_{j=1}^{i} \frac{(2j-1)(2j+2\alpha)-2}{(2j+1)(2j+2)-2} \\ &= \prod_{j=1}^{i} \frac{2j+2\alpha-1}{2j} \cdot \frac{2j-1+(2j-3)/(2j+2\alpha-1)}{2j+3} . \end{aligned}$$

We claim that

$$c_{1,2i} < \prod_{j=1}^{i} \frac{2j+2\alpha-1}{2j} \tag{9}$$

or, equivalently,

$$4 - \frac{2j-3}{2j+2\alpha-1} > 0 . \tag{10}$$

In the case $\alpha \geqq 1$ under consideration this is clearly so for all $j \geqq 1$. Consequently

$$f_1(r) < r + \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i} \frac{2j+2\alpha-1}{2j} \right) r^{1+2i} .$$

We compare this with the expansion

$$\begin{aligned} r(1-r)^{-2\alpha} &= r + \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i} \frac{j+2\alpha-1}{j} \right) r^{1+i} \\ &> r + \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i} \frac{2j+2\alpha-1}{2j} \right) r^{1+2i} \end{aligned}$$

and obtain the lemma.

4. We shall make use of Lemma 1 to prove:

LEMMA 2. *$B_\alpha \notin O_{H^2D}$ for $\alpha \geqq 1$.*

*Proof.* A necessary and sufficient condition for the existence of an $H^2D$-function $u$ is that the Laplacian $\Delta u = h$ satisfies

$$|(h, \varphi)| \leqq K\sqrt{D(\varphi)} \tag{11}$$

for all $\varphi \in C_0^\infty$ and some constant $K$ independent of $\varphi$ (Nakai-Sario [5]).

Let $h = f_1(r)S_{11} = f_1(r) \cos \theta^1$, and take any $\varphi \in C_0^\infty(B_\alpha)$. By Lemma 1 and the Fourier expansion

$$\varphi = \sum_{n=0}^{\infty} \sum_{m=1}^{2n+1} b_{nm}(r)S_{nm}(\theta^1, \theta^2) ,$$

we obtain

$$\begin{aligned} |(h, \varphi)| &= \left| \text{const} \int_0^1 b_{11}(r)f_1(r)r^2(1 - r^2)^{3\alpha}dr \right| \\ &< \text{const} \int_0^1 |b_{11}(r)|(1 - r)^\alpha dr . \end{aligned}$$

By Schwarz's inequality,

$$\begin{aligned} |(h, \varphi)|^2 &\leqq \text{const} \int_0^1 (1 - r)^\alpha dr \cdot \int_0^1 b_{11}^2(r)(1 - r)^\alpha dr \\ &= \text{const} \int_0^1 b_{11}^2(r)(1 - r)^\alpha dr . \end{aligned} \tag{12}$$

On the other hand,

$$\begin{aligned} &D(\varphi) \\ &= \int_{B_\alpha} |\text{grad}\, \varphi|^2 dV \geqq \text{const} \int_{B_\alpha} r^{-2}(1 - r^2)^{-2\alpha}\left(\frac{\partial \varphi}{\partial \theta^1}\right)^2 r^2(1 - r^2)^{3\alpha}dr \\ &\geqq \text{const} \int_0^1 b_{11}^2(r)(1 - r)^\alpha dr . \end{aligned} \tag{13}$$

5. Denote by $Q$ the class of quasiharmonic functions $u$, characterized by $\Delta_\lambda u = 1$. We recall (Sario-Wang [9]) that $B_\alpha \notin O_{QD}$ if and only if $\alpha \in (-3/5, 1)$. Since $QD \subset H^2D$, we have trivially:

LEMMA 3. *$B_\alpha \notin O_{H^2D}$ if $\alpha \in (-3/5, 1)$.*

6. Next we consider the case $\alpha < -3/5$.

LEMMA 4. *$B_\alpha \in O_{H^2D}$ if $\alpha < -3/5$.*

*Proof.* Suppose there exists an $H^2D$-function $u$ on $B_\alpha$, that is,

$\Delta u = h$ satisfies (11). By Theorem 2, $h$ has the expansion

$$h = \sum_{n=0}^{\infty} \sum_{m=1}^{2n+1} \sum_{i=0}^{\infty} a_{nm} c_{n,2i} r^{n+2i} S_{nm} .$$

If $a_{nm} \neq 0$ for some $(n, m)$, choose for our testing functions $\varphi_t$, $0 < t \leqq 1$,

$$\varphi_t(r, \theta) = \rho_t(r) S_{nm}(\theta), \qquad \rho_t(r) = g\left(\frac{1-r}{t}\right),$$

where $g(r)$ is a fixed nonnegative $C_0^\infty$-function with $\operatorname{supp} g \subset (\beta, \gamma)$, $0 < \beta < \gamma < 1$. Since $\lim_{r \to 1} f_n(r) \neq 0$,

$$\int_{1-\gamma t}^{1-\beta t} \rho_t(r) dr = t \int_{\beta}^{\gamma} g(r) dr ,$$

and $(1 - r^2)^{3\alpha} > 2^{3\alpha}(1 - r)^{3\alpha} \geqq 2^{3\alpha}(\gamma t)^{3\alpha}$ for $\alpha < 0$, we have for sufficiently small $t$,

$$\begin{aligned} |(h, \varphi_t)| &= \text{const} \left| \int_{1-\gamma t}^{1-\beta t} f_n(r) \rho_t(r) r^2 (1 - r^2)^{3\alpha} dr \right| \\ &\geqq \text{const}\, (1 - \gamma)^2 (\gamma t)^{3\alpha} \int_{1-\gamma t}^{1-\beta t} \rho_t(r) dr = \text{const}\, t^{3\alpha+1} . \end{aligned} \tag{14}$$

On the other hand,

$$\begin{aligned} D(\varphi_t) &= \int_{B_\alpha} |\operatorname{grad} \varphi_t|^2 dV \\ &= \int_{1-\gamma t}^{1-\beta t} (1 - r^2)^{-2\alpha} (c_1(\rho'(r))^2 + c_2 r^{-2} \rho^2(r)) r^2 (1 - r^2)^{3\alpha} dr \\ &< \text{const}\, (\gamma t)^{\alpha} \int_{1-\gamma t}^{1-\beta t} (c_1(\rho'(r))^2 + c_2 \rho^2(r)) dr \\ &= d_1 t^{\alpha-1} + d_2 t^{\alpha+1} < d t^{\alpha-1} , \end{aligned} \tag{15}$$

$\alpha < 0$, where $d_1$, $d_2$, and $d$ are independent of $t$. If $\alpha < -3/5$, then (11) is violated as $t \to 0$, a contradiction. Thus $B_\alpha \in O_{H^2 D}$ for $\alpha < -3/5$.

7. It remains to consider the case $\alpha = -3/5$.

LEMMA 5. $B_{-3/5} \in O_{H^2 D}$.

*Proof.* We choose a decreasing sequence of real numbers $t_j \in (0, 1]$ tending to 0 such that $1 - \beta t_j < 1 - \gamma t_{j+1}$ and (14) is satisfied for each $t_j$. Set $q_j = t_j^{-3\alpha-1} j^{-1} \cdot \operatorname{sign}\, (h, \varphi_{t_j})$ and take for the testing functions $\varphi_n = \sum_{j=1}^{n} q_j \varphi_{t_j}$. We obtain by (14)

$$|(h, \varphi_n)| = \left| \sum_{j=1}^{n} q_j (h, \varphi_{t_j}) \right| > \text{const} \sum_{j=1}^{n} j^{-1}$$

and by (15)

$$D(\varphi_n) = \sum_{j=1}^{n} q_j^2 D(\varphi_{t_j}) < \text{const} \sum_{j=1}^{n} j^{-2}(t_j^{-5\alpha-3}) .$$

For $\alpha = -3/5$, we have $D(\varphi_n) < \text{const} \sum_1^n j^{-2}$, which stays bounded as $n \to \infty$ whereas $|(h, \varphi_n)| \to \infty$. Thus (11) is violated, and we conclude that $B_{-3/5} \in O_{H^2D}$.

The proof of Theorem 1 is herewith complete.

8. Since $B_\alpha \notin O_{QD}$ if and only if $\alpha \in (-3/5, 1)$, Theorem 1 has the following applications to the classification of Riemannian manifolds, with $\tilde{O}$ standing for the complement of $O$:

THEOREM 3. *The totality of Riemannian* 3*-manifolds has the decomposition*

$$\{R\} = O_{H^2D} \cup (O_{QD} \cap \tilde{O}_{H^2D}) \cup \tilde{O}_{QD}$$

*into three disjoint nonempty subclasses.*

THEOREM 4. *There exist parabolic Riemannian* 3*-manifolds which carry $H^2D$-functions, and hyperbolic Riemannian* 3*-manifolds which do not carry $H^2D$-functions.*

For dimension 2, this was shown in Nakai-Sario [5], but for higher dimensions it has been an open problem.

For the proof of Theorem 4, let $O_G$ be the class of parabolic Riemannian manifolds. It was proved in Sario-Wang [9] that $B_\alpha \in O_G$ if and only if $\alpha \geqq 1$. As a consequence,

$$B_\alpha \in O_G \cap \tilde{O}_{H^2D} \Longleftrightarrow \alpha \geqq 1 ,$$

$$B_\alpha \in \tilde{O}_G \cap O_{H^2D} \Longleftrightarrow \alpha \leqq -\frac{3}{5} .$$

We shall return to the classification of higher dimensional Riemannian manifolds in further studies.

We are indebted to Mr. Dennis Hada, who preused the manuscript and made his valued comments.

## REFERENCES

1. Y. K. Kwon, L. Sario, and B. Walsh, *Behavior of biharmonic functions on Wiener's and Royden's compactifications*, Ann. Inst. Fourier, (Grenoble), **21** (1971), 217–226.
2. C. Müller, *Spherical Harmonics*, Lecture Notes in Mathematics 17, Springer, 1966.

3. M. Nakai and L. Sario, *Biharmonic classification of Riemannian manifolds*, Bull. Amer. Math. Soc., **77** (1971), 432-436.
4. ————, *Quasiharmonic classification of Riemannian manifolds*, Proc. Amer. Math. Soc., **31** (1972), 165-169.
5. ————, *Existence of Dirichlet finite biharmonic functions*, Ann. Acad. Sci. Fenn. A. I. **532** (1973), 1-33.
6. H. O'Malla, *Dirichlet finite biharmonic functions on the unit disk with distorted metrics*, Proc. Amer. Math. Soc., **32** (1972), 521-524.
7. R. Narasimhan, *Analysis on Real and Complex Manifolds*, Masson, Paris, 1968, 246 pp.
8. L. Sario, *Biharmonic and quasiharmonic functions on Riemannian manifolds*, Duplicated lecture notes 1969-1970, University of California, Los Angeles.
9. L. Sario and C. Wang, *Quasiharmonic functions on the Poincaré N-ball*, Rend. Mat., (to appear).

Received July 19, 1972. The work was sponsored by the U. S. Army Research Office—Durham, Grant DA-ARO-D-31-124-71-G181, University of California, Los Angeles.

UNIVERSITY OF CALIFORNIA, LOS ANGELES

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# ON A GENERALIZATION OF MARTINGALES DUE TO BLAKE

R. SUBRAMANIAN

**It is shown that any uniformly integrable fairer with time game (stochastic process) converges in $L_1$.**

**1. Introduction.** Let $(\Omega, \mathscr{U}, P)$ be a probability space and $\{\mathscr{U}_n\}_{n\geqq 1}$ an increasing family of sub $\sigma$-algebras of $\mathscr{U}$. Let $\{X_n\}_{n\geqq 1}$ be a stochastic process adapted to $\{\mathscr{U}_n\}_{n\geqq 1}$ (see, [2, p. 65]). Following Blake [1] we refer to $\{X_n\}_{n\geqq 1}$ as a game and define

DEFINITION. The game $\{X_n\}_{n\geqq 1}$ will be said to become *fairer with time* if for every $\varepsilon > 0$

$$P[|E(X_n/\mathscr{U}_m) - X_m| > \varepsilon] \to 0$$

as $n, m \to \infty$ with $n \geqq m$. Any martingale is, trivially, a fairer with time game and thus this concept generalizes that of martingales. Blake, in [1], gave a set of sufficient conditions under which any uniformly integrable fairer with time game $\{X_n\}_{n\geqq 1}$ is convergent in $L_1$. We show that these sufficient conditions are not needed; in fact, we show that any uniformly integrable, fairer with time game converges in $L_1$.

2. THEOREM 2.1. *Any uniformly integrable fairer with time game $\{X_n\}_{n\geqq 1}$ converges in $L_1$.*

*Proof.* To facilitate understanding, we break up the proof into a few important steps numbered (S1) through (S5). For every $m$ and $n \geqq m$ define $Y_{m,n} = E(X_n/\mathscr{U}_m)$. Let $\Gamma$ stand for the family $\{Y_{m,n},$ for all $m$ and $n \geqq m\}$.

(S1) $\Gamma$ is uniformly integrable.

Since $\{X_n\}_{n\geqq 1}$ is uniformly integrable there exists a function $f$ defined on the nonnegative real axis which is positive, increasing and convex, such that

$$\lim_{t\to\infty} \frac{f(t)}{t} = +\infty$$

and $\sup_n E[f \circ |X_n|] < \infty$. (See [2, II T 22].) Now,

$$\begin{aligned} E[f \circ |Y_{m,n}|] &= E[f \circ |E(X_n/\mathscr{U}_m)|] \\ &\leqq E[f \circ E(|X_n|/\mathscr{U}_m)] \text{ (since } f \text{ is nondecreasing)} \\ &\leqq E[E(f \circ |X_n|/\mathscr{U}_m)] \\ &= E[f \circ |X_n|] . \end{aligned}$$

Therefore,

$$\sup_{Y_{m,n}\in\Gamma} E[f\circ|Y_{m,n}|] \leqq \sup_n E[f\circ|X_n|] < \infty .$$

Another application of II T 22 of [2] ensures that $\Gamma$ is uniformly integrable. Hence (S1).

(S2) Given $\varepsilon > 0$, there exists $M$ such that for all $m \geqq M$, one has

$$E(|X_m - Y_{m,n}|) \leqq 2\varepsilon \quad \text{for all} \quad n \geqq m .$$

Since $\Gamma$ is uniformly integrable given $\varepsilon > 0$ there exists $\delta > 0$ such that $P(A) < \delta$ implies $\int_A |Y_{m,n}|\,dP \leqq \varepsilon/2$, for all $Y_{m,n}\in\Gamma$. Choose $M$ so large that $m \geqq M$ and $n \geqq m$ implies $P[|X_m - E(X_n/U_m)| > \varepsilon] < \delta$. Then, it is not difficult to see that

$$E[|X_m - Y_{m,n}|] \leqq 2\varepsilon \quad \text{for all } m \geqq M \quad \text{and} \quad n \geqq m .$$

(S3) For every fixed $m$, the sequence $\{Y_{m,n}\}$ converges in $L_1$ to an $\mathscr{U}_m$ measurable random variable $Z_m$.

Let $m \leqq n < n'$.

$$\begin{aligned} E[|Y_{m,n} - Y_{m,n'}|] &= E[|E(X_n/\mathscr{U}_m) - E(X_{n'}/\mathscr{U}_m)|] \\ &= E[|E(X_n - X_{n'}/\mathscr{U}_m)|] \\ &= E[|E(\{E(X_n - X_{n'}/\mathscr{U}_n)\}/\mathscr{U}_m)|] \\ &\leqq E[E(\{|E(X_n - X_{n'}/\mathscr{U}_n)|\}/\mathscr{U}_m)] \\ &= E[|E(X_n - X_{n'}/\mathscr{U}_n)|] \\ &= E[|X_n - Y_{n,n'}|] . \end{aligned}$$

Now from (S2) it follows that given $\varepsilon > 0$ for all sufficiently large $n$ and $n'$

$$E[|Y_{m,n} - Y_{m,n'}|] \leqq E[|(X_n - Y_{n,n'})|] \leqq 2\varepsilon .$$

Hence, for $m$ fixed, the sequence $\{Y_{m,n}\}$ is Cauchy in the $L_1$-norm. So, there exists, an integrable random variable $Z_m$, such that, $Y_{m,n} \xrightarrow[n\to\infty]{L_1} Z_m$. Without loss of generality we can take $Z_m$ to be $\mathscr{U}_m$ measurable. (Note that each $Y_{m,n}$ is $\mathscr{U}_m$ measurable and there is a subsequence $\{Y_{m,n'}\}$ converging almost surely to $Z_m$.)

(S4) $\{Z_m, \mathscr{U}_m\}_{m\geqq 1}$ is a uniformly integrable martingale.

The fact that $\{Z_m\}_{m\geqq 1}$ is uniformly integrable follows trivially because the closure in $L_1$ of a uniformly integrable collection is uniformly integrable. (See, [2, II T20].) To show $\{Z_m, \mathscr{U}_m\}$ is a martingale it is enough to show that for every $m$, $E(Z_{m+1}/\mathscr{U}_m) = Z_m$ a.s. Since

$$
\begin{aligned}
E[| E(Y_{m+1,n}/\mathscr{U}_m) &- E(Z_{m+1}/\mathscr{U}_m) |] \\
&= E[| E\{(Y_{m+1,n} - Z_{m+1})/\mathscr{U}_m\} |] \\
&\leqq E[E\{| (Y_{m+1,n} - Z_{m+1}) |/\mathscr{U}_m\}] \\
&= E[| Y_{m+1,n} - Z_{m+1} |] \longrightarrow 0 \quad \text{as} \quad n \longrightarrow \infty ,
\end{aligned}
$$

there exists a subsequence $n'$ of $\{n: n \geqq m\}$ such that

$$E(Y_{m+1,n'}/\mathscr{U}_m) \xrightarrow{\text{a.s.}} E(Z_{m+1}/\mathscr{U}_m) .$$

We can assume (— if necessary, by choosing a further subsequence, —) that $Y_{m,n'} \xrightarrow{\text{a.s.}} Z_m$. Now,

$$
\begin{aligned}
E(Z_{m+1}/\mathscr{U}_m) &= \lim_{n' \to \infty} E(Y_{m+1,n'}/\mathscr{U}_m) \quad \text{a.s.} \\
&= \lim_{n' \to \infty} E(\{E(X_{n'}/\mathscr{U}_{m+1})\}/\mathscr{U}_m) \quad \text{a.s.} \\
&= \lim_{n' \to \infty} E(X_{n'}/\mathscr{U}_m) \quad \text{a.s.} \\
&= \lim_{n' \to \infty} Y_{m,n'} \quad \text{a.s.} \\
&= Z_m \quad \text{a.s.}
\end{aligned}
$$

Hence (S4). (S5) $\{X_n\}_{n\geqq 1}$ converges in $L_1$.

Since $\{Z_n, \mathscr{U}_n\}_{n\geqq 1}$ is an uniformly integrable martingale, there exists an integrable random variable $Z_\infty$ such that $Z_n \xrightarrow[n\to\infty]{L_1} Z_\infty$. We shall show that $X_n \xrightarrow[n\to\infty]{L_1} Z_\infty$. From (S3) and (S2) it is easy to check that given $\varepsilon > 0$ there exists $M$ such that for all $m \geqq M$

$$\int | X_m - Z_m | \, dP \leqq 2\varepsilon .$$

Therefore, for sufficiently large $m$,

$$\int | X_m - Z_\infty | \, dP \leqq \int | X_m - Z_m | \, dP + \int | Z_m - Z_\infty | \, dP \leqq 3\varepsilon ,$$

say. Hence (S5) and the theorem.

Since any game (stochastic process) $\{X_n\}_{n\geqq 1}$ converging in $L_1$ can be taken to be a game fairer with time, by setting $\mathscr{U}_n \equiv \mathscr{U}$ in $n$, we get the following corollary.

COROLLARY 2.1. *Let $\{X_n\}_{n\geqq 1}$ be a game. It converges in $L_1$ if and only if it is uniformly integrable and fairer with time with respect to some increasing family of sub $\sigma$-algebras $\{\mathscr{U}_n\}_{n\geqq 1}$ to which it is adapted.*

Let $p > 1$.

THEOREM 2.2. *Let $\{X_n\}_{n\geqq 1}$ be a fairer with time game with $\{|X_n|^p\}_{n\geqq 1}$ uniformly integrable. Then $\{X_n\}_{n\geqq 1}$ converges in Lp.*

*Proof.* Noting that the function $f$ defined on the nonnegative real axis by $f(t) = t^p$ is positive, increasing and convex and $\lim_{t\to\infty}(f(t)/t) = +\infty$, in view of II T 22 of [2], it is clear that $\{X_n\}_{n\geqq 1}$ is uniformly integrable. Hence by Theorem 2.1 it converges in $L_1$; in particular, $\{X_n\}_{n\geqq 1}$ converges in probability. Therefore, $\{X_n\}_{n\geqq 1}$ converges in $L_p$. (See Proposition II 6.1 of [3].)

COROLLARY 2.2. *The game $\{X_n\}_{n\geqq 1}$ converges in $L_p$ if and only if $\{|X_n|^p\}_{n\geqq 1}$ is uniformly integrable and $\{X_n\}_{n\geqq 1}$ is fairer with time with respect to some increasing family of sub $\sigma$-algebras $\{\mathscr{U}_n\}_{n\geqq 1}$ to which it is adapted.*

REMARK. In view of our Theorem 2.1, the second convergence theorem of Blake in [1] becomes redundant.

## REFERENCES

1. L. H. Blake, *A generalization of martingales and two consequent convergence theorems*, Pacific J. Math., **35** (1970), 279-283.

2. P. A. Meyer, *Probability and Potentials*, Blaisdell Publishing Company, Waltham, Massachusetts, 1966.

3. J. Neveu, *Mathematical Foundations of the Calculus of Probability*, Holden-Day, Inc., 1965.

Received March 10, 1972.

INDIAN STATISTICAL INSTITUTE
CALCUTTA-35, INDIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# ON STRONGLY NONLINEAR ELLIPTIC VARIATIONAL INEQUALITIES

BUI AN TON

**The existence of bounded solutions of nonlinear elliptic variational inequalities is shown. The nonlinear second order elliptic operator involved has at most an exponential growth in $u$ and a polynomial growth in $Du$. The regularity of the solutions is studied.**

Let $K$ be a closed convex subset of a reflexive Banach space $V$ and let $A$ be a pseudo-monotone coercive operator from $V$ into $V^*$. Then for any $f$ in $V^*$ it is known that there exists $u$ in $K$ such that:

$$(Au - f, v - u) \geqslant 0 \qquad \text{for all } v \text{ in } K .$$

The existence of solutions of nonlinear elliptic variational inequalities has been shown by Brezis [1], Browder [4], Lions-Stampacchia [10] and others. The regularity of the solutions when $A$ is a linear second order elliptic operator written in divergence form has been studied by Brezis [2], Lewy-Stampacchia [8], Lions [9] and an abstract regularity result has been obtained by Brezis-Stampacchia [3] when $A$ is nonlinear.

It is the purpose of this paper to show the existence of *bounded* solutions of nonlinear variational inequalities when $A$ is a pseudo-monotone coercive operator defined on $V \cap L^\infty(G)$ and mapping $V \cap L^\infty(G)$ into $V^*$. $V$ is a given closed subspace of $W^{1,p}(G)$ with $W_0^{1,p}(G) \subset V \subset W^{1,p}(G)$. The nonlinear elliptic operator $A$ may have an *exponential growth in $u$* and a polynomial growth in $Du$. Abstract existence theorems are proved in §2. The applications and the regularity of the solutions are studied in §3. The notations and the basic assumptions are given in §1.

1. Let $G$ be a bounded open subset of $R^n$ with a smooth boundary $\partial G$. Set: $D_j = \partial/\partial x_j$, $j = 1, \cdots, n$ and

$$W^{1,p}(G) = \{u\colon u \quad \text{in} \quad L^p(G), D_j u \quad \text{in} \quad L^p(G), j = 1, \cdots, n\} .$$

$W^{1,p}(G)$ is a real reflexive separable Banach space with the norm:

$$||u||_{1,p} = \left\{||u||^p_{L^p(G)} + \sum_{j=1}^n ||D_j u||^p_{L^p(G)}\right\}^{1/p} ; 2 \leqslant p < \infty .$$

$W_0^{1,p}(G)$ is the completion of $C_0^\infty(G)$, the family of all infinitely differentiable functions with compact support in $G$, in the $||\cdot||_{1,p}$ norm.

Let $V$ be a closed linear subspace of $W^{1,p}(G)$ with $W_0^{1,p}(G) \subset V \subset W^{1,p}$ and let $H = L^2(G)$. The pairing between $V$ and its dual $V^*$ is denoted by $(\cdot, \cdot)$.

Let $W$ be a closed linear subspace of $W^{1,p+n}(G)$ with $W_0^{1,p+n}(G) \subset W \subset W^{1,p+n}(G)$. The Sobolev imbedding theorem gives $W \subset C(\mathrm{cl}\, G)$. The pairing between $W$ and $W^*$ is denoted by $((\cdot, \cdot))$. Throughout the paper we shall assume:

(a) $W \subset V$.

(b) If $u$ is in $W$ then $|u|^{s-2}u$ is also in $W$ for all $s$ with $2 \leqslant s < \infty$. The assumption is verified if $V = W_0^{1,p}(G)$, $W = W_0^{1,p+n}(G)$ and if $V = W^{1,p}(G)$ with $W = W^{1,p+n}(G)$.

$C^{0,\alpha}(G)$, $0 < \alpha < 1$, is the family of all functions $u$ which are Hölder-continuous with exponent $\alpha$ on any compact subset of $G$.

Set:

$$F = V \cap L^\infty(G) .$$

In this paper nonlinear operators $A(u, v)$ mapping $F \times V$ into $V^*$ satisfying the following assumption are considered.

*Assumption* (I): (i) $A(u, v)$ *maps bounded sets of* $F \times V$ *into bounded sets of* $V^*$.

(ii) *If* $u_n \to u$ *weakly in* $V$, $u_n \to u$ *in the weak*-topology of* $L^\infty(G)$ *and* $\limsup\, (A(u_n, u_n), u_n - u) \leqslant 0$ *then*:

$$\liminf\, (A(u_n, u_n), u_n - v) \geqslant (A(u, u), u - v) \quad \textit{for all } v \textit{ in } V .$$

The pseudo-monotone operators from $V$ into $V^*$ satisfy all the conditions of Assumption (I).

Let $A(u, v)$ be the operator defined by:

$$(A(u, v), w) = \sum_{j=1}^{n} \int_G (1 + c(x) \exp u)\, |D_j v|^{p-2} D_j v D_j w dx + \int_G b(x) u \exp u \cdot w dx$$

$u$ is in $F$, $v$, $w$ in $V$, $c(x)$ and $b(x)$ are two nonnegative bounded functions on $G$. $A(u, v)$ is the prototype of operators considered in this paper. It is not difficult to check that $A(u, v)$ satisfies all the conditions of Assumption (I).

In the proof of the existence theorems we need the following auxiliary nonlinear monotone operator:

$$A_2 v = -\sum_{j=1}^{n} D_j(|D_j v|^{p+n-2} D_j v) + v .$$

2. In this section the existence of bounded solutions of nonlinear elliptic variational inequalities is shown.

The following theorem gives a generalization of a result of Browder [4], Hartman-Stampacchia [6] and others when $V$ is a closed linear subspace of $W^{1,p}(G)$.

THEOREM 1. *Let $A(u, v)$ be a nonlinear operator mapping $F \times V$ into $V^*$ and satisfying Assumption* (I). *Let $K$ be a closed convex subset of $V$ with $0 \in K$ and let $\beta$ be the penalty operator associated with $K$. Suppose that:*

(i) $(A(u, u), u) \geqslant c\|u\|_V^p$ *for all $u$ in $F$.*

(ii) $(A(u, u), |u|^{s-2}u) \geqslant c\|u\|_{L^s(G)}^s$ *for all $u$ in $W$ and all $s$ with $2 \leqslant s < \infty$. The positive constant $c$ is independent of $s$.*

(iii) $(\beta(u), |u|^{s-2}u) \geqslant 0$ *for all $u$ in $W$ and all $s$, $2 \leqslant s < \infty$. Then for any $f$ in $L^\infty(G)$ there exists $u$ in $F \cap K$ such that:*

$$(A(u, u) - f, v - u) \geqslant 0 \qquad \textit{for all } v \textit{ in } K.$$

THEOREM 2. *Let $A(u, w)$ be as in Theorem 1 and let $K$ be a closed convex subset of both $V$ and $W$ with $0 \in K$. Let $\beta$ be the penalty operator associated with $K$ considered as a subset of $W$. Suppose all the hypotheses of Theorem 1 are satisfied with* (iii) *replaced by:*

$$((\beta(u), u))/\|u\|_W \longrightarrow +\infty \quad \textit{as} \quad \|u\|_W \longrightarrow +\infty .$$

*Then for any $f$ in $L^\infty(G)$ there exists $u$ in $F \cap K$ such that:*

$$(A(u, u) - f, v - u) \geqslant 0 \qquad \textit{for all } v \textit{ in } K.$$

For a smaller class of nonlinear operators $A(u, v)$ we have the following theorem which extends a result of Dubinskii [5].

THEOREM 3. *Let $A(u, v)$ be a nonlinear operator mapping bounded sets of $L^\infty(G) \times V$ into bounded sets of $V^*$. Suppose that:*

(i) $(A(u, v) - A(u, w), v - w \geqslant 0$ *for all $u$ in $L^\infty(G)$ and all $v, w$ in $V$.*

(ii) *For fixed $u$ in $L^\infty(G)$, $A(u, \cdot)$ is continuous from the strong topology of $V$ to the weak topology of $V^*$.*

(iii) *For fixed $v$ in $V$, if $u_n \to u$* a.e. *on $G$ and $u_n \to u$ in the weak*-topology of $L^\infty(G)$ then: $A(u_n, v) \to A(u, v)$ in $V^*$.*

(iv) $(A(|u|^{s-2}u, u), u) \geqslant c\|u\|_V^p$ *for all $u$ in $W$ and all $s$ with $1 < s < \infty$.*

(v) $(A(|u|^{s-2}u, u), |u|^{r-2}u) \geqslant c\|u\|_{L^{r+s-2}(G)}^{r+s-2}$ *for all $u$ in $W$ with $1 < s < \infty$ and $2 \leqslant r < \infty$. The constant $c$ is independent of $r$.*

*Then for any $f$ in $L^\infty(G)$ there exists $u$ in $L^\infty(G)$ with $|u|^{\alpha-2}u$ in $V$ such that:*

$$A(u, |u|^{\alpha-2}u) = f .$$

*$\alpha$ is any number with* $1 < \alpha < \infty$.

*Proof of Theorem* 1. (1) Let $0 < \eta < \varepsilon < 1$ and let $\mathscr{A}(u, v)$ be the nonlinear operator:

$$\mathscr{A}(u, v) = \eta A_2 v + A(u, u) + \varepsilon^{-1}\beta(u) .$$

$\mathscr{A}(u, v)$ maps bounded sets of $W \times W$ into bounded sets of $W^*$ and is coercive. Being the sum of a monotone operator and a pseudo-monotone operator, $\mathscr{A}$ is a pseudo-monotone operator from $W$ into $W^*$. It follows from the theory of coercive pseudo-monotone operators that there exists $u_{\varepsilon\eta}$ in $W$, solution of the equation:

$$\eta A_2 u_{\varepsilon\eta} + A(u_{\varepsilon\eta}, u_{\varepsilon\eta}) + \varepsilon^{-1}\beta(u_{\varepsilon\eta}) = f .$$

It is clear that:

$$\eta \|u_{\varepsilon\eta}\|_W^{p+n} + \|u_{\varepsilon\eta}\|_V \leqslant C .$$

$C$ is a constant independent of both $\varepsilon$ and $\eta$.
Since $u_{\varepsilon\eta}$ is in $W$, $|u_{\varepsilon\eta}|^{s-2}u_{\varepsilon\eta}$ lies also in $W$ for all $s$ with $2 \leqslant s < \infty$. Thus:

$$\eta((A_2 u_{\varepsilon\eta}, |u_{\varepsilon\eta}|^{s-2}u_{\varepsilon\eta})) + (A(u_{\varepsilon\eta}, u_{\varepsilon\eta}), |u_{\varepsilon\eta}|^{s-2}u_{\varepsilon\eta}) \\ + \varepsilon^{-1}(\beta(u_{\varepsilon\eta}), |u_{\varepsilon\eta}|^{s-2}u_{\varepsilon\eta}) = (f, |u_{\varepsilon\eta}|^{s-2}u_{\varepsilon\eta}) .$$

An elementary computation gives:

$$((A_2 u_{\varepsilon\eta}, |u_{\varepsilon\eta}|^{s-2}u_{\varepsilon\eta})) \geqslant 0 .$$

From the hypotheses of the theorem, we get:

$$\|u_{\varepsilon\eta}\|_{L^s(G)} \leqslant C \|f\|_{L^\infty(G)} .$$

Since $u_{\varepsilon\eta}$ is in $W$ hence in $L^\infty(G)$ we may let $s \to +\infty$.
So:

$$\|u_{\varepsilon\eta}\|_{L^\infty(G)} \leqslant C \|f\|_{L^\infty(G)} .$$

$C$ is a constant independent of both $\varepsilon$ and $\eta$.

(2) From the weak compactness of the unit ball in a reflexive Banach space we obtain by taking subsequences if necessary: $u_{\varepsilon\eta} \to u_\varepsilon$ weakly in $V$, $u_{\varepsilon\eta} \to u_\varepsilon$ in the weak*-topology of $L^\infty(G)$, $\eta^{1/p+n}u_{\varepsilon\eta} \to 0$ weakly in $W$, $A(u_{\varepsilon\eta}, u_{\varepsilon\eta}) \to g_\varepsilon$ weakly in $V^*$ and $\beta(u_{\varepsilon\eta}) \to h_\varepsilon$ weakly in $V^*$ as $\eta \to 0$.

So:

$$\|u_\varepsilon\|_V + \|u_\varepsilon\|_{L^\infty(G)} \leqslant C .$$

Moreover:

$$g_\varepsilon + \varepsilon^{-1} h_\varepsilon = f .$$

We show that $g_\varepsilon = A(u_\varepsilon, u_\varepsilon)$ and $h_\varepsilon = \beta(u_\varepsilon)$.
We have:

$$\limsup_{\eta \to 0} (A(u_{\varepsilon\eta}, u_{\varepsilon\eta}) + \varepsilon^{-1}\beta(u_{\varepsilon\eta}), u_{\varepsilon\eta}) \leqslant (g_\varepsilon + \varepsilon^{-1} h_\varepsilon, u_\varepsilon) .$$

On the other hand:

$$\begin{aligned}(A(u_{\varepsilon\eta}, u_{\varepsilon\eta}), u_{\varepsilon\eta} - u_\varepsilon) &= (A(u_{\varepsilon\eta}, u_{\varepsilon\eta}) + \varepsilon^{-1}\beta(u_{\varepsilon\eta}), u_{\varepsilon\eta} - u_\varepsilon) \\ &\quad - \varepsilon^{-1}(\beta(u_{\varepsilon\eta}) - \beta(u_\varepsilon), u_{\varepsilon\eta} - u_\varepsilon) - \varepsilon^{-1}(\beta(u_\varepsilon), u_{\varepsilon\eta} - u_\varepsilon) .\end{aligned}$$

Taking into account the monotonicity of the penalty operator, we get:

$$\begin{aligned}(A(u_{\varepsilon\eta}, u_{\varepsilon\eta}), u_{\varepsilon\eta} - u_\varepsilon) \leqslant (A(u_{\varepsilon\eta}, u_{\varepsilon\eta}) &+ \varepsilon^{-1}\beta(u_{\varepsilon\eta}), u_{\varepsilon\eta} - u_\varepsilon) \\ &- \varepsilon^{-1}(\beta(u_\varepsilon), u_{\varepsilon\eta} - u_\varepsilon) .\end{aligned}$$

Thus:

$$\limsup_{\eta \to 0} (A(u_{\varepsilon\eta}, u_{\varepsilon\eta}), u_{\varepsilon\eta} - u_\varepsilon) \leqq 0 .$$

Assumption (I) gives:

$$A(u_\varepsilon, u_\varepsilon) = g_\varepsilon .$$

It is now easy to show that $h_\varepsilon = \beta(u_\varepsilon)$.

(3) Again from the weak compactness of the unit ball in a reflexive Banach space we get by taking subsequences if necessary: $u_\varepsilon \to u$ weakly in $V$, $u_\varepsilon \to u$ in the weak*-topology of $L^\infty(G)$ and $A(u_\varepsilon, u_\varepsilon) \to g$ weakly in $V^*$ as $\varepsilon \to 0$.

Since:

$$A(u_\varepsilon, u_\varepsilon) + \varepsilon^{-1}\beta(u_\varepsilon) = f$$

it is clear that

$$\beta(u_\varepsilon) \longrightarrow 0 \quad \text{in} \quad V^* \quad \text{as} \quad \varepsilon \longrightarrow 0 .$$

The penalty operator $\beta$ is a monotone hemi-continuous operator from $V$ into $V^*$. A standard argument gives:

$$\beta(u) = 0 \text{ i.e., } u \in K .$$

We have:

$$\begin{aligned}(A(u_\varepsilon, u_\varepsilon), u_\varepsilon - u) &= (f - \varepsilon^{-1}\beta(u_\varepsilon), u_\varepsilon - u) \\ &= (f, u_\varepsilon - u) - \varepsilon^{-1}(\beta(u_\varepsilon) - \beta(u), u_\varepsilon - u) .\end{aligned}$$

Thus:

$$(A(u_\varepsilon, u_\varepsilon), u_\varepsilon - u) \leqq (f, u_\varepsilon - u) .$$

Hence:

$$\limsup_{\varepsilon \to 0} (A(u_\varepsilon, u_\varepsilon), u_\varepsilon - u) \leqq 0 .$$

Assumption (I) gives:

$$A(u, u) = g$$

and

$$\liminf_{\varepsilon \to 0} (A(u_\varepsilon, u_\varepsilon), u_\varepsilon - v) \geqslant (A(u, u), u - v) .$$

Let $v \in K$ and we have:

$$\begin{aligned}(A(u_\varepsilon, u_\varepsilon) - f, v - u_\varepsilon) &= -\varepsilon^{-1}(\beta(u_\varepsilon), v - u_\varepsilon) \\ &= \varepsilon^{-1}(\beta(v) - \beta(u_\varepsilon), v - u_\varepsilon) .\end{aligned}$$

So:

$$(A(u_\varepsilon, u_\varepsilon) - f, v - u_\varepsilon) \geqslant 0 \qquad \text{for all } v \text{ in } K .$$

Let $\varepsilon \to 0$ and we obtain:

$$(A(u, u) - f, v - u) \geqq 0 \qquad \text{for all } v \text{ in } K .$$

The theorem is proved.

*Proof of Theorem* 2. The proof is similar to that of Theorem 1, we shall not reproduce it.

*Proof of Theorem* 3. (1) Let $0 < \varepsilon < 1$ and consider the nonlinear operator $\mathscr{A}(u, v)$ defined by:

$$\mathscr{A}(u, v) = \varepsilon A_2 v + A(|u|^{\alpha'-2}u, u)$$

$u$ and $v$ are in $W$ and $1/\alpha + 1/\alpha' = 1$.

$\mathscr{A}$ is coercive and maps bounded sets of $W \times W$ into bounded sets of $W^*$. Being the sum of a monotone operator and a pseudo-monotone operator, $\mathscr{A}$ is pseudo-monotone. Therefore, there exists $v_\varepsilon$ in $W$, solution of the equation:

$$\varepsilon A_2 v_\varepsilon + A(|v_\varepsilon|^{\alpha'-2} v_\varepsilon, v_\varepsilon) = f .$$

It is easy to show that:

$$\varepsilon \|v_\varepsilon\|_W^{p+n} + \|v_\varepsilon\|_V \leqslant C .$$

$C$ is a constant independent of $\varepsilon$.

Since $v_\varepsilon$ is in $W$, $|v_\varepsilon|^{s-2}v_\varepsilon$ lies in $W$ for all $s$ with $2 \leqslant s < \infty$. Thus:

$$\varepsilon((A_2 v_\varepsilon, |v_\varepsilon|_\varepsilon^{s-2} v_\varepsilon)) + (A(|v_\varepsilon|^{\alpha'-2}v_\varepsilon, v_\varepsilon), |v_\varepsilon|^{s-2}v_\varepsilon) = (f, |v_\varepsilon|^{s-2}v_\varepsilon) .$$

An elementary computation shows that the first term of the left hand side of the above equation is positive. It follows from the hypotheses of the theorem that:

$$||v_\varepsilon||^{\alpha'-1}_{L^{s+\alpha'-2}(G)} \leqslant C\,||f||_{L^\infty(G)} .$$

Since $v_\varepsilon$ is in $W$ we may let $s \to +\infty$. Hence:

$$||v_\varepsilon||_{L^\infty(G)} \leqslant M .$$

$M$ is independent of $\varepsilon$.

(2) The weak compactness of the unit ball in a reflexive Banach space gives by taking subsequences if necessary: $v_\varepsilon \to v$ weakly in $V$, $v_\varepsilon \to v$ in the weak*-topology of $L^\infty(G)$, $\varepsilon^{1/p+n}v_\varepsilon \to 0$ weakly in $W$, $|v_\varepsilon|^{\alpha'-2}v_\varepsilon \to h$ weakly in $L^\alpha(G)$ and

$$A(|v_\varepsilon|^{\alpha'-2}v_\varepsilon, v_\varepsilon) \longrightarrow g \quad \text{weakly in} \quad V^* \quad \text{as} \quad \varepsilon \longrightarrow 0 .$$

Since the injection mapping of $V$ into $H$ is compact, the Lebesgue convergence theorem yields:

$$v_\varepsilon \longrightarrow v \quad \text{in} \quad L^s(G) \quad \text{for any } s \text{ with} \quad 1 < s < \infty .$$

Thus:

$$|v_\varepsilon|^{\alpha'-2}v_\varepsilon \longrightarrow |v|^{\alpha'-2}v \text{ a.e.} \quad \text{on} \quad G \quad \text{and} \qquad h = |v|^{\alpha'-2}v .$$

We have:

$$\limsup\, (A(|v_\varepsilon|^{\alpha'-2}v_\varepsilon, v_\varepsilon), v_\varepsilon) \leqslant (g, v) .$$

So:

$$\limsup\, (A(|v_\varepsilon|^{\alpha'-2}v_\varepsilon, \varphi), v_\varepsilon - \varphi) \leqslant (g, v - \varphi) \qquad \text{for all } \varphi \text{ in } V .$$

It follows from the hypotheses of the theorem that:

$$(g - A(|v|^{\alpha'-2}v, \varphi), v - \varphi) \geqslant 0 \qquad \text{for all } \varphi \text{ in } V .$$

Hence:

$$A(|v|^{\alpha'-2}v, v) = g = f .$$

Set:

$$u = |v|^{\alpha'-2}v .$$

Then:

$$u \in L^\infty(G) \quad \text{and} \quad v = |u|^{\alpha-2}u \in V .$$

Moreover:

$$A(u, |u|^{\alpha-2}u) = f .$$

The theorem is proved.

3. We shall now give some applications of the theorems proved in the previous section and study the regularity of the solutions. First consider the case when the constraint is in $G$.

THEOREM 4. *Let* $V = W_0^{1,p}(G)$, $K = \{u\colon u \textit{ in } L^p(G), u \geqslant 0 \text{ a.e. } \textit{on } G\}$ *and let* $A(u, v) = -\sum_{j=1}^n D_jA_j(x, u, Dv)$.
*Suppose that*:

( i ) *For each* $j$, $A_j(x, u, Dv)$ *is continuous with respect to* $x$, $u$, $Dv$.
(ii) $|A_j(x, u, Dv)| \leqslant C\{1 + \exp(|u|) + |Dv|^{p-1}\exp(|u|)\}$.
(iii) $\sum_{j=1}^n (A_j(x, u, Dv) - A_j(x, u, Dw))(D_jv - D_jw) \geqslant 0$.
(iv) $\sum_{j=1}^n A_j(x, u, Du)D_ju \geqslant c\sum_{j=1}^n |D_ju|^p$.

*Then for any* $f$ *in* $L^\infty(G)$ *there exists* $u$ *in* $K \cap V \cap L^\infty(G)$ *such that*:

$$(A(u, u) - f, v - u) \geqslant 0 \quad \textit{for all } v \textit{ in } V \cap K .$$

*Moreover*:

$$A(u, u) \in L^\infty(G) .$$

*Proof*. (1) $K$ is a closed convex subset of $V$ and $0 \in K$. Let $\beta(v) = -|v^-|^{p-2}v^-$ where $v^- = 0$ if $v(x) \geqslant 0$ and $v^- = -v(x)$ otherwise. It is easy to check that $\beta$ verifies the hypothesis of Theorem 1. Let $0 < \eta < 1$ and consider the operator:

$$A_1(u, v) = A(u, v) + \eta u .$$

$A_1(u, v)$ and $\eta A_2v + A_1(u, v)$ satisfy all the hypotheses of Theorem 1. Therefore, from Theorem 1, we get:

$$\eta A_2u_{\varepsilon\eta} + A_1(u_{\varepsilon\eta}, u_{\varepsilon\eta}) + \varepsilon^{-1}\beta(u_{\varepsilon\eta}) = f . \quad 0 < \eta < \varepsilon < 1 .$$

Moreover, from the proof of Theorem 1 we have:

$$\eta\|u_{\varepsilon\eta}\|_W^{p+n} + \|u_{\varepsilon\eta}\|_V + \eta\|u_{\varepsilon\eta}\|_{L^\infty(G)} \leqslant C .$$

$C$ is a constant independent of both $\varepsilon$ and $\eta$.

Since $u_{\varepsilon\eta}$ is in $W$ and hence in $C(\mathrm{cl}\, G)$ we get:

$$\operatorname*{ess\,sup}_{\partial G} |u_{\varepsilon\eta}| = 0 .$$

It follows from Theorem 7.1 of [7], p. 287 that $\|u_{\varepsilon\eta}\|_{L^\infty(G)}$ is majorized by $\|u_{\varepsilon\eta}\|_V$, $c$, the $L^\infty(G)$-norm of $f$.
Hence:

$$\|u_{\varepsilon\eta}\|_{L^\infty(G)} \leqslant C .$$

$C$ is again a constant independent of both $\varepsilon$ and $\eta$.

(2) We have:

$$\eta((A_2 u_{\varepsilon\eta}, -|u_{\varepsilon\eta}^-|^{s-2}u_{\varepsilon\eta}^-)) + (A_1(u_{\varepsilon\eta}, u_{\varepsilon\eta}), -|u_{\varepsilon\eta}^-|^{s-2}u_{\varepsilon\eta}^-) + \varepsilon^{-1}\|u_{\varepsilon\eta}^-\|_{L^{s+p-2}(G)}^{s+p-2} \leqslant \|f\|_{L^\infty(G)}\|u_{\varepsilon\eta}^-\|_{L^{s+p-2}(G)}^{s-1} .$$

$s$ is any number with $2 \leqslant s < \infty$.

It is not difficult to see that the first two expressions of the left hand side of the above inequality are nonnegative.
Hence:

$$\varepsilon^{-1}\|u_{\varepsilon\eta}^-\|_{L^{s+p-2}(G)}^{p-1} \leqslant \|f\|_{L^\infty(G)} .$$

Since $u_{\varepsilon\eta}^- \in L^\infty(G)$, we may let $s \to +\infty$ and get:

$$\varepsilon^{-1}\|u_{\varepsilon\eta}^-\|_{L^\infty(G)}^{p-1} \leqslant C .$$

$C$ is independent of both $\varepsilon$ and $\eta$.

(3) Let $\eta \to 0$ and the proof of Theorem 1 gives:

$$A(u_\varepsilon, u_\varepsilon) + \varepsilon^{-1}\beta(u_\varepsilon) = f .$$

Moreover:

$$\|u_\varepsilon\|_{L^\infty(G)} + \|u_\varepsilon\|_V + \varepsilon^{-1}\|u_\varepsilon^-\|_{L^\infty(G)}^{p-1} \leqslant C .$$

$u_\varepsilon$ is the weak limit in $V$ of $u_{\varepsilon\eta}$ as $\eta \to 0$.
Let $\varepsilon \to 0$ then again the proof of Theorem 1 gives:

$$(A(u, u) - f, v - u) \geqslant 0 \qquad \text{for all } v \text{ in } K\ V .$$

Since $\varepsilon^{-1}\beta(u_\varepsilon) \to g$ in the weak*-topology of $L^\infty(G)$, we have:

$$A(u, u) = f - g .$$

The theorem is proved.

With some further hypotheses on $A_j$, we have a regularity result.

THEOREM 5. *Let $V, K$ be as in Theorem 4 and let*

$$A(u, v) = -\sum_{j=1}^{n} D_j A_j(x, u, Dv) .$$

*Suppose that*:

(i) *$A_j(x, u, Dv)$ is continuously differentiable with respect to $x$, $u$ and $Dv$.*

(ii) $(1 + |Du|)(|A_j(x, u, Du)| + |\partial A_j/\partial u| + |A_{jk}(x, u, Du)|) + |\partial A_j/\partial x_k| \leqq C(1 + \exp(|u|) + |Du|^{p-1}\exp(|u|))$.

$$A_{jk}(x, u, Du) = \partial A_j(x, u, Du)/\partial(D_k u)\ .\quad j, k = 1, \cdots, n\ .$$

(iii) $\sum_{j,k=1}^n A_{jk}(x, u, Dv)D_j v D_k v \geqslant c\sum_{j=1}^n (1 + |D_j u|^{p-2})|D_j v|^2$.

*Then for any $f$ in $L^\infty(G)$, there exists $u$ in $K \cap V \cap L^\infty(G)$ such that:*

$$(A(u, u) - f, v - u) \geqslant 0 \qquad \text{for all } v \text{ in } K\ V\ .$$

*Moreover:*

$$u \in C^{0,\alpha}(G) \cap C^{1,\lambda}(G') \cap W^{2,p'}(G)\ .$$

$0 < \alpha, \gamma < 1$ *and $G'$ is any subset of $G$ with* $\mathrm{cl}\, G' \subset G$.

*Proof.* With the above hypotheses on $A_j$, $A(u, v)$ is a semi-monotone operator and satisfies all the hypotheses of Theorem 4. Therefore, there exists $u$ in $K \cap V \cap L^\infty(G)$ such that:

$$(A(u, u) - f, v - u) \geqslant 0 \qquad \text{for all } v \text{ in } V\ .$$

Moreover:

$$A(u, u) = f - g \in L^\infty(G)\ .$$

From the theory of second order elliptic equations, we get:

$$u \in C^{0,\alpha}(G)\ .$$

E.g. cf. [7] Theorem 1.1, p. 251.

It has been shown by Dubinskii [5] that the solution of $A(u, u) = f - g$ is in $W^{2,p'}(G)$ and moreover:

$$\sum_{j,k=1}^n \int_{G'} (1 + |\operatorname{grad} u|)^{p-2}|D_j D_k u|^2 dx < \infty\ .$$

The above relation together with the Hölder-continuity of $u$ gives:

$$\int_{G'} |\operatorname{grad} u|^{p+2} dx < \infty\ .$$

$G'$ is any subset of $G$ with $\mathrm{cl}\, G' \subset G$.

It follows from [7] that $\operatorname{ess\,sup}_{G'} |\operatorname{grad} u| < \infty$ and therefore Theorem 6.2 of [7], p. 282 gives:

$$u \in C^{1,\lambda}(G') \cap W^{2,2}(G')\ .$$

The theorem is proved.

If the principal part of $A$ is linear, stronger results could be obtained.

PROPOSITION 1. *Let $V$ and $K$ be as in Theorem 4 with $p = 2$. Let:*

$$A(u, v) = - \sum_{j,k=1}^{n} D_j(a_{jk}(x)D_k v) + u \exp(u) .$$

*Suppose that:*

(i) $a_{jk} \in C^2(\mathrm{cl}\, G)$.

(ii) $\sum_{j,k=1}^{n} a_{jk} D_j v D_k v \geqslant c \sum_{j=1}^{n} |D_j v|^2$ .

*Then for any $f$ in $L^\infty(G)$ there exists $u$ in $K \cap V \cap W^{2,s}(G)$ such that:*

$$(A(u, u) - f, v - u) \geqslant 0 \qquad \textit{for all } v \textit{ in } K\ V .$$

*$s$ is any number with $1 < s < \infty$.*

*Proof.* We may write

$$A(u, v) = \mathscr{A} v + u \exp(u) .$$

$A(u, v)$ satisfies all the hypotheses of Theorem 4. Thus there exists $u$ in $L^\infty(G) \cap K \cap V$ such that:

$$(A(u, u) - f, v - u) \geqslant 0 \qquad \text{for all } v \text{ in } K\ V .$$

Moreover:

$$\mathscr{A} u + u \exp(u) = g \in L^\infty(G) .$$

So:

$$\mathscr{A} u = g - u \exp(u) \in L^\infty(G) .$$

It follows from the theory of linear elliptic operators that $u \in W^{2,s}(G)$ for any $s$ with $1 < s < \infty$.

THEOREM 6. *Let $V = W_0^{1,p}(G)$, $K = \{u\colon u$ in $V$, $|\mathrm{grad}\, u| \leqslant 1$ a.e. on $G\}$ and let $A(u, v)$ be as in Theorem 4. Then for any $f$ in $L^\infty(G)$ and for any $\lambda > 0$, there exists $u$ in $K \cap L^\infty(G)$ such that:*

$$(A(u, u) + \lambda u - f, v - u) \geqslant 0 \qquad \textit{for all } v \textit{ in } K .$$

*Proof.* $K$ is a closed convex subset of both $V$ and $W$ with $0 \in K$. Let

$$\beta(v) = - \sum_{j=1}^{n} D_j\{(1 - |\mathrm{grad}\, v|^{p+n-2})^- D_j v\} .$$

$\beta$ maps bounded sets of $W = W_0^{1,p+n}(G)$ into bounded sets of $W^*$ and is a monotone hemi-continuous operator. Cf. Lions [9].
Moreover:

$$((\beta(v), v)) \geqslant ||v||_W^{p+n} - (c_1 + c_2 ||v||_V^{p+n-2}) .$$

It is not difficult to check that $\beta$ satisfies all the hypotheses of Theorem 2. Thus the result follows from Theorem 2.

Along the lines of Theorem 2, we may consider the case when

$$K = \{u: u \text{ in } W_0^{1,p}(G), u \geqslant 0 \text{ a.e. on } G, |\operatorname{grad} u| \leqslant 1 \text{ a.e. on } G\} .$$

It suffices to apply Theorem 2 with

$$\beta(v) = -|v^-|^{p-2}v^- - \sum_{j=1}^{n} D_j\{(1 - |\operatorname{grad} v|^{p+n-2})^- D_j v\} .$$

We shall now consider the case when constraint imposed on the solution is at the boundary of $G$.

THEOREM 7. *Let $V = W^{1,p}(G)$, $W = W^{1,p+n}(G)$, $K = \{u: u \text{ in } V, u \geqslant 0$ a.e. on $\partial G\}$ and let $A(u, v)$ be as in Theorem 5. Then for any $\lambda > 0$ and for any $f$ in $L^\infty(G)$ there exists $u \in C^{0,\alpha}(G) \cap C^{1,\mu}(G') \cap K$ such that:*

$$(A(u, u) - f, v - u) \geqslant 0 \qquad \text{for all } v \text{ in } K .$$

*$G'$ is any subset of $G$ with $\mathrm{cl}\, G' \subset G$ and $0 < \alpha, \mu < 1$.*

*Proof.* (1) Let $\beta$ be the penalty operator defined by:

$$(\beta(v), \varphi) = -\int_{\partial G} v^- \varphi \, dx , \qquad \varphi \text{ in } V .$$

$\beta$ satisfies all the hypotheses of Theorem 1. Cf. Lions [9]. Let $0 < \varepsilon < 1$, then from Theorem 1 we have:

$$A(u_\varepsilon, u_\varepsilon) + \lambda u_\varepsilon = f \quad \text{on} \quad G, \partial u_\varepsilon/\partial\gamma_A = -\varepsilon^{-1} u_\varepsilon^- \quad \text{on} \quad \partial G .$$

Moreover, from the proof of Theorem 1 we get:

$$||u_\varepsilon||_V + ||u_\varepsilon||_{L^\infty(G)} \leqslant C .$$

$C$ is a constant independent of $\varepsilon$.

(2) Let $\varepsilon \to 0$ then:

$$(A(u, u) + \lambda u - f, v - u) \geqslant 0 \qquad \text{for all } v \in K .$$

Moreover:

$$A(u, u) + \lambda u = f \quad \text{on} \quad G .$$

$u$ is the weak limit in $V$ of $u_\varepsilon$ as $\varepsilon \to 0$.

Thus:

$$A(u, u) = f - \lambda u \in L^\infty(G) .$$

Hence:

$$u \in C^{0,\alpha}(G) .$$

A standard argument of the theory of elliptic operators gives:

$$\sum_{j,k=1}^{n} \int_{G'} (1 + |\operatorname{grad} u|^{p-2}) |D_j D_k u|^2 dx < \infty .$$

$G'$ is any subset of $G$ with $\operatorname{cl} G' \subset G$.
The Hölder-continuity of $u$ together with the above relation implies that:

$$\int_{G'} |\operatorname{grad} u|^{p+2} dx < \infty .$$

Thus:

$$u \in C^{1,\mu}(G'') \quad \text{where} \quad \operatorname{cl} G'' \subset G' \subset \operatorname{cl} G' \subset G .$$

The theorem is proved.

## References

1. H. Brezis, *Equations et inéquations dans les espaces vectoriels en dualité*, Ann. Inst. Fourier, **18** (1968), 115-175.
2. ———, *Problèmes unilatéraux*, J. Math. Pure Appl., **51** (1972), 1-168.
3. H. Brezis and G. Stampacchia, *Sur la régularité de la solution d'inéquations elliptiques*, Bull. Soc. Math. France, **96** (1968), 153-180.
4. F. E. Browder, *Nonlinear monotone operators and convex sets in Banach spaces*, Bull. Amer. Math. Soc., **71** (1965), 780-785.
5. Ju. A. Dubinskii, *Weak convergence in nonlinear elliptic and parabolic equations*, Mat. Sb., **67** (109), (1965), 609-642 = Amer. Math. Soc. Transl., **2** (67), (1968), 226-258.
6. P. Hartman and G. Stampacchia. *On some nonlinear elliptic differential functional equations*, Acta Math., **115** (1966), 271-310.
7. O. A. Ladyzenskaya and N. N. Uraltseva, *Linear and Quasilinear Elliptic Equations*, Acad. Press, 1968.
8. H. Lewy and G. Stampacchia, *On the regularity of the solution of a variational inequality*, Comm. Pure Appl. Math., **22** (1969), 153-188.
9. J. L. Lions, *Quelques méthodes de résolution des problemes aux limites non linéaires*, Dund. Paris 1969.
10. J. L. Lions and G. Stampacchia, *Variational inequalities*, Comm. Pure Appl. Math., **20** (1967), 493-519.

Received July 25, 1972.

UNIVERSITY OF BRITISH COLUMBIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# A TOPOLOGICAL CHARACTERIZATION OF COMPLETE, DISCRETELY VALUED FIELDS

SETH WARNER

**It is shown that the topology of a topological field $F$ is given by a complete, discrete valuation if and only if $F$ is locally strictly linearly compact. More generally, the topology of a topological division ring $K$ is given by a complete, discrete valuation and $K$ is finite dimensional over its center if and only if $K$ is locally centrally linearly compact, that is, if and only if $K$ contains an open subring $B$, the open left ideals of which form a fundamental system of neighborhoods of zero, such that $B$, regarded as a module over its center, is strictly linearly compact.**

In [5], Jacobson showed that the topology of an indiscrete, totally disconnected, locally compact division ring is given by a discrete valuation (that is, a valuation whose value group is isomorphic to the cyclic group of integers). Consequently, an indiscrete topological division ring $K$ is locally compact and totally disconnected if and only if its topology is given by a complete, discrete valuation whose residue field is finite [4, Prop. 2, p. 118, Prop. 1, p. 156]. From this, it follows rather readily that the center $C$ of $K$ is indiscrete, that $K$ is finite dimensional over $C$, and that $C$ is either a finite extension of the $p$-adic number field for some prime $p$ or the field of formal power series over a finite field [4, Theorem 1, p. 158].

Our purpose here is to generalize Jacobson's theorem by characterizing those topological fields whose topology is given by a complete, discrete valuation, and more generally, those topological division rings $K$ such that $K$ is finite dimensional over its center and the topology of $K$ is given by a complete, discrete valuation.

For this purpose, we assume some familiarity with basic properties of linearly compact and strictly linearly compact modules and rings, as developed in [10] or [3, Exercises 14-22, pp. 108-112]. We recall that a (left) topological $A$-module $E$ (it is not assumed that $E$ is unitary) is *linearly topologized* if the open submodules of $E$ form a fundamental system of neighborhoods of zero; $E$ is *linearly compact* if $E$ is Hausdorff, linearly topologized, and every filter base of cosets of submodules has an adherent point; $E$ is *strictly linearly compact* if $E$ is linearly compact and every continuous epimorphism from $E$ onto a Hausdorff, linearly topologized $A$-module is open (equivalently, if $E/U$ is an artinian $A$-module for every open submodule $U$ of $E$). A topological ring $A$ is respectively linearly topologized, linearly compact, or strictly linearly compact if the associated left $A$-module $A$ is.

DEFINITION. A topological ring $A$ is *locally strictly linearly compact* if $A$ has an open subring $B$ that is strictly linearly compact for its induced topology.

To handle the noncommutative case, we need the following definition:

DEFINITION. A topological ring $B$ is *centrally linearly compact* if the open left ideals of $B$ form a fundamental system of neighborhoods of zero and if $B$, regarded as a module over its center $C_B$, is a strictly linearly compact $C_B$-module. A topological ring $A$ is *locally centrally linearly compact* if $A$ contains an open subring that is centrally linearly compact for its induced topology.

Thus a commutative topological ring is (locally) centrally linearly compact if and only if it is (locally) strictly linearly compact. Note that if $B$ is a centrally linearly compact ring, then for any subring $B_0$ of $B$ that contains the center $C_B$, $B$ is a strictly linearly compact $B_0$-module (in particular, $B$ is a strictly linearly compact ring); indeed, since the open left ideals of $B$ form a fundamental system of neighborhoods of zero, $B$ is a linearly topologized $B_0$-module, and since a $B_0$-submodule is also a $C_B$-submodule, every filter base of cosets of $B_0$-submodules necessarily has an adherent point.

By a topological division ring (field) $K$ we mean a topological ring that is algebraically a division ring (field); that is, we do not assume that $x \mapsto x^{-1}$ is continuous on the set $K^*$ of nonzero elements.

LEMMA 1. *If $B$ is an open, centrally linearly compact subring of an indiscrete topological division ring $K$, then there is an open, centrally linearly compact subring $B_1$ of $K$ that contains* 1.

*Proof.* Let $B_1$ be the closure of the subring generated by $B$ and 1. The open left ideals of $B$ then form a fundamental system of neighborhoods of zero in $B_1$; each open left ideal $\mathfrak{a}$ of $B$ is a left ideal of $B_1$, for as $\mathfrak{a}$ is closed, $\{x \in B_1 : x\mathfrak{a} \subseteqq \mathfrak{a}\}$ is a closed subring of $B_1$ containing $B$ and 1 and hence is all of $B_1$.

Since $B$ is open, $B \neq (0)$; let $b$ be some nonzero element of $B$, and let $c$ be its inverse in $K$. Then, $B_1 = B_1bc \subseteqq B_1Bc$, so $B_1 \subseteqq Bc$ since, as we saw above, $B$ is a left ideal of $B_1$. Thus $Bc \supseteqq B_1 \supseteqq B$, so $Bc$ is a linearly topologized $C_B$-module, where $C_B$ is the center of $B$. Hence $Bc$ is a strictly linearly compact $C_B$-module, as it is the image of the strictly linearly compact $C_B$-module $B$ under the continuous homomorphism $x \mapsto xc$. Consequently, the closed $C_B$-submodule

$B_1$ of $Bc$ is strictly compact; as $C_B$ is contained in the center of $B_1$, $B_1$ is *a fortiori* strictly linearly compact over its center.

We recall that an element $a$ of a topological ring is *topologically nilpotent* if $\lim a^n = 0$.

LEMMA 2. *Let $K$ be a Hausdorff topological division ring, let $B$ be an open subring of $K$ that contains* 1, *and let $\mathfrak{r}$ be the radical of $B$. If $B$ is strictly linearly compact, then $B$ is a (left) noetherian ring, $B/\mathfrak{r}$ is a division ring, the topology of $B$ is the $\mathfrak{r}$-adic topology, and $\mathfrak{r}$ is the set of all topological nilpotents of $B$.*

*Proof.* As $B$ is open and as $y \mapsto yx$ is a homeomorphism for each $x \in K^*$, $Bx$ is open for every $x \in K^*$, and hence every nonzero left ideal of $B$ is open. Let $\mathfrak{s} = \bigcap_{n=1}^{\infty} \mathfrak{r}^n$. Assume that $\mathfrak{s} \neq (0)$. Then $\mathfrak{s}$ is open, so $B/\mathfrak{s}$ is an artinian $B$-module and hence an artinian ring. Consequently, its radical $\mathfrak{r}/\mathfrak{s}$ is nilpotent, so there exists $n$ such that $\mathfrak{r}^n = \mathfrak{s}$. Hence $(0) \neq \mathfrak{r}^n = \mathfrak{r}^{n+1} = \cdots$, in contradiction to [10, Theorem 9]. Therefore, $\bigcap_{n=1}^{\infty} \mathfrak{r}^n = (0)$.

Since every nonzero left ideal of $B$ is open and hence closed, $B$ is a (left) noetherian ring, $B/\mathfrak{r}$ is an artinian ring, and the topology of $B$ is its $\mathfrak{r}$-adic topology by [13, Theorem 16]. Consequently, every element of $\mathfrak{r}$ is a topological nilpotent. Therefore, as $B$ is complete, $B$ is suitable for building idempotents [11, Lemma 4; 6, Definition 1, p. 53]. Thus every idempotent of $B/\mathfrak{r}$ is the coset of $\mathfrak{r}$ determined by an idempotent of $B$ [6, Proposition 4, p. 54]. But as $K$ is a division ring, $B$ has no idempotents other than 0 and 1. Thus $B/\mathfrak{r}$ is an artinian, semisimple ring whose only idempotents are 0 and 1. By the Wedderburn-Artin theorem, therefore, $B/\mathfrak{r}$ is a division ring. In particular, if $x \notin \mathfrak{r}$, then $x + \mathfrak{r}$ is not a nilpotent of $B/\mathfrak{r}$, so $x$ is not a topological nilpotent.

THEOREM 1. *If $K$ is an indiscrete, Hausdorff topological field, then the topology of $K$ is given by a complete, discrete valuation if and only if $K$ is locally strictly linearly compact.*

*Proof. Necessity.* It is well known that a complete, semilocal noetherian ring, equipped with its natural $\mathfrak{r}$-adic topology, is strictly linearly compact [cf. 13, Corollary of Lemma 2]. In particular, the valuation ring of a complete discrete valuation is strictly linearly compact.

*Sufficiency.* By Lemma 1, there is an open, strictly linearly compact subring $B$ of $K$ that contains 1. By Lemma 2, $B$ is a complete, local noetherian domain, and its topology is its natural $\mathfrak{m}$-adic topology, where $\mathfrak{m}$ is the maximal ideal of $B$. In particular, $B$ is not a field since $B$ is not discrete. Therefore, as $B$ is open in the topological field $K$, the topology of $K$ is defined by a complete, discrete valuation [12, Theorem 6].

THEOREM 2. *If $K$ is an indiscrete, Hausdorff topological division ring, then the topology of $K$ is given by a complete, discrete valuation and $K$ is finite-dimensional over its center $C$ if and only if $K$ is locally centrally linearly compact; in this case, $C$ is indiscrete, and hence its topology is given by a complete, discrete valuation.*

*Proof. Necessity.* As $K$ is finite-dimensional over $C$, the valuation induced on $C$ by that of $K$ is not the improper valuation; hence as $C$ is closed, the topology of $C$ is given by a complete, discrete valuation $v$. Let $e_1, \cdots, e_n$ be a basis of $K$ over $C$ such that $e_1 = 1$, and let $e_i e_j = \sum_{k=1}^{n} \alpha_{ijk} e_k$. Let $\lambda \in C$ be such that $v(\lambda) \geqq 0$ and $v(\lambda) \geqq -\min\{v(\alpha_{ijk}): 1 \leqq i, j, k \leqq n\}$. Let $f_1 = 1$ and $f_k = \lambda e_k$ for $2 \leqq k \leqq n$. Let $V$ be the valuation ring of $C$, and for each $m \geqq 0$ let $V_m = \{x \in V: v(x) \geqq m\}$. Let $B = Vf_1 + \cdots + Vf_n$, and for each $m \geqq 0$ let $\mathfrak{b}_m = V_m f_1 + \cdots + V_m f_n$. Easy calculations establish that $B$ is a ring and that $\mathfrak{b}_m$ is an ideal of $B$ for each $m \geqq 0$. By [2, Theorem 2, p. 18], $F: (\lambda_1, \cdots, \lambda_n) \mapsto \sum_{i=1}^{n} \lambda_i f_i$ is a topological isomorphism from the $C$-vector space $C^n$ onto the $C$-vector space $K$. Hence $B$ is an open subring of $K$, and $(\mathfrak{b}_m)_{m \geqq 0}$ is a fundamental system of neighborhoods of zero in $B$, each an ideal of $B$. We saw earlier that $V$ is strictly linearly compact; hence as $B = F(V^n)$, $B$ is a strictly linearly compact $V$-module and, *a fortiori*, is a centrally linearly compact ring.

*Sufficiency.* By Lemma 1, there is an open, centrally linearly compact subring $B$ that contains 1. Let $\mathfrak{r}$ be the radical of $B$. As the $\mathfrak{r}$-adic topology is the given indiscrete topology of $B$ by Lemma 2, there exists a nonzero $a \in B$ such that $\lim a^n = 0$. Let $K_0$ be the closed subfield generated by $C$ and $a$, let $B_0 = K_0 \cap B$, and let $\mathfrak{r}_0$ be the radical of $B_0$. Since the open left ideals of $B$ form a fundamental system of neighborhoods of zero for $B$, the open ideals of $B_0$ form a fundamental system of neighborhoods of zero for $B_0$. Moreover, the center $C_B$ of $B$ is simply $C \cap B$; indeed, if $c \in C_B$ and if $x \in K$, then $a^n x \in B$ for some $n$ as $\lim a^n x = 0$, whence $(a^n x)c = c(a^n x) = (ca^n)x = (a^n c)x$, so $xc = cx$. Thus $C_B = C \cap B \subseteqq K_0 \cap B = B_0$, so $B_0$ is a closed $C_B$-submodule of $B$ and hence is a strictly linearly compact $C_B$-module. Consequently, $B_0$ is a strictly linearly compact ring, so by Lemma 2, the topology of $B_0$ is the $\mathfrak{r}_0$-adic topology, and $\mathfrak{r}$ and $\mathfrak{r}_0$ are respectively the set of topological nilpotents in $B$ and $B_0$, whence $\mathfrak{r}_0 = \mathfrak{r} \cap B_0$. Hence $\bigcap_{n=1}^{\infty} (\mathfrak{r}_0^n B)^- \subseteqq \bigcap_{n=1}^{\infty} \mathfrak{r}^n = (0)$. As the topology of $B_0$ is indiscrete, $\mathfrak{r}_0^2 \neq (0)$, so $\mathfrak{r}_0^2 B$ is open as it contains a nonzero left ideal of $B$. By

[13, Theorem 10], $\mathfrak{x}_0 B$ is a finitely generated $B_0$-module; let $\mathfrak{x}_0 B = B_0 x_1 + \cdots + B_0 x_m$. Also as $B$ is a strictly linearly compact $C_B$-module and as $\mathfrak{x}_0 B$ is open, $B/\mathfrak{x}_0 B$ is an artinian $C_B$-module, hence an artinian $B_0$-module; now $B/\mathfrak{x}_0 B$ admits the structure of $B_0/\mathfrak{x}_0$-module, and $B_0/\mathfrak{x}_0$ is a field by Lemma 2; consequently $B/\mathfrak{x}_0 B$ is an artinian, therefore, finite-dimensional, and hence noetherian $B_0/\mathfrak{x}_0$-vector space; thus $B/\mathfrak{x}_0 B$ is a noetherian $B_0$-module. Let $x_{m+1}, \cdots, x_n \in B$ be such that $B = B_0 x_{m+1} + \cdots + B_0 x_n + \mathfrak{x}_0 B$. Then $B = B_0 x_1 + \cdots + B_0 x_n$. Consequently, $x_1, \cdots, x_n$ is a set of generators of the $K_0$-vector space $K$, for if $z \in K$, there exists $t$ such that $a^t z \in B$, whence $a^t z = b_1 x_1 + \cdots + b_n x_n$ where $b_i \in B_0$, and thus $z = (a^{-t} b_1) x_1 + \cdots + (a^{-t} b_n) x_n \in K_0 x_1 + \cdots + K_0 x_n$. By [1, Theorem 16], the centralizer $K_0'$ of $K_0$ has degree $\leqq n$ over $C$. But $K_0' \supseteqq K_0$ as $K_0$ is commutative. Moreover, the topology of $K_0$ is given by a discrete valuation by Theorem 1, as $B_0$ is an open, strictly linearly compact subring. Therefore, as $[K_0: C] \leqq n$, the valuation induced on $C$ is not the improper valuation; hence the topology of $C$ is given by a complete, discrete valuation. As

$$[K: C] = [K: K_0][K_0: C] \leqq n^2 ,$$

the given topology of $K$ is the only topology for which $K$ is a Hausdorff topological vector space over $C$ [2, Theorem 2, p. 18]; by valuation theory, that topology is given by a complete, discrete valuation.

The idea of using [1, Theorem 16] is suggested by Kaplansky's treatment of locally compact division rings in [8].

Jacobson's theorem concerning totally disconnected locally compact division rings follows at once from Theorem 2. Indeed, if $K$ is an indiscrete, totally disconnected, locally compact division ring, then $K$ contains a compact open subring $B$ [9, Lemma 4]; the open ideals of $B$ form a fundamental system of neighborhoods of zero [7, Lemmas 9 and 10], and therefore the compact ring $B$ is clearly centrally linearly compact; by Theorem 2, $K$ is finite-dimensional over its center, which is indiscrete, and the topology of $K$ is given by a complete, discrete valuation.

## REFERENCES


1. E. Artin and G. Whaples, *The theory of simple rings*, Amer. J. Math., **65** (1943), 87-107.
2. N. Bourbaki, *Espaces vectoriels topologiques*, Ch. 1-2, 2nd ed., Hermann, Paris, 1966.
3. ———, *Algèbre commutative*, Ch. 3-4, Hermann, Paris, 1961.
4. ———, *Algèbre commutative*, Ch. 5-6, Hermann, Paris, 1964.
5. N. Jacobson, *Totally disconnected locally compact rings*, Amer. J. Math., **58** (1936), 433-449.
6. ———, *Structure of Rings*, Amer. Math. Soc. Colloq. Publ., vol. 37, Providence, R. I., 1956.
7. Irving Kaplansky, *Topological rings*, Amer. J. Math., **69** (1947), 153-183.

8. Irving Kaplansky, *Topological methods in valuation theory*, Duke Math. J., **14** (1947), 527–541.
9. ———, *Locally compact rings*, Amer. J. Math., **70** (1948), 447–459.
10. Horst Leptin, *Linear kompakte Moduln und Ringe*, Math. Z., **62** (1955), 241–267.
11. Seth Warner, *Compact rings*, Math. Ann., **145** (1962), 52–63.
12. ———, *Openly embedding local noetherian domains*, J. Reine Angew. Math., **253** (1972), 146–151.
13. ———, *Linearly compact rings and modules*, Math. Ann., **197** (1972), 29–43.

Received June 12, 1972.

DUKE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

# COMMON FIXED POINTS OF TWO MAPPINGS

CHI SONG WONG

**Let $S$, $T$ be functions on a nonempty complete metric space $(X, d)$. The main result of this paper is the following. $S$ or $T$ has a fixed point if there exist decreasing functions $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ of $(0, \infty)$ into $[0, 1)$ such that (a) $\sum_{i=1}^{5} \alpha_i < 1$; (b) $\alpha_1 = \alpha_2$ or $\alpha_3 = \alpha_4$, (c) $\lim_{t \downarrow 0} (\alpha_1 + \alpha_2) < 1$ and $\lim_{t \downarrow 0} (\alpha_3 + \alpha_4) < 1$ and (d) for any distinct $x$, $y$ in $X$,**

$$d(S(x), T(y)) \leqq a_1 d(x, S(x)) + a_2 d(y, T(y)) + a_3 d(x, T(y)) + a_4 d(y, S(x)) + a_5 d(x, y) ,$$

**where $a_i = \alpha_i(d(x, y))$. A number of related results are obtained.**

1. Introduction. Let $(X, d)$ be a nonempty complete metric space and let $S$, $T$ be mappings of $X$ into itself which are not necessarily continuous nor commuting. Suppose that there are nonnegative real numbers $a_1, a_2, a_3, a_4, a_5$ such that

$$\text{(a)} \qquad a_1 + a_2 + a_3 + a_4 + a_5 < 1 ,$$

$$\text{(b)} \qquad a_1 = a_2 \quad \text{or} \quad a_3 = a_4 ,$$

and for any $x$, $y$ in $X$,

$$\text{(c)} \qquad d(S(x), T(y)) \leqq a_1 d(x, S(x)) + a_2 d(y, T(y)) + a_3 d(x, T(y)) + a_4 d(y, S(x)) + a_5 d(x, y) .$$

It is proved in this paper that each of $S$, $T$ has a unique fixed point and these two fixed points coincide. Among others, a generalization is obtained by replacing $a_1, a_2, a_3, a_4, a_5$ with nonnegative real-valued functions on $(0, \infty)$. This result generalizes the Banach contraction mapping theorem and some results of G. Hardy and T. Rogers [5], R. Kannan [7], E. Rakotch [8], S. Reich [9], P. Srivastava, and V. K. Gupta [10]. It also gives a different proof for these special cases. Note that even if $X = [0, 1]$ and if $T_1$, $T_2$ are commuting continuous functions of $X$ into itself, $T_1$, $T_2$ need not have a common fixed point [1], [2], and [6].

2. Basic results.

THEOREM 1. *Let $S$, $T$ be mappings of a complete metric space $(X, d)$ into itself. Suppose that there exist nonnegative real numbers $a_1, a_2, a_3, a_4, a_5$ which satisfy* (a), (b), *and* (c). *Then each of $S$, $T$*

*has a unique fixed point and these two fixed points coincide.*

*Proof.* Let $x_0 \in X$. Define

$$x_{2n+1} = S(x_{2n}),\ x_{2n+2} = T(x_{2n+1}) , \qquad n = 0, 1, 2, \cdots .$$

From (c),

$$\begin{aligned} d(x_1, x_2) &= d(S(x_0),\ T(x_1)) \\ &\leqq (a_1 + a_5)d(x_0, x_1) + a_2 d(x_1, x_2) + a_3 d(x_0, x_2) \\ &\leqq (a_1 + a_5)d(x_0, x_1) + a_2 d(x_1, x_2) + a_3(d(x_0, x_1) + d(x_1, x_2)) . \end{aligned}$$

So

$$(1) \qquad d(x_1, x_2) \leqq \frac{a_1 + a_3 + a_5}{1 - a_2 - a_3} d(x_0, x_1) .$$

Similarly,

$$(2) \qquad d(x_2, x_3) \leqq \frac{a_2 + a_4 + a_5}{1 - a_1 - a_4} d(x_1, x_2) .$$

Let

$$r = \frac{a_1 + a_3 + a_5}{1 - a_2 - a_3} , \qquad s = \frac{a_2 + a_4 + a_5}{1 - a_1 - a_4} .$$

Repeating the above argument, we obtain, for each $n = 0, 1, 2, \cdots$,

$$(3) \qquad d(x_{2n+1}, x_{2n+2}) \leqq rd(x_{2n+1}, x_{2n}) ,$$

$$(4) \qquad d(x_{2n+3}, x_{2n+2}) \leqq sd(x_{2n+2}, x_{2n+1}) .$$

By (3), (4), and induction, we have, for each $n = 0, 1, 2, \cdots$,

$$(5) \qquad d(x_{2n+1}, x_{2n+2}) \leqq r(rs)^n d(x_0, x_1) ,$$

$$(6) \qquad d(x_{2n+2}, x_{2n+3}) \leqq (rs)^{n+1} d(x_0, x_1) .$$

Since $rs < 1$ and

$$\sum_{n=0}^{\infty} d(x_n, x_{n+1}) \leqq (1 + r) \sum_{n=0}^{\infty} (rs)^n d(x_0, x_1) ,$$

$\{x_n\}$ is Cauchy. By completeness of $(X, d)$, $\{x_n\}$ converges to some point $x$ in $X$. We shall now prove that $x$ is a fixed point of $S$ and $T$. Let $n$ be given. Then

$$(7) \qquad \begin{aligned} d(x, S(x)) &\leqq d(x, x_{2n+2}) + d(S(x), x_{2n+2}) \\ &= d(x, x_{2n+2}) + d(S(x), T(x_{2n+1})) . \end{aligned}$$

By (c),

$$d(S(x), T(x_{2n+1})) \leqq a_1 d(x, S(x)) + a_2 d(x_{2n+1}, x_{2n+2}) + a_3 d(x, x_{2n+2}) \\ + a_4 d(x_{2n+1}, S(x)) + a_5 d(x, x_{2n+1}) . \tag{8}$$

Combining (7) and (8) and letting $n$ tend to infinity, we obtain

$$d(x, S(x)) \leqq (a_1 + a_4) d(x, S(x)) .$$

Since $a_1 + a_4 < 1$, $S(x) = x$. Similarly $T(x) = x$. Let $y$ be a fixed point of $T$. Then from $d(x, y) = d(S(x), T(y))$ and (c), we obtain

$$d(x, y) \leqq (a_3 + a_4 + a_5) d(x, y) .$$

Since $a_3 + a_4 + a_5 < 1$, $d(x, y) = 0$. So $T$ has a unique fixed point. Similarly, $S$ has a unque fixed point.

When $a_3 = a_4 = a_5 = 0$, $S = T$ and $T$ is continuous (or even $x \rightarrow d(x, T(x))$ is lower semicontinuous) on $X$, Theorem 1 can be obtained by an earlier result of the author [11, Theorem 1].

From the proof of Theorem 1, we know that $S$, $T$ still have a common fixed point if conditions (a), (b) are replaced by the following conditions:

$$(a_1 + a_3 + a_5)(a_2 + a_4 + a_5) < (1 - a_2 - a_3)(1 - a_1 - a_4) , \tag{9}$$

$$a_1 + a_4 < 1 . \tag{10}$$

If in addition,

$$a_3 + a_4 + a_5 < 1 , \tag{11}$$

then the common fixed point of $S$, $T$ is the unique fixed point of $S$ (and $T$). Note that conditions (a) and (b) imply (9), but (a) alone does not. Indeed, for any $a_1$, $a_2$, $a_5$ in $[0, \infty)$ with $a_1 \neq a_2$ and $a_1 + a_2 + a_5 < 1$, we can always find $a_3$, $a_4$ in $[0, \infty)$ such that (a) holds but (9) does not. This can be seen by considering the affine function $f$:

$$f(x, y) = (1 - a_2 - x)(1 - a_1 - y) - (a_1 + x + a_5)(a_2 + y + a_5)$$

defined on the compact convex set

$$K = \{(x, y) \in [0, 1] \times [0, 1] \colon a_1 + a_2 + x + y + a_5 \leqq 1\} .$$

$f$ takes its minimum value at one of the extreme points of $K$. With some computation, we conclude that

$$\min f(K) = -\,|a_1 - a_2|\,(1 - a_1 - a_2 - a_5) .$$

Since $a_1 + a_2 + a_5 > 1$, $\min f(K) < 0$ if and only if $a_1 \neq a_2$. Thus if $a_1 \neq a_2$, then by continuity of $f$, there exists a point $(a_3, a_4)$ in

$$K \backslash \{(x, y) \in K\colon a_1 + a_2 + x + y + a_5 = 1\}$$

such that $f(a_3, a_4) < 0$.

COROLLARY 1. R. Kannan [7, Theorem 1]. *Let $S$ be a mapping of a complete metric space $(X, d)$ into itself. Suppose that there exists a number $r$ in $[0, 1/2)$ such that*

$$d(S(x), S(y)) \leqq r(d(x + S(x)) + d(y, S(y)))$$

*for all $x$, $y$ in $X$. Then $S$ has a unique fixed point.*

COROLLARY 2. P. Srivastava and V. K. Gupta [10, Theorem 1]. *Let $S$, $T$ be mappings of a complete metric space $(X, d)$ into itself. Suppose that there exists nonnegative real numbers $a_1$, $a_2$ such that*

$$\text{(a)} \qquad a_1 + a_2 < 1$$

and

$$\text{(b)} \qquad d(S(x), T(y)) \leqq a_1 d(x, S(x)) + a_2 d(y, T(y)) \quad \textit{for all } x, y \textit{ in } X.$$

*Then $S$, $T$ have a unique common fixed point.*

Srivastava and Gupta stated the above result in a more general form with $S$, $T$ replaced by $S^p$, $T^q$ for some positive integers $p$, $q$. Since the unique fixed point of $S^p$ (similarly $T^q$) is the unique fixed point of $S$, this result is equivalent to Corollary 2.

For Corollaries 1 and 2, we have the following related result.

PROPOSITION. *Let $S$, $T$ be self-maps of a nonempty complete metric space $(X, d)$. Suppose that there exist nonnegative real numbers $a_1$, $a_2$ such that $a_1 + a_2 < 1$ and*

$$(*) \qquad d(S(x), T(y)) \leqq a_1 d(x, S(x)) + a_2 d(y, T(y)), \quad x, y \in X.$$

*Then either (*) is true when all of its $S$ are replaced by $T$ or (*) is true when all of its $T$ are replaced by $S$.*

The following example proves that our result is actually more general than that of Srivastava and Gupta.

EXAMPLE. Let $X = \{1, 2, 3\}$. Let $d$ be the metric for $X$ determined by

$$d(1, 2) = 1,\ d(2, 3) = \frac{4}{7}, \quad d(1, 3) = \frac{5}{7}.$$

Let $S$, $T$ be the function on $X$ such that

$$S(1) = S(2) = S(3) = 1;$$

$$T(1) = T(3) = 1 , \quad T(2) = 3 .$$

Let $a_1 = 0$, $a_2 = 0$, $a_3 = 0$, $a_4 = 5/7$, $a_5 = 0$. Then the conditions of Theorem 1 are satisfied. However, no nonnegative real numbers $a_1$, $a_2$, $a_3$, $a_5$ can be chosen such that $a_1 + a_2 + a_3 + a_5 < 1$ and for $x, y \in X$,

$$d(S(x), T(y)) \leqq a_1 d(x, S(x)) + a_2 d(y, T(y)) + a_3 d(x, T(y)) + a_5 d(x, y) .$$

For if there exist such $a_1$, $a_2$, $a_3$, $a_5$, then

$$d(S(3), T(2)) \leqq a_1 d(3, S(3)) + a_2 d(2, T(2)) + a_3 d(3, T(2)) + a_5 d(3, 2) .$$

So

$$\frac{5}{7} \leqq \frac{5a_1}{7} + \frac{4a_2}{7} + \frac{4a_5}{7} \leqq \frac{5}{7}(a_1 + a_2 + a_5) < \frac{5}{7} ,$$

a contradiction.

COROLLARY 3. G. Hardy and T. Rogers [5, Theorem 1]. *Let $S$ be a mapping of a nonempty complete metric space $(X, d)$ into itself. Suppose that there exist nonnegative real numbers $a_1$, $a_2$, $a_3$, $a_4$, $a_5$ such that*

(a) $$a_1 + a_2 + a_3 + a_4 + a_5 < 1$$

*and*

(b) $$d(S(x), S(y)) \leqq a_1 d(x, S(x) + a_2 d(y, S(y)) + a_3 d(x, S(y)) + a_4 d(y, S(x)) + a_5 d(x, y) \quad \text{for all } x, y \text{ in } X .$$

*Then $S$ has a unique fixed point.*

Note that in the above case, we may without loss of generality assume that $a_1 = a_2$, $a_3 = a_4$ (replace $a_1$, $a_2$, $a_3$, $a_4$, $a_5$ respectively by

$$\frac{a_1 + a_2}{2}, \frac{a_1 + a_2}{2}, \frac{a_3 + a_4}{2}, \frac{a_3 + a_4}{2}, a_5$$

if necessary). So the above result follows from Theorem 1. The above example shows that there is no such symmetry ($a_1 = a_2$, $a_3 = a_4$) for the general case. Indeed, we cannot even assume $a_3 = a_4$. For if $a_3 = a_4$, then for the above example, we have

$$\begin{aligned}\frac{5}{7} = d(S(3),\, T(3)) &\leqq \frac{5}{7} a_1 + \frac{4}{7} a_2 + a_4 + \frac{4}{7} a_5 \,. \\ &= \frac{5}{7} a_1 + \frac{4}{7} a_2 + \frac{1}{2} a_3 + \frac{1}{2} a_4 + \frac{4}{7} a_5 \\ &< \frac{5}{7}(a_1 + a_2 + a_3 + a_4 + a_5) < \frac{5}{7}\,,\end{aligned}$$

a contradiction.

**2. Extensions and some ralated results.** The following result generalizes Theorem 1. Its proof is different from the one we gave for Theorem 1.

THEOREM 2. *Let $S$, $T$ be functions on a nonempty complete metric space $(X, d)$. Suppose that there exist decreasing functions $\alpha_1$, $\alpha_2$, $\alpha_3$, $\alpha_4$, $\alpha_5$ of $(0, \infty)$ into $[0, 1)$ such that*

(a) $\sum_{i=1}^{5} \alpha_i < 1$;

(b) $\alpha_1 = \alpha_2$ *or* $\alpha_3 = \alpha_4$;

(c) $\lim_{t\downarrow 0}(\alpha_2 + \alpha_3) < 1$ and $\lim_{t\downarrow 0}(\alpha_1 + \alpha_4) < 1$;

(d) *for any distinct $x$, $y$ in $X$,*

$$\begin{aligned}d(S(x),\, T(y)) \leqq a_1 d(x, S(x)) &+ a_2 d(y, T(y)) + a_3 d(x, T(y)) \\ &+ a_4 d(y, S(x)) + a_5 d(x, y)\,,\end{aligned}$$

*where $a_i = \alpha_i(d(x, y))$.*

*Then at least one of $S$, $T$ has a fixed point. If both $S$ and $T$ have fixed points, then each of $S$, $T$ has a unique fixed point and these two fixed points coincide.*

*Proof.* Let $x_0 \in X$. Define for each $n = 0, 1, 2, \cdots$,

$$x_{2n+1} = S(x_{2n})\,, \quad x_{2n+2} = T(x_{2n+1})\,, \quad b_n = d(x_n, x_{n+1})\,.$$

We may assume that $b_n > 0$ for each $n$, for otherwise some $x_n$ is a fixed point of $S$ or $T$. Let

$$r(t) = \frac{\alpha_1(t) + \alpha_3(t) + \alpha_5(t)}{1 - \alpha_2(t) - \alpha_3(t)}\,, \qquad t > 0\,,$$

$$s(t) = \frac{\alpha_2(t) + \alpha_4(t) + \alpha_5(t)}{1 - \alpha_1(t) - \alpha_4(t)}\,, \qquad t > 0\,.$$

Then $r$, $s$ are decreasing. From (a) and (c), the limits

$$r_0 = \lim_{t\downarrow 0} r(t)\,, \quad s_0 = \lim_{t\downarrow 0} s(t)$$

are nonnegative real numbers. Let

$$f(t) = r(t)s(t) , \qquad t > 0 .$$

Then $f$ is decreasing and $f(t) < 1$ for each $t > 0$. As in the proof of Theorem 1, we have for each $n = 0, 1, 2, \cdots$,

$$b_{2n+1} \leqq r(b_{2n})b_{2n} , \tag{12}$$

$$b_{2n+2} \leqq s(b_{2n+1})b_{2n+1} . \tag{13}$$

Let $n$ be given. Then

$$b_{2n+3} \leqq r(b_{2n+2})\, s(b_{2n+1})b_{2n+1} , \tag{14}$$

$$b_{2n+2} \leqq s(b_{2n+1})\, r(b_{2n})b_{2n} . \tag{15}$$

Since $r$, $s$ are decreasing,

$$b_{2n+3} \leqq f(\min\{b_{2n+2}, b_{2n+1}\})b_{2n+1} , \tag{16}$$

$$b_{2n+2} \leqq f(\min\{b_{2n+1}, b_{2n}\})b_{2n} . \tag{17}$$

Since $f(t) < 1$ for each $t > 0$, $\{b_{2n+1}\}$, $\{b_{2n}\}$ are decreasing sequences. So $\{b_{2n+1}\}$, $\{b_{2n}\}$ converge respectively to some points $c_1$, $c_2$. We shall prove that $c_1 = 0$, $c_2 = 0$. From (12) and (13),

$$c_1 \leqq r_0 c_2 , \quad c_2 \leqq s_0 c_1 .$$

So either both $c_1$, $c_2$ are zero or both $c_1$, $c_2$ are not zero. Suppose to the contrary that $c_1 \neq 0$, $c_2 \neq 0$. Then from (16) and (17),

$$b_{n+2} \leqq f(\min\{c_1, c_2\})b_n , \qquad n = 0, 1, 2, \cdots . \tag{18}$$

By induction,

$$b_{2n} \leqq (f(\min\{c_1, c_2\}))^n b_0 \qquad n = 0, 1, 2, \cdots . \tag{19}$$

So $c_2 = 0$, a contradiction. Therefore, $c_1 = c_2 = 0$. This proves that $\{b_n\}$ converges to 0.

Now we shall prove that $\{x_n\}$ is Cauchy. Suppose not. Then there exist $\varepsilon \in (0, \infty)$ and sequences $\{p(n)\}$, $\{q(n)\}$ such that for each $n \geqq 0$,

$$p(n) > q(n) > n , \tag{20}$$

$$d(x_{p(n)}, x_{q(n)}) \geqq \varepsilon , \tag{21}$$

and (by the well-ordering principle),

$$d(x_{p(n)-1}, x_{q(n)}) < \varepsilon . \tag{22}$$

Let $n \geqq 0$ be given, $c_n = d(x_{p(n)}, x_{q(n)})$. Then

$$\begin{aligned}\varepsilon &\leqq c_n \\ &\leqq d(x_{p(n)}, x_{p(n)-1}) + d(x_{p(n)-1}, x_{q(n)}) < b_{p(n)-1} + \varepsilon\ .\end{aligned} \tag{23}$$

From $c_1 = c_2 = 0$, we conclude that $\{c_n\}$ converges to $\varepsilon$ from the right. Let

$$I_1 = \{n\colon\ p(n), q(n) \text{ are odd}\}\ ,$$
$$I_2 = \{n\colon\ p(n) \text{ is odd, } q(n) \text{ is even}\}\ .$$
$$I_3 = \{n\colon\ p(n) \text{ is even, } q(n) \text{ is odd}\}\ ,$$
$$I_4 = \{n\colon\ p(n), q(n) \text{ are even}\}\ .$$

Then at least one of $I_1$, $I_2$, $I_3$, $I_4$ is infinite. Suppose first that $I_1$ is infinite. Let

$$d_n = d(x_{p(n)-1}, x_{q(n)})\ , \qquad n = 0, 1, 2, \cdots\ .$$

Since $\{c_n\}$ converges to $\varepsilon$ and $\{b_n\}$ converges to 0, we conclude from (22) that $\{d_n\}$ converges to $\varepsilon$ from the left. Thus

$$J_1 = \{n \in I_1\colon\ x_{p(n)-1} \neq x_{q(n)}\}$$

is infinite. Let $n \in J_1$, $u_n = d(x_{p(n)-1}, x_{q(n)+1})$. Then

$$\begin{aligned} c_n = d(x_{p(n)}, x_{q(n)}) &\leqq d(x_{p(n)}, x_{q(n)+1}) + d(x_{q(n)+1}, x_{q(n)}) \\ &\leqq d(S(x_{p(n)-1}),\ T(x_{q(n)})) + b_{q(n)}\ .\end{aligned} \tag{24}$$

From $(d)$,

$$\begin{aligned} d(S(x_{p(n)-1}),\ T(x_{q(n)})) \leqq\ & \alpha_1(d_n)b_{p(n)-1} + \alpha_2(d_n)b_{q(n)} + \alpha_3(d_n)u_n \\ & + \alpha_4(d_n)c_n + \alpha_5(d_n)d_n\ .\end{aligned} \tag{25}$$

From (24) and (25),

$$\begin{aligned} c_n \leqq\ & \alpha_1(d_n)b_{p(n)-1} + \alpha_2(d_n)b_{q(n)} + \alpha_3(d_n)u_n + \alpha_4(d_n)c_n \\ & + \alpha_5(d_n)d_n + b_{q(n)}\ .\end{aligned} \tag{26}$$

Without loss of generality, we may assume that each $\alpha_i$ is continuous from the left, for we can replace the $\alpha_i$'s by

$$\beta_i(t) = \lim_{s \uparrow t} \alpha_i(s)\ , \quad t > 0\ , \qquad i = 1, 2, 3, 4, 5$$

and conditions (a), (b), (c), and (d) still hold. Thus

$$\lim_{n \to \infty} \alpha_i(d_n) = \alpha_i(\varepsilon)\ , \qquad i = 1, 2, 3, 4, 5\ .$$

So from (26),

$$\varepsilon \leqq (\alpha_3(\varepsilon) + \alpha_4(\varepsilon) + \alpha_5(\varepsilon))\varepsilon < \varepsilon\ ,$$

a contradiction. Now suppose that $I_2$ is infinite. By a similar argument, $J_2 = \{n \in I_2 \colon x_{p(n)-1} \neq x_{q(n)-1}\}$ is infinite. Let $n \in J_2$,

$$v_n = d(x_{p(n)-1}, x_{q(n)-1}) , \quad w_n = d(x_{p(n)}, x_{q(n)-1}) .$$

Then

$$\begin{aligned} c_n &= d(S(x_{p(n)-1}), T(x_{q(n)-1})) \\ &\leqq \alpha_1(v_n)b_{p(n)-1} + \alpha_2(v_n)b_{q(n)-1} + \alpha_3(v_n)d_n + \alpha_4(v_n)w_n + \alpha_5(v_n)v_n . \end{aligned} \tag{27}$$

Since $\{v_n\}$ converges to $\varepsilon$ (not necessarily from the left or right), we obtain the same contradiction from (27). The other two cases are similar to the above two except the roles of $S$, $T$ interchange. Hence $\{x_n\}$ is Cauchy. By completeness, $\{x_n\}$ converges to a point $x$ in $X$. Since $b_n > 0$ for each $n$, $J = \{n \colon x \neq x_{2n+1}\}$ or $K = \{n \colon x \neq x_{2n}\}$ is infinite. Suppose that $K$ is infinite. Let $n \in K$,

$$l_n = d(x, x_{2n}) , \quad h_n = d(x, x_{2n+1}) .$$

Then

$$\begin{aligned} d(x, T(x)) &\leqq d(x, x_{2n+1}) + d(x_{2n+1}, T(x)) \\ &= h_n + d(S(x_{2n}), T(x)) \\ &\leqq h_n + \alpha_1(l_n)b_{2n} + \alpha_2(l_n)d(x, T(x)) + \alpha_3(l_n)d(x_{2n}, T(x)) \\ &\quad + \alpha_4(l_n)h_n + \alpha_5(l_n)l_n \\ &\leqq h_n + \alpha_1(l_n)b_{2n} + \alpha_2(l_n)d(x, T(x)) + \alpha_3(l_n)[l_n + d(x, T(x))] \\ &\quad + \alpha_4(l_n)h_n + \alpha_5(l_n)l_n . \end{aligned}$$

So

$$\begin{aligned} d(x, T(x)) \leqq \frac{1 + \alpha_4(l_n)}{1 - \alpha_2(l_n) - \alpha_3(l_n)} h_n &+ \frac{\alpha_3(l_n) + \alpha_5(l_n)}{1 - \alpha_2(l_n) - \alpha_3(l_n)} l_n \\ &+ \frac{\alpha_1(l_n)}{1 - \alpha_2(l_n) - \alpha_3(l_n)} b_{2n} . \end{aligned} \tag{28}$$

From (a) and (c), the sequences

$$\frac{1 + \alpha_4(l_n)}{1 - \alpha_2(l_n) - \alpha_3(l_n)} , \quad \frac{\alpha_3(l_n) + \alpha_5(l_n)}{1 - \alpha_2(l_n) - \alpha_3(l_n)} , \quad \frac{\alpha_1(l_n)}{1 - \alpha_2(l_n) - \alpha_3(l_n)}$$

are bounded. So from (28), $T(x) = x$. Similarly, $S(x) = x$ if $J$ is infinite. Hence $S$ or $T$ has a fixed point.

The following result follows easily from Theorem 2.

THEOREM 3. *With the conditions of Theorem 2, if further,*

$$d(S(x), T(x)) \leqq \alpha\,[d(x, S(x)) + d(x, T(x))] , \quad x \in X$$

*for some* $\alpha \in [0, 1)$, *then each of* $S$, $T$ *has a unique fixed point and these two fixed points coincide.*

We remark that the conditions of Theorem 1 imply the conditions of Theorem 3. Also, G. Hardy and T. Rogers [5, Theorem 2] gave a different proof for the case $S = T$. Their proof cannot be modified for the general case. To see that the conclusion of Theorem 2 is best possible, we note that if $X = \{0, 1\}$ with the usual distance and if $S$, $T$ are two distinct functions of $X$ onto $X$, then $S$, $T$ satisfy the conditions of Theorem 2 (and Theorem 3 with $\alpha = 1$), but one has two fixed points and the other has none.

THEOREM 4. *Let* $(X, d)$ *be a nonempty compact metric space. Let* $S$, $T$ *be functions of* $X$ *into itself. Suppose that* $S$ *or* $T$ *is continuous. Suppose further that there exist nonnegative real-valued decreasing functions* $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ on $(0, \infty)$ *such that*

(a) $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \leqq 1$,

(b) $\alpha_1 = \alpha_2$ *and* $\alpha_3 = \alpha_4$,

(c) *for any distinct* $x$, $y$ *in* $X$,

$$d(S(x), T(y)) < a_1 d(x, S(x)) + a_2 d(y, T(y)) + a_3 d(x, T(y)) + a_4 d(y, S(x)) + a_5 d(x, y) ,$$

*where* $a_i = \alpha_i(d(x, y))$.

*Then* $S$ *or* $T$ *has a fixed point. If both* $S$ *and* $T$ *have fixed points, then each of* $S$ *and* $T$ *has a unique fixed point and these two fixed points coincide.*

*Proof.* By symmetry, we may assume that $S$ is continuous. Let $f$ be the function on $X$ such that

$$f(x) = d(x, S(x)) , \quad x \in X .$$

Then $f$ is continuous (we merely need the fact that $f$ is lower semi-continuous) on $X$. So $f$ takes its minimum value at some $x_0$ in $X$. We claim that $x_0$ is a fixed point of $S$ or $S(x_0)$ is a fixed point of $T$. Suppose not. Let

$$x_1 = S(x_0) , \quad x_2 = T(x_1) , \quad x_3 = S(x_2) ,$$
$$b_0 = d(x_0, x_1), \quad b_1 = d(x_2, x_3), \quad b_2 = d(x_2, x_3) .$$

Then $b_0 > 0$, $b_1 > 0$. From (c), we can prove that

$$(1 - \alpha_2(b_0) - \alpha_3(b_0))b_1 < (\alpha_1(b_0) + \alpha_3(b_0) + \alpha_5(b_0))b_0 . \tag{29}$$

Let

$$p(t) = 1 - \alpha_2(t) - \alpha_3(t)\ , \quad q(t) = \alpha_1(t) + \alpha_3(t) + \alpha_5(t)\ , \quad t > 0\ .$$

From (a) and (b), $p(b_0) > 0$. So

$$b_1 < \frac{q(b_0)}{p(b_0)} b_0\ . \tag{30}$$

Similarly,

$$b_2 < \frac{v(b_1)}{u(b_1)} b_1\ , \tag{31}$$

where

$$u(t) = 1 - \alpha_1(t) - \alpha_4(t), v(t) = \alpha_2(t) + \alpha_4(t) + \alpha_5(t)\ , \quad t > 0\ .$$

From (30) and (31),

$$b_2 < \frac{v(b_1)}{u(b_1)} \frac{q(b_0)}{p(b_0)} b_0\ . \tag{32}$$

It suffices to prove that $(v(b_1)q(b_0)/u(b_1)p(b_0)) < 1$, for then, $b_2 < b_0$, a contradiction to the minimality of $b_0$. Let $b = \min\{b_0, b_1\}$. Then

$$v(b_1)q(b_0) - u(b_1)p(b_0) \leqq v(b)q(b) - u(b)p(b) < 0$$

if $\alpha_1 = \alpha_2$ and $\alpha_3 = \alpha_4$. So $S$ or $T$ has a fixed point. Now suppose that $x$ is a fixed point of $S$ and $y$ is a fixed point of $T$. Then $x = y$, otherwise, from (c),

$$d(x, y) = d(S(x), T(y)) < d(x, y)\ ,$$

a contradiction.

The following result is stated without proof.

THEOREM 5. *Let $(X, d)$ be complete metric space. Let $\{S_n\}$, $\{T_n\}$ be sequence of functions of $X$ into $X$ which converge pointwise to $S$, $T$ respectively. Suppose that the pairs $(S_n, T_n)$ satisfy the conditions of Theorem 3 with the same $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$. Then $S$, $T$ have a unique common fixed point $x$ and $x$ is the limit of the sequence $\{x_n\}$ of the fixed points $x_n$ of $S_n$.*

THEOREM 6. *Let $(X, d)$ be a nonempty compact metric space. Let $\{S_n\}$, $\{T_n\}$ be sequences of functions of $X$ into itself which converge pointwise to the functions $S$, $T$ on $X$ respectively. Suppose that for each $n$, there exist decreasing functions $\alpha_1^n, \alpha_2^n, \alpha_3^n, \alpha_4^n, \alpha_5^n$ of $(0, \infty)$ into $[0, \infty)$ such that*

(a) $\alpha_1^n + \alpha_2^n + \alpha_3^n + \alpha_4^n + \alpha_5^n \leqq 1$,
(b) $\alpha_1^n = \alpha_2^n$ *and* $\alpha_3^n = \alpha_4^n$,
(c) *for any distinct* $x$, $y$ *in* $X$,

$$d(S_n(x), T_n(y)) < a_1^n d(x, S_n(x)) + a_2^n d(y, T_n(y)) + a_3^n d(x, T_n(y)) + a_4^n d(y, S_n(x)) + a_5^n d(x, y) \ ,$$

*where*

$$a_i^n = \alpha_i^n(d(x, y)) \ .$$

*Then $S$ or $T$ has a fixed point. Indeed, every cluster point of a sequence $\{x_n\}$ of fixed points $x_n$ of $S_n$ or $T_n$ is a fixed point of $S$ or $T$.*

*Proof.* By Theorem 4, for each $n$, either $S_n$ or $T_n$ has a fixed point. By symmetry, we may assume that $S_n$ has a fixed point for infinitely many of $n$'s. So there is a subsequence $\{S_{n(k)}\}$ of $\{S_n\}$ such that each $S_{n(k)}$ has a fixed point, say $x_k$. By compactness, we may (by taking a subsequence) assume that $\{x_k\}$ converges to some $x$ in $X$. We shall prove that $x$ is a fixed point of $S$ or $T$. If $x_k \neq x$ for only finitely many of $k$'s, then

$$\begin{aligned} S(x) &= \lim_{k\to\infty} S_{n(k)}(x) \\ &= \lim_{k\to\infty} S_{n(k)}(x_k) \\ &= \lim_{k\to\infty} x_k \\ &= x \ . \end{aligned}$$

So we may assume that $x_k \neq x$ for infinitely many of $k$'s. By taking a subsequence, we may assume that $x_k \neq x$ for each $k$. Let $k \geqq 1$ and $b_k = d(x, x_k)$. Then

$$\begin{aligned} d(x, T(x)) &\leqq d(x, x_k) + d(x_k, T_{n(k)}(x)) + d(T_{n(k)}(x), T(x)) \\ &= d(x, x_k) + d(S_{n(k)}(x_k), T_{n(k)}(x)) + d(T_{n(k)}(x), T(x)) \ . \end{aligned} \tag{33}$$

From (c),

$$d(S_{n(k)}(x_k), T_{n(k)}(x)) < \alpha_2^k(b_k)d(x, T_{n(k)}(x)) + \alpha_3^k(b_k)d(x_k, T_{n(k)}(x)) + \alpha_4^k(b_k)d(x, x_k) + \alpha_5^k(b_k)b_k \ . \tag{34}$$

Combining (33) and (34) and letting $k$ tend to the infinity, we have

$$\begin{aligned} d(x, T(x)) &\leqq \limsup_{k\to\infty} (\alpha_2^k(b_k) + \alpha_3^k(b_k))d(x, T(x)) \\ &\leqq \limsup_{k\to\infty} \lim_{t\downarrow 0} (\alpha_2^k(t) + \alpha_3^k(t))d(x, T(x)) \ . \end{aligned} \tag{35}$$

From (b), $\alpha_2^k(t) + \alpha_3^k(t) \leqq 1/2$ for each $t > 0$, $k = 1, 2, \cdots$. So

$$\limsup_{k\to\infty} \lim_{t\downarrow 0} (\alpha_2^k(t) + \alpha_3^k(t)) \leqq \frac{1}{2} . \tag{36}$$

From (35) and (36), we conclude that $T(x) = x$.

From the proof, we know that the same conclusion holds if in Theorem 6, we replace (b) by the following weaker conditions:

$$\alpha_1^n = \alpha_2^n \quad \text{or} \quad \alpha_3^n = \alpha_4^n ,$$

$$\limsup_{k\to\infty} \lim_{t\downarrow 0} (\alpha_2^k(t) + \alpha_3^k(t)) < 1 ,$$

and

$$\limsup_{k\to\infty} \lim_{t\downarrow 0} (\alpha_1^n(t) + \alpha_4^n(t)) < 1 .$$

We note that, unlike Theorem 5, $S$, $T$ in Theorem 6 need not satisfy the condition required for the pairs $(S_n, T_n)$.

THEOREM 7. *Let $(X, d)$ be a nonempty compact metric space. Let $\{S_n\}$ be a sequence of functions of $X$ into itself which converges pointwise to some function $S$ on $X$. Suppose that for each $n$, there exist decreasing functions $\alpha_1^n$, $\alpha_2^n$, $\alpha_3^n$, $\alpha_4^n$, $\alpha_5^n$ of $(0, \infty)$ into $[0, \infty)$ such that*

(a) $\alpha_1^n + \alpha_2^n + \alpha_3^n + \alpha_4^n + \alpha_5^n \leqq 1$,

(b) *for any distinct $x$, $y$ in $X$,*

$$d(S_n(x), S_n(y)) < a_1 d(x, S_n(x)) + a_2 d(y, S_n(y)) + a_3 d(x, S_n(y)) + a_4 d(y, S_n(x)) + a_5 d(x, y) ,$$

*where*

$$a_i = \alpha_i(d(x, y)) .$$

*Then $S$ has a fixed point. Indeed, every cluster point of the sequence of fixed points of $S_n$ is a fixed point of $S$.*

The above result follows from Theorem 6 by averaging two applications of condition (b).

We shall now give a simple example to show that the conclusion of Theorem 7 is best possible. Let $X$ be a star-shaped [4] compact subset of a normed linear space $B$. Then there exists a point $z$ in $X$ such that for any $y$ in $X$, the line segment

$$\{tz + (1 - t)y\colon\ t \in [0, 1]\}$$

is contained in $X$. For each $n$, let

$$S_n(x) = \frac{1}{n} z + \left(1 - \frac{1}{n}\right) x , \qquad x \in X .$$

Then $\{S_n\}$ is a sequence of mappings of $X$ into $X$ which satisfy the conditions of Theorem 7. $\{S_n\}$ converges pointwise to the identity function $S$ on $X$. Every point of $X$ is a fixed point of $S$. So unlike Theorem 5, it is too much to ask that $S$ in Theorem 7 has a unique fixed point.

## REFERENCES


1. W. M. Boyce, *Commuting functions with no common fixed point*, Trans. Amer. Math. Soc., **137** (1969), 77-92.
2. ———, *Γ-compact mappings on an interval and fixed points*, Trans. Amer. Math. Soc., **160** (1971), 87-102.
3. M. Edelstein, *On fixed and periodic points under contractive mappings*, J. London Math. Soc., **37** (1962), 74-79.
4. B. Halpern, *The kernel of a starshaped subset of the plane*, Proc. Amer. Math. Soc., **23** (1969), 692-696.
5. G. Hardy and T. Rogers, *A generalization of a fixed point theorem of Reich*, (to appear).
6. J. P. Huneke, *On common fixed point of commuting continuous functions on an interval*, Trans. Amer. Math. Soc., **139** (1969), 371-381.
7. R. Kannan, *Some results on fixed points-II*, Amer. Math. Monthly, **76** (1969), 405-408.
8. E. Rakotch, *A note on contractive mappings*, Proc. Amer. Math. Soc., **13** (1962), 469-465.
9. S. Reich, *Some remarks concerning contraction mappings*, Canad. Math. Bull. **14** (1971), 121-124.
10. P. Srivastava and V. K. Gupta, *A note on common fixed points*, Yokohama Math. J., **XIX** (1971), 91-95.
11. Chi Song Wong, *Fixed point theorems for nonexpansive mappings*, J. Math. Anal. Appl., **37** (1972), 142-150.


Received June 13, 1972. This research was partially supported by the National Research Council of Canada Grant A8518 and a grant from the Canadian Mathematical Congress. It was prepared while the author was at the Summer Research Institute, University of Alberta.

UNIVERSITY OF WINDSOR

# Pacific Journal of Mathematics

Vol. 48, No. 1 March, 1973