ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 47

FASC. 3—4

SZEGED, 1984

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

47. KÖTET

FASC. 3—4.

SZEGED, 1984

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

Acta Sci. Math. 47 (1984), 265—275

Amalgamated free product of lattices. III. Free generating sets

G. GRÄTZER* and A. P. HUHN

1. Introduction

In G. GRÄTZER and A. P. HUHN [4] it was proved that for a finite lattice Q any two Q-free products have a common refinement. This means that, whenever L, A_0, A_1, B_0, B_1 are lattices such that $L=A_0 *_Q A_1=B_0 *_Q B_1$, then

$$L = (A_0 \cap B_0) *_Q (A_0 \cap B_1) *_Q (A_1 \cap B_0) *_Q (A_1 \cap B_1)$$

$$A_i = (A_i \cap B_0) *_Q (A_i \cap B_1), \quad i = 0, 1,$$

and

$$B_j = (A_0 \cap B_j) *_Q (A_1 \cap B_j), \quad j = 0, 1.$$

It is still an open question whether there is any lattice Q not having this property. In this paper, we shall prove a related weaker statement.

By a *free generating set* of a lattice L we mean any relative sublattice freely generating L. The following question arises:

Is it true, that a free generating set of an amalgamated free product always contains free generating sets of the components?

In case of an affirmative answer it would follow that, for arbitrary Q, any two Q-free products have a common refinement, thus the above property is, indeed, stronger than the Common Refinement Property. In fact, assume that $L=A_0 *_Q A_1= =B_0 *_Q B_1$. Then $B_0 \cup B_1$ is a free generating set of L. Hence $A_i \cap (B_0 \cup B_1)= =(A_i \cap B_0) \cup (A_i \cap B_1)$ is a generating set of A_i. Thus, by Section 5 of [4], $A_i= =(A_i \cap B_0) *_Q (A_i \cap B_1)$, $i=0, 1$, whence, by the Main Theorem of [4], it follows that the two Q-free products have a common refinement.

We shall give a negative answer by proving the following theorem.

Theorem 1. *There exist lattices L, A_0, A_1, Q with $L=A_0 *_Q A_1$ and a free generating set G of L such that $[G \cap A_i]$ is a proper part of A_i, $i=0, 1$.*

Received February 25, 1983

* The research of the first author was supported by the NSERC of Canada.

In fact, in our example $[G\cap A_i]=Q$, and $\bar{Q}$ is a proper part of A_i. The generating set $\bar{G}$ will be of the form $\bar{B}_0\cup\bar{B}_1$, where $\bar{B}_0$, $\bar{B}_1$ are relative sublattices of L, and $L=B_0*_Q B_1$ with $B_i=[\bar{B}_i]$. Therefore, it is natural to ask whether Theorem 1 can be developed into a counterexample showing that the two Q-free products $A_0*_Q A_1$ and $B_0*_Q B_1$ have no common refinement. Theorem 2 in Section 5 shows that this is not the case.

2. The construction of Q, A_i and G

First we shall define a partial lattice P and relative sublattices $\bar{Q}$, $\bar{A}_i$, and $\bar{B}_i$ of P, which will serve as generating sets of Q, A_i, and B_i, respectively. For a set X, let $S_e(X)$ denote the free semigroup on X with unit element e. P is defined as a subset of $S_e(\{0, 1, l, r\})$:

$$P= S_e(\{0, 1\})\cup\{sl|s\in S_e(\{0, 1\})\}\cup\{sr|s\in S_e(\{0, 1\})\}.$$

The elements of P will be referred to as words, the elements $0, 1, r, l$ will be called letters. The last letter of a word s will be denoted by $\breve{s}$. $|s|$ will denote the number of letters in s. e will be considered the empty word. We shall use the convention that $\breve{e}=0$. Now we start defining joins and meets in P.

(i) For any $s\in S_e(\{0, 1\})$, define $s=s0\vee s1=s1\vee s0$.

(ii) For any $s\in S_e(\{0, 1\})$, define

$$s = sr\vee sl = sl\vee sr,$$
$$sl = s00\vee s10 = s10\vee s00,$$
$$sr = s01\vee s11 = s11\vee s01.$$

(iii) For any s, $p_0, p_1\in P$ with $\breve{s}=0$, define

$$sr = s01p_0\vee s11p_1 = s11p_1\vee s01p_0,$$

and, for any s, $p_0, p_1\in P$ with $\breve{s}=1$, define

$$sl = s00p_0\vee s10p_1 = s10p_1\vee s00p_0.$$

Now let

$$\bar{Q} = \{sr\,|\,\breve{s} = 0\}\cup\{sl\,|\,\breve{s} = 1\},$$
$$\bar{A}_i = \bar{Q}\cup\{s\,|\,\breve{s} = i,\ |s| \text{ is even},\ x\in\{e, l, r\}\},\quad i = 0, 1,$$
$$\bar{B}_i = \bar{Q}\cup\{s\,|\,\breve{s} = i,\ |s| \text{ is odd},\ x\in\{e, l, r\}\},\quad i = 0, 1.$$

(iv) For $a, b\in P$, define $a\leqq b$ if and only if either $a=b$ or there exist a positive integer n, and elements $a_0, a_1, \dots, a_{n-1}, a_n, c_0, c_1, \dots, c_{n-1}\in P$, such that $a=a_0$, $a_n=b$ and the relations $a_i\vee c_i=a_{i+1}$, $i=0, 1, \dots, n-1$ hold by (i), (ii), or (iii).

This relation is a partial ordering on P. If $a \leqq b$, define $a \vee b = b \vee a = b$ and $a \wedge b = b \wedge a = a$.

A part of P together with all non-trivial joins (there are only trivial meets) is illustrated in the Figure.

Finally, let $L = F(P)$, the free lattice over P, let $Q = [\bar{Q}]$, $A_i = [\bar{A}_i]$, and $B_i = [\bar{B}_i]$, $i = 0, 1$, in L, and let $G = \bar{B}_0 \cup \bar{B}_1$.

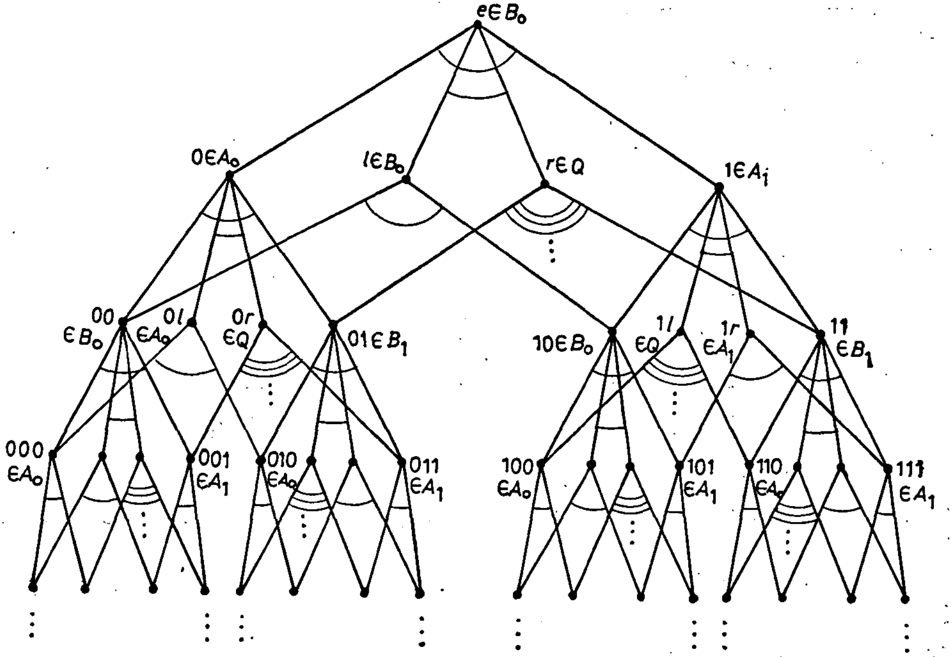


Fig. 1.

3. *P* is a partial lattice

This statement is of primary importance in the proof of Theorem 1 (see the proof of Lemma 6). In this section we shall give a proof. The following lemma will be used to prove that P is a weak partial lattice.

Lemma 1. *For any $a, b, c \in P$, if $a \leqq c$, $b \leqq c$, and $a \vee b$ is defined, then $a \vee b \leqq c$.*

The proof of this lemma proceeds via checking all the possible cases (i), (ii), (iii), and (iv) of how $b \vee c$ is defined and establishing the assertion in these separate cases. We omit the details.

Lemma 2. *P is a weak partial lattice.*

Proof. The following four statements and their duals are to be proved.

(a) for any $a \in P$, $a \vee a$ is defined and $a \vee a = a$;

(b) for any $a, b \in P$, if $a \vee b$ is defined, then $b \vee a$ is defined and $a \vee b = b \vee a$;

(c) for any $a, b \in P$, if $a \vee b$, $(a \vee b) \vee c$, $b \vee c$ are defined, then $a \vee (b \vee c)$ is defined and $(a \vee b) \vee c = a \vee (b \vee c)$;

(d) for an $a, b \in P$, if $a \wedge b$ is defined, then $a \vee (a \wedge b)$ is defined and $a \vee (a \wedge b) = a$.

Of these only (c) is non-trivial. We consider the following five cases.

First case: $a \vee b = b$. Then $a \leqq b \leqq b \vee c$, thus the right hand side in (c) exists and equals $b \vee c (= (a \vee b) \vee c)$.

Second case: $b < a \vee b$, $c \| b$, and $c \| a \vee b$. Observe that, under these conditions, the joins $b \vee c$ and $(a \vee b) \vee c$ can only be defined if, for suitable elements $p_0, p_1, p_2, s \in P$, one of the following four subcases holds:

$$\breve{s} = 0, \quad c = s01p_0, \quad b = s11p_1, \quad a \vee b = s11p_2;$$

$$\breve{s} = 0, \quad c = s11p_0, \quad b = s01p_1, \quad a \vee b = s01p_2;$$

$$\breve{s} = 1, \quad c = s00p_0, \quad b = s10p_1, \quad a \vee b = s10p_2;$$

$$\breve{s} = 1, \quad c = s10p_0, \quad b = s00p_1, \quad a \vee b = s00p_2.$$

In the first two subcases $a \vee (b \vee c)$ exists and equals sr, which is also the value of $(a \vee b) \vee c$. The last two subcases are similar, only the common value of the two sides is s.

Third case: $b < a \vee b$, $b \leqq c$, and $c \| a \vee b$. This case is impossible, for $(a \vee b) \vee c$ is defined and two incomparable elements whose join is defined cannot have a common lower bound (check the definitions (i), (ii), and (iii)).

Fourth case: $b < a \vee b$, $b \| c$, $c \leqq a \vee b$. Applying Lemma 1, we have that $b \leqq b \vee c \leqq a \vee b$. If the join $a \vee b$ was defined in (i) or (ii), then $b \vee c = b$, or $b \vee c = a \vee b$. But $b \vee c = b$ contradicts $b \| c$, thus $b \vee c = a \vee b$. Then $a \vee (b \vee c)$ is defined and $a \vee (b \vee c) = a \vee (a \vee b) = a \vee b = (a \vee b) \vee c$. If $a \vee b$ was defined in (iv), then $a \geqq b$, thus $a = a \vee b$. Hence $a \vee (b \vee c) = a = (a \vee b) \vee c$. Finally, if $a \vee b$ was defined in (iii), then we again have to consider four subcases as in the second case; we check only one of these:

$$a = s01p_0, \quad b = s11p_1, \quad \breve{s} = 0.$$

Then $a \vee b = sr$, whence $s11p_1 \leqq b \vee c \leqq sr$. Thus either $b \vee c = sr = a \vee b$, which can be handled similarly as the cases (i) or (ii), or there is a factorization $p_1 = p_2 p_3$ such that $b \vee c = s11p_2$ ($q_1 = e$ is allowed, too). But then, (iii) applies again, whence $a \vee (b \vee c) = sr = a \vee b = (a \vee b) \vee c$.

Fifth case: $b < a \vee b$ and c is comparable with both b and $a \vee b$. Then the subcases $c \leqq b$ and $a \vee b \leqq c$ are trivial and $b \leqq c \leqq a \vee b$ can be handled similarly as the fourth case.

These five cases exhaust all possibilities.

To finish the proof of the statement formulated in the heading of this section, we have to prove the following lemma (and its dual, but the latter is obvious).

Lemma 3. *If $a, b, c \in P$ and $(a] \vee (b] = (c]$ in the ideal lattice of P, then $a \vee b = c$ in P.*

Proof (by R. W. Quackenbush). Suppose that $(a] \vee (b] = (c]$ and $a \vee b$ is *not* defined. Let $a * b = s$ be the largest common initial segment of a and b. Then $(a] \vee (b] \subseteq (s]$, so $c \leqq s$. Now $a, b \in \{sr, sl, s0p, s1q\}$ for some p, q.

Case 1. $a = sl$. Then $b = s0p$ or $s1q$ since $sl \vee sr = s$.

1.1: $b = s0p$. Since $s \succ sl$, $c = s$.

Claim. $(s0] \vee (sl] = (s0] \cup (sl]$.

Proof. Let $d \leqq s0$ and $e \leqq sl$. Thus $d = s0p$; we assume that $d \vee e$ is defined. Thus $e \neq sl$; so $e = s00q$ or $s10q$. If $e = s00q$ then $d \vee e \leqq s0$. Thus let $e = s10q$. Then $d \vee e = sl$. This contradicts $(a] \vee (b] = (s]$ since $a \leqq s0$ and $b = sl$.

1.2: $b = s1q$. Similar to 1.1 using $(s1] \vee (sl] = (s1] \cup (sl]$.

Case 2: $a = sr$. By symmetry with Case 1.

Case 3: $a = s0p$, $b = s1q$. By symmetry, this is the last case.

3.1: $a = s0$. Thus $q \neq \emptyset$. We compute $(s0] \vee (s1q]$. Let $d \leqq s0$ and $e \leqq s1q$ and let us assume that $d \vee e$ is defined.

3.11: $q = 1q'$. The only possibilities are:

$$d \vee e = s01 \vee s11 = sr, \quad d \vee e = s01p' \vee s11q' = sr.$$

Thus either $sr \in (s0] \vee (s1q]$ and so $(s0] \vee (s1q] = (s0] \vee (sr] = (s0] \cup (sr]$ or $(s0] \vee (s1q] = (s0] \cup (s1q]$.

3.12: $q = 0q'$. Then similarly to 1.11, $(s0] \vee (s1q] = (s0] \cup (s1q]$ or $(s0] \cup (sl]$.

3.2: $b = s1$. So $p = \emptyset$. By symmetry with 1.1.

3.3: $a = s00p'$, $b = s11q'$. If $d \leqq a$ and $e \leqq b$, then $d = s00p''$, $e = s11q''$ and $d \vee e$ is not defined. Thus $(a] \vee (b] = (a] \cup (b]$.

3.4: $a = s01p'$, $b = s10q'$. Similar to 3.3.

3.5: $a = s00p'$, $b = s10q'$. Let $d \leqq a$ and $e \leqq b$. Since $a \vee b$ is not defined we must have $p' \neq \emptyset$ or $q' \neq \emptyset$ and we must have $\bar{s} = 0$. But then $d \vee e$ is not defined, since $d = s00p''$, $e = s10q''$. Hence $(a] \vee (b] = (a] \cup (b]$.

3.6: $a = s01p'$, $b = s11q'$. Similar to 3.5.

Now the above results, together with Funayama's characterization of partial lattices (see G. Grätzer [3]), gurantee that P is a partial lattice.

4. Proof of Theorem 1

We shall need a description of the free lattice generated by a partial lattice. The description we use is due to R. A. Dean [2] (see also H. Lakser [5]). Let $\langle X; \wedge, \vee \rangle$ (or briefly X) be a partial lattice, and let $F(X)$ denote the free lattice generated by X. Denote by $FP(X)$ the algebra of polynomial symbols in the two binary operation symbols $\wedge$ and $\vee$ generated by the set X. Then $F(X)$ is the image of $FP(X)$ under a homomorphism $\varrho\colon FP(X)\to F(X)$ with $x\varrho=x$ for $x\in X$. For each $p\in FP(X)$, we define an ideal p_X and a dual ideal p^X of X as follows.

$$p_X = \{x\in X | x \leqq p\varrho \text{ in } F(X)\}, \quad p^X = \{x\in X | p\varrho \leqq x \text{ in } F(X)\}.$$

Now the description of $F(X)$ is found in the following three propositions. Actually, we need here only Propositions 2 and 3; Proposition 1 will be used in Section 5.

Proposition 1. *If $p, q\in FP(X)$, then $p\varrho\leqq q\varrho$ iff it follows by applying the following five rules.*

(W_C) $p^X\cap q_X \neq \emptyset$;

$({}_\vee W)$ $p = p_0\vee p_1$, $p_0\varrho \leqq q\varrho$ *and* $p_1\varrho \leqq q\varrho$;

$({}_\wedge W)$ $p = p_0\wedge p_1$, $p_0\varrho \leqq q\varrho$ *or* $p_1\varrho \leqq q\varrho$;

$(W_\vee)$ $q = q_0\vee q_1$, $p\varrho \leqq q_0\varrho$ *or* $p\varrho \leqq q_1\varrho$;

$(W_\wedge)$ $q = q_0\wedge q_1$, $p\varrho \leqq q_0\varrho$ *and* $p\varrho \leqq q_1\varrho$.

If $p\in P(X)$, then p_X and p^X can be calculated as follows.

Proposition 2. *For $p\in X$, $p_X=(p]$ (in $\langle X; \wedge, \vee\rangle$) and $p^X=[p)$. For $p=p_0\vee p_1$,*

$$p_X = (p_0)_X\vee (p_1)_X, \quad p^X = (p_0)^X\wedge (p_1)^X,$$

and, for $p=p_0\wedge p_1$,

$$p_X = (p_0)_X\wedge (p_1)_X, \quad p^X = (p_0)^X\vee (p_1)^X$$

where the $\vee$ and $\wedge$ on the right hand sides are to be formed in the lattice of all ideals (respectively, dual ideals) of $\langle X; \wedge, \vee\rangle$.

By a *binary tree* we mean a finite poset T with greatest element such that every element of T is either minimal or has exactly two lower covers. Now the join and meet of a set of ideals of $\langle X; \wedge, \vee\rangle$ can be formed as follows. The operations on the dual ideals are analogous.

Proposition 3. *Let $I_j, j\in J$ be ideals of $\langle X; \wedge, \vee\rangle$. Then $x\in \bigvee (I_j | j\in J)$ iff there is a binary tree T and there exist elements $x_t\in X$, $t\in T$ such that*

(1) $x=x_{\sup T}$;

(2) *if* t *is a minimal element in* T, *then* $x_t \in I_j$ *for some* $j \in J$;

(3) *if* u *and* v *are different lower covers of* t, *then* $x_u \vee x_v$ *is defined in* $\langle X; \wedge, \vee \rangle$, *and* $x_t \leqq x_u \vee x_v$.

$\wedge (I_j | j \in J)$ *is the intersection of* $\{I_j | j \in J\}$.

The proof of Theorem 1 will be completed by the following three lemmas.

Lemma 4. *L is freely generated by* $\bar{A}_0 \cup \bar{A}_1$ *as well as by* $\bar{B}_0 \cup \bar{B}_1$.

Proof. It is enough to show that all the elements of P can be expressed by elements of $\bar{A}_0 \cup \bar{A}_1$, and these expressions obey all the relations (i) to (iv) in $F(\bar{A}_0 \cup \bar{A}_1)$, that is, (i) to (iv) can be derived from the relations valid in $\bar{A}_0 \cup \bar{A}_1$. (The statement concerning $\bar{B}_0 \cup \bar{B}_1$ can be proved analogously.) In fact, let $s \in P$, $s \notin A_0 \cup A_1$. Then an expression of s by elements of $\bar{A}_0 \cup \bar{A}_1$ is

$$s = s0 \vee s1 \quad \text{if} \quad \check{s} \in \{0, 1\}, \tag{4}$$

$$s = s'000 \vee s'001 \vee s'100 \vee s'101 \quad \text{if} \quad s = s'l, \tag{5}$$

$$s = s'010 \vee s'011 \vee s'110 \vee s'111 \quad \text{if} \quad s = s'r. \tag{6}$$

It is straightforward to check the relations (i) to (iv). Let us consider only one example: $s = s0 \vee s1$, $s \in \bar{A}_0 \cup \bar{A}_1$. In fact, applying (ii) within $\bar{A}_0 \cup \bar{A}_1$ and (4) we have

$$s = sl \vee sr = (s00 \vee s10) \vee (s01 \vee s11) = (s00 \vee s01) \vee (s10 \vee s11) = s0 \vee s1.$$

Lemma 5. $L = A_0 *_Q A_1 = B_0 *_Q B_1$.

Proof. Let $\bar{A}_0 \cup_{\bar{Q}} \bar{A}_1$ be the weakest partial lattice defined on the set $\bar{A}_0 \cup \bar{A}_1$ having $\bar{A}_0$ and $\bar{A}_1$ as sublattices. The same proof as that of Lemma 4 yields that $L = F(\bar{A}_0 \cup_{\bar{Q}} \bar{A}_1)$, for every join defined in $\bar{A}_0 \cup \bar{A}_1$ is defined either within A_0 or within A_1. Now $Q \subseteqq A_0$ and $Q \subseteqq A_1$, thus we can form the union $A_0 \dot{\cup} A_1$ subject to the condition $A_0 \cap A_1 = Q$. Let $A_0 \cup_Q A_1$ be the weakest partial lattice on $A_0 \dot{\cup} A_1$ extending the operations defined in A_0 or A_1. Since $A_0 \cup_Q A_1$ contains a copy of $\bar{A}_0 \cup_{\bar{Q}} \bar{A}_1$, there is a homomorphism φ of $L = F(\bar{A}_0 \cup_{\bar{Q}} \bar{A}_1)$ onto $F(A_0 \cup_Q A_1)$. Since L contains copies of A_0 and A_1 with $Q \subseteqq A_0$, $Q \subseteqq A_1$, there is a homomorphism ψ of $F(A_0 \cup_{\bar{Q}} A_1)$ into L. $\varphi\psi$ is the identity on $A_0 \cup_Q A_1$, hence it is the identity on L. Thus φ is one-to-one. Summarizing

$$L = F(\bar{A}_0 \cup_{\bar{Q}} \bar{A}_1) \cong F(A_0 \cup_Q A_1) = A_0 *_Q A_1.$$

This isomorphism is the identity on A_0 and on A_1, therefore L is the Q-free product of its sublattices A_0 and A_1. Analogously, $L = B_0 *_Q B_1$, completing the proof.

By Lemma 4, L has the free generating set $G = \bar{B}_0 \cup \bar{B}_1$, and, by Lemma 5, it has the Q-free decomposition $L = A_0 *_Q A_1$. Thus the following lemma proves Theorem 1.

Lemma 6. *$G\cap A_i \subseteq Q$, that is, $[G\cap A_i]$ is a proper part of A_i.*

Proof. By symmetry it is enough to show that $\bar{B}_0\cap A_0 \subseteq Q$. Let us assume that an element $b_0 \in \bar{B}_0$ can be expressed by elements of $\bar{A}_0$, that is, $b_0 = p(a_0, \ldots, a_n)$, where p is a polynomial and $a_0, \ldots, a_n \in \bar{A}_0$. Then, by Proposition 2,

$$[b_0) = p^\delta([a_0), \ldots, [a_n))$$

holds in the lattice of all dual ideals of P, where p^δ is the polynomial dual to p. This lattice is distributive and, by distributivity, p^δ can be rearranged in such a way that all the joins precede all the meets in it:

$$[b_0) = \bigvee(\bigwedge([a_j)|j\in J_i)|i\in I) \tag{7}$$

with $J_i \subseteq \{0, 1, \ldots, n\}$, for all $i\in I$, while, by the distributive inequality,

$$b_0 \leqq \bigwedge\bigvee(a_j|j\in J_i|i\in I) \tag{8}$$

holds in L. Since $[b_0)$ is a principal dual ideal, from (7) we obtain that there exists i in I such that

$$[b_0) = \bigwedge([a_j)|j\in J_i).$$

By Proposition 2, we have

$$b_0 \geqq \bigvee(a_j|j\in J_i).$$

This, together with (8) yields

$$b_0 = \bigvee(a_j|j\in J_i).$$

Again, by Proposition 2, we have

$$(b_0] = \bigvee((a_j]|j\in J_i).$$

Now we show that this is impossible unless $b_0\in Q$. We carry out the proof for $b_0 = e$; for other choices of b_0 there is no essential difference in the proof.

We show that $e\notin \bigvee(a_j]$ if a_j runs over all elements of $\bar{A}_0$. Consider a binary tree $\bar{T}$ and a set $X=\{x_t|t\in T\}$ with the properties (1) to (3), with $I_j=(a_j]$. There are only two joins with the value e, namely $e=0\vee 1$ and $e=l\vee r$. Thus X contains 0 and 1 or l and r. Of these $1\not\leqq a_j$ (respectively, $l\not\leqq a_j$) for all j, therefore, by (2), there is a $t\in T$ (t not minimal), such that $1=x_t$ (respectively, $l=x_t$). Thus (3) can be applied: 10 (and 11) or $1r$ (and $1l$) (respectively, 10 (and 00)) are contained in X. (2) does not apply for 10 and $1r$, thus we can proceed by (3): $101\in X$ or $10l\in X$. Now, by induction, we obtain that $101\ldots01\in X$ or $101\ldots0l\in X$, which contradicts the fact that X is finite. This contradiction completes the proof.

5. Some remarks

First of all, we prove the statement already announced in the introduction that the example is no counterexample for the common refinement property. It is worth mentioning that it is exactly the characterization theorem of the existence of common refinements in [4] that will be used to prove this assertion.

Theorem 2. *The two Q-free products* $L=A_0*_QA_1=B_0*_0B_1$ *have a common refinement.*

We need the following lemma.

Lemma 7. *Let* $b_0, \ldots, b_m \in B_0$ *and let p be a polynomial in* $b_0, \ldots, b_m$. *Then for any* $x \in \bar{A}_0$ *satisfying* $x \leqq p(b_0, \ldots, b_m)$ *in* $F(P)$, *there exists an element* $c \in A_0 \cap B_0$ *with* $x \leqq c \leqq p(b_0, \ldots, b_m)$.

Before proceeding to the proof, we present another lemma, which will be used in the proof of Lemma 7.

Lemma 8. *Let* $b_0, \ldots, b_m \in B_0$ *and let* $p=p_0 \vee p_1$ *be a polynomial in* $b_0, \ldots, b_m$. *Assume that, for any* $x \in \bar{A}_0$ *satisfying* $x \leqq p_i(b_0, \ldots, b_m)$ *for* $i=0$ *or* $i=1$, *there exists an element* $c \in A_0 \cup B_0$ *such that* $x \leqq c \leqq p_i$. *Let, furthermore, T be a binary tree and let* x_t, $t \in T$, *be elements of P satisfying the condition* $x_{\sup T} \in \bar{A}_0$ *as well as the conditions (2) and (3) of Section 4 with* $(p_j)_P$, $j=0, 1$, *and P in the place of* I_j, $j \in J$, *and* $\langle X; \wedge, \vee \rangle$, *respectively. Then there exists an element* $c \in A_0 \cap B_0$ *such that* $x_{\sup T'} \leqq c \leqq p(b_0, \ldots, b_m)$.

Proof of Lemma 8. We proceed by an induction. Set $b=p(b_0, \ldots, b_m)$. If $T=\{t\}$ is a singleton, then, by (2), $x_t \leqq p_i(b_0, \ldots, b_m)$, for $i=0$ or $i=1$. By one of our assumptions $x_t \leqq c \leqq p_i(t_0, \ldots, b_m)$ for a suitable $c \in A_0 \cap B_0$, whence $x_t \leqq c \leqq$ $\leqq p(b_0, \ldots, b_m)=b$. Assume that T consists of more than one element and the statement is valid for any proper binary subtree of T. Let u and v denote the different maximal elements of $T-\{\sup T\}$. Now $x_{\sup T} \leqq x_u \vee x_v \leqq b$. If $x_u \vee x_v \in \bar{A}_1$ (respectively, $x_u \vee x_v \in \bar{B}_1$), then, by Lemma 5, there exists an element $q \in Q$ such that $x_u \vee x_v \leqq q \leqq$ $\leqq x_{\sup T}$ (respectively, $x_{\sup T} \leqq q \leqq b$), proving the statement of the lemma. If $x_u \vee x_v \in$ $\in \bar{B}_0$, then we may assume that there exists no $y \in \bar{A}_1 \cup \bar{B}_1$ with $x_{\sup T} \leqq y \leqq x_u \vee x_v$, else we could find an element $q \in Q$ with $x_{\sup T} \leqq q \leqq x_u \vee x_v$ similarly as above. Thus it follows that the interval $[x_{\sup T}, x_u \vee x_v]$ contains a prime interval $[y_0, y_1]$ of P with $y_0 \in \bar{A}_0, y_1 \in \bar{B}_0$. Then, using the notation of Section 2, $y_0=y_10$. Let $c=y_10 \vee y_1r$. Obviously, $c \in A_0$. Compute:

$$c = y_10 \vee y_1r = y_100 \vee y_101 \vee y_1r = y_100 \vee y_1r.$$

Now $y_1r \in Q$ and $y_100 \in B_0$, hence $c \in B_0$, which again proves the lemma. We may assume that $x_u \vee x_v \in \bar{A}_0$. We may also assume that $x_u \vee x_v \neq x_u, x_v$. Thus, by the definition of P, either $x_u, x_v \in \bar{A}_0$ or $x_u \in \bar{B}_0, x_v \in \bar{B}_1$. In the former case we can apply the induction hypothesis for the subtrees $(u], (v] \subseteq T$, whence there exist elements $c_0, c_1 \in A_0 \cap B_0$ with $x_u \leqq c_0 \leqq b$, $x_v \leqq c_1 \leqq b$. Thus $x_u \vee x_v \leqq c_0 \vee c_1 \leqq b$ and $c_0 \vee c_1 \in \in A_0 \cap B_0$. In the latter case, using again the notations introduced in Section 2, $x_u = x_u 0 \vee x_u 1$, $x_v = x_v 0 \vee x_v 1$, and $x_u 1 \vee x_v 1 \in Q$. Now, replacing the element x_u by $x_u 0$ and x_v by $x_v 0$, we may apply the induction hypothesis for the subtrees $(u]$ and $(v]$. Hence we obtain that there exist elements $c_0, c_1 \in A_0 \cap B_0$, with $x_u 0 \leqq c_0 \leqq b$, $x_v 0 \leqq \leqq c_1 \leqq b$. Therefore

$$x_u \vee x_v = (x_u 0 \vee x_u 1) \vee (x_v 0 \vee x_v 1) \leqq c_0 \vee c_1 \vee (x_u 1 \vee x_v 1) \in A_0 \cap B_0,$$

completing the proof of Lemma 8.

Proof of Lemma 7. We again use an induction. Set $b = p(b_0, \dots, b_m)$. If p is a projection, that is $b \in \bar{B}_0$, then we may assume that there exists no $y \in \bar{A}_1 \cup \bar{B}_1$ with $x \leqq y \leqq b$. In fact, for example the existence of such an $y \in A_1$ would imply the existence of a $q \in Q$ with $x \leqq q \leqq y \leqq b$, proving the lemma. Thus the interval $[x, b]$ contains a prime interval $[y_0, y_1]$ with $y_0 \in \bar{A}_0$, $y_1 \in \bar{B}_0$, and we can proceed similarly as in the proof of Lemma 8. Consider the case $p = p_0 \wedge p_1$. By the induction hypothesis, there are elements $c_0, c_1 \in A_0 \cap B_0$ with $x \leqq c_i \leqq p_i(b_0, \dots, b_m)$. Hence $x \leqq c_0 \wedge c_1 \leqq \leqq p(b_0, \dots, b_m)$. Thus we may assume that $p = p_0 \vee p_1$, and the polynomials p_i have the property described in the lemma. By Proposition 2, we have $x \in (p_0)_p \vee (p_1)_p$. By Proposition 3, there exists a binary tree T and elements $x_t \in P$, $t \in T$, satisfying conditions (1) to (3) of Section 4, with $(p_j)_p$, $j = 0, 1$, and $\bar{P}$ in the place of I_j, $j \in J$, and $\langle X; \wedge, \vee \rangle$ respectively. Now an application of (1) and Lemma 8 completes the proof.

Proof of Theorem 2. By the main theorem of GRÄTZER, HUHN [4] and by symmetry, it suffices to prove that, for any $a \in A_0$ and $b \in B_0$ with $a \leqq b$ in L, there is an element $c \in A_0 \cap B_0$ with $a \leqq c$ and $c \leqq b$. Let $a = p'(a_0, \dots, a_n)$, $b = p(b_0, \dots \dots, b_m)$, $a_0, \dots, a_n \in \bar{A}_0$, $b_0, \dots, b_m \in \bar{B}_0$, $p, p' \in FP(P)$. We apply an induction following the description in Proposition 1. Assume $a \leqq b$ by $({}_\vee W)$, that is $p' = p'_0 \vee p'_1$ and $p'_i(a_0, \dots, a_n) \leqq p(b_0, \dots, b_m)$, $i = 0, 1$. Then, by the induction hypothesis there are elements $c_0, c_1 \in A_0 \cap B_0$, with $p'_i(a_0, \dots, a_n) \leqq c_i \leqq p(b_0, \dots, b_m)$. Hence

$$p'(a_0, \dots, a_n) = p'_0(a_0, \dots, a_n) \vee p'_1(a_0, \dots, a_n) \leqq c_0 \vee c_1 \leqq p(b_0, \dots, b_m),$$

as claimed. The proof is similar if $a \leqq b$ by $({}_\wedge W)$, $W_\vee)$, or $(W_\wedge)$. Thus we may assume that $a \leqq b$ follows from (W_C), that is, there is an element $x \in P$ with $a \leqq x \leqq b$. If $x \in \bar{A}_1$ (respectively, $x \in \bar{B}_1$), then, by Lemma 5, there exists an element $q \in Q$ with $a \leqq q \leqq x$ (respectively, $x \leqq q \leqq b$), and we can choose $c = q$. If $x \in \bar{A}_0$, then, by Lemma 7, $[x, b] \cap (A_0 \cap B_0) \neq \emptyset$. If $x \in \bar{B}_0$, then the dual of Lemma 7 yields that $[a, x] \cap$

$\bigcap(A_0 \cap B_0) \neq \emptyset$. (The dual of Lemma 7 could be proved similarly as Lemma 7 but the proof is much easier, for the operations on the dual ideals of P are the set operations.) This completes the proof.

We conclude this paper by mentioning an open problem. There is an obvious similarity between our main theorem and M. E. ADAMS' theorem [1] that a generating set of a free product (without amalgamation) need not contain generating sets of the components. This gives rise to the following question.

Problem. Need a free generating set of a free product always contain free generating sets of the components?

References

[1] M. E. ADAMS, Generators of free products of lattices, *Algebra Universalis*, **7** (1977), 409—410.

[2] R. A. DEAN, Free lattices generated by partially ordered sets and preserving bounds, *Canad. J. Math.*, **16** (1964), 136—148.

[3] G. GRÄTZER, *General Lattice Theory*, Pure and Applied Mathematics Series, Academic Press (New York, N. Y.); Mathematische Reihe, Band 52, Birkhäuser Verlag (Basel); Akademie Verlag (Berlin, 1978).

[4] G. GRÄTZER and A. P. HUHN, Amalgamated free product of lattices. I. The common refinement property, *Acta Sci. Math.*, **44** (1982), 53—66.

[5] H. LAKSER: *Free lattices generated by partially ordered sets*, Ph. D. Thesis, University of Manitoba, 1968.

(G. G.)
UNIVERSITY OF MANITOBA
WINNIPEG, MANITOBA
R3T 2N2 CANADA

(A. P. H.)
BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Acta Sci. Math. 47 (1984), 277—288

On self-injectivity and strong regularity

ROGER YUE CHI MING

A generalization of quasi-injectivity, called I-injectivity, is introduced and various properties are derived. Semi-prime left q-rings (studied in [10]) are characterized in terms of I-injectivity. Left non-singular left I-injective rings are proved to be left continuous regular. Fully left idempotent rings whose essential left ideals are two sided (which effectively generalize semi-prime left q-rings and strongly regular rings) are studied. Characteristic properties of strongly regular rings are given. Certain rings having von Neumann regular centre are considered.

Introduction

Throughout, A represents an associative ring with identity and A-modules are unitary. J, Z, Y will denote respectively the Jacobson radical, the left singular ideal and the right singular ideal of A. As usual, a left (right) ideal of A is called reduced iff it contains no non-zero nilpotent element. An ideal of A will always mean a two-sided ideal. A is called a left V-ring iff every simple left A-module is injective (cf. [5]). Recall that (1) A is ELT (resp. MELT) iff every essential (resp. maximal essential, if it exists) left ideal of A is an ideal; (2) A is a left CM-ring iff for any maximal essential left ideal M of A (if it exists), every complement left subideal is an ideal of M (cf. [21]). ELT (MELT) rings generalize left q-rings [10], left duo rings while left CM-rings generalize left PCI rings [5, p. 140], left uniform rings and left duo rings.

It is well known that A is von Neumann regular iff every left (right) A-module is flat. A theorem of I. KAPLANSKY asserts that a commutative ring is regular iff it is a V-ring [5, Corollary 19.53]. For completeness, recall that a left A-module M is p-injective iff for any principal left ideal P of A, any left A-homomorphism $g\colon P\to M$, there exists $y\in M$ such that $g(b)=by$ for all $b\in P$. Then A is regular iff every left (right) A-module is p-injective. If I is a p-injective left ideal of A, then A/I is a flat

Received December 21, 1982, and in revised form March 18, 1983.

left A-module [20, Remark 1]. Consequently, a finitely generated p-injective left ideal is a direct summand of ${}_AA$. For several years, von Neumann regular rings (introduced in [18]), self-injective rings, V-rings and associated rings have been studied by many authors (cf. [1] to [17]).

Rings whose left ideals are quasi-injective, called left q-rings, are studied in [10], where they are characterized as ELT left self-injective rings. We now introduce the following generalization of quasi-injectivity.

Definition. A left A-module M is called I-injective if, for all left submodules N, P which are isomorphic, any left A-homomorphism of N into P extends to an endomorphism of ${}_AM$.

(If Q, R are non-isomorphic quasi-injective non-injective left A-modules such that $Q \cap R = 0$ and their injective hulls are isomorphic, then $Q \oplus R$ is I-injective but not quasi-injective (cf. [7, p. 53, ex. 1].)

If every simple left A-module is p-injective, then A is fully left idempotent (cf. [14, Proposition 6]). Since any simple left A-module is I-injective, we see that I-injectivity does not even imply p-injectivity. The converse is not true either (cf. Remark 3 below).

1. I-injectivity

Our first result characterizes semi-prime left q-rings in terms of I-injectivity. A is called left I-injective iff ${}_AA$ is I-injective.

Theorem 1. *The following conditions are equivalent:*

(1) *A is an* ELT *left and right self-injective regular, left and right V-ring of bounded index;*

(2) *A is a semi-prime left q-ring;*

(3) *A is a semi-prime* ELT *left I-injective ring;*

(4) *A is a* MELT *left T-injective ring whose simple right modules are flat;*

(5) *A is an* ELT *left non-singular left I-injective ring.*

Proof. By [10, Theorem 2.3], (1) implies (2) while (2) implies (3) and (4). Since a semi-prime ELT ring is left non-singular, (3) implies (5).

If A is a MELT ring whose simple right modules are flat, then any simple left A-module is either injective or projective which implies that A is ELT (because any proper essential left ideal is an intersection of maximal left ideals). Consequently, (4) implies (5).

Assume (5). Let I be an essential left ideal of A, $g\colon I \to A$ a non-zero left A-homomorphism. For any $b \in I$, let K be a complement left ideal such that $L = l(b) \oplus K$ is an essential left ideal. If $f\colon Kb \to K$ is the map given by $f(kb) = k$ for all $k \in K$,

f is an isomorphism and by hypothesis, f extends to an endomorphism h of ${}_AA$. If $h(1)=d$, then $k=f(kb)=h(kb)=kbh(1)=kbd$ for all $k\in K$, which implies $L\subseteq l(b-bdb)$, whence $b-bdb\in Z=0$. Now $g(b)=g(bdb)=bg(db)\in I$ (because A is ELT), which shows that g is an endomorphism of ${}_AI$ and by hypothesis, g extends to an endomorphism of ${}_AA$. This proves that A is left self-injective and then (5) implies (1) by [21, Lemma 1.1].

The next corollary improves [10, Theorem 2.13].

Corollary 1.1. *A is simple Artinian iff A is a prime* ELT *left I-injective ring.*

Corollary 1.2. *The following conditions are equivalent:*

(1) *A is a direct sum of a semi-simple Artinian ring and a left and right self-injective strongly regular ring;*

(2) *A is a semi-prime* ELT *left I-injective ring.*

(Apply [10, Theorem 2.19] to Theorem 1.)

Since a prime ELT fully idempotent ring is primitive fully left idempotent, therefore [8, Theorem 6.10] and Corollary 1.1 imply

Corollary 1.3. *Suppose that A is an* ELT *fully idempotent ring such that any primitive factor ring is left I-injective. Then A is a unit-regular left and right V-ring.*

(Following [6], A is called fully idempotent (resp. fully left idempotent) iff every ideal (resp. left ideal) of A is idempotent.)

It is well-known that if A is left self-injective, then $Z=J$ (cf. for example [5, p. 78]). This is generalized in our first remark.

Remark 1. (a) If A is left I-injective, then $Z=J$ and every left or right A-module is divisible; (b) A left I-injective left Noetherian ring is left Artinian.

The following question is due to the referee: when do the rings of Remark 1 (b) coincide with quasi-Frobeniusean rings?

Theorem 2. *The following conditions are equivalent:*

(1) *A is left and right self-injective strongly regular;*

(2) *A is left non-singular left I-injective such that every maximal left ideal is an ideal;*

(3) *A is left non-singular left I-injective such that every maximal right ideal is an ideal;*

(4) *A is a reduced left I-injective ring.*

Proof. (1) implies (2) and (3) evidently.

If $J=0$ and every maximal left (resp. right) ideal of A is an ideal, then A is reduced. Consequently, either of (2) or (3) implies (4) by Remark 1 (a).

Assume (4). Since $Z=0$, the proof of Theorem 1 shows that A is von Neumann regular. Since A is reduced, A is strongly regular and hence (4) implies (1) by Theorem 1.

Corollary 2.1. *The following conditions are equivalent:*

(1) *A is either semi-simple Artinian or left and right self-injective strongly regular;*

(2) *A is a left non-singular left* CM *left I-injective ring.*

Quasi-injective left A-modules are I-injective. The proof of [7, Theorem 2.16] yields the following analogue of a well-known theorem of C. Faith—Y. Utumi concerning quasi-injective modules.

Theorem 3. *Let M be an I-injective left A-module, $E=\text{End}\,({}_AM)$, $J(E)=$ the Jacobson radical of E. Then $E/J(E)$ is von Neumann regular and $J(E)=\{f\in E\,|\ker f$ is essential in ${}_AM\}$.*

Recall that A is a left QI-ring iff each quasi-injective left A-module is injective [5]. Left QI-rings are left Noetherian left V-rings [5, p. 114]. ELT left QI-rings are then semi-simple Artinian by [21, Theorem 1.11].

The next proposition shows that, in general, a direct sum of I-injective left A-modules need not be I-injective.

Proposition 4. *The following conditions are equivalent:*

(1) *Each direct sum of I-injective left A-modules is I-injective;*

(2) *A is a left QI-ring and each I-injective left A-module is injective.*

Proof. Assume (1). Let M be an I-injective left A-module, $\hat{M}$ the injective hull of ${}_AM$. If $S={}_AM\oplus{}_A\hat{M}$, $j\colon M\to\hat{M}$ and $t\colon \hat{M}\to S$ are the inclusion maps, $u\colon M\to S$ the natural injection, $p\colon S\to M$ the natural projection, $i\colon M\to M$ the identity map, then i extends to an endomorphism h of ${}_AS$, since ${}_AS$ is I-injective. Hence $htj(m)=$ $=ui(m)$ for all $m\in M$, which implies that $htj=ui$ and hence $phtj=pui=i$. Thus $g=pht\colon \hat{M}\to M$ such that $gj=$the identity map on M which implies that ${}_AM$ is a direct summand of ${}_A\hat{M}$, whence $M=\hat{M}$ is injective. Since any quasi-injective left A-module is I-injective, therefore A is a left QI-ring and hence (1) implies (2).

(2) implies (1) by [5, Theorem 20.1].

It is well-known that A is left hereditary iff the sum of any two injective left A-modules is injective. The next corollary then follows.

Corollary 4.1. *If the sum of any two I-injective left A-modules is I-injective, then A is a left Noetherian, left hereditary, left V-ring.*

Since any direct sum of p-injective left A-modules is p-injective, then the proof of Proposition 4 yields

Remark 2. Suppose that every p-injective left A-module is I-injective. Then A is a left Noetherian ring whose p-injective left modules are injective.

Applying [5, Theorem 24.5] to Remark 2, we get

Remark 3. If A is a left p-injective ring whose p-injective left modules are I-injective, then A is quasi-Frobeniusean.

We now proceed to prove that a left non-singular left I-injective ring is left continuous regular. Recall that A is left continuous (in the sense of Utumi [17, p. 158]) iff every left ideal of A which is isomorphic to a complement left ideal is a direct summand of ${}_AA$.

Lemma 5. *Let M be an I-injective left A-module. K a complement left submodule of M. Then*

(1) *If N is a left submodule of M containing K, then any left A-homomorphism f of N into K extends to one of M into K;*

(2) *${}_AK$ is a direct summand of ${}_AM$.*

Proof. (1) The set of left submodules P of M containing N such that f extends to a left A-homomorphism of P into K has a maximal member U by Zorn's Lemma. Let $h\colon U\to K$ be the extension of f to U. If $j\colon K\to U$ is the inclusion map, then by hypothesis, jh extends to an endomorphism t of ${}_AM$. If $t(M)\nsubseteq K$, and D is a left submodule of M which is maximal with respect to $K\cap D=0$, then $\big(t(M)+K\big)\cap \cap D\neq 0$. If $0\neq d\in\big(t(M)+K\big)\cap D$, $d=t(m)+k$, $m\in M$, $k\in K$, then $t(m)=d-k\in \in D\oplus K$, $t(m)\notin K$ and therefore $m\notin U$. If $E=\{b\in M\,|\,t(b)\in D\oplus K\}$, then E strictly contains U. If p is the natural projection of $D\oplus K$ onto K, then $pt\colon E\to K$ extends f to E, which contradicts the maximality of U. This proves that t maps M into K and for any $n\in N$, $t(n)=jh(n)=h(n)=f(n)$.

(2) If C is a complement left ideal of A such that $K\oplus C$ is an essential left ideal, $p\colon K\oplus C\to K$ the natural projection, then by (1), p extends to a left A-homomorphism $g\colon M\to K$. Since $K\cap\ker g=0$, then for any $m\in M$, $m=g(m)+\big(m-g(m)\big)$, where $g(m)\in K$, $\big(m-g(m)\big)\in\ker g$, which proves that $M=K\oplus\ker g$.

If A is left I-injective, then A/Z is von Neumann regular (cf. the proof of Theorem 1). Consequently, Lemma 5(2) yields

Proposition 6. *If A is left non-singular, left I-injective, then A is left continuous regular.*

Corollary 6.1. *A left I-injective, left or right V-ring is left continuous regular.*

Corollary 6.2. *A left I-injective ring whose I-injective left modules are p-injective is left continuous regular.*

Applying [6, Theorem 16] to Proposition 6, we get

2

Corollary 6.3. *A semi-prime left I-injective ring which satisfies a polynomial identity is a left continuous regular, left and right V-ring.*

[16, Theorem 3] and a theorem of K. GOODEARL [5], Corollary 19.67] yield

Corollary 6.4. *A is primitive left self-injective regular iff A is prime left non-singular left I-injective.*

If M is a left A-module, N a left submodule of M, the usual closure of N in M is $Cl_M(N)=\{y\in M|Ly\subseteq N$ for some essential left ideal L of $A\}$. $Z(M)=Cl_M(0)$ is the singular submodule of M.

Proposition 7. *If A is left non-singular, then any quotient module Q of an I-injective left A-module contains its singular submodule Z(Q) as a direct summand.*

Proof. Let M be an I-injective left A-module, Q a quotient module of M, $f\colon M\to Q$ the canonical projection. Since $Z=0$, $Cl_M(\ker f)$ is a complement left submodule of ${}_AM$ and therefore $f^{-1}(Z(Q))=Cl_M(\ker f)$ is a direct summand of ${}_AM$ by Lemma 5(2). If $M=f^{-1}(Z(Q))\oplus N$, then $Q=f(M)=Z(Q)\oplus f(N)$.

2. Strongly regular rings

We now turn to characterizations of strongly regular rings.

Lemma 8. *The following conditions are equivalent:*

(1) *A is a division ring;*

(2) *A is a prime ring containing a non-zero reduced p-injective right ideal.*

Proof. Obviously (1) implies (2).

Assume (2). Let I be a non-zero reduced p-injective right ideal of A, $0\neq b\in I$, $i\colon bA\to I$ the inclusion map. Then there exists $c\in I$ such that $b=i(b)=cb$ and since I is reduced, $l(b)\subseteq r(b)$ which implies $A(1-c)\subseteq l(b)\subseteq r(b)$, whence $AbA(1-c)=0$. Since A is prime, therefore $1=c\in I$ which implies $A=I$ is a right p-injective integral domain. For any $0\neq c\in A$, if $f\colon cA\to A$ is the map $f(ca)=a$ for all $a\in A$, then there exists $d\in A$ such that $1=f(c)=dc$ which proves that (2) implies (1).

Lemma 9. *Let A be an* ELT *fully left idempotent ring. Then*

(1) *Any non-zero-divisor of A is invertible. Consequently, every left or right A-module is divisible;*

(2) *Any reduced principal left ideal is a direct summand of ${}_AA$;*

(3) *Any reduced principal right ideal is a direct summand of ${}_AA$.*

Proof. (1) Let c be a non-zero-divisor of A. If $Ac\neq A$, let M be a maximal left ideal containing Ac. If $M=l(e)$, where $e=e^2\in A$, then $ce=0$ implies $e=0$, whence

$M=A$, which is impossible. Therefore M is an essential left ideal and hence an ideal of A. Since A is fully left idempontent, $c=dc$ for some $d\in AcA\subseteqq M$ and then $1=d\in M$, again contradicting $M\neq A$. This proves that c is left invertible and since c is a non-zero-divisor, c is invertible in A. For any left A-module M, $M=cbM\subseteqq cM\subseteqq M$, where $cb=bc=1$, which yields $M=cM$. Similarly, any right A-module is divisible.

(2) Let $a\in A$ be such that Aa is reduced. Suppose that $Aa+l(a)\neq A$. If M is a maximal left ideal containing $Aa+l(a)$, and if $M=l(e)$, $e=e^2\in A$, then $e\in r(a)\subseteqq l(a)$ (because Aa is reduced) which implies $e=e^2=0$, contradicting $M\neq A$. Thus M is a maximal essential left ideal which is therefore an ideal of A. Since A is fully left idempotent, therefore A/M_A is flat [13, Lemma 2.3] which implies that $u\in Mu$ for all $u\in M$. In particular, $a=da$ for some $d\in M$ which yields $1-d\in l(a)\subseteqq M$, whence $1\in M$, again a contradiction. This proves that $Aa+l(a)=A$ and therefore $a=ca^2$ for some $c\in A$ and since Aa is reduced, $(a-aca)^2=0$ implies $a=aca$, whence Aa is a direct summand of ${}_AA$.

(3) Let $b\in A$ be such that bA is reduced and K a complement left ideal such that $L=Ab\oplus K$ is an essential left ideal. Then A/L_A is flat which implies $b=db$ for some $d\in L$, whence $b=bd$ (since bA is reduced). If $d=cb+k$, $c\in A$, $k\in K$, then $b-bcb=bk\in Ab\cap K=0$ which proves that bA is a direct summand of A_A.

Corollary 9.1. *If A is an* ELT *left V-ring, then* (a) *any non-zero-divisor is invertible;* (b) *any reduced principal left or right ideal is generated by an idempotent.*

Corollary 9.2. *If A is a prime* ELT *left idempotent ring, then A is either a division ring or a primitive ring with non-zero socle such that every non-zero left or right ideal contains a non-zero nilpotent element.*

Remark 4. If A is ELT fully left idempotent, then $J=Z=Y=0$.

Remark 5. [2, Corollary 6] holds for the following classes of rings: (1) ELT fully left idempotent rings; (2) Fully right idempotent rings whose essential right ideals are ideals; (3) Right I-injective rings.

Note that (a) rings whose essential left ideals are idempotent need not be semiprime (cf. for example, V. S. Ramamurthi and K. M. Rangaswamy, *Math. Scand.*, **31** (1972), 69—77); (b) reduced V-rings need not be regular (even when they are prime) [6, p. 109, Example 2].

Theorem 10. *The following conditions are equivalent:*

(1) *A is strongly regular;*

(2) *A is reduced such that any prime factor ring is left I-injective;*

(3) *A is regular such that every non-zero factor ring contains a non-zero reduced right ideal;*

(4) *A is left V-ring such that every non-zero factor ring contains a non-zero reduced p-injective right ideal;*

(5) *A is right V-ring such that every non-zero factor ring contains a non-zero reduced p-injective right ideal;*

(6) *Every non-zero factor ring of A is semi-prime containing a non-zero reduced p-injective right ideal;*

(7) *A is a reduced ring such that every non-zero factor ring contains a non-zero p-injective right ideal;*

(8) *A is an* ELT *reduced fully idempotent ring;*

(9) *A is a reduced* MELT *ring whose essential left ideals are idempotent;*

(10) *A is a reduced* MELT *ring whose essential right ideals are idempotent;*

(11) *A is an* ELT *fully idempotent ring whose proper prime ideals are completely prime.*

Proof. It is easy to see that (1) implies (2) through (5).

Assume (2). Let P be a proper prime ideal such that A/P is an integral domain. Then A/P is a division ring by Theorem 2 and (2) implies (6) by [8, Theorem 1.21].

Any one of (3), (4) or (5) implies (6).

Assume (6). Then A is a fully idempotent ring such that any non-zero prime factor ring is a division ring by Lemma 8. A is therefore strongly regular by [8. Corollary 1.18 and Theorem 3.2]. Thus (6) implies (7).

(7) implies (8) by [8, Theorem 1.21] and Lemma 8.

It is clear that (8) implies (9).

Assume (9). Let B be a prime factor ring of A, $0\neq b\in B$, $T=BbB$. Let K be a complement left subideal of T such that $L=Bb\oplus K$ is an essential left subideal of T. Since ${}_BT$ is essential in ${}_BB$, then so is ${}_BL$, whence $L=L^2$ (because every essential left ideal of B is idempotent). Now $b\in L^2$ implies $b=\sum_{i=1}^{n}(b_ib+k_i)(d_ib+c_i)$, where b_i, $d_i\in B$, k_i, $c_i\in K$, whence $b-\sum_{i=1}^{n}(b_ib+k_i)d_ib=\sum_{i=1}^{n}(b_ib+k_i)c_i\in Bb\cap K=0$ and therefore $b=\sum_{i=1}^{n}(b_ib+k_i)d_ib\in Tb=(Bb)^2$ which proves that B is fully left idempotent. If, further, B is an integral domain, then B is a division ring by Lemma 9(2) (because a MELT fully left idempotent ring is ELT). Thus (9) implies (10) by [8, Theorem 1.21].

Similarly, (10) implies (11).

Assume (11). If B is a non-zero prime factor ring of A, then B is an ELT fully idempotent domain which implies that B is a division ring. Consequently, (11) implies (1) by [8, Theorem 3.2].

Applying [16, Theorem 3] to Theorem 10(2), we get

Corollary 10.1. *If A is a left continuous regular ring such that any proper non-zero factor ring contains a non-zero reduced right ideal, then A is either left self-injective or right continuous strongly regular.*

Then Theorem 2 and Proposition 6 yield

Corollary 10.2. *If A is left non-singular left I-injective such that any proper non-zero factor ring contains a non-zero reduced right ideal, then A is left self-injective regular.*

We now consider rings having von Neumann regular centre. The centre of A will always be denoted by C. Rings whose simple left modules are either p-injective or flat need not be semi-prime (the converse is not true either).

Proposition 11. *Let A belong to any one of the following classes of rings:* (1) *A is semi-prime such that every essential left ideal is idempotent* (2) *A is such that each factor ring B satisfies one of the following conditions:* (a) *B is semi-prime;* (b) *The intersection of the Jacobson radical, the left singular ideal and the right singular ideal of B is zero;* (c) *Every simple left B-module is either p-injective or flat;* (3) *A is semi-prime such that for any non-zero element a of A, there exists a positive integer n such that Aa^n is a non-zero left annihilator. Then C, the centre of A, is von Neumann regular.*

Proof. (1) Let $c\in C$. If K is a complement left ideal such that $L=Ac\oplus K$ is an essential left ideal of A, then $c\in L^2=L$ and since $AcK\subseteq Ac\cap K=0$, $(KAc)^2=0$ implies $KAc=0$ (A being semi-prime), whence $c\in(Ac)^2+K^2$ which yields $c\in(Ac)^2$. Thus $c=cdc$ for some $d\in A$ and it follows from the proof of [18, Theorem 3] that $c=cvc$ for some $v\in C$.

(2) Suppose that $c\in C$ such that $c^2=0$.

(a) If A is semi-prime, then $(Ac)^2=Ac^2=0$ implies $c=0$.

(b) Let $J\cap Z\cap Y=0$. If K is a complement right ideal of A such that $R=r(c)\oplus K$ is an essential right ideal, then $Kc\subseteq Ac=cA\subseteq r(c)$ implies $ck=Kc\subseteq \subseteq r(c)\cap K=0$, whence $K\subseteq r(c)$ and therefore $K=0$, implying that $c\in Y$. Similarly, $c\in Z$. Also, for any $a\in A$, $(1+ac)(1-ac)=1$ which proves that $c\in J$. Thus $c\in J\cap \cap Z\cap Y=0$.

(c) Suppose that every simple left A-module is either p-injective or flat. If $c\neq 0$, M a maximal left ideal containing $l(c)$, then ${}_AA/M$ is either p-injective or flat. If ${}_AA/M$ is flat, the proof of Lemma 9(2) shows that we shall end with a contradiction. If ${}_AA/M$ is p-injective, the map $Ac\to A/M$ given by $ac\to a+M$ for all $a\in A$ leads again to a contradiction. Thus $c^2=0$ implies $c=0$ in (2) which proves that C (and hence the centre of any factor ring) is reduced. In particular, for any $u\in C$, $u+Au^2$ is a nilpotent element of the centre of A/Au^2 which implies $u\in Au^2$, whence $u=uvu$ for some $v\in C$.

(3) Since A is semi-prime, C is reduced (cf. (2)). If $0 \neq c \in C$, Ac^n is a non-zero left annihilator for some positive integer n. For any $b \in r(Ac^n)$, $(Acb)^n \subseteq Ac^n b = 0$ implies $b \in r(Ac)$ and hence $r(Ac^n) = r(Ac)$. Now $c \in l(r(Ac)) = l(r(Ac^n)) = Ac^n$. If $n > 1$, then $c = cac^{n-1}$, $a \in A$, which proves that Ac is a direct summand of ${}_AA$. Thus, whether $n = 1$ or $n > 1$, Ac is always a left annihilator for any non-zero $c \in C$. In particular, Ac^2 is a left annihilator and the preceding argument yields $c \in Ac^2$, whence $c = cvc$ for some $v \in C$.

Applying [1, Theorem 3] to Proposition 11, we get

Corollary 11.1. *Suppose that for each maximal ideal M of C, A/AM is regular. Then A is regular iff A satisfies any one of conditions* (1), (2), (3) *of Proposition* 11.

The proof of Proposition 11(2) and Corollary 11.1 yield

Proporisiton 12. *Suppose that A is semi-prime such that the centre C is not a field. Then A is regular iff for each non-zero ideal T of C, A/AT is regular.*

For any left A-module M, any left submodule N, write $K_M(N) = \{y \in M \mid cy \in N$ for some non-zero-divisor c of $A\}$. In general, $K_M(N) \neq Cl_M(N)$. If A has a classical left quotient ring, then $K_M(N)$ is a left submodule of M. Note that A has a classical left quotient ring iff A satisfies the left Ore condition (cf. for example [7, p. 101]). By [7, Theorem 3.34], the two "closures" $K_M(N)$ and $Cl_M(N)$ coincide over semi-prime left Goldie rings. To simplify the notation, write $K_A(I) = K(I)$ and $Cl_A(I) = Cl(I)$ for any left ideal I of A. If A is either left p-injective or a ring whose simple left modules are flat, then $K_M(N) = N$ for all left A-modules M and submodules N. Note that A is semi-simple Artinian iff $Cl_M(N) = N$ for all left A-modules M and submodules N.

Proposition 13. *The following conditions are equivalent:*

(1) *A is semi-simple Artinian;*

(2) *A is an* ELT *left hereditary left I-injective ring;*

(3) *A is an* ELT *fully left idempotent ring such that $K(I) = Cl(I)$ for any left ideal I of A;*

(4) *A is a left I-injective ring such that $K(I)$ is a complement left ideal for any left ideal I;*

(5) *A is semi-prime left I-injective satisfying the maximum condition on left annihilators;*

(6) *The direct sum of a projective and an I-injective left A-modules is I-injective.*

Proof. Obviously, (1) implies (2) through (6).

Since a well-known result of B. Osofsky asserts that a left self-injective left hereditary ring is semi-simple Artinian, (2) implies (1) by Theorem 1.

Assume (3). By Lemma 9(1), A is its own classical left quotient ring. Since a semi-prime ELT ring is left non-singular, then $K(I)=Cl(I)$ is a complement left ideal for any left ideal I. In particular, if L is an essential left ideal, $K(L)=A$ which implies that L contains a non-zero-divisor c. By Lemma 9(1), c is invertible in A which yields $L=A$. This proves that (3) implies (1).

Similarly, (4) implies (1) by Remark 1(a).

(5) implies (1) by Proposition 6.

Assume (6). If P is a projective left A-module, H the injective hull of ${}_AP$, then $P\oplus H$ is a left I-injective A-module and the proof of Proposition 4 shows that ${}_AP$ is injective. Therefore every injective left A-module is projective by [5, Theorem 24.20] and from Proposition 4, every I-injective left A-module is injective which implies that every simple left A-module is projective. Thus (6) implies (1).

Remark 6. The following conditions are equivalent for a left CM-ring A: (1) is semi-prime left Goldie; (2) For any left A-module M and every left submodule N, $K_M(N)=Cl_M(N)$; (3) Every essential left ideal of A contains a non-zero-divisor.

We add a last remark on rings whose essential left ideals are idempotent.

Remark 7. Suppose that every essential left ideal of A is idempotent. If A is either ELT or left CM, then the centre of A is von Neumann regular.

The referee has kindly drawn my attention to the following papers:

(1) V. E. Govorov, Semi-injective modules, *Algebra i Logika,* **2** (1963), 21—49.

(2) A. A. Tuganbaev, Quasi-injective and weakly injective modules, *Bull. Moscow Univ. Math. Mech.,* Series № 2 (1977).

Acknowledgement. I am extremely grateful to the referee for many helpful comments and suggestions, in particular, for the present version of Proposition 4.

References

[1] E. P. Armendariz, J. W. Fisher and S. A. Steinberg, Central localizations of regular rings, *Proc. Amer. Math. Soc.,* **46** (1974), 315—321.

[2] G. F. Birkenmeier, Idempotents and completely semi-prime ideals, *Comm. Algebra,* **11** (1983), 567—580.

[3] K. Chiba and H. Tominaga, On strongly regular rings, *Proc. Japan Acad.,* **49** (1973), 435—437.

[4] K. Chiba and H. Tominaga, On strongly regular rings. II, *Proc. Japan Acad.,* **50** (1974), 444—445.

[5] C. Faith, *Algebra II. Ring Theory,* Grundlehren der math. Wissenschaften, Band 191, Springer-Verlag (Berlin, 1976).

[6] J. W. Fischer, Von Neumann regular rings versus V-rings, in: *Ring Theory* (Proc. Oklahoma Conference), Lecture note № 7, Dekker (New York, 1974), pp. 101—119.

[7] K. R. Goodearl, *Ring Theory. Non-singular rings and modules*, Pure and Appl. Math. № 33, Dekker (New York, 1976).

[8] K. R. Goodearl, *Von Neumann regular rings*, Monographs and studies in Math. 4, Pitman (London, 1979).

[9] Y. Hirano and H. Tominaga, Regular rings, V-rings and their generalizations, *Hiroshima Math. J.*, **9** (1979), 137—149.

[10] S. K. Jain, S. H. Mohamed and S. Singh, Rings in which every right ideal is quasi-injective, *Pacific J. Math.*, **31** (1969), 73—79.

[11] S. Lajos and F. Szász, Characterizations of strongly regular rings, *Proc. Japan Acad.*, **46** (1970), 38—40.

[12] S. Lajos and F. Szász, Characterizations of strongly regular rings. II, *Proc. Japan Acad.*, **46** (1970), 287—289.

[13] G. D. Michler and O. E. Villamayor, On rings whose simple modules are injective, *J. Algebra*, **25** (1973), 185—201.

[14] H. Tominaga, On s-unital rings, *Math. J. Okayama Univ.*, **18** (1976), 117—134.

[15] H. Tominaga, On s-unital rings. II, *Math. J. Okayama Univ.*, **19** (1977), 171—182.

[16] Y. Utumi, On continuous regular rings and semi-simple self-injective rings, *Canad. J. Math.*, **12** (1960), 597—605.

[17] Y. Umuti, On continuous rings and self-injective rings, *Trans. Amer. Math. Soc.*, **118** (1965), 158—173.

[18] J. von Neumann, On regular rings, *Proc. Nat. Acad. Sci.*, **22** (1936), 707—713.

[19] R. Yue Chi Ming, On V-rings and prime rings, *J. Algebra*, **62** (1980), 13—20.

[20] R. Yue Chi Ming, On von Neumann regular rings. V, *Math. J. Okayama Univ.*, **22** (1980), 151—160.

[21] R. Yue Chi Ming, On regular rings and self-injective rings, *Monatsh. Math.*, **91** (1981), 153—166.

UNIVERSITE PARIS VII
U. E. R. DE MATHEMATIQUE ET INFORMATIQUE
2, PLACE JUSSIEU
75251 PARIS CEDEX 05, FRANCE

Acta Sci. Math. 47 (1984), 289—295

Classification of finite minimal non-metacyclic groups

MARIO CURZIO

Dedicated to the memory of Carlo Miranda

Let (MC) be the class of *metacyclic* groups $\big(G\in(\mathrm{MC})$ if and only if it has a cyclic factor group G/H with H cyclic$\big)$. A group G is said to be *minimal non-metacyclic* $\big($*minimal non-*(MC)$\big)$ if and only if $G\notin(\mathrm{MC})$ and $H\in(\mathrm{MC})$ for every subgroup $H<G$.

N. Blackburn [1] (Theorem 3.2) determined all finite minimal non-(MC) p-groups (p prime). In the present paper we construct all other finite minimal non-(MC) groups (see Theorems 1.2, 2.7, 2.8 and 2.10). They are generally monomial (see 2.12), and for the set $\pi(G)$ of all prime divisors of the order of G we have $|\pi(G)|=$ $=2$, 3 (see 2.11). Moreover, every $G_p\in\mathrm{Syl}_p(G)$ $\big(p\neq\min\pi(G)\big)$ is either cyclic, or of order p^2 and exponent p. Finally, the metacyclic p-group G_p is rather general for $p=\min\pi(G)$.

All groups we shall deal with are finite.

Notation and terminology are the usual ones in group theory (see for instance [3], [6] and [7]). We just point out that $G/\mathscr{L}(G)$ will denote the largest nilpotent factor group of G, Q_8 the quaternion group.

1. A minimal non-(MC) group is either supersolvable or minimal non-supersolvable. In this section we shall determine the structure of non-supersolvable and minimal non-(MC) groups.

1.1. *Let G be a non-supersolvable and minimal non-*(MC) *group, and G_p its normal Sylow subgroup*[1]*. Then:*

(1) *if $\Phi(G)=1$, then G is minimal non-abelian and its order is p^2q (q prime);*

(2) *$G=G_pG_q$ with G_q cyclic and $\Phi(G_q)\lhd G$;*

(3) *if $p>2$, then $|G_p|=p^2$ and $\exp G_p=p$;*

(4) *if $p=2$, then either $G_2\cong Q_8$ or $|G_2|=4$ and $\exp G_2=2$.*

[1]) A minimal non-supersolvable group has a unique normal Sylow subgroup (see [2], Hilfssatz C).

Proof. (1) Suppose $\Phi(G)=1$. G_p is the only minimal normal subgroup of G (see [2], Satz 1a); it is elementary abelian, not cyclic and metacyclic, hence $|G_p|=p^2$. In contrary to our claim, assume there is a $G_{p'}$ of composite order; then there exists a subgroup $H=G_pM\lhd G$ with $1<M<G_{p'}$. If H' is cyclic and non-trivial, there is an $N\leqq H'$, minimal normal in G, so $N=G_p$, which is a contradiction, as G_p is not cyclic. If $H'=1$, then $H=G_p\times M$, and G has a minimal normal subgroup $N\leqq M$, hence $N\neq G_p$, again a contradiction.

(2) By (1) we have $G=G_pG_q$ and $|G/\Phi(G)|=p^2q$. The Sylow q-subgroup S of $\Phi(G)$ is maximal in some G_q and normal in G. If $x\in G_q-S$, then $G=\langle G_p, S, x\rangle=$ $=G_p\langle x\rangle$ and G_q is cyclic; thus $\Phi(G_q)=S\lhd G$.

(3) Suppose $p>2$, hence (see [2], Satz 1f) $\exp G_p=p$. Since G_p is metacyclic of order greater than p, it follows that $|G_p|=p^2$.

(4) Suppose $p=2$, hence (see [2], Satz 1f) $\exp G_2\leqq 4$. If G_2 is abelian, then by (2) G is minimal non-abelian, so $\exp G_2=2$, whence $|G_2|=4$. Suppose now G_2 is not abelian, hence $\exp G_2=4$; G being metacyclic, it follows that either $|G_2|=8$ or $|G_2|=16$; whether the latter case occurs or G_2 is dihedral (of order 8), $|\mathrm{Aut}(G_2)|$ is a power of 2, thus $G=[G_2]G_q$, whence the contradiction $G=G_2\times G_q$. This proves that $G_2\cong Q_8$.

Theorem 1.2. *A non-supersolvable group G is minimal non-*(MC) *if, and only if, one of the following holds:*

(a) *G is minimal non-abelian of order p^2q^n ($p\neq q$ primes, $G_p\lhd G$);*

(b) *$G=\langle Q_8, x\rangle$, where $|x|=3^n$ and x induces on Q_8 an automorphism of order* 3.

Proof. Let G be a minimal non-(MC) group. By 1.1, $G=[G_p]G_q$; if G_p is abelian, then (a) holds. If G_p is not abelian, by 1.1 we get $G_p=G_2\cong Q_8$, and (b) holds.

2. Minimal non-(MC) p-groups were classified by Blackburn [1]. In this section we construct all other supersolvable minimal non-(MC) groups (see Theorems 2.7, 2.8 and 2.10).

2.1. *Let $G=MN$ be a metacyclic p-group ($p>2$) with $M\neq 1$, $N\neq 1$ subgroups such that $M\cap N=1$. Then both M and N are cyclic.*

Proof. G is modular (see [8], Proposition 1.8), so $\Omega_1=\{x\in G|x^p=1\}$ is a metacyclic p-group of exponent p; then $|\Omega_1|=p^2$, whence the assertion follows.

2.2. *Let $G=[A]B$ with A cyclic of odd order, B nilpotent, $|\pi(G)|>1$, and suppose each $H<G$ is metacyclic. Then*

(1) *$B_p\in\mathrm{Syl}(B)$ is cyclic for any $p\in\pi(A)\cap\pi(B)$;*

(2) *G is metacyclic if $|\pi(B)|>1$;*

(3) *G is metacyclic if $|\pi(A)|>2$.*

Proof. (1) follows from 2.1.

(2) Suppose $|\pi(B)|>1$ and let $p\in\pi(B)-\pi(A)$. The subgroup $K=AB_p<G$ is metacyclic, so there exists a cyclic subgroup $C\lhd K$ with cyclic factor K/C. The Sylow p-subgroup C_p of C is normal in G and B_p/C_p is cyclic. By (1) we have $B=(\bigtimes_{p\nmid|A|} B_p)\times T$, with T cyclic. The normal subgroup $H=A\times(\bigtimes_{p\nmid|A|} C_p)$ is cyclic, as is $G/H\cong T\times(\bigtimes_{p\nmid|A|} B_p/C_p)$.

(3) Assume $|\pi(A)|>2$. After (2) we may suppose B is a p-group. If p divides $|A|$, then B is cyclic (see (1)) and G is metacyclic. So let p be relatively prime to $|A|$. Since $|\pi(A)|>2$, we get that $A=R\times S\times T$, with R, S and T non-trivial Hall subgroups. $K=(R\times S)B<G$ is metacyclic, so there exists $N\lhd G$ cyclic with K/N cyclic. $N_p\in\mathrm{Syl}\,(N)$ is contained in $C_B(R)\cap C_B(S)$; moreover, B/N_p is a cyclic p-group, hence $C_B(R)$ and $C_B(S)$ are comparable. Arguing as before, we see that $C_B(R)$, $C_B(S)$ and $C_B(T)$ are pairwise comparable; assuming $C_B(R)$ is the smallest one, N_p centralizes $R\times S\times T=A$. The normal subgroup $H=A\times N_p$ is cyclic, as is $G/H\cong B/N_p$.

2.3. *Suppose G has a modular subgroup $G_p\in\mathrm{Syl}\,(G)$ with $p>2$. Then $G_p\cap\cap Z(G)\cap\mathscr{L}(G)=1$.*

Proof. See Huppert [5], 3.2. Satz.

Lemma 2.4. *Let G be a supersolvable group such that for each $p\neq\min\pi(G)$, $G_p\in\mathrm{Syl}\,(G)$ is modular. Then $G=[\mathscr{L}(G)]M$, with M a system normalizer.*

Proof. Let $q=\min\pi(G)$. Then $(G_q)'$ is normal in G (see [4], Satz 4) and $G/(G_q)'$ has abelian Sylow q-subgroups. Now apply the following theorem of Huppert [5] (3.3. Satz): For a solvable group G with each G_p modular and each G_2 abelian, $G=[\mathscr{L}(G)]M$, with M a system normalizer. Thus we have $G/(G_q)'=[A/(G_q)']M/(G_q)'$ with $A/(G_q)'=\mathscr{L}(G/(G_q)')$ and $M/(G_q)'$ a system normalizer. $A/(G_q)'\leq(G/(G_q)')'$ is nilpotent and its order is relatively prime to q. As G is supersolvable, M is nilpotent and $A=B\times(G_q)'$ with B nilpotent and $|B|$ relatively prime to q. We have $G=AM=(B\times(G_q)')M=[B]M$, and $\mathscr{L}(G)\leq B$; on the other hand, $G/(\mathscr{L}(G)\times\times(G_q)')$ is nilpotent, hence $\mathscr{L}(G)\times(G_q)'\geq A=B\times(G_q)'$. From this we get $B=\mathscr{L}(G)$. In a similar way the assertion about M can also be proved.

2.5. *Let G be a non-primary, supersolvable and minimal non-*(MC) *group. Then either $\mathscr{L}(G)=G_p$ ($|G_p|\leq p^2$ and $\exp G_p=p$), or $\mathscr{L}(G)=G_p\times G_q$ ($|G_p|=p$ and $|G_q|=q$).*

Proof. A metacyclic p-group of odd order is modular (see [8], Proposition 1.8) hence $G=[\mathscr{L}(G)]M$ (by Lemma 2.4) and $\mathscr{L}(G)\leq G'$ is nilpotent of odd order. Suppose there is a non-abelian $\mathscr{L}_p\in\mathrm{Syl}\,(\mathscr{L}(G))$. If $\mathscr{L}_p<G_p$, then $G_p=\mathscr{L}_pM_p<$

$<G$ is metacyclic, hence $\mathscr{L}_p$ is cyclic (see 2.1), a contradiction. If $\mathscr{L}_p=G_p$, there exists a non-trivial factor G/H which is a p-group (see [5], 1.5. Satz, saying that for a group G with $G_p\in\mathrm{Syl}(G)$ of odd order, metacyclic and not abelian, there is a non-trivial factor group G/H which is a p-group); this is again a contradiction, as $G_p=\mathscr{L}_p\leqq\mathscr{L}(G)$. Thus $\mathscr{L}(G)$ is abelian, and we can consider the following cases:

(i) There exists $\mathscr{L}_p\in\mathrm{Syl}\,(\mathscr{L}(G))$ having a socle of order p^2. Arguing by contradiction, let $S<\mathscr{L}(G)$. If $K=SM$, $\mathscr{L}(K)\leqq K'$ is cyclic, and K splits on $\mathscr{L}(K)$. On the other hand, $\mathscr{L}(K)\leqq S$ hence either $K=S\times M$, or $K=[N_1](N_2\times M)$ with $|N_i|=p$. In both cases, G has a non-trivial central subgroup which is contained in $G_p\cap Z(G)\cap\mathscr{L}(G)$, contradicting 2.3. Then we have $|\mathscr{L}(G)|=p^2$ and $\exp\mathscr{L}(G)=p$. Assume $\mathscr{L}(G)<G_p$; then $[\mathscr{L}(G)]M_p$ should be metacyclic, hence $\mathscr{L}(G)$ cyclic. which contradicts the hypothesis. Thus $\mathscr{L}(G)=G_p$.

(ii) $\mathscr{L}(G)$ is cyclic. We have $|\pi(\mathscr{L}(G))|\leqq 2$ and $|\pi(M)|=1$ (see 2.2), hence either $\mathscr{L}(G)=G_p$, or $\mathscr{L}(G)=G_p\times G_q$. In both cases, let $P\leqq G_p$ be of order p. $K=[G_p]C_M(P)$ splits on $\mathscr{L}(K)\leqq G_p$, and either $\mathscr{L}(K)=G_p$ or $\mathscr{L}(K)=1$, since G_p is cyclic. In the first case $1\neq P\leqq K_p\cap\mathscr{L}(K)\cap Z(K)$, which contradicts 2.3. Hence $K=G_p\times C_M(P)$ and $C_M(P)=C_M(G_p)$.

Suppose now $\mathscr{L}(G)=G_p$ and, by contradiction, $\mathscr{L}(G)>P$. The subgroup $PM<G$ is metacyclic, so there exists a cyclic subgroup $X\leqq C_M(P)$, $X\lhd G$ with M/X cyclic; but $C_M(P)=C_M(G_p)$ and G is metacyclic, a contradiction.

Suppose finally $\mathscr{L}(G)=G_p\times G_q$ and consider $P\leqq G_p$ and $Q\leqq G_q$ of order p and q respectively; as before, $C_M(P)=C_M(G_p)$ and $C_M(Q)=C_M(G_q)$. If $(P\times Q)M$ were metacyclic, there should exist a cyclic subgroup $X\lhd M$, with cyclic factor M/X, such that $X\leqq C_M(P\times Q)=C_M(G_p)\cap C_M(G_q)$; then G should be metacyclic.

This proves that $P\times Q=\mathscr{L}(G)$.

2.6. *Let G be a supersolvable minimal non-*(MC) *group and suppose* $|\pi(G)|=3$. *Then:*

(1) *either* $G=(G_p\times G_q)G_r$ *with* $\mathscr{L}(G)=G_p\times G_q$ *and* $|\mathscr{L}(G)|=pq$, *or* $G=G_p(G_q\times G_r)$ *with* $\mathscr{L}(G)=G_p$ $(|G_p)|=p^2$ *and* $\exp G_p=p)$;

(2) *if* $\mathscr{L}(G)=G_p\times G_q$, *then* $G_r=M_1M_2$ $(M_i\lhd G$, M_i *cyclic*, $M_1<C_G(G_p)$ *and* $M_2<C_G(G_q))$;

(3) *if* $\mathscr{L}(G)=G_p\times G_q$, *then* $C_{G_r}(G_p)$ *and* $C_{G_r}(G_q)$ *are maximal subgroup of* G_r.

Proof. (1) By 2.5, either $\mathscr{L}(G)=G_p\times G_q$ $(|\mathscr{L}(G)|=pq)$, or $\mathscr{L}(G)=G_p$ $(|G_p|\leqq p^2$ and $\exp G_p=p)$. If $|\mathscr{L}(G)|=p$, from 2.2 we would have $|\pi(G)|=2$, a contradiction; now (1) readily follows.

(2) Let $\mathscr{L}(G)=G_p\times G_q$. G_pG_r and G_qG_r are metacyclic; then there are cyclic subgroups $M_i\lhd G_r$, with cyclic factor groups G_r/M_i, such that $M_1<C_G(G_p)$ and $M_2<C_G(G_q)$. Arguing by contradiction, suppose $M_1M_2<G_r$. $(G_p\times G_q)M_1M_2<G$ is metacyclic; then we can find a cyclic subgroup $X\lhd M_1M_2$ with M_1M_2/X

cyclic and $X<C_G(G_p\times G_q)$; M_1M_2/X is primary, so that one can suppose $M_2X\geqq\geqq M_1X\geqq M_1$, thus $M_1<C_G(G_p\times G_q)$ and G is metacyclic.

(3) After (2), assuming $\mathscr{L}(G)=G_p\times G_q$, one has $G_r=M_1C_{G_r}(G_q)$. Denoting by N_1 the maximal subgroup of M_1, suppose $N_1\nleqq C_{G_r}(G_q)$. Then $N_1C_{G_r}(G_q)<G_r$, so that $(G_p\times G_q)N_1C_{G_r}(G_q)$ is metacyclic; hence there is a cyclic subgroup $X\leqq\leqq C_{G_r}(G_p\times G_q)$, normal in $N_1C_{G_r}(G_q)$ with primary cyclic quotient. Thus either $C_{G_r}(G_q)=XC_{G_r}(G_q)\geqq XN_1\geqq N_1$, or $XN_1\geqq XC_{G_r}(G_q)=C_{G_r}(G_q)\geqq M_2$. In the first case the contradiction is clear. In the second case we get $M_2\leqq C_{G_r}(G_p\times G_q)$ and G is metacyclic, again a contradiction.

Theorem 2.7. *Let G be a supersolvable group with $|\pi(G)|=3$. Then G is a minimal non-*(MC) *group if and only if it has one of the following structures:*

(a) $G=[G_p\times G_q]G_r$, *where* $|G_pG_q|=pq$, $G_r=M_1M_2$ ($M_i\lhd G_r$, M_i *cyclic*), $C_{G_r}(G_p)\geqq M_1$ *and* $C_{G_r}(G_q)\geqq M_2$ *are maximal subgroups of* G_r;

(b) $G=G_p(G_q\times G_r)$, *where* $G_q\times G_r$ *is cyclic,* $G_p=N_1\times N_2$ ($N_i\lhd G$ *and* $|N_i|=p$), $G_q<C_G(N_1)$, $\Phi(G_q\times G_r)<C_G(G_p)$, N_1G_r *and* N_2G_q *are non-abelian.*

Proof. Assume G is a minimal non-(MC) group. Then either $\mathscr{L}(G)=G_p\times G_q$ and $|\mathscr{L}(G)|=pq$, or $G=G_p(G_q\times G_r)$ with $\mathscr{L}(G)=G_p$ of order p^2 and exponent p (see 2.7 (1)). In the first case, (a) holds (see 2.6 (2) and (3)). Let us look at the second possibility. We have $G_p=N_1\times N_2$ with N_i minimal normal in G. $(N_1\times N_2)G_q<G$ has a cyclic commutator subgroup, so G_q centralizes only one of the N_i's. Indeed, were $G_q<C_G(N_1\times N_2)$, $G=G_q\times G_pG_r$ would be metacyclic since G_q and G_pG_r are metacyclic and of coprime orders. Suppose G_p centralizes N_1. We cannot have $G_r<<C_G(N_1)$ for this implies $G=N_2(G_q\times G_r)\times N_1$, which contradicts the meaning of $\mathscr{L}(G)=N_1\times N_2$. Thus $G_r<C_G(N_2)$. Neither G_q nor G_r centralizes $G_p=N_1\times N_2$, hence $x\notin C_G(N_2)$ and $y\notin C_G(N_1)$ for suitable $x\in G_q$ and $y\in G_r$. $\langle G_p; x, y\rangle$ has a non-cyclic commutator subgroup; hence it coincides with G; so $G_q=\langle x\rangle$ and $G_r=\langle y\rangle$. Denoting by M the maximal subgroup of $\langle y\rangle$, $(N_1\times N_2)(\langle x\rangle\times M)$ has a cyclic commutator subgroup, thus $M<C_G(G_p)$; similarly, the maximal subgroup of $\langle x\rangle$ centralizes G_p, so G is like in (b).

Vice versa, if (b) holds, G is clearly minimal non-(MC). Assume (a) holds. G is not metacyclic, since, modulo $G_{G_r}(G_p)\cap C_{G_r}(G_q)$, G_r is not cyclic. Suppose now $M<G$ is a maximal subgroup. If $(G:M)=q$, $M=G_pG_r$ is metacyclic as $G_p\times M_1$ and $M/(G_p\times M_1)\cong G_r/M_1$ are cyclic; similarly M turns out to be metacyclic when $(G:M)=p$. Finally, suppose $(G:M)=r$, so that $M=(G_p\times G_q)X$ with X maximal in G_r. We can assume $M_1\nleqq X$, since $G_r=M_1M_2$. Then $M_1\cap X\leqq C_{G_r}(G_p)$ is the maximal subgroup of M_1 and we also get $M_1\cap X\leqq C_{G_r}(G_q)$, since $G_r=M_1C_{G_r}(G_q)$ with M_1 cyclic and $C_{G_r}(G_q)$ maximal in G_r, implying that the maximal subgroup of M_1 is contained in $C_{G_r}(G_q)$. Hence it follows that $H=(G_p\times G_q)\times(X\cap M_1)$ is cyclic, as is $M/H\cong X(M_1\cap X)\cong G_r/M_1$.

Theorem 2.8. *Let G be a supersolvable group with* $|\pi(G)|=2$ *and* G_p *not cyclic* $(p=\max\pi(G))$. *Then G is minimal non-*(MC) *if and only if it has the following structure:* $G=(N_1\times N_2)G_q$, *where is G_q cyclic,* $N_i\lhd G$ *and* $|N_i|=p$, $\Phi(G_q)<C_G(N_1)$, N_1G_q *and* N_2G_q *are non-abelian.*

Proof. A group with the above structure is clearly minimal non-(MC).

Vice versa, suppose G is minimal non-(MC). By 2.5 we have $\mathcal{L}(G)=G_p=N_1\times N_2$ ($N_i\lhd G$ and $|N_i|=p$), $G=G_pG_q$. If G_q centralizes N_1, then $G=N_1\times N_2G_q$, which contradicts the meaning of $\mathcal{L}(G)=N_1\times N_2$; similarly, $G_q\nleqq C_G(N_2)$.

Let M be a maximal subgroup of G_q; the commutator subgroup of $(N_1\times N_2)M<G$ is cyclic, so it centralizes at least one of the N_i's. If G_q were not cyclic, there should be at least three maximal subgroups in G_q, hence two maximal subgroups of G should centralize the same N_i (for instance N_1); hence we get the contradiction $G_q<C_G(N_1)$.

Definition 2.9. Let G_p be a group of prime order $p>2$, G_q a q-group (q prime), metacyclic with a subgroup $C\lhd G_q$ such that G_q/C is a cyclic and $|G_q/C|$ divides $p-1$. Moreover, suppose there is no cyclic quotient G_q/X with X cyclic and $X\leqq C$, while for every maximal subgroup $M<G_q$ there exists a cyclic factor M/X_M with X_M cyclic and $X_M\leqq C\cap M$. Under these hypotheses, there exists an homomorphism $\alpha: G_q\to \operatorname{Aut} G_p$ such that $\operatorname{Ker}\alpha=C$. We shall call the semidirect product $G=[G_p]G_q$ (determined by α) a *group of type G_α*.

An easy example of such a group can be obtained in the following way. Let us denote by G_2 the dihedral group of order 8 and by G_p a group of prime order $p>2$. Then for any maximal non-cyclic subgroup $C<G_2$, the hypotheses of Definition 2.9 hold, hence the semidirect product $G=[G_p]G_2$ determined by the homomorphism $\alpha: G_2\to\operatorname{Aut} G_p$ with kernel C is of type G_α.

Remark. Let $G_q\ncong Q_8$ be a metacyclic non-abelian q-group (q prime). With standard calculations (omitted here for the sake of brevity) we can prove the existence of a subgroup $C\lhd G_q$ such that: G_q/C is cyclic and there is no cyclic quotient G_q/X with X cyclic and $X\leqq C$, while for every maximal subgroup $M<G_q$ there is a cyclic factor M/X_M with X_M cyclic and contained in $C\cap M$. From this it follows that in Definition 2.9 the q-Sylow subgroups of G can be almost arbitrary.

We thank Mercede Maj for this remark.

Theorem 2.10. *Let G be a supersolvable group with* $|\pi(G)|=2$ *and G_p cyclic* $(p=\max\pi(G))$. *Then G is a minimal non-*(MC) *group if and only if it is of type G_α.*

Proof. Let G be minimal non-(MC). By 2.5, $\mathcal{L}(G)=G_p$ and $|G_p|=p$; $G=G_pG_q$ is of type G_α (see Definition 2.9), where $C=C_{G_q}(G_p)$.

The converse statement is trivial.

2.11. *Let G be a supersolvable and minimal non-*(MC) *group. Then* $|\pi(G)|\leqq 3$.

Proof. If $\mathscr{L}(G)$ is cyclic, the statement follows from 2.2 and 2.4. Assume $\mathscr{L}(G)$ is not cyclic; then (see 2.5) $G=[\mathscr{L}(G)]M$ and $\mathscr{L}(G)=G_p=N_1\times N_2$ ($N_i \lhd G$ and $|N_i|=p$). Arguing by contradiction, suppose $M=A\times B\times C$ with A, B and C non-trivial Hall subgroups. The commutator subgroup of $(N_1\times N_2)(A\times B)$ is cyclic; hence we can assume $A\times B\leqq C_G(N_1)$. Similarly, either $A\times C< C_G(N_2)$ or $A\times C< C_G(N_1)$, whence either $G=A\times G_p(B\times C)$, or $C=N_1\times N_2M$; in the first case G is metacyclic, since A and $G_p(B\times C)$ are metacyclic of coprime orders; in the second case we get a contradiction to the meaning of $\mathscr{L}(G)=N_1\times N_2$.

By 2.11, Theorems 1.2, 2.7, 2.8 and 2.10 characterize the non-primary and minimal non-metacyclic groups; thus the theory of group extentions allows us to give an effective construction of these groups. Furthermore:

2.12. *Let G be a minimal non-*(MC) *group, without any normal $G_2\in \mathrm{Syl}(G)$ isomorphic to Q_8. Then any irreducible representation of G over an algebraically closed field K such that* $\mathrm{ch}\, K\not| |G|$ *is monomial.*

Proof. G is either supersolvable or metabelian (see Theorem 1.2), hence the assertion is an immediate consequence of the following well-known result by Huppert [6] (V. 18.4. Satz): Every solvable group G having a supersolvable quotient G/H such that H has abelian Sylow subgroups is monomial.

References

[1] N. Blackburn, Generalizations of certain elementary theorems on p-groups, *Proc. London Math. Soc.*, **11** (1961), 1—22.

[2] K. Doerk, Minimal nicht überauflösbare Gruppen, *Math. Z.*, **91** (1966), 198—205.

[3] D. Gorenstein, *Finite groups*, Harper and Row (New York, 1968).

[4] B. Huppert, Normalteiler und maximale Untergruppen endlicher Gruppen, *Math. Z.*, **60** (1954), 409—434.

[5] B. Huppert, Gruppen mit modularen Sylowgruppen, *Math. Z.*, **75** (1961), 140—153.

[6] B. Huppert, *Endliche Gruppen. I*, Springer (Berlin, 1967).

[7] D. J. S. Robinson, *Finiteness conditions and generalized soluble groups. I—II*, Springer (Berlin, 1972).

[8] M. Suzuki, *Structure of a group and the structure of its lattice of subgroups*, Springer (Berlin, 1956).

ISTITUTO MATEMATICO DELL'UNIVERSITÀ
VIA MEZZOCANNONE, 8
80 134 NAPOLI, ITALY

Acta Sci. Math. 47 (1984), 297—301

Многообразия квазиортодоксальных полугрупп

Н. Г. ТОРЛОПОВА

Регулярная полугруппа называется ортодоксальной, если множество всех ее идемпотентов образует подполугруппу в ней. Произвольную полугруппу с таким же свойством множества всех ее идемпотентов назовем квазиортодоксальной. Многообразие полугрупп V назовем квазиортодоксальным, если каждая полугруппа из V квазиортодоксальна.

В настоящей работе дан критерий, позволяющий по совокупности тождеств Φ, задающей многообразие полугрупп V, выяснить, является ли V квазиортодоксальным. Кроме того, описаны минимальные неквазиортодоксальные многообразия полугрупп, т.е. такие неквазиортодоксальные многообразия, каждое собственное подмногообразие которых является квазиортодоксальным.

Все необходимые сведения из теории полугрупп можно найти в [1] и [5].

1. Через S_0 обозначим следующую четырехэлементную полугруппу:

	1	2	3	0
1	1	3	3	0
2	0	2	0	0
3	0	3	0	0
0	0	0	0	0

Через S_p обозначим вполне простую полугруппу над циклической группой простого порядка p с матрицей

$$\begin{pmatrix} 1 & g \\ 1 & 1 \end{pmatrix},$$

где g — образующий элемент группы, 1 — ее единица.

2. Предложение 1. *Если многообразие полугрупп V не является квазиортодоксальным, то либо $S_p \in V$ для некоторого простого p, либо $S_0 \in V$.*

Поступило 10 ноября 1982, в переработанной форме 10 марта 1983.

Доказательство. Так как V не является квазиортодоксальным, то V содержит полугруппу S, порожденную идемпотентами e и f, такими, что $ef\neq(ef)^2$. Рассмотрим идеал I полугруппы S, порожденный элементами efe и fef. Возможен один из случаев: 1) $ef\in I$ и $fe\in I$; 2) $ef\notin I$ или $fe\notin I$.

Используя результаты работы [2], заключаем, что в первом случае существует натуральное число k, такое, что $ef=(ef)^k$, $fe=(fe)^k$, $k>2$ и k — наименьшее из всех чисел с таким свойством. Обозначим через $H_{e,e}=\{(ef)^\alpha e\}$, $H_{f,e}=\{(fe)^\alpha\}$, $H_{e,f}=\{(ef)^\alpha\}$, $H_{f,f}=\{(fe)^\alpha f\}$, где $\alpha=1, 2, \ldots, k-1$. Каждое из этих множеств есть, очевидно, подгруппа в I; $H_{i,j}\cap H_{m,l}=\emptyset$ при $i\neq m$ или $j\neq l$ $(i, j, m, l\in\{e, f\})$, а поэтому $I=\bigcup\limits_{i,j\in\{e,f\}} H_{i,j}$ есть прямоугольная связка групп, а значит (см. [5], стр. 114), I — вполне простая полугруппа. Обозначим через $\Lambda=\{e, f\}$, через G циклическую группу порядка $k-1$: $G=\{a, a^2, \ldots, a^{k-1}=1\}$. Тогда $I\cong\mathcal{M}(G, \Lambda, \Lambda, P)$, где $P=\begin{pmatrix}1 & a\\ 1 & 1\end{pmatrix}$, $\mathcal{M}(G, \Lambda, \Lambda, P)$ — регулярная рисовская полугруппа матричного типа над группой G.

Так как $k-1>1$, то существует простое число p, такое, что $k-1$ делится на p. А значит, существует подгруппа H группы G порядка p. Пусть h — образующий элемент группы H. Тогда существует гомоморфизм φ группы G на H, при котором $\varphi(a)=h$. А значит, как известно (см. [5], 3.11), регулярная рисовская полугруппа $\mathcal{M}(H, \Lambda, \Lambda, P^*)$, где $P^*=\begin{pmatrix}1 & h\\ 1 & 1\end{pmatrix}$, является гомоморфным образом полугруппы $\mathcal{M}(G, \Lambda, \Lambda, P)$. Так как $S\in V$, V — многообразие полугрупп, то $\mathcal{M}(H, \Lambda, \Lambda, P^*)\in V$. Но полугруппа $\mathcal{M}(H, \Lambda, \Lambda, P^*)$ есть ничто иное, как S_p. Итак, если $ef\in I$ и $fe\in I$, то нашлось такое простое число p, что $S_p\in V$.

Если $ef\notin I$, $fe\in I$, то фактор-полугруппа Риса S/I изоморфна полугруппе S_0. Если $ef\notin I$ и $fe\notin I$, то фактор-полугруппа Риса $(S/I)/I^*$ полугруппы S/I по идеалу $I^*=I\cup\{fe\}$ изоморфна S_0. Таким образом, если $ef\notin I$ или $fe\notin I$, то $S_0\in V$.

Предложение 1 доказано полностью.

3. Пусть $u=v$ — тождество над счетным алфавитом X. Через $l_z(u)$ обозначим число вхождений буквы z в слово u; $l_{xy}(u)$ — число вхождений слова xy в слово u; $h(u)$ — первую букву слова u; $t(u)$ — последнюю букву слова u; $\chi(u)$ — множество букв алфавита X, участвующих в записи слова u; если $\chi(u)=\chi(v)$, то, как обычно, назовем тождество $u=v$ нормальным. Через Var A обозначим многообразие полугрупп, порожденное полугруппой A. Равенство элементов в полугруппе слов над алфавитом X будем обозначать так: $\equiv$.

4. Используя результаты работы [7] или [8], нетрудно убедиться в том, что справедливо

Предложение 2. $\mathrm{Var}\, S_p=\Pi\big((xy)^p x=x,\ x^2yx=xyx^2\big)$.

5. Определение. Тождество $u=v$ над счетным алфавитом X назовем квазиортодоксальным, если $u=v$ не является нормальным, либо $u=v$ нормально и для него имеет место дизъюнкция следующих двух условий:

1) одно из слов, например, u, можно представить в виде $u\equiv u_1u_2$, где $\chi(u_1)\cap\chi(u_2)=\emptyset$, а слово $v\equiv v_1yxv_2$, где $y\in\chi(u_2)$, $x\in\chi(u_1)$.

2) одно из слов u и v, например, u, представимо в виде: $u\equiv u_1zu_2$, $\chi(u_1)\cap\chi(u_2)=\emptyset$, $l_z(u)=1$, причем либо $l_z(v)>1$, либо $l_z(v)=1$, $v\equiv v_1zv_2$, $\chi(v_1)\cap\chi(u_2)\neq\emptyset$ или $\chi(v_2)\cap\chi(u_1)\neq\emptyset$.

6. Предложение 3. *Тождество $u=v$ выполняется в полугруппе S_0 тогда и только тогда, когда оно не является квазиортодоксальным.*

Доказательсто. Необходимость. Пусть тождество $u=v$ выполняется в полугруппе S_0. Допустим, что $u=v$ не является нормальным, т.е. существует буква $z\in\chi(u)$ и $z\notin\chi(v)$. Строим отображение φ: $X\to S_0$ следующим образом: $\varphi z=2$, $\varphi x=1$, $\forall x\neq z$. Тогда $\varphi v=1$, $\varphi u=3$ или $\varphi u=0$ или $\varphi u=2$.

Значит, тождество $u=v$ нормально. Допустим, что $u=v$ квазиортодоксально. Тогда имеет место хотя бы одно из условий определения квазиортодоксального тождества. Пусть имеет место первое условие, т.е. слово u представимо в виде: $u\equiv u_1u_2$, $\chi(u_1)\cap\chi(u_2)=\emptyset$, а $v\equiv v_1yxv_2$, где $y\in\chi(u_2)$, $x\in\chi(u_1)$. Строим отображение φ: $X\to S_0$ следующим образом: полагаем $\varphi x_i=1$ $\forall x_i\in\chi(u_1)$, $\varphi y_j=2$ $\forall y_j\in\chi(u_2)$. Тогда $\varphi u=1\cdot 2=3$, $\varphi v=\varphi v_1\cdot 2\cdot 1\cdot\varphi v_2=0$. Противоречие.

Пусть теперь для тождества $u=v$ имеет место второе условие определения квазиортодоксального тождества. Пусть $u\equiv u_1zu_2$, где $\chi(u_1)\cap\chi(u_2)=\emptyset$, $z\notin\chi(u_1)\cup\chi(u_2)$, а $l_z(v)>1$. Зададим отображение φ: $X\to S_0$, положив $\varphi z=3$, $\varphi x=1$ $\forall x\in\chi(u_1)$, $\varphi y=2$ $\forall y\in\chi(u_2)$. Тогда $\varphi u=3$, $\varphi v=0$. Если же $v\equiv v_1zv_2$, где $l_z(v)=1$, $\chi(v_1)\cap\chi(u_2)\neq\emptyset$ или $\chi(v_2)\cap\chi(u_1)\neq\emptyset$, то опять будем иметь $\varphi u=3$, $\varphi v=0$.

Необходимость доказана.

Достаточность. Пусть тождество $u=v$ нормально и не является квазиортодоксальным. Допустим, что $u=v$ не выполняется в S_0. Значит, существует отображение φ: $X\to S_0$ такое, что $\varphi u\neq\varphi v$. Не может быть $\varphi u=1$ или $\varphi u=2$, так как это означало бы, что все буквы из $\chi(u)$ отображаются при φ в 1 или 2, а так как $\chi(u)=\chi(v)$, то это означало бы, что $\varphi u=\varphi v$. Аналогично, $\varphi v\neq 1$ и $\varphi v\neq 2$. Значит, один из элементов φu и φv, например φu, равен 3, а другой — 0. Итак, пусть $\varphi u=3$, $\varphi v=0$. Если никакая буква из $\chi(u)=\chi(v)$ не отображается при φ в 3, то $\chi(u)$ есть объединение двух непересекающихся множеств $\{x_1, \ldots, x_m\}$, $\{y_1, \ldots, y_t\}$, при этом $\varphi x_i=1$, $\varphi y_j=2$, $i=1, \ldots, m$; $j=1, \ldots, t$. А так как $\varphi u=3$, то $u\equiv u_1u_2$, $\chi(u_1)=\{x_1, \ldots, x_m\}$, $\chi(u_2)=\{y_1, \ldots, y_t\}$. Так как $\varphi v=0$, то в этом случае $v \equiv v_1 y_{j_0} x_{i_0} v_2$, где $j_0 \in \{1, ..., t\}$, $i_0 \in \{1, ..., m\}$. А это означает, что $u=v$ является квазиортодоксальным.

Если существует буква $z \in \chi(u)=\chi(v)$ такая, что $\varphi z=3$, то $l_z(u)=1$, так как $\varphi u=3$; $u \equiv u_1 z u_2$, $\chi(u_1) \cap \chi(u_2)=\emptyset$, причем $\varphi(\chi(u_1))=\{1\}$, $\varphi(\chi(u_2))=\{2\}$, если $u_1 \not\equiv \emptyset$ или $u_2 \not\equiv \emptyset$. Так как $\varphi v=0$, то либо $l_z(v)>1$, либо $l_z(v)=1$, но $v \equiv \equiv v_1 z v_2$, где $\chi(v_1) \cap \chi(u_2) \neq \emptyset$ или $\chi(v_2) \cap \chi(u_1) \neq \emptyset$. А это опять означает, что тождество $u=v$ квазиортодоксально.

Предложение 3 доказано.

7. В работах [4], [9] указан базис тождеств полугруппы S_0, а именно:

$$\operatorname{Var} S_0 = \Pi(x^2 = x^3,\ xyx = yxy,\ xyzx = xzyx,\ xyx = xyx^2).$$

Отметим, что доказательство этого факта вытекает из Предложения 3.

8. Введем еще несколько определений. Пусть $u=v$ — тождество над алфавитом X. Через $d(u=v)$ обозначим наибольший общий делитель разностей $|l_x(u)-l_x(v)|$ по всем $x \in X$. Если Φ — некоторая совокупность тождеств, то через $D(\Phi)$ обозначим наибольший общий делитель всех чисел $d(u=v)$ по всем тождествам $u=v$ из Φ. (Эта характеристика была рассмотрена в [3]).

Для каждого слова xy (x и y не обязательно различные буквы алфавита) находим $|l_{xy}(u)-l_{xy}(v)|$ и наибольший общий делитель всех этих чисел назовем двухбуквенной характеристикой тождества $u=v$. Наибольший общий делитель двухбуквенных характеристик всех тождеств из совокупности Φ назовем двухбуквенной характеристикой Φ и обозначим $D^*(\Phi)$. ($D^*(\Phi)$ рассматривалась в [6]). Наибольший общий делитель чисел $D(\Phi)$ и $D^*(\Phi)$ назовем характеристикой совокупности тождеств Φ.

9. Теорема 1. *Следующие свойства для многообразия полугрупп $V=\Pi(\Phi)$ эквивалентны:*

1) *V — квазиортодоксальное многообразие полугрупп;*
2) *V не содержит полугрупп S_0, S_p, где p — произвольное простое число;*
3) *совокупность тождеств Φ удовлетворяет двум условиям:*
   а) *$h(u) \neq h(v)$ или $t(u) \neq t(v)$ для некоторого тождества $u=v$ из Φ, или же характеристика Φ равна* 1;
   б) *среди тождеств Φ есть хотя бы одно квазиортодоксальное тождество.*

Доказательство. Согласно Предложению 1 многообразие V квазиортодоксально (поскольку полугруппы S_0, S_p по всем простым p не квазиортодоксальны) тогда и только тогда, когда V не содержит полугрупп S_0, S_p (по всем простым p). Согласно Предложению 3, V не содержит S_0 тогда и только тогда, когда среди тождеств из Φ есть хотя бы одно квазиортодоксальное тождество.

Согласно Предложению 2 и результатам работ [6], [7], [8], V не содержит полугрупп S_p тогда и только тогда, когда V удовлетворяет условию а).

10. Теорема 2. *Минимальными неквазиортодоксальными многообразиями полугрупп являются следующие:*

$$\mathrm{Var}\, S_p = \Pi\big((xy)^p x = x,\ x^2yx = xyx^2\big),$$

где p-произвольное простое число, и

$$\mathrm{Var}\, S_0 = \Pi(x^2 = x^3,\ xyx = yxy,\ xyzx = xzyx,\ xyx = xyx^2).$$

Доказательство. 1) Каждое из перечисленных многообразий не является квазиортодоксальным, поскольку неквазиортодоксальны полугруппы S_0, S_p.

2) Допустим, что $\mathrm{Var}\, S_p$ не является минимальным. Тогда существует собственное неквазиортодоксальное подмногообразие V' многообразия $\mathrm{Var}\, S_p$. А значит V', а тогда и $\mathrm{Var}\, S_p$, содержит S_0 или S_q, где q — некоторое простое число. Но $S_0 \notin \mathrm{Var}\, S_p$, так как в полугруппе S_0 не выполняется тождество $(xy)^p x = x$. $S_q \notin \mathrm{Var}\, S_p$ при $q \neq p$, так как тождество $(xy)^p x = x$ не выполняется и в полугруппе S_q. Значит, $\mathrm{Var}\, S_p \subset V'$, а тогда $\mathrm{Var}\, S_p = V'$. Значит, $\mathrm{Var}\, S_p$ минимально.

Поскольку тождество $xyx = xyx^2$ не выполняется ни в какой полугруппе S_p, то $\mathrm{Var}\, S_0$ — минимально.

Других минимальных неквазиортодоксальных многообразий полугрупп нет. Это непосредственно следует из Предложения 1.

Литература

[1] А. Я. Айзенштат, В. К. Богута, О решетке многообразий полугрупп, *Полугрупповые многообразия и полугруппы эндоморфизмов* (Ленинград, 1979), 3—46.

[2] C. Benzaken et H. C. Mayr, Notion de demi-bande: demi-bandes de type deux, *Semigroup Forum*, **10** (1975), 115—128.

[3] А. П. Бирюков, Полугруппы, заданные тождествами, *Уч. зап. Новосибирского Гос. Пед. ин-та, Физика и математика, Новосибирск*, **18** (1963), 139—169.

[4] C. C. Edmunds, On certain finitely based varieties of semigroups, *Semigroup Forum*, **15** (1977), 21—39.

[5] A. H. Clifford and G. B. Preston, *The Algebraic Theory of Semigroups. I*, Amer. Math. Soc. (Providence, R. I., 1968).

[6] Г. И. Машевицкий, О тождествах вполне простых полугрупп над абелевыми группами, *XXXI Герценовские чтения* (Ленинград, 1978), 33—36.

[7] Г. И. Машевицкий, О тождествах в многообразиях вполне простых полугрупп над абелевыми группами, — В сб.: *Современная алгебра* (Ленинград, 1978), 81—89.

[8] V. V. Rasin, On the lattice of varieties of completely simple semigroups, *Semigroup Forum*, **17** (1979), 113—122.

[9] Е. П. Симельгор, О тождествах четырехэлементных полугрупп, — В сб.: *Современная алгебра* (Ленинград, 1978), 146—153.

УЛ. МАРКСА 227, КВ. 15
167000 СЫКТЫВКАР, СССР

Acta Sci. Math. 47 (1984), 303—306.

Term functions and subalgebras

EMIL W. KISS

Answering a question of A. F. Pixley this note shows that the class of para primal algebras cannot be characterised by the preservation properties of the term functions. Moreover all classes are described which can be characterised in such a way.

1. The characterisability result

Let $\mathfrak{A}$ be a finite algebra with underlying set A and Q a collection of finitary relations on A. A $\varrho\in Q$ is called compatible on $\mathfrak{A}$ if each term function f of $\mathfrak{A}$ preserves ϱ. Now if each finitary function f that preserves all the compatible elements of Q is a term function of $\mathfrak{A}$ then Q is said to characterise the term functions of $\mathfrak{A}$. The class of the clones of all such $\mathfrak{A}$ is denoted by Q^* and we say that *a class $\mathcal{K}$ of algebras on A can be characterised by the preservation properties of the term functions* if the set of clones of all the algebras in $\mathcal{K}$ is of the form Q^* for an appropriate collection Q.

This complicated definition can lead to very useful characterisations of classes $\mathcal{K}$ when Q is a concrete collection. The most important example is the class of quasi primal algebras where Q consists of the partial bijections on A (cf. WERNER [7] also for other examples).

In order to give an internal description of characterisible classes let us call a clone F *cocyclic* if $F=\mathrm{Pol}\,\varrho$ for some (finitary) relation ϱ on A (for notation and elementary results concerning the Pol—Inv connection see Pöschel—Kalužnin [5]).

Theorem. *A class $\mathcal{C}$ of clones on a finite set A is of the form Q^* iff*

(i) *$\mathcal{C}$ is closed under intersection (in particular the clone of all operations is in $\mathcal{C}$);*

(ii) *Each element of $\mathcal{C}$ is the intersection of cocyclic elements of $\mathcal{C}$.*

$\mathcal{C}$ is of the form Q^* for some finite Q iff (i);

Recevied November 22, 1981, and in revised form November 10, 1983.

(iii) *Each element of $\mathscr{C}$ is cocyclic;*
(iv) *$\mathscr{C}$ is finite.*

Proof. The key observation is the following:

(∗) $F \in Q^*$ iff $F = \text{Pol}\, Q'$ for some $Q' \subseteqq Q$.

Indeed, suppose $F \in Q^*$ and let Q' be the set of all compatible elements of Q. Then $F = \text{Pol}\, Q'$ by the definition of Q^*. Conversely, suppose $F = \text{Pol}\, Q'$ for some $Q' \subseteqq Q$. If an operation f preserves all the elements of Q that are compatible with F then, in particular, f preserves those of Q' (by $Q' \subseteqq Q$ and $F \subseteqq \text{Pol}\, Q'$) so by $F \supseteqq \text{Pol}\, Q'$ we have $f \in F$ as desired.

Now the Theorem is obvious by using the rule

$$\text{Pol}\,\{\cup Q_i\} = \cap\, \text{Pol}\, Q_i$$

and the following observation which gives also an intrinsic characterisation of cocyclic clones (see e.g. Jablonskiĭ [3]):

Proposition. Pol $\{\varrho_1, \ldots, \varrho_k\}$ *is always cocyclic. A clone F is cocyclic iff there is an integer n such that $f \in F$ if and only if every at most n-ary function resulting from f by identifying certain (maybe no) variables is contained in F.*

Proof. Let $F = \text{Pol}\,\{\varrho_1, \ldots, \varrho_k\}$ and choose n to be the maximum of the cardinalities of the ϱ_i-s. Then F satisfies the property in the second assertion. Conversely, if F is such and $\varrho \subseteqq A^{A^n}$ is the set of all n-ary elements of F then clearly $F = \text{Pol}\, \varrho$.

2. Para primal algebras

We prove

Corollary. *Suppose A is a finite set of at least two elements. Then the class of all para primal algebras on A can be defined by the preservation properties of the term functions iff A has two elements. In this case this class can be defined by finitely many relations.*

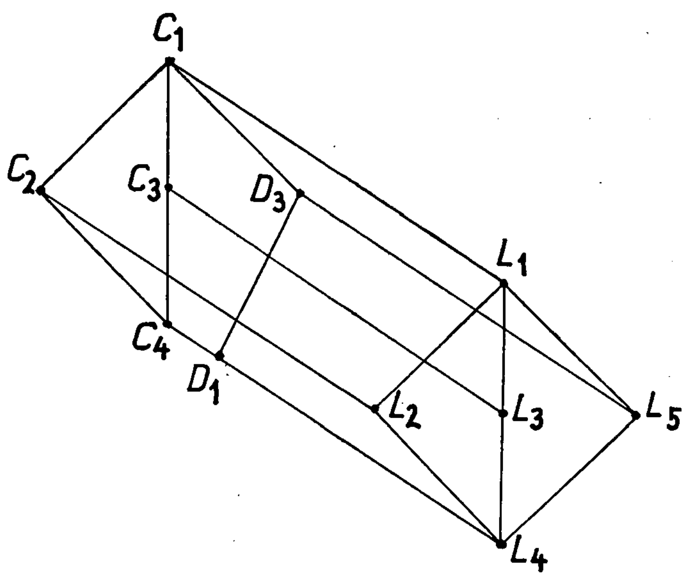
Proof. In the case $|A| \geqq 3$ let F_c be the clone of a cyclic group on A and F_t be the clone generated by the ternary discriminator. As the elements of F_t preserve all subsets of A and the elements of F_c are of the form

$$f(x_1, \ldots, x_n) = k_1 x_1 + \ldots + k_n x_n$$

where the k_i-s are integers, an easy calculation shows that $F_c \cap F_t$ consists of the projections. Thus the class of para primal clones does not satisfy (i) of the Theorem.

The case $|A| = 2$ could be settled by an elementary argument: all para primal clones are either quasi primal or affine by McKenzie [4], such clones are always cocyclic by the Proposition and easy calculations. However, for the sake of better visibility of the situation we derive the poset (in fact the lattice) of para primal clones on the set $\{0, 1\}$ from Post's classification ([6], for a considerably shorter proof see [1]). The clones D_1 (generated by the discriminator), D_3, C_1, C_2, C_3, C_4 defined below are quasi primal and L_1, L_2, L_3, L_4, L_5 are affine. These are eleven clones but C_2 and C_3 as well as L_2 and L_3 give cryptomorphic algebras by $0 \leftrightarrow 1$, so one can obtain the list of two element para primal algebras found in CLARK—KRAUSS [2].

Finally I wish to say thanks to B. Csákány and Á. Szendrei for their remarks that made possible to simplify the paper.



$C_1 = \{\text{all finitary functions on } \{0, 1\}\}$;

+	0	1
0	0	1
1	1	0

$\mathbf{1}\colon \{0, 1\} \to \{1\}, \quad \bar{0} = 1,$

$\mathbf{0}\colon \{0, 1\} \to \{0\}, \quad \bar{1} = 0,$

$C_4 = \{f \in C_1 | f(x, \ldots, x) = x\},$

$C_3 = \{f \in C_1 | f(0, \ldots, 0) = 0\},$

$C_2 = \{f \in C_1 | f(1, \ldots, 1) = 1\},$

$D_1 = \{f \in C_4 | f(\bar{x}_1, \ldots, \bar{x}_n) = \overline{f(x_1, \ldots, x_n)}\},$

$D_3 = \{f \in C_1 | f(\bar{x}_1, \ldots, \bar{x}_n) = \overline{f(x_1, \ldots, x_n)}\},$

$L_1 = [x+y, \mathbf{1}]$ (that is, the clone generated by these operations),

$L_2 = [x+y+1], \quad L_3 = [x+y], \quad L_4 = [x+y+z], \quad L_5 = [x+y+z+1].$

References

[1] J. Demetrovics, A kétértékű logika strukturális vizsgálata, *Alkalmazott Matematikai Lapok*, **1** (1975), 405—424.

[2] D. M. Clark and P. H. Krauss, Plain para primal algebras, *Algebra Universalis*, **11** (1980), 365—388.

[3] S. V. Jablonskiĭ, The structure of the upper neighborhood for predicately describable classes in P_k, *Dokl. Akad. Nauk SSSR*, **218** (1974), 304—307.

[4] R. McKenzie, Para primal varieties: a study of finite axiomatizability and definable principal congruences in locally finite varieties, *Algebra Universalis*, **8** (1978), 336—348.

[5] R. Pöschel and L. A. Kalužnin, *Funktionen- und Relationenalgebren*, Deutscher Verlag der Wissenschaften (Berlin, 1978).

[6] E. L. Post, *The Two-Valued Iterative Systems of Mathematical Logic*, Annals of Math. Studies, No. 5, Princeton Univ. Press (1941).

[7] H. Werner, *Discriminator Algebras*, Studien zur Algebra und ihre Anwendungen, Band. 6., Akademie-Verlag (Berlin, 1978).

MATHEMATICAL INSTITUTE OF THE
HUNGARIAN ACADEMY OF SCIENCES
REÁLTANODA U. 13—15
1053 BUDAPEST, HUNGARY

Acta Sci. Math. 47 (1984), 307—319

On questions of hereditariness of radicals

L. C. A. van LEEUWEN

Introduction

All rings considered are associative. We shall use the following notation: $\mathcal{R}$ is a radical class, $\mathcal{S}\mathcal{R}$ the corresponding semisimple class; $\lhd$ indicates an ideal; ann (A) is the two-sided annihilator of a ring A; $\mathcal{B}$ is the lower Baer radical; $L(\ )=$ $=$lower radical class, for instance, $\mathcal{B}=L$(zero-rings).

A radical class $\mathcal{R}$ is said to be a *hereditary class* if $\mathcal{R}$ satisfies:

$$B \lhd A, \quad A \in \mathcal{R} \Rightarrow B \in \mathcal{R}.$$

In [1] a weak version of hereditariness was introduced, which arose in connection with the finite closure property of radicals under subdirect sums. If a radical class $\mathcal{R}$ is closed under finite subdirect sums, then $\mathcal{R}$ has the property:

$$I \lhd A, \quad A \in \mathcal{R}, \quad I \subseteqq \operatorname{ann} A \Rightarrow I \in \mathcal{R}.$$

Such a radical is said to be *hereditary for annihilator ideals* ([1], Proposition 1.7). Although this condition is not sufficient for the finite closure property of $\mathcal{R}$, very little is needed to make $\mathcal{R}$ hereditary. Hereditary radical classes are closed under finite subdirect sums. We investigate these questions in §2.

In [3] a new characterization was found for the *maximal hereditary subradical* $h_{\mathcal{R}}$ of a radical $\mathcal{R}$, in fact

$$h_{\mathcal{R}} = \bar{\bar{\mathcal{R}}} = \{A| \quad \text{any ideal of} \quad A \quad \text{is in} \quad \bar{\mathcal{R}}\}$$

where $\bar{\mathcal{R}} = \{A|$ any ideal of A is in $\mathcal{R}\}$. We use this result to sharpen Proposition 1.6 of [1], where $h_{\mathcal{R}}$ was given as an intersection of an infinite number of radical classes. We show that the chain, used in [1], stops at the second step. We also show that, for any radical $\mathcal{R}$ containing $\mathcal{B}$ or being subidempotent,

$$h_{\mathcal{R}} = \bar{\mathcal{R}} = \{A| \quad \text{any ideal of} \quad A \quad \text{is in} \quad \mathcal{R}\},$$

Received January 5, 1983.

i.e. $\bar{\mathcal{R}}$ is hereditary. Here a radical class $\mathcal{R}$ is called *subidempotent* if any ring A in $\mathcal{R}$ is idempotent.

Our terminology for radical theory is the usual one. Both a radical and a radical class are denoted by $\mathcal{R}$. A ring A is in the radical class $\mathcal{R}$ or A is an $\mathcal{R}$-ring if $A = \mathcal{R}(A)$, where $\mathcal{R}(A)$ is the radical of the ring A. The semisimple class $\mathcal{S}\mathcal{R}$ of the radical $\mathcal{R}$ consists of all rings A, such that $\mathcal{R}(A)=0$, i.e.

$$\mathcal{S}\mathcal{R} = \{A \mid A \text{ has no non-zero ideal in } \mathcal{R}\}.$$

A class $\mathbf{M}$ is said to be *closed under finite subdirect sums* if $A_1, \ldots, A_n \in \mathbf{M}$ implies that $A_1 + \ldots + A_n \in \mathbf{M}$ (subdirect sum) for any finite number n of rings $A_1, \ldots, A_n$. In order to show closure under finite subdirect sums one needs only consider $n=2$.

I would like to thank Dr. R. Wiegandt for his criticism and valuable remarks in preparing this paper. Originally I tried to do something with quasi-radicals, but he remarked that an order-preserving quasi-radical is complete, which, together with idempotency, makes it a radical (cf. [2]).

1. In our first result we deal with sums of ideals (cf. Problem 12 in [4]).

Theorem 1. *Let A be a ring with ideals B, C and $B \cap C \in \mathcal{R}$ for some radical $\mathcal{R}$. Then $\mathcal{R}(B+C) = \mathcal{R}(B) + \mathcal{R}(C)$.*

Proof. The inclusion $\mathcal{R}(B)+\mathcal{R}(C) \subseteqq \mathcal{R}(B+C)$ is clear. Obviously, we have the direct decomposition

$$B+C/B \cap C = B/B \cap C \oplus C/B \cap C.$$

By the assumption $B \cap C \subseteqq \mathcal{R}(B+C)$, therefore the above direct decomposition yields

$$\mathcal{R}(B+C)/B \cap C = K/B \cap C \oplus L/B \cap C$$

for ideals K resp. L in B resp. C. Clearly $K/B \cap C$ is an $\mathcal{R}$-ring and contained in $\mathcal{R}(B/B \cap C) = \mathcal{R}(B)/B \cap C$. Similarly

$$L/B \cap C \subseteqq \mathcal{R}(C/B \cap C) = \mathcal{R}(C)/B \cap C.$$

Hence

$$\mathcal{R}(B+C)/B \cap C \subseteqq \mathcal{R}(B)/B \cap C \oplus \mathcal{R}(C)/B \cap C$$

giving

$$\mathcal{R}(B+C) \subseteqq \mathcal{R}(B) + \mathcal{R}(C).$$

In addition we have

Theorem 2. *For any ring A with arbitrary ideals B, C and I, J and for any radical $\mathcal{R}$ the following two statements are equivalent:*

(i) A/B, $A/C \in \mathcal{R}$, $\mathcal{R}(B) = \mathcal{R}(C)$ *implies* $A/(B \cap C) \in \mathcal{R}$,

(ii) A/I, $A/J \in \mathcal{R}$, $\mathcal{R}(I) = \mathcal{R}(J) = 0$ *implies* $A/(I \cap J) \in \mathcal{R}$.

Proof. (i)$\Rightarrow$(ii) is trivial.

Let A/B, $A/C \in \mathcal{R}$ with $\mathcal{R}(B) = \mathcal{R}(C)$. Then

$$\frac{A/\mathcal{R}(B)}{B/\mathcal{R}(B)} \cong A/B \in \mathcal{R}, \quad \frac{A/\mathcal{R}(B)}{C/\mathcal{R}(B)} \cong A/C \in \mathcal{R}$$

with

$$\mathcal{R}\left(\frac{B}{\mathcal{R}(B)}\right) = \mathcal{R}\left(\frac{C}{\mathcal{R}(B)}\right) = 0 \quad (\mathcal{R}(B) = \mathcal{R}(C)).$$

Hence

$$\frac{A/\mathcal{R}(B)}{B/\mathcal{R}(B) \cap C/\mathcal{R}(B)} = \frac{A/\mathcal{R}(B)}{(B \cap C)/\mathcal{R}(B)} \cong \frac{A}{B \cap C} \in \mathcal{R}.$$

In order to show that a radical class $\mathcal{R}$ is closed under finite subdirect sums we might use the following reduction:

Theorem 3. *If for any ring A and arbitrary ideals I, J in A with $I \cap J = 0$ the condition $A/I, A/J \in \mathcal{R}$ implies that $A/(I \cap J) \cong A \in \mathcal{R}$, then $\mathcal{R}$ is closed under finite subdirect sums.*

Proof. The symbol $\oplus$ will mean "direct sum". Let I, J be ideals of A such that $I \cap J = 0$, $A/I \in \mathcal{R}$ and $A/J \in \mathcal{R}$. By $I \cap J = 0$ we have

$$(1) \qquad \big(I \oplus \mathcal{R}(J)\big) \cap \big(\mathcal{R}(I) \oplus J\big) = \mathcal{R}(I) \oplus \mathcal{R}(J).$$

and also

$$(2) \qquad \big(I \oplus \mathcal{R}(J)\big)/\big(\mathcal{R}(I) \oplus \mathcal{R}(J)\big) \cong I/\mathcal{R}(I) \in \mathcal{S}\mathcal{R}$$

$$(3) \qquad \big(\mathcal{R}(I) \oplus J\big)/\big(\mathcal{R}(I) \oplus \mathcal{R}(J)\big) \cong J/\mathcal{R}(J) \in \mathcal{S}\mathcal{R}.$$

In (2) and (3) the left hand sides are ideals of $A/\big(\mathcal{R}(I) \oplus \mathcal{R}(J)\big)$ and by (1) these ideals have zero intersection. Since

$$\frac{A/\big(\mathcal{R}(I) \oplus \mathcal{R}(J)\big)}{\big(I \oplus \mathcal{R}(J)\big)/\big(\mathcal{R}(I) \oplus \mathcal{R}(J)\big)} \cong A/\big(I \oplus \mathcal{R}(J)\big) \in \mathcal{R}$$

and

$$\frac{A/\big(\mathcal{R}(I) \oplus \mathcal{R}(J)\big)}{\big(\mathcal{R}(I) \oplus J\big)/\big(\mathcal{R}(I) \oplus \mathcal{R}(J)\big)} \cong A/\big(\mathcal{R}(I) \oplus J\big) \in \mathcal{R}$$

the imposed condition is applicable yielding

$$A/\big(\mathcal{R}(I) \oplus \mathcal{R}(J)\big) \in \mathcal{R};$$

and so the extension property of $\mathcal{R}$ implies $A \in \mathcal{R}$.

Lemma 4a. *Let A be a ring with ideals I, J such that $I \cap J = 0$, $A/I \in \mathcal{R}$* and *$A/J \in \mathcal{R}$. If $\mathcal{R}$ is hereditary for annihilator ideals, then* ann $A \in \mathcal{R}$. *Moreover, if, in addition, $I, J \in \mathcal{S}\mathcal{R}$, then $I \cap \operatorname{ann} A = J \cap \operatorname{ann} A = 0$.*

Proof. $\operatorname{ann} A/\operatorname{ann} A\cap I \cong (\operatorname{ann} A+I)/I$ is an annihilator ideal of $A/I\in\mathcal{R}$, so $\operatorname{ann} A/\operatorname{ann} A\cap I\in\mathcal{R}$. Also

$$\operatorname{ann} A\cap I \cong \frac{(\operatorname{ann} A\cap I)+J}{J} \subseteq \frac{J+\operatorname{ann} A}{J}\in\mathcal{R},$$

since $(J+\operatorname{ann} A)/J\subseteq \operatorname{ann} A/J$, $A/J\in\mathcal{R}$. Again, since $((\operatorname{ann} A\cap I)+J)/J$ is an annihilator ideal of $(J+\operatorname{ann} A)/J$,

$$\frac{(\operatorname{ann} A\cap I)+J}{J} \cong \operatorname{ann} A\cap I\in\mathcal{R}.$$

The extension property of $\mathcal{R}$ implies $\operatorname{ann} A\in\mathcal{R}$. Now assume that $I, J\in\mathcal{S}\mathcal{R}$. Then $\operatorname{ann} A\cap I$ is an ideal in $I\in\mathcal{S}\mathcal{R}$ implies $\operatorname{ann} A\cap I=0$. Similarly $\operatorname{ann} A\cap J=0$.

From the proof of Lemma 4a we see that $I\in\mathcal{S}\mathcal{R}$ implies that $\operatorname{ann} A\cap I=0$. Clearly $\operatorname{ann} A\subseteq I^*=\{a\in A\mid aI=Ia=(0)\}$, as $I\subseteq A$. We can say more if we assume that $\mathcal{B}\subseteq\mathcal{R}$.

Lemma 4b. *Let $\mathcal{B}\subseteq\mathcal{R}$; A is a ring with ideals I, J such that $I\in\mathcal{S}\mathcal{R}$, $I\cap J=0$, $A/I\in\mathcal{R}$ and $A/J\in\mathcal{R}$. Also $\mathcal{R}$ is hereditary for annihilator ideals. Then any ideal K in A such that $K\cap I=0$ is contained in I^* and I^* is maximal with respect to $I^*\cap I=0$. Moreover $A/I^*\in\mathcal{R}$, whereas* $\operatorname{ann}(A/I^*)=0$.

Proof. $(I^*\cap I)^2\subseteq I^*\cdot I=0$, so $(I^*\cap I)\lhd I\in\mathcal{S}\mathcal{R}$ gives $I^*\cap I=0$. As $J\cap I=0$, by Zorn's lemma there exists an ideal M, maximal relative to $M\cap I=0$. Since $MI=IM=0$, $M\subseteq I^*$ and the maximality of M ensures $M=I^*$.

Let K be any ideal in A such that $K\cap I=0$. If K is not contained in I^*, then $(K+I^*)\cap I\neq 0$. Now let x, y be arbitrary elements in $(K+I^*)\cap I$, then $x=k+a$ $(k\in K, a\in I^*)$, $y\in I$. Hence $xy=(k+a)y=ky+ay=0$, as $K\cap I=I^*\cap I=0$. So $[(K+I^*)\cap I]^2=0$. But $[(K+I^*)\cap I]\lhd I\in\mathcal{S}\mathcal{R}$ and $\mathcal{S}\mathcal{R}\subseteq\mathcal{S}\mathcal{B}$ implies $(K+I^*)\cap I$ is a semiprime ring, consequently $(K+I^*)\cap I=0$. This contradicts $(K+I^*)\cap I\neq 0$, so $K\subseteq I^*$. In particular, $J\subseteq I^*$ and $A/J\in\mathcal{R}$ implies $A/I^*\in\mathcal{R}$. The ideal $I\cong(I+I^*)/I^*$ is essential in A/I^*: if $B/I^*\neq 0$ is an ideal of A/I^*, then $B\cap I\nsubseteq I^*$, otherwise $B\cap I\subseteq I^*\cap I=0$ implies $B\cap I=0$ which is impossible by the maximality of I^*. Hence

$$0\neq((B\cap I)+I^*)/I^*\subseteq B/I^*\cap(I+I^*)/I^*.$$

As $\mathcal{R}$ is hereditary for annihilator ideals and $(I+I^*)/I^*\cap\operatorname{ann} A/I^*\subseteq\operatorname{ann} A/I^*$, it follows that $(I+I^*)/I^*\cap\operatorname{ann} A/I^*\in\mathcal{R}$. On the other hand $I\cong(I+I^*)/I^*\in\mathcal{S}\mathcal{R}$, so $(I+I^*)/I^*\cap\operatorname{ann} A/I^*\in\mathcal{S}\mathcal{R}$ yielding $(I+I^*)/I^*\cap\operatorname{ann} A/I^*=0$. The essential property of $(I+I^*)/I^*$ in A/I^* implies $\operatorname{ann} A/I^*=0$.

For any ring A and any ideal I in A we define $[I:A]:=\{x\in A\mid xA\subseteq I,\ Ax\subseteq I\}$.

Theorem 5. *Let $\mathcal{R}$ be an arbitrary radical class. $\mathcal{R}$ is closed under finite subdirect sums if and only if*

(i) *Whenever I and J are ideals in a ring A with $I\cap J=0$, then $A/[I:A], A/[J:A]\in\mathcal{R}$ implies $A/[I:A]\cap[J:A]\in\mathcal{R}$.*

(ii) *$\mathcal{R}$ is hereditary for annihilator ideals.*

Proof. Suppose that (i) and (ii) are satisfied. Let I, J be ideals in A with $I\cap J=0$ and suppose A/I, $A/J\in\mathcal{R}$. Since $I\subseteqq[I: A]$ and $J\subseteqq[J: A]$ it follows that $A/[I: A]$, $A/[J: A]\in\mathcal{R}$. It can easily be seen that $I\cap J=0$ implies $\operatorname{ann} A=[I: A]\cap[J: A]$. Hence (i) implies that $A/\operatorname{ann} A\in\mathcal{R}$. From Lemma 4a we get, using (ii), that $\operatorname{ann} A\in\mathcal{R}$. The extension property of $\mathcal{R}$ implies $A\in\mathcal{R}$.

The converse is clear by Proposition 1.7 [1].

Note that $\operatorname{ann}(A/I)=[I: A]/I$, so we may replace (i) by

$$\frac{A/I}{\operatorname{ann}(A/I)},\quad \frac{A/J}{\operatorname{ann}(A/J)}\in\mathcal{R}\quad \text{implies}\quad \frac{A}{\operatorname{ann} A}\in\mathcal{R}.$$

Corollary 6. *Let $\mathcal{R}$ be a radical class such that $\mathcal{B}\subseteqq\mathcal{R}$. Then $\mathcal{R}$ is closed under finite subdirect sums if and only if*

$$\frac{A}{I^*},\quad \frac{A}{[I: A]}\in\mathcal{R}\quad \textit{implies}\quad \frac{A}{I^*\cap[I: A]}\in\mathcal{R}$$

for any ideal I in any ring A.

Proof. Obviously $\mathcal{B}\subseteqq\mathcal{R}$ implies that $\mathcal{R}$ is hereditary for annihilator ideals. Let A be a ring with ideals I, J such that $I\cap J=0$; A/I, $A/J\in\mathcal{R}$. We have to show that $A\in\mathcal{R}$. If $I\notin\mathcal{S}\mathcal{R}$, then $I/\mathcal{R}(I)$, $(J+\mathcal{R}(I))/\mathcal{R}(I)$ are ideals in $A/\mathcal{R}(I)$ and $I\cap(J+\mathcal{R}(I))=\mathcal{R}(I)+(I\cap J)=\mathcal{R}(I)$. So $A/\mathcal{R}(I)$ is a ring with ideals $I/\mathcal{R}(I)$, $(J+\mathcal{R}(I))/\mathcal{R}(I)$ having zero-intersection; also A/I, $A/(J+\mathcal{R}(I))\in\mathcal{R}$, as $A/J\in\mathcal{R}$. Now $I/\mathcal{R}(I)\in\mathcal{S}\mathcal{R}$. If we can show that $A/\mathcal{R}(I)\in\mathcal{R}$, we are done by the extension property.

Hence, without loss of generality, we may assume: $I\lhd A, J\lhd A$; A/I, $A/J\in\mathcal{R}$ and $I\in\mathcal{S}\mathcal{R}$.

Now apply Lemma 4b. Then $J\subseteqq I^*$ and $I\subseteqq[I: A]$ imply A/I^*, $A/[I: A]\in\mathcal{R}$. Hence $A/(I^*\cap[I: A])\in\mathcal{R}$. By Lemma 4b we know that $\operatorname{ann}(A/I^*)=0$, i.e. $[I^*: A]=I^*$. From $I\cap I^*=0$, as $I\in\mathcal{S}\mathcal{R}$, it follows that $\operatorname{ann} A=[I: A]\cap[I^*: A]=I^*\cap[I^*: A]$. Hence $A/\operatorname{ann} A\in\mathcal{R}$. Then Lemma 4a implies that $\operatorname{ann} A\in\mathcal{R}$ and consequently $A\in\mathcal{R}$. So the condition is sufficient. The converse is obvious.

The above proof of Corollary 6 suggests the next result which is a further reduction for the question of finite subdirect closure for radicals (cf. Theorem 3).

Theorem 7. *If for any ring A and arbitrary ideals I, J in A with $I \cap J = 0$, $I, J \in \mathcal{S}\mathcal{R}$ the condition A/J, $A/I \in \mathcal{R}$ implies that $A \in \mathcal{R}$, then $\mathcal{R}$ is closed under finite subdirect sums.*

Proof. Let A be a ring with ideals I, J such that $I \cap J = 0$; A/I, $A/J \in \mathcal{R}$. By Theorem 3 we have to show that $A \in \mathcal{R}$. Now the ring $A/(\mathcal{R}(I) \oplus \mathcal{R}(J))$ has ideals $(I \oplus \mathcal{R}(J))/(\mathcal{R}(I) \oplus \mathcal{R}(J))$, $(\mathcal{R}(I) \oplus J)/(\mathcal{R}(I) \oplus \mathcal{R}(J))$ with zero intersection and both ideals are in $\mathcal{S}\mathcal{R}$ (see the proof of Theorem 3). Hence $A/(\mathcal{R}(I) \oplus \mathcal{R}(J)) \in \mathcal{R}$ and $A \in \mathcal{R}$.

Theorem 8. *Let $\mathcal{R}$ be a radical class. Then $\mathcal{R}$ is hereditary for annihilator ideals if and only if AI, $IA \in \mathcal{R}$ imply $I \in \mathcal{R}$ for any ring $A \in \mathcal{R}$ and any ideal I in A.*

Proof. Let $I \lhd A$ with $A \in \mathcal{R}$ and $I \subseteqq \operatorname{ann} A$. Then $AI = IA = 0 \in \mathcal{R}$ implies $I \in \mathcal{R}$. Conversely, let $I \lhd A$ with $A \in \mathcal{R}$ such that AI, $IA \in \mathcal{R}$. Now

$$\frac{I}{AI+IA} \lhd \frac{A}{AI+IA}$$

and clearly

$$\frac{I}{AI+IA} \subseteqq \operatorname{ann}\left(\frac{A}{AI+IA}\right),$$

so

$$\frac{A}{AI+IA} \in \mathcal{R} \quad \text{implies} \quad \frac{I}{AI+IA} \in \mathcal{R}.$$

Also

$$\frac{AI+IA}{AI} \cong \frac{IA}{AI \cap IA} \in \mathcal{R},$$

as $IA \in \mathcal{R}$. Hence

$$\left(\frac{I}{AI}\right) \Big/ \left(\frac{AI+IA}{AI}\right) \cong \frac{I}{AI+IA} \in \mathcal{R}$$

implies $I/AI \in \mathcal{R}$. But $AI \in \mathcal{R}$, so $I \in \mathcal{R}$.

2. In a number of cases we get that $\mathcal{R}$ is hereditary for annihilator ideals implies that $\mathcal{R}$ is hereditary. We need some kind of extra condition, otherwise the condition of hereditariness for annihilator ideals would be sufficient for closure under finite subdirect sums. In [1] a counter-example is given.

Theorem 9. *Let $\mathcal{R}$ be a radical class which is hereditary for annihilator ideals. Then $\mathcal{R}$ is hereditary if and only if $I \lhd A \in \mathcal{R}$ implies AI, $IA \in \mathcal{R}$.*

Proof. From the above proof in Theorem 8 we infer that $I \lhd A \in \mathcal{R}$ together with $AI, IA \in \mathcal{R}$ implies $I \in \mathcal{R}$. Hence $\mathcal{R}$ is hereditary. The converse is trivial.

Theorem 10. *Let $\mathcal{R}$ be a radical class which is hereditary for annihilator ideals. Then $\mathcal{R}$ is hereditary if and only if $I \lhd A \in \mathcal{R}$, $I \subseteq A^2$ implies $I \in \mathcal{R}$.*

Proof. Again let $I \lhd A \in \mathcal{R}$. Now $AI \subseteq A^2$, $IA \subseteq A^2$ with both AI and IA ideals in $\mathcal{R}$ imply AI, $IA \in \mathcal{R}$. As $\mathcal{R}$ is hereditary for annihilator ideals, it follows that $I \in \mathcal{R}$ (Theorem 8), so $\mathcal{R}$ is hereditary. The converse is trivial.

Another condition which ensures hereditariness of $\mathcal{R}$ is contained in the following

Theorem 11. *A radical class $\mathcal{R}$ is hereditary if and only if $I \lhd A \in \mathcal{R}$ implies $I \in \mathcal{R}$ whenever $I^2 = (0)$ or $I \subseteq A^2$.*

Proof. This is a direct consequence of Theorem 10, since the condition

$$I \lhd A \in \mathcal{R}, \quad I^2 = (0) \Rightarrow I \in \mathcal{R}$$

yields also

$$I \lhd A \in \mathcal{R}, \quad AI = 0 = IA \Rightarrow I \in \mathcal{R}$$

so that $\mathcal{R}$ is hereditary for annihilator ideals.

Corollary 12. *Let $\mathcal{R}$ be a radical class which contains $\mathcal{B}$. Then $\mathcal{R}$ is hereditary if and only if*

$$I \lhd A \in \mathcal{R}, \quad I \subseteq A^2 \Rightarrow I \in \mathcal{R}.$$

Proof. Let $I \lhd A \in \mathcal{R}$. Now $I/I^2 \in \mathcal{B} \subseteq \mathcal{R}$. But $I^2 \subseteq A^2$, so $I^2 \in \mathcal{R}$, hence $I \in \mathcal{R}$ and $\mathcal{R}$ is hereditary.

We might remark that Corollary 12 is an easy consequence of Theorem 10, since any radical class $\mathcal{R}$ which contains $\mathcal{B}$ is hereditary for annihilator ideals (see the proof of Corollary 6).

The proof of Corollary 12 also indicates the next result:

Corollary 13. *Let $\mathcal{R}$ be a radical class which contains $\mathcal{B}$. Then $\mathcal{R}$ is hereditary if and only if*

$$I \lhd A \in \mathcal{R} \Rightarrow I^2 \in \mathcal{R}.$$

Proof. See Corollary 12.

Theorem 14. *A radical class $\mathcal{R}$ is hereditary if and only if $\mathcal{R}$ is hereditary for annihilator ideals and*

$$\mathcal{R}(A)\big(I \cap \mathcal{R}(A)\big) \subseteq \mathcal{R}(I), \quad \big(I \cap \mathcal{R}(A)\big)\mathcal{R}(A) \subseteq \mathcal{R}(I)$$

for any ideal I in any ring A.

Proof. Obviously if $\mathcal{R}$ is hereditary, then using $I \cap \mathcal{R}(A) = \mathcal{R}(I)$ for any ideal I in any ring A, we get the conditions.

Conversely, let I be an ideal in a ring A. Then $(I\cap\mathcal{R}(A))/\mathcal{R}(I)\lhd\mathcal{R}(A)/\mathcal{R}(I)$ and the second condition implies that $(I\cap\mathcal{R}(A))/\mathcal{R}(I)\subseteqq\operatorname{ann}\mathcal{R}(A)/\mathcal{R}(I)$. Hence, since $\mathcal{R}(A)/\mathcal{R}(I)\in\mathcal{R}$, the first condition gives $(I\cap\mathcal{R}(A))/\mathcal{R}(I)\in\mathcal{R}$. This says $I\cap\mathcal{R}(A)\in\mathcal{R}$ or $I\cap\mathcal{R}(A)\subseteqq\mathcal{R}(I)$. Always $\mathcal{R}(I)\subseteqq I\cap\mathcal{R}(A)$, whence $I\cap\mathcal{R}(A)=\mathcal{R}(I)$ and $\mathcal{R}$ is hereditary.

Corollary 15. *A radical class $\mathcal{R}$ is hereditary if and only if $\mathcal{R}$ is hereditary for annihilator ideals and*

$$I\lhd A\in\mathcal{R},\quad AI+IA\subseteqq\mathcal{R}(I)\Rightarrow I\in\mathcal{R}$$

for any ring $A\in\mathcal{R}$ and any ideal I in A.

Proof. The necessity being trivial, let $I\lhd A\in\mathcal{R}$. Then $\mathcal{R}(A)(I\cap\mathcal{R}(A))=A(I\cap A)\subseteqq AI\subseteqq\mathcal{R}(I)$ and $(I\cap\mathcal{R}(A))\mathcal{R}(A)=(I\cap A)A\subseteqq IA\subseteqq\mathcal{R}(I)$, if $AI+IA\subseteqq\mathcal{R}(I)$ is assumed. Now apply Theorem 14.

It might be noted that Theorem 9 follows directly from Corollary 15. For, if $I\lhd A\in\mathcal{R}$, then AI, $IA\in\mathcal{R}$ implies AI, $IA\subseteqq\mathcal{R}(I)$, so $AI+IA\subseteqq\mathcal{R}(I)$. Corollary 15 gives $I\in\mathcal{R}$ or $\mathcal{R}$ is hereditary.

We conclude this section with a more general result.

Theorem 16. *Let $\mathcal{R}$ and S resp. be radicals such that S-semi-simple rings are $\mathcal{R}$-radical. Then $\mathcal{R}$ is hereditary if and only if*

$$I\lhd A\in\mathcal{R},\quad I\subseteqq\mathrm{S}(A)\Rightarrow I\in\mathcal{R}$$

for any ring $A\in\mathcal{R}$ and any ideal I in A.

Proof. Suppose the condition be satisfied and assume that $I\lhd A\in\mathcal{R}$. As $I/\mathrm{S}(I)$ is S-semi-simple, we have $I/\mathrm{S}(I)\in\mathcal{R}$. Now $\mathrm{S}(I)\lhd A\in\mathcal{R}$ and $\mathrm{S}(I)\subseteqq\mathrm{S}(A)$, so $\mathrm{S}(I)\in\mathcal{R}\Rightarrow I\in\mathcal{R}$. Then $\mathcal{R}$ is hereditary. The converse is obvious.

Example. Let $\mathcal{R}$ be the class of idempotent rings, i.e. the rings A with $A^2=A$. Let S be the upper radical determined by the Boolean rings. A ring A is called a Boolean ring if $a^2=a$ for every element $a\in A$. Since Boolean rings form a special class of rings, S is a special radical and the S-semi-simple rings are subdirect sums of Boolean rings, so they are again Boolean rings. Any Boolean ring is idempotent, hence any S-semi-simple ring is $\mathcal{R}$-radical. It is known that $\mathcal{R}$ is not hereditary. If we take the subradical class $\mathcal{R}'$ (of $\mathcal{R}$) of the hereditarily idempotent rings, we get a hereditary radical $\mathcal{R}'$. Again any S-semi-simple ring is $\mathcal{R}'$-radical, as any Boolean ring is hereditarily idempotent. (If A is a Boolean ring and $I\lhd A$, then I is again a Boolean ring and idempotent).

3. It is known that for any radical $\mathcal{R}$ there exists a unique maximal hereditary radical $h_{\mathcal{R}}$, contained in $\mathcal{R}$. In [3] it is shown that $h_{\mathcal{R}}=\overline{\overline{\mathcal{R}}}$, where $\overline{\mathcal{R}}=\{A|$any ideal of A is in $\mathcal{R}\}$. It can easily be proved that $\bar{\mathcal{R}}$ is a radical and $\mathcal{R}$ is hereditary if and only if $\mathcal{R}=\bar{\mathcal{R}}$. Let $(\mathcal{S}\mathcal{R})_k$ be the essential closure of the semisimple class $\mathcal{S}\mathcal{R}$ of the radical $\mathcal{R}$. A ring $A\in(\mathcal{S}\mathcal{R})_k$ if A has an essential ideal $B\in\mathcal{S}\mathcal{R}$.

Lemma 17. *For any radical* $\mathcal{R}$, $\bar{\mathcal{R}}=\mathcal{U}(\mathcal{S}\mathcal{R})_k$ *(upper radical).*

Proof. Let $A\in\bar{\mathcal{R}}$ and suppose that $A\notin\mathcal{U}(\mathcal{S}\mathcal{R})_k$. Then there exists a non-zero homomorphic image $A/I\in(\mathcal{S}\mathcal{R})_k$ and A/I has an essential ideal $B/I\in\mathcal{S}\mathcal{R}$. But $A\in\bar{\mathcal{R}}$, so $A/I\in\bar{\mathcal{R}}$. By definition of $\bar{\mathcal{R}}$, it follows that $B/I\in\mathcal{R}$, which implies $B/I\in\mathcal{R}\cap\mathcal{S}\mathcal{R}=0$. Since this is impossible for an essential ideal, we get that $A\in\mathcal{U}(\mathcal{S}\mathcal{R})_k$.

Conversely, assume that $A\in\mathcal{U}(\mathcal{S}\mathcal{R})_k$. If $A\notin\bar{\mathcal{R}}$, A has a non-zero ideal I, $I\notin\mathcal{R}$. Then $0\neq I/\mathcal{R}(I)$ is an ideal in $A/\mathcal{R}(I)$ and $I/\mathcal{R}(I)\in\mathcal{S}\mathcal{R}$. Now there exists a homomorphic image A/J of $A/\mathcal{R}(I)$ containing an isomorphic copy of $I/\mathcal{R}(I)$, such that this copy is an essential ideal in A/J. But $A\in\mathcal{U}(\mathcal{S}\mathcal{R})_k$ implies that $A/J\in\mathcal{U}(\mathcal{S}\mathcal{R})_k$, hence $A/J\in\mathcal{U}(\mathcal{S}\mathcal{R})_k\cap(\mathcal{S}\mathcal{R})_k=0$ or $A=J$. Contradiction, so $A\in\bar{\mathcal{R}}$ and $\bar{\mathcal{R}}=\mathcal{U}(\mathcal{S}\mathcal{R})_k$.

For our next result we use the notation of [1]. $\mathcal{R}$ is a radical class.

$$\mathcal{G}^0_{\mathcal{R}}:=\{(S,A)|\quad S\lhd A \quad\text{and}\quad S\in\mathcal{S}\mathcal{R}\},$$

$$\bar{\mathcal{G}}^0_{\mathcal{R}}:=\{A|\quad\text{every}\quad 0\neq A/I\quad\text{has no nonzero ideals in}\quad\mathcal{S}\mathcal{R}\}.$$

$\bar{\mathcal{G}}^0_{\mathcal{R}}$ is a radical class [1].

$$\mathcal{G}^1_{\mathcal{R}}:=\{(S,A)|\quad S\lhd A\quad\text{and}\quad S\in\mathcal{S}(\bar{\mathcal{G}}^0_{\mathcal{R}})\},$$

$$\bar{\mathcal{G}}^1_{\mathcal{R}}:=\{A|\quad\text{every}\quad 0\neq A/I\quad\text{has no nonzero ideals in}\quad\mathcal{S}(\bar{\mathcal{G}}^0_{\mathcal{R}})\};$$

$\bar{\mathcal{G}}^1_{\mathcal{R}}$ is a radical class [1].

Continuing in this way, one gets a chain of radical classes:

$$\mathcal{R}\supseteqq\bar{\mathcal{G}}^0_{\mathcal{R}}\supseteqq\ldots\supseteqq\bar{\mathcal{G}}^n_{\mathcal{R}}\supseteqq\ldots.$$

In [1] it was shown that $\bigcap_n \bar{\mathcal{G}}^n_{\mathcal{R}}$ is the unique maximal hereditary radical subclass of $\mathcal{R}$. An improvement of this result is given in the next theorem.

Theorem 18. *For any radical class* $\mathcal{R}$ *we have:* $\bar{\mathcal{G}}^1_R$ *is the unique maximal hereditary radical subclass of* $\mathcal{R}$.

Proof. We show that, with the above notation, $\bar{\mathcal{G}}^0_{\mathcal{R}}=\bar{\mathcal{R}}$. Let $A\in\bar{\mathcal{G}}^0_{\mathcal{R}}$. Since for any $I\lhd A$ we have $\mathcal{R}(I)\lhd A$ and $I/\mathcal{R}(I)\in\mathcal{S}\mathcal{R}$, the assumption $A\in\bar{\mathcal{G}}^0_{\mathcal{R}}$ yields $I/\mathcal{R}(I)=0$. Thus $A\in\bar{\mathcal{R}}$.

Conversely, let $A\in\overline{\mathscr{R}}$ and take any $0\neq A/I$. If A/I has a nonzero ideal $B(I)\in\in\mathscr{S}\mathscr{R}$, then $A/I\in\overline{\mathscr{R}}$ yields that $B/I\in\mathscr{R}\cap\mathscr{S}\mathscr{R}=0$, which is a contradiction. Hence $0\neq A/I$ has no nonzero ideals in $\mathscr{S}\mathscr{R}$, i.e. $A\in\bar{\mathscr{G}}^0_{\mathscr{R}}$. Using Lemma 17 we have established: $\overline{\mathscr{R}}=\mathscr{U}(\mathscr{S}\mathscr{R})_k=\bar{\mathscr{G}}^0_{\mathscr{R}}$. Apply now Lemma 17 again to the radical $\bar{\mathscr{G}}^0_{\mathscr{R}}$: $\overline{\overline{\mathscr{R}}}=\mathscr{U}(\mathscr{S}\overline{\mathscr{R}})_k=\bar{\mathscr{G}}^0_{\overline{\mathscr{R}}}$. From $\overline{\mathscr{R}}=\bar{\mathscr{G}}^0_{\mathscr{R}}$ and the definitions of $\mathscr{G}^0_{\mathscr{R}}$ and $\mathscr{G}^1_{\mathscr{R}}$ resp. we infer that $\mathscr{G}^0_{\overline{\mathscr{R}}}=\mathscr{G}^1_{\mathscr{R}}$. Hence we get: $\bar{\mathscr{G}}^0_{\overline{\mathscr{R}}}=\bar{\mathscr{G}}^1_{\mathscr{R}}$ or $\bar{\mathscr{G}}^1_{\mathscr{R}}=\overline{\overline{\mathscr{R}}}$, which is the unique maximal hereditary subradical of $\mathscr{R}$.

Note that the above chain now reads:

$$\mathscr{R}\supseteqq\overline{\mathscr{R}}\supseteqq\overline{\overline{\mathscr{R}}}=\overline{\overline{\overline{\mathscr{R}}}}=\dots$$

since $\bigcap_n\bar{\mathscr{G}}^n_{\mathscr{R}}=\bar{\mathscr{G}}^1_{\mathscr{R}}=\overline{\overline{\mathscr{R}}}$.

An example in [1] shows that, in general, $\bar{\mathscr{G}}^0_{\mathscr{R}}=\overline{\mathscr{R}}$ need not be hereditary. In fact, $\bar{\mathscr{G}}^0_{\mathscr{R}}$ is hereditary if and only if $\bar{\mathscr{G}}^0_{\mathscr{R}}=\bar{\mathscr{G}}^1_{\mathscr{R}}$ or, in our notation, $\overline{\mathscr{R}}$ is hereditary if and only if $\overline{\mathscr{R}}=\overline{\overline{\mathscr{R}}}$.

Theorem 19. *If a radical class $\mathscr{R}$ is hereditary for annihilator ideals, then $\overline{\mathscr{R}}$ is hereditary.*

Proof. Let A be a zero-ring and suppose that $A\in\mathscr{R}$. Then any ideal I of A is in $\mathscr{R}$, so $A\in\overline{\mathscr{R}}$. Therefore any zero-ring in $\mathscr{R}$ is in $\overline{\mathscr{R}}$, which implies $\overline{\mathscr{R}}=\overline{\overline{\mathscr{R}}}$ ([3], Proposition 1 and Corollary 1).

The next result is well-known. For a radical class $\mathscr{R}$ the following are equivalent.

a) $\mathscr{R}$ contains all zero-rings;

b) $\mathscr{R}$ contains all nilpotent rings;

c) $\mathscr{B}\subseteqq\mathscr{R}$.

The above proof of Theorem 19 indicates that any radical class $\mathscr{R}$ containing all zero-rings satisfies: $\overline{\mathscr{R}}$ is hereditary. So we get

Corollary 20. *Let $\mathscr{R}$ be a radical with $\mathscr{B}\subseteqq\mathscr{R}$. Then $\overline{\mathscr{R}}$ is the maximal hereditary subradical of $\mathscr{R}$.*

Proof. Obviously $\mathscr{B}\subseteqq\mathscr{R}$ implies that $\mathscr{R}$ is hereditary for annihilator ideals, so Corollary 20 is a direct consequence of Theorem 19.

Remark. We will see that the condition of Theorem 19 for hereditariness of $\overline{\mathscr{R}}$ is not necessary (after Theorem 24).

The counterpart is formed by the radicals $\mathscr{R}$ containing no nonzero zero-rings.

Lemma 21. *For a radical class $\mathscr{R}$ the following are equivalent:*

a) *$\mathscr{R}$ contains no nonzero zero-rings;*

b) $\mathcal{R}$ *contains no nonzero nilpotent rings;*
c) $\mathcal{R}$ *is subidempotent i.e. any ring* A *in* $\mathcal{R}$ *is idempotent.*

Proof. Since the proof is straightforward, we omit it.

In order to study radicals $\mathcal{R}$ with the above property, we introduce

$$\mathcal{G}_{\mathcal{R}} := \{(S, A) |\quad S \in \mathcal{S}\mathcal{R} \quad \text{and} \quad S \subseteqq \operatorname{ann} A\},$$

where S is a subring of A. This implies $S \lhd A$.

$\bar{\mathcal{G}}_{\mathcal{R}} := \{A| \quad \text{every} \quad 0 \neq A/I \quad \text{has no nonzero ideals in} \quad \operatorname{ann}(A/I) \quad \text{and in} \quad \mathcal{S}\mathcal{R}\}$.
Then $\bar{\mathcal{G}}_{\mathcal{R}}$ is a radical class and $\mathcal{R} \cap \bar{\mathcal{G}}_{\mathcal{R}}$ is the maximal radical subclass of $\mathcal{R}$ which is hereditary for annihilator ideals ([1], Proposition 1.8).

Define

$$\mathcal{E}_6 := \{A| \quad \text{every} \quad 0 \neq A/I \quad \text{has} \quad \operatorname{ann}(A/I) = 0\}.$$

Then $\mathcal{E}_6$ is a radical class (cf. [4]). It is clear that for any radical $\mathcal{R}$ one has: $\mathcal{E}_6 \subseteqq \bar{\mathcal{G}}_{\mathcal{R}}$. The next lemma shows that equality holds for subidempotent radicals $\mathcal{R}$.

Lemma 22. *Let* $\mathcal{R}$ *be a subidempotent radical. Then* $\bar{\mathcal{G}}_{\mathcal{R}} = \mathcal{E}_6$.

Proof. We only need to prove that $\bar{\mathcal{G}}_{\mathcal{R}} \subseteqq \mathcal{E}_6$. Let $A \in \bar{\mathcal{G}}_{\mathcal{R}}$ and take any $0 \neq$ $\neq A/I = \bar{A}$. Then $\operatorname{ann} \bar{A}/\mathcal{R}(\operatorname{ann} \bar{A}) \subseteqq \operatorname{ann}(\bar{A}/\mathcal{R}(\operatorname{ann} \bar{A}))$. Since $A \in \bar{\mathcal{G}}_{\mathcal{R}}$, it follows that $\operatorname{ann} \bar{A}/\mathcal{R}(\operatorname{ann} \bar{A}) = 0$, so $\operatorname{ann} \bar{A} \in \mathcal{R}$. But $(\operatorname{ann} \bar{A})^2 = 0$, so $\operatorname{ann} \bar{A} = 0$, as $\mathcal{R}$ is subidempotent. Hence $A \in \mathcal{E}_6$.

In general one can show that

$$\bar{\mathcal{G}}_{\mathcal{R}} = \{A| \quad \text{any} \quad 0 \neq A/I \quad \text{has the property:} \ J/I \lhd A/I,\ J/I \subseteqq \operatorname{ann}(A/I) \Rightarrow J/I \in \mathcal{R}\}.$$

From the definitions of $\mathcal{G}_{\mathcal{R}}$ and $\mathcal{G}^0_{\mathcal{R}}$ resp. we get immediately: $\mathcal{G}_{\mathcal{R}} \subseteqq \mathcal{G}^0_{\mathcal{R}}$ yielding $\bar{\mathcal{G}}^0_{\mathcal{R}} \subseteqq \bar{\mathcal{G}}_{\mathcal{R}}$ for any radical $\mathcal{R}$. Always $\bar{\mathcal{G}}^0_{\mathcal{R}} \subseteqq \mathcal{R}$, hence $\bar{\mathcal{R}} = \bar{\mathcal{G}}^0_{\mathcal{R}} \subseteqq \mathcal{R} \cap \bar{\mathcal{G}}_{\mathcal{R}}$ for any radical $\mathcal{R}$.

In the following theorem we will give a sufficient condition in order that $\bar{\mathcal{R}} = \mathcal{R} \cap \bar{\mathcal{G}}_{\mathcal{R}}$.

Theorem 23. *Let* $\mathcal{R}$ *be a radical class such that* $A \in \mathcal{R}$ *implies* $AS, SA \in \mathcal{R}(S)$ *for any ring* A *and any ideal* S *in* A. *Then* $\bar{\mathcal{R}} = \mathcal{R} \cap \bar{\mathcal{G}}_{\mathcal{R}}$ *and* $\bar{\mathcal{R}}$ *is the unique maximal radical subclass of* $\mathcal{R}$ *which is hereditary for annihilator ideals.*

Proof. We have to show that $A \in \mathcal{R} \cap \bar{\mathcal{G}}_{\mathcal{R}}$ implies $A \in \bar{\mathcal{R}}$. Assume $A \in \mathcal{R} \cap \bar{\mathcal{G}}_{\mathcal{R}}$ and let $S \lhd A$. Then $S/\mathcal{R}(S) \lhd A/\mathcal{R}(S) \in \bar{\mathcal{G}}_{\mathcal{R}}$, as $A \in \bar{\mathcal{G}}_{\mathcal{R}}$. Also $S/\mathcal{R}(S) \subseteqq \operatorname{ann}(A/\mathcal{R}(S))$, as $AS, SA \in \mathcal{R}(S)$. Hence $S/\mathcal{R}(S) \in \mathcal{R}$, as $A \in \bar{\mathcal{G}}_{\mathcal{R}}$ (see the above characterization of $\bar{\mathcal{G}}_{\mathcal{R}}$). Therefore $S = \mathcal{R}(S)$ or $S \in \mathcal{R}$. It follows that $A \in \bar{\mathcal{R}}$. By Proposition 1.8 [1] $\bar{\mathcal{R}} = \mathcal{R} \cap \bar{\mathcal{G}}_{\mathcal{R}}$ has the required property of maximality.

We have seen that radicals $\mathcal{R}$ with $\mathcal{B} \subseteq \mathcal{R}$ have the property that $\bar{\mathcal{R}}$ is the maximal hereditary subradical of $\mathcal{R}$. Our final result contains another class of radicals $\mathcal{R}$ for which this phenomenon occurs.

Theorem 24. *Let $\mathcal{R}$ be a subidempotent radical. Then $\mathcal{R}$ is hereditary for annihilator ideals if and only if $\mathcal{R} \subseteq \mathcal{E}_6$.*

For any subidempotent radical $\mathcal{R}$ we have that $\bar{\mathcal{R}}$ is the maximal hereditary subradical of $\mathcal{R}$. $\bar{\mathcal{R}}$ is a hereditarily idempotent radical.

Proof. From [1], Proposition 1.8 it follows that $\mathcal{R}$ is hereditary for annihilator ideals if and only if $\mathcal{R} \subseteq \bar{\mathcal{G}}_{\mathcal{R}}$ for any radical $\mathcal{R}$. So for a subidempotent radical we get the first result immediately from Lemma 22. Now let $\mathcal{R}$ be an arbitrary subidempotent radical. Take any ring $A \in \mathcal{R}$. If $A^2=0$, then $A=0$, so any zero-$\mathcal{R}$-ring is in $\bar{\mathcal{R}}$, hence $\bar{\mathcal{R}}$ is hereditary ([3], Proposition 1) and $\bar{\mathcal{R}}$ is a hereditarily idempotent radical.

Remark. As not every subidempotent radical $\mathcal{R}$ is contained in $\mathcal{E}_6$, it follows that a subidempotent radical $\mathcal{R}$ need not be hereditary for annihilator ideals. This shows that the sufficient condition in Theorem 19 is not necessary.

In the light of the previous results we examine the Examples 1.4 and 1.5 in [1]. Consider the ring R whose additive group is $Q+Q$ (direct sum) and whose multiplication is given by

$$(a, b)(c, d) = (ac, ad+bc).$$

The homomorphic images of R are 0, Q and R, while the ideals of R are 0, $I\ (\cong Q^0)$ and R (Q^0 is the zero-ring on Q).

Let $\mathcal{D}$ be the (radical) class of rings with divisible additive groups. Then both R and I are in $\mathcal{D}$. Since I is the only non-trivial ideal in R, we get that $R \in \bar{\mathcal{D}}$. However, $I \notin \bar{\mathcal{D}}$, as $I(\cong Q^0)$ has non-zero reduced ideals. So $\bar{\mathcal{D}}$ is not hereditary. By Theorem 19 we get that $\mathcal{D}$ is not hereditary for annihilator ideals. Note that $\mathcal{B}$ is not contained in $\mathcal{D}$, since $Z^0 \notin \mathcal{D}$, $Z^0 \in \mathcal{B}$ (Z^0 is the zero-ring on Z). In addition, $\mathcal{D}$ is not subidempotent, since $I \in \mathcal{D}$, but $I^2=0$. This is in accordance with Corollary 20 and Theorem 24, since any radical $\mathcal{R}$ containing $\mathcal{B}$ or being subidempotent has a hereditary subradical $\bar{\mathcal{R}}$.

We also consider the lower radical class $L(\{R\})$, determined by R. Now R is a non-simple ring with identity $(1, 0)$. Since I is the only non-trivial ideal of R and $R/I \cong Q$, Q not isomorphic to R, we see that R satisfies the conditions (i) and (ii) of Theorem 3.5 in [1]. Hence $L(\{R\})$ is not closed under finite subdirect sums.

On the other hand, R is idempotent and $R/I \cong Q$ is idempotent, so that $R \in \mathcal{E}_6$. Therefore $L(\{R\}) \subseteq \mathcal{E}_6$. Also $L(\{R\})$ is a subidempotent radical, as any radical contained in $\mathcal{E}_6$ is subidempotent. Hence $L(\{R\})$ is hereditary for annihilator ideals (Theorem 24).

References

[1] B. J. Gardner and P. N. Stewart, The closure of radical classes under finite subdirect products, *Compositio Math.*, **46** (1982), 145—158.
[2] L. Márki and R. Wiegandt, Remarks on radicals in categories, in: *Category Theory*, Lect. Notes in Math., 962, Springer (Berlin—Heidelberg—New York, 1982); 190—196.
[3] E. Puczyłowski, A note on hereditary radicals, *Acta Sci. Math.*, **44** (1982), 133—135.
[4] F. Szász, *Radikale der Ringe*, Akadémiai Kiadó (Budapest, 1976).
[5] R. Wiegandt, *Radical and semisimple classes of rings*, Queen's Papers in Pure and Applied Mathematics, № 37 (Kingston, Ontario, 1974).

RIJKSUNIVERSITEIT TE GRONINGEN
MATHEMATISCH INSTITUUT
POSTBUS 800
GRONINGEN, THE NETHERLANDS

Acta Sci. Math. 47 (1984), 321—348

Classification and construction of complete hypersurfaces satisfying $R(X, Y)\cdot R=0$

Z. I. SZABÓ

In one of his papers K. NOMIZU [3] examined the immersed hypersurfaces in $\mathbf{R}^{n+1}$ satisfying $R(X, Y)\cdot R=0$ for all tangent vectors X; Y, where the curvature endomorphism $R(X, Y)$ operates on R as a derivation of the tensor algebra at each point of the manifold. The main theorem of Nomizu's paper is the following.

Theorem (K. Nomizu). *Let M be an n-dimensional, connected, complete Riemannian manifold, which is isometrically immersed in $\mathbf{R}^{n+1}$ so that the type number is greater than 2 at least at one point. If M satisfies the condition $R(X, Y)\cdot R=0$ then it is of the form $M=S^k\times\mathbf{R}^{n-k}$, where S^k is a hypersphere in a euclidean subspace $\mathbf{R}^{k+1}$ of $\mathbf{R}^{n+1}$ and $\mathbf{R}^{n-k}$ is a euclidean subspace orthogonal to $\mathbf{R}^{k+1}$.*

This theorem inspired the so called Nomizu conjecture: Every irreducible complete space with $\dim\geqq 3$ and $R(X, Y)\cdot R=0$ is locally symmetric.

But the answer for this conjecture was negative as H. TAKAGI [6] constructed a 3-dimensional counterexample. This counterexample is a connected complete immersed hypersurface in $\mathbf{R}^4$. Thus the problem is to determine all the connected complete n-dimensional immersed hypersurfaces in $\mathbf{R}^{n+1}$ satisfying $R(X, Y)\cdot R=0$, the description of which completes Nomizu's theorem. The main purpose of this paper is to give a complete description and classification of these hypersurfaces.

1. Basic formulas

A C^∞ Riemannian manifold*) (M^n, g) with the property $R(X, Y)\cdot R=0$ is called a semisymmetric manifold. Let us assume that the semisymmetric manifold (M^n, g) is an immersed hypersurface in $\mathbf{R}^{n+1}$. Let $\mathbf{n}$ be a normal unit vector field on a connected orientable neighbourhood U of M^n. If D resp. ∇ denotes the Riemannian

Received February 7, 1983.

*) The notion differentiable is used in the meaning C^∞.

covariant derivative in $\mathbf{R}^{n+1}$ resp. in M^n, then

$$(1.1)\qquad \begin{aligned} D_X Y &= \nabla_X Y + H(X, Y)\mathbf{n}, \\ D_X \mathbf{n} &= A(X), \quad H(X, Y) = -g(A(X), Y) \end{aligned}$$

holds, for all differentiable vector fields X, Y on U tangent to M^n. $H(X, Y)$ is the so-called second fundamental form of the hypersurface, and $A(X)$ is the so-called Weingarten field. The $A(X)$ is a symmetric endomorphism's field on the manifold. The rank of A at a point $p \in M^n$ is called the type number at p and it is denoted by $k(p)$.

The curvature tensor field $R(X, Y)Z$ of M^n is of the form

$$(1.2)\qquad R(X, Y)Z = -g(A(X), Z)A(Y) + g(A(Y), Z)A(X)$$

by the Gauss' equation.

The nullspace of the cuvature operator at a point p consists of vetors $Z \in T_p(M)$ for which $R(X, Y)Z = 0$ holds for all vectors $X; Y \in T_p(M)$. The dimension of the nullspace at p is called the index of nullity, and it is denoted by $i(p)$. If $k(p)$ is 0 or 1, then $R_p = 0$ holds, and $i(p) = n$ in this case. But if $k(p) > 1$ holds, then $k(p) = n - i(p)$ (see in [2], p. 42).

It is not hard to see, that all the hypersurfaces with $k(p) \leqq 2$ (or equivalently $i(p) \geqq n-2$) are semisymmetric. By Nomizu's theorem every connected, complete immersed semisymmetric hypersurface M^n in $\mathbf{R}^{n+1}$ is a cylinder, if at least at one point p, $k(p) > 2$ holds, so in what follows we examine only the hypersurfaces for which $k(p) \leqq 2$ holds at every point $p \in M^n$.

If at a point $k(p) = 2$ holds, then $i(p) = n-2$. Let λ_1 and λ_2 be the two non-trivial eigenvalues of A_p, and let $\mathbf{x}_1, \mathbf{x}_2$ be the corresponding orthogonal unit eigenvectors. If V_p^1 denotes the 2-dimensional subspace spanned by $\mathbf{x}_1$ and $\mathbf{x}_2$, then the orthogonal complement V_p^0 of V_p^1 is just the nullspace of the curvature operator, and also

$$T_p(M) = V_p^0 + V_p^1$$

holds. This direct sum is called the *V-decomposition* of the tangent space $T_p(M)$. Since $k(p) \leqq 2$ holds everywhere, and the eigenvalue functions $\lambda_1(q) \leqq \lambda_2(q)$ are continuous, so $k(q) = 2$ holds in a neighbourhood of p. I.e. the set, where $k(q) = 2$ holds, is an open set U in M^n. If we consider the above V-decomposition on U, then the distributions V^i, $i = 0; 1$, are differentiable, since V^1 is spanned by the vector fields of the form $R(X, Y)Z$.

The V-decomposition is defined at the points p with $k(p) < 2$ by the trivial decomposition $T_p(M) = V_p^0$.

Further on we examine the hypersurface on the open set U, where $k(q) = 2$ holds.

The following relations are simple consequences of the Bianchi identity

$\sigma(\nabla_X R)(Y, Z)=0$:

$$\nabla_{V^0}V^1 \subseteqq V^1, \quad \nabla_{V^0}V^0 \subseteqq V^0, \quad \nabla_{V^1}V^1 \subseteqq V^0+V^1 = T(M), \tag{1.3}$$

where the formula $\nabla_{V^i}V^j \subseteqq V^k$ means that for the differentiable vector fields X_l, tangent to V^l, the vector field $\nabla_{X_i}X_j$ is tangent to V^k.

We mention that the distribution V^1 is in general not integrable, but by the second relation in (1.3) it follows, that the distribution V^0 on U is always integrable and the integral manifolds are totally geodesic and locally euclidean submanifolds. From the first formula in (1.3) we can see too, that the distribution V^1 is parallel along the curves which are going in the above totalgeodesic integral manifolds of V^0.

Now let us consider a local system $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_{n-2}$ of differentiable unit vector fields tangent to V^0 which are paarwise orthogonal, furthermore, also $\nabla_{\mathbf{m}_\alpha}\mathbf{m}_\beta=0$ hold. From the above considerations it follows, that such a vector field system exists around every point of U.

Next we introduce some basic formulas w.r.t. the system $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_{n-2}$. For the differentiable vector fields X; Y tangent to V^1 we can write

$$\nabla_X \mathbf{m}_\alpha = B_\alpha(X) + \sum_\beta M_\alpha^\beta(X)\mathbf{m}_\beta, \quad \text{where} \quad B_\alpha(X)_{/p} \in V_p^1, \tag{1.4}$$

$$\nabla_X Y = \tilde{\nabla}_X Y + \sum_\alpha M^\alpha(X, Y)\mathbf{m}_\alpha, \quad \text{where} \quad \tilde{\nabla}_X Y_{/p} \in V_p^1. \tag{1.5}$$

Using these formulas we define the tensor fields B_α, M^α, M_α^β and the covariant derivative $\tilde{\nabla}$ only on the distribution V^1.

But let us extend these tensor fields and this covariant derivative over the whole tangent bundle in such a way that $B_\alpha(\mathbf{m}_p)=0$, $M_\alpha^\beta(\mathbf{m}_\gamma)=0$, $M^\alpha(\mathbf{m}_\beta, X)= =M^\alpha(\mathbf{m}_\beta, \mathbf{m}_\gamma)=0$ and $\tilde{\nabla}_{\mathbf{m}_\alpha}X=\nabla_{\mathbf{m}_\alpha}X, \tilde{\nabla}_{\mathbf{m}_\alpha}\mathbf{m}_\beta=0$ hold. Then the fields B_α, M^α, M_α^β are differentiable tensor fields indeed, furthermore, $\tilde{\nabla}$ is a metrical covariant derivative, i.e: $\tilde{\nabla}g=0$ holds. The following formulas are also obvious:

$$M^\alpha(X, Y) = -g\big(B_\alpha(X), Y\big), \quad M_\beta^\alpha(X) = -M_\alpha^\beta(X). \tag{1.6}$$

We leave the proof of these facts to the reader. Let $\tilde{R}(X, Y)Z$ be the curvature tensor of $\tilde{\nabla}$.

Proposition 1.1. *For differentiable vector fields X, Y, Z tangent to V^1 the tensor fields $B_\alpha, M^\alpha, M_\beta^\alpha, \tilde{R}$ satisfy the following basic formulas:*

$$R(X, Y)Z = \tilde{R}(X, Y)Z + \sum_\alpha \{M^\alpha(Y, Z)B_\alpha(X) - M^\alpha(X, Z)B_\alpha(Y)\}, \tag{1.7}$$

$$(\tilde{\nabla}_X B_\alpha)(Y) - (\tilde{\nabla}_Y B_\alpha)(X) = \sum_\beta \{M_\alpha^\beta(X)B_\beta(Y) - M_\alpha^\beta(Y)B_\beta(X)\}, \tag{1.8}$$

$$\begin{gathered} dM_\alpha^\beta(X, Y) = \\ = \sum_\gamma M_\gamma^\beta(X)\wedge M_\alpha^\gamma(Y) - (1/2)\big\{M^\beta(X, B_\alpha(Y)) - M^\beta(Y, B_\alpha(X))\big\}, \end{gathered} \tag{1.9}$$

(1.10) $$(\nabla_{\mathbf{m}_\alpha} B_\beta)(X) = -B_\beta \circ B_\alpha(X),$$

(1.11) $$(\nabla_{\mathbf{m}_\alpha} M_\beta^\gamma)(X) = -M_\beta^\gamma(B_\alpha(X)),$$

(1.12) $$\tilde{R}(\mathbf{m}_\alpha, X)Y = 0,$$
$$\text{i.e.}\quad \nabla_{\mathbf{m}_\alpha}\tilde{\nabla}_X Y = \tilde{\nabla}_X \nabla_{\mathbf{m}_\alpha} Y + \tilde{\nabla}_{\nabla_{\mathbf{m}_\alpha} X} Y - \tilde{\nabla}_{B_\alpha(X)} Y - \sum_\beta M_\alpha^\beta(X)\nabla_{\mathbf{m}_\beta} Y,$$

(1.13) $$(\nabla_{\mathbf{m}_\alpha} R)(X, Y) = R(Y, B_\alpha(X)) + R(B_\alpha(Y), X),$$

where d is the exterior derivative and the symbol $\wedge$ denotes the skew-product.

The complete proof of these formulas is contained in [4]. But we mention, that (1.7) follows by (1.4) and (1.5) from the formula $R(X, Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z$, the formulas (1.8)—(1.12) are equivalent to the identities $R(X, Y)\mathbf{m}_\alpha = 0$, $R(\mathbf{m}_\alpha, X)Y = 0$, $R(\mathbf{m}_\alpha, X)\mathbf{m}_\beta = 0$, and formula (1.13) follows from the Bianchi identity and from (1.4) in the following manner:

$$(\nabla_{\mathbf{m}_\alpha} R)(X, Y) = -(\nabla_X R)(Y, \mathbf{m}_\alpha) - (\nabla_Y R)(\mathbf{m}_\alpha, X) = R(Y, B_\alpha(X)) + R(B_\alpha(Y), X).$$

Here the details are also left to the reader.

2. Reduction of the basic formulas

Further on let us examine the complete connected semisymmetric hypersurface M^n in $\mathbf{R}^{n+1}$ with $k(p) \leqq 2$ on the open set U, where $k(p)=2$ and thus $R_p(X, Y)Z \neq 0$ holds. Let us consider also the V-decomposition $T(M) = V^0 + V^1$ on U and for a point $p \in U$ let us consider the maximal connected integral manifold N of V^0 through a point p. If $c(s)$ is a differentiable curve in N, parametrized by arc-length and if $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_{n-2}$ is a vector field system around $c(s)$ defined in the previous chapter, then for the tangent vector $\dot{c}(s) = \sum_\alpha a^\alpha(s)\mathbf{m}_\alpha$ the tensor, defined by

(2.1) $$B_{\dot{c}(s)} := \sum_\alpha a^\alpha(s) B_{\alpha/c(s)},$$

is uniquely determined, and it is independent from the choice of the system $\mathbf{m}_1, \ldots, \mathbf{m}_{n-2}$ around $c(s)$. Indeed if $\tilde{\mathbf{m}}_1, \ldots, \tilde{\mathbf{m}}_{n-2}$ is another system around $c(s)$ with $\tilde{\mathbf{m}}_\alpha = \sum_\beta b_\alpha^\beta \mathbf{m}_\beta$, and the corresponding tensors w.r.t. this system are denoted by $\tilde{B}_\alpha$, then from

$$\tilde{B}_\alpha = \sum_\beta b_\alpha^\beta B_\beta, \quad \mathbf{m}_\beta = \sum_\alpha (b^{-1})_\beta^\alpha \tilde{\mathbf{m}}_\alpha, \quad \dot{c}(s) = \sum_\beta a^\beta(s)\mathbf{m}_\beta = \sum_{\alpha,\beta} a^\beta (b^{-1})_\beta^\alpha \mathbf{m}_\alpha$$

we get

$$\sum_{\alpha,\beta} a^\beta (b^{-1})_\beta^\alpha \tilde{B}_\alpha = \sum_{\alpha,\beta,\gamma} a^\beta (b^{-1})_\beta^\alpha b_\alpha^\gamma B_\gamma = \sum_\alpha a^\alpha B_\alpha,$$

which proves the above statement.

Let us notice too, that the curvature tensor $R(X, Y)Z$ is of the form

$$R(X, Y)Z = K\big(g(Y, Z)X - g(X, Z)Y\big), \quad X; Y; Z \in V^1, \tag{2.2}$$

on V^1, where $K(p)$ is the sectional curvature w.r.t. the section V_p^1 at p. From (1.13) it follows that the function $K(s)=K\big(c(s)\big)$ satisfies the differential equation

$$\frac{dK}{ds} = -(\mathrm{Tr}\, B_{\dot c})K, \tag{2.3}$$

and thus we have

$$K(s) = K(0)e^{-\int\limits_0^s \mathrm{Tr}\, B_{\dot c}\, ds}. \tag{2.4}$$

From this formula we get, that K is zero neither on N nor on the boundary of N, and thus the boundary of N is inside of U. But N is maximal, thus N cannot have boundary points. As the space is complete, N is a complete, connected, locally euclidean and totally geodesic submanifold in the maifold M. On the other hand the second fundamental form A vanishes on the tangent spaces of N, further V^1 is totally parallel along N, thus N is an open subset in an $(n-2)$-dimensional euclidean subspace $\mathbf{R}^{n-2}$ of $\mathbf{R}^{n+1}$. But because of the completeness of N it must be equal to the whole euclidean subspace $\mathbf{R}^{n-2}$, and thus we have

Proposition 2.1. *Every maximal integral manifold N of V^0, through a point p, where $R_p \neq 0$ holds, is complete, totally geodesic and isometric with $\mathbf{R}^{n-2}$. In addition N is an $(n-2)$-dimensional euclidean subspace in $\mathbf{R}^{n+1}$. The curvature tensor R_p of the space M^n never vanishes at the points of such a submanifold N.*

Now let $c(s)$, $-\infty < s < \infty$, be a complete geodesic in a subspace N, considered in the above proposition and parametrised by arc-length s. Let us consider also $B_{\dot c}$ along $c(s)$ defined in (2.1). Then

$$\nabla_{\dot c} B_{\dot c} = -B_{\dot c}^2 \tag{2.5}$$

holds. From this equation it follows, that $B_{\dot c}$ never vanishes along $c(s)$ if it is non-zero at a point $c(s_0)$, and so it is a zero-field, if it is zero at a point. Let us remember too, that V^1 is invariant under the action of $B_{\dot c}$, and that also $B_{\dot c}(V^0)=0$ holds.

Next we solve the differential equtiaon (2.5). We can distinguish two cases.

Accordingly let $\dot c$ and $B_{\dot c}$ be as above in a connected and complete semisymmetric hypersurface M^n with $k(p) \leqq 2$.

Proposition 2.2. *If the endomorphism $B_{\dot c}$ degenerates at a point $c(s_0)$ in $V^1_{c(s_0)}$, then $B_{\dot c}^2=0$ holds along the whole $c(s)$ and $B_{\dot c}$ is parallel along $c(s)$.*

Proposition 2.3. *If the endomorphism $B_{\dot c}$ is non-singular at one point $c(s_0)$ in $V^1_{c(s_0)}$, then it is non-singular along $c(s)$ in $V^1_{c(s_0)}$, and at every point $c(s)$ the eigenvalues of $B_{\dot c}$ are non-real complex numbers in V^1_c:*

As a consequence we get, that in a complete semisymmetric hypersurface with $k(p)\leqq 2$ the endomorphisms $B_{\dot c}$ cannot have real non-zero eigenvalues.

In the following proofs the completeness of the manifold is important.

Proof of Proposition 2.2. Let $\mathbf{x}_1(s_0)$ be the unit vector in $V^1_{c(s_0)}$ belonging to the image set of $B_{\dot c(s_0)}$ and let $\mathbf{x}_2(s_0)$ be the orthogonal unit vector in $V^1_{c(s_0)}$. Let us extend these vectors into parallel vector fields $\mathbf{x}_1(s)$, $\mathbf{x}_2(s)$ along $c(s)$. Then these are tangent to $V^1_{c(s)}$.

The restriction of $B_{\dot c(s_0)}$ onto $V^1_{c(s_0)}$ has the matrix in $\{\mathbf{x}_1(s_0), \mathbf{x}_2(s_0)\}$ of the form

$$\begin{bmatrix}\lambda(s_0), & \gamma(s_0)\\ 0, & 0\end{bmatrix}, \tag{2.6}$$

where $\lambda(s_0)=0$ holds iff $B^2_{\dot c(s_0)}=0$ is satisfied. The solutions of (2.5) are uniquely determined by the initial value (2.6), so if $\lambda(s_0)=0$ holds, then the solution of (2.5) has the matrix of the form

$$\begin{bmatrix}0, & \gamma(s)=\gamma(s_0)\\ 0, & 0\end{bmatrix}, \tag{2.7}$$

w.r.t. the basis $\{\mathbf{x}_1(s), \mathbf{x}_2(s)\}$ in $V^1_{c(s)}$, since (2.7) is a solution of (2.5) with the above initial conditions.

Now if $\lambda(s_0)\neq 0$ holds, then the solution of (2.5) has the matrix of the form

$$\begin{bmatrix}\dfrac{1}{s+c_1}, & \gamma(s_0)e^{-\int_{s_0}^{s} dt/(t+c_1)}\\ 0, & 0\end{bmatrix}, \tag{2.8}$$

w.r.t. $\{\mathbf{x}_1(s), \mathbf{x}_2(s)\}$ in $V^1_{c(s)}$, where $c_1=(1-s_0\lambda(s_0))/\lambda(s_0)$ is constant. But in this case the functions $\lambda(s)$, $\gamma(s)$, $K(s)$ have infinity value at $-c_1$ which contradicts the completeness of the manifold. Thus this case doesn't occur and $\lambda(s_0)=0$ holds, which proves the proposition.

Proof of Proposition 2.3. Let $\{\mathbf{x}_1(s_0), \mathbf{x}_2(s_0)\}$ be an orthonormed basis in $V^1_{c(s_0)}$ such that the vectors $\mathbf{x}_i(s_0)$ are the eigenvectors of the symmetric part of $B_{\dot c(s_0)}$. The matrix of $B_{\dot c(s_0)}$ restricted onto $V^1_{c(s_0)}$ is of the form

$$\begin{bmatrix}\alpha_1(s_0), & -\beta(s_0)\\ \beta(s_0), & \alpha_2(s_0)\end{bmatrix}, \tag{2.9}$$

w.r.t. this basis. Let $\{\mathbf{x}_1(s), \mathbf{x}_2(s)\}$ be the extension of $\{\mathbf{x}_1(s_0), \mathbf{x}_2(s_0)\}$ onto $c(s)$ by parallel displacement. If we consider $B_{\dot c(s)}$ only in $V^1_{c(s)}$, then from (2.5) we get the following:

$$B_{\dot c}^{-1}\nabla_{\dot c}B_{\dot c}=-B_{\dot c}, \quad \nabla_{\dot c}B_{\dot c}^{-1}=I.$$

Thus the matrix of the solution $B_{\dot c}$ of (2.5) with initial condition (2.9) is

$$\begin{bmatrix} \dfrac{s+c_1}{(s+c_1)(s+c_2)+c_3^2}, & \dfrac{-c_3}{(s+c_1)(s+c_2)+c_3^2} \\ \dfrac{c_3}{(s+c_1)(s+c_2)+c_3^2}, & \dfrac{s+c_2}{(s+c_1)(s+c_2)+c_3^2} \end{bmatrix}, \tag{2.10}$$

w.r.t. $\{\mathbf{x}_1(s), \mathbf{x}_2(s)\}$, where

$$c_1 = ((\alpha_1(0)/\det B_{\dot c(s_0)}) - s_0, \quad c_2 = (\alpha_2(0)/\det B_{\dot c(s_0)}) - s_0, \quad c_3 = (\beta(0)/\det B_{\dot c(s_0)}) - s_0.$$

Because of the completeness of the hypersurfaces the equation

$$(s+c_1)(s+c_2)+c_3^2 = 0$$

of second order can't have real solution, i.e. for it's discriminant Δ

$$\Delta = (c_1-c_2)^2-4c_3^2 < 0$$

holds. It is easy to see from (2.10) that by this conditon the eigenvalues of the restricted $B_{\dot c(s)}$ are non-real along $c(s)$ which proves the proposition.

After these propositions we examine the orthogonal projection of vector fields $\nabla_X Y$ onto V^0, where X and Y are tangent to V^1. We denote this projected vector field by $v(\nabla_X Y)$.

Proposition 2.4. *Let M^n be a connected complete semisymmetric hypersurface with $k(p)\leqq 2$. Then the vectors $v(\nabla_X Y)$ span an at most 1-dimensional subspace S_p in V_p^0 for every point p.*

Proof. We start with the indirect assumption $\dim S_p \geqq 2$ for a point p. By the assumption the V-decomposition is of the form $T_q(M)=V_q^0+V_q^1$ around p, where $\dim V_q^0=n-2$. Let $\{\mathbf{x}_1, \mathbf{x}_2\}$ be an orthonormed differentiable basic field around p in V^1. Let us denote the vector $v(\nabla_{\mathbf{x}_i}\mathbf{x}_j)_{/p}$ by $\mathbf{x}_{ij}$. Then for arbitrary unit vector $\mathbf{m}$, tangent to V_p^0, the matrix of $B_{\mathbf{m}}$ w.r.t. $(\mathbf{x}_1, \mathbf{x}_2)$ is the following:

$$\begin{bmatrix} -g(\mathbf{x}_{11}, \mathbf{m}), & -g(\mathbf{x}_{21}, \mathbf{m}) \\ -g(\mathbf{x}_{12}, \mathbf{m}), & -g(\mathbf{x}_{22}, \mathbf{m}) \end{bmatrix}.$$

The characteristic equation of this matrix is

$$\lambda^2+\{g(\mathbf{x}_{11}, \mathbf{m})+g(\mathbf{x}_{22}, \mathbf{m})\}\lambda+\{g(\mathbf{x}_{11}, \mathbf{m})g(\mathbf{x}_{22}, \mathbf{m})-g(\mathbf{x}_{12}, \mathbf{m})g(\mathbf{x}_{21}, \mathbf{m})\} = 0,$$

which has the discriminant

$$\Delta = \{g(\mathbf{x}_{11}, \mathbf{m})-g(\mathbf{x}_{22}, \mathbf{m})\}^2+4g(\mathbf{x}_{12}, \mathbf{m})g(\mathbf{x}_{21}, \mathbf{m}).$$

If $\mathbf{x}_{11}\neq 0$ or $\mathbf{x}_{22}\neq 0$ holds and $\mathbf{m}$ is orthogonal to $\mathbf{x}_{12}$ or to $\mathbf{x}_{21}$, then the eigenvalues are $-g(\mathbf{x}_{11}, \mathbf{m})$, $-g(\mathbf{x}_{22}, \mathbf{m})$. And if $\mathbf{x}_{11}=\mathbf{x}_{22}=0$ holds, furthermore $\mathbf{m}$

halves the angle of $\mathbf{x}_{12}$ and $\mathbf{x}_{21}$, then the eigenvalues are $\pm\sqrt{g(\mathbf{x}_{12},\mathbf{m})g(\mathbf{x}_{21},\mathbf{m})}\neq\neq 0$, and these eigenvalues are also reals. Consequently we can choose such a vector $\mathbf{m}$ for which $B_{\mathbf{m}}$ has real, non-zero eigenvalue. This contradicts the previous proposition and the proof is complete.

Let p be a point for which $\dim S_p=1$ holds. Then $\dim S_p=1$ holds in a neighbourhood of p. Let M^2 be such a 2-dimensional submanifold through p in the points of which

$$T_q(M^n) = T_q(M^2)+V_q^0, \quad \dim S_q = 1$$

hold. Let us choose such a system $\mathbf{m}_1, \dots, \mathbf{m}_{n-2}$ around p for which the vectors $\mathbf{m}_1(q)$, $q\in M^2$, are pointing in the direction of S_q. Then in the points $q\in M^2$

$$B_1(q)\neq 0, \quad B_2(q)=\dots=B_{n-2}(q)=0$$

holds. Since the differential equation (1.10) is of first order, so

$$B_1\neq 0, \quad B_2=\dots=B_{n-2}=0$$

hold everywhere, and $\mathbf{m}_1$ is pointing in the direction of S.

A system $\mathbf{m}_1, \dots, \mathbf{m}_{n-2}$ constructed in this way is called a *reduced system.* For such a system only the first tensor B_1 is non-trivial, which we denote by B. Also the basic formulas (1.8) and (1.9) are more simple w.r.t. such a system, and we get for them:

(2.12) $$(\tilde\nabla_X B)(Y)-(\tilde\nabla_Y B)(X) = 0,$$

(2.13) $$M_\alpha^1(X)B(Y)-M_\alpha^1(Y)B(X) = 0,$$

(2.14) $$dM_\alpha^\beta-\sum_\gamma M_\alpha^\gamma\wedge M_\gamma^\beta = 0.$$

The other basic formulas are unchanged.

At the end we give some definitions.

Let M^n be a connected complete immersed hypersurface in $\mathbf{R}^{n+1}$ with $k(p)\leqq 2$ everywhere. Let $\mathcal{V}_1$ be the open set, where $k(p)=2$, i.e. $K(p)\neq 0$ holds for the Riemannian curvature scalar K. Then in the interior $\mathcal{V}_0$ of $M^n\setminus\mathcal{V}_1$ the Riemann curvature $R(X, Y)Z$ vanishes. Let $\mathcal{V}_2\subseteqq\mathcal{V}_1$ be the open set where the subspace S_p (defined in Proposition 2.4) is 1-dimensional. Then the tensor B vanishes in the interior $\mathcal{V}_t$ of $\mathcal{V}_1\setminus\mathcal{V}_2$. The open set $\mathcal{V}_t$ is called the pure trivial part of M^n. At the end let $\mathcal{V}_h\subseteqq\mathcal{V}_2$ be the open set where B has two non-real eigenvalues. Then in the interior $\mathcal{V}_p$ of $\mathcal{V}_2\setminus\mathcal{V}_h$ B doesn't vanish and it has only zero eigenvalues on $\mathcal{V}_p$. The open sets $\mathcal{V}_p$ resp. $\mathcal{V}_h$ are called the pure parabolic resp. pure hyperbolic part of M^n.

It is rather trivial that the open set

(2.15) $$\mathcal{V}_0\cup\mathcal{V}_t\cup\mathcal{V}_p\cup\mathcal{V}_h$$

is everywhere dense in M^n. Furthermore the open sets $\mathcal{V}_t, \mathcal{V}_p$ resp. $\mathcal{V}_h$ always contain the complete integral manifolds of V^0, i.e. the type of the hypersurface is uniquely determined along a maximal integralmanifold of V^0, where $\dim V_q^0=n-2$ holds.

Now let M^n be a general (not necessarily complete) immersed hypersurface, with $k(p)\leqq 2$ everywhere. The V-decomposition is defined for it in the same way as in § 1. This decomposition is of the form

$$T_p(M^n) = V_p^0+V_p^1, \quad \dim V_p^0 = n-2,$$

iff the Riemannian curvature scalar $K(p)$ doesn't vanish. The maximal integral manifold of V^0 through such a point p is always an open set in an euclidean subspace $\mathbf{R}^{n-2}$ of $\mathbf{R}^n$. The M^n is called *vertically complete* iff all these integral manifolds are complete euclidean suspaces $\mathbf{R}^{n-2}$ in $\mathbf{R}^n$.

We can define the open sets $\mathcal{V}_0, \mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_t, \mathcal{V}_p, \mathcal{V}_h$ for vertically complete hypersurfaces with $k(p)\leqq 2$ in some way as before, since propositions (2.2), (2.3) and (2.4) hold for such hypersurfaces also. The type of hypersurfaces along an integral manifold of V^0 (where $\dim V_q^0=n-2$) is also uniquely determined.

Definition. A vertically complete immersed hypersurface M^n with $k(p)\leqq 2$ is said to be of

1) trivial type if $\mathcal{V}_2=\emptyset$ holds, i.e. M^n contains only $\mathcal{V}_0$ resp. pure trivial parts,

2) parabolic type if $\mathcal{V}_t=\mathcal{V}_h=\emptyset$, $\mathcal{V}_p\neq\emptyset$, hold, i.e. M^n contains only $\mathcal{V}_0$ and nonempty pure parabolic part,

3) hyperbolic type if $M^n=\mathcal{V}_h$, i.e. M^n contains only pure hyperbolic part.

By formula (2.15) all complete hypersurfaces with $k(p)\leqq 2$ can be built up from vertically complete hypersurfaces of the above types. In the next sections we give general procedures for the construction of vertically complete immersed hypersurfaces of the above types.

3. Hypersurfaces of trivial type

Strong theorems are known — local or global — which describe all the hypersurfaces with zero Riemannian curvature. For example a complete connected hypersurface M^n with zero Riemannian curvature is a cylinder of the form $M^n=c\times\mathbf{R}^{n-1}$ where c is a curve in an euclidean plane $\mathbf{R}^2$ and $\mathbf{R}^{n-1}$ is the orthogonal complement of $\mathbf{R}^2$ [1]. So by the description of hypersurfaces of trivial type we assume that the open set $\mathcal{V}_t$ is nonempty.

Proposition 3.1. *Let U be a connected component of $\mathcal{V}_t$ in a hypersurface of trivial type. Then U is a cylinder of the form $U=M^2\times\mathbf{R}^{n-2}$, where M^2 is a hypersurface in a euclidean subspace $\mathbf{R}^3$ and $\mathbf{R}^{n-2}$ is the orthogonal complement to $\mathbf{R}^3$.*

Proof. The tensor fields B_α are zero in the considered case, so $\nabla_{V^1}V^1 \subseteqq V^1$ holds. So the distribution V^1 is integrable and the integral manifolds are totally geodesic. Let M^2 be an integral manifold of V^1. From $B_\alpha=0$ and $A(V^0)=0$,

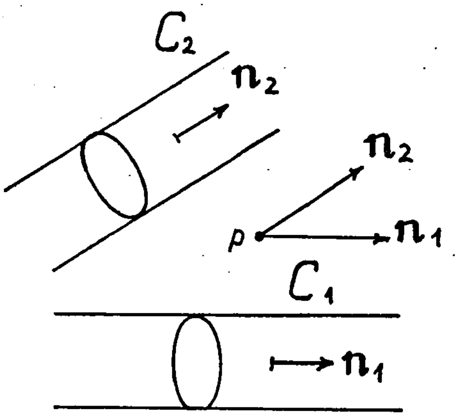
$$D_{V^1}V^0 \subseteqq V^0$$

follows, where D is the covariant derivative of $\mathbf{R}^{n+1}$. Thus the integral manifolds of V^0 are parallel euclidean subspaces, and M^2 is contained in the orthogonal complement $\mathbf{R}^3$ of these parallel subspaces. It is rather trivial, that U is of the form $U=M^2\times\mathbf{R}^{n-2}$ indeed.

The following theorem is obvious.

Theorem 3.1. *For a hypersurface of trivial type there exists an everywhere dense open subset, on the connected component of which the space is of zero Riemannian curvature or it is a cylinder described in the above proposition.*

Generally a hypersurface of trivial type doesn't split into a global direct product of the form $M^2\times\mathbf{R}^{n-2}$. To show this fact we construct a 3-dimensional irreducible hypersurface of trivial type.



Let C_1 and C_2 be two infinite closed circle-cylindrical domains without common points in $\mathbf{R}^3$, which are pointing in different directions $\mathbf{n}_1$ resp. $\mathbf{n}_2$. Furthermore let $f(x,y,z)$ be such a differentiable real function on $\mathbf{R}^3$ which has zero value on $\mathbf{R}^3\setminus(C_1\cup C_2)$ and f is positive inside of C_i, $i=1,2$, such that it is constant along the lines parallel to $\mathbf{n}_i$. Such functions obviously exist.

Proposition 3.2. *The hypersurface M^3 represented by $(x,y,z,f(x,y,z))$ in $\mathbf{R}^4$ is a complete irreducible hypersurface of trivial type, diffeomorphic to $\mathbf{R}^3$.*

Proof. The open sets $\mathcal{V}^2\subset M^3$, $i=1;2$, represented by $(x,y,z,f(x,y,z))$, $(x,y,z)\in C_i$, are cylindrical of the form $\mathcal{V}^i=M_i^2\times\mathbf{R}$, furthermore the Riemannian curvature vanishes on $M^3\setminus(\mathcal{V}^1\cup\mathcal{V}^2)$. Thus M^3 is of trivial type.

Let p be arbitrary point of $\mathbf{R}^3\setminus(C_1\cup C_2)$. Then p is a point of M^3. It is easy to show, that the holonomy group H_p of M^3 is generated by the rotation groups $SO(2)_1$, $SO(2)_2$, where $SO(2)_i$, $i=1; 2$, acts around the axis through p pointing in the direction of $\mathbf{n}_i$. Thus $H_p\cong SO(3)$ holds, and M^3 is irreducible. The other statement in the proposition is obvious.

Since the above example is not locally symmetric, so it is also a counterexample to Nomizu's conjecture.

With the above method one can construct n-dimensional complete irreducible hypersurfaces of trivial type for any dimension n.

4. Hypersurfaces of parabolic type

Let us consider the hypersurface M^n on the open set $\mathscr{V}_p$, where $R\neq 0$, $B\neq 0$ with $B^2=0$. The system $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_{n-2}$ is by assumption a reduced system. Let $\{\partial_0, \partial_1\}$ be an orthonormed basis in V^1 such that ∂_1 is tangent to the image space of B.

By $\nabla_{\mathbf{m}_\alpha}B=0$ we get that ∂_0 and ∂_1 are parallel vector fields along any integral manifold of V^0, i.e. $\nabla_{\mathbf{m}_\alpha}\partial_i=0$ holds. Furthermore from $B^2=0$ we have that the matrix of the restricted B (onto V^1) is of the form

$$\begin{bmatrix} 0, & 0 \\ b, & 0 \end{bmatrix}, \tag{4.1}$$

w.r.t. $\{\partial_0, \partial_1\}$.

Let us introduce also the functions λ, λ_1 by

$$\tilde{\nabla}_{\partial_0}\partial_0=\lambda\partial_1,\quad \tilde{\nabla}_{\partial_1}\partial_1=\lambda_1\partial_0,\quad \tilde{\nabla}_{\partial_0}\partial_1=-\lambda\partial_0,\quad \tilde{\nabla}_{\partial_1}\partial_0=-\lambda_1\partial_1. \tag{4.2}$$

Proposition 4.1. *The above functions satisfy the following equations:*

$$\lambda_1=0,\quad \partial_1(b)=\lambda b, \tag{4.3}$$

$$\nabla_{\partial_1}\partial_0=\nabla_{\mathbf{m}_\alpha}\partial_i=0. \tag{4.4}$$

Proof. From (2.12), (4.1) and (4.2) we have

$$(\tilde{\nabla}_{\partial_1}B)(\partial_0)=\partial_1(b)\partial_1+b\lambda_1\partial_0=(\tilde{\nabla}_{\partial_0}B)(\partial_1)=B(\lambda\partial_0)=\lambda b\partial_1$$

so we get (4.3). (4.4) is obvious by $\lambda_1=0$ and by the above considerations.

Now let us examine the Weingarten field A of the hypersurface. As for it $A(V^0)=0$, $A(V^1)=V^1$ hold, so let $\tilde{A}$ be the restriction of A onto V^1. The matrix of $\tilde{A}$ w.r.t. $\{\partial_0, \partial_1\}$ is of the form

$$\begin{bmatrix} \gamma_0, & \delta \\ \delta, & \gamma_1 \end{bmatrix}. \tag{4.5}$$

Proposition 4.2. *The Weingarten field A satisfies the following relations:*

(4.6) $$\nabla_{\mathbf{m}_1} A = -A \circ B, \quad \nabla_{\mathbf{m}_\alpha} A = 0 \quad \textit{for} \quad \alpha \geqq 2,$$

(4.7) $$A \circ B \quad \textit{is symmetric}, \quad (\tilde{\nabla}_{\partial_0} A)(\partial_1) = (\tilde{\nabla}_{\partial_1} A)(\partial_0),$$

(4.8) $$\gamma_1 = 0, \quad \mathbf{m}_1(\gamma_0) + \delta b = 0, \quad \mathbf{m}_\alpha(\gamma_0) = 0 \quad \text{for} \quad \alpha \geqq 2,$$

(4.9) $$\mathbf{m}_\alpha(\delta) = 0 \quad \textit{if} \quad \alpha \geqq 1,$$

thus δ is constant along the integral manifolds of V^0,

(4.10) $$\partial_1(\gamma_0) = \partial_0(\delta) + \lambda\gamma_0, \quad \partial_1(\delta) = 2\lambda\delta.$$

Proof. Equations (4.6) and (4.7) come from the Codazzi—Mainardi equation

$$(\nabla_X A)(Y) = (\nabla_Y A)(X),$$

using the vector fields $\partial_0, \partial_1, \mathbf{m}_1, \dots, \mathbf{m}_{n-2}$. The equation $\gamma_1 = 0$ comes from symmetry of $A \circ B$, and the others are equivalent to (4.6) and (4.7) using the formulas (4.1)—(4.5).

By (4.8) and (4.5) the curvature scalar K of M^n is

(4.11) $$K = \det \tilde{A} = -\delta^2 < 0$$

on $\mathcal{V}_p$, so the matrix of $\tilde{A}$ in $\{\partial_0, \partial_1\}$ is of the form

(4.12) $$\tilde{A}^i_j = \begin{bmatrix} \gamma_0, & \sqrt{-K} \\ \sqrt{-K}, & 0 \end{bmatrix}.$$

By the second equation of (4.10) also the equation

(4.13) $$\partial_1(K) = 4\lambda K$$

holds.

Let us notice too, that the sectional curvature K_σ is non-positive in a hypersurface of parabolic type so from the Hadamard—Cartan theorem we get:

Proposition 4.3. *The sectional curvature K_σ of a hypersurface M^n of parabolic type is non-positive. Thus if M^n is complete and simply connected then it is diffeomorphic to $\mathbf{R}^n$.*

Proposition 4.4. *The distribution W^0, spanned by ∂_1 and V^0, is involutive, and the integral manifolds of W^0 are open sets in $(n-1)$-dimensional euclidean subspaces of $\mathbf{R}^{n+1}$. In addition if the hypersurface is complete, then the maximal integral manifolds of W^0 are complete $(n-1)$-dimensional euclidean subspaces in $\mathbf{R}^{n+1}$.*

Proof. For the Lie derivative $[\partial_1, \mathbf{m}_\alpha]$ resp. $[\mathbf{m}_\alpha, \mathbf{m}_\beta]$ we have

$$[\partial_1, \mathbf{m}_\alpha] = \nabla_{\partial_1}\mathbf{m}_\alpha - \nabla_{\mathbf{m}_\alpha}\partial_1 = \nabla_{\partial_1}\mathbf{m}_\alpha = B_\alpha(\partial_1) + \sum_\gamma M^\gamma_\alpha(\partial_1)\mathbf{m}_\gamma = \sum_\gamma M^\gamma_\alpha(\partial_1)\mathbf{m}_\gamma,$$

$$[\mathbf{m}_\alpha, \mathbf{m}_\beta] = 0,$$

thus W^0 is involutive.

Let H be an integral manifold of W^0. Then H is a hypersurface in M^n with normal vector field ∂_0. H is by (4.4) a totally goedesic hypersurface in M^n with zero Riemannian curvature as well.

Let D be the covariant derivative in $\mathbf{R}^{n+1}$. By (4.1) and (4.5) we have

$$D_{\partial_1}\mathbf{n} = \delta\partial_0, \quad D_{\partial_1}\partial_0 = -\delta\mathbf{n}, \quad D_{\mathbf{m}_\alpha}\mathbf{n} = 0, \quad D_{\mathbf{m}_\alpha}\partial_0 = 0.$$

Thus the planes spanned by $\mathbf{n}$ and ∂_0 (along H) are parallel, and so H is an open set in the euclidean subspace which is orthogonal to the above parallel planes.

Now let M^n be a complete hypersurface of parabolic type and let H be a maximal integral manifold of W^0. From the second equation of (4.3) and from (4.13) we get, that K resp. B vanishes neither on H nor on the boundary of H. Thus H is without boundary points and so it is a complete $(n-1)$-dimensional euclidean subspace in $\mathbf{R}^{n+1}$.

By the above proposition every connected component $\mathscr{V}_p^i$ of $\mathscr{V}_p$ in a complete M^n can be considered as a fibred space $\Pi\colon \mathscr{V}_p^i\to\mathbf{R}$, where the fibres $\Pi^{-1}(q)$, $q\in\mathbf{R}$, are $(n-1)$-dimensional euclidean spaces. In the following proposition we make this fibration into a global fibration.

Theorem 4.1. *Let M^n be a simply connected and complete immersed hypersurface of parabolic type in $\mathbf{R}^{n+1}$. Then M^n is in a natural manner a fibred space $\Pi\colon M^n\to\mathbf{R}$, where the fibres $\Pi^{-1}(q)$, $q\in\mathbf{R}$, are $(n-1)$-dimensional euclidean subspaces in $\mathbf{R}^{n+1}$.*

Proof. Let us examine M^n on the open set $\mathscr{V}_0$. The rank of the Weingarten field A on $\mathscr{V}_0$ is 1 or 0. Let $\mathscr{V}_0^1\subseteq\mathscr{V}_0$ be the open set, where $\operatorname{rank} A=1$ holds, and let $\mathscr{V}_0^0$ be the interior of $\mathscr{V}_0\setminus\mathscr{V}_0^1$. If ∂_0 is the unit vector field on $\mathscr{V}_0^1$, tangent to the image-space of A, then

$$A(\partial_0) = \gamma_0\partial_0 \quad\text{with}\quad \gamma_0\neq 0$$

holds. Let $\overset{*}{W}{}_q^0\subset T_q(\mathscr{V}_0^1)$, $q\in\mathscr{V}_0^1$, be the subspace orthogonal to $\partial_0(q)$. It is well known that the distribution $\overset{*}{W}{}^0$ is involutive and the integral manifolds of it are open sets in the $(n-1)$-dimensional euclidean subspaces of $\mathbf{R}^{n+1}$. In the following we prove the completeness of these integral manifolds.

First of all let us notice, that the fibration described in Proposition 4.4. can be extended continuously onto the boundary of $\mathscr{V}_p$. In fact, in the opposite case two sequences $p_i, q_i\in\mathscr{V}_p$ could be chosen such that $p=\lim p_i=\lim q_i=q$ is on the boundary of $\mathscr{V}_p$, the integral manifolds H_{p_i} resp. H_{q_i} of W^0 through p_i resp. q_i converge to H_p resp. H_q, but $H_p\neq H_q$ holds. As the spaces $H_{p_i}, H_{q_i}, H_p, H_q$ are hypersurfaces in M^n thus

$$\dim(H_{p_i}\cap H_{q_i}) = n-2.$$

would hold for large numbers i, which is a contradiction. Thus the proof of the statement is complete.

Let us return to the investigation of $\overset{*}{W}^0$'s integral manifolds. Let H be a maximal integral mainfold. For a vector field X tangent to H we have

$$(\nabla_{\partial_0} A)(X) = (\nabla_X A)(\partial_0),$$

from which we get

$$\nabla_X \partial_0 = 0, \quad X(\gamma_0) = \gamma_0 g(X, \nabla_{\partial_0} \partial_0). \tag{4.14}$$

So if $x(t)$ denotes an integral curve of X, then along it

$$\gamma_0(t) = \gamma_0(0) e^{\int_0^t g(\dot{x}, \nabla_{\partial_0} \partial_0)}$$

holds. From this we have, that A vanishes neither on H nor on the boundary of H. So every boundary point of H is a boundary point of $\mathscr{V}_p$, too. We prove, that such a boundary point doesn't exist for H.

We start with the indirect assumption. If q would be such a boundary point, then let H_q be the subspace through q which we get by the extension of the fibration, described in Proposition 4.4, onto the boundary of $\mathscr{V}_p$. Then $\dim(H \cap H_q) = n-2$ holds obviously. Let $\bar{\partial}_0$ be the normal vector of H_q in $T_q(M^n)$. Since $K(q)=0$, $A(q) \neq 0$ hold, so by (4.12) we get, that $\bar{\partial}_0$ is the unique non-trivial eigenvector of $A(q)$. But by (4.14) the non-trivial eigenvector ∂_0 is parallel along H, so the vector $\partial_0(q)$ is also a non-trivial eigenvector of $A(q)$. This is contradiction, because $\partial_0(q) \neq \neq \bar{\partial}_0$ holds.

So we get, that the maximal integral manifolds of $\overset{*}{W}^0$ are also complete $(n-1)$-dimensional euclidean subspaces in $\mathbf{R}^{n+1}$. Now let us consider a connected component $\mathscr{V}_0^{0i}$ of $\mathscr{V}_0^0$. From the above considerations it follows, that $\mathscr{V}_0^{0i}$ is an open set in an n-dimensional euclidean hyperspace, such that the boundary of $\mathscr{V}_0^{0i}$ is either an $(n-1)$-dimensional euclidean subspace, or two parallel $(n-1)$-dimensional subspaces. Thus the extension of the fibration onto $\mathscr{V}_0^0$ is trivial, which proves the proposition.

The above statements suggest a simple constructional method for hypersurfaces of parabolic type.

Proposition 4.5. *Let $c(s)$ be an immersed curve in $\mathbf{R}^{n+1}$, parametrised by arc-length. Furthermore let $H_{c(s)}$ be a differentiable field of $(n-1)$-dimensional euclidean subspaces along $c(s)$ such that $H_{c(s)}$ is orthogonal to $\dot{c}(s)$. Then the subspaces $H_{c(s)}$ cover an immersed hypersurface with $k(p) \leqq 2$ around $c(s)$.*

Proof. It is trivial, that the subspaces $H_{c(s)}$ cover an immersed hypersurface M^n in a neighbourhood of $c(s)$. Let $\mathbf{n}$ be the normal vector field of this hypersurface

M^n, and let ∂_0 be the unit vector field in M^n, orthogonal to the subspaces $H_{c(s)}$. Since the vector $D_X\mathbf{n}$, where X is tangent to $H_{c(s)}$, is pointing always in the direction of ∂_0, so the image-space of Weingarten map A is spanned by the vectors ∂_0 and $D_{\partial_0}\mathbf{n}$. Thus rank $A \leqq 2$ holds, and the proof is finished.

The spaces constructed in the previous proposition are in general not complete. But in many cases a field $H_{c(s)}$ described above covers globally a complete immersed hypersurface M^n. This is the case, if we consider an arbitrary differentiable field $H_{c(s)}$ of orthogonal $(n-1)$-dimensional euclidean subspaces along a line $c(s)$ of $\mathbf{R}^{n+1}$. Of course there can be given more complicated cases. Since such a hypersurface is in general not of the form

$$c \times H_c,$$

where c is a plane curve in a euclidean subplane $\mathbf{R}^2$ and H_c is orthogonal to $\mathbf{R}^2$, so these hypersurfaces have non-zero curvature in general.

Theorem 4.2. *Let $c(s)$, $-\infty < s < \infty$, be an immersed curve in $\mathbf{R}^{n+1}$ and let $H_{c(s)}$ be such a differentiable field of orthogonal $(to\ \dot{c}(s))$, $(n-1)$-dimensional euclidean subspaces along $c(s)$, which cover a complete hypersurface M^n. Then for M^n we have $k(p) \leqq 2$, $B^2=0$ and*

$$K = -(D_{\partial_0}\mathbf{n}, D_{\partial_0}\mathbf{n}) + (D_{\partial_0}\mathbf{n}, \partial_0)^2 \leqq 0. \tag{4.15}$$

Furthermore if $K(p)<0$ holds in a point $p \in H_p$, then $K<0$ is satisfied along H_p.

Proof. By Proposition 4.5 $k(p) \leqq 2$ holds for M^n, and if $K(p) \neq 0$ (i.e. $k(p)=2$) is satisfied, then the image space of the Weingarten field A_p is spanned by ∂_0 and $D_{\partial_0}\mathbf{n}$, where $D_{\partial_0}\mathbf{n}$ has non-zero projection onto the fibre H_p. Let ∂_1 be the unit vector pointing in the direction of this projected vector. Then the non-trivial subspace of A_p is spanned by ∂_0 and ∂_1. Since for $D_{\partial_1}\mathbf{n}$ the relation $D_{\partial_1}\mathbf{n} = \delta\, \partial_0 = A(\partial_1)$ holds, so the matrix of A_p w.r.t. $\{\partial_0, \partial_1\}$ is of the form

$$\begin{bmatrix} \gamma_0, & \delta \\ \delta, & 0 \end{bmatrix}$$

with $\delta \neq 0$. Since $D_{\partial_0}\mathbf{n} = \gamma_0 \partial_0 + \delta \partial_1$ holds, so by $K=-\delta^2$ we get the relation (4.15). Of course (4.15) holds also in the case $K(p)=0$, as in this case $D_{\partial_0}\mathbf{n}$ is pointing in the direction of ∂_0.

The subspaces $H_{c(s)}$ are totally geodesic so $\nabla_{\partial_1}\partial_0 = 0$ follows. From this we get $g(B(\partial_1), \partial_0)=0$ i.e. ∂_1 is an eigenvector of B. But the space is complete so B has only zero real eigenvalue. Thus $B(\partial_1)=0$ and $B^2=0$ follows.

The integral manifolds of V^0 are parallel hyperspaces in the fibres $H_{c(s)}$, and so the integral curves of ∂_1 are lines in $H_{c(s)}$. From (2.4) and (4.13) we get, that $K<0$ holds along $H_{c(s)}$ if in a point $p \in H_{c(s)}$, $K(p)<0$ is satisfied.

We are going to investigate the irreducibility of the previously described spaces. Let M^n be a complete simple connected immersed hypersurface as in Theorem 4.2 with $K<0$, and let $c(s)$, $-\infty<s<\infty$, be an arbitrary fixed integral curve of ∂_0. The subspaces $H_{c(s)}$ can be described uniquely by the normal vector field $\mathbf{n}(s)$ along $c(s)$.

Theorem 4.3*). *The hypersurface M^n with $K<0$ is reducible iff a euclidean subspace $\mathbf{R}^k$ with $k<n+1$ exists, which contains $c(s)$ with the vector field $\mathbf{n}(s)$ as well. If $\mathbf{R}^k$ is the smallest such subspace, then M^n is of the form*

$$M^n = M^{k-1}\times\mathbf{R}^{n-k+1}, \tag{4.16}$$

where M^{k-1} is an irreducible complete hypersurface in $\mathbf{R}^k$ covered by a one-parametrized family $H^*_{c(s)}$ of $(k-1)$-dimensional euclidean subspaces, furthermore $\mathbf{R}^{n-k+1}$ is euclidean subspace in $\mathbf{R}^{n+1}$ orthogonal to $\mathbf{R}^k$.

Proof. If $c(s)$ with $\mathbf{n}(s)$ is contained in a subspace $\mathbf{R}^k$, $k<n+1$, then M^n is obviously of the form (4.16). Thus we examine the other direction, and let us assume that M^n is reducible, and it is of the form

$$M^n = Q^{k-1}\times Q^{n-k+1} \tag{4.17}$$

with $k<n$.

First we prove that (4.17) is a cylindrical decomposition. Let T^1 resp. T^2 be the tangent space of Q^{k-1} resp. Q^{n-k+1}. Since for the curvature tensor R the equation $R(T^1, T^2)X=0$ holds, so by the Gauss equation we get

$$g(X, A(T^1))A(T^2) = g(X, A(T^2))A(T^1), \tag{4.18}$$

for every tangent vector $X\in T(M)$. We show, that A vanishes on one of the tangent spaces T^i.

In fact, if there were tangent vectors $X^i\in T^i_p$, $i=1;2$ for which $A(X^i)\neq 0$ holded, then by (4.18) the vectors $A(X^i)$ would point in the same direction, and so A would be of rank 1. But this is imposible, because $K<0$ holds.

So we get, that one of the spaces Q^{k-1}, Q^{n-k+1} has negative scalar curvature, and the other is of zero curvature. Let Q^{k-1} be the space with $K<0$. Since $A(T^2)=0$ holds, so $T^2\subseteq V^0$ and the integral manifolds of T^2 are complete $(n-k+1)$-dimensional euclidean subspaces. Because of the decomposition (4.17) these euclidean subspaces must be parallel subspaces in $\mathbf{R}^{n+1}$. So (4.17) is a cylindrical decomposition of the form

$$M^n = Q^{k-1}\times\mathbf{R}^{n-k+1},$$

where Q^{k-1} is a hypersurface in $\mathbf{R}^k$ orthogonal to $\mathbf{R}^{n-k+1}$. Since $\mathbf{R}^{n-k+1}$ is orthogonal to $c(s)$ and $\mathbf{n}(s)$ as well, so $c(s)$ and $\mathbf{n}(s)$ are contained in $\mathbf{R}^k$.

---

*) The theorem is true also in case $K\leqq 0$.

The last statement in the theorem is obvious.

We mention, that the above theorem is true also in the case, when we consider M^n only for an open interval $a<s<b$.

By Theorem 4.2 the hypersurfaces described in the theorem can contain also pure trivial part $\mathscr{V}_t$, i.e. on which $K<0, B=0$ hold. It is clear by the above remark, that $\mathscr{V}_t$ is non-empty iff an open interval $a<s<b$ exists, for which $c(s)$ with $\mathbf{n}(s)$ is contained in a 3-dimensional subspace $\mathbf{R}^3$, but a smaller subspace doesn't contain the system $\{c(s), \mathbf{n}(s)\}$. So excluding this possibility the other hypersurfaces described in Theorem 4.2 are of parabolic type.

It is very easy to construct such complete, irreducible hypersurfaces which contain pure parabolic part only.

For example let us consider a differentiable field of unit vectors $\mathbf{n}(s)$ along a line $c(s)$, $-\infty<s<\infty$, in $\mathbf{R}^{n+1}$ for which

1. the vector $D_{\dot{c}}\mathbf{n}$ is non-zero along $c(s)$,
2. the system $\{c(s), \mathbf{n}(s)\}$, $-\infty<s<\infty$, is not contained in a subspace $\mathbf{R}^k$ with $k<(n+1)$.
3. There is no interval $a<s<b$, for which $\{c(s), \mathbf{n}(s)\}$ is in a subspace $\mathbf{R}^3$.

Then the euclidean subspaces $H_{c(s)}$, orthogonal to $c(s)$ and $\mathbf{n}(s)$, inscribe in $\mathbf{R}^{n+1}$ an irreducible complete hypersurface with pure parabolic part only.

It is very easy to contruct also such hypersurfaces which contain only pure trivial and pure parabolic parts.

5. Hypersurfaces of hyperbolic type

Theorem 5.1. *Every connected and simply connected immersed hypersurface M^n of hyperbolic type is of the form $M^n=M^3\times\mathbf{R}^{n-3}$, where M^3 is an immersed hypersurface of hyperbolic type in a euclidean subspace $\mathbf{R}^4$ and $\mathbf{R}^{n-3}$ is euclidean subspace orthogonal to $\mathbf{R}^4$.*

Proof. By (2.13)

$$M_\alpha^1(X)B_1(Y)-M_\alpha^1(Y)B_1(Y)=0$$

holds. Since B_1 is non-degenerate thus $M_\alpha^1=-M_1^\alpha=0$ holds for $\alpha\geqq 2$. This means that $\nabla_X\mathbf{m}_1$ is contained in V_p^1 for every vector $X\in V_p^1$. By formulas (1.3) and Proposition 2.4 the distribution V_p^*, spanned by V_p^1 and $\mathbf{m}_1/p$, is involutive and the integral manifolds of this distribution are totally geodesic. It is also trivial, that the orthogonal complement V_p^{**} of V_p^* is also involutive, and the maximal integral manifolds of it are $(n-3)$-dimensional euclidean subspaces in $\mathbf{R}^{n+1}$. Let M^3 be a maximal integral manifold of V^*. Then for every vector field Y tangent to V^{**} and for every vector field X tangent to V^* the vector field D_XY is also tangent to V^{**}, where D is the covariant derivative in $\mathbf{R}^{n+1}$. This means, that the integral manifolds of V^{**} are parallel euclidean subspaces in $\mathbf{R}^{n+1}$ and that M^3 is an immersed hypersurface of hyperbolic type in an orthogonal complement $\mathbf{R}^4$ of the above parallel euclidean spaces. From the basic formulas it is rather trivial, that the metric of M^n is of the form $M^n = M^3 \times \mathbf{R}^{n-3}$ indeed.

From the above theorem we can see, that for the construction of hyperbolic hypersurfaces we must construct only the 3-dimensional cases. In the following we describe a general construction for such hypersurfaces.

At first let us consider a one-fold covering of a simply connected open set U of $\mathbf{R}^3$ with complete lines such that the unit vector field $\mathbf{u}$ tangent to these lines is differentiable. We call such a covering a *line-fibration of* U. For a point $p \in U$ let $\overset{*}{V}{}_p^1$ be the orthogonal complement of $\mathbf{u}_p$ and let $\overset{*}{V}{}_p^0$ be the 1-dimensional subspace in $T_p(U)$ spanned by $\mathbf{u}_p$. The following relations are obvious for the covariant derivative D of $\mathbf{R}^3$:

$$D_{\overset{*}{V}{}^0}\overset{*}{V}{}^1 \subseteqq \overset{*}{V}{}^1, \quad D_{\overset{*}{V}{}^0}\overset{*}{V}{}^0 \subseteqq \overset{*}{V}{}^0, \quad D_{\overset{*}{V}{}^1}\overset{*}{V}{}^1 \subseteqq \overset{*}{V}{}^0 + \overset{*}{V}{}^1. \tag{5.1}$$

Furthermore let $\overset{*}{B}(X) := D_X\mathbf{u}$ be the derived tensor field of $\mathbf{u}$ and let $\overset{*}{\nabla}$ be the covariant derivative defined by

$$\begin{gathered}\overset{*}{\nabla}_X Y := D_XY - (D_XY, \mathbf{u})\mathbf{u} = D_XY + (\overset{*}{B}(X), Y)\mathbf{u}, \quad X_p;\ Y_p \in \overset{*}{V}{}_p^1, \\ \overset{*}{\nabla}_X\mathbf{u} := 0 \quad \text{for every vector field } X, \quad \text{and} \quad \overset{*}{\nabla}_{\mathbf{u}}X := D_{\mathbf{u}}X \ \text{ if } \ X_p \in \overset{*}{V}{}_p^1\end{gathered} \tag{5.2}$$

on U, where (X, Y) denotes the inner product in $\mathbf{R}^3$. It is rather trivial that $\overset{*}{\nabla}$ is metrical w.r.t. (X, Y). If $\overset{*}{R}$ denotes the curvature tensor of $\overset{*}{\nabla}$, then the following basic formulas hold for the given line fibration:

$$\begin{gathered}\overset{*}{R}(X, Y)Z = (\overset{*}{B}(Y), Z)\overset{*}{B}(X) - (\overset{*}{B}(X), Z)\overset{*}{B}(Y), \\ (\overset{*}{\nabla}_X\overset{*}{B})(Y) - (\overset{*}{\nabla}_Y\overset{*}{B})(X) = 0 \quad \text{if} \quad X_p;\ Y_p \in \overset{*}{V}{}_p^1, \\ \nabla_{\mathbf{u}}B^* = -B^* \circ B^*, \\ \overset{*}{R}(X, \mathbf{u})Y = \overset{*}{R}(X, Y)\mathbf{u} = 0.\end{gathered} \tag{5.3}$$

These formulas can be proved in a similar way as the formulas of Proposition 1.1. Since the lines in the fibration are complete lines so it can be proved (similarly to Proposition 2.2 and 2.3) that along a line either $B^{*2} = 0$ holds or B^* is non-degenerated on $\overset{*}{V}{}^1$ and it has two non-real eigenvalues.

Now let $U_1 \subseteqq U$ be the maximal open set where $B^{*2}=0$ holds and let $U_2 \subseteqq U$ be the open set where B^* is non-degenerated in $\overset{*}{V}^1$. Then the open set $U_1 \cup U_2$ is everywhere dense in U, and both open sets are line-fibred open sets. Thus for the line fibrations we can give the following local classification. One class of such fibrations contains the fibrations for which $B^{*2}=0$ holds, and the other class contains the fibrations for which B^* is non-degenerated in $\overset{*}{V}^1$. We describe this classification form a more geometric point of view.

First let us consider the case $B^{*2}=0$. If $\overset{*}{B}=0$ holds on an open set, then this open set is fibred with parallel lines. And if $B^* \neq 0$ holds, then let $\overset{*}{\partial}_0$, $\overset{*}{\partial}_1$ be the orthogonal unit vector fields tangent to $\overset{*}{V}^1$, such that $\overset{*}{\partial}_1$ is tangent to the kernel of $\overset{*}{B}$. The following statement can be proved in the same way as Proposition 4.2.

Proposition 5.1. *The distribution $\overset{*}{W}^0$ spanned by* **u** *and $\overset{*}{\partial}_1$ is involutive.*

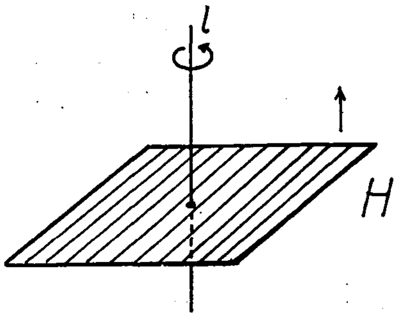
A maximal integral manifold $\overset{*}{H}$ of $\overset{*}{W}^0$ is an open set in a euclidean hyperplane of $\mathbf{R}^3$ such that the lines of fibration, which have common point with $\overset{*}{H}$, are parallel lines in this hyperplane and the integral curves of $\overset{*}{\partial}_1$ in $\overset{*}{H}$ are parallel line segments in the plane.

Conversely, if through every line l of a line-fibration there exists a euclidean hyperplane H such that H covers parallel lines from the fibration around l then the equation $\overset{*}{B}^2=0$ holds for the line-fibration.

The last statement of the above proposition is also obvious.

Thus the above local classification of line-fibrations is the following. One class contains the line-fibrations which can be covered with one parametric family of hyperplanes in the sense of Proposition 5.1 and the elements of other class cannot be covered in such a way. So we call the elements of the first class *plane-coverable line-fibrations* and the elements of the second class *plane-uncoverable line-fibrations.*

It is easy to give plane-coverable line-fibrations. For example let us consider a family of parallel lines in a hyperplane H of $\mathbf{R}^3$.



Let us move H along a line l (perpendicular to H) in such a way that H also turns around l. In this way we get a plane-coverable line-fibration of the whole $\mathbf{R}^3$. In order to show the existence of fibrations belonging to the second class we also give an example of a plane-uncoverable line-fibration of whole $\mathbf{R}^3$.

Let us consider the unit vector field

(5.4)
$$\mathbf{u} = (z^2+1)^{-1/2}(x^2+y^2+z^2+1)^{-1/2}\left\{(xz-y)\frac{\partial}{\partial x}+(yz+x)\frac{\partial}{\partial y}+(z^2+1)\frac{\partial}{\partial z}\right\}$$

defined in a Cartesian coordinate neighbourhood (x, y, z) of $\mathbf{R}^3$. A simple computation shows the equation $D_{\mathbf{u}}\mathbf{u}=0$, thus the maximal integral curves of u are lines and these lines define a line-fibration of $\mathbf{R}^3$. Every line intersects the (x, y)-plane $(z=0)$ just in one point. It can be simply computed that the eigenvalues of $\overset{*}{B}(X)=D_X\mathbf{u}$ at the point of the (x, y)-plane are

$$0, \quad (x^2+y^2+1)^{-1/2}\mathbf{i}, \quad -(x^2+y^2+1)^{-1/2}\mathbf{i}, \tag{5.5}$$

where $\mathbf{i}$ is the imaginary number. Thus $\overset{*}{B}$ has two non-real eigenvalues at every point of $\mathbf{R}^3$ and the fibration is a plane-uncoverable line-fibration.

Now let us consider a 3-dimensional hypersurface M^3 of hyperbolic type in $\mathbf{R}^4$. The integral curves of the vector field $\mathbf{m}$ in M^3 are lines in $\mathbf{R}^4$ and the tangent hyperspaces $T_p(M^3)$ coincide along such an integral curve l. Let us denote this constant hyperspace by $T_l(M^3)$. If S is such a euclidean hyperspace in $\mathbf{R}^4$, which is not orthogonal to l, then the orthogonal projection $\Pi\colon M^3 \to S$ maps an open neighbourhood U of l diffeomorphically onto an open set U^* of S such that the image of $\mathbf{m}$'-s integral curves form a line-fibration of U^*. This line-fibration is called the projected line-fibration of U^*.

Proposition 5.2. *The projected line-fibration of U^* is plane-uncoverable if M^3 is of hyperbolic type.*

Proof. Let α be the angle between the line l and the projected line l'. Then α can be cosidered as a differentiable function on U^* which is constant along the projected lines l'. If $\lambda_i(p)$, $p\in U^*$, $i=1, 2, 3$ denotes the eigenvalues of $B(X)=\nabla_X\mathbf{m}$ at the point $\Pi^{-1}(p)\in U$ then by a simple computation we get, that the eigenvalues of $\overset{*}{B}(X)=D_X\mathbf{u}$ are $\lambda_i^*=\cos\alpha\lambda_i$, $i=1, 2, 3$, which proves the proposition.

By the above considerations every hypersurface M^3 of hyperbolic type can be represented locally as the position of the points

$$(x, y, z, f(x, y, z)), \tag{5.6}$$

where $f(x, y, z)$ is a differentiable function on an open set $U^*\subseteqq\mathbf{R}^3$, where U^* is an open set, line-fibred in a plane-uncoverable way.

We mention, that the unit normal vector field $\mathbf{n}$ of M^3 is represented by

$$\mathbf{n} = \frac{1}{h}(-f_x, -f_y, -f_z, 1), \tag{5.7}$$

where $h=(1+f_x^2+f_y^2+f_z^2)^{1/2}$, furthermore the second fundamental form is represented by

$$H = \frac{1}{h}\begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix}. \tag{5.8}$$

(For details see [6].) Let $\mathbf{u}$ be the unit vector field referring to the line-fibration of U^*. Then the covariant vector field df is parallel along the integral curves of $\mathbf{u}$, i.e. $D_{\mathbf{u}}df=0$ holds, furthermore rank $H=2$ holds at every point $p\in U^*$, and the nullspace of H is spanned by $\mathbf{u}$.

Now we turn to the reversed problem, and we give a general construction for hyper-surfaces M^3 of hyperbolic type.

Theorem 5.2. *Let $U^*\subseteqq \mathbf{R}^3$ be an open set which is line-fibred in a plane-uncoverable way. Then around every line of the fibration there exist differentiable functions $f(x, y, z)$ such that the points*

$$\big(x, y, z, f(x, y, z)\big)$$

represent hypersurfaces of hyperbolic type.

Proof. Let $\mathbf{u}$ be the vector field referring to the fibration of U^*.

Lemma 5.2.1. *The hypersurface $\big(x, y, z, f(x, y, z)\big)$ is of hyperbolic type referring to the fibration of U^* iff*

$$D_{\mathbf{u}}df = 0, \quad \operatorname{rank} D^2f = 2 \tag{5.9}$$

hold.

The proof is obvious by Proposition 5.2 and formula (5.8).

Let $M^2\subset U^*$ be such a hypersurface in $\mathbf{R}^3$ for which the tangent spaces $T_p(M^2)$ are complements of $\mathbf{u}_p$, i.e. $T_p(M^2)+S_p=T_p(\mathbf{R}^3)$ holds, where S_p is the 1-dimensional subspace spanned by $\mathbf{u}_p$. Thus M^2 can be considered as a cross-section of U^*'s fibration. If (x^1, x^2) is a coordinate neighbourhood of M^2, then it can be extended uniquely onto a coordinate neighbourhood (x^1, x^2, t) of U^* such that $\partial/\partial t=\mathbf{u}$ holds, and $(x^1, x^2, 0)$ is just (x^1, x^2) on M^2. The vector fields $\partial/\partial x^i$ can be written in the form

$$\frac{\partial}{\partial x^i} = E_i+\Phi_i\mathbf{u}, \tag{5.10}$$

where E_i is orthogonal to $\mathbf{u}$ and thus also

$$\Phi_i = \left(\frac{\partial}{\partial x^i}, \mathbf{u}\right) \tag{5.11}$$

holds. For the tensor field $\overset{*}{B}$ the following holds:

$$\overset{*}{B}\left(\frac{\partial}{\partial x^i}\right) = \overset{*}{B}(E_i+\Phi_i\mathbf{u}) = \overset{*}{B}(E_i) = \overset{*}{B}{}^r_i E_r = \overset{*}{B}{}^r_i \frac{\partial}{\partial x^r} - \overset{*}{B}{}^r_i \Phi_r \mathbf{u}. \tag{5.12}$$

Lemma 5.2.2. *The fields E_i, Φ_i, $\overset{*}{B}{}^r_i$ fulfill the following formulas:*

$$\begin{gathered}\frac{\partial \Phi_i}{\partial t} = 0, \quad D_{\mathbf{u}} E_i = \overset{*}{B}(E_i) = \overset{*}{B}{}^r_i E_r, \quad \frac{\partial \overset{*}{B}{}^j_i}{\partial t} = -\overset{*}{B}{}^j_r \overset{*}{B}{}^r_i, \\ (\overset{*}{B}(E_j), E_i)-(\overset{*}{B}(E_i), E_j) = E_j(\Phi_i)-E_i(\Phi_j) = \partial\Phi_i/\partial x^j - \partial\Phi_j/\partial x^i.\end{gathered} \tag{5.13}$$

Proof. From $[\partial/\partial x^i, \mathbf{u}]=[\partial/\partial x^i, \partial/\partial t]=0$ we get

$$0 = \left[\frac{\partial}{\partial x^i}, \mathbf{u}\right] = [E_i+\Phi_i\mathbf{u}, \mathbf{u}] = [E_i, \mathbf{u}] - \frac{\partial \Phi_i}{\partial t}\mathbf{u}.$$

On the other hand

$$[E_i, \mathbf{u}] = D_{E_i} u - D_{\mathbf{u}} E_i = \overset{*}{B}(E_i) - D_{\mathbf{u}} E_i.$$

Since both components of these equations are orthogonal to $\mathbf{u}$, so we get the first two equations in (5.13). We get the third equation form $D_{\mathbf{u}}\overset{*}{B} = -\overset{*}{B}{}^2$ and from the second equation. We get the last equation in the following way:

$$0 = \left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] = [E_i+\Phi_i\mathbf{u}, E_j+\Phi_j\mathbf{u}] =$$

$$= [E_i, E_j]^{\sim} + \{(\overset{*}{B}(E_j), E_i) - (\overset{*}{B}(E_i), E_j) - E_j(\Phi_i) + E_i(\Phi_j)\}\mathbf{u},$$

thus the last equation is also satisfied.

Every solution f of $D_{\mathbf{u}} df=0$ satisfies $\mathbf{u}\cdot\mathbf{u}(f)=0$, thus f must be of the form $f=\varrho(x^1, x^2)t+\lambda(x^1, x^2)$ in the above coordinate neighbourhood (x^1, x^2, t), where the functions ϱ, λ are the functions of the variables (x^1, x^2) only.

Lemma 5.2.3. *A function $f=\varrho(x^1, x^2)t+\lambda(x^1, x^2)$ is the solution of $D_{\mathbf{u}} df=0$ iff for ϱ and λ the differential equation*

$$\frac{\partial \varrho}{\partial x^i} - \overset{*}{B}{}^r_i\left(\frac{\partial \varrho}{\partial x^r}t + \frac{\partial \lambda}{\partial x^r} - \Phi_r \varrho\right) = 0, \quad i = 1;\ 2, \tag{5.14}$$

holds.

Proof. This equation comes from (5.13) by

$$(D_{\mathbf{u}}df)\left(\frac{\partial}{\partial x^i}\right) = \frac{\partial^2 f}{\partial t\partial x^i} - df\left(D_{\mathbf{u}}\frac{\partial}{\partial x^i}\right) = \frac{\partial^2 f}{\partial t\partial x^i} - df(D_{\mathbf{u}}E_i + \Phi_i\mathbf{u}) =$$

$$= \frac{\partial^2 f}{\partial t\partial x^i} - df(\overset{*}{B}(E_i)) = \frac{\partial \varrho}{\partial x^i} - \overset{*}{B}{}^r_i\left(\frac{\partial \varrho}{\partial x^r}t + \frac{\partial \lambda}{\partial x^r} - \varrho\Phi_r\right).$$

For every solution f the restrictions of ϱ and λ onto M^2 satisfy the differential equation

$$\frac{\partial \varrho}{\partial x^i} - \overset{*}{B}{}^r_i\left(\frac{\partial \lambda}{\partial x^r} - \varrho\Phi_r\right) = 0. \tag{5.15}$$

Lemma 5.2.4. *Let $\varrho(x^1, x^2)$ and $\lambda(x^1, x^2)$ be the solutions of* (5.15) *on M^2. Then the function $f=\varrho t+\lambda$ defined on (x^1, x^2, t) is a solution of $D_{\mathbf{u}}df=0$.*

Proof. Let $\omega_i(t)$ be the functions defined by the left side of (5.14) along a line of the fibration. Since $\overset{*}{B}{}^r_i$ is of the form (2.10) along a line thus $\omega_i(t)$ are analytical functions with $\omega_i(t)=0$. A simple computation shows the equation

$$\frac{d^n\omega_i}{dt^n} = (-1)^n \overset{*}{B}{}^{l_1}_i \overset{*}{B}{}^{l_2}_{l_1} \dots \overset{*}{B}{}^{l_n}_{l_{n-1}}\omega_{l_n},$$

so $d^n\omega_i/dt^n_{/0}=0$, i.e. $\omega_i=0$ everywhere. This proves the statement.

Now let us assume that M^2 is a hyperplane in $\mathbf{R}^3$ and that (x^1, x^2) is a Descartesian coordinate system on it.

Lemma 5.2.5. *The covariant vector field $p_i=\overset{*}{B}{}^r_i\Phi_r$ is a closed form on a hyperplane M^2.*

Proof. It can be seen from (5.3) that the equation

$$(D_X\overset{*}{B})(Y) = (D_Y\overset{*}{B})(X) \tag{5.16}$$

holds for every vector field X, Y in $\mathbf{R}^3$. By this formula we get

$$0 = D_{\partial/\partial x^i}\overset{*}{B}\left(\frac{\partial}{\partial x^j}\right) - D_{\partial/\partial x^j}\overset{*}{B}\left(\frac{\partial}{\partial x^i}\right) =$$

$$= D_{\partial/\partial x^i}\left(\overset{*}{B}{}^r_j\frac{\partial}{\partial x^r} - \overset{*}{B}{}^r_j\Phi_r\mathbf{u}\right) - D_{\partial/\partial x^j}\left(\overset{*}{B}{}^r_i\frac{\partial}{\partial x^r} - B^r_i\Phi_r\mathbf{u}\right) =$$

$$= \left\{\frac{\partial \overset{*}{B}{}^r_j}{\partial x^i} - \frac{\partial \overset{*}{B}{}^r_i}{\partial x^j} - \overset{*}{B}{}^q_j\Phi_q\overset{*}{B}{}^r_i + \overset{*}{B}{}^q_i\Phi_q\overset{*}{B}{}^r_j\right\}\frac{\partial}{\partial x^r} + \left\{\frac{\partial \overset{*}{B}{}^r_i\Phi_r}{\partial x^j} - \frac{\partial \overset{*}{B}{}^r_j\Phi_r}{\partial x^i}\right\}\mathbf{u},$$

and so

$$(5.17)\qquad \frac{\partial \overset{*}{B}{}^r_j}{\partial x^i}-\frac{\partial \overset{*}{B}{}^r_i}{\partial x^j}=\overset{*}{B}{}^q_j\Phi_q\overset{*}{B}{}^r_i-\overset{*}{B}{}^q_i\Phi_q\overset{*}{B}{}^r_j,\qquad \frac{\partial \overset{*}{B}{}^r_i\Phi_r}{\partial x^j}-\frac{\partial \overset{*}{B}{}^r_j\Phi_r}{\partial x^i}=0.$$

By the last formula the proof is complete.

Let us define the matrix field

$$(5.18)\qquad a^{ij}:=\begin{bmatrix}-\overset{*}{B}{}^1_2, & (1/2)(\overset{*}{B}{}^1_1-\overset{*}{B}{}^2_2)\\ (1/2)(\overset{*}{B}{}^1_1-\overset{*}{B}{}^2_2), & \overset{*}{B}{}^2_1\end{bmatrix}$$

on M^2. This matrix field is positive definite as by the plane-uncoverable fibration

$$(5.19)\qquad \det(a^{ij})=-\overset{*}{B}{}^2_1\overset{*}{B}{}^1_2-(1/4)(\overset{*}{B}{}^1_1-\overset{*}{B}{}^2_2)^2>0$$

holds, since the discriminant $\Delta(=-\det(a^{ij}))$ of the characteristic equation

$$\lambda^2-\operatorname{Tr}\overset{*}{B}\lambda+\det\overset{*}{B}=0$$

is negative.

Lemma 5.2.6. *In a hyperplane M^2 the differential equation* (5.15) *is equivalent to the equations*

$$(5.20)\qquad a^{ij}\frac{\partial^2\lambda}{\partial x^i\partial x^j}=0,\quad \det(a^{ij})>0,$$

$$(5.21)\qquad \frac{\partial\varrho}{\partial x^i}+\overset{*}{B}{}^r_i\Phi_r\varrho=\overset{*}{B}{}^r_i\frac{\partial\lambda}{\partial x^r}.$$

Furthermore for a fixed solution λ of (5.20) *the differential equation* (5.21) *is completely integrable w.r.t. ϱ.*

Proof. We can write the equation (5.15) also in the following invariant form

$$(5.22)\qquad d\varrho+\varrho\delta-\omega=0,$$

where δ resp. ω are the covariant vector fields $\overset{*}{B}{}^r_i\Phi_r$ resp. $\overset{*}{B}{}^r_i\partial\lambda/\partial x^r$. As the operator d acts on the left side of this equation so we get by Lemma 5.2.5:

$$(5.23)\qquad d\omega=d\varrho\wedge\delta=\omega\wedge\delta.$$

We show, that this equation is equivalent to (5.20). Indeed, the equation (5.23) is just the following:

$$\frac{\partial \overset{*}{B}{}^r_i\lambda_r}{\partial x^j}-\frac{\partial \overset{*}{B}{}^r_j\lambda_r}{\partial x^i}=\overset{*}{B}{}^r_i\Phi_r\overset{*}{B}{}^p_j\lambda_p-\overset{*}{B}{}^r_j\Phi_r\overset{*}{B}{}^p_i\lambda_p,$$

where $\lambda_r:=\lambda/\partial x^r$. By the first equation of (5.23) we get

$$\overset{*}{B}{}^r_i \frac{\partial^2\lambda}{\partial x^r \partial x^j} - \overset{*}{B}{}^r_j \frac{\partial^2\lambda}{\partial x^r \partial x^i} = 0, \tag{5.24}$$

which is equivalent to (5.20) indeed. Since (5.23) is the condition of integrability for (5.21) thus the last statement is in the lemma also obvious.

Now let l be a line from the line-fibration of U^*. For a point $p\in l$ let M^2 be a hyperplane such that l is not belonging to M^2. Then there exists a neighbourhood V of p in M^2 such that the lines going through points of V are not belonging to M^2. Let (x^1, x^2) be a Descartesian coordinate neighbourhood on M^2 and let λ be a non-linear solution of (5.20) around p. Then λ is non-linear in a neighbourhood V^* of p, i.e. the matrix field $\partial^2\lambda/\partial x^i\partial x^j$ is non-trivial on V^*. Let ϱ be a solution of (5.21) w.r.t. the fixed λ. Then ϱ is uniquely determined by the initial value $\varrho(p)$. By the above considerations the function $f(x^1, x^2, t)=\varrho(x^1, x^2)t+\lambda(x^1, x^2)$ satisfies the differential equation $D_u df=0$. On the other hand the rank of D^2f is 2 in a neighbourhood of l. To prove this statement we only have to show that the matrix field $\partial^2\lambda/\partial x^i\partial x^j$ is non-singular on V^*. Indeed, by (5.24) the field $\partial^2\lambda/\lambda x^i\partial x^j$ cannot be of rank 1, on V^*, because in the opposite case the null-space would be an eigen direction of $\overset{*}{B}{}^i_j$ by (5.24). This is impossible, because the two eigenvalues of $\overset{*}{B}{}^i_j$ are non-real. So for a neighbourhood of l the points $(x, y, z, f(x, y, z))$ represent a hypersurface of hyperbolic type and the proof of Theorem is complete.

Now we turn to Takagi's counterexample. Let us consider the line-fibration (5.4). Then every line of the fibration intersects the (x, y)-plane only in one point. Let us denote this canonical coordinate neighbourhood on this plane by (x^1, x^2). A simple computation shows, that the matrix field $\overset{*}{B}{}^j_i$ is of the form

$$\overset{*}{B}{}^j_i = ((x^1)^2+(x^2)^2+1)\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

on this plane and so the function $\lambda(x^1, x^2):=-x^1x^2$ satisfies the differential equation (5.20) with $\det(\partial^2\lambda/\partial x^i\partial x^j)=-1$. From (5.21) we get the solution

$$\varrho = (1/2)\big((x^2)^2-(x^1)^2\big)\big((x^1)^2+(x^2)^2+1\big)^{-1/2}.$$

If we compute the function $f(x^1, x^2, t)=\varrho t+\lambda$ in the Descartesian coordinate neighbourhood (x, y, z) of $\mathbf{R}^3$, we have

$$f(x, y, z) = \frac{x^2z-y^2z-2xy}{2(z^2+1)},$$

and so the points $(x, y, z, f(x, y, z))$ represent a complete irreducible hypersurface of hyperbolic type which is of course irreducible and non-symmetric. But this is just Takagi's counterexample, so we have:

Proposition 5.3. *Takagi's counterexample is a complete hypersurface of hyperbolic type.*

Proposition 5.4. *The sectional curvature K_σ is non-positive for every plane σ in a hypersurfaces of hyperbolic type. So every complete and simple connected immersed hypersurface M^n of hyperbolic type is diffeomorphic to $\mathbf{R}^n$.*

Proof. It is enough to prove, that the sectional curvature w.r.t. $\sigma=V_p^1$ is negative. If (A_j^i), $i;j=1;2$, is the Weingarten field, restricted onto $\sigma=V_p^1$, then $K_\sigma=$ $=\det(A_j^i)$ holds. On the other hand $\nabla_m A=-A\circ B$ holds, thus we get

$$B_i^r A_{rj} = B_j^r A_{ri}.$$

If A_{ij} were positive definite, then B would have two non-zero real eigenvalues. So the signature of A_{ij} is 1, and thus $K_\sigma=\det(A_j^i)<0$ holds.

6. Classification of complete semisymmetric hypersurfaces

At the end we can summarize the results of the paper in the following manner.

Theorem 6.1. *Let M^n be a complete semisymmetric immersed hypersurface in $\mathbf{R}^{n+1}$. Then M^n is one of the following types.*

1. *M^n is of zero curvature, and it is of the form $M^n=c\times\mathbf{R}^{n-1}$, where c is a curve in a hyperplane $\mathbf{R}^2$ and $\mathbf{R}^{n-1}$ is orthogonal to $\mathbf{R}^2$.*
2. *M^n is a straight cylinder of the form $M^n=S^k\times\mathbf{R}^{n-k}$ described in Nomizu's theorem.*
3. *M^n is pure trivial of the form $M^n=M^2\times\mathbf{R}^{n-2}$, where M^2 is a hypersurface in a 3-dimensional euclidean subspace $\mathbf{R}^3$ and $\mathbf{R}^{n-2}$ is orthogonal to $\mathbf{R}^3$.*
4. *M^n is pure parabolic of the form $M^n=M^k\times\mathbf{R}^{n-k}$, where M^k is an irreducible pure parabolic hypersurface in a euclidean subspace $\mathbf{R}^{k+1}$ and $\mathbf{R}^{n-k}$ is orthogonal to $\mathbf{R}^{k+1}$.*
5. *M^n is pure hyperbolic of the form $M^n=M^3\times\mathbf{R}^{n-3}$, where M^3 is a pure hyperbolic irreducible hypersurface in a 4-dimensional euclidean subspace $\mathbf{R}^4$ and $\mathbf{R}^{n-3}$ is orthogonal to $\mathbf{R}^4$.*
6. *M^n satisfies the relation $k(p)\leqq 2$ and it is mixed with $\mathcal{V}_0$, $\mathcal{V}_t$, $\mathcal{V}_p$, $\mathcal{V}_h$ parts.*

Theorem 6.2. *A complete semisymmetric immersed hypersurface with $K>0$ is one of the following types.*

1. *M^n is a cylinder $M^n=S^{k-1}\times\mathbf{R}^{n-k}$ described in Nomizu's theorem.*
2. *M^n is pure trivial of the form $M^n=M^2\times\mathbf{R}^{n-2}$ described above in point 3.*

Theorem 6.3. *Let M^n be a complete immersed semisymmetric hypersurface with $|K|\geqq\varepsilon>0$ for a constant ε. Then M^n is also one of the types described in the above theorem.*

Proof. Let M^n have the property $k(p)\leqq 2$. Then M^n can't have hyperbolic part, because on an integral line of $\mathbf{m}_1$ on this part the function $K(s)$ is of the form

$$K(s)=\frac{Q}{(s+c_1)(s+c_2)+c_3^2},\quad Q=\text{constant},$$

by (2.4) and (2.10).

But M^n can't have pure parabolic part either. Indeed, on this part the integral manifolds of W^0 would be complete $(n-1)$-dimensional euclidean subspaces in $\mathbf{R}^{n+1}$ by (2.4), (2.7), (4.3) and (4.13), and the maximal integral curves of ∂_1 would be complete lines in these subspaces.

On the other hand B degenerates on this part, so by (1.7) $R(\partial_1, \partial_0)\partial_0=\tilde{R}(\partial_1, \partial_0)\partial_0$ holds. From this relation we get

$$\partial_1(\lambda)=K+\lambda^2,$$

so along an integral curve of ∂_1

$$\frac{dK}{ds}=4\lambda K,\quad \frac{\partial\lambda}{ds}=K+\lambda^2$$

hold. The general solutions of this system with $K<0$ are the following:

$$K(t)=\frac{Q_1}{(Q_1-(t+Q_2)^2)},\quad \lambda(t)=\frac{t+Q_2}{-(t+Q_2)^2+Q_1},$$

where Q_1 and Q_2 are constants with $Q_1<0$. So this case is also impossible and M^n contains only pure trivial part. By Proposition 3.1 the proof is finished.

References

[1] P. Hartman, L. Nirenberg, On spherical image maps whose Jacobians do not change sign, *Amer. J. Math.*, **81** (1959), 901—920.

[2] S. Kobayashi, K. Nomizu, *Foundations of Differential Geometry*, Vol. II, Interscience Publisher (New York, 1969).

[3] K. Nomizu, On hypersurfaces satisfying a certain condition on the curvature tensor, *Tôhoku Math. J.*, **20** (1968), 46—59.

[4] Z. I. Szabó, Structure theorems on Riemannian manifold satisfying $R(X, Y)\cdot R=0$, I, Local version, *J. Differential Geom.*, **17** (1982), 531—582.

[5] Z. I. Szabó, Structure theorems on Riemannian manifolds satisfying $R(X, Y)\cdot R=0$, II, Global version, to appear.

[6] H. Takagi, An example of Riemannian manifold satisfying $R(X, Y)\cdot R=0$ but not $\nabla R=0$, *Tôhoku Math. J.*, **24** (1972), 105—108.

JÓZSEF ATTILA UNIVERSITY
BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Acta Sci. Math. 47 (1984), 349—369

On the a.e. convergence of multiple orthogonal series. II (Unrestricted convergence of the rectangular partial sums)

F. MÓRICZ and K. TANDORI

1. Preliminaries and notations

Let Z_+^d be the set of all d-tuples $k=(k_1, ..., k_d)$ with positive integral coordinates. In case $d=1$, Z_+^1 is the set of the positive integers, which is well-ordered. For $d\geqq 2$, Z_+^d is only partially ordered by agreeing that for $k=(k_1, ..., k_d)$ and $n==(n_1, ..., n_d)$ we write $k\leqq n$ iff $k_j\leqq n_j$ for each $j(=1, 2, ..., d)$. Further, sometimes we write 1 for the d-tuple $(1, ..., 1)$.

Let $\varphi=\{\varphi_k(x): k\in Z_+^d\}$ be an orthonormal system (in abbreviation: ONS) on the unit interval $I=(0, 1)$. Since we are interested in the questions of almost everywhere (in abbreviation: a.e.) convergence behaviour, in this paper we do not make any distinction among open, half-closed, and closed intervals. Consider the d-multiple orthogonal series

(1) $$\sum_{k\in Z_+^d} a_k\varphi_k(x) = \sum_{k_1=1}^{\infty} \cdots \sum_{k_d=1}^{\infty} a_{k_1, ..., k_d}\varphi_{k_1, ..., k_d}(x),$$

where $\mathfrak{a}=\{a_k: k\in Z_+^d\}$ is a d-multiple sequence of real numbers (coefficients), for which

(2) $$\sum_{k\in Z_+^d} a_k^2 < \infty.$$

By the well-known Riesz—Fischer theorem, there exists a function $f(x)\in L^2(I)$ such that the rectangular partial sums

$$s_n(x) = \sum_{k\leqq n} a_k\varphi_k(x) = \sum_{k_1=1}^{n_1} \cdots \sum_{k_d=1}^{n_d} a_{k_1, ..., k_d}\varphi_{k_1, ..., k_d}(x)$$

Received January 10, 1983.

of series (1) converge to $f(x)$ in L^2-metric:

$$\int_0^1 [s_n(x)-f(x)]^2\,dx \to 0 \quad \text{as} \quad \min_{1\leqq j\leqq d} n_j \to \infty.$$

It is a fundamental fact that condition (2) itself does not ensure the pointwise convergence of $s_n(x)$ to $f(x)$ (see [2] for $d=1$ and [5] for $d\geqq 2$). Our goal is to give a necessary and sufficient condition in order to ensure the a.e. convergence of the rectangular partial sums $s_n(x)$ of series (1) for every ONS φ on I. The case $d=1$ was elaborated by the second author in [6] and [7]. Some of the results for $d\geqq 2$ were announced by the first author in [4].

In this paper we do not suppose any restriction on the ratios n_j/n_i, $1\leqq i,j\leqq d$, that is, we are concerned ourselves with the a.e. unrestricted convergence of the rectangular partial sums $s_n(x)$ of series (1).

Given a d-multiple sequence $\mathfrak{a}=\{a_k\colon k\in Z_+^d\}$, let us introduce the following quantity:

$$\|\mathfrak{a}\| = \sup_\varphi \left\{ \int_0^1 \Big(\sup_{m,n\in Z_+^d : m\leqq n} \Big| \sum_{m\leqq k\leqq n} a_k \varphi_k(x) \Big| \Big)^2 dx \right\}^{1/2},$$

where the first supremum is extended over all ONS φ on I. Here and in the sequel

$$\sum_{m\leqq k\leqq n} a_k\varphi_k(x) = \sum_{k_1=m_1}^{n_1} \cdots \sum_{k_d=m_d}^{n_d} a_{k_1,\dots,k_d}\varphi_{k_1,\dots,k_d}(x).$$

Given an arbitrary subset Q of Z_+^d, we consider another d-multiple sequence $\mathfrak{a}(Q)=\{a_k(Q)\colon k\in Z_+^d\}$ defined as follows

$$a_k(Q) = \begin{cases} a_k & \text{for} \quad k\in Q, \\ 0 & \text{for} \quad k\in Z_+^d\setminus Q. \end{cases}$$

In particular, we write

$$Q_N=\{k\in Z_+^d\colon k_j\leqq N \ \text{for each}\ j\} \quad (N=1,2,\dots).$$

In this case we may write

$$\|\mathfrak{a}(Q_N)\| = \sup_\varphi \Big\{ \int_0^1 \Big(\max_{m,n\in Q_N : m\leqq n} \Big| \sum_{m\leqq k\leqq n} a_k\varphi_k(x) \Big| \Big)^2 dx \Big\}^{1/2}. \tag{3}$$

It is clear that $\|\mathfrak{a}(Q_N)\| \leqq \|\mathfrak{a}\|$ for every $N=1,2,\dots$. On the other hand, by Beppo Levi's theorem, it follows that

$$\lim_{N\to\infty} \|\mathfrak{a}(Q_N)\| = \|\mathfrak{a}\|. \tag{4}$$

Denote by

$$\mathfrak{M} = \{\mathfrak{a}\colon \|\mathfrak{a}\| < \infty\}.$$

It will turn out that $\mathfrak{M}$ is the very class of those d-multiple sequences $\mathfrak{a}=\{a_k: k\in Z_+^d\}$, for which series (1) converges a.e. for every ONS φ on I.

Remark 1. Let us observe that

$$\sum_{m\leqq k\leqq n} a_k\varphi_k(x) = \sum_{\delta_1=0}^{1}\cdots\sum_{\delta_d=0}^{1}(-1)^{\delta_1+\dots+\delta_d}s_{\delta_1(m_1-1)+(1-\delta_1)n_1,\dots,\delta_d(m_d-1)+(1-\delta_d)n_d}(x)$$

with the agreement of taking $s_{k_1,\dots,k_d}(x)=0$ if $k_j=0$ for at least one j. Thus, introducing another quantity:

$$\|\mathfrak{a}\|_* = \sup_{\varphi}\Big\{\int_0^1\big(\sup_{n\in Z_+^d}\big|\sum_{1\leqq k\leqq n} a_k\varphi_k(x)\big|\big)^2\,dx\Big\}^{1/2}$$

for every d-multiple sequence $\mathfrak{a}$ we have

$$\|\mathfrak{a}\|_* \leqq \|\mathfrak{a}\| \leqq 2^d\|\mathfrak{a}\|_*.$$

This means that the corresponding classes $\mathfrak{M}$ and $\mathfrak{M}_*=\{\mathfrak{a}: \|\mathfrak{a}\|_*<\infty\}$ coincide. However, the use of $\|\mathfrak{a}\|$ is more convenient for our purposes.

Remark 2. The definition of $\|\mathfrak{a}\|$ and the theorems below remain valid if the interval I of orthogonality is replaced by any finite, nonatomic, positive measure space $(X, \mathcal{F}, \nu)$, in particular $X=I^d$. In addition, the treatment can be extended, with some simple modifications, to the case when we consider ONS φ of complex-valued functions and d-multiple sequences $\mathfrak{a}$ of complex numbers.

2. Auxiliary results

We begin with

Lemma 1. *For every positive integer N we have*

$$\Big\{\sum_{k\in Q_N} a_k^2\Big\}^{1/2} \leqq \|\mathfrak{a}(Q_N)\| \leqq \sum_{k\in Q_N}|a_k|. \tag{5}$$

Proof. It immediately follows from the following inequalities:

$$\big|\sum_{k\in Q_N} a_k\varphi_k(x)\big| \leqq \max_{m,n\in Q_N: m\leqq n}\big|\sum_{m\leqq k\leqq n} a_k\varphi_k(x)\big| \leqq \sum_{k\in Q_N}|a_k\varphi_k(x)|.$$

Theorem 1. *The mapping $\|\cdot\|$: $\mathfrak{a}(\in\mathfrak{M})\to\|\mathfrak{a}\|$ is a norm, and $\mathfrak{M}$ is a Banach space with respect to the usual vector operations and the norm $\|\cdot\|$.*

Proof. Obviously $\|\mathfrak{a}\|\geqq 0$. By (4) and (5),

$$\Big\{\sum_{k\in Z_+^d} a_k^2\Big\}^{1/2} \leqq \|\mathfrak{a}\| \leqq \sum_{k\in Z_+^d}|a_k|. \tag{5'}$$

Hence it follows that $\|\mathfrak{a}\|=0$ if and only if $a_k=0$ for each $k\in Z_+^d$.

It is also clear that $\|\alpha\mathfrak{a}\|=|\alpha|\|\mathfrak{a}\|$ for every real number α and sequence $\mathfrak{a}$.

Now let two sequences $\mathfrak{a}=\{a_k\colon k\in Z_+^d\}$ and $\mathfrak{b}=\{b_k\colon k\in Z_+^d\}$ be given. Then for every positive integer N

$$\max\Big|\sum_{m\leqq k\leqq n}(a_k+b_k)\,\varphi_k(x)\Big|\leqq\max\Big|\sum_{m\leqq k\leqq n}a_k\varphi_k(x)\Big|+\max\Big|\sum_{m\leqq k\leqq n}b_k\varphi_k(x)\Big|,$$

where all the three maxima are taken under the conditions $m,n\in Q_N$ and $m\leqq n$. Applying the Bunjakovskiĭ—Schwartz inequality and definition (3), we get that

$$\|(\mathfrak{a}+\mathfrak{b})(Q_N)\|\leqq\|\mathfrak{a}(Q_N)\|+\|\mathfrak{b}(Q_N)\|.$$

Hence, via (4),

$$\|\mathfrak{a}+\mathfrak{b}\|\leqq\|\mathfrak{a}\|+\|\mathfrak{b}\|.$$

Thus we have shown that $\mathfrak{M}$ is a linear space. Now we prove the completeness with respect to the norm $\|\cdot\|$. To this effect, let $\mathfrak{a}^{(p)}=\{a_k^{(p)}\colon k\in Z_+^d\}$ $(p=1,2,\ldots)$ be an ordinary sequence of elements from $\mathfrak{M}$ satisfying the Cauchy convergence criterion:

$$\|\mathfrak{a}^{(p)}-\mathfrak{a}^{(q)}\|\to 0\quad\text{as}\quad p,q\to\infty.$$

By (5′),

$$\sum_{k\in Z_+^d}(a_k^{(p)}-a_k^{(q)})^2\to 0\quad\text{as}\quad p,q\to\infty.$$

So there exists an $\mathfrak{a}=\{a_k\colon k\in Z_+^d\}$ such that

$$a_k^{(p)}\to a_k\quad\text{as}\quad p\to\infty\quad\text{for each}\quad k\in Z_+^d.$$

Let an $\varepsilon>0$ be given. By assumption there exists a positive integer $p_0=p_0(\varepsilon)$ such that

$$\|\mathfrak{a}^{(p)}-\mathfrak{a}^{(q)}\|\leqq\varepsilon\quad\text{whenever}\quad p,q\geqq p_0.$$

Given a positive integer N, a fortiori

$$\|\mathfrak{a}^{(p)}(Q_N)-\mathfrak{a}^{(q)}(Q_N)\|\leqq\varepsilon\quad\text{whenever}\quad p,q\geqq p_0.$$

By (5) and the triangle inequality,

$$\|\mathfrak{a}^{(p)}(Q_N)-\mathfrak{a}(Q_N)\|\leqq\|\mathfrak{a}^{(p)}(Q_N)-\mathfrak{a}^{(q)}(Q_N)\|+\|\mathfrak{a}^{(q)}(Q_N)-\mathfrak{a}(Q_N)\|\leqq$$
$$\leqq\varepsilon+\sum_{k\in Q_N}|a_k^{(q)}-a_k|.$$

Letting q tend to infinity, hence

$$\|\mathfrak{a}^{(p)}(Q_N)-\mathfrak{a}(Q_N)\|\leqq\varepsilon\quad\text{whenever}\quad p\geqq p_0.$$

This holds true for each $N=1,2,\ldots$. Thus, by (4)

$$\|\mathfrak{a}^{(p)}-\mathfrak{a}\|\leqq\varepsilon\quad\text{whenever}\quad p\geqq p_0,$$

in particular, $\mathfrak{a}\in\mathfrak{M}$. Being $\varepsilon>0$ arbitrary,

$$\|\mathfrak{a}^{(p)}-\mathfrak{a}\|\to 0\quad\text{as}\quad p\to\infty.$$

Remark 3. By (5′), if $\mathfrak{a}\in\mathfrak{M}$, then condition (2) is necessarily satisfied.

Theorem 2. *If* $\mathfrak{a}=\{a_k\colon k\in Z_+^d\}$ *and* $\mathfrak{b}=\{b_k\colon k\in Z_+^d\}$ *are such that*

$$|a_k| \leqq |b_k| \quad \text{for every} \quad k\in Z_+^d, \tag{6}$$

then $\|\mathfrak{a}\|\leqq\|\mathfrak{b}\|$.

This immediately yields

Corollary 1. *Let* $\mathfrak{a}$ *and* $\mathfrak{b}$ *be such that* (6) *is satisfied. If* $\mathfrak{b}\in\mathfrak{M}$, *then* $\mathfrak{a}\in\mathfrak{M}$; *and consequently, if* $\mathfrak{a}\notin\mathfrak{M}$, *then* $\mathfrak{b}\notin\mathfrak{M}$.

Proof of Theorem 2. By (4), it is enough to prove that for every positive integer N

$$\|\mathfrak{a}(Q_N)\| \leqq \|\mathfrak{b}(Q_N)\|. \tag{7}$$

By (6), if $b_k=0$ for every $k\in Q_N$, then also $a_k=0$ for every $k\in Q_N$. Thus, (7) is trivially satisfied:

$$\|\mathfrak{a}(Q_N)\| = \|\mathfrak{b}(Q_N)\| = 0.$$

Now assume that the set

$$R_N = \{k\in Q_N\colon \quad b_k \neq 0\}$$

is non-empty. If $k\in Q_N\setminus R_N$, then $b_k=0$ and $a_k=0$. For a given $\varepsilon>0$, let us choose an ONS $\{\varphi_k(x)\colon k\in Q_N\}$ in such a way that

$$\|\mathfrak{a}(Q_N)\|^2-\varepsilon \leqq \int_0^1 \Big(\max_{m,n\in Q_N:m\leqq n} \Big|\sum_{m\leqq k\leqq n} a_k\varphi_k(x)\Big|\Big)^2\,dx. \tag{8}$$

We define for $k\in R_N$

$$\bar\varphi_k(x) = \begin{cases} \sqrt{3}\, a_k b_k^{-1}\varphi_k(3x) & \text{for} \quad x\in(0,\,1/3),\\ \sqrt{3}\,(1-a_k^2 b_k^{-2})^{1/2}\varphi_k(3x-1) & \text{for} \quad x\in(1/3,\,2/3),\\ 0 & \text{for} \quad x\in(2/3,\,1);\end{cases}$$

and for $k\in Q_N\setminus R_N$

$$\bar\varphi_k(x) = \begin{cases} 0 & \text{for} \quad x\in(0,\,2/3),\\ \sqrt{3}\,\varphi_k(3x-2) & \text{for} \quad x\in(2/3,\,1).\end{cases}$$

It is easy to check that $\{\bar\varphi_k(x)\colon k\in Q_N\}$ is also an ONS on I. Further, (8) implies that

$$\|\mathfrak{b}(Q_N)\|^2 \geqq \int_0^1 \Big(\max\Big|\sum_{m\leqq k\leqq n} b_k\bar\varphi_k(x)\Big|\Big)^2\,dx \geqq 3\int_0^{1/3} \Big(\max\Big|\sum_{m\leqq k\leqq n} a_k\varphi_k(3x)\Big|\Big)^2\,dx =$$

$$= \int_0^1 \Big(\max\Big|\sum_{m\leqq k\leqq n} a_k\varphi_k(x)\Big|\Big)^2\,dx \geqq \mathfrak{a}(Q_N)\|^2-\varepsilon,$$

where all the three maxima are taken under the conditions $m, n \in Q_N$ and $m \leqq n$. Being $\varepsilon > 0$ arbitrary, hence the wanted inequality (7) follows.

In the sequel we shall need the following

Lemma 2. *Let* $\mathfrak{a}(Q_N) = \{a_k \colon k \in Q_N\}$ *be given, where* N *is a positive integer. Then there exist an ONS* $\psi = \{\psi_k(x) \colon k \in Q_N\}$ *of step functions on* I *and a simple subset* E *of* I *having the following properties:*

$$\text{mes}\, E \geqq C_1 \tag{9}$$

and

$$\max_{m,n \in Q_N : m \leqq n} \Big| \sum_{m \leqq k \leqq n} a_k \psi_k(x) \Big| \geqq \|\mathfrak{a}(Q_N)\| \quad \textit{for every} \quad x \in E, \tag{10}$$

where C_1 *is a positive constant.*

A set E is said to be simple if it is the union of finitely many disjoint intervals and mes E stands for the sum of the lengths of these intervals (i.e. the Lebesgue measure of E). In the following, by $C_2, C_3, \ldots$ we shall denote positive constants, sometimes depending on d.

Proof. If $\|\mathfrak{a}(Q_N)\| = 0$, then statements (9) and (10) are satisfied for $E = (0, 1)$, $C_1 = 1$, and arbitrary ONS ψ of step functions.

From now on we assume that $\|\mathfrak{a}(Q_N)\| > 0$. Without loss of generality, we may also assume that $\|\mathfrak{a}(Q_N)\| = 1$. By definition, there exists on ONS φ on I, for which

$$\int_0^1 \Big(\max_{m,n \in Q_N : m \leqq n} \Big| \sum_{m \leqq k \leqq n} a_k \varphi_k(x) \Big| \Big)^2 dx \geqq \frac{1}{2}. \tag{11}$$

Let $\varepsilon > 0$ be arbitrary, and let $\chi_k(x)$, $k \in Q_N$, be step functions on I such that

$$\int_0^1 [\varphi_k(x) - \chi_k(x)]^2 \, dx \leqq \varepsilon \quad (k \in Q_N).$$

We set

$$\alpha_{k,m} = \int_0^1 \chi_k(x) \chi_m(x) \, dx$$

and

$$\eta_k = \sum_{m \in Q_N : m \neq k} |\alpha_{k,m}| \quad (k, m \in Q_N).$$

It is not hard to see that if $\varepsilon > 0$ is small enough, then we have

$$\int_0^1 \Big(\max_{m,n \in Q_N : m \leqq n} \Big| \sum_{m \leqq k \leqq n} a_k \chi_k(x) \Big| \Big)^2 dx \geqq \frac{1}{4} \tag{12}$$

and

$$\int_0^1 \Big(\max_{m,n\in Q_N:\, m\leqq n} \Big|\sum_{m\leqq k\leqq n} a_k\Big(1-\frac{1}{\sqrt{\alpha_{k,k}+\eta_k}}\Big)\chi_k(x)\Big|\Big)^2 dx \leqq \frac{1}{8}. \tag{13}$$

We shall define an ONS $\{\bar\chi_k(x)\colon k\in Q_N\}$ of step functions on the interval $(0, 2)$ in the following way. We divide the interval $(1, 2)$ into $N^d(N^d-1)$ subintervals $I_{k,m}$ of equal length, where $k, m\in Q_N$ and $k\neq m$. Then, for $k\in Q_N$, we set

$$\bar\chi_k(x) = \begin{cases} \chi_k(x) & \text{for } x\in(0,1), \\ \left\{\dfrac{|\alpha_{k,m}|}{2\operatorname{mes} I_{k,m}}\right\}^{1/2} & \text{for } x\in I_{k,m}, \\ -\left\{\dfrac{|\alpha_{k,m}|}{2\operatorname{mes} I_{k,m}}\right\}^{1/2} \operatorname{sign}\alpha_{k,m} & \text{for } x\in I_{m,k}, \\ 0 & \text{otherwise,} \end{cases}$$

where in the second and third lines m runs over Q_N except k. Taking into account that

$$\int_0^2 \bar\chi_k^2(x)\,dx = \alpha_{k,k}+\eta_k,$$

it is obvious that the step functions

$$\bar\psi_k(x) = \frac{\bar\chi_k(x)}{\sqrt{\alpha_{k,k}+\eta_k}} \quad (k\in Q_N)$$

constitute an ONS on the interval $(0, 2)$. Furthermore, by (12) and (13)

$$\int_0^2 \Big(\max_{m,n\in Q_N:\, m\leqq n} \Big|\sum_{m\leqq k\leqq n} a_k\bar\psi_k(x)\Big|\Big)^2 dx \geqq \frac{1}{8}. \tag{14}$$

Now we set

$$F(x) = \max_{m,n\in Q_N:\, m\leqq n} \Big|\sum_{m\leqq k\leqq n} a_k\bar\psi_k(x)\Big|.$$

Since $F(x)$ is a step function, we can divide the interval $(0, 2)$ into disjoint subintervals $J_1, J_2, \ldots, J_\varrho$ such that it is constant on each J_r; denote by w_r this constant value $(r=1, 2, \ldots, \varrho)$. Then (14) can be rewritten into the following form:

$$S = \sum_{r=1}^{\varrho} w_r^2 \operatorname{mes} J_r \geqq \frac{1}{8}.$$

Taking ε sufficiently small, we may assume that $S\leqq 2$. We set

$$u_0 = 0, \quad u_r = \frac{1}{2}\sum_{s=1}^{r} w_s^2 \operatorname{mes} J_s \quad (r = 1, 2, \ldots, \varrho),$$

and, for $k\in Q_N$,

$$\psi_k(x)=\begin{cases}\dfrac{\sqrt{2}}{w_{r+1}}\bar{\psi}_k\left(\dfrac{2}{w_{r+1}^2}(x-u_r)+\dfrac{1}{2}\sum_{s=1}^{r}\operatorname{mes} J_s\right) & \text{for } x\in(u_r, u_{r+1}),\\ & r=0,1,\dots,\varrho-1, \text{ provided } w_r\neq 0;\\ 0 & \text{otherwise.}\end{cases}$$

It is easy to verify that these functions $\psi_k(x)$, $k\in Q_N$, the simple set $E=\bigcup_{r=0}^{\varrho-1}(u_r, u_{r+1})$ with $C_1=1/8$ satisfy all requirements of Lemma 2.

Theorem 3. *Let* $\mathfrak{a}=\{a_k:\ k\in Z_+^d\}$ *be given. If* Q' *and* $Q''\subseteqq Z_+^d$ *are such that*

$$Q'\cap Q''=\emptyset \quad \text{and} \quad Q'\cup Q''=Z_+^d,$$

then

$$\|\mathfrak{a}(Q')\|^2+\|\mathfrak{a}(Q'')\|^2\leqq\|\mathfrak{a}\|^2.$$

Proof. Given an $\varepsilon>0$, there exist two ONS $\{\varphi_k'(x):\ k\in Z_+^d\}$ and $\{\varphi_k''(x):\ k\in Z_+^d\}$ such that

$$(15)\quad \begin{aligned}&\int_0^1\Big(\sup_{m,n\in Z_+^d:\,m\leqq n}\Big|\sum_{m\leqq k\leqq n:\,k\in Q'}a_k\varphi_k'(x)\Big|\Big)^2dx\geqq\|\mathfrak{a}(Q')\|^2-\varepsilon,\\ &\int_0^1\Big(\sup_{m,n\in Z_+^d:\,m\leqq n}\Big|\sum_{m\leqq k\leqq n:\,k\in Q''}a_k\varphi_k''(x)\Big|\Big)^2dx\geqq\|\mathfrak{a}(Q'')\|-\varepsilon.\end{aligned}$$

We define for $k\in Q'$

$$\varphi_k(x)=\begin{cases}\sqrt{2}\,\varphi_k'(2x) & \text{for } x\in(0,1/2),\\ 0 & \text{for } x\in(1/2,1);\end{cases}$$

and for $k\in Q''$

$$\varphi_k(x)=\begin{cases}0 & \text{for } x\in(0,1/2),\\ \sqrt{2}\,\varphi_k''(2x-1) & \text{for } x\in(1/2,1).\end{cases}$$

It is clear that $\{\varphi_k(x):\ k\in Z_+^d\}$ is an ONS on I. Furthermore, by (15)

$$\|\mathfrak{a}\|^2\geqq\int_0^1\Big(\sup\Big|\sum_{m\leqq k\leqq n}a_k\varphi_k(x)\Big|\Big)^2dx=2\int_0^{1/2}\Big(\sup\Big|\sum_{m\leqq k\leqq n:\,k\in Q'}a_k\varphi_k''(2x)\Big|\Big)^2dx+$$

$$+2\int_{1/2}^1\Big(\sup\Big|\sum_{m\leqq k\leqq n:\,k\in Q''}a_k\varphi_k''(2x-1)\Big|\Big)^2dx=$$

$$=\int_0^1\Big(\sup\Big|\sum_{m\leqq k\leqq n:\,k\in Q'}a_k\varphi_k'(x)\Big|\Big)^2dx+\int_0^1\Big(\sup\Big|\sum_{m\leqq k\leqq n:\,k\in Q''}a_k\varphi_k''(x)\Big|\Big)^2dx\geqq$$

$$\geqq\|\mathfrak{a}(Q')\|^2+\|\mathfrak{a}(Q'')\|^2-2\varepsilon,$$

where all the five suprema are taken over all $m, n \in Z_+^d$ such that $m \leqq n$. Being $\varepsilon > 0$ arbitrary, the proof is complete.

Corollary 2. *If* $\mathfrak{a} \in \mathfrak{M}$, *then*

$$\lim_{N\to\infty} \|\mathfrak{a}(Z_+^d \setminus Q_N)\| = 0.$$

Proof. Given $\varepsilon > 0$, by (4) there exists a positive integer N_0 such that

$$\|\mathfrak{a}(Q_N)\|^2 \geqq \|\mathfrak{a}\|^2 - \varepsilon \quad \text{whenever} \quad N \geqq N_0.$$

On the other hand, in virtue of Theorem 3

$$\|\mathfrak{a}(Q_N)\|^2 + \|\mathfrak{a}(Z_+^d \setminus Q_N)\|^2 \leqq \|\mathfrak{a}\|^2 < \infty.$$

Combining the two estimates above, we find that

$$\|\mathfrak{a}(Z_+^d \setminus Q_N)\|^2 \leqq \varepsilon \quad \text{whenever} \quad N \geqq N_0.$$

Corollary 3. $\mathfrak{M}$ *is separable.*

Proof. On the one hand, by Corollary 2,

$$\|\mathfrak{a} - \mathfrak{a}(Q_N)\| = \|\mathfrak{a}(Z_+^d \setminus Q_N)\| \leqq \varepsilon$$

if N is large enough. On the other hand, we can choose $\mathfrak{b}(Q_N) = \{b_k : k \in Q_N\}$ in such a way that all b_k, $k \in Q_N$, are rational numbers and by (5)

$$\|\mathfrak{a}(Q_N) - \mathfrak{b}(Q_N)\| \leqq \sum_{k \in Q_N} |a_k - b_k| \leqq \varepsilon.$$

Since the class $\bigcup_{N=1}^{\infty} \{\mathfrak{b}(Q_N)\colon$ all b_k are rational numbers for $k \in Q_N\}$ is countable, the proof is complete.

Theorem 4. *If* $\mathfrak{a} \in \mathfrak{M}$, *then there exists a d-multiple sequence* $\lambda = \{\lambda_k : k \in Z_+^d\}$ *of positive numbers such that*

$$\lambda_k \to \infty \quad \text{as} \quad \max_{1 \leqq j \leqq d} k_j \to \infty \quad \text{and} \quad \lambda\mathfrak{a} \in \mathfrak{M}. \tag{16}$$

If $\mathfrak{a} \notin \mathfrak{M}$, *then there exists a d-multiple sequence* $\mu = \{\mu_k : k \in Z_+^d\}$ *of positive numbers such that*

$$\mu_k \to 0 \quad \text{as} \quad \max_{1 \leqq j \leqq d} k_j \to \infty \quad \text{and} \quad \mu\mathfrak{a} \notin \mathfrak{M}. \tag{17}$$

Proof. If $\mathfrak{a} \in \mathfrak{M}$, then by Corollary 2 there exists a sequence $(0=)N_0 < N_1 < \ldots$ $\ldots < N_p < \ldots$ of integers for which

$$\|\mathfrak{a}(Q_{N_p} \setminus Q_{N_{p-1}})\| \leqq p^{-3} \quad (p = 2, 3, \ldots).$$

We set

$$\lambda_k = p \quad \text{for} \quad k \in Q_{N_p} \setminus Q_{N_{p-1}} \quad (p = 1, 2, \ldots).$$

The first assertion in (16) is clearly satisfied. On the other hand, using the triangle inequality and (4),

$$\|\lambda \mathfrak{a}\| = \lim_{q\to\infty} \|\lambda \mathfrak{a}(Q_{N_q})\| \leqq \lim_{q\to\infty} \sum_{p=1}^{q} \|\lambda \mathfrak{a}(Q_{N_p} \setminus Q_{N_{p-1}})\| =$$

$$= \lim_{q\to\infty} \sum_{p=1}^{q} p\|\mathfrak{a}(Q_{N_p} \setminus Q_{N_{p-1}})\| \leqq \mathfrak{a}(Q_{N_1})\| + \sum_{p=1}^{\infty} p^{-2} < \infty.$$

This is the second assertion in (16).

If $\mathfrak{a} \notin \mathfrak{M}$, then by (4), (5) and the triangle inequality there exists a sequence $(0=)N_0 < N_1 < \ldots < N_p < \ldots$ of integers such that

$$\|\mathfrak{a}(Q_{N_p} \setminus Q_{N_{p-1}})\| \geqq p^2 \quad (p = 1, 2, \ldots).$$

Now we set

$$\mu_k = p^{-1} \quad \text{for} \quad k \in Q_{N_p} \setminus Q_{N_{p-1}} \quad (p = 1, 2, \ldots).$$

The fulfilment of the first assertion in (17) is obvious. Applying Theorem 2, we find that

$$\|\mu \mathfrak{a}\| \geqq \|\mu \mathfrak{a}(Q_{N_p} \setminus Q_{N_{p-1}})\| \geqq p \quad (p = 1, 2, \ldots),$$

which implies $\mu \mathfrak{a} \notin \mathfrak{M}$.

3. Two convergence notions for multiple series

Let us consider a d-multiple series

(18) $$\sum_{k\in Z_+^d} u_k = \sum_{k_1=1}^{\infty} \cdots \sum_{k_d=1}^{\infty} u_{k_1,\ldots,k_d}$$

of real numbers, with the rectangular partial sums

$$s_n = \sum_{k\leqq n} u_k = \sum_{k_1=1}^{n_1} \cdots \sum_{k_d=1}^{n_d} u_{k_1,\ldots,k_d} \quad (n \in Z_+^d).$$

More generally, given a rectangle R in Z_+^d with edges of finite length and parallel to the coordinate axis, i.e. $R = \{k \in Z_+^d : m \leqq k \leqq n\}$, set

$$s(R) = \sum_{k\in R} u_k = \sum_{m\leqq k\leqq n} u_k =$$

$$= \sum_{k_1=m_1}^{n_1} \cdots \sum_{k_d=m_d}^{n_d} u_{k_1,\ldots,k_d} \quad (m, n \in Z_+^d;\ m \leqq n).$$

It is clear that $s(R)=s_n$ in the special case $m=1$. On the other hand, it will be useful to notice that

$$(19)\qquad s(R) = \sum_{\delta_1=0}^{1} \cdots \sum_{\delta_d=0}^{1} (-1)^{\delta_1+\dots+\delta_d} s_{\delta_1(m_1-1)+(1-\delta_1)n_1, \dots, \delta_d(m_d-1)+(1-\delta_d)n_d}$$

with the agreement $s_{k_1,\dots,k_d}=0$ if $k_j=0$ for at least one j.

We remind that series (18) is said to be convergent in *Pringsheim's sense* if there exists a finite number s with the following property: for every $\varepsilon>0$ there exists a number $N=N(\varepsilon)$ so that

$$|s_n-s| < \varepsilon \quad \text{whenever} \quad \min_{1\leqq j\leqq d} n_j \geqq N.$$

The number s is said to be the sum of (18). It is well-known that a necessary and sufficient condition that series (18) converge in Pringsheim's sense is that for every $\varepsilon>0$ there exist a number $M=M(\varepsilon)$ so that

$$|s_m-s_n| < \varepsilon \quad \text{whenever} \quad \min_{1\leqq j\leqq d} m_j \geqq M \quad \text{and} \quad \min_{1\leqq j\leqq d} n_j \geqq M$$

(the Cauchy convergence principle).

It is also known from the literature that series (18) is said to be *regularly convergent* if for every $\varepsilon>0$ there exists a number $N=N(\varepsilon)$ so that for every rectangle $R=\{k\in Z_+^d\colon m\leqq k\leqq n\}$

$$|s(R)| < \varepsilon \quad \text{whenever} \quad \max_{1\leqq j\leqq d} m_j > N \quad \text{and} \quad n \geqq m,$$

i.e. $m\in Z_+^d \setminus Q_N$ and $n\geqq m$.

It is an exercise to show that convergence in Pringsheim's sense follows from regular convergence, but the converse statement is not true.

The notion of regular convergence is due to HARDY [1]. Much later this kind of convergence was rediscovered by the first author and called in [3] convergence in a restricted sense. (As to a relatively complete history of the question, we refer to [4], where some of the results of the present paper were already stated.)

4. The main results

One of our main results is that the class $\mathfrak{M}$ introduced in Section 1 contains exactly those d-multiple sequences $\mathfrak{a}=\{a_k\colon k\in Z_+^d\}$ of coefficients for which the orthogonal series (1) regularly converges a.e. for every ONS φ on I.

Theorem 5. *If* $\mathfrak{a}\in\mathfrak{M}$, *then series* (1) *regularly converges a.e. for every d-multiple ONS* φ *on* I.

Proof. Let us fix an ONS φ on I and set

$$G_N(x) = \Big(\sup_{m,n \in Z_+^d \setminus Q_N : m \leqq n} \Big| \sum_{m \leqq k \leqq n} a_k \varphi_k(x) \Big| \Big)^2 \quad (N = 1, 2, \ldots).$$

It is plain that

$$G_N(x) \geqq G_{N+1}(x) \geqq 0 \quad (N = 1, 2, \ldots).$$

Since

$$\int_0^1 G_N(x)\,dx \leqq \|\mathfrak{a}(Z_+^d \setminus Q_N)\|^2,$$

Corollary 2 yields

$$\lim_{N \to \infty} \int_0^1 G_N(x)\,dx = 0.$$

Hence, via Fatou's lemma, we obtain that

$$\lim_{N \to \infty} G_N(x) = 0 \quad \text{a.e.}$$

and this is equivalent to the a.e. regular convergence of series (1).

Theorem 6. *If* $\mathfrak{a} \notin \mathfrak{M}$, *then there exists an ONS* $\Phi = \{\Phi_k(x) \colon k \in Z_+^d\}$ *of step functions on I such that series* (1) *for* $\varphi = \Phi$ *does not converge regularly a.e. on I; even we have*

$$\limsup \Big| \sum_{k \leqq n} a_k \Phi_k(x) \Big| = \infty \quad \text{a.e. as} \quad \max_{1 \leqq j \leqq d} n_j \to \infty. \tag{20}$$

Proof. By (4) and (5) there exists a sequence $(0=)N_0 < N_1 < \ldots < N_p < \ldots$ of integers such that

$$\|\mathfrak{a}(Q_{N_p} \setminus Q_{N_{p-1}})\| \geqq p \quad (p = 1, 2, \ldots).$$

For each p we consider the sequence $\mathfrak{a}(Q_{N_p} \setminus Q_{N_{p-1}})$ and apply Lemma 2. As a result we obtain an ONS $\{\psi_k(p; x) \colon k \in Q_{N_p}\}$ of step functions and a simple set E_p for each $p = 1, 2, \ldots$ with the properties stated in Lemma 2.

By induction we will define an ONS $\Phi = \{\Phi_k(x) \colon k \in Z_+^d\}$ of step functions and a sequence $\{H_p \colon p = 1, 2, \ldots\}$ of stochastically independent, simple subsets of I having the following properties:

$$\max_{m,n \in Q_{N_p} \setminus Q_{N_{p-1}} : m \leqq n} \Big| \sum_{m \leqq k \leqq n} a_k \Phi_k(x) \Big| \geqq 2^{-d} p \quad \text{for} \quad x \in H_p \tag{21}$$

and

$$\operatorname{mes} H_p \geqq C_1 \quad (p = 1, 2, \ldots) \tag{22}$$

with the same constant as in Lemma 2.

For $p = 1$ we set

$$H_1 = E_1 \quad \text{and} \quad \Phi_k(x) = \psi_k(1; x) \quad (k \in Q_{N_1}).$$

Then (21) and (22) are obviously satisfied $(Q_0 = \emptyset)$.

Now let p_0 be a positive integer and assume that the step functions $\Phi_k(x)$ for $k\in Q_{N_{p_0}}$ and the simple sets $H_1, H_2, ..., H_{p_0}$ have been defined in such a way that these functions constitute an ONS on I, these sets are stochastically independent and relations (21) and (22) are satisfied for $p=1, 2, ..., p_0$. Then there exists a partition $\{J_r: r=1, 2, ..., \varrho\}$ of the interval I into disjoint subintervals such that each function $\Phi_k(x)$, $k\in Q_{N_{p_0}}$, assumes a constant value on each J_r, $r=1, 2, ..., \varrho$, and each set H_p, $p=1, 2, ..., p_0$, is the union of a certain number of J_r. Let us divide each J_r into two subintervals J_r' and J_r'' of equal length.

We shall use the following notations. Given a function $f(x)$ defined on I, a subset H and a subinterval $J=(a, b)$ of I, we define

$$f(J; x) = \begin{cases} f\left(\dfrac{x-a}{b-a}\right) & \text{for } x\in J, \\ 0 & \text{for } x\in I\setminus J; \end{cases}$$

and $H(J)$ to be the set, into which H is carried over by the linear transformation $y=(b-a)x+a$.

Now we define the functions $\Phi_k(x)$ for $k\in Q_{N_{p_0+1}}\setminus Q_{N_{p_0}}$ and the set H_{p_0+1} as follows:

$$\Phi_k(x) = \sum_{r=1}^{\varrho} [\psi_k(p_0+1; J_r'; x)-\psi_k(p_0+1; J_r''; x)]$$

and

$$H_{p_0+1} = \bigcup_{r=1}^{\varrho} [E_{p_0+1}(J_r')\cup E_{p_0+1}(J_r'')].$$

Obviously, these $\Phi_k(x)$, $k\in Q_{N_{p_0+1}}\setminus Q_{N_{p_0}}$, are step functions and H_{p_0+1} is a simple set. It is a routine to verify that the functions $\Phi_k(x)$, $k\in Q_{N_{p_0+1}}$, form an ONS on I, the sets H_p, $p=1, 2, ..., p_0+1$, are stochastically independent, and relations (21) and (22) are satisfied for $p=p_0+1$. (To deduce (21) from (10) one has to use a representation similar to (19).)

The above induction scheme shows that the ONS $\Phi=\{\Phi_k(x): k\in Z_+^d\}$ and the sequence $\{H_p: p\in Z_+^1\}$ of stochastically independent sets can be defined in such a way that conditions (21) and (22) hold true.

We set

$$H = \limsup_{p\to\infty} H_p.$$

By (22), the second Borel—Cantelli lemma implies that mes $H=1$. If $x\in H$, then $x\in H_p$ and consequently (21) holds true for an infinite number of p. In other words, this means that

$$\limsup \Big|\sum_{m\le k\le n} a_k\Phi_k(x)\Big| = \infty \quad \text{a.e. as} \quad \max_{1\le j\le d} m_j\to\infty.$$

Hence it is clear that series (1) for $\varphi=\Phi$ does not converge regularly a.e. Taking into account the representation of $\sum_{m\leq k\leq n} a_k\Phi_k(x)$ corresponding to (19), assertion (20) also follows.

Theorems 5 and 6 immediately yield the following two corollaries.

Corollary 4. *A necessary and sufficient condition that a d-multiple sequence* $\mathfrak{a}$ *of numbers be such that series* (1) *regularly converge a.e. for every ONS* φ *on I is that* $\mathfrak{a}\in\mathfrak{M}$.

Corollary 5. *If a d-multiple sequence* $\mathfrak{a}$ *of numbers is such that series* (1) *regularly converges a.e. for every ONS* φ *on I, then for every ONS* φ *the rectangular partial sums* $s_n(x)$ *of series* (1) *are majorized by a square integrable function* $F(x)=$ $=F(x;\ \mathfrak{a},\varphi)$ *on I, the square integral of which depends only on* $\mathfrak{a}$*, but not on* φ.

Indeed, the condition of Corollary 5 is equivalent to the fact that $\mathfrak{a}\in\mathfrak{M}$. In this case, setting

$$F(x)=\sup_{m,n\in Z_+^d:\,m\leq n}\Big|\sum_{m\leq k\leq n} a_k\varphi_k(x)\Big|,$$

we have

$$\int_0^1 F^2(x)\,dx\leq\|\mathfrak{a}\|^2<\infty,$$

as stated in Corollary 5.

Using a previous result of the second author, we are able to prove a stronger assertion than that is stated in Theorem 6. This makes possible to deduce our second main result; if the a.e. convergence of series (1) is considered for every ONS on I, then regular convergence and convergence in Pringsheim's sense are equivalent, up to a set of measure zero. This will be a corollary of the following

Theorem 7. *If* $\mathfrak{a}\notin\mathfrak{M}$, *then there exist an ONS* $\Phi=\{\Phi_k(x):\ k\in Z_+^d\}$ *of step functions on I such that*

$$\limsup\Big|\sum_{k\leq n} a_k\Phi_k(x)\Big|=\infty\quad \textit{a.e. as}\quad \min_{1\leq j\leq d} n_j\to\infty. \tag{23}$$

Consequently, series (1) *for* $\varphi=\Phi$ *does not converge a.e. even in Pringsheim's sense.*

Proof. It will be done by induction with respect to d. If $d=1$, Theorem 7 is a result of the second author [7].

For the sake of simplicity, we present the induction step from $d=1$ to $d+1=2$. In this case we write (k,l) instead of (k_1,k_2). For given positive integers k_0 and l_0 let us put

$$T_{k_0}^{(1)}=\{(k_0,l):\ l=1,2,\ldots\}\quad\text{and}\quad T_{l_0}^{(2)}=\{(k,l_0):\ k=1,2,\ldots\}$$

and consider the norms $\|\mathfrak{a}(T_{k_0}^{(1)})\|$ and $\|\mathfrak{a}(T_{l_0}^{(2)})\|$, respectively. We distinguish two cases.

Case (i). For all positive integers k_0 and l_0 we have respectively

$$\|\mathfrak{a}(T_{k_0}^{(1)})\| < \infty \quad \text{and} \quad \|\mathfrak{a}(T_{l_0}^{(2)})\| < \infty.$$

Applying the above mentioned theorem of the second author, we obtain that for every positive integer k_0 the single series

$$\sum_{l=1}^{\infty} a_{k_0,l}\,\varphi_l(x)$$

(a so-called "column") converges a.e. on I for every ONS $\{\varphi_l(x)\colon l=1,2,\ldots\}$; and for every positive integer l_0 the single series

$$\sum_{k=1}^{\infty} a_{k,l_0}\,\varphi_k(x)$$

(a so-called "row") converges a.e. on I for every ONS $\{\varphi_k(x)\colon k=1,2,\ldots\}$. Consequently, for every double ONS $\varphi=\{\varphi_{kl}(x)\colon k, l=1,2,\ldots\}$ and for every positive integer N we have

(24)
$$\limsup \Big|\sum_{k=1}^{m}\sum_{l=1}^{n} a_{kl}\varphi_{kl}(x)\Big| < \infty \quad \text{a.e. as} \quad \max(m,n)\to\infty \quad \text{and} \quad \min(m,n)\leq N.$$

In virtue of Theorem 6, there exists a double ONS $\Phi=\{\Phi_{kl}(x)\colon k, l=1,2,\ldots\}$ such that relation (20) holds true. Taking into account observation (24) we can strengthen (20) as follows:

$$\limsup \Big|\sum_{k=1}^{m}\sum_{l=1}^{n} a_{kl}\Phi_{kl}(x)\Big| = \infty \quad \text{a.e. as} \quad \min(m,n)\to\infty.$$

This is statement (23) for $d=2$.

Case (ii). There exists at least one positive integer k_0 or l_0, for which

$$\|\mathfrak{a}(T_{k_0}^{(1)})\| = \infty \quad \text{or} \quad \|\mathfrak{a}(T_{l_0}^{(2)})\| = \infty.$$

For definiteness, let us assume the fulfilment of the first relation. Again applying the theorem of the second author [7], we can find an ONS $\{\Psi_l(x)\colon l=1,2,\ldots\}$ of step functions on I such that the single series

$$\sum_{l=1}^{\infty} a_{k_0,l}\,\Psi_l(x)$$

diverges a.e. on I in the sense that

$$\limsup_{N\to\infty} \left|\sum_{l=1}^{N} a_{k_0,l}\,\Psi_l(x)\right| = \infty \quad \text{a.e.}$$

From here it follows that there exist a sequence $(0=)N_0<N_1<\ldots<N_p<\ldots$ of integers and a sequence $\{E_p\colon p=1,2,\ldots\}$ of simple subsets of I such that

$$\max_{N_{p-1}<N\leqq N_p} \left|\sum_{l=N_{p-1}+1}^{N} a_{k_0,l}\,\Psi_l(x)\right| \geqq p \quad \text{for} \quad x\in E_p \tag{25}$$

and

$$\operatorname{mes} E_p \geqq 1-2^{-p-1} \quad (p=1,2,\ldots). \tag{26}$$

We may assume that $N_1\geqq k_0$.

We are going to construct a double ONS $\Phi=\{\Phi_{kl}(x)\colon k,l=1,2,\ldots\}$ of step functions and another sequence $\{H_p\colon p=1,2,\ldots\}$ of simple subsets of I in such a way that

$$\max_{N_{p-1}<N\leqq N_p} \left|\sum_{k=N_{p-1}+1}^{N}\ \sum_{l=N_{p-1}+1}^{N} a_{kl}\,\Phi_{kl}(x)\right| \geqq p \quad \text{for} \quad x\in H_p \tag{27}$$

and

$$\operatorname{mes} H_p \geqq 1-2^{-p} \quad (p=1,2,\ldots). \tag{28}$$

We use again an induction argument, this time with respect to p. If $p=1$, we set for $l=1,2,\ldots,N_1$

$$\Phi_{k_0,l}(x) = \begin{cases} \sqrt{2}\,\Psi_l(2x) & \text{for} \quad x\in(0,1/2),\\ 0 & \text{for} \quad x\in(1/2,1); \end{cases}$$

and define the other functions $\Phi_{kl}(x)$ for $(k,l)\in Q_{N_1}=\{(k,l)\colon k,l=1,2,\ldots,N_1\}$, $k\neq k_0$, in such a way that they be zero on $(0,1/2)$ and they form an ONS on $(1/2,1)$ consisting of step functions. Furthermore, set $H_1=E_1$. It is clear that (27) and (28) are satisfied for $p=1$.

Now let p_0 be a positive integer and suppose that the step functions $\Phi_{kl}(x)$ for $(k,l)\in Q_{N_{p_0}}$ and the simple sets H_p for $p=1,2,\ldots,p_0$ have been defined in such a way that these functions form an ONS on I, and relations (27) and (28) are satisfied for $p=1,2,\ldots,p_0$. Then there exists a partition $\{J_s\colon s=1,2,\ldots,\sigma\}$ of the interval I into disjoint subintervals such that each function $\Phi_{kl}(x)$, $(k,l)\in Q_{N_{p_0}}$, assumes a constant value on each J_s, $s=1,2,\ldots,\sigma$.

Let us divide each J_s into three subintervals J_s', J_s'' and J_s''' with the following lengths:

$$\operatorname{mes} J_s' = \operatorname{mes} J_s'' = 2^{-1}(1-2^{-p_0-2}) \operatorname{mes} J_s \tag{29}$$

and

$$\text{mes } J_s''' = 2^{-p_0-2} \text{ mes } J_s \quad (s = 1, 2, \ldots, \sigma).$$

Now we define the functions $\Phi_{k_0,l}(x)$ for $l=N_{p_0}+1, N_{p_0}+2, \ldots, N_{p_0+1}$ and the set H_{p_0+1} as follows:

$$\Phi_{k_0,l}(x) = (1-2^{-p_0-2})^{-1/2} \sum_{s=1}^{\sigma} [\Psi_l(J_s'; x) - \Psi_l(J_s''; x)]$$

and

$$H_{p_0+1} = \bigcup_{s=1}^{\sigma} [E_{p_0+1}(J_s') \cup E_{p_0+1}(J_s'')].$$

Relation (27) follows from (25), while (28) follows from (26) and (29). It is clear that each function $\Phi_{k_0,l}(x)$, $N_{p_0} < l \leqq N_{p_0+1}$, vanishes on $\bigcup_{s=1}^{\sigma} J_s'''$ and H_{p_0+1} is also disjoint from $\bigcup_{s=1}^{\sigma} J_s'''$. Finally, we define the other functions $\Phi_{kl}(x)$ for $Q_{N_{p_0+1}} \setminus Q_{N_{p_0}}$, $k \neq k_0$, in such a way that they vanish on $\bigcup_{s=1}^{\sigma} (J_s' \cup J_s'')$ and they form an ONS on $\bigcup_{s=1}^{\sigma} J_s'''$, consisting of step functions with zero mean on each interval J_s''', $s=1, 2, \ldots, \sigma$.

By construction, the step functions $\Phi_{kl}(x)$, $(k, l) \in Q_{N_{p_0+1}}$, form and ONS on I, the sets $H_1, H_2, \ldots, H_{p_0+1}$ are simple, and relations (27) and (28) are satisfied for $p=1, 2, \ldots, p_0+1$. This completes the proof of the induction step.

We set

$$H = \limsup_{p\to\infty} H_p.$$

By (28), the first Borel—Cantelli lemma implies that

$$\text{mes}\left[\liminf_{p\to\infty} (I \setminus H_p)\right] = 0, \quad \text{or equivalently,} \quad \text{mes } H = 1.$$

If $x \in H$, then (27) holds true for an infinite number of p, consequently,

$$\limsup_{p\to\infty} \max_{N_{p-1} < N \leqq N_p} \Big| \sum_{k=N_{p-1}+1}^{N} \sum_{l=N_{p-1}+1}^{N} a_{kl} \Phi_{kl}(x) \Big| = \infty \quad \text{a.e.}$$

Taking into account that

$$\sum_{k=N_{p-1}+1}^{N} \sum_{l=N_{p-1}+1}^{N} a_{kl} \Phi_{kl}(x) =$$
$$= \left\{ \sum_{k=1}^{N} \sum_{l=1}^{N} - \sum_{k=1}^{N} \sum_{l=1}^{N_{p-1}} - \sum_{k=1}^{N_{p-1}} \sum_{l=1}^{N} + \sum_{k=1}^{N_{p-1}} \sum_{l=1}^{N_{p-1}} \right\} a_{kl} \Phi_{kl}(x),$$

assertion (23) for $d=2$ immediately follows.

The proof of Theorem 7 is complete.

We emphasize the significance of the following two consequences of Theorems 5, 6 and 7.

Corollary 6. *If a d-multiple sequence* $\mathfrak{a}$ *of numbers is such that for every ONS* φ *series* (1) *converges in Pringsheim's sense on a set of positive measure, perhaps depending on* φ, *then series* (1) *for every ONS* φ *regularly converges a.e.*

Corollary 7. *If a d-multiple sequence* $\mathfrak{a}$ *of numbers is such that for an ONS* φ *series* (1) *does not converge regularly on a set of positive measure, then there exists another ONS* Φ *such that series* (1) *for* $\varphi=\Phi$ *does not converge in Pringsheim's sense a.e.*

We note that for an individual ONS the notions of a.e. regular convergence and a.e. convergence in Pringsheim's sense can essentially differ from each other. In [4, pp. 214—215] a double sequence $\{a_{kl}: k, l=1, 2, \ldots\}$ of real numbers and on $I^2=[0; 1]^2$ a double ONS $\{\varphi_{kl}(x): k, l=1, 2, \ldots\}$ are constructed in such a way that

$$\sum_{k=1}^{\infty}\sum_{l=1}^{\infty} a_{kl}^2 < \infty,$$

the double orthogonal series

$$\sum_{k=1}^{\infty}\sum_{l=1}^{\infty} a_{kl}\varphi_{kl}(x)$$

converges in Pringsheim's sense a.e on I^2, but does not converge regularly on a set of measure at least 1/2. It is not hard to modify this example so as the resulting orthogonal series converges in Pringsheim's sense a.e. and does not converge regularly a.e.

5. Estimation of the norm $\|\mathfrak{a}\|$

Using the d-multiple generalization of the famous Rademacher—Menšov inequality, it is not hard to give an upper bound for $\|\mathfrak{a}\|$ (see [3, Corollary 2]).

Theorem 8. *For every d-multiple sequence* $\mathfrak{a}$ *we have*

$$\|\mathfrak{a}\| \leqq C_2\Big\{\sum_{k\in Z_+^d} a_k^2 \prod_{j=1}^{d}(\log 2k_j)^2\Big\}^{1/2}, \tag{30}$$

where $C_2=C_2(d)$.

Here and in the sequel the logarithms are to the base 2.

A nontrivial lower bound for $\|\mathfrak{a}\|$ is not known in general. In the special case when $\mathfrak{a}$ is such that $\{|a_k|: k\in Z_+^d\}$ is nonincreasing in the sense that

$$|a_k| \geqq |a_n| \quad \text{whenever} \quad k, n\in Z_+^d \quad \text{and} \quad k \leqq n, \tag{31}$$

an opposite inequality to (30) holds also true.

Theorem 9. *If a d-multiple sequence* $\mathfrak{a}$ *is such that* (31) *is satisfied, then we have*

$$\|\mathfrak{a}\| \geqq C_3\{\sum_{k\in Z_+^d} a_k^2 \prod_{j=1}^d (\log 2k_j)^2\}^{1/2}, \tag{32}$$

where $C_3=C_3(d)$.

The proof of Theorem 9 is based on the following basic result of Menšov [2].

Lemma 3. *For every positive integer N there exist an ONS* $\{\psi_k^{(N)}(x): k=1, 2, \ldots, N\}$ *of step functions on the interval I and a simple subset* $E^{(N)}$ *of I such that*

$$\operatorname{mes} E^{(N)} \geqq C_4 \tag{33}$$

and for every $x\in E^{(N)}$ *there exists an integer* $n=n(x)$ *between* 1 *and N such that* $\psi_k^{(N)}(x)\geqq 0$ *for* $k=1, 2, \ldots, n$ *and*

$$\sum_{k=1}^n \psi_k^{(N)}(x) \geqq C_5 \sqrt{N} \log 2N. \tag{34}$$

A trivial consequence of (33) and (34) is that

$$\int_0^1 (\max_{1\leqq n\leqq N} |\sum_{k=1}^n \psi_k^{(N)}(x)|)^2\, dx \geqq C_6 N (\log 2N)^2. \tag{35}$$

This inequality will be enough for our purpose.

Proof of Theorem 9. For the sake of simplicity in notations, we present the proof again for the case $d=2$.

Denote by T a measure-preserving transformation of the square I^2 onto the interval I: $T(y_1, y_2)=x$, where $(y_1, y_2)\in I^2$ and $x\in I$. Given two positive integers N_1 and N_2, we define for $k=1, 2, \ldots, N_1$; $l=1, 2, \ldots, N_2$

$$\varphi_{kl}^{(N_1, N_2)}(x) = \psi_k^{(N_1)}(y_1)\, \psi_l^{(N_2)}(y_2).$$

Then (35) yields

$$\begin{aligned}\int_0^1 (\max_{1\leqq m\leqq N_1} \max_{1\leqq n\leqq N_2} |\sum_{k=1}^m \sum_{l=1}^n \varphi_{kl}^{(N_1, N_2)}(x)|)^2\, dx \geqq \\ \geqq C_6^2 N_1 N_2 (\log 2N_1)^2 (\log 2N_2)^2.\end{aligned} \tag{36}$$

After these preliminaries, let us consider the partition of Z_+^2 into the following "dyadic" rectangles:

$$Q_{mn} = \{(k, l)\in Z_+^2 : 2^{m-1} \leqq k < 2^m \quad \text{and} \quad 2^{n-1} \leqq l < 2^n\},$$

where (m, n) runs over Z_+^2. According to this partition we modify the original sequence $\mathfrak{a}$ into another $\mathfrak{a}^*$ so as it should be constant on each Q_{mn}: $\mathfrak{a}^* = \{a_{kl}^* : (k, l)\in Z_+^2\}$, where

$$a_{kl}^* = a_{2^m, 2^n} \quad \text{for} \quad (k, l)\in Q_{mn}, \quad (m, n)\in Z_+^2.$$

Due to Theorem 2, inequality (36), and the monotony of $|a_{kl}|$, for every $(m, n)\in Z_+^2$

$$\|\mathfrak{a}(Q_{mn})\| \geqq \|\mathfrak{a}^*(Q_{mn})\| \geqq C_6^2 2^{m-1} 2^{n-1} m^2 n^2 a_{2^m, 2^n}^2 \geqq$$

$$\geqq 3^{-4} C_6^2 \sum_{k=2^m}^{2^{m+1}-1} \sum_{l=2^n}^{2^{n+1}-1} a_{kl}^2 (\log 2k)^2 (\log 2l)^2.$$

Applying Theorem 3, we obtain that

$$\|\mathfrak{a}\|^2 \geqq \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \|\mathfrak{a}(Q_{mn})\|^2 \geqq 3^{-4} C_6^2 \sum_{k=2}^{\infty} \sum_{l=2}^{\infty} a_{kl}^2 (\log 2k)^2 (\log 2l)^2. \tag{37}$$

Now we examine the cases $m=1$ or (and) $n=1$ once more. A more accurate calculation gives

$$\|\mathfrak{a}(Q_{1n})\| \geqq \|\mathfrak{a}^*(Q_{1n})\| \geqq C_6 2^{n-1} n^2 a_{1, 2^n}^2 \geqq 3^{-2} C_6 \sum_{l=2^n}^{2^{n+1}-1} a_{1l}^2 (\log 2l)^2,$$

whence we get that

$$\|\mathfrak{a}\|^2 \geqq \sum_{n=1}^{\infty} \|\mathfrak{a}(Q_{1n})\|^2 \geqq 3^{-2} C_6 \sum_{l=2}^{\infty} a_{1l}^2 (\log 2l)^2. \tag{38}$$

Analogously,

$$\|\mathfrak{a}\|^2 \geqq 3^{-2} C_6 \sum_{k=2}^{\infty} a_{k1}^2 (\log 2k)^2. \tag{39}$$

Finally, it is obvious that

$$\|\mathfrak{a}\|^2 \geqq a_{11}^2. \tag{40}$$

Now the statement of Theorem 9 immediately follows from relations (37)—(40).

Remark 4. If one treats each "finite" sequence $\mathfrak{a}(Q_N)$, $N=1, 2, \ldots$, separately instead of the whole sequence $\mathfrak{a}$ and makes use of the fact that all $\psi_k^{(N)}(x)$ are step functions, one can prove Theorem 9 without taking a measure-preserving transformation T of the unit sequare I^2 onto I.

References

[1] G. H. Hardy, On the convergence of certain multiple series, *Proc. Cambridge Philosoph. Soc.*, **19** (1916—1919), 86—95.
[2] D. E. Menchoff, Sur les séries de fonctions orthogonales, *Fund. Math.*, **4** (1923), 82—105.
[3] F. Móricz, On the convergence in a restricted sense of multiple series, *Analysis Math.*, **5** (1979), 135—147.
[4] F. Móricz, The regular convergence of multiple series, in: *Functional Analysis and Approximation (Proc. Conf. Oberwolfach, 1980)*, Birkhäuser (Basel, 1981), 203—218.
[5] F. Móricz and K. Tandori, On the divergence of multiple orthogonal series, *Acta Sci. Math.*, **42** (1980), 133—142.
[6] K. Tandori, Über die Konvergenz der Orhogonalreihen, *Acta Sci. Math.*, **24** (1963), 135—151.
[7] K. Tandori, Über die Konvergenz der Orthogonalreihen. II, *Acta Sci. Math.*, **25** (1964), 219—232.

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Embedding theorems and strong approximation

L. LEINDLER* and A. MEIR

1. Let $f(x)$ be a continuous and 2π-periodic function and let

$$f(x)\sim\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx) \tag{1}$$

be its Fourier series. We denote by $s_n=s_n(x)=s_n(f;\ x)$ the n-th partial sum of (1), the usual supremum norm by $\|\cdot\|$ and by $E_n=E_n(f)$ the best approximation of f by trigonometric polynomials of order at most n. Let $\omega(\delta)$ be a nondecreasing continuous function on the interval $[0, 2\pi]$ having the properties: $\omega(0)=0$, $\omega(\delta_1+\delta_2)\leqq\omega(\delta_1)+\omega(\delta_2)$ for any $0\leqq\delta_1\leqq\delta_2\leqq\delta_1+\delta_2\leqq 2\pi$. Such a function is called a modulus of continuity.

In order to quote the result of [3], which has initiated our present investigation, we define two classes of functions:

$$H^{\omega}:=\{f:\ \omega(f;\ \delta)=O(\omega(\delta))\}$$

and

$$S_p(\lambda):=\{f:\ \|\sum_{n=0}^{\infty}\lambda_n|s_n-f|^p\|<\infty\},$$

where $\lambda=\{\lambda_n\}$ is a monotonic sequence of positive numbers and $0<p<\infty$. V. G. Krotov and L. Leindler [3] proved the following result.

Theorem A. *If $\{\lambda_n\}$ is a monotonic sequence, ω is a modulus of continuity and $0<p<\infty$, then*

$$\sum_{k=1}^{n}(k\lambda_k)^{-1/p}=O\left(n\omega\left(\frac{1}{n}\right)\right) \tag{2}$$

Received March 31, 1983.

*) The research of the first author was supported by an NSERC of Canada grant while visiting the University of Alberta, Edmonton.

implies

(3) $$S_p(\lambda) \subset H^\omega.$$

Conversely, if there exists a number Q *such that* $0 \leqq Q < 1$ *and*

(4) $$n^Q \lambda_n \uparrow,$$

then (3) *implies* (2).

It is well known that the classical de la Vallée Poussin means

$$\tau_n = \tau_n(f;\, x) := \frac{1}{n} \sum_{k=n+1}^{2n} s_k(x), \quad n = 1, 2, \ldots$$

usually approximate the function f, in the sup norm, better than the partial sums do. Hence, if in analogy to $S_p(\lambda)$ we consider the class of functions

$$V_p(\lambda) := \{f\colon \|\sum_{n=0}^{\infty} \lambda_n |\tau_n - f|^p\| < \infty\},$$

we may expect that under reasonable conditions the following embedding relations will hold

(5) $$S_p(\lambda) \subset V_p(\lambda) \subset H^\omega.$$

In the present paper we establish that condition (2) does imply the inclusion $V_p(\lambda) \subset H^\omega$ for all positive p. We further show that the embedding relation $S_p(\lambda) \subset V_p(\lambda)$ also holds if $p \geqq 1$ and the sequence $\{\lambda_n\}$ satisfies the mild restriction

(6) $$\frac{\lambda_n}{\lambda_{2n}} \leqq K, \quad n = 1, 2, \ldots,$$

with a fixed positive K ($K, K_1, K_2, \ldots$ will denote positive constants, not necessarily the same at each occurrence).

We were unable to decide whether $S_p(\lambda) \subset V_p(\lambda)$ holds when $0 < p < 1$; it is left as an open problem.

2. We shall establish the following results.

Theorem 1. *If* $p \geqq 1$ *and* $\{\lambda_n\}$ *is a monotonic (nondecreasing or nonincreasing) sequence of positive numbers satisfying* (6), *then*

(7) $$S_p(\lambda) \subset V_p(\lambda)$$

holds.

Theorem 2. *Let* $\{\lambda_n\}$ *be a monotonic sequence of positive numbers, furthermore let* ω *be a modulus of continuity and* $0 < p < \infty$. *Then condition* (2) *implies*

(8) $$V_p(\lambda) \subset H^\omega.$$

If $p \geqq 1$ *and there exists a number* Q *such that* $0 \leqq Q < 1$ *and* (4) *holds, then, conversely,* (8) *implies* (2).

3. To prove our theorems we require the following lemmas.

Lemma 1 ([1, p. 534]). *For any continuous function* f *we have the following inequality*

$$\omega\left(f; \frac{1}{n}\right) \leqq Kn^{-1} \sum_{k=1}^{n} E_k(f). \tag{9}$$

Lemma 2. *Let* $a = \{a_n\}_0^\infty$ *be a nonincreasing sequence of positive numbers,* $q > 0$ *and* $\gamma > 0$. *Then there exists a positive constant* $C = C(a, \gamma, q)$ *such that for every* m

$$\sum_{n=0}^{m} q^n a_n \leqq C \cdot \sum_{n=0}^{m} q^n a_n \left(\frac{a_{n+1}}{a_n}\right)^{\gamma}. \tag{10}$$

Proof. We let $\beta = \min(a_1/a_2, 1/2q)$. We define the (possibly finite) sequence of integers $N_0 < N_1 < \ldots$ as follows. Let $N_0 = 0$. For $i \geqq 1$ let N_i be the smallest integer such that $N_i > N_{i-1}$ and $a_{N_i+1} \geqq \beta a_{N_i}$; if no such integer exists we set $N_i = \infty$. Now, if $N_i < n < N_{i+1}$, then $a_{n+1} < \beta a_n$ and so $a_{N_i+r}\beta^{r-1}$ for $r = 1, 2, \ldots, N_{i+1} - N_i$. Therefore, we have for $i = 0, 1, \ldots$

$$\sum_{n=N_i+1}^{N_{i+1}} q^n a_n \leqq q^{N_i+1} a_{N_i+1} \cdot (1 + q\beta + q^2\beta^2 + \ldots) \leqq 2q^{N_i+1} a_{N_i+1} \leqq 2q^{N_i+1} a_{N_i} \leqq$$
$$\leqq 2q\beta^{-\gamma} q^{N_i} a_{N_i} \left(\frac{a_{N_i+1}}{a_{N_i}}\right)^{\gamma}, \tag{11}$$

on using, in the last inequality, the definition of the sequence $\{N_i\}$. Now, for any given integer m, let j be the largest integer so that $N_j < m$. We then have, on using (11), and the fact that $\beta \leqq a_1/a_0$,

$$\sum_{n=0}^{m} q^n a_n \leqq \beta^{-\gamma} a_0 \left(\frac{a_1}{a_0}\right)^{\gamma} + 2q\beta^{-\gamma} \sum_{i=0}^{j} q^{N_i} a_{N_i} \left(\frac{a_{N_i+1}}{a_{N_i}}\right)^{\gamma} \leqq C \cdot \sum_{n=0}^{m} q^n a_n \left(\frac{a_{n+1}}{a_n}\right)^{\gamma},$$

with $C = \beta^{-\gamma}(1 + 2q)$.

This completes the proof of Lemma 2.

4. Proof of Theorem 1. For $p \geqq 1$ we have, by the "power sum inequality",

$$|\tau_n - f|^p \leqq \frac{1}{n} \sum_{k=n+1}^{2n} |s_k - f|^p.$$

Hence

$$(12)\qquad \begin{aligned}\sum_{n=1}^{\infty}\lambda_n|\tau_n-f|^p &\leqq \sum_{n=1}^{\infty}(\lambda_n/n)\sum_{k=n+1}^{2n}|s_k-f|^p \leqq\\ &\leqq \sum_{k=2}^{\infty}|s_k-f|^p\sum_{n=k/2}^{k-1}(\lambda_n/n) \leqq K\sum_{k=2}^{\infty}\lambda_k|s_k-f|^p,\end{aligned}$$

where the last inequality follows from (6). Inequality (12) clearly implies (7).

Proof of Theorem 2. First we consider the case $p\geqq 1$. Suppose $f\in V_p(\lambda)$. Then we have for $n=1,2,\ldots$

$$(13)\qquad E_{4n}(f)\leqq\left\|\frac{1}{n}\sum_{k=n+1}^{2n}|\tau_k-f|\right\|\leqq\left\|\left\{\frac{1}{n}\sum_{k=n+1}^{2n}|\tau_k-f|^p\right\}^{1/p}\right\|\leqq K_1(n\lambda_n^*)^{-1/p},$$

where $\lambda_n^*=\min(\lambda_{n+1},\lambda_{2n})$ and the last inequality follows from the assumption $f\in V_p(\lambda)$. Now, from (13), both if $\{\lambda_n\}$ is increasing or decreasing we can deduce that

$$(14)\qquad \sum_{\nu=1}^{m}4^{\nu}E_{4^{\nu}}(f)\leqq K_2\sum_{\nu=0}^{m}4^{\nu}(4^{\nu}\lambda_{4^{\nu}})^{-1/p},$$

with a suitable $K_2>0$.

Hence, by Lemma 1 and (2), for $m=1,2,\ldots$

$$(15)\qquad \omega(f;\,4^{-m})\leqq K_3\omega(4^{-m}),$$

which proves that $f\in H^{\omega}$.

We turn now to the case $0<p<1$. We have for $n=1,2,\ldots$

$$(16)\qquad nE_{4n}(f)\leqq\left\|\sum_{k=n+1}^{2n}|\tau_k-f|\right\|=\left\|\sum_{k=n+1}^{2n}|\tau_k-f|^p\cdot|\tau_k-f|^{1-p}\right\|.$$

It is known [see e.g. [2], p. 58] that $\|\tau_k-f\|\leqq KE_k(f)$ for all k; hence, in particular, for $n+1\leqq k\leqq 2n$, $\|\tau_k-f\|\leqq KE_n(f)$. Therefore, from (16) we obtain that

$$nE_{4n}(f)\leqq K(E_n(f))^{1-p}\left\|\sum_{k=n+1}^{2n}|\tau_k-f|^p\right\|$$

which, since $f\in V_p(\lambda)$, implies that $E_{4n}(f)\leqq K_1(E_n(f))^{1-p}(n\lambda_n^*)^{-1}$, with $\lambda_n^*=\min(\lambda_{n+1},\lambda_{2n})$. If we rewrite the last inequality as

$$E_n(f)\cdot\left(\frac{E_{4n}(f)}{E_n(f)}\right)^{1/p}\leqq K_2(n\lambda_n^*)^{-1/p},$$

and use it for $n=4^{\nu}$, $\nu=0,1,\ldots,m$, we see that

$$\sum_{\nu=0}^{m}4^{\nu}E_{4^{\nu}}(f)\left(\frac{E_{4^{\nu+1}}(f)}{E_{4^{\nu}}(f)}\right)^{1/p}\leqq K_3\sum_{\nu=0}^{m}4^{\nu}(4^{\nu}\lambda_{4^{\nu}})^{-1/p}$$

holds. Applying Lemma 2 now with $a_n = E_{4^n}$, $q=4$ and $\gamma = 1/p$, we get that (14) is satisfied in this case as well. Hence, as before, f satisfies (15) and so $f \in H^\omega$.

This completes the proof of (8) for all positive p.

In order to prove that, under the stated assumptions, (8) implies (2), it is sufficient to note that, because of (7), relation (3) of Theorem A is satisfied; therefore Theorem A provides the proof of the required assertion.

References

[1] N. BARY, *A treatise on trigonometric series* (Oxford, 1964).

[2] R. P. FEINERMAN and D. J. NEWMAN, *Polynomial Approximation,* Waverly Press (1974).

[3] V. G. KROTOV and L. LEINDLER, On the strong summability of Fourier series and the class H^ω, *Acta. Sci. Math.,* **40** (1978), 93—98.

(L. L.)
BOLYAI INTÉZET
ARADI VÉRTANÚK TERE 1
SZEGED, HUNGARY

(A. M.)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ALBERTA
EDMONTON, CANADA

Acta Sci. Math. 47 (1984), 377—389

On the unicity of best Chebyshev approximation of differentiable functions

ANDRÁS KROÓ*

Let X be a normed linear space, U_n an n-dimensional subspace of X. The problem of best approximation consists in determining for each $x\in X$ its best approximations in U_n, i.e. such elements $p\in U_n$ for which $\|x-p\|=\text{dist}(x, U_n)=\inf\{\|x-g\|: g\in U_n\}$. Let us denote by $G(x)=\{p\in U_n: \|x-p\|=\text{dist}(x, U_n)\}$ the set of best approximations of x. Evidently, for every $x\in X$ the set $G(x)$ is nonempty and convex. Recall that the convex set $G(x)$ is said to have dimension k if there exist $k+1$ elements $p_0, \ldots, p_k\in G(x)$ such that p_i-p_0, $1\leqq i\leqq k$, are linearly independent and $G(x)$ does not contain $k+2$ elements satisfying this property ($k\geqq 0$). The subspace U_n is called k-Chebyshev if the dimension of $G(x)$ is at most k for any $x\in X$ ($0\leqq k\leqq n-1$). In particular when $k=0$, i.e. each $x\in X$ possesses a unique best approximation in U_n, we say that U_n is a Chebyshev subspace of X.

Let us consider the classical case of Chebyshev approximation when $X=C(Q)$ is the space of complex valued continuous functions on the compact Hausdorff space Q endowed with the supremum norm $\|f\|_C=\sup\{|f(x)|: x\in Q\}$. (The subspace of real valued functions in $C(Q)$ will be denoted by $C_0(Q)$.) The characterization of Chebyshev subspaces of $C(Q)$ is given by the celebrated Haar—Kolmogorov theorem: the n-dimensional subspace U_n is a Chebyshev subspace of $C(Q)$ if and only if each $p\in U_n\setminus\{0\}$ has at most $n-1$ distinct zeros at Q. (This theorem was given at first by HAAR [3] in the real case $X=C_0(Q)$ and then by KOLMOGOROV [4] in the complex case $X=C(Q)$.) Later RUBINSTEIN [8] gave the characterization of k-Chebyshev subspace of $C_0(Q)$ and ROMANOVA [7] generalized it for $C(Q)$. Their result reads as follows: U_n is a k-Chebyshev subspace of $C(Q)$ if and only if any $k+1$ linearly independent elements of U_n have at most $n-k-1$ common zeros at Q ($0\leqq k\leqq n-1$). (For $k=0$ this result immediately implies the Haar—Kolmogorov theorem.)

Received December 3, 1982.

*) A part of this paper was written during the author's visit at the Royal Institute of Technology, Department of Mathematics, Stockholm, Sweden.

In the present paper we shall investigate the unicity of best Chebyshev approximation in the spaces of differentiable functions. This problem was posed by S. B. Stechkin and considered in the real case by GARKAVI [2].

Let $C^r[a, b]$ ($C_0^r[a, b]$) denote the space of r-times continuously differentiable complex (resp. real) functions on $[a, b]$ endowed with the supremum norm, $1 \leqq r \leqq \infty$. (In what follows $c \in [a, b]$ will be called a special zero of $f \in C_0^1[a, b]$ if either $f'(c) = f(c) = 0$ or $f(c) = 0$ and c coincides with one of the endpoints of the interval $[a, b]$.) GARKAVI [2] gave the following characterization of k-Chebyshev subspaces of $C_0^r[a, b]$: $U_n \subset C_0^r[a, b]$ ($r \geqq 1$) is a k-Chebyshev subspace of $C_0^r[a, b]$ if and only if for any s linearly independent elements $p_1, \ldots, p_s \in U_n$ ($k+1 \leqq s \leqq n$) among their common zeros there are not more than $n-s$ special zeros common to any $k+1$ of the elements $p_1, \ldots, p_s$. In particular in order that U_n be a Chebyshev subspace of $C_0^r[a, b]$ it is necessary and sufficient that for any s linearly independent elements $p_1, \ldots, p_s$ of U_n ($1 \leqq s \leqq n$) among their common zeros there are at most $n-s$ special zeros of any of $p_1, \ldots, p_s$. (Remark, that the characterization of k-Chebyshev subspaces of $C_0^r[a, b]$ turns out to be independent of $1 \leqq r \leqq \infty$.)

In this paper we shall present another approach to the study of k-Chebyshev and Chebyshev subspaces of differentiable functions. This approach is based on the so-called "extremal sets" which are essential in the study of unicity of best Chebyshev approximation of *complex valued* differentiable functions. Our method gives a possibility to generalize Garkavi's result to the complex case. In the last sections of the paper we shall give several applications for the study of unicity of best Chebyshev approximation of differentiable functions by real and complex lacunary polynomials.

1. First of all let us formulate a lemma characterizing best approximants. Recall, that the sign of a complex number $c \in \mathbf{C}$ is given by $\bar{c}/|c|$ if $c \neq 0$ and 0 if $c = 0$.

Lemma 1 ([9], p. 178). *Let U_n be an n-dimensional subspace of $C(Q)$ $(C_0(Q))$. Then $p \in U_n$ is a best approximation of $f \in C(Q)$ $(C_0(Q))$ if and only if there exist m points $x_1, \ldots, x_m \in Q$, where $1 \leqq m \leqq n+1$ in the real case and $1 \leqq m \leqq 2n+1$ in the complex case, and m numbers $a_1, \ldots, a_m \neq 0$ such that*

$$\sum_{j=1}^{m} a_j g(x_j) = 0 \tag{1}$$

for each $g \in U_n$ and

$$f(x_j) - p(x_j) = \operatorname{sign} a_j \|f-p\|_C \quad (1 \leqq j \leqq m).$$

This lemma suggests the following definition.

Definition. The set of m distinct points $x_1, \ldots, x_m \in Q$, where $1 \leqq m \leqq n+1$ in the real case and $1 \leqq m \leqq 2n+1$ in the complex case, is called an extremal set of

$U_n \subset C(Q)$ if there exist nonzero complex numbers $a_1, \ldots, a_m \neq 0$ (real if $U_n \subset C_0(Q)$) such that (1) holds for any $g \in U_n$.

If $\{x_i\}_{i=1}^m$ is an extremal set of U_n then the corresponding numbers $\{a_i\}_{i=1}^m$ are called the coefficients of this extremal set. Evidently the coefficients of an extremal set are defined in general nonuniquely (even with a normalization). Note that extremal sets are closely related to the set Q, on which the functions of U_n are considered. (The idea of the above definition comes essentially from REMEZ [6] who was the first to give a proposition like Lemma 1.)

Using the notion of extremal sets we can formulate the Rubinstein—Romanova (and in particular the Haar—Kolmogorov) theorem in the following way: U_n is a k-Chebyshev subspace of $C(Q)$ if and only if the points of an extremal set of U_n cannot be common zeros of $k+1$ linearly independent elements of U_n $(0 \leqq k \leqq n-1)$. In particular U_n is Chebyshev if and only if no $p \in U_n \setminus \{0\}$ can vanish on an extremal set of U_n. The proof is left to the reader. Similar characterizations of Chebyshev subspaces of $C(Q)$ were also given by CHENEY—WULBERT [1] and PHELPS [5].

The next theorem characterizing the k-Chebyshev (in particular Chebyshev) subspaces of $C^r[a, b]$ is our principal result. This characterization is essentially based on extremal sets since it also involves the coefficients of extremal sets.

Theorem 1. *Let U_n be a subspace of $C^r[a, b]$, $1 \leqq r \leqq \infty$, $0 \leqq k \leqq n-1$. Then U_n is a k-Chebyshev subspace of $C^r[a, b]$ if and only if there does not exist an extremal set $\{x_i\}_{i=1}^m \subset [a, b]$ of U_n with coefficients $\{a_i\}_{i=1}^m$ and $k+1$ linearly independent elements $p_1, \ldots, p_{k+1} \in U_n$ such that $p_j(x_i)=0$ $(1 \leqq i \leqq m,\ 1 \leqq j \leqq k+1)$ and $\operatorname{Re} a_i p_j'(x_i) = 0$ for each $1 \leqq j \leqq k+1$ and $x_i \in (a, b)$.*

In particular U_n is a Chebyshev subspace of $C^r[a, b]$ if and only if there does not exist an extremal set $\{x_i\}_{i=1}^m$ of U_n with coefficients $\{a_i\}_{i=1}^m$ and $p \in U_n \setminus \{0\}$ such that $p(x_i)=0$ $(1 \leqq i \leqq m)$ and $\operatorname{Re} a_i p'(x_i)=0$ for each $x_i \in (a, b)$.

In the real case the coefficients of the extremal set do not appear in the characterization theorem and therefore its formulation is much simpler.

Corollary 1. *In order that U_n be a k-Chebyshev subspace of $C_0^r[a, b]$ it is necessary and sufficient that the points of an extremal set of U_n cannot be common special zeros of $k+1$ linearly independent elements of U_n.*

In particular U_n is a Chebyshev subspace of $C_0^r[a, b]$ if and only if the points of an extremal set of U_n cannot be special zeros of a nontrivial element of U_n.

The above corollary is equivalent to Garkavi's result. It also follows from a result of BROSOWSKI—STOER [11] where an extension of Garkavi's result for *real* rational families was given.

Proof of Theorem 1. Sufficiency. Assume that U_n is not a k-Chebyshev subspace of $C^r[a, b]$ $(1 \leqq r \leqq \infty)$. Then there exists $f \in C^r[a, b]$ with best approximants

$0, p_1, \ldots, p_{k+1} \in U_n$, where $p_1, \ldots, p_{k+1}$ are linearly independent. Since 0 is a best approximation of f it follows from Lemma 1 that we can find an extremal set $\{x_i\}_{i=1}^m$ of U_n with coefficients $\{a_i\}_{i=1}^m$ such that

$$f(x_j) = \operatorname{sign} a_j \|f\|_C, \quad 1 \leqq j \leqq m. \tag{2}$$

Moreover, $\|f-p_s\|_C = \|f\|_C$ for each $1 \leqq s \leqq k+1$. Hence and by (2) we have

$$\begin{aligned} \|f\|_C^2 &\geqq |f(x_j)-p_s(x_j)|^2 = \big|\|f\|_C - \operatorname{sign} \bar{a}_j p_s(x_j)\big|^2 = \\ &= \left(\|f\|_C - \frac{1}{|a_j|} \operatorname{Re} a_j p_s(x_j)\right)^2 + \left(\frac{1}{|a_j|} \operatorname{Im} a_j p_s(x_j)\right)^2 \\ &\qquad (1 \leqq j \leqq m, \quad 1 \leqq s \leqq k+1). \end{aligned} \tag{3}$$

We can easily derive from (3) that $\operatorname{Re} a_j p_s(x_j) \geqq 0$ for each $1 \leqq j \leqq m$ and $1 \leqq s \leqq$ $\leqq k+1$. On the other hand by the definition of extremal sets $\sum_{j=1}^{m} a_j p_s(x_j) = 0$ for every $1 \leqq s \leqq k+1$. Hence $\operatorname{Re} a_j p_s(x_j) = 0$ $(1 \leqq j \leqq m,\ 1 \leqq s \leqq k+1)$. Moreover, this and (3) imply that $\operatorname{Im} a_j p_s(x_j) = 0$, too. Since all coefficients $a_j \neq 0$ we finally obtain that

$$p_s(x_j) = 0 \quad (1 \leqq j \leqq m,\ 1 \leqq s \leqq k+1). \tag{4}$$

Now we shall use the differentiability of the functions involved. Consider an arbitrary $x_j \in (a, b)$ and set $\tilde{f}(x) = (1/|a_j|) \operatorname{Re} a_j f(x)$, $\tilde{p}_s(x) = (1/|a_j|) \operatorname{Re} a_j p_s(x)$, $1 \leqq s \leqq$ $\leqq k+1$. Obviously, $\tilde{f}, \tilde{p}_1, \ldots, \tilde{p}_{k+1} \in C_0^r[a, b]$; $\tilde{f}(x_j) = \|\tilde{f}\|_C = \|f\|_C$, $\tilde{p}_s(x_j) = 0$ $(1 \leqq s \leqq$ $\leqq k+1)$ and $\|\tilde{f} - \tilde{p}_s\|_C = \|\tilde{f}\|_C$ $(1 \leqq s \leqq k+1)$. Since $x_j \in (a, b)$ is an extremum point of the real function $\tilde{f}$ it follows that $\tilde{f}'(x_j) = 0$. Therefore for any $h \in \mathbf{R}$ such that $|h| < \min\{x_j - a, b - x_j\}$ we have

$$\begin{aligned} \tilde{p}_s(x_j+h) &\geqq \tilde{f}(x_j+h) - \|\tilde{f}\|_C = \tilde{f}(x_j+h) - \tilde{f}(x_j) \geqq \\ &\geqq -|h|\omega(\tilde{f}', |h|) \quad (1 \leqq s \leqq k+1). \end{aligned} \tag{5}$$

(Here and in what follows we denote by $\omega(F, \delta) = \max\{|F(x_1) - F(x_2)|:\ x_1, x_2 \in$ $\in [a, b],\ |x_1 - x_2| \leqq \delta\}$ the modulus of continuity of $F \in C[a, b]$.) Combining (4) and (5) we can easily derive, that $\tilde{p}_s'(x_j) = 0$, i.e.

$$\operatorname{Re} a_j p_s'(x_j) = 0 \quad (1 \leqq s \leqq k+1)$$

if $x_j \in (a, b)$. This and (4) imply that for the extremal set $\{x_i\}_{i=1}^m$ with coefficients $\{a_i\}_{i=1}^m$ and $k+1$ linearly independent elements $p_1, \ldots, p_{k+1} \in U_n$ the condition of the theorem is violated, which proves its sufficiency.

Necessity. Assume that the condition of theorem does not hold, i.e. there exists an extremal set $\{x_i\}_{i=1}^m$ of U_n with coefficients $\{a_i\}_{i=1}^m$ and $k+1$ linearly independent functions $p_1, \ldots, p_{k+1} \in U_n$ such that $p_j(x_i) = 0$ $(1 \leqq i \leqq m,\ 1 \leqq j \leqq k+1)$ and

$\operatorname{Re} a_i p_j'(x_i)=0$ for any $1\leqq j\leqq k+1$ and $x_i\in(a, b)$. Without loss of generality we may assume that $\|p_j'\|_C\leqq 1$ for each $1\leqq j\leqq k+1$.

Let $0<h$ be small enough so that $[-h+x_i, x_i+h]\cap[-h+x_j, x_j+h]=\emptyset$ if $i\neq j$, and set $A_h=[a, b]\cap(\bigcup_{i=1}^{m}(-h+x_i, x_i+h))$. Evidently, there exists a function $g\in C^\infty[a, b]$ such that $\|g\|_C=1$ and $g(x)=\operatorname{sign} a_i$ for $x\in[-h+x_i, x_i+h]\cap[a, b]$ $(1\leqq i\leqq m)$. (This function can be chosen real if $a_i\in\mathbf{R}$.)

Consider at first the case $r=1$. Set

$$\varphi(\delta)=\delta+\sum_{j=1}^{k+1}\omega(p_j', \delta)\quad(0\leqq\delta\leqq b-a); \tag{6}$$

$$\psi_i(x)=\begin{cases}\int\limits_0^{|x-x_i|}\varphi(t)\,dt, & \text{if}\quad x_i\in(a, b)\\ |x-x_i|, & \text{if}\quad x_i=a\quad\text{or } b\end{cases}\qquad(1\leqq i\leqq m); \tag{7}$$

$$\psi(x)=\prod_{i=1}^{m}\psi_i(x)\quad(x\in[a, b]). \tag{8}$$

It is easy to see that $\psi_i\in C_0^1[a, b]$ $(1\leqq i\leqq m)$. Furthermore, we have by (6) and (7) that if $x_i\in(a, b)$ then for any $x\in[a, b]$

$$\begin{aligned}\psi_i(x)&=\int\limits_0^{|x-x_i|}\varphi(t)\,dt\geqq\frac{|x-x_i|}{2}\varphi\left(\frac{|x-x_i|}{2}\right)\geqq\\ &\geqq\frac{|x-x_i|}{2}\omega\left(p_j', \frac{|x-x_i|}{2}\right)\geqq\frac{|x-x_i|}{4}\omega(p_j', |x-x_i|)\quad(1\leqq j\leqq k+1).\end{aligned} \tag{9}$$

On the other hand, since $p_j(x_i)=0$ $(1\leqq i\leqq m,\ 1\leqq j\leqq k+1)$ and $\operatorname{Re} a_i p_j'(x_i)=0$ for any $1\leqq j\leqq k+1$ if $x_i\in(a, b)$ we have by (9) and (7) that

$$|\operatorname{Re} a_i p_j(x)|/|a_i|\leqq|x-x_i|\omega(p_j', |x-x_i|)\leqq 4\psi_i(x)$$

if $x_i\in(a, b)$ and

$$|\operatorname{Re} a_i p_j(x)|/|a_i|\leqq|x-x_i|\,\|p_j'\|_C\leqq\psi_i(x)$$

if $x_i=a$ or b $(x\in[a, b], 1\leqq j\leqq k+1)$. Thus for any $1\leqq i\leqq m, 1\leqq j\leqq k+1$ and $x\in[a, b]$

$$|\operatorname{Re} a_i p_j(x)|/|a_i|\leqq 4\psi_i(x). \tag{10}$$

Furthermore the function ψ/ψ_i is positive on $[a, b]\cap[-h+x_i, x_i+h]$ $(1\leqq i\leqq m)$, hence $\psi(x)/\psi_i(x)\geqq c_0>0$ for any $x\in[a, b]\cap[-h+x_i, x_i+h]$ and $1\leqq i\leqq m$. This and (10) imply that for each $x\in[a, b]\cap[-h+x_i, x_i+h]$

$$\psi(x)=\prod_{i=1}^{m}\psi_i(x)\geqq\frac{c_0|\operatorname{Re} a_i p_j(x)|}{4|a_i|}=\frac{K_1}{|a_i|}|\operatorname{Re} a_i p_j(x)| \tag{11}$$
$$(1\leqq i\leqq m,\quad 1\leqq j\leqq k+1).$$

In addition we can derive from (6) and (7) that

$$\psi_i(x) \geqq \min\{|x-x_i|, (x-x_i)^2/2\} \geqq (x-x_i)^2 \min\{1/(b-a), 1/2\} = c_1(x-x_i)^2.$$

Hence estimating as in (11) we have for $x\in[a,b]\cap[-h+x_i, x_i+h]$

$$\psi(x) \geqq c_0c_1(x-x_i)^2 = K_2(x-x_i)^2 \quad (1 \leqq i \leqq m). \tag{12}$$

Let us consider now the function

$$f(x) = g(x)\bigl(1-\lambda\psi(x)\bigr) \quad (x\in[a,b]), \tag{13}$$

where $\lambda=1/2\|\psi\|_C$. Obviously $f\in C^1[a,b]$ and $\|f\|_C=1$. Moreover $f(x_i)= =g(x_i)\bigl(1-\lambda\psi(x_i)\bigr)=g(x_i)=\operatorname{sign} a_i$ and $|f(x)|<1$ if $x\neq x_i$ $(1\leqq i\leqq m)$. Since $\{x_i\}_{i=1}^m$ is an extremal set of U_n with coefficients $\{a_i\}_{i=1}^m$ it follows from Lemma 1 that 0 is a best approximation of f. We state that $\varepsilon p_1, \ldots, \varepsilon p_{k+1}$ are also best approximants of f for $\varepsilon>0$ small enough. Using that $|f(x)|<1$ if $x\neq x_i$, and $x_i\in A_h$, $1\leqq i\leqq m$, we can find a constant $n>0$ such that $|f(x)|\leqq 1-\eta$ if $x\in[a,b]\setminus A_h$. Then if $0<\varepsilon\leqq\eta/M$, where $M=\max_{1\leqq j\leqq k+1}\|p_j\|_C$ we have for $x\in[a,b]\setminus A_h$

$$|f(x)-\varepsilon p_j(x)| \leqq 1-\eta+\varepsilon M \leqq 1 \quad (1 \leqq j \leqq k+1). \tag{14}$$

Assume now that $x\in A_h$, i.e. $x\in(-h+x_i, x_i+h)\cap[a,b]$ for some $1\leqq i\leqq m$. In this case $g(x)=\operatorname{sign} a_i$, hence and by (13)

$$\begin{aligned}|f(x)-\varepsilon p_j(x)|^2 &= |\operatorname{sign} a_i\bigl(1-\lambda\psi(x)\bigr)-\varepsilon p_j(x)|^2 = |1-\lambda\psi(x)-\varepsilon(a_i/|a_i|)\,p_j(x)|^2 = \\ &= \bigl(1-\lambda\psi(x)-(\varepsilon/|a_i|)\operatorname{Re} a_ip_j(x)\bigr)^2+\bigl((\varepsilon/|a_i|)\operatorname{Im} a_ip_j(x)\bigr)^2 \quad (1 \leqq j \leqq m).\end{aligned} \tag{15}$$

Since $p_j(x_i)=0$ $(1\leqq j\leqq k+1)$ it follows that

$$|\operatorname{Im} a_ip_j(x)|/|a_i| \leqq \|p_j'\|_C|x-x_i| \leqq |x-x_i| \quad (1 \leqq j \leqq m). \tag{16}$$

Assume now in addition that $\varepsilon<\lambda K_1/2$. Then (11) yields that for any $1\leqq j\leqq m$

$$0 \leqq 1-(3\lambda/2)\psi(x) \leqq 1-\lambda\psi(x)-(\varepsilon/|a_i|)\operatorname{Re} a_ip_j(x) \leqq 1-(\lambda/2)\psi(x). \tag{17}$$

Applying inequalities (17), (16) and (12) in (15) we have

$$\begin{aligned}|f(x)-\varepsilon p_j(x)|^2 &\leqq \bigl(1-(\lambda/2)\psi(x)\bigr)^2+\varepsilon^2(x-x_i)^2 \leqq 1-(\lambda/2)\psi(x)+\varepsilon^2(x-x_i)^2 \leqq \\ &\leqq 1-(\lambda K_2/2)(x-x_i)^2+\varepsilon^2(x-x_i)^2 \leqq 1 \quad (1 \leqq j \leqq k+1),\end{aligned}$$

if we assume also that $\varepsilon<\sqrt{\lambda K_2/2}$. Hence and by (14) we finally obtain that for ε small enough $\|f-\varepsilon p_j\|_C\leqq 1$ $(1\leqq j\leqq k+1)$. This means that $k+1$ linearly independent elements $p_1, \ldots, p_{k+1}\in U_n$ are also best approximants of f (in addition to 0), i.e. U_n is not k-Chebyshev.

If $r\geqq 2$ then setting $\varphi(\delta)=\delta$ and constructing f by (7), (8) and (13) we can analogously verify that U_n is not k-Chebyshev.

The proof of Theorem 1 is completed.

The corollary follows immediately from Theorem 1 since in the real case the coefficients of extremal sets and therefore the function (13) are real.

Let us now show that the characterization of k-Chebyshev subspaces of $C_0^r[a, b]$ given by Corollary 1 is equivalent to Garkavi's characterization.

Proposition 1. *Let U_n be a subspace of $C_0^1[a, b]$. Then for any $0 \leqq k \leqq n-1$ the following statements are equivalent:*

(i) *for any extremal set of U_n its points cannot be common special zeros of $k+1$ linearly independent elements of U_n;*

(ii) *for any s linearly independent elements $p_1, \ldots, p_s$ in U_n $(k+1 \leqq s \leqq n)$ among their common zeros there are at most $n-s$ special zeros common to any $k+1$ of the elements $p_1, \ldots, p_s$.*

Proof. (i)⇒(ii). Let $\{\varphi_i\}_{i=1}^n$ be a basis in U_n and consider the matrix

$$M = M(x_1, \ldots, x_{n-s+1}) = \begin{pmatrix} \varphi_1(x_1) & \ldots & \varphi_n(x_1) \\ \vdots & & \\ \varphi_1(x_{n-s+1}) & \ldots & \varphi_n(x_{n-s+1}) \end{pmatrix}$$

where $x_1, \ldots, x_{n-s+1}$ are arbitrary distinct points at $[a, b]$. If $x_1, \ldots, x_{n-s+1}$ are common zeros of s linearly independent elements in U_n then it follows that $\operatorname{rank} M \leqq n-s$. Therefore for some $b_i \in \mathbf{R}$, $\sum_{i=1}^{n-s+1} |b_i| = 1$, we have $\sum_{i=1}^{n-s+1} b_i \varphi_j(x_i) = 0 \quad (1 \leqq j \leqq n)$. This means that the set $\{x_i\}_{i=1}^{n-s+1}$ or a proper subset of it is an extremal set of U_n. Hence if s linearly independent elements in U_n have $n-s+1$ common zeros $x_1, \ldots, x_{n-s+1}$ then the set $\{x_i\}_{i=1}^{n-s+1}$ or a proper subset of it is an extremal set of U_n. This observation proves the implication (i)⇒(ii).

(ii)⇒(i). Assume that (i) is not true. Then there exists an extremal set $\{x_i\}_{i=1}^m$ and $k+1$ linearly independent elements $p_1, \ldots, p_{k+1} \in U_n$ such that each x_i is a special zero of p_j $(1 \leqq i \leqq m,\ 1 \leqq j \leqq k+1)$. Consider the matrix $M^* = M(x_1, \ldots, x_m)$. Then the set of functions in U_n vanishing on $\{x_i\}_{i=1}^m$ is a subspace of dimension $s = n - \operatorname{rank} M^*$. Evidently, $s \geqq k+1$. Then we can find elements $p_{k+2}, \ldots, p_s \in U_n$ such that $p_1, \ldots, p_s$ are linearly independent and $p_{k+2}, \ldots, p_s$ also vanish on $\{x_i\}_{i=1}^m$. It follows from (ii) that $m \leqq n-s = \operatorname{rank} M^*$, i.e. $\operatorname{rank} M^* = m$. But since $\{x_i\}_{i=1}^m$ is an extremal set of U_n the rows of M^* are linearly dependent. This implies that $\operatorname{rank} M^* \leqq m-1$, a contradiction.

2. In [2] there are given different examples of real polynomial spaces which are Chebyshev subspaces of $C_0^1[a, b]$ but do not satisfy this property with respect to $C_0[a, b]$. Let P_n denote the space of real algebraic polynomials of degree at most $n-1$. It is shown in [2] that if for the subspace U the embeddings $P_{[n/2]} \subset U \subset P_n$ hold then U is a Chebyshev subspace of $C_0^1[a, b]$. In this section applying Theorem 1 we shall give a similar statement for complex polynomials. (Since the characterization of Chebyshev subspaces of $C^r[a, b]$ does not depend on $r \geqq 1$, in what follows we shall consider only the case $r=1$.)

Let $T_n = \{\sum_{s=0}^{n-1} c_s e^{isx},\ c_s \in \mathbf{C}\}$ be the space of complex polynomials of degree at most $n-1$ and real variable $x \in [a, b]$, where $0 \leqq a < b < 2\pi$. Evidently, each extremal set of T_n consists of at least $n+1$ points (and at most $2n+1$ points by definition). In order to apply Theorem 1 we shall also need some information on the coefficients of extremal sets of T_n.

Lemma 2 (Vidensky [10]). *Let $\{x_j\}_{j=1}^m \subset [a, b]$ be an extremal set of T_n with coefficients $\{a_j\}_{j=1}^m$ $(n+1 \leqq m \leqq 2n+1)$. Then there exists $u \in T_{m-n-1}$ such that for any $j=1, 2, \ldots, m$*

$$a_j = u(x_j) / \prod_{\substack{s=1 \\ s \neq j}}^{m} (e^{ix_j} - e^{ix_s}). \tag{18}$$

Let now $0 = r_0 < r_1 < \ldots < r_{p-1}$ be a sequence of integers and set $U_p = \{\sum_{s=0}^{p-1} c_s e^{ir_s x},\ c_s \in \mathbf{C}\}$. Since $g(x) = 1 - e^{ir_{p-1}x} \in U_p$ may have r_{p-1} distinct zeros at $[a, b]$ it follows by the Haar—Kolmogorov theorem that U_p is in general a Chebyshev subspace of $C[a, b]$ only if $r_{p-1} = p-1$ and thus $U_p = T_p$. But for the space $C^1[a, b]$ we have a much more general statement.

Theorem 2. *Assume that $T_r \subset U_p \subset T_n$, where $r \leqq p \leqq n$ and $r = [2n/3]$ $(n \geqq 4)$. Then U_p is a Chebyshev subspace of $C^1[a, b]$ for any $0 \leqq a < b < 2\pi$.*

Proof. Assume that U_p is not a Chebyshev subspace of $C^1[a, b]$. Then by Theorem 1 there exists an extremal set $\{x_j\}_{j=1}^m$ of U_p with coefficients $\{a_j\}_{j=1}^m$ and $g \in U_p \setminus \{0\}$ such that $g(x_j) = 0$ $(1 \leqq j \leqq m)$ and $\operatorname{Re} a_j g'(x_j) = 0$ for each $x_j \in (a, b)$. Without loss of generality we may assume that $x_j \in (a, b)$ for every $2 \leqq j \leqq m-1$. Since $U_p \supset T_r$, $\{x_j\}_{j=1}^m$ is an extremal set of T_r, too. Hence $m \geqq r+1$. On the other hand, $g \in T_n \setminus \{0\}$ vanishes on x_j, $1 \leqq j \leqq m$. Thus $r+1 \leqq m \leqq n-1 \leqq 2r+1$. Therefore by Lemma 2 we can find a polynomial $u \in T_{m-r-1}$ such that for any $j = 1, 2, \ldots, m$ (18) holds. Furthermore, using that $g(x_j) = 0$ $(1 \leqq j \leqq m)$ we can write

$$g(x) = \prod_{j=1}^{m} (e^{ix} - e^{ix_j})\, \tilde{g}(x),$$

where $\tilde{g} \in T_{n-m}$. This yields that

$$g'(x_j) = i e^{ix_j} \prod_{\substack{s=1 \\ s \neq j}}^{m} (e^{ix_j} - e^{ix_s})\, \tilde{g}(x_j), \quad 1 \leqq j \leqq m. \tag{19}$$

Since $\{x_j\}_{j=2}^{m-1}\subset(a,b)$, $\operatorname{Re} a_j g'(x_j)=0$ for $2\leqq j\leqq m-1$. This together with (18) and (19) imply that for each $2\leqq j\leqq m-1$

$$0=\operatorname{Re} a_j g'(x_j)=\operatorname{Re} u(x_j)\, ie^{ix_j}\tilde{g}(x_j)=\operatorname{Re}\tilde{u}(x_j), \tag{20}$$

where $\tilde{u}(x)=ie^{ix}u(x)\tilde{g}(x)\in T_{n-r-1}$ and $\tilde{u}$ does not contain the constant term. Moreover (20) yields that $t(x)=\operatorname{Re}\tilde{u}(x)$ has $m-2$ distinct zeros at (a,b), where $m-2\geqq\geqq r-1$. On the other hand t is a trigonometric polynomial of degree at most $n-r-2$. Thus either t is identically zero or it has not more than $2n-2r-4$ distinct zeros at (a,b). But since $r=[2n/3]$ it follows that $2n-2r-4<3r+3-2r-4=r-1$. Hence $t(x)=\operatorname{Re}\tilde{u}(x)$ is the zero function. Using that $\tilde{u}\in T_{n-r-1}$ does not contain the constant term we finally obtain that $\tilde{u}$ is identically zero, a contradiction. The theorem is proved.

3. In this final section of our paper we shall solve some extremal problems connected with the unicity of best Chebyshev approximation of real differentiable functions by lacunary polynomials. Consider the space $C_0[-1,1]$. Then $P_n=\operatorname{span}\{1,x,\dots,x^{n-1}\}$ is a simple example of a Chebyshev subspace of $C_0[-1,1]$. Here and in what follows we denote by span $\{\dots\}$ the real linear span of functions specified in the brackets. Let us now omit the basis function x^k $(0<k<n-1)$ and consider the resulting space of lacunary polynomials $P_{n-1}^{(k)}=\operatorname{span}\{1,\dots,x^{k-1},x^{k+1},\dots,x^{n-1}\}$. The polynomials in $P_{n-1}^{(k)}$ may still have $n-1$ distinct zeros at $[-1,1]$, while the dimension of this space is only $n-1$. Thus $P_{n-1}^{(k)}$ is not a Chebyshev subspace of $C_0[-1,1]$. On the other hand it was shown in [2] that $P_{n-1}^{(k)}$ is a Chebyshev subspace of $C_0^1[-1,1]$, if $n\geqq 4$. Analogously, if we add to P_n an arbitrary power function x^r $(r\in\mathbf{N}, r\geqq n+1)$ then the resulting space $\bar{P}_{n+1}^{(r)}=\operatorname{span}\{1,x,\dots,x^{n-1},x^r\}$ is Chebyshev in $C_0[-1,1]$ only if $r-n$ is even but nevertheless it is Chebyshev in $C_0^1[-1,1]$ for any r (see [2]). Thus deleting from P_n or adding to P_n a power function we in general violate the Haar property and hence obtain nonuniqueness of best Chebyshev approximation in $C_0[-1,1]$. On the other hand the unicity with restriction to the space $C_0^1[-1,1]$ still holds. This observation raises the following questions:

A) Determine the maximal integer $\gamma=\gamma(n)$ such that omitting from P_n arbitrary γ basis functions $x^{r_1},\dots,x^{r_\gamma}$ $(1\leqq r_1<\dots<r_\gamma\leqq n-2,\ r_i\in\mathbf{N})$ the resulting set of lacunary polynomials $P_{n-\gamma}^{*}=\operatorname{span}\{x^i,\ 0\leqq i\leqq n-1,\ i\neq r_j,\ 1\leqq j\leqq\gamma\}$ is still a Chebyshev subspace of $C_0^1[-1,1]$.

B) Determine the maximal integer $\mu=\mu(n)$ such that adding to P_n arbitrary μ powers $x^{t_1},\dots,x^{t_\mu}$ $(n+1\leqq t_1<\dots<t_\mu,\ t_i\in\mathbf{N})$ the resulting set of lacunary polynomials $P'_{n+\mu}=\operatorname{span}\{1,x,\dots,x^{n-1},x^{t_1},\dots,x^{t_\mu}\}$ is still a Chebyshev subspace of $C_0^1[-1,1]$.

We shall verify in this section that $\gamma(n)=[n/4]$ and $\mu(n)=[n/2]$. Thus omitting (or adding) from P_n a considerable number of power functions we can still guarantee the unicity of best Chebyshev approximation in $C_0^1[-1,1]$.

In what follows the finite-dimensional Chebyshev subspaces of $C_0[-1, 1]$ will be called Haar spaces.

We shall need the following simple lemma.

Lemma 3. *Let $r\in\mathbf{N}$ and $0=m_0<m_1<\ldots<m_r$ be a sequence of integers such that m_j-m_{j-1} is odd for each $j=1, 2, \ldots, r$. Then the space $P^*_{r+1}=\operatorname{span}\{1=x^{m_0}, x^{m_1}, \ldots, x^{m_r}\}$ is a Haar space.*

Proof. We shall prove the lemma by induction. For $r=1$ the statement is evident. Assume that it holds for $r-1$ $(r\geqq 2)$. For any $p\in P^*_{r+1}$ which is not a constant function $p'(x)=x^{m_1-1}\tilde{p}(x)$, where $\tilde{p}\in\tilde{P}_r=\operatorname{span}\{1, x^{m_2-m_1}, \ldots, x^{m_r-m_1}\}$. By our assumption $\tilde{P}_r$ is an r-dimensional Haar space, hence $\tilde{p}$ has at most $r-1$ distinct zeros at $[-1, 1]$. Moreover, m_1-1 is even, therefore p' has at most $r-1$ points of change of sign at $[-1, 1]$. This yields that p has not more than r distinct zeros at $[-1, 1]$. The lemma is proved.

By the well-known interpolatory property of Haar spaces it follows that each extremal set of an n-dimensional Haar space consists of exactly $n+1$ points on $[-1, 1]$. In particular if U_n contains a k-dimensional Haar subspace $(k\leqq n)$ then each extremal set of U_n consists of at least $k+1$ points $(U_n\subset C_0[-1, 1])$. We shall frequently use this simple observation.

Theorem 3. *For any $n\geqq 4$, $\gamma(n)=[n/4]$.*

Proof. Let us prove at first that $\gamma(n)\geqq[n/4]$. Set $m=[n/4]$ and let $1\leqq r_1<\ldots<r_m\leqq n-2$ be arbitrary integers. Omitting from P_n the basis functions x^{r_i} $(1\leqq i\leqq m)$ we obtain the space $P^*_{n-m}=\operatorname{span}\{x^{t_0}, x^{t_1}, \ldots, x^{t_{n-m-1}}\}$, where $0=t_0<t_1<\ldots<t_{n-m-1}=n-1$ and $t_i\neq r_j$ for every $0\leqq i\leqq n-m-1$, $1\leqq j\leqq m$. Set $c_j=t_j-t_{j-1}$, $1\leqq j\leqq n-m-1$. Evidently, at most m of these $n-m-1$ integers are even. Deleting from the sequence $0=t_0<t_1<\ldots<t_{n-m-1}=n-1$ those integers t_j for which c_j is even we obtain a sequence $0=t'_0<t'_1<\ldots<t'_s\leqq n-1$, where $s\geqq n-2m-1$. Let us prove that for any $1\leqq j\leqq s$, $t'_j-t'_{j-1}$ is odd. Indeed, we have for some $q<r$, that $t'_{j-1}=t_q<t_{q+1}<\ldots<t_r=t'_j$, where c_i is even for every $q+1\leqq i\leqq r-1$, while c_r is odd. Therefore $t'_j-t'_{j-1}=t_r-t_q=\sum_{i=q+1}^{r} c_i$ is odd. Applying Lemma 3 we can conclude that $\operatorname{span}\{x^{t'_0}, \ldots, x^{t'_s}\}$ is a Haar space. Thus P^*_{n-m} contains a Haar space of dimension $s+1\geqq n-2m$. Therefore each extremal set of P^*_{n-m} consists of at least $n-2m+1$ points. If the points of an extremal set of P^*_{n-m} are special zeros of $g\in P^*_{n-m}$ then g has at least $(n-2m+1)+(n-2m-1)=2n-4m\geqq n$ zeros counting double zeros twice. Since $g\in P_n$ it follows that g is identically zero. Thus we obtain by Corollary 1 that P^*_{n-m} is a Chebyshev subspace of $C^1_0[-1, 1]$, i.e. $\gamma(n)\geqq m=[n/4]$.

Now we shall verify that $\gamma(n)\leqq m=[n/4]$. We have $n=4m+i$ $(i=0, 1, 2, 3)$. Assume that in contrary $\gamma(n)\geqq m+1$, i.e. omitting from P_n arbitrary $m+1$ basis functions we still have a Chebyshev subspace of $C_0^1[-1, 1]$. Set

$$P_{n-m-1}^* = \operatorname{span}\{1, x, \ldots, x^{2m+i-3}, x^{2m+i-1}, \ldots, x^{4m+i-1}\}.$$

P_{n-m-1}^* is generated from $P_{4m+i}=P_n$ by deleting $m+1$ powers $x^{2m+i-2+2s}$, $0\leqq s\leqq m$. Thus by our assumption P_{n-m-1}^* is a Chebyshev subspace of $C_0^1[-1, 1]$. Consider the function $f(x)=x^{2m+i-2}$ $(x\in[-1, 1])$.

Case 1: $i=1$ or 3. Then f is odd. Since f possesses a unique best approximation q in P_{n-m-1}^*, q is odd, too. But the powers $x^{2m+i-1+2s}$ $(0\leqq s\leqq m)$ are even, hence $q\in P_{2m+i-3}$. Therefore $f-q=t_{2m+i-2}$, where $t_k(x)=2^{-k+1}\cos k \arccos x$ denotes the Chebyshev polynomial of degree k. Consider the extremas of t_{2m+i-2}: $x_j= =\cos j\pi/(2m+i-2)$ $(0\leqq j\leqq 2m+i-2)$. Since q is the best approximation of f in P_{n-m-1}^* it follows from Lemma 1 that the set $\{x_j, 0\leqq j\leqq 2m+i-2\}$ or a proper subset of it is an extremal set of P_{n-m-1}^*. On the other hand $P_{n-m-1}^*\supset P_{2m+i-2}$, hence each extremal set of P_{n-m-1}^* contains at least $2m+i-1$ points. Thus the set $\{x_j, 0\leqq j\leqq 2m+i-2\}$ is an extremal set of P_{n-m-1}^*. Consider now the polynomial

$$\tilde{p}(x) = (1-x^2)\prod_{i=1}^{2m+i-3}(x-x_i)^2. \tag{21}$$

Evidently, each x_j is a special zero of $\tilde{p}$ $(0\leqq j\leqq 2m+i-2)$ and $\deg\tilde{p}=4m+2i-4\leqq \leqq 4m+i-1$. Furthermore, since $x_j=-x_{2m+i-2-j}$ $(0\leqq j\leqq 2m+i-2)$ it follows that $\tilde{p}$ is even. Thus finally we obtain that $\tilde{p}\in P_{n-m-1}^*$, which contradicts Corollary 1.

Case 2: $i=0$ or 2. In this case instead of polynomial $\tilde{p}$ given by (21) we should consider the polynomial $p^*(x)=x\tilde{p}(x)$. Then we can derive a contradiction analogously to Case 1, the details are left to the reader.

Thus the assumption $\gamma(n)\geqq m+1$ leads to a contradiction. This completes the proof of the equality $\gamma(n)=[n/4]$.

Theorem 4. *For any* $n\geqq 2$, $\mu(n)=[n/2]$.

Proof. Let us verify that $\mu(n)\geqq m=[n/2]$. Take arbitrary integers $n+1\leqq \leqq t_1<t_2<\ldots<t_m$ and consider the space $P'_{n+m}=\operatorname{span}\{1, x, \ldots, x^{n-1}, x^{t_1}, \ldots, x^{t_m}\}$. Obviously, each extremal set of P'_{n+m} consists of at least $n+1$ points. We state that P'_{n+m} is a Chebyshev subspace of $C_0^1[-1, 1]$. Assume the contrary. Then for some extremal set of P'_{n+m} and some $g\in P'_{n+m}\setminus\{0\}$ the points of the extremal set are special zeros of g, hence g' has at least $2n-1$ distinct zeros at $[-1, 1]$. Furthermore $g'\in\operatorname{span}\{1, x, \ldots, x^{n-2}, x^{t_1-1}, \ldots, x^{t_m-1}\}\setminus\{0\}$. By Lemma 3 the space $\operatorname{span}\{1, x, \ldots, x^{n-2}, x^{t_1-1}, \ldots, x^{t_m-1}\}$ can be imbedded to a Haar space of dimension at most $n+2m-1\leqq 2n-1$. This means that each element of this space, in particular g', may have at most $2n-2$ distinct zeros at $[-1, 1]$, a contradiction. By this contradic-

tion we obtain that P'_{n+m} is a Chebyshev subspace of $C_0^1[-1,1]$, i.e. $\mu(n)\geqq m=[n/2]$.

Assume now that $\mu(n)\geqq m+1$. Set $n=2m+i$ $(i=0,1)$,

$$P'_{n+m+1}=\operatorname{span}\{1, x, \ldots, x^{2m+i-1}, x^{2m+i+1}, \ldots, x^{4m+i+1}\}.$$

Then P'_{n+m+1} is generated from $P_{2m+i}=P_n$ by adding $m+1$ basis functions $x^{2m+i+1+2s}$ $(0\leqq s\leqq m)$. Since $\mu(n)\geqq m+1$, it follows that P'_{n+m+1} is a Chebyshev subspace of $C_0^1[-1,1]$. Now we can derive a contradiction analogously to the proof of Theorem 3. We omit the details.

This completes the proof of Theorem 4.

Consider now the general case of lacunary polynomials. Let $0=m_0<m_1<\ldots<m_r$ be arbitrary integers and set

$$\bar{P}=\bar{P}_{r+1}=\operatorname{span}\{1=x^{m_0}, x^{m_1}, \ldots, x^{m_r}\} \quad (r\in\mathbf{N}). \tag{22}$$

Furthemore denote by $\delta(\bar{P})$ the number of those j-s for which m_j-m_{j-1} is even, $1\leqq j\leqq r$. Then $0\leqq\delta(\bar{P})\leqq r=\dim(\bar{P})-1$. By Lemma 3 if $\delta(\bar{P})=0$ then $\bar{P}$ is a Haar space on $[-1,1]$. It can be easily shown that this condition is also necessary for the Haar property. The next theorem gives a sufficient condition for $\bar{P}$ to be a Chebyshev subspace of $C_0^1[-1,1]$.

Theorem 5. *Let* $\dim(\bar{P})\geqq 4$ *and assume that*

$$\delta(\bar{P})\leqq[(\dim(\bar{P})-1)/3] \tag{23}$$

holds. Then $\bar{P}$ *is a Chebyshev subspace of* $C_0^1[-1,1]$.

Proof. Consider the space $\bar{P}^*$ which results from $\bar{P}$ after deleting in (22) all basis functions x^{m_j} such that m_j-m_{j-1} is even. Obviously, $\bar{P}^*$ is a space of dimension $\dim(\bar{P})-\delta(\bar{P})$. Moreover, similarly as in the proof of Theorem 3 we can show that $\delta(\bar{P}^*)=0$, thus by Lemma 3 $\bar{P}^*$ is a Haar space. Therefore each extremal set of $\bar{P}$ consists of at least $\dim(\bar{P}^*)+1=\dim(\bar{P})-\delta(\bar{P})+1$ points. Assume that (23) holds but $\bar{P}$ is not a Chebyshev subspace of $C_0^1[-1,1]$. Then there exists a $p\in\bar{P}\setminus\{0\}$ such that the set of special zeros of p contains at least $\dim(\bar{P})-\delta(\bar{P})+1$ points. This means that p' has at least $2\dim(\bar{P})-2\delta(\bar{P})-1$ distinct zeros at $[-1,1]$. Furthermore, $p'=x^{m_1-1}g$, where $g\in\operatorname{span}\{1, x^{m_2-m_1}, \ldots, x^{m_r-m_1}\}=\tilde{P}^*$ and g is not identically zero. It is evident, that $\delta(\tilde{P}^*)\leqq\delta(\bar{P})$. Hence adding to $\tilde{P}^*$ at most $\delta(\bar{P})$ power functions we can obtain (by Lemma 3) a Haar space. This means that $\tilde{P}^*$ can be enbedded to a Haar space of dimension at most $\dim(\bar{P})+\delta(\bar{P})-1$. Hence $g\in\tilde{P}^*\setminus\{0\}$ can have not more than $\dim(\bar{P})+\delta(\bar{P})-2$ zeros, i.e. p' has at most $\dim(\bar{P})+\delta(\bar{P})-1$ distinct zeros at $[-1,1]$. Since we have shown that p' has at least $2\dim(\bar{P})-2\delta(\bar{P})-1$ distinct zeros, it follows that $2\dim(\bar{P})-2\delta(\bar{P})-1\leqq\dim(\bar{P})+\delta(\bar{P})-1$, i.e. $\dim(\bar{P})\leqq 3\delta(\bar{P})$. But this contradicts (23). The theorem is proved.

Remark. The converse of Theorem 5 is not true in general. There exist Chebyshev subspaces of $C_0^1[-1, 1]$ of the form (22) such that (23) does not hold. Indeed, let $n=2k$ and add to P_n k odd powers greater than $n-1$. Then for the resulting space $\bar{P}$ the realtion $\delta(\bar{P})=k$ holds. By Theorem 4 $\bar{P}$ is a Chebyshev subspace of $C_0^1[-1, 1]$. On the other hand $\delta(\bar{P})=k>k-1=[(\dim(\bar{P})-1)/3]$.

References

[1] E. W. Cheney, D. E. Wulbert, The existence and unicity of best approximations, *Math. Scand.*, **24** (1969), 118—140.

[2] A. A. Гаркави, О размерности многогранника наилучшего приближения дифференцируемых функций, *Изв. Акад. Наук СССР, сер. матем.*, **23** (1959), 93—114.

[3] A. Haar, Die Minkowskische Geometrie und die Annäherung an stetige Funktionen, *Math. Ann.*, **78** (1918), 294—311.

[4] A. H. Колмогоров, Замечание по поводу многочленов П. Л. Чебышева, наименее уклоняющихся от заданной функции, *Успехи Мат. Наук*, **3**, 1 (23) (1948), 216—221.

[5] R. R. Phelps, Cebysev subspaces of finite codimension in $C(X)$, *Pacific J. Math.*, **13** (1963), 647—655.

[6] E. Я. Ремез, Некоторые вопросы чебышевского приближения в комплексной области, *Укр. Матем. Ж.*, **5** (1953), 3—49.

[7] З. С. Романова, О размерности многогранника наилучших приближений в пространстве непрерывных функций, *Литовский Мат. Сб.*, **2** (1963), 181—191.

[8] Л. Ш. Рубинштейн, Об одном методе исследования выпуклых множеств, *Доклады Акад. Наук СССР*, **103** (1955), 451—454.

[9] I. Singer, *Best approximation in normed linear spaces by elements of linear subspaces*, Springer (Berlin—Heidelberg—New York, 1970).

[10] B. C. Виденский, О равномерном приближении в комплексной области, *Успехи Мат. Наук*, **11** (1956), 169—175.

[11] B. Brosowski, J. Stoer, Zur rationalen Tschebyscheff-Approximation differenzierbarer und analytischer Funktionen, *Numer. Math.*, **12** (1968), 57—65.

INSTITUTE OF MATHEMATICS OF THE
HUNGARIAN ACADEMY OF SCIENCES
REÁLTANODA U. 13—15.
1053 BUDAPEST, HUNGARY

Acta Sci. Math. 47 (1984), 391—403

Lacunarity with respect to orthogonal polynomial sequences

RUPERT LASSER

Lacunarity has been studied in a variety of settings: on the circle group T with dual $\mathbf{Z}$, on compact abelian groups G with dual $\hat{G}$, on compact (nonabelian) groups resp. on the space of conjugacy classes of compact groups and on compact hypergroups with dual Σ. For references we recommend [11]. In view of the classical case T and $\mathbf{Z}$ a most natural setting to study lacunarity are orthogonal polynomial sequences. In fact to many orthogonal polynomial sequences there corresponds a hypergroup structure on $\mathbf{N}_0=\mathbf{N}\cup\{0\}$ having as dual a compact subset D_S of $\mathbf{R}$, see [8]. In this way a set $E\subseteq\mathbf{N}_0$ is a Sidon set if each bounded sequence can be represented on E as a (generalized) Fourier—Stieltjes transform. We emphasize that D_S is in general not a hypergroup under pointwise operations. Thus only $\mathbf{N}_0$ bears an algebraic structure in constrast to the situations above.

Combining two recent results, Theorem 3.2 of [14] and Chapter 4, ad(a) of [8], we can deduce that with respect to Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$, where $\alpha\geqq\beta>-1$ and in addition $\beta\geqq-1/2$ or $\alpha+\beta\geqq0$, but $\alpha\neq-1/2$, a set E is a Sidon set if and only if E is finite. This result suggests to perform further investigations on the subject.

We assume that the polynomial sequences satisfy a certain positivity property. This property and its consequences are presented in Section I. Sidonicity is the subject of II. In III there is shown that $\mathbf{N}_0$ is never a Sidon set. The fact that some orthogonal polynomial sequences admit only finite Sidon sets is established in IV. The existence of infinite Sidon sets is studied in section V.

Received November 2, 1982

I. Property (P)

At first we have to set up some notation. Let $(a_n)_{n\in\mathbf{N}}$, $(b_n)_{n\in\mathbf{N}}$, $(c_n)_{n\in\mathbf{N}}$ be three real-valued sequences such that $a_n>0$, $c_n>0$, $b_n\geqq 0$ and $a_n+b_n+c_n=1$. Further fixing $a_0>0$, $b_0\in\mathbf{R}$ such that $a_0+b_0=1$ define

$$P_0(x)=1,\quad P_1(x)=\frac{1}{a_0}x-\frac{b_0}{a_0},$$

$$P_{n+1}(x)=\frac{1}{a_n}P_1(x)P_n(x)-\frac{b_n}{a_n}P_n(x)-\frac{c_n}{a_n}P_{n-1}(x),\quad n\in\mathbf{N}.$$

Then $(P_n(x))$ is an orthogonal polynomial sequence. Write the linearization of the products $P_m(x)P_n(x)$, $1\leqq m\leqq n$, by

$$P_m(x)P_n(x)=\sum_{k=0}^{2m} g(m,n,n+m-k)P_{n+m-k}(x).$$

The coefficients $g(m,n,n+m-k)$ are uniquely determined by the sequences $(a_n)_{n\in\mathbf{N}}$, $(b_n)_{n\in\mathbf{N}}$, $(c_n)_{n\in\mathbf{N}}$. We require throughout this paper that the positivity property

$$\text{(P)}\qquad g(m,n,n+m-k)\geqq 0$$

is satisfied.

This assumption yields that $(P_n(x))$ is closely related to a commutative hypergroup structure on $\mathbf{N}_0=\mathbf{N}\cup\{0\}$. The convolution on $\mathbf{N}_0$ is defined by

$$p_m * p_n=\sum_{k=0}^{2m} g(m,n,n+m-k)p_{n+m-k},\quad 1\leqq m\leqq n,$$

where p_n is the point measure of $n\in\mathbf{N}_0$. The involution is the identity on $\mathbf{N}_0$ and the zero is the unit element. Each character on $\mathbf{N}_0$ is given by α_x: $\mathbf{N}_0\to\mathbf{R}$, where $x\in D_s$,

$$D_S=\{x\in\mathbf{R}\colon (P_n(x))_{n\in\mathbf{N}} \text{ is bounded}\}\quad\text{and}\quad \alpha_x(n)=P_n(x).$$

Further the character space $\hat{\mathbf{N}}_0$ is homeomorphic to D_S. For details we refer to [8]. Many prominent examples of $(P_n(x))$ satisfying property (P) can be found in [8], [9], [10].

The Haar measure h on the hypergroup $\mathbf{N}_0$ is given by

$$h(0)=1,\quad h(1)=\frac{1}{c_1},\quad h(n)=\prod_{k=1}^{n-1}a_k\Big/\prod_{k=1}^{n}c_k,\quad n=2,3,\ldots.$$

The Plancherel measure π on D_S is the orthogonalization measure of $(P_n(x))$:

$$\int_{D_S} P_n(x)P_m(x)\,d\pi(x)=\begin{cases}1/h(n) & \text{if } n=m\\ 0 & \text{if } n\neq m.\end{cases}$$

We have $\operatorname{supp}\pi\subseteqq D_S\subseteqq[1-2a_0,1]$.

For an absolutely convergent function $f \in l^1(\mathbf{N}_0) = l^1(\mathbf{N}_0, h)$ define the Fourier transform $\hat{f}$ on D_S by

$$\hat{f}(x) = \sum_{n \in \mathbf{N}_0} f(n) P_n(x) h(n).$$

For a Radon measure $\mu \in M(D_S)$ denote the inverse Fourier—Stieltjes transform $\check{\mu}$ on $\mathbf{N}_0$ by

$$\check{\mu}(n) = \int_{D_S} P_n(x) d\mu(x).$$

We shall say that D_S is a hypergroup with respect to pointwise multiplication, if for $x, y \in D_S$ there exists a probability measure $p_x * p_y \in M(D_S)$ such that

(i) $\quad P_n(x) P_n(y) = \int_{D_S} P_n(z) d p_x * p_y(z)$ for each $n \in \mathbf{N}_0$, and

(ii) D_S is a hypergroup with this convolution, the identity as involution and $1 \in D_S$ as unit;

compare Chapters 1, 4 of [8]. We recall that D_S is in general not a hypergroup with respect to pointwise multiplication.

II. Sidon sets

We assume throughout this paper that $(P_n(x))$ is an orthogonal polynomial sequence defined by (a_n), (b_n), (c_n) satisfying property (P). We abbreviate $S = \operatorname{supp} \pi$ and interpret $M(S)$ as a subspace of $M(D_S)$ and $L^1(D_S) = L^1(D_S, \pi)$ as a subspace of $M(S)$. We write $\|f\|_S = \sup \{|f(x)|: x \in S\}$ for a function $f \in C(D_S)$. Let E be a subset of $\mathbf{N}_0$. As usual $l^\infty(E)$ denotes the space of all bounded functions on E, $c_0(E)$ the space of all functions on E vanishing at infinity, $M(E)$ the space of all bounded measures on E and $\mathrm{Trig}_E(D_S)$ the linear span of $\{P_n(x): n \in E\}$. We shall call E a *Sidon set* if $l^\infty(E) = M(S)^\vee | E$.

Proposition 1. *Let* $E \subseteq \mathbf{N}_0$. *The following are equivalent:*

(a) *E is a Sidon set.*

(b) $L^1(D_S)^\vee | E = c_0(E)$.

(c) *$M(E)^\wedge | S$ is sup-norm closed in $C(S)$.*

(d) *There exists a constant $B > 0$ such that $\|\mu\| \leqq B \|\hat{\mu}\|_S$ for each $\mu \in M(E)$.*

(e) *There exists a constant $B > 0$ such that $\|\check{f}\|_1 \leqq B \|f\|_S$ for each $f \in \mathrm{Trig}_E(D_S)$.*

(f) *Given $\varphi: E \to \{-1, 1\}$, there exists some $\mu \in M(S)$ such that* $\sup \{|\check{\mu}(n) - \varphi(n)|: n \in E\} < 1$.

Proof. Since the set of measures having finite support in E is norm-dense in $M(E)$, property (d) is equivalent to (e). Using Proposition 1 of [10] define the operator

$A\colon L^1(D_S)\to c_0(E)$, $A(g)=\check{g}|E$, $g\in L^1(D_S)$. The adjoint operator $A^*\colon M(E)\to$ $\to L^\infty(D_S)$ satisfies for $g\in C(D_S)$ and $\mu\in M(E)$:

$$\int_{D_S} A^*(\mu)(x)\cdot g(x)\,d\pi(x) = \sum_{n\in E} A(g)(n)\mu(n) = \int_{D_S} \hat{\mu}(x)\cdot g(x)\,d\pi(x).$$

Thus $A^*(\mu)|S=\hat{\mu}|S$. By Lemma 12.2B of [6] the operator A^* is injective. Now Theorem (E.9) of [5] yields the equivalence of (b), (c) and (d). Define $B\colon M(E)\to C(S)$, $B(\mu)=\hat{\mu}|S$. For the adjoint operator $B^*\colon M(S)\to l^\infty(E)$ we deduce that $B^*(\nu)=$ $=\check{\nu}|E$ for $\nu\in M(S)$. The injectivity of B, Corollary (E.8) and Theorem (E.10) of [5] imply the equivalence of (a) and (c). There remains to prove that (f) implies (d). First deduce that in (d) it is sufficient to consider only real-valued measures $\mu\in M^{\mathbf{R}}(E)$ having finite support. Now assume that (d) does not hold. Then there exists for each $n\in\mathbf{N}$ a measure $\lambda_n\in M^{\mathbf{R}}(E)$ such that $\|\lambda_n\|=1$, $\|\hat{\lambda}_n\|_S<1/n$, and the sets $F_n=\operatorname{supp}\lambda_n$ are finite and pairwise disjoint. In fact having already chosen appropriate $\lambda_1, \dots, \lambda_m\in$ $\in M^{\mathbf{R}}(E)$ observe that $E'=E\setminus\bigcup_{k=1}^{m} F_k$ is not a Sidon set, too. Hence there exists a measure $\lambda_{m+1}\in M^{\mathbf{R}}(E')\subseteq M^{\mathbf{R}}(E)$ such that $\|\lambda_{m+1}\|=1$, $\|\lambda_{m+1}\|_S<1/(m+1)$ and $F_{m+1}=\operatorname{supp}\lambda_{m+1}$ finite. Define $\varphi\colon E\to\{-1,1\}$ by $\varphi(k)\lambda_n(K)=|\lambda_n(k)|$ for $k\in F_n$ and $\varphi(k)=1$ for k elsewhere. By (f) there exist $\mu\in M(S)$ and $\delta>0$ such that $|\check{\mu}(k)-\varphi(k)|\leqq 1-\delta$ for each $k\in E$. We may assume that $\check{\mu}(k)\in\mathbf{R}$. One obtains that

$$\big|\check{\mu}\lambda_n(k)-|\lambda_n(k)|\big| = |\check{\mu}(k)-\varphi(k)|\,|\lambda_n(k)| \leqq (1-\delta)\,|\lambda_n(k)|$$

and then $0\leqq\delta|\lambda_n(k)|\leqq\check{\mu}(k)\lambda_n(k)$ for each $k\in E$, $n\in\mathbf{N}$. Hence for each $n\in\mathbf{N}$ we have

$$\int_S \hat{\lambda}_n(x)\,d\mu(x) = \sum_{k\in F_n}\lambda_n(k)\check{\mu}(k) \geqq \delta\sum_{k\in F_n}|\lambda_n(k)| = \delta.$$

This is in contradiction to

$$\Big|\int_S \hat{\lambda}_n(x)\,d\mu(x)\Big| \leqq \int_S |\hat{\lambda}_n(x)|\,d|\mu|(x) \leqq \frac{|\mu|(S)}{n},$$

and we have shown that (f) implies (d).

Remark. There holds an appropriate version of Proposition 1 for any discrete hypergroup K.

If $f\in C(D_S)$ satisfies $\check{f}(n)=0$ for each $n\notin E$ we write $f\in C_E(D_S)$. Comparing with the group case, see e.g. Theorem 1.3 of [11], one might notice the failure of the following property $(*)$ in the above list of equivalences

$$(*)\qquad C_E(D_S)^{\vee}\subseteq l^1(\mathbf{N}_0).$$

We know the following partial results:

Proposition 2. *Let* $E \subseteqq \mathbf{N}_0$.

(a) *If E satisfies property* $(*)$ *then E is a Sidon set.*

(b) *If* D_S *is a hypergroup with respect to pointwise multiplication, then E is a Sidon set if and only if E fulfils property* $(*)$.

Proof. (a) Using $(*)$ the map $f|S \to \check{f}$, $C_E(D_S)|S \to l^1(E)$ is an isomorphism such that $\|f\|_S \leqq \|\check{f}\|_1$. By the open mapping theorem there exists a constant $B>0$ such that $\|\check{f}\|_1 \leqq B\|f\|_S$ for each $f \in C_E(D_S)$. In particular condition (e) of Proposition 1 is valid.

(b) We refer to Theorem 2.2 of [13]. Note that $\hat{D}_S = \mathbf{N}_0$, see Proposition 2 of [8].

For $n \in \mathbf{N}$ and $f \in C(D_S)$ denote $S_n(f)(x) = \sum_{k=0}^{n} \check{f}(k) P_k(x) h(k)$. Further for $E \subseteqq \mathbf{N}_0$ let

$$U_E(D_S) = \{f \in C_E(D_S) \colon S_n f \to f \text{ uniformly on } S\}.$$

Proposition 3. *Let* $E \subseteqq \mathbf{N}_0$. *The following are equivalent:*

(a) *E is a Sidon set.*

(b) $U_E(D_S)^{\check{}} \subseteqq l^1(\mathbf{N}_0)$.

Proof. At first assume that E is a Sidon set. Let $f \in U_E(D_S)$. Since $S_n f \in M(E)^{\hat{}}$ we have that $f|S$ is an element of the uniform closure of $M(E)^{\hat{}}|S$. By condition (c) of Proposition 1 it follows that $f|S \in M(E)^{\hat{}}|S$. Hence $\check{f} \in l^1(\mathbf{N}_0)$.

Now assume that (b) is valid and E is not a Sidon set. Write $E = \{n_1, n_2, \ldots\}$. Let $N_0 = 0$. For $j \in \mathbf{N}$ there exist $N_j \in \mathbf{N}$, $\lambda_j \in M(E)$ such that $\lambda_j = \sum_{k=N_{j-1}+1}^{N} c_k p_{n_k}$, $\|\lambda_j\| = 1/j$, $\|\hat{\lambda}_j\|_S \leqq 1/2^j$. Define $g(x) = \sum_{j=1}^{\infty} \hat{\lambda}_j(x)$ for $x \in S$, and let f be a continuous extension of g to D_S. Then $\check{f}(n) = 0$ for $n \notin E$ and $\check{f}(n_k) = c_k / h(n_k)$. Hence $S_{N_j}(f) \xrightarrow{j} f$ uniformly on S. For $N_j < n \leqq N_{j+1}$ we obtain

$$\|S_n(f) - S_{N_j}(f)\|_S \leqq \Big\| \sum_{k=N_j+1}^{n} \check{f}(k) P_k h(k) \Big\|_S \leqq \sum_{k=N_j+1}^{n} |c_k| \leqq 1/(j+1).$$

Thus $S_n(f) \xrightarrow{n} f$ uniformly on S, i.e. $f \in U_E(D_S)$. But $\sum_{n=0}^{\infty} |\check{f}(n)| h(n) = \infty$, a contradiction.

III. $\mathbf{N}_0$ is not a Sidon set

If we assume that D_S is a hypergroup with respect to pointwise multiplication, Theorem 2.11 of [13] or Theorem 2.5 of [10] yield that $\mathbf{N}_0$ is not a Sidon set. We shall show that this is true without any assumption on D_S. Our proof is motivated by [3]. In $l^\infty(\mathbf{N}_0)^*$ let τ denote the weak-$*$ topology. Let j be the canonical embedding of

$l^1(\mathbf{N}_0)$ into $l^\infty(N_0)^*$. The set

$$\mathcal{M} = \{\varphi \in l^\infty(\mathbf{N}_0)^*\colon \varphi(f) \geqq 0 \quad \text{for} \quad f \geqq 0,\ \varphi(1) = 1\}$$

is convex and τ-compact. $M^1(\mathbf{N}_0) = \{g \in l^1(\mathbf{N}_0)\colon g(n) \geqq 0,\ \|g\|_1 = 1\}$ acts as a commutative semigroup of τ-continuous operators of $\mathcal{M}$ in $\mathcal{M}$, where

$$g * \varphi(f) = \varphi(g * f), \quad f \in l^\infty(\mathbf{N}_0), \quad g \in M^1(\mathbf{N}_0), \quad \varphi \in \mathcal{M}.$$

The Markov—Kakutani fixed point theorem yields $\psi \in \mathcal{M}$ such that $g * \psi = \psi$, i.e. $\psi(g * f) = \psi(f)$ for each $g \in M^1(\mathbf{N}_0)$, $f \in l^\infty(\mathbf{N}_0)$. Using the notation of means we have shown that there exists a mean on $l^\infty(\mathbf{N}_0)$ which is invariant under $f \mapsto g * f$, $g \in M^1(\mathbf{N}_0)$.

Lemma 1. *There exists a sequence (g_k), $g_k \in M^1(\mathbf{N}_0)$, such that $\hat{g}_k(x) \xrightarrow{k} 0$ for each $x \in D_S$, $x \neq 1$.*

Proof. Let ψ be an invariant mean according to the above arguments. By Goldstine's theorem [2, p. 424], there is a sequence (h_k), $h_k \in l^1(\mathbf{N}_0)$, $\|h_k\|_1 \leqq 1$ such that $jh_k \xrightarrow{k} \psi$ in the τ-topology. Note that $l^1(\mathbf{N}_0)$ is separable. Consider the Jordan decompositions $h_k = h_{1k} - h_{2k} + ih_{3k} - ih_{4k}$. Since $\psi \in \mathcal{M}$ we may assume that $h_{3k} = h_{4k} = 0$. Further $1 \geqq \|h_{1k} - h_{2k}\|_1 = \hat{h}_{1k}(1) + \hat{h}_{2k}(1)$ and $\hat{h}_k(1) = \hat{h}_{1k}(1) - \hat{h}_{2k}(1) \xrightarrow{k} 1$ imply that $\hat{h}_{1k}(1) \xrightarrow{k} 1$ and $\hat{h}_{2k}(1) \xrightarrow{k} 0$. Let $g_k = h_{1k}/\hat{h}_{1k}(1)$, k sufficiently large. Fix $x \in D_S$, $x \neq 1$. Then

$$\sum_{n \in \mathbf{N}_0} g_k(n) p_1 * \alpha_x(n) \xrightarrow{k} \psi(p_1 * \alpha_x) = \psi(\alpha_x)$$

and

$$\sum_{n \in \mathbf{N}_0} g_k(n) p_1 * \alpha_x(n) = P_1(x) \sum_{n \in \mathbf{N}_0} g_k(n) \alpha_x(n) \xrightarrow{k} P_1(x) \psi(\alpha_x).$$

Since $P_1(x) \neq 1$, we have $\hat{g}_k(x) \xrightarrow{k} 0$.

Proposition 4. *Let $\varphi \in \mathcal{M}$ be invariant on $M(D_S)^\vee$, i.e. $\varphi(g * \check{\nu}) = \varphi(\check{\nu})$ for each $g \in M^1(\mathbf{N}_0)$, $\nu \in M(D_S)$. Then $\varphi(\check{\nu}) = \nu(\{1\})$ for $\nu \in M(D_S)$.*

Proof. The argument in Lemma 1 yields a sequence (h_k), $h_k \in M^1(\mathbf{N}_0)$, such that $\sum_{k \in \mathbf{N}_0} h_k(n) f(n) \xrightarrow{k} \varphi(f)$, $f \in l^\infty(\mathbf{N}_0)$. In particular $\hat{h}_k(1) = 1$ and $\hat{h}_k(x) \to 0$ for $x \neq 1$. Since

$$\sum_{n \in \mathbf{N}_0} h_k(n) \check{\nu}(n) = \int_{D_S} \hat{g}_k(x)\, d\nu(x),$$

we obtain by the dominated convergence theorem $\varphi(\check{\nu}) = \nu(\{1\})$.

Now given $f \in l^\infty(\mathbf{N}_0)$ let $\mathcal{O}(f)$ be the weak-$*$ closure of $\{g * f\colon g \in M^1(\mathbf{N}_0)\}$.

Proposition 5. *Let $f \in l^\infty(\mathbf{N}_0)$ such that the constant function $c \in \mathcal{O}(f)$. Then there exists $\psi \in \mathcal{M}$ such that ψ is invariant on $M(D_S)^\vee$ and $\psi(f)=c$.*

Proof. Let (h_k) be a sequence such that $h_k * f \xrightarrow{k} c$ in the weak-$*$ topology. By Lemma 1 we may assume that in addition $\hat{h}_k(x) \xrightarrow{k} 0$ for $x \neq 1$. Let ψ be a τ-cluster point of (jh_k) in $l^\infty(\mathbf{N}_0)^*$. Then $\psi \in \mathcal{M}$ and for $g \in M^1(\mathbf{N}_0)$, $v \in M(D_S)$ we obtain

$$\psi(g * \check{v}) = \lim \sum_{n \in \mathbf{N}_0} h_k(n) g * \check{v}(n) = \lim \sum_{n \in \mathbf{N}_0} h_k * g(n) \check{v}(n) =$$

$$= \lim \int_{D_S} \hat{h}_k(x) \hat{g}(x) dv(x) = v(\{1\}) = \psi(\check{v}).$$

Further $\psi(f) = \lim \sum_{n \in \mathbf{N}_0} h_k(n) f(n) = \lim h_k * f(0) = c$.

Theorem 1. *$M(D_S)^\vee$ is a proper subspace of $l^\infty(\mathbf{N}_0)$, i.e. $\mathbf{N}_0$ is not a Sidon set.*

Proof. We present a function $f \in l^\infty(\mathbf{N}_0)$ such that $\mathcal{O}(f)$ contains the two constants 1 and 0. Then by Proposition 4 and 5 the assertion follows. Let

$$f(n) = \begin{cases} 1 & \text{if} \quad n = 5^i, \quad 5^i+1, \ldots, 5^i + 2 \cdot 5^i - 1 \quad \text{and} \quad n = 0 \\ 0 & \text{if} \quad n = 5^i + 2 \cdot 5^i, \ldots, 5^{i+1} - 1, \quad \text{where} \quad i \in \mathbf{N}_0. \end{cases}$$

Let $n_i = 2 \cdot 5^i$. One easily obtains that $p_{n_i} * f(m) \xrightarrow{i} 1$. For $n_i = 4 \cdot 5^i$ we have $p_{n_i} * f(m) \xrightarrow{i} 0$. In fact choose $i \in \mathbf{N}$ such that $m+1 \leqq 5^i$.

IV. Orthogonal polynomial sequences admitting only finite Sidon sets

Let A be a finite subset of D_S. Denote $M_A(D_S) = \{\mu \in M(D_S) \colon |\mu|(A) = 0\}$. Obviously $M(D_S) = M(A) \oplus M_A(D_S)$.

Proposition 6. *Assume that there exists a finite subset A of D_S such that $M_A(D_S)^\vee \subseteqq c_0(\mathbf{N}_0)$. Then the Sidon sets are exactly the finite subsets of $\mathbf{N}_0$.*

Proof. Assume that $E \subseteqq \mathbf{N}_0$ is an infinite Sidon set. Since $M(D_S) = M(A) \oplus \oplus M_A(D_S)$ we obtain that

$$l^\infty(E) = M(S)^\vee | E \subseteqq V + c_0(E),$$

where V is a space with dimension at most $|A|$. But E being infinite, $c_0(E)$ has infinite codimension in $l^\infty(E)$.

Assume that for each $x, y \in D_S$ there exists a (not necessarily positive) measure $\mu_{x,y} \in M(D_S)$ such that

(i) $P_n(x) P_n(y) = \int_{D_S} P_n(z) d\mu_{x,y}(z)$,

(ii) $\|\mu_{x,y}\| \leqq M$, M a constant independent of x, y.

Using conditions (i) and (ii) we can show that given $f\in C(D_S)$ the map $(x, y)\to\mu_{x,y}(f)$ is continuous, compare e.g. Proposition 1 of [8]. Hence we can define a "quasi-convolution" of two measures $\mu, \nu\in M(D_S)$ by

$$\mu*\nu(f)=\int_{D_S}\int_{D_S}\mu_{x,y}(f)\,d\mu(x)\,d\nu(y).$$

By (i) $\mu*\nu(P_n)=\mu(P_n)\nu(P_n)$ is valid for each $n\in\mathbf{N}_0$.

We present now examples for which Proposition 6 applies. The Jacobi-polynomials $P_n^{(\alpha,\beta)}(x)$ are orthogonal on $D_S=[-1,1]$ with respect to $d\pi(x)=(1-x)^\alpha\cdot(1+x)^\beta dx$ (up to normalization). The sequences $(P_n^{(\alpha,\beta)}(x))$ satisfy property (P) for $\alpha\geqq\beta>-1$, $\alpha+\beta+1\geqq 0$, see Chapter 3(a) of [8]. The generalized Tchebichef polynomials $T_n^{(\alpha,\beta)}(x)$ are orthogonal on $D_S=[-1,1]$ with respect to $d\pi(x)=(1-x^2)^\alpha|x|^{2\beta+1}dx$ (up to normalization) and satisfy property (P) for $\beta>-1$, $\alpha\geqq\beta+1$, see Chapter 3(f) of [8]. Finally we consider polynomials $G_n^a(x)$ studied by Geronimus. They are orthogonal on $D_S=[-1,1]$ with respect to $(1-x^2)^{1/2}/(1-\mu x^2)$, $\mu=a-a^2/4$ and satisfy property (P) for $a\geqq 2$, see Chapter 3(g)(i) of [8].

Theorem 2. *The set $E\subseteqq\mathbf{N}_0$ is a Sidon set if and only if E is finite in case*

(a) $P_n(x)=P_n^{(\alpha,\beta)}(x)$ *the Jacobi polynomials with* $\alpha\geqq\beta>-1$, $\alpha+\beta+1\geqq 0$ *and* $\alpha\neq -1/2$.

(b) $P_n(x)=T_n^{(\alpha,\beta)}(x)$ *the generalized Tchebichef polynomials with* $\beta>-1$, $\alpha\geqq\beta+1$.

(c) $P_n(x)=G_n^a(x)$ *with* $a>2$.

Proof. (a) Fix $\alpha,\beta\in\mathbf{R}$ such that $\alpha\geqq\beta>-1$, $\alpha+\beta+1\geqq 0$ and choose $A=\{-1,1\}$. By Gasper's theorem of [4] and (2.3), (2.4) of [7] there exist for $x, y\in D_S=[-1,1]$ measures $\mu_{x,y}\in M(D_S)$ such that the above conditions (i) and (ii) are satisfied. First consider the case $\alpha+\beta+1>0$. If $x,y\in]-1,1[$ then $d\mu_{x,y}(z)=K(x,y,z)d\pi(z)$, see [4]. Let $\mu\in M_A(D_S)$. We show that $\mu*\mu\in L^1(D_S,\pi)$. Let $B\subseteqq D_S$ be a Borel set such that $\pi(B)=0$. Then

$$|\mu*\mu(B)|\leqq\int_{D_S\setminus A}\int_{D_S\setminus A}|\mu_{x,y}(B)|\,d|\mu|(x)\,d|\mu|(y)+$$
$$+\int_A\int_{D_S}|\mu_{x,y}(B)|\,d|\mu|(x)\,d|\mu|(y)+\int_{D_S}\int_A|\mu_{x,y}(B)|\,d|\mu|(x)\,d|\mu|(y).$$

Since $\mu\in M_A(D_S)$, the second and third integrals are zero. Since the measures $\mu_{x,y}$, $-1<x, y<1$, are absolutely continuous, the first integral is zero. Hence $\check{\mu}\check{\mu}=(\mu*\mu)^\vee\in c_0(\mathbf{N}_0)$ by Proposition 1 of [10]. Then obviously $\check{\mu}\in c_0(\mathbf{N}_0)$ and Proposition 6 applies. If $\beta>-1$, $\alpha>-1/2$ and $\alpha+\beta+1=0$ then given $x,y\in]-1,1[$ we have $d\mu_{x,y}(z)=K(x,y,z)d\pi(z)+d\nu_{x,y}(z)$, where $\nu_{x,y}=0$ if $x\neq -y$ and $\nu_{x,y}=p_{-1}/2$ if $x=-y$. If $\mu\in M_A(D_S)$ and $\pi(B)=0$ we obtain now $\mu*\mu(B)-$

$-cp_{-1}(B)=0, \quad c=\int\limits_{D_S\setminus A}\int\limits_{D_S\setminus A} v_{x,y}(\{-1\})d\mu(x)d\mu(y).$ Hence $\check{\mu}\check{\mu}-c\alpha_{-1}\in c_0(\mathbf{N}_0)$. Using the recurrence formula of $P_n(x)$ and $\alpha+\beta+1=0$ an induction argument shows that

$$\alpha_{-1}(n)=P_n(-1)=\prod_{k=0}^{n-1}(\alpha-k)/\prod_{k=1}^{n}(\alpha+k), \quad n\in\mathbf{N}.$$

Hence $\lim|\alpha_{-1}(n)|=(\Gamma(1-|\alpha|)/\Gamma(|\alpha|))\lim\Gamma(|\alpha|+n)/\Gamma(1-|\alpha|+n)=0$, because of $-1/2<\alpha<0$. Thus $\alpha_{-1}\in c_0(\mathbf{N}_0)$ and consequently $\check{\mu}\in c_0(\mathbf{N}_0)$.

(b) Choose again $A=\{-1,1\}$. Using Theorem 1 of [7] an argument as in (a) yields that $M_A(D_S)^{\check{}}\subseteq c_0(\mathbf{N}_0)$.

(c) Derive from [8] or from Chapter VI, (13.4) of [1] that

$$G_n^a(x)=\big(a/(n(a-2)+2)\big)P_n^{(-1/2,-1/2)}(x)+\big(((a-2)(n-1))/(n(a-2)+2)\big)P_n^{(1/2,1/2)}(x).$$

For $A=\{-1,1\}$ and $\mu\in M_A(D_S)$ we have

$$\check{\mu}(n)=\big(a/(n(a-2)+2)\big)\int\limits_{D_S}P_n^{(-1/2,-1/2)}(x)d\mu(x)+$$

$$+\big(((a-2)(n-1))/(n(a-2)+2)\big)\int\limits_{D_S}P_n^{(1/2,1/2)}(x)d\mu(x).$$

Since $|\int\limits_{D_S}P_n^{(-1/2,-1/2)}(x)d\mu(x)|\leqq\|\mu\|$ for $n\in\mathbf{N}_0$ and $\int\limits_{D_S}P_n^{(1/2,1/2)}(x)d\mu(x)\xrightarrow{n}0$ by (a), we have $\check{\mu}\in c_0(\mathbf{N}_0)$ provided $a>2$.

Remark. The assertion of Theorem 2(a) follows by Theorem 3.2 of [14] and Chapter 4(a) of [8] provided we require that in addition $\beta\geqq-1/2$ or $\alpha+\beta\geqq0$.

V. Infinite Sidon sets

Finally we consider orthogonal polynomial sequences $(P_n(x))$ having infinite Sidon sets. Let $E=\{n_1,n_2,\ldots\}\subseteq\mathbf{N}_0$ and m, $N\in\mathbf{N}$, $N\geqq m$. Denote by

$$E_N^m=\{p_{n_{i_1}}*p_{n_{i_2}}*\ldots*p_{n_{i_m}}\in M^1(\mathbf{N}_0)\colon 1\leqq i_1<i_2<\ldots<i_m\leqq N\}$$

and call E a *Rider set* if there exists a constant $B\geqq1$ such that

$$\lim_N\sum_{\mu\in E_N^m}\mu(\{0\})\leqq B^m \quad \text{for each } m\in\mathbf{N}.$$

Lemma 2. *Let E be a Rider set. There exists a constant $C\geqq1$ such that $\lim\limits_N\sum\limits_{\mu\in E_N^m}\mu(\{k\})\leqq C^m h(k)$ for each $k\in\mathbf{N}_0$, $m\in\mathbf{N}$.*

Proof. Write $E=\{n_1, n_2, \dots\}$. Let $\beta=1/(2B)$, where B is the constant of the Rider set E. Consider for $N\in\mathbf{N}$ the Riesz products

$$R_N(x) = \prod_{k=1}^{N}(1+\beta P_{n_k}(x)).$$

Obviously $R_N(x)=1+\sum_{k=0}^{\infty} c_N(k)P_k(x)$, where $c_N(k)=\sum_{m=1}^{N}\left(\sum_{\mu\in E_N^m}\mu(\{k\})\right)\beta^m$. Now conclude as in the proof of Lemma 3.1 of [12], compare also [11, p. 28], that

$$\|R_N\|_1 = 1+c_N(0) \leqq 1+\sum_{m=1}^{\infty} 2^{-m} = 2.$$

Hence for $k\in\mathbf{N}$, $c_N(k)/h(k)=|\check{R_N}(k)|\leqq\|R_N\|_1\leqq 2$ and then $\lim_N \sum_{\mu\in E_N^m}\mu(\{k\})\leqq$ $\leqq(2B)^m 2h(k)\leqq C^m h(k)$, where $C=4B$.

Lemma 3. *Let E be a Rider set with $0\notin E$, and let $C\geqq 1$ be the corresponding constant of Lemma 2. Let $0<\varepsilon<1$. Given $\varphi\colon E\to\mathbf{R}$, $\|\varphi\|_E\leqq 1$ there exists a positive measure $\mu\in M(S)$ such that $\|\mu\|\leqq\varepsilon+2C^2/\varepsilon$, $|\check{\mu}(k)|\leqq\varepsilon$ for each $k\notin E\cup\{0\}$ and $|\check{\mu}(k)-\varphi(k)/h(k)|\leqq\varepsilon$ for $k\in E$.*

Proof. We have again to modify the arguments of Theorem 3.2 in [12] or of [11, p. 28—29]. Write $E=\{n_1, n_2, \dots\}$. Let $\beta=\varepsilon/(2C^2)$ and

$$R_N(x) = \prod_{k=1}^{N}(1+\beta\varphi(n_k)P_{n_k}(x)) = 1+\sum_{k=1}^{N}\beta\varphi(n_k)P_{n_k}(x)+\sum_{k=0}^{\infty} d_N(k)P_k(x),$$

where $|d_N(k)|\leqq\sum_{m=2}^{N}\left(\sum_{\mu\in E_N^m}\mu(\{k\})\right)\beta^m$. Then

$$\|R_N\|_1 \leqq 1+|d_N(0)| \leqq 1+\sum_{m=2}^{\infty}(C\beta)^m \leqq 1+\varepsilon\beta.$$

Further for $n_k\in E$ and $N\geqq k$

$$|h(n_k)\check{R_N}(n_k)-\beta\varphi(n_k)| \leqq |d_N(n_k)| \leqq \sum_{m=2}^{\infty}(C\beta)^m h(n_k) \leqq \varepsilon\beta h(n_k).$$

For $n\notin E\cup\{0\}$ we have $|\check{R_N}(n)|=|d_N(n)|/h(n)\leqq\varepsilon\beta$. Alaoglu's theorem and a normalization by $1/\beta$ yields the appropriate positive measure μ.

Theorem 3. *Let E be a finite union of Rider sets and assume that* $\sup\{h(k)\colon k\in E\}<\infty$. *Then E is a Sidon set.*

Proof. Assume $0\notin E$. As in [11, pp. 29—30] one obtains that Lemma 3 is valid for a finite union of Rider sets. Let $M=\sup\{h(k)\colon k\in E\}$. Given $\psi\colon E\to\{-1, 1\}$ consider $\varphi\colon E\to\mathbf{R}$, $\varphi(n)=\psi h(n)/M$, $n\in E$. There exists a positive measure $\mu\in M(S)$

such that

$$|\check{\mu}(k)-\psi(k)/M| \leqq 1/(2M) \quad \text{for} \quad k\in E.$$

Define $\nu=M\mu$. By Proposition 1(f) E is a Sidon set. Finally let $0\in E$. Given $\psi\colon E\to\{-1,1\}$ we know that there exists a measure such that $|\check{\mu}(k)-\psi(k)|\leqq 1/2$ for $k\in E\setminus\{0\}$. Define $\alpha=\psi(0)-\check{\mu}(0)$. Replace μ by $\mu+\alpha\pi$ establishing that E is a Sidon set.

Corollary. *Assume that* $\sup\{h(k)\colon k\in\mathbf{N}\}<\infty$. *If* $E=\{n_1, n_2, \ldots\}\subseteq\mathbf{N}$ *satisfies* $n_{k+1}/n_k\geqq q$ *for* $k\in\mathbf{N}$, *where* $q>1$ *is a constant, then E is a Sidon set.*

Proof. It is sufficient to consider the case $q\geqq 3$, compare e.g. [11, p. 23]. Then for $m, N\in\mathbf{N}$, $2\leqq m\leqq N$ and $1\leqq n_{i_1}<n_{i_2}<\ldots<n_{i_m}\leqq N$ we obtain $n_{i_m}-(n_{i_{m-1}}+\ldots+n_{i_1})\geqq n_{i_m}(q-2)/(q-1)\geqq 0$. Hence $0\notin\operatorname{supp} p_{n_{i_1}}*\ldots*p_{n_{i_m}}$. Consequently E is a Rider set.

We present now examples with bounded Haar function h (and property (P)). Of course the Tchebichef polynomials of first kind, $T_n(x)=P_n^{(-1/2,\,-1/2)}(x)=\cos n\varphi$, $\cos\varphi=x$, $\varphi\in[0,\pi]$, have the Haar function $h(0)=1$, $h(n)=2$, $n\in\mathbf{N}$. A class containing $(T_n(x))$ is studied in Chapter 3(g) (ii) of [8]. For $a\geqq 2$ these polynomials $T_n(x;\,a)$ have the representation

$$T_1(x;\,a)=x,\quad T_n(x;\,a)=\big(a/2(a-1)\big)T_n(x)+\big((a-2)/2(a-1)\big)T_{n-2}(x),\quad n\geqq 2.$$

The Haar function is $h(0)=1$, $h(1)=a$, $h(n)=2(a-1)$, $n\geqq 2$. We introduce an extension depending on two parameters a, $b\geqq 2$. Let $a_1=(a-1)/a$, $c_1=1/a$, $a_2=(b-1)/b$, $c_2=1/b$, $a_n=c_n=1/2$ if $n=3, 4, \ldots$ and $b_n=0$, $n\in\mathbf{N}$. Further let $a_0=1$, $b_0=0$. By the recursion formula, see Section I, there is defined an orthogonal polynomial sequence $(T_n(x;\,a,b))$ with the representation

$$T_1(x;\,a,b)=x,\quad T_2(x;\,a,b)=\big(a/2(a-1)\big)T_2(x)+\big((a-2)/2(a-1)\big)T_0(x),$$

$$T_3(x;\,a,b)=\big(ab/4(a-1)(b-1)\big)T_3(x)+$$

$$+\big(((a-2)(b-2)+(a-2)b+(b-2)a)/4(a-1)(b-1)\big)T_1(x),$$

$$T_n(x;\,a,b)=\big(ab/4(a-1)(b-1)\big)T_n(x)+\big(2(a-2)/4(a-1)\big)T_{n-2}(x)+$$

$$+\big(a(b-2)/4(a-1)(b-1)\big)T_{n-4}(x)\quad \text{if}\quad n=3,4,\ldots$$

$(T_n(x;\,a,b))$ satisfies property (P). In fact the coefficients $g(m,n,n+m-k)$, $m\leqq n$, $0\leqq k\leqq 2m$, can be computed directly using formula (1) of [8]. Obviously $g(m,n,n+m-k)=0$ for $k=1,3,\ldots,2m-1$. We omit the coefficients $g(m,n,n+m-k)$

for $m=2,3,4$ noting only their positivity. The general formulas for $5 \leqq m \leqq n$ are

$$8(a-1)(b-1)\cdot g(m,n,n+m-2k) = \begin{cases} ab & \text{if } k=0 \\ (a-2)(b-2)+b(a-2) & \text{if } k=1 \\ a(b-2) & \text{if } k=2 \\ 0 & \text{if } k=3,\dots,m-3 \end{cases}$$

and for $k=m-2$

$$8(a-1)(b-1)\cdot g(m,n,n-m+4) = \begin{cases} 2a(b-2) & \text{if } n=m \\ a(b-2) & \text{if } n=m+1,\dots, \end{cases}$$

for $k=m-1$

$$8(a-1)(b-1)\cdot g(m,n,n-m+2) =$$

$$= \begin{cases} 2b(a-2) & \text{if } n=m \\ (a-2)(b-2)+b(a-2)+a(b-2) & \text{if } n=m+1 \\ (a-2)(b-2)+b(a-2) & \text{if } n=m+2,\dots, \end{cases}$$

and for $k=m$

$$8(a-1)(b-1)\cdot g(m,n,n-m) = \begin{cases} 4 & \text{if } n=m \\ 2a & \text{if } n=m+1 \\ ab & \text{if } n=m+2,\dots. \end{cases}$$

The Haar function is

$$h(0)=1, \quad h(1)=a, \quad h(2)=b(a-1), \quad h(n)=2(a-1)(b-1) \quad \text{if} \quad n=3,4,\dots.$$

Remark. The above example suggests the way how to define a general class of polynomials depending on an arbitrary number of parameters and having bounded Haar function. A study of this class, such as representations and orthogonality relations, will be given in another paper.

References

[1] T. S. Chihara, *An Introduction to Orthogonal Polynomials*, Gordon and Breach (New York, 1978).

[2] N. Dunford and J. T. Schwartz, *Linear Operators. I: General Theory*, Interscience (New York, 1958).

[3] C. F. Dunkl and D. E. Ramirez, Existence and nonuniqueness of invariant means on $\mathcal{L}^\infty(\hat{G})$, *Proc. Amer. Math. Soc.*, **32** (1972), 525—530.

[4] G. Gasper, Banach algebras for Jacobi series and positivity of a kernel, *Ann. of Math.*, **95** (1972), 261—280.

[5] E. Hewitt and K. A. Ross, *Abstract harmonic analysis, II*, Springer-Verlag (Berlin, 1970).

[6] R. I. Jewett, Spaces with an abstract convolution of measures, *Adv. in Math.*, **18** (1975), 1—101.

[7] T. P. LAINE, The product formula and convolution structure for the generalized Chebyshev polynomials, *SIAM J. Math. Anal.*, **11** (1980), 133—146.
[8] R. LASSER, Orthogonal polynomials and hypergroups, *Rend. Mat.*, **3** (1983), 185—209.
[9] R. LASSER, Bochner theorems for hypergroups and their applications to orthogonal polynomial expansions, *J. Approx. Theory*, **37** (1983), 311—325.
[10] R. LASSER, Fourier—Stieltjes transforms on hypergroups, *Analysis*, **2** (1982), 281—303.
[11] J. M. LOPEZ and K. A. ROSS, *Sidon sets*, Marcel Dekker (New York, 1975).
[12] W. A. PARKER, Central Sidon and central Λ_p sets, *J. Austral. Math. Soc.*, **14** (1972), 62—74.
[13] R. C. VREM, Lacunarity on compact hypergroups, *Math. Z.*, **164** (1978), 93—104.
[14] R. C. VREM, Continuous measures and lacunarity on hypergroups, *Trans. Amer. Math. Soc.*, **269** (1982), 549—556.

INSTITUT FÜR MATHEMATIK
DER TECHNISCHEN UNIVERSITÄT MÜNCHEN
ARCISSTR. 21
8000 MÜNCHEN 2, WEST GERMANY

Acta Sci. Math. 47 (1984), 405—412

О разрывности сопряженной функции

В. И. КОЛЯДА

Пусть $f(x)$ — 2π-периодическая суммируемая функция, и

$$\tilde{f}(x) = \lim_{\varepsilon\to+0}\left\{-(1/\pi)\int_{\varepsilon}^{\pi}\left(f(x+t)-f(x-t)\right)/(2\operatorname{tg} t/2)\,dt\right\}$$

— функция, сопряженная к $f(x)$. Н. Н. Лузин впервые обратил внимание на то, что сопряженная функция $\tilde{f}(x)$ может быть несуммируемой ни на одном интервале $\Delta\subset[0, 2\pi]$. Более того, Н. Н. Лузин [1] доказал существование абсолютно непрерывной функции $F(x)$, сопряженная к которой существенно неограничена на любом интервале.

Измеримую на интервале Δ функцию $\varphi(x)$ будем называть существенно непрерывной в точке $x_0\in\Delta$, если существует функция $\varphi^*(x)$, эквивалентная $\varphi(x)$ и непрерывная в точке x_0; в противном случае говорят, что $\varphi(x)$ существенно разрывна в точке x_0.

В примере Лузина функция $\tilde{F}(x)$ существенно разрывна всюду; однако она является функцией 1-го класса Бэра, и $\lim_{x\to\xi}\tilde{F}(x)=-\infty$ в каждой точке ξ некоторого множества 2-ой категории. В связи с этим возникает вопрос: если функция $F(x)$ абсолютно непрерывна, а сопряженная к ней функция $\tilde{F}(x)$ ограничена, то не обязана ли $\tilde{F}(x)$ иметь точки существенной непрерывности?

Ответ на этот вопрос отрицателен. Именно, в предлагаемой статье (теорема 1) строится пример абсолютно непрерывной функции, сопряженная к которой существенно ограничена и всюда существенно разрывна.

Далее, пусть E — множество всех тех точек $x\in[0, 2\pi]$, в которых существует сопряженная функция $\tilde{f}(x)$. В статье устанавливается (теорема 2), что для каждой суммируемой функции $f(x)$ сопряженная к ней $\tilde{f}(x)$ обладает следующими свойствами: для любого интервала $\Delta\subset[0, 2\pi]$ точная верхняя грань функции $\tilde{f}(x)$ на множестве $\Delta'=E\cap\Delta$ не изменяется при выбрасывании из Δ' произвольного нуль-множества; существенная непрерывность функции $\tilde{f}(x)$ в точке x_0

Поступило 7 марта 1983.

равносильна ее обычной непрерывности в этой точке по множеству E. Эти факты интересно сопоставить с теоремой Шеффера—Лебега ([2], стр. 77), выражающей аналогичные свойства производной функции.

Таково содержание работы; перейдем теперь к детальному изложению ее результатов.

Лемма 1. *Пусть ε и η — положительные числа. Тогда существует неотрицательная 2π-периодическая функция $f_{\varepsilon,\eta}\equiv f\in L$, обладающая следующими свойствами:*

$$\int_0^{2\pi} f(t)\,dt < \varepsilon; \tag{1}$$

$$g(x) \equiv \int_0^{2\pi} f(t)\big|\ln|\sin(x-t/2)|\big|\,dt < 2 \quad \text{для всех} \quad x; \tag{2}$$

$$g(x) \lesssim \varepsilon \quad \text{для всех} \quad x\in[\eta, 2\pi]; \tag{3}$$

$$g(x) \quad \text{непрерывна на} \quad (0, 2\pi); \tag{4}$$

$$\text{для любого} \quad \delta > 0, \quad \inf_{0<x<\delta} g(x) < \varepsilon, \quad \sup_{0<x<\delta} g(x) > 1. \tag{5}$$

Доказательство. Будем считать, что $\varepsilon<1$ и $\eta<1$. Положим $\varphi(t)=$ $=|\ln|\sin t/2||$, $\alpha_k=\eta 2^{-k^2}$ $(k=1, 2, \ldots)$, и выберем положительную убывающую последовательность $\{\delta_k\}$ так, чтобы было

$$\varphi(\alpha_k/4) < (\varepsilon/2^{k+2})\varphi(\delta_k), \tag{6}$$

$$\delta_k < \alpha_k/8, \quad k = 1, 2, \ldots. \tag{7}$$

Обозначим $I_k=[\alpha_k-\delta_k, \alpha_k]$, $N_k=1/\delta_k\varphi(\delta_k)$. Отрезки I_k попарно не пересекаются, поскольку

$$\alpha_k-\alpha_{k+1} \geqq 7\alpha_k/8. \tag{8}$$

Положим

$$f(t) = \begin{cases} N_k, & t\in I_k \quad (k = 1, 2, \ldots) \\ 0, & t\in[0, 2\pi]\setminus \bigcup\limits_{j=1}^{\infty} I_j, \end{cases}$$

$f(t+2\pi)=f(t)$. Покажем, что функция f обладает требуемыми свойствами.

Прежде всего, в силу (6),

$$\int_0^{2\pi} f(t)\varphi(t)\,dt = \sum_{k=1}^{\infty} N_k \int_{I_k} \varphi(t)\,dt < \sum_{k=1}^{\infty} N_k\delta_k\varphi(\alpha_k/4) < \varepsilon\sum_{k=1}^{\infty} 2^{-k-2} < \varepsilon; \tag{9}$$

отсюда, в частности, следует (1).

Далее, заметим, что для любого $x\in[0, 2\pi]$

$$\int\limits_{I_k} f(t)\varphi(x-t)\,dt \leqq 2N_k \int\limits_0^{\delta_k/2} \varphi(u)\,du < 3/2.$$

Пусть $x\in[(\alpha_k+\alpha_{k+1})/2, \alpha_k]$. Тогда

$$\int\limits_0^{2\pi} f(t)\varphi(x-t)\,dt = \sum_{j=1}^{\infty} \int\limits_{I_j} f(t)\varphi(x-t)\,dt < (3/2)+\sum_{j\neq k} \int\limits_{I_j} f(t)\varphi(x-t)\,dt.$$

При $j>k$ и $t\in I_j$ имеем $x-t>\alpha_k/4$ (см. (8)), и, в силу (6),

$$\begin{aligned}\sum_{j=k+1}^{\infty} \int\limits_{I_j} f(t)\varphi(x-t)\,dt &< \varphi(\alpha_k/4) \sum_{j=k+1}^{\infty} \int\limits_{I_j} f(t)\,dt = \\ &= \varphi(\alpha_k/4) \sum_{j=k+1}^{\infty} [\varphi(\delta_j)]^{-1} < \varepsilon \sum_{j=k+1}^{\infty} 2^{-j-2} \leqq \varepsilon/4.\end{aligned} \tag{10}$$

Далее, если $1\leqq j<k$, и $t\in I_j$, то (см. (7) и (8)) $|x-t|>\alpha_j/4$, и

$$\begin{aligned}\sum_{j=1}^{k-1} \int\limits_{I_j} f(t)\varphi(x-t)\,dt &\leqq \sum_{j=1}^{k-1} \varphi(\alpha_j/4) \int\limits_{I_j} f(t)\,dt = \\ &= \sum_{j=1}^{k-1} \varphi(\alpha_j/4)/\varphi(\delta_j) < \varepsilon \sum_{j=1}^{k-1} 2^{-j-2} < \varepsilon/4.\end{aligned} \tag{11}$$

Таким образом, $g(x)=\int\limits_0^{2\pi} f(t)\varphi(x-t)\,dt<2$ для всех $x\in[(\alpha_k+\alpha_{k+1})/2, \alpha_k]$. Аналогично убеждаемся в справедливости этого неравенства в случае $x\in[\alpha_{k+1}, (\alpha_k+\alpha_{k+1})/2]$, а также в случае $x\in[\alpha_1, \eta]$.

Пусть теперь $x\in[\eta, 2\pi]$. Если $\eta\leqq x<\pi$, то для всех $t\in I_j$ $(j=1, 2, \ldots)$, $\alpha_1\leqq x-t<\pi$, $\varphi(x-t)\leqq\varphi(\alpha_1)$, и

$$g(x) \leqq \varphi(\alpha_1) \sum_{j=1}^{\infty} [\varphi(\delta_j)]^{-1} \leqq \varepsilon \sum_{j=1}^{\infty} 2^{-j-2} < \varepsilon.$$

Если же $\pi\leqq x\leqq 2\pi$, то, как легко видеть, $\varphi(x-t)<\varphi(t)$ для $t\in[0, \alpha_1]$; следовательно (см. (9)), $g(x)\leqq g(0)<\varepsilon$. Таким образом, свойство (3) выполняется.

Далее, чтобы установить непрерывность функции $g(x)$ в произвольной точке $\xi\in(0, 2\pi)$, возьмем отрезок $[\alpha, \beta]\subset(0, 2\pi)$, такой, что $\alpha<\xi<\beta$. Тогда для любого $\sigma>0$ найдется такое $\tau>0$, что при всех $x\in[\alpha, \beta]$

$$\int\limits_0^{\tau} f(t)\varphi(x-t)\,dt < \delta.$$

Остается учесть еще, что функция

$$g^*(x) = \int_{\tau}^{2\pi} f(t)\varphi(x-t)\,dt$$

непрерывна, поскольку $f(t)$ ограничена на $[\tau, 2\pi]$.

Наконец, заметим, что при любом k

$$g(\alpha_k) > \int_{I_k} f(t)\varphi(\alpha_k - t)\,dt = N_k \int_0^{\delta_k} \varphi(u)\,du > 1.$$

С другой стороны, если $\beta_k = (\alpha_k + \alpha_{k+1})/2$, то, пользуясь оценками (10) и (11), и учитывая, что $t - \beta_k > \alpha_k/4$ для всех $t \in I_k$ (см. (7) и (8)), получаем, в силу (6):

$$g(\beta_k) < (\varepsilon/2) + \int_{I_k} f(t)\varphi(\beta_k - t)\,dt < (\varepsilon/2) + \varphi(\alpha_k/4)/\varphi(\delta_k) < \varepsilon.$$

Таким образом, имеет место свойство (5). Лемма доказана.

Теорема 1. *Существует 2π-периодическая абсолютно непрерывная функция $F(x)$, такая, что сопряженная функция $\tilde{F}(x)$ существенно ограничена и всюду существенно разрывна.*

Доказательство. Пусть $\{\varrho_k\}$ — последовательность всех рациональных точек отрезка $[0, 2\pi]$, $\varrho_0 = 0$, и $\varepsilon_k = 2^{-2k-1}$ $(k=0, 1, \ldots)$, $\eta_0 = 1/2$. Применяя лемму 1, положим $f_0(t) = f_{\varepsilon_0, \eta_0}(t)$, и по индукции построим последовательность положительных чисел $\{\eta_k\}_{k=1}^{\infty}$, последовательность $\{\sigma_k\}_{k=0}^{\infty}$ с $|\sigma_k| = 1$ $(\sigma_0 = 1)$, и последовательность функций $f_k(t) = f_{\varepsilon_k, \eta_k}(t - \varrho_k)$; выбор этих последовательностей будем производить, исходя из свойств (4) и (5) функций

$$g_k(x) = \int_0^{2\pi} f_k(t)\varphi(x-t)\,dt \quad (\varphi(t) = \big|\ln|\sin(t/2)|\big|)$$

так, чтобы выполнялись следующие условия:

(a) $\Delta_k \equiv [\varrho_k, \varrho_k + \eta_k] \subset (0, 2\pi)$, $k = 1, 2, \ldots$;

(b) $\eta_k < 2^{-k} \min\limits_{0 \leqq j < k} \mu_j$, где μ_j — наименьшая из мер множеств $I_j' = \{x \in \Delta_j : g_j(x) < \varepsilon_j\}$, $I_j'' = \{x \in \Delta_j : g_j(x) > 1\}$;

(c) колебание функции $S_{k-1}(x) = \sum\limits_{j=0}^{k-1} \sigma_j g_j(x)$ на отрезке Δ_k меньше ε_k;

(d) если $S_{k-1}(\varrho_k) \geqq 0$, то $\sigma_k = -1$; в противном случае $\sigma_k = 1$.

Поскольку $\int_0^{2\pi} f_k(t)\,dt < \varepsilon_k$ (см. (1)), то ряд $\sum\limits_{k=0}^{\infty} \sigma_k f_k(t)$ сходится в L к некоторой суммируемой функции $f(t)$. Пусть

$$g(x) = \int_0^{2\pi} f(t)\varphi(x-t)\,dt.$$

Ясно, что ряд $\sum_{k=0}^{\infty}\sigma_k g_k(x)$ сходится к $g(x)$ в L. Покажем, что $g(x)$ существенно ограничена и всюду существенно разрывна.

Прежде всего установим, что для всех x

(12) $$|S_n(x)| \leqq 2+\varepsilon_0+\ldots+\varepsilon_n \quad (n = 0, 1, \ldots).$$

Для $n=0$ (12) выполнено (см. (2)). Предположим, что (12) имеет место для некоторого $n\geqq 1$. Если $x\in[0, 2\pi]\setminus\Delta_{n+1}$, то, в силу (3), $g_{n+1}(x)<\varepsilon_{n+1}$, и

$$|S_{n+1}(x)| \leqq |S_n(x)|+|g_{n+1}(x)| \leqq 2+\varepsilon_0+\ldots+\varepsilon_{n+1}.$$

Пусть $x\in\Delta_{n+1}$. Если $S_n(\varrho_{n+1})\geqq 0$, то, в силу (12) и свойства (c),

$$-\varepsilon_{n+1} < S_n(x) \leqq 2+\varepsilon_0+\ldots+\varepsilon_n.$$

Поскольку $S_{n+1}=S_n-g_{n+1}$ (см. (d)), а в силу (2) $0\leqq g_{n+1}(x)<2$, то

$$-2-\varepsilon_{n+1} < S_{n+1}(x) \leqq 2+\varepsilon_0+\ldots+\varepsilon_n.$$

Аналогично, в случае, когда $S_n(\varrho_{n+1})<0$, по свойству (c) имеем для $x\in\Delta_{n+1}$

$$-(2+\varepsilon_0+\ldots+\varepsilon_n) \leqq S_n(x) < \varepsilon_{n+1};$$

поскольку $S_{n+1}=S_n+g_{n+1}$, то получаем

$$-(2+\ldots+\varepsilon_0+\varepsilon_n)\leqq S_{n+1}(x) < 2+\varepsilon_{n+1}.$$

Таким образом, по индукции установлена справедливость неравенства (12). В силу этого неравенства, $|S_n(x)|<3$ при всех n и всех x. Следовательно, $|g(x)|\leqq 3$ почти всюду.

Пусть теперь $\Delta\subset[0, 2\pi]$ — произвольный интервал. Покажем, что существенное колебание функции $g(x)$ на интервале Δ больше 1/2. Очевидно, существует номер $k\geqq 1$, такой, что $\Delta_k\subset\Delta$. Полагая $E_k=\Delta_k-\bigcup_{j=k+1}^{\infty}\Delta_j$, получим, в силу свойства (b)

(13) $$|E_k| \geqq \eta_k-2^{-k}\mu_k \geqq \eta_k-\mu_k/2.$$

Но для всех $x\in E_k$ при любом $n>k$

$$|S_n(x)-S_k(x)| \leqq \sum_{j=k+1}^{\infty}\varepsilon_j < \varepsilon_k/2.$$

Следовательно, $|g(x)-S_k(x)|\leqq\varepsilon_k/2$ почти всюду на E_k. В силу (13) (см. также (b)) множества $I'_k\cap E_k$ и $I''_k\cap E_k$ имеют положительные меры, причем на первом из них выполняется неравенство $|g(x)-S_{k-1}(x)|<3\varepsilon_k/2$, а на втором $|g(x)-S_{k-1}(x)|>1-\varepsilon_k/2$. Следовательно, существенное колебание $g(x)-S_{k-1}(x)$ на Δ_k больше, чем $1-2\varepsilon_k$. Учитывая, что колебание $S_{k-1}(x)$ на Δ_k меньше ε_k

(см. (с)), получаем, что существенное колебание функции $g(x)$ на интервале Δ_k (а следовательно, и на Δ) больше 1/2. В силу произвольности интервала Δ, отсюда следует, что функция $g(x)$ существенно разрывна в каждой точке $[0, 2\pi]$.

Положим теперь

$$F(x) = \int_0^x f(t)\,dt - c_0 x, \quad \text{где} \quad c_0 = (1/2\pi)\int_0^{2\pi} f(t)\,dt.$$

Функция $F(x)$ абсолютно непрерывна и имеет период 2π. Сопряженная к ней функция $\tilde{F}(x)$, представляемая формулой Лузина ([3], стр. 556)

$$\tilde{F}(x) = -(1/\pi)\int_0^{2\pi} [f(x_0+t)-c_0]\varphi(t)\,dt =$$

$$= -(1/\pi)\int_1^{2\pi} f(x+t)\varphi(t)\,dt + c_1 = -(1/\pi)g(x)+c_1$$

существенно ограничена и всюду существенно разрывна. Теорема доказана.

Лемма 2. *Пусть* $f\in L$, *и существует* $\tilde{f}(x_0)=y_0$. *Тогда для любых положительных чисел* ε *и* δ

$$\operatorname{mes}\{x\in(x_0-\delta, x_0+\delta)\colon \tilde{f}(x) < y_0+\varepsilon\} > 0.$$

Доказательство. Будем предполагать, что $x_0=0$. Пусть существуют $\varepsilon>0$ и $\delta>0$, такие, что $\tilde{f}(x)\geqq y_0+\varepsilon$ для почти всех $x\in(-\delta,\delta)$. Из теоремы Титчмарша о Q-интегрируемости сопряженной функции ([4], теорема 6) и существенной ограниченности снизу на интервале $(-\delta,\delta)$ функции $\tilde{f}(x)$ следует суммируемость $\tilde{f}(x)$ на любом отрезке, содержащемся в $(-\delta,\delta)$.

Пусть $0<\delta'<\delta$, и $\lambda(x)$ — непрерывная 2π-периодическая функция, равная 1 для $x\in[-\delta'/2, \delta'/2]$, нулю для $\delta'\leqq|x|\leqq\pi$, и линейная на отрезках $[-\delta', -\delta'/2]$, $[\delta'/2, \delta']$. Очевидно, что

$$|\lambda(x_1)-\lambda(x_2)| \leqq K|x_1-x_2|. \tag{14}$$

Обозначим $g(x)=\lambda(x)f(x)$. Тогда функция $\tilde{g}(x)$ суммируема на $[-\pi,\pi]$. Действительно, пусть $\delta'<\delta''<\delta$. Поскольку $g(x)=0$ для $\delta'\leqq|x|\leqq\pi$, то $\tilde{g}(x)$ ограничена для значений $\delta''\leqq|x|\leqq\pi$. Далее, в силу (14),

$$|\tilde{g}(x)| = (1/\pi)\left|\int_0^{\pi}(\lambda(x+t)f(x+t)-\lambda(x-t)f(x-t))/(2\operatorname{tg}t/2)\,dt\right| \leqq$$

$$\leqq (K/\pi)\int_0^{\pi}[|f(x+t)|+|f(x-t)|]\,dt+\lambda(x)|\tilde{f}(x)| \leqq (K/\pi)\|f\|_1+|\tilde{f}(x)|.$$

Так как $\tilde{f}(x)$ суммируема на $[-\delta'', \delta'']$, то отсюда следует суммируемость $\tilde{g}(x)$.

Если $h(x)=g(x)-f(x)$, то сопряженная функция $\tilde{h}(x)$ существует и непрерывна в некоторой окрестности нуля. Найдется такое $0<\delta_1<\delta$, что для почти всех $x\in(-\delta_1, \delta_1)$

$$\tilde{g}(x) \geqq \tilde{h}(0)+\tilde{f}(0)+\varepsilon/2 = \tilde{g}(0)+\varepsilon/2.$$

Следовательно, для интеграла Пуассона*) $\tilde{g}(r, x)$ суммируемой функции $\tilde{g}$ выполняется неравенство

$$\underline{\lim}_{r\to 1-0} \tilde{g}(r, 0) \geqq \tilde{g}(0)+\varepsilon/2.$$

Но из существования $\tilde{g}(0)$ следует, что

$$\lim_{r\to 1-0} \tilde{g}(r, 0) = \tilde{g}(0)$$

([5], стр. 172). Полученное противоречие доказывает лемму.

Теорема 2. *Пусть $f\in L$, и E — множество всех тех точек $x\in[-\pi, \pi]$' в которых существует сопряженная функция $\tilde{f}(x)$. Тогда:*

1) *для любого интервала $\Delta\subset[-\pi, \pi]$ и любого подмножества $E'\subset E$ с мерой $|E'|=|E|=2\pi$*

$$\sup_{x\in E'\cap\Delta} \tilde{f}(x) = \sup_{x\in E\cap\Delta} \tilde{f}(x);$$

2) *функция $\tilde{f}(x)$ существенно непрерывна в точке $x_0\in(-\pi, \pi)$ тогда и только тогда, когда $\tilde{f}(x_0)$ существует, и*

$$\lim_{x\to x_0, x\in E} \tilde{f}(x) = \tilde{f}(x_0).$$

Доказательство. Утверждение 1) непосредственно следует из леммы 2. Далее, предположим, что $x_0=0$ и $\tilde{f}(x)$ эквивалентна функции, непрерывной в нуле. Тогда, в силу утверждения 1), существует предел $\lim_{x\to 0, x\in E} \tilde{f}(x)$. Докажем существование $\tilde{f}(0)$. Согласно предположению, найдется такое $\delta>0$, что функция $\tilde{f}(x)$ существенно ограничена на интервале $(-\delta, \delta)$. Определим функцию $\lambda(x)$ так же, как в доказательстве леммы 2. Тогда, полагая $g(x)=\lambda(x)f(x)$, получим, что $\tilde{g}(x)$ существенно ограничена на $[-\pi, \pi]$. Далее, для функции $h(x)=g(x)-f(x)$ сопряженная функция $\tilde{h}(x)$ существует и непрерывна в некоторой окрестности нуля. Стало быть, существует предел

$$\lim_{x\to 0, x\in E} \tilde{g}(x) = s,$$

*) В силу теоремы Смирнова ([3], стр. 583) интеграл Пуассона функции $\tilde{g}$ совпадает с сопряженным интегралом Пуассона функции g.

и для завершения доказательства достаточно установить, что существует $\tilde{g}(0)=s$.

Положим

$$\tilde{g}(x;\ \eta) = -(1/\pi) \int_{\eta}^{\pi} \big(g(x+t)-g(x-t)\big)/(2\operatorname{tg} t/2)\,dt, \quad 0<\eta<\pi.$$

Согласно формуле М. Рисса [6] (см. также [5], стр. 467),

$$\tilde{g}(x;\ \eta) = (1/\pi^2) \int_{-\pi}^{\pi} \tilde{g}(x+t)(1/2)\operatorname{ctg}(t/2)\ln\big|(\sin(t+\eta)/2)/(\sin(t-\eta)/2)\big|\,dt.$$

При этом [6]

$$(15) \qquad (1/\pi^2) \int_{-\pi}^{\pi} (1/2)\operatorname{ctg}(t/2)\ln\big|(\sin(t+\eta)/2)/(\sin(t-\eta)/2)\big|\,dt = 1-\eta/\pi.$$

Обозначим подынтегральную функцию в левой части равенства (15) через $\varphi(t,\eta)$. Ясно, что $\varphi(t,\eta)$ неотрицательна, и для любого $t_0>0$ равномерно стремится к нулю при $\eta\to+0$ на каждом из отрезков $[+\pi, -t_0]$ и $[t_0, \pi]$.

Зададим произвольное $\varepsilon>0$. Тогда найдется $\delta_1>0$, такое, что для почти всех $x\in(-\delta_1, \delta_1)$, $|\tilde{g}(x)-s|<\varepsilon$. Учитывая (15) и ограниченность функции $\tilde{g}(x)$, получим при $\eta\to 0$

$$\tilde{g}(0;\ \eta)-s = (1/\pi^2) \int_{-\delta_1}^{\delta_1} [\tilde{g}(t)-s]\varphi(t,\eta)\,dt+o(1).$$

Следовательно, для достаточно малых $\eta>0$ (см. (15))

$$|\tilde{g}(0;\ \eta)-s| < 2\varepsilon,$$

и существует предел $\lim_{\eta\to+0}\tilde{g}(0;\ \eta)=s$. Теорема доказана.

Литература

[1] Н. Н. Лузин, *Интеграл и тригонометрический ряд*, Гостехиздат (Москва—Ленинград, 1951).

[2] А. Лебег, *Интегрирование и отыскание примитивных функций*, Гостехиздат (Москва—Ленинград, 1934).

[3] Н. К. Бари, *Тригонометрические ряды*, Физматгиз (Москва, 1961).

[4] E. C. Titchmarsh, On conjugate functions, *Proc. London Math. Soc.*, **29** (1929), 49—80.

[5] А. Зигмунд, *Тригонометрические ряды. I*, Мир (Москва, 1965).

[6] M. Riesz, Sur les fonctions conjuguées, *Math. Z.*, **27** (1927), 218—244.

ОДЕССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМ. И. И. МЕЧНИКОВА
МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ, КАФЕДРА МАТЕМАТИЧЕСКОГО АНАЛИЗА
УЛ. П. ВЕЛИКОГО 2
270 000 ОДЕССА, СССР

Acta Sci. Math. 47 (1984), 413—417

Some discrete inequalities of Opial's type

GRADIMIR V. MILOVANOVIĆ and IGOR Ž. MILOVANOVIĆ

1. Introduction

Let us given an index set $I=\{1, 2, \ldots, n\}$ and weight sequences $\mathbf{r}=(r_k)_{k\in I}=$ $=(r_1, \ldots, r_n)$ and $\mathbf{p}=(p_k)_{k\in I}=(p_1, \ldots, p_n)$. For a sequence $\mathbf{x}=(x_k)_{k\in I}=(x_1, \ldots, x_n)$

(1) $$\|\mathbf{x}\|_{\mathbf{r}} = \Big(\sum_{k=1}^{n} r_k x_k^2\Big)^{1/2}$$

and

(2) $$(\mathbf{x}, \nabla\mathbf{x}) = \sum_{n=1}^{n} p_k x_k \nabla x_k,$$

where the sequence $\nabla\mathbf{x}$ is given by $\nabla\mathbf{x}=(x_1, x_2-x_1, \ldots, x_n-x_{n-1})$. If we put $x_0=0$ and $\nabla x_k=x_k-x_{k-1}$ $(k=1, \ldots, n)$, then the sequence $\nabla\mathbf{x}$ can be expressed in the form $\nabla\mathbf{x}=(\nabla x_1, \nabla x_2, \ldots, \nabla x_n)$.

In this paper we determine the best constants A_n and B_n in the inequalities

(3) $$A_n\|\mathbf{x}\|_{\mathbf{r}}^2 \leqq (\mathbf{x}, \nabla\mathbf{x}) \leqq B_n\|\mathbf{x}\|_{\mathbf{r}}^2,$$

which are a discrete analogue of inequalities of Opial's type (see, for example, [1, pp. 154—162]). The idea for this paper came from the papers [2] and [3].

2. Main results

Theorem. *Define a sequence* $(Q_k(x))$ *of polynomials for the given weight sequences* $\mathbf{r}$ *and* $\mathbf{p}$ *using the recursive relation*

(4) $$xQ_{k-1}(x) = b_kQ_k(x)+a_kQ_{k-1}(x)+b_{k-1}Q_{k-2}(x) \quad (k = 1, 2, \ldots),$$
$$Q_0(x) = Q_0 \neq 0, \quad Q_{-1}(x) \stackrel{\text{def}}{=} 0,$$

Received February 16, 1983

where

(5) $a_k=(p_k/r_k)\ (k=1,\dots,n)$ *and* $b_k=-(p_{k+1}/2\sqrt{r_kr_{k+1}})\ (k=1,\dots,n-1)$.

For each sequence $\mathbf{x}=(x_k)_{k\in I}$ *of real numbers the inequalities* (3) *hold, where* A_n *and* B_n *are the minimum and the maximum zeros of polynomial* $Q_n(x)$, *respectively.*

Equality holds in the left-hand (right-hand) inequality in (3) *if and only if* $x_k=(C/\sqrt{r_k})Q_{k-1}(\lambda)$ $(k=1,\dots,n)$, *where* $\lambda=A_n$ $(\lambda=B_n)$ *and C is an arbitrary real constant different from zero.*

Proof. Let X be an n-dimensional euklidean space with scalar product $(\vec{z},\vec{w})=\sum_{k=1}^{n}z_kw_k$, where $\vec{z}=[z_1,\dots,z_n]^{\mathrm{T}}$ and $\vec{w}=[w_1,\dots,w_n]^{\mathrm{T}}$. Let, further, $\mathbf{a}=(a_1,\dots,a_n)$, $\mathbf{b}=(b_1,\dots,b_{n-1})$, and define a three-diagonal matrix by

$$H_n(\mathbf{a},\mathbf{b})=\begin{bmatrix} a_1 & b_1 & 0 & \dots & 0 & 0\\ b_1 & a_2 & b_2 & & 0 & 0\\ \vdots & & & & & \\ 0 & 0 & 0 & & a_{n-1} & b_{n-1}\\ 0 & 0 & 0 & & b_{n-1} & a_n \end{bmatrix}.$$

Introducing $z_k=\sqrt{r_k}\,x_k$ $(k=1,\dots,n)$, from (1) and (2) we get

$$\|\mathbf{x}\|_r^2=\sum_{k=1}^{n}r_kx_k^2=\sum_{k=1}^{n}z_k^2=(\vec{z},\vec{z}),$$

and

$$(\mathbf{x},\nabla\mathbf{x})=\sum_{k=1}^{n}p_kx_k\nabla x_k=\sum_{k=1}^{n}(p_kz_k/\sqrt{r_k})\nabla(z_k/\sqrt{r_k})=$$

$$=(p_1z_1^2/r_1)+\sum_{k=2}^{n}(p_kz_k/r_k\sqrt{r_{k-1}})(\sqrt{r_{k-1}}\,z_k-\sqrt{r_k}\,z_{k-1}).$$

Thus by (5),

$$(\mathbf{x},\nabla\mathbf{x})=(H_n(\mathbf{a},\mathbf{b})\vec{z},\vec{z}).$$

On the other hand, let us consider the sequence $(Q_k(x))$ of polynomials defined by (4). For $k=1,2,\dots,n$, we obtain from (4) the equality

$$x\vec{v}=H_n(\mathbf{a},\mathbf{b})\vec{v}+b_nQ_n(x)\vec{e}, \tag{6}$$

where $\vec{v}=[Q_0(x),Q_1(x),\dots,Q_{n-1}(x)]^{\mathrm{T}}$ and $\vec{e}=[0,0,\dots,0,1]^{\mathrm{T}}$. Setting $x=\lambda$ in (6), we conclude: If λ is such that $Q_n(\lambda)=0$, then λ is an eigenvalue of the matrix $H_n(\mathbf{a},\mathbf{b})$ and $\vec{v}=[Q_0(\lambda),Q_1(\lambda),\dots,Q_{n-1}(\lambda)]^{\mathrm{T}}$ is the corresponding eigenvector of the matrix $H_n(\mathbf{a},\mathbf{b})$, and conversely, according to (6), if λ is an eigenvalue of the matrix $H_n(\mathbf{a},\mathbf{b})$, then $Q_n(\lambda)=0$, i.e. λ is a zero of the polynomial $Q_n(x)$.

Thus, the eigenvalues of the matrix $H_n(\mathbf{a}, \mathbf{b})$ are exactly the zeros the of polynomial $Q_n(x)$. Since $H_n(\mathbf{a}, \mathbf{b})$ is a three-diagonal matrix ($b_i^2>0$, $i=1, \dots, n-1$) all its eigenvalues λ_i ($i=1, \dots, n$) are real and distinct, and

$$A_n(\vec{z}, \vec{z}) \leqq \big(H_n(\mathbf{a}, \mathbf{b})\vec{z}, \vec{z}\big) \leqq B_n(\vec{z}, \vec{z})$$

hold, with equality for eigenvectors corresponding to the eigenvalues $A_n=\min \lambda_i$, $B_n=\max \lambda_i$.

This completes the proof of the theorem.

Corollary 1. *Let the sequences* $\mathbf{r}$ *and* $\mathbf{p}$ *be given recursively by*

$$r_{k+1} = \big(4k(k+s)/(2k+s+1)^2\big)r_k \quad (k = 1, \dots, n-1),$$
$$p_k = (2k+s-1)r_k \quad (k = 1, \dots, n),$$

with $r_1=1$ *and* $s>-1$. *Then for every sequence* $\mathbf{x}=(x_k)_{k\in I}$ *of real numbers the inequalities* (3) *hold, where* A_n *and* B_n *are the minimal and the maximal zeros of the normalized generalized Laguerre polynomials* $\bar{L}_n^s(x)=L_n^s(x)/\|L_n^s\|$. *Here*

$$L_n^s(x) = \sum_{m=0}^{n} \binom{n+s}{n-m}\big((-x)^m/m!\big) \quad \text{and} \quad \|L_n^s\| = \sqrt{\Gamma(n+s+1)/n!}.$$

Equality holds in the left-hand (right-hand) inequality in (3) *if and only if* $x_k=(C_k/\sqrt{r_k})L_{k-1}^s(\lambda)$ $(k=1, \dots, n)$, *where* $\lambda=A_n$ $(\lambda=B_n)$ *and* $C(\neq 0)$ *is an arbitrary constant.*

Proof. For the proof of this result it is enough to show that in this case (4) reduces to the recurrence relation for generalized Laguerre polynomials. Since

$$a_k = (p_k/r_k) = 2k+s-1 \quad \text{and} \quad b_k = -\big(p_{k+1}/2\sqrt{r_k r_{k+1}}\big) = -\sqrt{k(k+s)},$$

(4) becomes

$$xQ_{k-1}(x) = -\sqrt{k(k+s)}Q_k(x)+(2k+s-1)Q_{k-1}(x)-\sqrt{(k-1)(k+s-1)}Q_{k-2}(x),$$

which is the recurrence relation for normalized generalized Laguerre polynomials $\big(Q_k(x)=\bar{L}_k^s(x)\big)$.

In the special case $p_k=r_k=1$ $(k=1, \dots, n)$, we have the following result:

Corollary 2. *For every sequence* $\mathbf{x}=(x_k)_{k\in I}$ *of real numbers and for* $x_0=0$, *the inequalities*

$$(7) \qquad 2\sin^2\big(\pi/2(n+1)\big)\sum_{k=1}^{n} x_k^2 \leqq \sum_{k=1}^{n} x_k(x_k-x_{k-1}) \leqq 2\cos^2\big(\pi/2(n+1)\big)\sum_{k=1}^{n} x_k^2,$$

are valid.

Equality holds in the left-hand inequality if and only if $x_k=C\sin(k\pi/(n+1))$ $(k=1, \ldots, n)$, *where* $C=const\neq 0$, *and in the right-hand inequality if and only if* $x_k=(-1)^{k-1}C\sin(k\pi/(n+1))$, $(k=1, \ldots, n)$, *where* $C=const\neq 0$.

Proof. In this case, we have $a_k=1$, $b_k=-1/2$ and

$$xQ_{k-1}(x)=-(1/2)Q_k(x)+Q_{k-1}(x)-(1/2)Q_{k-2}(x), \tag{8}$$

where $Q_0(x)$ can be $Q_0(x)=1$. If we put $t=1-x$, one can easily obtain the solution of the difference equation (8), for example for $|t|<1$, i. e. $0<x<2$,

$$Q_k(x)=(\sin(k+1)\theta/\sin\theta) \quad (k=1, \ldots, n), \tag{9}$$

where $e^{i\theta}=t+i\sqrt{1-t^2}$. Then, from $Q_n(x)=0$ it follows $\lambda_k=2\sin^2(k\pi/2(n+1))$ $(k=1, \ldots, n)$, implying

$$A_n=\min_k \lambda_k=2\sin^2(\pi/2(n+1)) \quad \text{and} \quad B_n=\max_k \lambda_k=2\cos^2(\pi/2(n+1)).$$

Using (9) the conditions for equality are simply obtained.

Also we note that the inequalities (7) can be written in the form

$$-\cos(\pi/(n+1))\sum_{k=1}^{n} x_k^2 \leqq \sum_{k=2}^{n} x_k x_{k-1} \leqq \cos(\pi/(n+1))\sum_{k=1}^{n} x_k^2,$$

i.e.,

$$\Big|\sum_{k=2}^{n} x_k x_{k-1}\Big| \leqq \cos(\pi/(n+1))\sum_{k=1}^{n} x_k^2. \tag{10}$$

Remark. The inequality (10) is related to an extremal problem occurring in the investigation of approximative properties of positive polynomial operators. Namely, let C_m be the class of all nonnegative trigonometric polynomials of order m

$$T_m(t)=1+2a_1\cos t+\ldots+2a_m\cos mt. \tag{11}$$

The problem is to determine a polynomial $T_m^*\in C_m$ which has the greatest coefficient a_1 (see, for example, [4, pp. 113—115]). If the polynomial (11) is written in the form

$$T_m(t)=|x_1+x_2e^{it}+\ldots+x_{m+1}e^{imt}|=\sum_{k=1}^{m+1} x_k^2+2\Big(\sum_{k=2}^{m+1} x_k x_{k-1}\Big)\cos t+\ldots,$$

where x_k $(k=1, \ldots, m+1)$ are real numbers, the determination of T_m^* is reduced to finding

$$\sup a_1=\sup\sum_{k=2}^{m+1} x_k x_{k-1}, \quad \sum_{k=1}^{m+1} x_k^2=1.$$

Putting $n=m+1$ in (10), we have $\sup a_1=\cos(\pi/(m+2))$.

References

[1] D. S. MITRINOVIĆ and P. M. VASIĆ, *Analytic inequalities,* Grundlehren der Math. Wiss., Band 165, Springer-Verlag (Berlin—Heidelberg—New York, 1970).

[2] K. FAN, O. TAUSSKY and J. TODD, Discrete analogs of inequalities of Wirtinger, *Monatsh. Math.,* **59** (1955), 73—79.

[3] G. V. MILOVANOVIĆ and I. Ž. MILOVANOVIĆ, On discrete inequalities of Wirtinger's type, *J. Math. Anal. Appl.,* **88** (1982), 378—387.

[4] V. M. TIHOMIROV, *Some problems of approximation theory,* Univ. Moscow (1976). (Russian)

FACULTY OF ELECTRONIC ENGINEERING
DEPARTMENT OF MATHEMATICS
BEOGRADSKA 14
18 000 NIŠ, YUGOSLAVIA

Acta Sci. Math. 47 (1984), 419—439

On distances between unitary orbits of self-adjoint operators

EDWARD A. AZOFF[1] and CHANDLER DAVIS[2]

Dedicated to Béla Szőkefalvi-Nagy

1. Introduction

In this paper we study distances between unitary equivalence classes of self-adjoint operators. Our starting point is the following fact, observed by H. Weyl [10, Theorem 1].

Theorem 1.1. *Let A and B be self-adjoint operators acting on a finite-dimensional Hilbert space, and write $\alpha_1 \leqq \alpha_2 \leqq \ldots \leqq \alpha_n$ and $\beta_1 \leqq \beta_2 \leqq \ldots \leqq \beta_n$ for their eigenvalues, repeated according to multiplicity. Then*

$$\|A-B\| \geqq \max_j |\alpha_j - \beta_j|. \tag{1.1}$$

There are several alternate expressions for the number $\max|\alpha_j - \beta_j|$, but for now, we only want to emphasize the fact that it can be computed from the multiplicity functions α and β of A and B respectively, so we denote it by $\delta(\alpha, \beta)$. In particular, (1.1) persists if A and B are replaced by unitary transforms. In fact, if these transforms are chosen to have a common basis of eigenvectors corresponding to the ordered sets of eigenvalues in the Theorem, then equality will hold in (1.1). This leads to the following restatement of Theorem 1.1.

Theorem 1.2. *Let A and B be self-adjoint operators acting on a finite-dimensional Hilbert space, and write α, β for their multiplicity functions. Then $\delta(\alpha, \beta)$ measures the distance between the unitary equivalence classes $\mathcal{U}(A)$ and $\mathcal{U}(B)$. Moreover, there exist commuting representatives A', B' of $\mathcal{U}(A)$ and $\mathcal{U}(B)$ respectively such that $\|A'-B'\| = \delta(\alpha, \beta)$.*

Received July 20, 1983.

[1] Research supported in part by a grant from the National Science Foundation.

[2] Research supported in part by a grant from NSERC (Canada).

In seeking to generalize Theorem 1.2 to infinite-dimensional spaces, it is important to realize that unitary orbits may fail to be closed. This is both good and bad news. It is good because the distance between two unitary orbits is the same as the distance between their closures, so the invariant α which we associate with A does not have to be a complete invariant for $\mathscr{U}(A)$ but only for $\overline{\mathscr{U}(A)}$. Such an invariant already exists in the literature — it is the function which assigns to each open set of real numbers the rank of the corresponding spectral projection of A. We call this function the crude multiplicity function of A. Crude multiplicity functions have pleasant properties and it is easy to define a natural distance δ between them.

The bad news is that we can't expect unitary orbits on infinite-dimensional spaces to have closest representatives. Indeed, if B belongs to the closure of $\mathscr{U}(A)$, but not to $\mathscr{U}(A)$ itself, then the distance between $\mathscr{U}(A)$ and $\mathscr{U}(B)$ will be zero, so the representatives A' and B', mentioned in the last sentence of Theorem 1.2, cannot be found in $\mathscr{U}(A)$ and $\mathscr{U}(B)$. The main result of the paper thus reads as follows.

Theorem 1.3. *Let A and B be self-adjoint operators acting on a common Hilbert space, and write α, β for their crude multiplicity functions. Then $\delta(\alpha, \beta)$ measures the distance between $\mathscr{U}(A)$ and $\mathscr{U}(B)$. Moreover, there exist commuting operators A', B' in the closures of these orbits such that $\|A'-B'\| = \delta(\alpha, \beta)$.*

Crude multiplicity functions are studied in Section 2. Most relevant to Theorem 1.3 are definition of the distance δ between them, and the proof of the fact that the distance between $\mathscr{U}(A)$ and $\mathscr{U}(B)$ is at least $\delta(\alpha, \beta)$, but we also digress to show how crude multiplicity functions can be viewed as cardinal-valued functions and measures on **R**.

In Section 3, we study operators with finite spectra. These have closed unitary orbits, and a slight generalization of a combinatorial result known as the Marriage Theorem is used to show that they satisfy the conclusion of Theorem 1.2. A redistribution of spectral measures argument is then employed to establish the first assertion of Theorem 1.3 for arbitrary operators.

Section 4 opens by introducing the notion of a monotone pair of operators — the idea is to generalize the observation, implicit in inequality (1.1), that $\|A'-B'\|$ is minimized when eigenvectors corresponding to the smaller eigenvalues of A' are simultaneously eigenvectors for the smaller eigenvalues of B'. Monotone pairs of operators always commute, and can be simultaneously decomposed as 'monotone' direct sums of operators with smaller spectra. Such decompositions correspond to 'monotone' decompositions of crude multiplicity functions, and the technical heart of the paper, Proposition 4.5, amounts to carrying out the simultaneous decomposition of pairs of crude multiplicity functions in an efficient manner. The proof of Theorem 1.3 is completed by using Proposition 4.5 to construct A' and B'.

Section 5 shows that the operators A', B' of Theorem 1.3 can always be chosen to be diagonal. It also provides a more geometric interpretation of the earlier sections of the paper. Briefly, the idea is that the joint spectral measure of a commuting pair A', B' of operators gives rise to a crude multiplicity function ϱ on $\mathbf{R}^2$ whose 'marginals' are the crude multiplicity functions of the original operators. Whether (A', B') form a monotone pair can be read off from the support of ϱ; so can the value of $\|A'-B'\|$. The correspondence $(A', B')\to\varrho$ is many-to-one, and it is this latitude that allows the modification of the A' and B' of Theorem 1.3 to diagonal operators.

The final section of the paper discusses the prospects for generalizing Theorem 1.3 to normal operators.

It is important to note that the number

$$\max_j |\alpha_j-\beta_j| \tag{1.2}$$

appearing in Theorem 1.1 can alternatively be written

$$\min_\pi \max_j |\alpha_j-\beta_{\pi j}| \tag{1.3}$$

where π ranges over the permutations of $1, 2, \dots, n$. The equality of (1.2) and (1.3) can of course be established directly, but it also follows from Theorem 1.2 and the fact that (1.3) represents the minimal distance between commuting representatives of $\mathcal{U}(A)$ and $\mathcal{U}(B)$. Whereas Theorem 1.1 was formulated in a way altogether dependent on the order of $\mathbf{R}$, (1.3) escapes reliance on order.

Let us emphasize that the spectral distance treated in this paper is different from the Hausdorff distance between spectra; see the discussion after Proposition 2.3. Our problem, in that it concerns unitary equivalence, is also to be distinguished from the study of similarity orbits [8], with which however it has some points of contact.

Acknowledgement. The authors would like to thank Catherine Olsen, Peter Rosenthal, and Randy Tuler for helpful conversations in the early stages of this research, and R.K. Davidson for careful criticisms.

2. Crude multiplicity functions

Our first task is to assign invariants to self-adjoint operators which can be used as a basis for measuring the distance between their unitary equivalence classes. Theoretically, any complete unitary invariant would serve this purpose, but as mentioned in the Introduction, we do not need to distinguish between unitary equivalence classes, but only between their closures.

Definition 2.1. Let A be a self-adjoint Hilbert space operator with spectral measure E. The function which assigns the cardinal number $\operatorname{rank} E(V)$ to each open subset V of $\mathbf{R}$ is called the *crude multiplicity function* of A.

This concept (but not the terminology) was discovered independently by D. Hadwin [6] and by R. Gellar and L. Page [5], and both of these references show that it is a complete invariant for closures of unitary equivalence classes. We will see this shortly, but one way to understand why it works on separable spaces is to recall Weyl's result that only the essential spectrum and the multiplicities of isolated eigenvalues are preserved under all the norm limits of the unitary transforms of a self-adjoint operator — this is precisely the information stored in the crude multiplicity function of the operator. To mention a specific example, all self-adjoint operators on separable spaces whose spectra are the unit interval share a common crude multiplicity function.

Spectral measures are countably subadditive in the sense that $E(\bigcup_{n=1}^{\infty} V_n) = \bigvee_{n=1}^{\infty} E(V_n)$ for every sequence of open sets. In particular, if the $\{V_n\}$ are monotone increasing, we have $\alpha(\bigcup_{n=1}^{\infty} V_n) = \sup_n \alpha(V_n)$. Thus α enjoys the regularity property $\alpha(V) = \sup\{\alpha(W) \mid W$ is compactly contained in $V\}$. This will prove useful later.

To motivate a notion of distance between crude multiplicity functions, consider the quantity $\max|\alpha_j - \beta_j|$ of (1.1). Suppose its value is r. Then if I is any open interval in $\mathbf{R}$, and I_r is obtained by extending it r units in each direction, then there must be at least as many β_j's in I_r as there are α_j's in I. In terms of the crude multiplicity functions α and β of A and B respectively, this means $\alpha(I) \leqq \beta(I_r)$, and of course by symmetry $\beta(I) \leqq \alpha(I_r)$. The argument is reversible in the sense that if $\alpha(I) \leqq \beta(I_r)$ and $\beta(I) \leqq \alpha(I_r)$ hold for every open interval I, then $\max|\alpha_j - \beta_j| \leqq r$.

Definition 2.2. Let α and β be crude multiplicity functions. Then the *distance* between them, denoted $\delta(\alpha, \beta)$, is the infimum of the numbers $r \geqq 0$ such that $\alpha(I) \leqq \beta(I_r)$ and $\beta(I) \leqq \alpha(I_r)$ hold for all open intervals I.

Several comments are in order here. First, for each $S \subseteqq \mathbf{R}$ and $r \geqq 0$, the notation S_r refers to $\{x \in \mathbf{R} \mid |x-y| \leqq r$ for some $y \in S\}$. If S is open, or closed, or an interval, then S_r will be the same; all three parts of the converse statement fail.

The infimum in the Definition is attained. Indeed, if $\alpha(I) \leqq \beta(I_{r+1/n})$ for all open intervals I and positive integers n, then $\alpha(J) \leqq \beta(I_r)$ for each open interval J compactly contained in I. Since $\alpha(I)$ is the supremum of $\{\alpha(J)\}$ for such J, we conclude $\alpha(I) \leqq \beta(I_r)$ as desired.

The truth of the equation $\alpha(I) \leqq \beta(I_r)$ for all open intervals I implies its validity for all open sets. Indeed, given V open, then V_r is the disjoint union of open intervals of the form I_r: $V_r = \dot{\bigcup\limits_n} I_r^n$, so that $V \subseteq \dot{\bigcup\limits_n} I^n$ and $\alpha(V) \leqq \sum \alpha(I^n) \leqq \sum \beta(I_r^n) = \beta(V_r)$. This argument makes enough use of monotonicity to be specific to **R**, but the strong notion of monotonicity implicit in (1.1) is muted in Definition 2.2. This will be rectified to some extent in Section 4, and a definition of δ which is a direct analogue of the quantity $\max|\alpha_j - \beta_j|$ will be presented in Section 5.

Finally, note that if $\alpha(\mathbf{R}) \neq \beta(\mathbf{R})$, then the distance between α and β is infinite. This is appropriate since if A and B act on spaces of different dimensions, there is no way to compare their unitary equivalence classes.

Proposition 2.3. *Let A and B be self-adjoint operators and write α and β for their crude multiplicity functions. Then the distance between (the closures of) the unitary equivalence classes $\mathcal{U}(A)$ and $\mathcal{U}(B)$ is at least $\delta(\alpha, \beta)$.*

Proof. Write E and F for the spectral measures of A and B respectively and suppose $r < \delta(\alpha, \beta)$. Then there is an interval I for which $\operatorname{rank} E(I) > \operatorname{rank} F(I_r)$ or $\operatorname{rank} F(I) > \operatorname{rank} E(I_r)$. Without loss of generality, assume the former, and also that $I = (-a, a)$ is centered at the origin. Choose a unit vector x in the range of $E(I)$, but orthogonal to the range of $F(I_r)$. Then $\|Ax\| < a$ while $\|Bx\| \geqq a + r$. This means $\|A - B\| > r$. Since r is arbitrary, we have $\|A - B\| \geqq \delta(\alpha, \beta)$. Since crude multiplicity is a unitary invariant, this inequality persists when A and B are replaced by unitary transforms, and the proof is complete.

Remark. Except for notation, the inequality $\|A - B\| \geqq \delta(\alpha, \beta)$ is essentially Theorem 7(i) of [3]*.

Remark. If S and T are compact subsets of **R** (or **C**), then the *Hausdorff* distance between them is given by $\theta(S, T) = \max\{\max\limits_{x \in S} \operatorname{dist}(x, T), \max\limits_{y \in T} \operatorname{dist}(S, y)\}$. It is known, even in the infinite-dimensional normal case, that $\|A - B\| \geqq \theta(\sigma(A), \sigma(B))$ and various further developments in this direction have recently been made [7], [2]. Although we will eventually show that $\operatorname{dist}(\mathcal{U}(A), \mathcal{U}(B))$ always equals $\delta(\alpha, \beta)$, equality with $\theta(\sigma(A), \sigma(B))$ rarely occurs. For example, operators $A = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & & \\ & 0 & \\ & & 0 \end{bmatrix}$ have the same spectrum, so $\theta(\sigma(A), \sigma(B)) = 0$, but $\delta(\alpha, \beta) = 1$.

*) The second author takes this occasion to call attention to errors in his paper [3]. The statement of the elementary Lemma on page 402 is too general (the second conclusion requires the hypothesis $Q = Q^* = Q^2$); this, however, is without effect on the rest of the paper. More serious, the proof of Theorem 3 is fallacious (the construction given is correct, but it does not establish the asserted inequality). This error invalidates Theorem 4, Theorem 5 (ii), Theorem 6 (iii)—(iv), and Theorem 7 (ii).

Definition 2.4. If α is a crude multiplicity function and S an arbitrary subset of $\mathbf{R}$, then $\alpha(S)\equiv\inf\{\alpha(V)|V$ an open set containing $S\}$.

This extension of the domain of α is basically a matter of convenience, but it has some surprising consequences, which will be explored after Proposition 2.5. In the meantime, two observations should be made.

(1) If $\alpha(S)\leqq\beta(S_r)$ holds for all open intervals, we have already noted that it remains valid for all open sets, and thus it holds for all subsets of $\mathbf{R}$.

(2) If E is the spectral measure of A, then rank $E(S)$ does *not* in general coincide with $\alpha(S)$ unless S is open; for example, $\alpha\{\lambda\}$ is non-zero for any λ in the spectrum of A, but $E\{\lambda\}=0$ unless λ is an eigenvalue of A.

We now prove, as promised earlier, that α is a complete invariant for the closure of $\mathcal{U}(A)$.

Proposition 2.5. *Let A and B be self-adjoint operators with crude multiplicity functions α and β respectively. Then the following are equivalent:*

(1) *the closures of $\mathcal{U}(A)$ and $\mathcal{U}(B)$ coincide;*

(2) *the distance between $\mathcal{U}(A)$ and $\mathcal{U}(B)$ is zero;*

(3) $\alpha=\beta$;

(4) $\delta(\alpha,\beta)=0$.

Proof. The implications (1)⇔(2) and (3)⇒(4) are clear. If $\delta(\alpha,\beta)=0$, then $\alpha(I)\leqq\beta(I)\leqq\alpha(I)$ for all intervals I since the infimum in Definition 2.2 is attained. This establishes (4)⇒(3).

That (2)⇒(4) follows from Proposition 2.3.

Suppose finally that $\alpha=\beta$. Call $\lambda\in\mathbf{R}$ *dispensable* for α if there is some open interval I containing λ with $\alpha(\lambda)=\inf_{\mu\in I}\alpha(\mu)$. Every open interval contains such points. Let $\lambda_0<\lambda_1<\ldots<\lambda_n$ be a partition of an interval containing $\sigma(A)=\sigma(B)$ and consisting of dispensable points. Then $\operatorname{rank} E(\lambda_{i-1},\lambda_i]=\operatorname{rank} F(\lambda_{i-1},\lambda_i]=$ $=\alpha(\lambda_{i-1},\lambda_i)$ for $i=1,\ldots,n$. In particular $\sum_{i=1}^{n}\lambda_i E(\lambda_{i-1},\lambda_i]$ and $\sum_{i=1}^{n}\lambda_i F(\lambda_{i-1},\lambda_i]$ are unitarily equivalent. Since these sums can be taken arbitrarily close to A, B respectively, we have established (3)⇒(2).

Let α be a crude multiplicity function. By the well ordering of the cardinal numbers the infimum in Definition 2.4 is always attained. Thus if S and T are disjoint compact sets in $\mathbf{R}$, there are disjoint open sets V and W containing them with $\alpha(S\cup T)=\alpha(V\cup W)=\alpha(V)+\alpha(W)=\alpha(S)+\alpha(T)$. It follows that $\alpha(S)=\sum_{x\in S}\alpha(x)$ for every finite set S. Outer regularity is built into Definition 2.4. The next result shows that α also enjoys a strong form of inner regularity. It implies that α can be reconstructed from its restriction to the collection of singleton sets, and in the sequel we will often regard α as a function on $\mathbf{R}$.

Proposition 2.6. *For any set S, we have $\alpha(S)=\sup\{\alpha(T)|T$ a finite subset of $S\}$.*

Proof. For each $x\in S$, choose an open set V containing x with $\alpha(x)=\alpha(V)$. These open sets cover S and thus admit a countable subcover $\{V_n\}$. Writing $\{x_n\}$ for the associated points in S, we have $\alpha(S)\leqq\alpha(\sum_{n=1}^{\infty}V_n)\leqq\sum_{n=1}^{\infty}\alpha(V_n)=\sum_{n=1}^{\infty}\alpha(x_n)$. This shows $\alpha(S)\leqq\sup\{\alpha(T)|T$ a finite subset of $S\}$. The reverse inequality is obvious.

Corollary 2.7. *α is countably additive.*

Proof. $\alpha(S)=\sup\{\alpha(T)|\ T$ a countable subset of $S\}$.

We close this section with an abstract characterization of crude multiplicity functions. Recall that a cardinal-valued function α is *upper semi-continuous* if $\{\lambda|\alpha(\lambda)<c\}$ is open for each cardinal number c.

Proposition 2.8. *A cardinal-valued function α defined on $\mathbf{R}$ is a crude multiplicity function if and only if*

(1) *α is compactly supported,*

(2) *α is upper semi-continuous, and*

(3) *the points at which α takes on finite non-zero values are isolated.*

Proof. The necessity of (1) is obvious, while (2) and (3) follow from the outer regularity built into Definition 2.4, and the inner regularity proved in Proposition 2.6.

Conversely, suppose α satisfies (1), (2) and (3). For each cardinal c in the range of α, choose a countable dense subset S_c of $\alpha^{-1}(c)$. There is a diagonal operator B with the nullity of $B-\lambda I$ being c iff $\lambda\in S_c$. The crude multiplicity function β of B is defined on open sets by $\beta(V)=\sum_{c}\sum_{\lambda\in S_c\cap V}\alpha(\lambda)$. We complete the proof by showing $\alpha=\beta$. Fix $\lambda_0\in\mathbf{R}$. Since every open set V containing λ_0 contains points in $S_{\alpha(\lambda_0)}$, we have $\beta(V)\geqq\alpha(\lambda_0)$ and hence $\beta(\lambda_0)\geqq\alpha(\lambda_0)$. If $\alpha(\lambda_0)$ is finite, (3) and (2) give $\beta(\lambda_0)=\alpha(\lambda_0)$. If on the other hand, $\alpha(\lambda_0)$ is infinite, use (2) to choose a neighborhood V_0 of λ_0 with $\alpha(\lambda)\leqq\alpha(\lambda_0)$ for all $\lambda\in V_0$. Then $\beta(\lambda_0)\leqq\beta(V_0)$, where $\beta(V_0)$ is a sum of cardinal numbers, each of which appears at most countably often, and all of which are $\leqq\alpha(\lambda_0)$. Thus we have $\beta(\lambda_0)\leqq\beta(V_0)\leqq\alpha(\lambda_0)$ and so $\alpha=\beta$ is a crude multiplicity function.

A totally different proof of this proposition will be outlined in Section 4, and will play an important role in establishing Theorem 1.3. The present simpler proof will be mimicked when we prove Proposition 5.5.

3. Operators with finite spectra

The separate treatment of operators with finite spectra presented in this section is not logically necessary for the sequel but the ideas involved are sufficiently different (and simpler!) to deserve exposition.

Proposition 3.1. *The unitary orbit of every self-adjoint operator with finite spectrum is closed.*

Proof. If the spectrum of A is finite and B belongs to the closure of $\mathcal{U}(A)$, then A and B have the same crude multiplicity function. This means $\sigma(A)=\sigma(B)$, and the corresponding eigenspaces have equal dimensions. This forces B to be unitarily equivalent to A.

The following combinatorial result was referred to in the Introduction. When X is finite (so that (1) is redundant) it is the classical result known as the Marriage Theorem and variously attributed to H. Weyl, J. Egerváry, P. Hall, and G. Pólya; see [11, Thm. 25A] or [9, Lemma 3.2].

Proposition 3.2. *Let $R\subseteq X\times Y$ be a relation with domain X satisfying:*

(1) *Only finitely many subsets of Y are of the form $R(x)$ for some $x\in X$, and*

(2) *For each subset S of X, the cardinality of $R(S)$ is at least as great as the cardinality of S.*

Then there is a one-to-one function $f\colon X\to Y$ whose graph is contained in R.

Proof. We use $|\dots|$ to denote cardinality.

Case 1: X is finite. We argue inductively on $|X|$. The result is clear if $|X|=1$. To effect the inductive step, note that if $|R(S)|=|S|$ for some proper subset of X, then $R\cap(S\times Y)$ and $R\cap[(X\setminus S)\times Y\setminus R(S)]$ again satisfy the hypothesis of the Proposition; on the other hand, if $|R(S)|>|S|$ for all proper subsets of X, then we could fix $x_0\in X$, $y_0\in R(x_0)$, and apply the inductive hypothesis to $R\cap[(X\setminus\{x_0\})\times\times Y\setminus\{y_0\}]$.

Case 2: The set $R(x)$ is infinite for each $x\in X$. Write $T_1,\dots,T_n$ for the various subsets of Y of the form $R(x)$ for some $x\in X$; and set $S_i=\{x\in X|R(x)=T_i\}$. Let $\mathcal{V}$ denote the collection of infinite subsets of Y which are obtained by intersecting some of the T_j's with the complements of the remaining T_j's. Express each $V\in\mathcal{V}$ as the disjoint union $V=\bigcup_{i=1}^{n} V_i$ of n sets of equal cardinality, and set $Y_i==\bigcup_{V\in\mathcal{V}} V_i$. Then $|T_i\cap Y_i|=|T_i|$ for each i; so there is a one-to-one map $f_i\colon S_i\to\to T_i\cap Y_i$. Take f to be the union of the $\{f_i\}$; this is injecitve since the $\{Y_i\}_{i=1}^{n}$ are disjoint.

Case 3: R is arbitrary. Let $S_1=\{x\in X \mid R(x)$ is finite$\}$. Then S_1 is finite since $R(S_1)$ must be the finite union of sets of the form $R(x)$ with $x\in S_1$, and $|R(S_1)|\geqq|S_1|$. Use Case 1 to define $f_1\colon S_1\to Y$ and apply Case 2 to the relation $R\cap[(X\setminus S_1)\times \times\{Y\setminus f_1(S_1)\}]$ to obtain a one-to-one f_2 on $X\setminus S_1$. Take $f=f_1\cup f_2$.

Remark. Let $X=Y$ be the positive integers and set $R=\{(x,y)\in X\times Y|(x=1$ and $y>1)$ or $x=y>1\}$. Although $|R(S)|\geqq|S|$ for every $S\subseteq X$, this R does not contain the graph of a one-to-one function. This example, which illustrates the necessity of hypothesis (1) in Proposition 3.2, was pointed out by Randy Tuler.

We can now extend Theorem 1.2 to operators with finite spectra.

Proposition 3.3. *Let A and B be self-adjoint operators with finite spectra which act on a common Hilbert space, and write α and β for their crude multiplicity functions. Then $\delta(\alpha,\beta)$ measures the distance between $\mathcal{U}(A)$ and $\mathcal{U}(B)$. Moreover there are commuting representatives A' and B' of $\mathcal{U}(A)$ and $\mathcal{U}(B)$ respectively such that $\|A'-B'\|=\delta(\alpha,\beta)$.*

Proof. Let X and Y be orthonormal bases of eigenvectors for A and B respectively and define a relation $R\subseteq X\times Y$ by $R\equiv\{(x,y)\in X\times Y|$ the eigenvalues corresponding to x and y differ by no more than $\delta(\alpha,\beta)\}$. Then R and R^{-1} satisfy the hypotheses of Proposition 3.2, so the Schroeder—Bernstein Theorem provides a bijection $\tau\colon X\to Y$ whose graph is contained in R. Let U be the unitary operator induced by (i.e. containing) τ. Set $A'=A$ and $B'=U^{-1}BU$. Then A' and B' commute and $\operatorname{dist}(\mathcal{U}(A),\mathcal{U}(B))\leqq\|A'-B'\|\leqq\delta(\alpha,\beta)$. Since we already know $\operatorname{dist}(\mathcal{U}(A),\mathcal{U}(B))\geqq\delta(\alpha,\beta)$, the proof is complete.

Proposition 3.3 leads to a quick proof of the first assertion of Theorem 1.3.

Proposition 3.4. *Let A and B be self-adjoint operators acting on a common Hilbert space, and write α,β for their crude multiplicity functions. Then $\delta(\alpha,\beta)$ measures the distance between $\mathcal{U}(A)$ and $\mathcal{U}(B)$.*

Proof. We already know $\operatorname{dist}(\mathcal{U}(A),\mathcal{U}(B))\geqq\delta(\alpha,\beta)$. Let $\varepsilon>0$ be given. By redistribution of spectral measures, we obtain self-adjoint operators A' and B' with finite spectra which are ε-perturbations of A and B respectively. Write α',β' for the crude multiplicity functions of A', B'. Then $\operatorname{dist}(\mathcal{U}(A),\mathcal{U}(B))<\operatorname{dist}(\mathcal{U}(A'),\mathcal{U}(B'))+2\varepsilon$ and $\delta(\alpha',\beta')<\delta(\alpha,\beta)+2\varepsilon$. Since ε was arbitrary and $\operatorname{dist}(\mathcal{U}(A'),\mathcal{U}(B'))=\delta(\alpha',\beta')$ by Proposition 3.3, we conclude that $\operatorname{dist}(\mathcal{U}(A),\mathcal{U}(B))\leqq\delta(\alpha,\beta)$, and the proof is complete.

For the sake of completeness, we close the section by characterizing the self-adjoint operators whose unitary orbits are closed.

Proposition 3.5. *Let A be self-adjoint with crude multiplicity function α. Then the following are equivalent:*

(1) *The unitary orbit of A is closed;*

(2) *The spectrum of A is countable, and each $\lambda\in\sigma(A)$ has a neighborhood U with $\alpha(\{\lambda\})>\alpha(U\setminus\{\lambda\})$.*

Proof. (1)⇒(2). Suppose first that $\lambda_0\in\sigma(A)$, but that the condition does not hold at λ_0. Then for all sufficiently small neighborhoods U of λ_0 we have $\alpha(U\setminus\{\lambda_0\})\geqq\alpha(\{\lambda_0\})$. If λ_0 is an eigenvalue of A, take B to be the restriction of A to the orthogonal complement of $\operatorname{Ker}(A-\lambda_0 I)$. If λ_0 is not an eigenvalue of A, set $B=A\oplus\lambda_0 I$ where I acts on a one-dimensional space. In either case, A and B have the same crude multiplicity function, but are not unitarily equivalent. This shows that (1) implies the second part of (2).

Suppose now that α satisfies the second part of condition (2). In this case each λ in $\sigma(A)$ is an eigenvalue of A. If A is not diagonal, let B be the restriction of A to $\bigvee\{\operatorname{Ker}(A-\lambda I)|\lambda\in\sigma(A)\}$. So A and B share a common crude multiplicity function, but they are not unitarily equivalent. If, on the other hand, A is diagonal and $\sigma(A)$ is uncountable, then let μ be a non-atomic measure supported on $\sigma(A)$, and take B to be the direct sum of A with the position operator on $L^2(\mu)$. Here too, α is the crude multiplicity function of the non-unitarily-equivalent operators A and B.

(2)⇒(1). If α satisfies (2), then every operator having α as its crude multiplicity function must be diagonal; the dimensions of the various eigenspaces are completely determined by α. All such operators are unitarily equivalent.

On separable spaces, condition (2) means $\sigma(A)$ is finite. On non-separable spaces, $\sigma(A)$ may have limit points, even infinitely many limit points.

The authors thank K.R. Davidson for correcting their faulty version of this Proposition.

4. Monotonicity and commuting representatives

The following definition will enable us to adapt the notion of monotonicity implicit in Theorem 1.1 to general pairs of self-adjoint operators.

Definition 4.1. Let A, B be self-adjoint operators on a common Hilbert space with spectral measures E, F respectively. We say the pair (A, B) is *monotone* if for each pair (a, b) of real numbers, either $E(-\infty, a)\leqq F(-\infty, b)$ of $F(-\infty, b)\leqq E(-\infty, a)$.

Proposition 4.2. *Let (A, B) be a monotone pair. Then there is a non-decreasing function $\tau\colon\mathbf{R}\to\mathbf{R}$ so that $F(-\infty, \tau(a))\leqq E(-\infty, a)\leqq F(-\infty, \tau(a)]$ for all $a\in\mathbf{R}$.*

Proof. For each $a\in\mathbf{R}$, set $\tau(a)=\inf\{b\geqq-\|B\| \mid E(-\infty, a)\leqq F(-\infty, b)\}$. For $b<\tau(a)$, we have $F(-\infty, b)\leqq E(-\infty, a)$ so the double inequality follows.

Corollary 4.3. *Every monotone pair of self-adjoint operators commutes.*

Proof. Let A, B, E, and F be as in Proposition 4.2. The conclusion of that result shows that $E(-\infty, a)$ commutes with every spectral projection of B. It follows that all the spectral projections of A and B commute with each other, and hence, so do A and B.

If the diagonal entries in two diagonal matrices are simultaneously non-decreasing, then the corresponding operators form a monotone pair. The operators A' and B' of Theorem 1.2, i.e., those which make equality hold in relation (1.1), can be taken to be a monotone pair, and we will use monotone pairs to establish the final assertion of Theorem 1.3.

Definition 4.4. The equation $\alpha=\alpha_1+\alpha_2$ represents a *monotone decomposition* of the crude multiplicity function α if α_1 and α_2 are also crude multiplicity functions and there is a real number a, called a *break-point* of the decomposition, such that $\alpha_1(x)=0$ for $x>a$ while $\alpha_2(x)=0$ for $x<a$.

It is easy to construct monotone decompositions — simply start with any number a, and choose appropriate values for $\alpha_i(a)$. (Beside the obvious restriction $\alpha_1(a)+$ $+\alpha_2(a)=\alpha(a)$, we must also have $\alpha_1(a)\geqq\limsup_{x\to a^-}\alpha(x)$ and $\alpha_2(a)\geqq\limsup_{x\to a^+}\alpha(x)$ to insure that the $\{\alpha_i\}$ are crude multiplicity functions — cf. Proposition 2.8 (2)). If A_1 and A_2 are operators with crude multiplicity functions α_1 and α_2 respectively, then α is the crude multiplicity function of the direct sum $A'\equiv A_1\oplus A_2$.

In fact, repeated monotone decomposition of α could be used to construct the implementing operator A' in the first place, thereby providing a (more technically complicated) proof of Proposition 2.8. To prove Theorem 1.3, we basically need to carry out this program on the crude multiplicity functions α and β simultaneously. The following proposition tells us how to get started, and Theorem 4.13 applies it to construct a monotone pair (A', B') which will satisfy Theorem 1.3.

Proposition 4.5. *Let $\beta_1+\beta_2$ be a monotone decomposition of a crude multiplicity function β, and suppose α is another crude multiplicity function with $\delta(\alpha, \beta)=$ $=r<\infty$. Then there is a monotone decomposition $\alpha_1+\alpha_2$ of α such that $\delta(\alpha_1, \beta_1)$ and $\delta(\alpha_2, \beta_2)$ are both less than or equal to r.*

Before embarking on the proof of this result, we illustrate its usefulness by establishing a special case of Theorem 1.3. It improves on Proposition 3.3 by only requiring A to have finite spectrum.

Corollary 4.6. *Let A and B be self-adjoint operators acting on a common Hilbert space, and write α, β for their crude multiplicity functions. Suppose A has finite spectrum. Then there is an operator $B' \in \overline{\mathcal{U}(B)}$ such that (A, B') is a monotone pair and $\|A-B'\| = \delta(\alpha, \beta) = \operatorname{dist}(\mathcal{U}(A), \mathcal{U}(B))$.*

Proof. We argue inductively on the cardinality of $\sigma(A)$. If $A=\lambda I$ is a scalar multiple of the identity, then (A, B) is itself a monotone pair, and $\|A-B\| = \operatorname{dist}(\mathcal{U}(A). \mathcal{U}(B))$ since $\mathcal{U}(A)=\{A\}$.

To establish the inductive step, write $A=A_1 \oplus A_2$ by splitting off the eigenspace corresponding to the smallest eigenvalue of A. Let $\alpha=\alpha_1+\alpha_2$ be the corresponding (monotone) decomposition of α, and decompose $\beta=\beta_1+\beta_2$ via Proposition 4.5. Choose operators B_1 and B_2 having these crude multiplicity functions. By the inductive hypothesis, it is possible to have $\|A_i-B_i'\| = \delta(\alpha_i, \beta_i)$ with (A_i, B_i') monotone pairs. Then $B'=B_1' \oplus B_2'$ satisfies the conclusion of the corollary.

We now work toward a proof of Proposition 4.5. Until this is completed, we will fix the notation of that proposition, i.e., α and β are crude multiplicity functions with $\delta(\alpha, \beta)=r<\infty$ and $\beta=\beta_1+\beta_2$ is a monotone decomposition of β. We seek a monotone decomposition $\alpha=\alpha_1+\alpha_2$ with both $\delta(\alpha_1, \beta_1) \leqq r$ and $\delta(\alpha_2, \beta_2) \leqq r$.

Consider first the problem of constructing α_1 — this must be a left restriction of α in the sense of the following definition.

Definition 4.7. Let γ_1 and γ be crude multiplicity functions, and write a for the largest x satisfying $\gamma_1(x) \neq 0$. We say γ_1 is a *left restriction* of γ and write $\gamma_1 \leqq \gamma$ if $\gamma_1(a) \leqq \gamma(a)$ and $\gamma_1(x)=\gamma(x)$ for $x<a$. The ordered pair $(a, \gamma_1(a))$ is called the *boundary point* of γ_1. *Right restrictions* are defined similarly.

If γ is understood, then γ_1 is completely determined by its boundary point. Note that $\leqq$ is a total order on the collection of left restrictions on γ; thought of in terms of boundary points, it is the usual dictionary order. Thus $\leqq$ has the least upper bound and greatest lower bound properties.

Returning to α_1, the requirement $\delta(\alpha_1, \beta_1) \leqq r$ means that α_1 must belong to the sets

$$\mathcal{S}^+ \equiv \{\gamma \leqq \alpha \mid \gamma(I) \leqq \beta_1(I_r) \text{ for all open intervals } I\},$$

and

$$\mathcal{S}^- \equiv \{\gamma \leqq \alpha \mid \beta_1(I) \leqq \gamma(I_r) \text{ for all open intervals } I\}.$$

Write α_1^+ for the supremum of $\mathcal{S}^+$. Since $\alpha_1^+(I) = \sup\{\gamma(I) \mid \gamma \in \mathcal{S}^+\}$ for every interval I, we see that α_1^+ belongs to $\mathcal{S}^+$. Similarly, $\alpha_1^- \equiv \inf \mathcal{S}^-$ belongs to $\mathcal{S}^-$. Thus $\mathcal{S}^+ \cap \mathcal{S}^- = \{\gamma \mid \alpha_1^- \leqq \gamma \leqq \alpha_1^+\}$ constitute our candidates for α_1. Lemma 4.9 shows that this set is nonempty.

Lemma 4.8. *Suppose* γ *is a left restriction of* α. *If* $\gamma(I)>\beta_1(I_r)$ *holds for* $I=(c,d)$, *then it holds for* $I=(c,\infty)$. *The same is true for the inequality* $\beta_1(I)>\gamma(I_r)$.

Proof. If $\gamma(c,d)>\beta_1(c-r,d+r)$, then β must have a break point below $d+r$, since otherwise $\gamma(c,d)\leqq\alpha(c,d)\leqq\beta(c-r,d+r)=\beta_1(c-r,d+r)$, the second inequality following from the assumption $\delta(\alpha,\beta)=r$. Thus replacing d by ∞ can only enlarge $\gamma(c,d)$ but will not change $\beta_1(c-r,d+r)$.

Similarly, the inequality $\beta_1(c,d)>\gamma(c-r,d+r)$ means that the boundary point $(a,\gamma(a))$ of γ satisfies $a<d+r$ so replacing d by ∞ leaves this intact as well.

Lemma 4.9. $\alpha_1^-\leqq\alpha_1^+$.

Proof. We argue by contradiction, assuming that $\alpha_1^->\alpha_1^+$. Then either there is a γ satisfying $\alpha_1^->\gamma>\alpha_1^+$ or α_1^- is an immediate successor of α_1^+. In the former case, set $\theta^+=\theta^-=\gamma$; in the latter, take $\theta^+=\alpha_1^-$ and $\theta^-=\alpha_1^+$. There are intervals $I=(c,\infty)$ and $J=(d,\infty)$ satisfying

$$\beta_1(I_r)<\theta^+(I) \tag{4.1}$$

and

$$\theta^-(J_r)<\beta_1(J). \tag{4.2}$$

If $|c-d|\leqq r$, we would have $I\subseteqq J_r$ and $J\subseteqq I_r$, so $\theta^-(I)\leqq\theta^-(J_r)<\beta_1(J)\leqq\leqq\beta_1(I_r)<\theta^+(I)$, a contradiction since θ^+ is at most an immediate successor of θ^-. Thus, if we assume for definiteness that $c\leqq d$, then we actually have $c<d-r$. By (4.2), there is a break point for β greater than d, so

$$\theta^+(c,d-r]\leqq\alpha(c,d-r]\leqq\beta(c-r,d]=\beta_1(c-r,d]. \tag{4.3}$$

Since θ^+ is at most an immediate successor of θ^-, we conclude from (4.2) that

$$\theta^+(d-r,\infty)\leqq\beta_1(d,\infty). \tag{4.4}$$

Adding (4.3) and (4.4), we contradict (4.1), and the proof is complete.

Of course, right restrictions of α are handled analogously to left restrictions. (The dictionary order on boundary points uses the order on **R** opposite to the usual one.) In particular, we take α_2^+ to be the maximal right restriction of α satisfying $\alpha_2^+(I)\leqq\beta_2(I_r)$ for all I and α_2^- to be the minimal right restriction of α satisfying $\beta_2(I)\leqq\alpha_2^-(I_r)$ for all I. The following analogue of Lemma 4.9 shows there are candidates for α_2.

Lemma 4.10. $\alpha_2^-\leqq\alpha_2^+$.

Proof. For each crude multiplicity function θ, write $\tilde\theta$ for its opposite, defined by $\tilde\theta(x)=\theta(-x)$. The operation $\sim$ converts right restrictions to left restrictions, so the present result is a corollary of Lemma 4.9.

We now have plenty of candidates for α_1 and α_2, but we must still choose carefully if $\alpha=\alpha_1+\alpha_2$ is to represent a monotone decomposition. Lemma 4.11 says that α_1^- and α_2^- are 'too small' to do the job; Lemma 4.12 says that α_1^+ and α_2^+ are 'too big'. We then complete the proof of Proposition 4.5 by 'interpolation'.

Lemma 4.11. *There is at most one number a such that $\alpha_1^-(a)$ and $\alpha_2^-(a)$ are simultaneously non-zero, and $\alpha_1^-(x)+\alpha_2^-(x)\leqq\alpha(x)$ for all x.*

Proof. We first show that if $\theta_i<\alpha_i^-$, then $\theta_1(x)+\theta_2(x)<\alpha(x)$ for some x. Indeed, by Lemma 4.8 (and its analogue for right restrictions), there are intervals satisfying

$$\theta_1(c-r,\infty) < \beta_1(c,\infty) \tag{4.5}$$

and

$$\theta_2(-\infty, d+r) < \beta_2(-\infty, d). \tag{4.6}$$

These inequalities force β to have a break-point between c and d. Adding them, we get

$$\theta_1(c-r,\infty)+\theta_2(-\infty,d+r) < \beta(c,d) \leqq \alpha(c-r,d+r). \tag{4.7}$$

This forces $\theta_1(x)+\theta_2(x)<\alpha(x)$ for some x, as desired.

Suppose there are three (or more) distinct numbers $a_1<a_2<a_3$ at which α_1^- and α_2^- are simultaneously non-zero. Let

$$\theta_1(x)=\begin{cases}\alpha(x) & \text{if } x\leqq a_2\\ 0 & \text{if } x>a_2\end{cases} \quad\text{and}\quad \theta_2(x)=\begin{cases}0 & \text{if } x<a_2\\ \alpha(x) & \text{if } x\geqq a_2.\end{cases}$$

Then $\theta_1<\alpha_1^-$ and $\theta_2<\alpha_2^-$ and $\theta_1(x)+\theta_2(x)\geqq\alpha(x)$ for all x. In view of the preceding paragraph, this case cannot occur.

The assumption that there are precisely two numbers $a_1<a_2$ at which α_1^- and α_2^- are both non-zero leads to the same contradiction by consideration of

$$\theta_1(x)=\begin{cases}\alpha(x) & \text{if } x\leqq a_1\\ 0 & \text{if } x>a_1,\end{cases} \qquad \theta_2(x)=\begin{cases}0 & \text{if } x<a_2\\ \alpha(x) & \text{if } x\geqq a_2.\end{cases}$$

We conclude there is at most one number at which α_1^- and α_2^- are both non-zero. If there are no such numbers, or if the number, a, satisfies $\alpha(a)$ infinite, the proof is complete. In the remaining case, i.e. $\alpha_1^-(a)$ and $\alpha_2^-(a)$ both finite, but non-zero, choose θ_1 and θ_2 to be immediate predecessors of α_1^- and α_2^- respectively. Reviewing the first paragraph of the proof, we note that the strict inequalities in (4.5), (4.6) and (4.7) all become equalities when θ_i is replaced by α_i^-. In particular all the numbers involved are finite and a must lie between $c-r$ and $d+r$. The revised (4.7) reads

$$\alpha_1^-(c-r,\infty)+\alpha_2^-(-\infty,d+r) \leqq \alpha(c-r,d+r), \tag{4.8}$$

or alternatively

$$(4.9)\quad \alpha(c-r, a)+\alpha_1^-(a)+\alpha_2^-(a)+\alpha(a, d+r) \leqq \alpha(c-r, a)+\alpha(a)+\alpha(a, d+r).$$

All numbers in this inequality are finite, and we conclude $\alpha_1^-(a)+\alpha_2^-(a)\leqq\alpha(a)$ as desired.

Lemma 4.12. $\alpha_1^+(x)+\alpha_2^+(x)\geqq\alpha(x)$ *for all* $x\in\mathbf{R}$.

Proof. We closely parallel the proof of Lemma 4.11. First, observe that if $\theta_i>\alpha_i^+$, then $\theta_1(x)+\theta_2(x)>\alpha(x)$ for some x. The relevant inequalities, replacing (4.5), (4.6) and (4.7), are:

$$(4.10)\qquad \beta_1(c-r, \infty) < \theta_1(c, \infty),$$

$$(4.11)\qquad \beta_2(-\infty, d+r) < \theta_2(-\infty, d),$$

and

$$(4.12)\qquad \alpha(c, d) \leqq \beta(c-r, d+r) < \theta_1(c, \infty)+\theta_2(-\infty, d).$$

Suppose now that $\alpha_1^+(a)+\alpha_2^+(a)<\alpha(a)$. If $\alpha(a)$ is infinite, set

$$\theta_1(x) = \begin{cases}\alpha(x) & \text{if } x \leqq a\\ 0 & \text{if } x > a,\end{cases} \qquad \theta_2(x) = \begin{cases}0 & \text{if } x < a\\ \alpha(x) & \text{if } x \geqq a\end{cases}$$

to obtain a contradiction with the preceding paragraph. On the other hand, if $\alpha(a)$ is finite, choose θ_i to be an immediate successor of α_i^+. Review of the first paragraph of the proof shows that if θ_i is replaced by α_i^+ in (4.12), we get

$$(4.13)\qquad \alpha(c, d) \leqq \alpha_1^+(c, \infty)+\alpha_2^+(-\infty, d).$$

Since a is between c and d, and the numbers in (4.13) are finite, this means $\alpha(a)\leqq \alpha_1^+(a)+\alpha_2^+(a)$.

Proof of Proposition 4.5. Lemmas 4.9 and 4.10 tell us $\alpha_i^-\leqq\alpha_i^+$. We will construct α_i such that $\alpha_i^-\leqq\alpha_i\leqq\alpha_i^+$ with $\alpha=\alpha_1+\alpha_2$ a monotone decomposition. The double inequalities force $\delta(\alpha_i, \beta_i)\leqq r$, so this will complete the proof.

We begin by choosing a break point a for our decomposition. Write $a_1=\sup\{x|\alpha_1^-(x)\neq0\}$ and $a_2=\inf\{x|\alpha_2^-(x)\neq0\}$. Lemma 4.12 shows that $a_1\leqq a_2$. We distinguish several (overlapping) cases.

Case 1: $\alpha_2^+(a_1)\neq0$. Take $a=a_1$.

Case 2: $\alpha_1^+(a_2)\neq0$. Take $a=a_2$.

Case 3: There is a number a between a_1 and a_2 such that both $\alpha_1^+(a)$ and $\alpha_2^+(a)$ are non-zero.

In all these cases, set:

$$\alpha_1(x) = \begin{cases} \alpha(x) & \text{if} \quad x < a, \\ 0 & \text{if} \quad x > a, \end{cases} \qquad \alpha_2(x) = \begin{cases} 0 & \text{if} \quad x < a \\ \alpha(x) & \text{if} \quad x > a, \end{cases}$$

and use the following recipe to define $\alpha_i(a)$:

Case A: $\alpha(a)$ is infinite. Set $\alpha_i(a)=\alpha_i^+(a)$.

Case B: $\alpha(a)$ is finite. Choose $\alpha_i(a)$ to satisfy $\alpha_i^-(a)\leqq\alpha_i(a)\leqq\alpha_i^+(a)$ and $\alpha_1(a)+ +\alpha_2(a)=\alpha(a)$. This is possible since $\alpha_1^-(a)+\alpha_2^-(a)\leqq\alpha(a)\leqq\alpha_1^+(a)+\alpha_2^+(a)$.

It is easy to check that in all these cases we have $\alpha_i^-\leqq\alpha_i\leqq\alpha_i^+$, the equation $\alpha_1+\alpha_2=\alpha$ is true, and α_1, α_2 are crude multiplicity functions by construction. There is one additional possibility not covered by Cases 1—3 above, namely when $\alpha_1^+(x)$ and $\alpha_2^+(x)$ are never simultaneously positive — but then $\alpha=\alpha_1^++\alpha_2^+$ by Lemma 4.12, so we may take $\alpha_i=\alpha_i^+$.

We are now in a position to prove the last assertion of Theorem 1.3. As mentioned earlier in the section, we will use a (necessarily commuting) monotone pair for (A', B'). In following the proof, the reader may want to keep the special cases $\alpha=\beta$ (Proposition 1.8) and α of finite support (Corollary 4.6) in mind.

Theorem 4.13. *Let α, β be crude multiplicity functions with $\delta(\alpha, \beta)<\infty$. Then there exists a monotone pair (A', B') of operators having α, β as their respective crude multiplicity functions and satisfying $\|A'-B'\|=\delta(\alpha, \beta)$.*

Proof. We first construct two families of crude multiplicity functions $\{\alpha_k\}$ and $\{\beta_k\}$ where k ranges over all finite sequences of 1's and 2's. We use the standard notations $k*j$ for the sequence k concatenated with (or followed by) j, and $|k|$ for the length of k, i.e., its number of terms. It is convenient to allow the empty sequence $k=\emptyset$ (of length zero) and to begin our construction by setting $\alpha_\emptyset=\alpha$ and $\beta_\emptyset=\beta$. We will also use the notations I_k and J_k for the support intervals of α_k and β_k respectively. (These are closed intervals whose endpoints are the smallest and largest points where α_k and β_k fail to vanish.)

Suppose α_k and β_k have been defined and $|k|$ is even. Then we choose a monotone decomposition $\alpha_k=\alpha_{k*1}+\alpha_{k*2}$ with the support intervals of α_{k*1} and α_{k*2} being at most half as long as I_k. Then we use Proposition 4.5 to construct a corresponding decomposition $\beta_n=\beta_{k*1}+\beta_{k*2}$. We proceed similarly if $|k|$ is odd, except that we first decompose β_k, controlling the lengths of J_{k*1} and J_{k*2}; and then apply Proposition 4.5 to decompose α_k.

If $\alpha_k=\beta_k=0$, take $a_k=b_k=0$; otherwise, fix points a_k and b_k in I_k and J_k respectively. For each integer n, write ε_n for the maximal length of the intervals I_k and J_k with $|k|=n$. By construction $\varepsilon_n\to 0$ as $n\to\infty$, and our application of Proposition 4.5 guarantees that $\delta(\alpha_k, \beta_k)\leqq\delta(\alpha, \beta)$ for all k. In particular $|a_k-b_k|\leqq \leqq\delta(\alpha, \beta)+2\varepsilon_{|k|}$.

Now fix a Hilbert space of dimension $\alpha(\mathbf{R})$, and construct a family $\{P_k\}$ of projections on it satisfying $\operatorname{rank} P_k=\alpha_k(\mathbf{R})$ and $P_k=P_{k*1}+P_{k*2}$ for each multi-index k. For each integer n, set

$$A_n=\sum_{|k|=n} a_k P_k \quad \text{and} \quad B_n=\sum_{|k|=n} b_k P_k.$$

Each pair (A_n, B_n) is monotone and we have $\|A_n-B_n\|\leqq\delta(\alpha,\beta)+2\varepsilon_n$ for each n. Since $a_{k*j}\in I_k$ for all j, we also have $\|A_n-A_m\|\leqq\varepsilon_n$ for $m\geqq n$. This means the sequences $\{A_n\}$ and $\{B_n\}$ converge (in norm) to operators A' and B' respectively. We have that (A', B') is a monotone pair and $\|A'-B'\|\leqq\delta(\alpha,\beta)$.

Write α^n for the crude multiplicity function of A_n. Then α^n is a 'redistribution' of α which concentrates all of $\alpha(I_k)$ at a_k whenever $|k|=n$. Thus $\delta(\alpha^n,\alpha)\leqq\varepsilon_n$. We conclude that α and β are the crude multiplicity functions of A' and B' respectively, and the proof is complete.

Remark. The construction in the proof is sufficiently general to produce all pairs (A', B') satisfying the conclusion of the Theorem, but it is difficult to predict a priori what these will be. We will see in the next section that they can always be chosen to be diagonal.

Proof of Theorem 1.3. Choose A' and B' as in Theorem 4.13. That they belong to the closures of $\mathcal{U}(A)$ and $\mathcal{U}(B)$ respectively follows from Proposition 2.5, that they commute from Corollary 4.3. Finally, $\operatorname{dist}(\mathcal{U}(A),\mathcal{U}(B))\geqq\delta(\alpha,\beta)$ by Proposition 2.3 while $\delta(\alpha,\beta)=\|A'-B'\|\geqq\operatorname{dist}(\mathcal{U}(A,\mathcal{U}(B))$ by definition of distance.

5. Diagonal representatives

In this section we introduce an additional characterization of the distance between crude multiplicity functions which is closer in spirit to the quantity $\max|\alpha_j-\beta_j|$ of Theorem 1.1. This characterization provides a geometric interpretation of monotonicity and leads to a proof of the fact that the representatives in Theorem 4.13 can be chosen to be diagonal.

Definiton 5.1. Let G be a spectral measure on $\mathbf{R}^2$. The *crude multiplicity function* of G is the function ϱ which assigns the cardinal number $\operatorname{rank} G(V)$ to each open subset V of $\mathbf{R}^2$.

As in Section 2, we extend the domain of ϱ by setting $\varrho(S)=\inf\{\varrho(V)|\ V \text{ open}, V\supseteqq S\}$ for every subset S of $\mathbf{R}^2$. The extended ϱ is countably additive and inner regular in the sense that $\varrho(S)=\sup\{\varrho(F)|\ F \text{ finite}, F\subseteqq S\}$ for $S\subseteqq\mathbf{R}^2$.

Definition 5.2. Let ϱ be a crude multiplicity function on $\mathbf{R}^2$: The *marginals* α and β of ϱ are defined by $\alpha(S)=\varrho(S\times\mathbf{R})$ and $\beta(S)=\varrho(\mathbf{R}\times S)$ for every $S\subseteqq\mathbf{R}^1$. Marginals are crude multiplicity functions (on $\mathbf{R}^1$).

Proposition 5.3. *Let A and B be commuting self-adjoint operators with spectral measures E, F, and crude multiplicity functions α, β respectively. Write G for their joint spectral measure on $\mathbf{R}^2$, and ϱ for the crude multiplicity function of G.*

(1) *The marginals of ϱ are α and β.*

(2) $\|A-B\|=\sup\{|x-y| \mid \varrho(x,y)\neq 0\}$.

(3) *The pair (A,B) is monotone iff $x_1<x_2$ and $y_1>y_2$ implies at least one of $\varrho(x_1,y_1)$, $\varrho(x_2,y_2)$ is zero.*

Proof. (1) Follows immediately from the definition.

(2) If $A=\sum_{i,j} a_i P_{ij}$ and $B=\sum_{i,j} b_j P_{ij}$ are diagonal operators, then $\|A-B\|= =\sup\{|a_i-b_j| \mid P_{ij}\neq 0\} = \sup\{|x-y| \mid \varrho(x,y)\neq 0\}$. The case of general A and B follows by redistribution of spectral measures.

(3) Suppose (A,B) is monotone, and $x_1<c<x_2$, $y_1>d>y_2$. If $E(-\infty,c)\leqq \leqq F(-\infty,d)$, then $\varrho((-\infty,c)\times(d,\infty))=0$ so $\varrho(x_1,y_1)=0$, while if $F(-\infty,d)\leqq \leqq E(-\infty,c)$, then $\varrho(x_2,y_2)=0$.

Suppose conversely ϱ is as stated in (3) and fix a, b. Then either $\varrho(x,y)=0$ for all $x<a$, $y>b$, or $\varrho(x,y)=0$ for all $x>a$, $y<b$. In the former case, we have $E(-\infty,a)\leqq F(-\infty,b)$; in the latter $F(-\infty,b)\leqq E(-\infty,a)$.

It is natural to call ϱ *monotone* if (3) of the Proposition holds — this means that the support of ϱ is a monotone relation in $\mathbf{R}^2$ in the usual sense. The number $\sup\{|x-y| \mid \varrho(x,y)\neq 0\}$ will be called the *departure* of ϱ — the smaller it is, the closer the support of ϱ is to the diagonal $x=y$.

Corollary 5.4. *Let α and β be crude multiplicity functions. The following numbers are equal:*

(1) *the distance $\delta(\alpha,\beta)$ between α and β,*

(2) *the minimum departure of all crude multiplicity functions on $\mathbf{R}^2$ having α and β as marginals,*

(3) *the minimum departure of all monotone crude multiplicity functions on $\mathbf{R}^2$ having α and β as marginals.*

Proof. By Propositions 2.3 and 2.5, we know that $\|A-B\|\geqq\delta(\alpha,\beta)$ for any operators A, B with crude multiplicity functions α, β respectively, and Theorem 4.13 tells us there is a monotone pair (A',B') with $\|A'-B'\|=\delta(\alpha,\beta)$. Application of Proposition 5.3 (2) completes the proof.

The numbers described in (2) and (3) of Corollary 5.4 are appropriate analogues of the expressions (1.3) and (1.2) of the Introduction. Indeed, let A and B be as in Theorem 1.1, and assume for simplicity that none of their eigenvalues $\alpha_1 < \ldots < \alpha_n$ or $\beta_1 < \ldots < \beta_n$ is repeated. Then the (crude) multiplicity functions α and β only take on the values 0 and 1. Every multiplicity function ϱ on $\mathbf{R}^2$ with these marginals must 'pair' the α_j's with the β_j's, i.e., there must be a permutation π so that ϱ takes on the value 1 at the points $(\alpha_j, \beta_{\pi j})$ and vanishes elsewhere. The number (2) of the Corollary is thus $\min_{\pi} \max_{j} |\alpha_j - \beta_{\pi j}|$, in agreement with (1.3). Since ϱ can only be monotone when π is the identity permutation, we also see that the expression in (3) of the Corollary reduces to $\max_{j} |\alpha_j - \beta_j|$.

The geometric appeal of Corollary 5.4 is somewhat offset by Definition 5.1, in which crude multiplicity functions on $\mathbf{R}^2$ are defined in terms of the somewhat elusive spectral measures on $\mathbf{R}^2$. The following analogue of Proposition 1.8 is intended to circumvent this problem.

Proposition 5.5. *Every crude multiplicity function on* $\mathbf{R}^2$ *is* (1) *compactly supported,* (2) *upper semi-continuous, and* (3) *vanishes in a deleted neighborhood of each point at which its value is finite. Conversely if* ϱ *is a cardinal-valued function on* $\mathbf{R}^2$ *having these properties, then there is a commuting pair* (A', B') *of diagonal operators such that* ϱ *is the crude multiplicity function of their joint spectral measure.*

Proof. The first assertion is a consequence of regularity. For the converse, suppose ϱ is a cardinal-valued function on $\mathbf{R}^2$ satisfying (1), (2) and (3). For each cardinal c, choose a countable dense subset S_c of $\varrho^{-1}(c)$. Let H be a Hilbert space of dimension $\varrho(\mathbf{R}^2)$, and choose an orthogonal supplementary family $\{P_p\}_{p \in \mathbf{R}^2}$ of projections on H such that rank $P_p = c$ iff $p \in S_c$. Define the (discrete) spectral measure G on $\mathbf{R}^2$ by $G(S) = \bigvee_{p \in S} P_p$. Then G is the joint spectral measure of the operators $A' \equiv \Sigma^{\oplus} x P_{xy}$ and $B' \equiv \Sigma^{\oplus} y P_{xy}$. Since $\operatorname{rank} G(V) = \sum_{p \in V} \operatorname{rank} P_p = \sum_{c} \sum_{\lambda \in S_c \cap V} \varrho(\lambda) = \varrho(V)$, we see ϱ is the crude multiplicity function of G, and the proof is complete.

Corollary 5.6. *The operators* (A', B') *of Theorem* 4.13 *can be chosen to be diagonal.*

Proof. Let G be the joint spectral measure for any pair of operators satisfying the conclusion of Theorem 4.13, and write ϱ for the crude multiplicity function of G. Take (A', B') to be the pair of operators associated with ϱ by the final statement of Proposition 5.5.

6. Normal operators

It is a long-standing question whether the analogue of (1.1), i.e.,

$$\|A-B\| \geqq \min_{\pi} \max_{j} |\alpha_j - \beta_{\pi j}| \tag{6.1}$$

is valid for (finite-dimensional) normal operators, and the present paper has nothing to add to the subject. For a history of the problem and a summary of known partial results, the reader should consult [1], [4].

Of course, if (6.1) turns out to be false, none of the Theorems stated in § 1 would generalize to the normal case. Even if (6.1) is valid, it is hard to imagine a normal analogue for the monotonicity notions of § 4, but it is possible to formulate a plan for generalizing the balance of the paper.

So assume (6.1) is true. There is little trouble in adapting §§ 2—3 to the normal case — it is only necessary to allow the sets V and I of Definitions 2.1 and 2.2 respectively to range over the open subsets of the plane. The proof of Proposition 2.3 would have to be changed, but it seems reasonable to assume that (6.1) would at least carry over to operators with finite spectra, and then one could apply the redistribution of spectral measures technique. The real challenge would be in proving a substitute for Proposition 4.5. The truth of the following conjecture would imply the normal analogues of Theorems 4.13 and 1.3.

Conjecture. Let $\beta=\beta_1+\beta_2$ be crude multiplicity functions on **C**, and assume $\beta_1(z)=0$ for $\mathrm{Re}\, z>0$ while $\beta_2(z)=0$ for $\mathrm{Re}\, z<0$. Then every α satisfying $\delta(\alpha, \beta)=$ $=r<\infty$ admits a decomposition $\alpha=\alpha_1+\alpha_2$ with $\delta(\alpha_i, \beta_i)\leqq r$ for $i=1, 2$.

This could perhaps be attacked via an 'exhaustion argument' similar to that used in the proof of the Hahn Decomposition Theorem for signed measures.

Bibliographical note. After our work was completed, we learned from E. C. Milner that a necessary and sufficient condition is now known for a relation between infinite sets to satisfy the conclusion of the Marriage Theorem. See R. Aharoni, C. St. J. A. Nash-Williams, S. Shelah, A general criterion for the existence of transversals, *Proc. London Mat. Soc.*, (3)**47** (1983), 43—68. However, this theorem does not seem to help in obtaining the conclusion we need in this paper (Proposition 3.2).

Note added in proof: For striking subsequent progress, see the forthcoming papers by K.R. Davidson, The distance between unitary orbits of normal operators, and The distance between unitary orbits of normal operators in the Calkin algebra.

References

[1] R. Bhatia, Ch. Davis, A. McIntosh, Perturbation of spectral subspaces and solution of linear operator equations, *Linear Algebra and Appl.*, **52/53** (1983), 45—67.

[2] R. Bouldin, Spectral approximations of a normal operator, *Proc. Amer. Math. Soc.*, **76** (1979), 279—284.

[3] Ch. Davis, Beyond the minimax principle, *Proc. Amer. Math. Soc.*, **81** (1981), 401—405.

[4] Ch. Davis, Perturbation of spectrum of normal operators and of commuting tuples, to appear.

[5] R. Gellar, L. Page, Limits of unitarily equivalent normal operators, *Duke Math. J.*, **41** (1974), 319—322.

[6] D. Hadwin, Closure of unitary classes, *Notices Amer. Math. Soc.*, **22** (1975), #720-47-20.

[7] P. R. Halmos, Spectral approximants of normal operators, *Proc. Edinburgh Math. Soc.*, **19** (1974), 51—58.

[8] D. Herrero, *Approximation of Hilbert Space Operators. I*, Pitman (Boston, 1983).

[9] D. J. Knuth, A permanent inequality, *Amer. Math. Monthly*, **88** (1981), 731—740.

[10] H. Weyl, Der asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen, *Math. Ann.*, **71** (1912), 441—479.

[11] R. J. Wilson, *Introduction to Graph Theory* (2nd ed.), Longman (London, 1979).

(E. A.)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF GEORGIA
ATHENS GA 30602, USA

(CH. D.)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TORONTO
TORONTO M5S 1A1, CANADA

Acta Sci. Math. 47 (1984), 441—444

Some characterizations of self-adjoint operators

FUAD KITTANEH

Let H be a separable, infinite dimensional complex Hilbert space, and let $B(H)$ denote the algebra of all bounded linear operators on H. For $T\in B(H)$, the absolute part of T, denoted by $|T|$, is defined as usual as the positive square root of T^*T. Each $T\in B(H)$ can be uniquely expressed as $A+iB$, where A, B are self-adjoint operators called the real part and the imaginary part respectively, denoted by $\operatorname{Re}T$ and $\operatorname{Im}T$, respectively. Note that $\operatorname{Re}T=(T+T^*)/2$ and $\operatorname{Im}T=(T-T^*)/2i$.

The following two theorems are characterization of self-adjoint and positive operators and were obtained by FONG and ISTRATESCU [1] and FONG and TSUI [2], respectively.

Theorem A. *An operator* $T\in B(H)$ *is self-adjoint if and only if* $|T|^2\leqq(\operatorname{Re}T)^2$.

Theorem B. *An operator* $T\in B(H)$ *is positive if and only if* $|T|\leqq\operatorname{Re}T$.

The purpose of this note is to generalize Theorem A as well as to present a new proof of Theorem B which may lead to further development in this direction. At the end of this paper we will give some characterization modulo C_p (the Schatten p-class) of self-adjoint operators.

Recall that $T\in B(H)$ is said to be hyponormal if $TT^*\leqq T^*T$ and in this case the spectral radius $r(T)=\|T\|$ (see [6]).

Theorem 1. *Let* $T\in B(H)$ *be hyponormal. If for some* $S\in B(H)$ *and a complex number* α, $|T|^2+\alpha(TS-ST)\leqq 0$, *then* $T=0$.

Proof. Since $r(T)=\|T\|$, there exists a sequence $\{x_n\}$ of unit vectors in H such that $(T-t)x_n\to 0$ where $|t|=\|T\|$. Now $(T^*Tx_n, x_n)+\alpha(TSx_n, x_n)-\alpha(STx_n, x_n)\leqq$ $\leqq 0$. Hence $\|Tx_n\|^2+\alpha(Sx_n, (T-t)^*x_n)-\alpha((T-t)x_n, S^*x_n)\leqq 0$. But since T is hyponormal and $\|(T-t)x_n\|\to 0$, it follows that $\|(T-t)^*x_n\|\to 0$. Letting $n\to\infty$, in the last inequality, we obtain $|t|^2\leqq 0$. Hence $T=0$ as required.

Received January 17, 1983.

Corollary 1. *If* $|T|^2 \leqq (\mathrm{Re}\,T)^2$, *then* $T = T^*$.

Proof. Let $T = A + iB$. Then $|T|^2 \leqq (\mathrm{Re}\,T)^2$ is equivalent to $B^2 + i(AB - BA) \leqq \leqq 0$. Now the corollary follows from Theorem 1.

The following proof of Theorem B was suggested to me by J. Stampfli.

Lemma 1. *Let* $T \in B(H)$ *be such that* $T = VP$ *where* V *is a contraction,* $P \geqq 0$ *and* $2P \leqq VP + PV^*$. *Let* $P = D + K$ *where* D *is diagonal and positive and* K *is arbitrary. If* $Dx = \lambda x$ *with* x *a unit vector in* H *and* $\lambda > 0$, *then* $\|(1-V)^*x\| \leqq \leqq (2/\lambda)\|Kx\|$.

Proof. Observe first that $\|(1-V)^*x\|^2 \leqq 2 - ((V+V^*)x, x)$. Now

$$\begin{aligned} 2\lambda + 2(Kx, x) &= 2(Px, x) \leqq (VPx, x) + (PV^*x, x) = \\ &= (VDx, x) + (DV^*x, x) + (VKx, x) + (KV^*x, x) = \\ &= \lambda((V+V^*)x, x) + (VKx, x) + (KV^*x, x). \end{aligned}$$

Therefore $\lambda[2 - ((V+V^*)x, x)] \leqq ((V-1)Kx, x) + (K(V^*-1)x, x)$ and so

$$2 - ((V+V^*)x, x) \leqq (2/\lambda)\,\|(V-1)^*x\|\;\|Kx\|.$$

Combining this inequality with the first inequality, we obtain

$$\|(1-V)^*x\| \leqq (2/\lambda)\,\|Kx\|$$

as required.

An alternative proof of Theorem B. Let $T = VP$ be the polar decomposition of T. Let $P = \int t\,dE(t)$ where $E(t)$ is the spectral measure of P. Fix $\alpha > 0$ and let $H_\alpha = E([\alpha, \infty))H$. If $\varepsilon > 0$ is given, then by Weyl—Von Neumann Theorem [3], $P_\alpha = D + K$ where D is diagonal and K is Hilbert—Schmidt with $\|K\|_2 < \varepsilon$ ($\|\cdot\|_2$ denotes the Hilbert—Schmidt norm). If $De_n = \lambda_n e_n$ where $\{e_n\}$ is a basis for H_α, then for any unit vector $y \in H_\alpha$, $y = \sum_{n=1}^{\infty} a_n e_n$ for some a_n with $\sum_{n=1}^{\infty} |a_n|^2 = 1$. Applying the lemma, we obtain

$$\begin{aligned} \|(1-V)^*y\| &\leqq \sum_{n=1}^{\infty} \|(1-V)^* a_n e_n\| \leqq \left(\sum_{n=1}^{\infty} |a_n|^2\right)^{1/2} \left(\sum_{n=1}^{\infty} \|(1-V)^* e_n\|^2\right)^{1/2} \leqq \\ &\leqq (2/\alpha)\left(\sum_{n=1}^{\infty} \|Ke_n\|^2\right)^{1/2} < (2/\alpha)\varepsilon. \end{aligned}$$

Since ε is arbitrary, $V = 1$ on H_α. Since $\alpha > 0$ is arbitrary we have $V = 1$ on $(\ker P)^\perp = = \overline{R(P)}$. Therefore $T = VP = P \geqq 0$ as required.

We remark that the above proof works for the following generalization of Theorem B.

Theorem 2. *If* $P\geqq 0$, V *is a contraction and* $2P\leqq VP+PV^*$, *then* $P=VP$ *and* $V|_{(\mathrm{Ker}\,P)^{\perp}}=1$.

In what follows we shall prove that if $|T|^2-(\mathrm{Re}\,T)^2\in C_p\,(p\geqq 1)$, then $T-T^*\in C_{2p}$. Recall that a compact operator C is in C_p if and only if $\|C\|_p^p=\sum_{i=1}^{\infty}s_i(C)^p<\infty$ where $s_1(C), s_2(C), \ldots$ denotes the sequence of eigenvalues of $|C|$ in decreasing order and repeated according to multiplicity. It is known (see [7]) that for $p\geqq 1$, $\|C\|_p^p\geqq \sum_{n=1}^{\infty}|(Ce_n, f_n)|^p$ for any orthonormal sets $\{e_n\}$ and $\{f_n\}$ in H. We refer to [5] or [7] for further properties of the Schatten p-classes.

Lemma 2. *Let* $T\in B(H)$ *be hyponormal. If for some* $S\in B(H)$ *and a complex number* α, $|T|^2+\alpha(TS-ST)$ *is compact, then* T *is compact.*

Proof. Let $K(H)$ denote the closed ideal of compact operators in $B(H)$, and let $\pi\colon B(H)\to B(H)/K(H)$ be the quotient map of $B(H)$ onto the Calkin algebra $B(H)/K(H)$. Therefore $|\pi(T)|^2+\alpha\big(\pi(T)\,\pi(S)-\pi(S)\,\pi(T)\big)=0$ and so by Theorem 1 we have $\pi(T)=0$, in other words, T is compact. (Recall that the Calkin algebra is a B^*-algebra and so it is representable as an operator algebra.)

Theorem 3. *Let* $T\in B(H)$ *be hyponormal. If for some* $S\in B(H)$ *and a complex number* α, $|T|^2+\alpha(TS-ST)\in C_p$ $(p\geqq 1)$, *then* $T\in C_{2p}$.

Proof. Since $C_p\subset K(H)$, we have by Lemma 2 that $T\in K(H)$. But it is known [6] that a compact hyponormal operator is diagonal, therefore $Te_n=\lambda_n e_n$ for some basis $\{e_n\}$ of H. Thus

$$\infty>\|\,|T|^2+\alpha(TS-ST)\|_p^p\geqq\sum_{n=1}^{\infty}\big|\big((|T|^2+\alpha(TS-ST)\,e_n, e_n\big)\big|^p=$$

$$=\sum_{n=1}^{\infty}\big|\|Te_n\|^2+\alpha(Se_n, T^*e_n)-\alpha(Te_n, S^*e_n)\big|^p=$$

$$=\sum_{n=1}^{\infty}\big||\lambda_n|^2+\alpha\lambda_n(Se_n, e_n)-\alpha\lambda_n(e_n, S^*e_n)\big|^p=\sum_{n=1}^{\infty}|\lambda_n|^{2p}$$

and so $T\in C_{2p}$ as required.

Corollary. *If* $|T|^2-(\mathrm{Re}T)^2\in C_p$ $(p\geqq 1)$, *then* $T-T^*\in C_{2p}$. *Hence T has a non-trivial invariant subspace.*

Proof. Observe that $|T|^2-(\mathrm{Re}T)^2=B^2+i(AB-BA)\in C_p$ and apply Theorem 3 to get $B\in C_{2p}$. The last assertion follows from Corollary 6.15 in [4] (which says that if $T-T^*\in C_p$ for some $p\geqq 1$, then T has a non-trivial invariant subspace).

Acknowledgment. I would like to thank Professor J. G. Stampfli for his kind assistance.

References

[1] C. K. Fong and V. I. Istratescu, Some characterization of Hermitian operators and related classes of operators. I, *Proc. Amer. Math. Soc.*, **76** (1979), 107—112.

[2] C. K. Fong and S. K. Tsui, A note on positive operators, *J. Operator Theory*, **5** (1981), 73—76.

[3] J. von Neumann, *Charakterisierung des spektrums eines integraloperators*, Actualities Sci. Indust., No. 229, Hermann (Paris, 1935).

[4] H. Radjavi and P. Rosenthal, *Invariant subspaces*, Springer-Verlag (Berlin, 1973).

[5] J. R. Ringrose, *Compact non-self-adjoint operators*, Van Nostrand Reinhold Co. (New York, 1971).

[6] J. G. Stampfli, Hyponormal operators, *Pacific J. Math.*, **12** (1962), 1453—1458.

[7] B. Simon, *Trace ideals and their applications*, University Press (Cambridge, 1979).

DEPARTMENT OF MATHEMATICS
UNITED ARAB EMIRATES UNIVERSITY
P. O. BOX: 15551
AL-AIN, U.A.E.

Acta Sci. Math. 47 (1984), 445—448

Normal composition operators

B. S. KOMAL and D. K. GUPTA

1. Preliminaries

Let $(X, \mathcal{S}, \lambda)$ be a σ-finite measure space and let T be a measurable non-singular $(\lambda T^{-1}(E)=0$ whenever $\lambda(E)=0)$ transformation from X into itself. Then the equation

$$C_T f = f \circ T \quad \text{for every} \quad f \in L^p(\lambda)$$

defines a composition transformation C_T from $L^p(\lambda)$ into the space of all complex valued functions on X. If the range of C_T is contained in $L^p(\lambda)$ and C_T turns out to be a bounded operator on $L^p(\lambda)$, then we call it a composition operator induced by T. Let $X=N$, the set of all (non-zero) positive integers and $\mathcal{S}=P(N)$, the power set of N. Suppose $w=\{w_n\}$ is a sequence of (non-zero) positive real numbers. Define a measure λ on $P(N)$ by

$$\lambda(E) = \sum_{n \in E} w_n \quad \text{for every} \quad E \in P(N).$$

Then for $p=2$, $L^p(\lambda)$ is a Hilbert space with the inner product defined on it by

$$\langle f, g \rangle = \sum_{n=1}^{\infty} w_n f(n) \overline{g(n)} \quad \text{for every} \quad f, g \in L^p(\lambda).$$

This space is called a weighted sequence space with weights $\{w_n : n \in N\}$ and is denoted by l_w^2. The symbol $B(l_w^2)$ denotes the Banach algebra of all bounded linear operators on l_w^2 and the symbol f_0 denotes the Radon—Nikodym derivative of the measure λT^{-1} with respect to the measure λ.

Received November 30, 1982.

2. Normal composition operators

A bounded linear operator A on a Hilbert space is called normal if it commutes with A^*. The operator A is called quasinormal if it commutes with A^*A. In [7] SINGH, KUMAR and GUPTA made the study of these operators on a weighted sequence space l_w^2 when $\sum_{n=1}^{\infty} w_n < \infty$. WHITLY [8] has studied the seoperators on $L^2(\lambda)$, when the underlying space is a finite measure space. He has proved that the class of unitary composition operators coincides with the class of normal composition operators. In our present note we have generalised this result to $L^2(\lambda)$, when the underlying measure space is a special type of σ-finite measure space. Some results on quasinormal, isometric and co-isometric composition operators are also reported.

We shall first give an example to show that a normal composition operator may not be unitary.

Example 2.1. Let $T: N \to N$ be the mapping defined by

$$T(n) = \begin{cases} 2 & \text{if } n = 1; \\ n+2 & \text{if } n \text{ is an even integer;} \\ n-2 & \text{if } n \text{ is an odd integer } > 1. \end{cases}$$

Let the sequence $\{w_n\}$ of weights be defined by

$$w_n = \begin{cases} 1 & \text{if } n = 1, \\ 1/2^n & \text{if } n \text{ is an ever integer,} \\ 2^{n-1} & \text{if } n \text{ is an odd integer } > 1. \end{cases}$$

Then $f_0(n)=4$ for every $n \in N$. Hence in view of Theorem 1 of [5] C_T is a bounded operator. Clearly $f_0(n)=f_0(T(n))$ for every $n \in N$. Since T is an injection, $T^{-1}(\mathcal{S})= =\mathcal{S}$. Hence by Lemma 2 of [8] C_T is normal. Since $C_T^* C_T = M_{f_0} = 4I$, it is clear that C_T is not unitary.

If the sequence $\{w_n\}$ is a convergent sequence of positive real numbers, then every normal operator becomes unitary. It is given in the following theorem. We shall first give a definition.

Definition. Let $T: N \to N$ be a mapping. Then two integers m and n are said to be in the same orbit of T if each can be reached from the other by composing T and T^{-1} (T^{-1} means a multivalued function) sufficiently many times.

Theorem 2.2. *Let $C_T \in B(l_w^2)$ and let $w = \{w_n\}$ be a convergent sequence of positive real numbers. Then the following are equivalent:*

(i) *C_T is unitary,*

(ii) *C_T is normal.*

Proof. The implication (i)⇒(ii) is true for any bounded operator A. We show that (ii)⇒(i). Assume that C_T is normal. Then by Lemma 2 of WHITLEY [8], $f_0=f_0\circ T$ and $T^{-1}(\mathcal{S})=\mathcal{S}$. From Lemma 1 of WHITLEY [8], C_T has dense range and hence C_T is onto in view of the normality of C_T. Thus by Corollary 2.3 and Corollary 2.5 of SINGH and KUMAR [6] T is invertible. Let $n_i\in T^{-1}(\{n\})$. Then $f_0(n_i)=f_0(T(n_i))=$ $=f_0(n)$. Similarly, we can show that f_0 is constant on the orbit of n. Further let $n_0\in N$. Then $O(n_0)$, the orbit of n_0 is either a finite set or it is an infinite set. We first suppose that $O(n_0)$ is a finite set. Then there is an integer m in N such that $T^m(n_0)=n_0$. Now $f_0(T(n))=f_0(T^2(n_0))=\ldots=f_0(T^m(n_0))$. Equivalently,

$$\frac{w_{n_0}}{w_{n_1}}=\frac{w_{n_1}}{w_{n_2}}=\ldots=\frac{w_{n_{m-1}}}{w_{n_m}}=\beta \quad \text{(say)}$$

where $T^k(n_0)=n_k$ for $k\leqq m$, and $n_m=n_0$. From this we get $w_{n_k}=w_{n_0}/\beta^k$ for $k\leqq m$ and hence $\beta^m=1$ which implies that $\beta=1$. Next, if $O(n_0)$ is an infinite subset of N, then $T^k(n_0)\neq n_0$ for every $k\in N$. Let $(T^k)^{-1}(n_0)=n_{-k}$. Then f_0 is constant on the set $\{n_k: k\in Z\}$, where Z is the set of all integers. Thus as shown in the first case $w_{n_k}=w_{n_0}/\beta^k$ (i) and $w_{n_{-k}}=\beta^k w_{n_0}$ (ii). If $\beta\neq 1$, then at least one of the subsequences (i) and (ii) is divergent. This contradicts the fact that every subsequence of a convergent sequence is convergent. Hence $\beta=1$. Thus $f_0(n)=1$ for every $n\in N$. This implies that C_T is an isometry by [3]. Since C_T has dense range, we can conclude that C_T is unitary.

Theorem 2.3. *Let $C_T\in B(l^2_w)$. Then C_T^* is an isometry if and only if $w=w\circ T$ and T is an injection.*

Proof. Suppose C_T is a co-isometry. Then $C_TC_T^*e'_n=e'_n$, where $e'_n=X_{\{n\}}/w_n$, $X_{\{n\}}$ being the sequence each terms of which is 0, except for the nth one which equals 1. Using the definition of C_T^* [5], we get $C_Te'_{T(n)}=e'_n$. This implies that

$$X_{T^{-1}(\{T(n)\})}/w_{T(n)}=X_{\{n\}}/w_n.$$

Hence we can conclude that T is an injection and $w=w\circ T$.

Conversely, if $w=w\circ T$ i.e. $w_n=w_{T(n)}$ for every $n\in N$ and T is an injection then $C_TC_T^*e'_n=e'_n$. Let $f\in l^2_w$. Then

$$(C_TC_T^*f)(n)=\langle C_TC_T^*f, e'_n\rangle=\langle f, C_TC_T^*e'_n\rangle=\langle f, e'_n\rangle=f(n)$$

for every $n\in N$. Hence $C_TC_T^*f=f$ for every $f\in l^2_w$. This completes the proof of the theorem.

Theorem 2.4. *Let $T\colon N\to N$ be an injection such that $C_T\in B(l^2_w)$. Then the following are equivalent:*

(i) *C_T^* is an isometry,*

(ii) *C_T is a partial isometry,*

(iii) $w=w\circ T$.

Proof. (i) implies (ii): Suppose C_T^* is an isometry. Then C_T^* is a partial isometry. Hence C_T is a partial isometry [1, p. 96]. (ii) implies (iii): If C_T is a partial isometry, then from a corollary to Problem 98 of [2] $C_T C_T^* C_T = C_T$. Since $C_T^* C_T = M_{f_0}$, this implies that $M_{f_0 \circ T} C_T = C_T$. Thus $f_0 \circ T = 1$ on the range of C_T. Now T is an injection, therefore by Corollary 2.4 of [6] C_T has dense range. Hence $(f_0 \circ T)(n) = 1$ for every $n \in N$. Thus $T^{-1}(\{T(n)\})/T(n) = 1$ for every $n \in N$ which implies that $w_n = w_{T(n)}$ for every $n \in N$. Hence $w = w \circ T$. (iii) implies (i): This proof is given in Theorem 2.3.

Whitley [8] has given an example to show that not every quasinormal composition operator is normal. We show that if T is an injection, then every quasinormal composition operator becomes normal. It is given in the following theorem.

Theorem 2.5. *Let $T: N \to N$ be an injection such that $C_T \in B(l_w^2)$. Then the following are equivalent:*

(i) *C_T is normal,*

(ii) *C_T is quasinormal,*

(iii) *C_T is an isometry.*

Proof. (i)⇒(ii) is trivial; (ii)⇒(iii) follows from Theorem 2 of [8]. (iii)⇒(i): By the assumption of the theorem, Corollary 2.4 of [6] guarantees that C_T has dense range.

References

[1] R. G. Douglas, *Banach Algebra Techniques in Operator Theory*, Academic Press Inc. (New York—London, 1972).

[2] P. R. Halmos, *A Hilbert space problem book*, Van Nostrand (Princeton, N. J., 1967).

[3] R. K. Singh, Normal and Hermitian composition operators, *Proc. Amer. Math. Soc.*, **47** (1975), 348—350.

[4] R. K. Singh, Composition operators induced by rational functions, *Proc. Amer. Math. Soc.*, **59** (1976), 329—333.

[5] R. K. Singh and D. K. Gupta, Normal and invertible composition operators, *Bull. Austral. Math. Soc.*, **18** (1978), 45—54.

[6] R. K. Singh and A. Kumar, Characterization of invertible, unitary and normal composition operators, *Bull. Austral. Math. Soc.*, **19** (1978), 81—95.

[7] R. K. Singh, A. Kumar and D. K. Gupta, Quasinormal composition operators, *Indian J. Pure Appl. Math.*, **11** (1980), 904—907.

[8] R. Whitley, Normal and quasinormal composition operators, *Proc. Amer. Math. Soc.*, **70** (1978), 114—118.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF JAMMU
JAMMU-180001, INDIA

Acta Sci. Math. 47 (1984), 449—455

Which $C_{\cdot 0}$ contraction is quasi-similar to its Jordan model?

PEI YUAN WU*

Dedicated to Professor Béla Szőkefalvi-Nagy on his 71st birthday

For certain $C_{\cdot 0}$ contractions on a Hilbert space, a Jordan model has been obtained by B. SZ.-NAGY [3] (cf. also [5]). It was shown that a $C_{\cdot 0}$ contraction T with the defect index $d_T = \operatorname{rank}(I-T^*T)^{1/2}$ finite is completely injection-similar to a unique *Jordan operator* of the form $J=S(\varphi_1)\oplus\ldots\oplus S(\varphi_k)\oplus S_l$, where φ_j's are non-constant inner functions satisfying $\varphi_j|\varphi_{j-1}$, $S(\varphi_j)$ denotes the compression of the unilateral shift $S(\varphi_j)f=P_j(e^{it}f)$ for $f\in H^2\ominus\varphi_jH^2$, P_j being the (orthogonal) projection onto $H^2\ominus\varphi_jH^2$, $j=1,\ldots,k$, and S_l denotes the unilateral shift on H_l^2. Moreover, if $n=d_T$ and $m=d_{T^*}=\operatorname{rank}(I-TT^*)^{1/2}$, then $k\leqq n$ and $l=m-n$. It is known that in general T is not quasi-similar to J even when $m<\infty$. (For an example, see [5], pp. 321—322.) In this paper, we find necessary and sufficient conditions under which they are quasi-similar at least in the case when both defect indices of T are finite. The main result (Theorem 2 below) is a generalization of the corresponding result for C_{10} contractions (cf. [13], Lemma 1). We also obtain other auxiliary results concerning the invariant subspaces and approximate decompositions of $C_{\cdot 0}$ contractions. Our treatments of contractions will be based on the dilation theory developed by B. Sz.-Nagy and C. Foiaş. The main reference is their book [4].

Recall that for operators T_1 and T_2 on H_1 and H_2, respectively, $T_1\overset{i}{\prec}T_2$ (resp. $T_1\overset{d}{\prec}T_2$) denotes that there exists an operator $X\colon H_1\to H_2$ which is injective (resp. has dense range) such that $T_2X=XT_1$. If X is both injective and with dense range (called *quasi-affinity*), then we denote this by $T_1\prec T_2$. T_1 is *quasi-similar* to T_2 ($T_1\sim T_2$) if $T_1\prec T_2$ and $T_2\prec T_1$. $T_1\overset{ci}{\prec}T_2$ denotes that there exists a family of injections $\{X_\alpha\}$ such that $T_2X_\alpha=X_\alpha T_1$ for each α and $\bigvee_\alpha X_\alpha H_1=H_2$. T_1 is *completely injection-similar* to T_2 ($T_1\overset{ci}{\sim}T_2$) if $T_1\overset{ci}{\prec}T_2$ and $T_2\overset{ci}{\prec}T_1$.

Received November 15, 1982.

* This research was partially supported by National Science Council of Taiwan, China.

We start by proving the following preliminary lemma.

Lemma 1. *Let T and S be $C_{\cdot 0}$ contractions with finite defect indices on H and K, respectively. Let $T=\begin{bmatrix} T_1 & * \\ 0 & T_2 \end{bmatrix}$ on $H=H_1\oplus H_2$ and $S=\begin{bmatrix} S_1 & * \\ 0 & S_2 \end{bmatrix}$ on $K=K_1\oplus K_2$ be the triangulations of type $\begin{bmatrix} C_{0\cdot} & * \\ 0 & C_{1\cdot} \end{bmatrix}$. If $T\sim S$, then $T_1\sim S_1$ and $T_2\sim S_2$.*

Proof. Let $X\colon H\to K$ be a quasi-affinity which intertwines T and S. Since $H_1=\{x\in H\colon T^n x\to 0 \text{ as } n\to\infty\}$ and $K_1=\{y\in K\colon S^n y\to 0 \text{ as } n\to\infty\}$, it is easily seen that $XH_1\subseteq K_1$. Hence X can be triangulated as $X=\begin{bmatrix} X_1 & * \\ 0 & X_2 \end{bmatrix}$. Note that X_1 is an injection which intertwines T_1 and S_1. Thus $T_1\overset{i}{\prec}S_1$. On the other hand, X_2 has dense range and intertwines T_2 and S_2 whence $T_2\overset{d}{\prec}S_2$. Similarly, from $S\prec T$ we infer that $S_1\overset{i}{\prec}T_1$ and $S_2\overset{d}{\prec}T_2$. Hence $T_1\sim S_1$ and $T_2\sim S_2$ as asserted (cf. [6], Theorem 1 (a) and [11], Theorem 2.11).

Now we are ready to prove our main result.

Theorem 2. *Let T be a $C_{\cdot 0}$ contraction with finite defect indices on H and let $T=\begin{bmatrix} T_1 & * \\ 0 & T_2 \end{bmatrix}$ on $H=H_1\oplus H_2$ be the triangulation of type $\begin{bmatrix} C_{0\cdot} & * \\ 0 & C_{1\cdot} \end{bmatrix}$. Then the following statements are equivalent:*

(1) T is quasi-similar to its Jordan model;

(2) T_2 is quasi-similar to a unilateral shift;

(3) there exists a bounded analytic function Ω such that $\Omega\Theta_{*e}=\delta I$ for some outer function δ, where Θ_{*e} is the $*$-outer factor of the characteristic function Θ_T of T.

Moreover, under these conditions we have $T\sim T_1\oplus S_{m-n}$ $(m=d_{T^*},\ n=d_T)$ and $T\sim T_1\oplus T_2$ and there exist quasi-affinities $X\colon H\to H_1\oplus H^2_{m-n}$ and $Y\colon H_1\oplus H^2_{m-n}\to H$ intertwining T and $T_1\oplus S_{m-n}$ and quasi-affinities $Z\colon H\to H_1\oplus H_2$ and $W\colon H_1\oplus H_2\to H$ intertwining T and $T_1\oplus T_2$ such that $XY=\delta(T_1\oplus S_{m-n})$, $YX=\delta(T)$, $ZW=\delta(T_1\oplus T_2)$ and $WZ=\delta(T)$.

Proof. (1)$\Rightarrow$(2). Let $J=J_1\oplus J_2$ be the Jordan model of T, where $J_1=S(\varphi_1)\oplus\ldots\oplus S(\varphi_k)$ and $J_2=S_{m-n}$. Certainly, $J=\begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix}$ is the triangulation of type $\begin{bmatrix} C_{0\cdot} & * \\ 0 & C_{1\cdot} \end{bmatrix}$. By Lemma 1, $T\sim J$ implies $T_1\sim J_1$ and $T_2\sim J_2=S_{m-n}$.

(2)$\Rightarrow$(3). Let $\Theta_T=\Theta_{*e}\Theta_{*i}$ be the $*$-canonical factorization of Θ_T. Then the characteristic function of T_2 coincides with the purely contractive part Θ^0_{*e} of Θ_{*e}. By [13], Lemma 1, there exists a bounded analytic function Ω^0 and an outer function δ^0 such that $\Omega^0\Theta^0_{*e}=\delta^0 I$. Condition (3) follows immediately.

(3)⇒(1). Note that Ω must be an outer function since $\overline{\Omega H_m^2} \supseteq \overline{\Omega \Theta_{*e} H_n^2} = \overline{\delta H_n^2} = H_n^2$ implies that $\overline{\Omega H_m^2} = H_n^2$. Consider the operator Ω_+ from H_m^2 to H_n^2 defined by $\Omega_+ f = \Omega f$ for $f \in H_m^2$. Let $K = \ker \Omega_+$. Then K is an invariant subspace for S_m, the unilateral shift on H_m^2. It follows that $K = \Phi H_l^2$ for some inner function Φ, where $0 \leqq l \leqq m$. We consider the functional model of T, that is, consider T as the operator defined on $H = H_m^2 \ominus \Theta_T H_n^2$ by $Tf = P(e^{it} f)$ for $f \in H$, where P denotes the (orthogonal) projection onto H. Similarly, consider T_1 as defined on $H_1 = H_n^2 \ominus \Theta_{*i} H_n^2$ by $T_1 g = P_1(e^{it} g)$ for $g \in H_1$, where P_1 denotes the (orthogonal) projection onto H_1. (Here T_1 is unitarily equivalent to the $C_{0\cdot}$ part of T.) Now define $X \colon H \to H_1 \oplus H_l^2$ by

$$Xf = P_1(\Omega f) \oplus \Phi^*(\delta f - \Theta_{*e} \Omega f) \quad \text{for } f \in H.$$

Note that

$$\Omega(\delta f - \Theta_{*e} \Omega f) = \delta \Omega f - \Omega \Theta_{*e} \Omega f = \delta \Omega f - \delta \Omega f = 0.$$

Hence $\delta f - \Theta_{*e} \Omega f$ is in $\ker \Omega_+ = K = \Phi H_l^2$. Thus $\Phi^*(\delta f - \Theta_{*e} \Omega f)$ is indeed in H_l^2. Next define $Y \colon H_1 \oplus H_l^2 \to H$ by

$$Y(g \oplus h) = P(\Theta_{*e} g + \Phi h) \quad \text{for} \quad g \oplus h \in H_1 \oplus H_l^2.$$

It is easily verified that X and Y intertwine T and $T_1 \oplus S_l$. Moreover, for $g \oplus h \in H_1 \oplus H_l^2$ we have

$$XY(g \oplus h) = XP(\Theta_{*e} g + \Phi h) = X(\Theta_{*e} g + \Phi h - \Theta_T u) =$$
$$= P_1(\Omega \Theta_{*e} g + \Omega \Phi h - \Omega \Theta_T u) \oplus \Phi^*(\delta \Theta_{*e} g + \delta \Phi h - \delta \Theta_T u - \Theta_{*e} \Omega \Theta_{*e} g - \Theta_{*e} \Omega \Phi h +$$
$$+ \Theta_{*e} \Omega \Theta_T u) = P_1(\delta g - \delta \Theta_{*i} u) \oplus \Phi^*(\delta \Phi h) = P_1(\delta g) \oplus \delta h = \delta(T_1 \oplus S_l)(g \oplus h),$$

where $u \in H_n^2$. On the other hand, for $f \in H$ we have

$$YXf = Y[P_1(\Omega f) \oplus \Phi^*(\delta f - \Theta_{*e} \Omega f)] = Y[(\Omega f - \Theta_{*i} v) \oplus \Phi^*(\delta f - \Theta_{*e} \Omega f)] =$$
$$= P[\Theta_{*e} \Omega f - \Theta_{*e} \Theta_{*i} v + \Phi \Phi^*(\delta f - \Theta_{*e} \Omega f)] = P(\Theta_{*e} \Omega f - \Theta_T v + \delta f - \Theta_{*e} \Omega f) =$$
$$= P(\delta f) = \delta(T) f,$$

where $v \in H_n^2$ and we made use of the fact that $\Phi \Phi^* w = w$ for $w \in \Phi H_l^2$ to simplify the expression. That δ is outer implies that both $\delta(T_1 \oplus S_l)$ and $\delta(T)$ are quasi-affinities. We conclude that so are X and Y. Thus $T \sim T_1 \oplus S_l$. As before, let $J = J_1 \oplus J_2$ be the Jordan model of T. Then J_1 is the Jordan model of T_1 (cf. [11], Lemma 2.7). From $T_1 \sim J_1$, we infer that $T \sim J_1 \oplus S_l$. If follows from the uniqueness of the Jordan model of T that $l = m - n$ (cf. [5], Theorem 3) and therefore $T \sim J_1 \oplus S_{m-n} = J_1 \oplus J_2 = J$.

From the proof above and the proof of (2)⇒(1) in [13], Lemma 1, we may deduce that $T \sim T_1 \oplus T_2$ and there are intertwining quasi-affinities Z' and W' such that $Z'W'=\delta^2(T_1\oplus T_2)$ and $W'Z'=\delta^2(T)$. In the following, we show that actually quasi-affinities Z and W can be found for which $ZW=\delta(T_1\oplus T_2)$ and $WZ=\delta(T)$.

As before, consider the functional model of T. Then $H=H_m^2\ominus\Theta_T H_n^2$, $H_1=\Theta_{*e}H_n^2\ominus\Theta_T H_n^2$ and $H_2=H_m^2\ominus\Theta_{*e}H_n^2$. Assume that T has the triangulation $T=\begin{bmatrix}T_1 & R\\ 0 & T_2\end{bmatrix}$ on the decomposition $H=H_1\oplus H_2$. Define $S\colon H_2\to H_1$ by

$$Sf=P(\Theta_{*e}\Omega f)\quad\text{for}\quad f\in H_2,$$

where P denotes the (orthogonal) projection onto H. We first check that $T_1S-ST_2=\delta(T_1)R$. For $f\in H_2$, assume that $T_2f=e^{it}f-\Theta_T u-\Theta_{*e}v$ and $Rf=\Theta_{*e}v$ for some $u, v\in H_n^2$. Then

$$\begin{aligned}(T_1S-ST_2)f&=T_1P(\Theta_{*e}\Omega f)-S(e^{it}f-\Theta_T u-\Theta_{*e}v)=\\ &=P(e^{it}\Theta_{*e}\Omega f)-P(\Theta_{*e}\Omega e^{it}f-\Theta_{*e}\Omega\Theta_T u-\Theta_{*e}\Omega\Theta_{*e}v)=\\ &=P(\delta\Theta_T u-\delta\Theta_{*e}v)=P(\delta\Theta_{*e}v).\end{aligned}$$

On the other hand,

$$\delta(T_1)Rf=\delta(T_1)(\Theta_{*e}v)=P(\delta\Theta_{*e}v).$$

Hence $T_1S-ST_2=\delta(T_1)R$ as asserted. Now define $Z\colon H\to H_1\oplus H_2$ and $W\colon H_1\oplus H_2\to H$ by

$$Z=\begin{pmatrix}\delta(T_1) & S\\ 0 & 1\end{pmatrix}\quad\text{and}\quad W=\begin{pmatrix}1 & V-S\\ 0 & \delta(T_2)\end{pmatrix},$$

where V is the operator appearing in $\delta(T)=\begin{bmatrix}\delta(T_1) & V\\ 0 & \delta(T_2)\end{bmatrix}$ on $H=H_1\oplus H_2$. The proof that Z and W intertwine T and $T_1\oplus T_2$ and that $ZW=\delta(T_1\oplus T_2)$ and $WZ=\delta(T)$ follows exactly the same as the one for Theorem 2.1 in [12]. We leave the verifications to the readers. This completes the proof.

We remark that the proof of (3)⇒(1) in the preceding theorem is valid even when $d_{T^*}=\infty$. Recall that for an arbitrary operator T, Alg T, $\{T\}''$ and $\{T\}'$ denote the weakly closed algebra generated by T and I, the double commutant and the commutant of T; Lat T, Lat″ T and Hyperlat T denote the lattices of invariant subspaces, bi-invariant subspaces and hyperinvariant subspaces of T, respectively.

Corollary 3. *Let T be a $C_{\cdot 0}$ contraction with finite defect indices and let $T=\begin{bmatrix}T_1 & *\\ 0 & T_2\end{bmatrix}$ be the triangulation of type $\begin{bmatrix}C_{0\cdot} & *\\ 0 & C_{1\cdot}\end{bmatrix}$. If T is quasi-similar to its Jordan model, then* $\operatorname{Lat} T\cong\operatorname{Lat}(T_1\oplus T_2)$ *and* $\operatorname{Lat}'' T\cong\operatorname{Lat}''(T_1\oplus T_2)$.

Proof. Since a $C_{\cdot 0}$ contraction T with $d_T<\infty$ satisfies $\operatorname{Alg} T=\{T\}''$ (cf. [10], Theorem 3.2 and [9], Theorem 1), we have $\operatorname{Lat} T=\operatorname{Lat}'' T$ and $\operatorname{Lat}(T_1\oplus T_2)=\operatorname{Lat}''(T_1\oplus T_2)$. Hence we only need to prove $\operatorname{Lat} T\cong\operatorname{Lat}(T_1\oplus T_2)$. It is easily verified that the lattice isomorphisms can be implemented by the mappings $K\to\overline{ZK}$ and $L\to\overline{WL}$ for $K\in\operatorname{Lat} T$ and $L\in\operatorname{Lat}(T_1\oplus T_2)$, where Z and W are the quasi-affinities given in Theorem 2.

For the hyperinvariant subspace lattice, more is true. If T is a $C_{\cdot 0}$ contraction with $d_T<\infty$ and $T=\begin{bmatrix} T_1 & * \\ 0 & T_2 \end{bmatrix}$ is of type $\begin{bmatrix} C_{0\cdot} & * \\ 0 & C_{1\cdot} \end{bmatrix}$, then T and $T_1\oplus T_2$ have the same Jordan model (cf. [11], Lemma 2.7) whence $\operatorname{Hyperlat} T\cong\operatorname{Hyperlat}(T_1\oplus T_2)$ (cf. [8], Theorem 2). This is true even without the quasi-similarity of T to its Jordan model.

If T is as above and $K\in\operatorname{Lat} T$, then, unlike the case for the more restrictive class of C_{10} contractions (cf. [13], Corollary 4 (2)), the quasi-similarity of T to its Jordan model does not imply that $T|K$ is quasi-similar to its Jordan model. The next example suffices to illustrate this.

Example 4. Let T be the $C_{\cdot 0}$ contraction $S(uv)\oplus S$, where u is the Blaschke product with zeros $1-1/n^2$, $n=1,2,\ldots$, v is the singular inner function $v(\lambda)=\exp((\lambda+1)/(\lambda-1))$ for $|\lambda|<1$, and S is the simple unilateral shift. Then the characteristic function of T is $\Theta_T=\begin{bmatrix} uv \\ 0 \end{bmatrix}$. Let $K\in\operatorname{Lat} T$ correspond to the regular factorization $\Theta_T=\Theta_2\Theta_1$, where

$$\Theta_2=\frac{1}{\sqrt{2}}\begin{pmatrix} v & u \\ v & -u \end{pmatrix} \quad\text{and}\quad \Theta_1=\frac{1}{\sqrt{2}}\begin{pmatrix} u \\ v \end{pmatrix}.$$

Note that T is itself a Jordan operator, but $T|K$ is not quasi-similar to its Jordan model (cf. [5], pp. 321—322).

Since it is known that if T is a C_{10} contraction with finite defect indices which is quasi-similar to its Jordan model or T is a C_0 contraction, then $\operatorname{Lat} T=\operatorname{Lat}'' T=\{\overline{\operatorname{ran} S}: S\in\{T\}'\}$ (cf. [13], Corollary 8 and [1], Corollary 2.11), we may be tempted to generalize this to $C_{\cdot 0}$ contractions. As it turns out, this is in general not true. The counterexample is provided by the operator T and its invariant subspace K in Example 4. Indeed, if $K=\overline{\operatorname{ran} S}$ for some $S\in\{T\}'$, then, by the main theorem of [7], there exist bounded analytic functions $\Phi=\begin{bmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{bmatrix}$ and $\Psi=[\psi]$ such that $\Phi\Theta_T=\Theta_1\Psi$ and $H_2^2=(\Phi H_2^2+\Theta_1 H^2)^-$. From the first equation we have $\varphi_{11}v=(1/\sqrt{2})\psi$ and $\varphi_{21}u=(1/\sqrt{2})\psi$. Thus $v|\psi$ and $u|\psi$. Since $u\wedge v=1$, these imply that $uv|\psi$. Say, $\psi=uvw$ for some $w\in H^\infty$. We obtain $\varphi_{11}=(1/\sqrt{2})uw$ and $\varphi_{21}=(1/\sqrt{2})vw$. For $\begin{bmatrix} f \\ g \end{bmatrix}\in H_2^2$ and $h\in H^2$,

$$\Phi\begin{pmatrix} f \\ g \end{pmatrix}+\Theta_1 h=\begin{pmatrix} (1/\sqrt{2})uwf+\varphi_{12}g \\ (1/\sqrt{2})vwf+\varphi_{22}g \end{pmatrix}+\frac{1}{\sqrt{2}}\begin{pmatrix} uh \\ vh \end{pmatrix}=\begin{pmatrix} u & \varphi_{12} \\ v & \varphi_{22} \end{pmatrix}\begin{pmatrix} (1/\sqrt{2})(wf+h) \\ g \end{pmatrix}.$$

Since these vectors are dense in H_2^2, we conclude that $\begin{bmatrix} u & \varphi_{12} \\ v & \varphi_{22} \end{bmatrix}$, together with its determinant $u\varphi_{22}-v\varphi_{12}$, is outer (cf. [4], Corollary V.6.3). The latter contradicts the main result proved in [2].

However, in such a situation, we still have something to say.

Theorem 5. *Let T be a $C_{.0}$ contraction with $d_T<\infty$ on H. Then* $\mathrm{Lat}\, T = \mathrm{Lat}''\, T = \{S_1 H \vee S_2 H\colon S_1, S_2 \in \{T\}'\}$.

Proof. Let $K \in \mathrm{Lat}\, T$ and let $J=S(\varphi_1)\oplus\ldots\oplus S(\varphi_n)\oplus S_p$ on H_1 and $J'=S(\psi_1)\oplus\ldots\oplus S(\psi_m)\oplus S_q$ on K_1 be the Jordan models of T and $T|K$, respectively. Since $J'\overset{i}{\prec}T|K\overset{i}{\prec}T\prec J$, we infer that $m\leqq n$, $q\leqq p$ and $\psi_j|\varphi_j$ for $j=1,\ldots,m$ (cf. [5], Theorem 4). Say, $\varphi_j=\psi_j\eta_j$ for each j. Note that $S(\varphi_j)|\overline{\mathrm{ran}\,\eta_j(S(\varphi_j))}\cong S(\psi_j)$ (cf. [5], pp. 315—316). For each j, let Z_j be the operator which implements this unitary equivalence and let $Z\colon H_1\to K_1$ be the operator

$$Z_1\eta_1(S(\varphi_1))\oplus\ldots\oplus Z_m\eta_m(S(\varphi_m))\oplus\underbrace{0\oplus\ldots\oplus 0}_{n-m}\oplus P,$$

where P denotes the (orthogonal) projection from H_p^2 onto H_q^2. Then Z intertwines J and J' and has dense range. Let $X\colon H\to H_1$ be the quasi-affinity which intertwines T and J and let $Y_1, Y_2\colon K_1\to K$ be the injections which intertwine J' and $T|K$ and are such that $K=Y_1K_1\vee Y_2K_1$. Let $S_1=Y_1ZX$ and $S_2=Y_2ZX$. Then S_1 and S_2 are in $\{T\}'$ and

$$K = Y_1K_1\vee Y_2K_1 = Y_1ZH_1\vee Y_2ZH_1 = Y_1ZXH\vee Y_2ZXH = S_1H\vee S_2H.$$

This completes the proof.

It is interesting to know whether the converse of Lemma 1 is true. It may turn out that a stronger assertion is true.

Open problem: If T is a $C_{.0}$ contraction with $d_T<\infty$ and $T=\begin{bmatrix} T_1 & * \\ 0 & T_2 \end{bmatrix}$ is the triangulation of type $\begin{bmatrix} C_{0\cdot} & * \\ 0 & C_{1\cdot} \end{bmatrix}$, is $T\sim T_1\oplus T_2$?

In this respect, we have the following partial result.

Theorem 6. *If T is a $C_{.0}$ contraction with $d_T<\infty$ and $T=\begin{bmatrix} T_1 & * \\ 0 & T_2 \end{bmatrix}$ on $H=H_1\oplus H_2$ is the triangulation of type $\begin{bmatrix} C_{0\cdot} & * \\ 0 & C_{1\cdot} \end{bmatrix}$, then $T_1\oplus T_2\overset{ci}{\prec}T\prec T_1\oplus T_2$.*

Proof. Let $J=J_1\oplus J_2$ be the Jordan model of T, where $J_1=S(\varphi_1)\oplus\ldots\oplus S(\varphi_k)$ and $J_2=S_{m-n}$ ($m=d_{T^*}$, $n=d_T$). Then $J\overset{ci}{\prec}T\prec J$. Since J_1 and J_2 are the Jordan models of T_1 and T_2, respectively (cf. [11], Lemma 2.7), we have $T_1\sim J_1$ and

$J_2 \overset{ci}{\prec} T_2 \prec J_2$. It follows that $T_1 \oplus T_2 \prec J_1 \oplus J_2 \overset{ci}{\prec} T$ and $T \prec J_1 \oplus J_2 \sim T_1 \oplus J_2$. Let X be a quasi-affinity which intertwines T and $T_1 \oplus J_2$. Then it is easily verified that on the decompositions $H = H_1 \oplus H_2$ and $H_1 \oplus H^2_{m-n}$, X can be triangulated as $X = \begin{bmatrix} X_1 & X_3 \\ 0 & X_2 \end{bmatrix}$. Consider the operator $X' = \begin{bmatrix} X_1 & X_3 \\ 0 & 1 \end{bmatrix}$ on $H = H_1 \oplus H_2$. It is easily seen that X' intertwines T and $T_1 \oplus T_2$. Moreover, since T_1 is a $C_0(N)$ contraction and X_1 is an injection in $\{T_1\}'$, X_1 must be a quasi-affinity (cf. [6], Theorem 2). It follows that X' is a quasi-affinity. This shows that $T \prec T_1 \oplus T_2$, completing the proof.

We would like to thank the referee for keeping us from making a foolish mistake. The arguments before Theorem 5 are due to him.

References

[1] H. Bercovici, On the Jordan model of C_0 operators. II, *Acta Sci. Math.*, **42** (1980), 43—56.

[2] E. A. Nordgren, The ring N^+ is not adequate, *Acta Sci. Math.*, **36** (1974), 203—204.

[3] B. Sz.-Nagy, Diagonalization of matrices over H^∞, *Acta Sci. Math.*, **38** (1976), 223—238.

[4] B. Sz.-Nagy and C. Foiaş, *Harmonic analysis of operators on Hilbert space*, North Holland—Akadémiai Kiadó (Amsterdam—Budapest, 1970).

[5] B. Sz.-Nagy and C. Foiaş, Jordan model for contractions of class $C_{\cdot 0}$, *Acta Sci. Math.*, **36** (1974), 305—322.

[6] B. Sz.-Nagy and C. Foiaş, On injections, intertwining operators of class C_0, *Acta Sci. Math.*, **40** (1978), 163—167.

[7] R. I. Teodorescu, Factorisations régulières et sous-espaces invariants, *Acta Sci. Math.*, **42** (1980), 325—330.

[8] M. Uchiyama, Hyperinvariant subspaces for contractions of class $C_{\cdot 0}$, *Hokkaido Math. J.*, **6** (1977), 260—272.

[9] M. Uchiyama, Double commutants of $C_{\cdot 0}$ contractions. II, *Proc. Amer. Math. Soc.*, **74** (1979), 271—277.

[10] P. Y. Wu, Commutants of $C_0(N)$ contractions, *Acta Sci. Math.*, **38** (1976), 193—202.

[11] P. Y. Wu, $C_{\cdot 0}$ contractions: cyclic vectors, commutants and Jordan models, *J. Operator Theory*, **5** (1981), 53—62.

[12] P. Y. Wu, Approximate decompositions of certain contractions, *Acta Sci. Math.*, **44** (1982), 137—149.

[13] P. Y. Wu, When is a contraction quasi-similar to an isometry?, *Acta Sci. Math.*, **44** (1982), 151—155.

DEPARTMENT OF APPLIED MATHEMATICS
NATIONAL CHIAO TUNG UNIVERSITY
HSINCHU, TAIWAN, CHINA

Acta Sci. Math. 47 (1984), 457—459

An elementary minimax theorem

ZOLTÁN SEBESTYÉN

A recent simple proof for von Neumann's minimax theorem by I. Joó [2] urged us to formulate a minimax principle as a direct property of the function in question. In consequence our approach omits the usual convextity requirements. However, our proof is simple by applying a finite dimensional separation result concerning convex sets. In fact we use a modified version of a proof taken from BALAKRISHNAN [1]. Our theorem generalizes some of the known results of this type.

Theorem. *Let $f(x, y)$ be a real-valued function on $X \times Y$ with the following three properties:*

$$(1^x) \qquad \min_{y \in B} \sum_{x \in A} \lambda(x) f(x, y) \leqq \sup_{x \in X} \min_{y \in B} f(x, y),$$

where $A \subset X$ and $B \subset Y$ are finite subsets and $\lambda\colon A \to \mathbf{R}_+$ is a discrete probability measure on A.

$$(2^y) \qquad \inf_{y \in Y} \sup_{x \in X} f(x, y) \leqq \sup_{x \in X} \sum_{y \in B} \mu(y) f(x, y),$$

where $B \subset Y$ is a finite subset and $\mu\colon B \to \mathbf{R}_+$ is a discrete probability measure on B.

(3) *There exist $y_0 \in Y$ and $c_0 \in \mathbf{R}$, $c_0 < \inf_{y \in Y} \sup_{x \in X} f(x, y) = c^*$ such that if $D \subset \subset (c_0, \infty) \times Y$ is a subset with the property that for any $x \in X$, $f(x, y_0) \geqq c_0$, there exists $(t_x, y_x) \in D$ with $f(x, y_x) < t_x$ then there exists a finite subset in D with the same property.*

Then

$$(4) \qquad c_* = \sup_{x \in X} \inf_{y \in Y} f(x, y) = \inf_{y \in Y} \sup_{x \in X} f(x, y) = c^*.$$

Proof. Since $\inf_{y \in Y} f(x, y) \leqq f(x, y)$ holds for any $x \in X$, $y \in Y$, the inequality $c_* \leqq \sup_{x \in X} f(x, y)$ follows for any $y \in Y$, showing that $c_* \leqq c^*$. To prove (4) we start with $c_* < c^*$ and for a c, $\max\{c_*, c_0\} < c < c^*$, write $H_y = \{x \in X : f(x, y) \geqq c\}$ for

Received December 14, 1982

any $y\in Y$. Showing that some $x_0\in X$ belongs to $\bigcap\{H_y: y\in Y\}$ we get a contradiction:

$$c \leqq \inf_{y\in Y} f(x_0, y) \leqq \sup_{x\in X} \inf_{y\in Y} f(x, y) = c_*.$$

To do this let first $B=\{y_1, \dots, y_n\}$ be a finite subset in Y, and suppose that $\bigcap\{H_y: y\in B\}$ is empty. Then for any $x\in X$ there exists a $y\in B$ such that $f(x, y)<c$. As a consequence, the function $\varphi: X\to \mathbf{R}^n$, given by

$$\varphi(x) = \big(f(x, y_1)-c, \dots, f(x, y_n)-c\big)$$

has the following property: $\varphi(A)\cap\mathbf{R}^n_+=\emptyset$, where $\varphi(A)$ is the range of φ and $\mathbf{R}^n_+$ is the positive cone of vectors with nonnegative coordinates in $\mathbf{R}^n$. But then $\mathrm{Co}\varphi(A)$, the convex hull of the range of φ, does not meet $\mathrm{int}\,\mathbf{R}^n_+$, the interior of $\mathbf{R}^n_+$. There were otherwise a discrete probability measure $\lambda: X\to\mathbf{R}_+$ with finite support A, $A=\{x_1, \dots, x_m\}\subset X$, such that $c<\sum_{j=1}^{m}\lambda_j f(x_j, y_i)$ holds for any $i=1, \dots, n$. But (1^x) implies then

$$c < \min_{1\leqq i\leqq n} \sum_{j=1}^{m} \lambda_j f(x_j, y_i) \leqq \sup_{x\in X} \min_{1\leqq i\leqq n} f(x, y_i),$$

contradicting the assumption that $\bigcap\{H_y: y\in B\}$ is empty. As a result we have a nonzero separating linear functional $\mu=(\mu_1, \dots, \mu_n)\in\mathbf{R}^n$ (see e.g. [2, 2.5.1]) such that

$$\sum_{i=1}^{n} \mu_i f(x, y_i) - c\sum_{i=1}^{n}\mu_i \leqq \sum_{i=1}^{n}\mu_i t_i \quad \text{for any} \quad x\in X, \quad t=(t_1, \dots, t_n)\in\mathbf{R}^n_+.$$

In this case $\mu\in\mathbf{R}^n_+$ is obvious so that we may assume that $\sum_{i=1}^{n}\mu_i=1$ also holds. As a consequence

$$c^* = \inf_{y\in Y}\sup_{x\in X} f(x, y) \leqq \sup_{x\in X}\sum_{i=1}^{n}\mu_i f(x, y_i) \leqq c;$$

a contradiction follows by (2^y) and the choice of c. Summing up, we have proved that $\bigcap\{H_y: y\in B\}$ is nonempty for any finite subset B in Y. For $B=Y$ we get the same conclusion if we topologize X by chosing the subsets $\{x\in X: f(x, y)<t\}$ ($t\in\mathbf{R}, y\in Y$) in X as a subbase for open sets such that $\{H_y\}_{y\in Y}$ are closed sets and $\{x\in X: f(x, y_0)\geqq c_0\}$ is compact by (3). Indeed, the finite intersection property of F. Riesz implies the desired conclusion. The proof is thus complete.

Corollary. *Let $f(x, y)$ be a real-valued function on $X\times Y$ with finite X such that (1^x) (with $A=X$) and (2^y) hold. Then* (4) *also holds.*

References

[1] A. V. Balakrishnan, *Applied Functional Analysis,* Springer (1976).
[2] I. Joó, A simple proof for von Neumann's minimax theorem, *Acta Sci. Math.*, 42 (1980), 91—94.

II. DEPARTMENT OF ANALYSIS
EÖTVÖS LORÁND UNIVERSITY
MÚZEUM KRT. 6—8
1088 BUDAPEST, HUNGARY

Acta Sci. Math. 47 (1984), 461—464

Note on a theorem of Dieudonné

J. JANAS

DIEUDONNÉ [2] has proved that for any $f \in L^1(A)$, $f * L^1(A) \neq L^1(A)$, where A is a nondiscrete, locally compact abelian group. Applying Banach algebra methods we shall prove the same result for $L^1(G)$ over a compact, connected Lie group G.

Dieudonné has proved the above result by applying the methods of harmonic analysis on LCA groups. Later this theorem was proved by GOLDBERG and BURNHAM [3] by applying Banach algebra methods. We shall follow their ideas, but since in our case the algebra $L^1(G)$ is not commutative in general, the proof is much more difficult.

We start by recalling a few notions from Banach algebras. Let B be a complex Banach algebra.

Definition 1. We say that $b \in B$ is a divisor of zero, if $rb = br = 0$ for some $r \in B$, $r \neq 0$.

Definition 2. We say that $a \in B$ is a topological divisor of zero, if there exists a sequence $\{g_n\} \subset B$ such that $\|g_n\| \geqq \delta > 0$ $(n = 1, 2, \dots)$ but $\|ag_n\| + \|g_n a\| \to 0$, as $n \to \infty$.

We have the following simple results on topological divisors of zero in Banach algebras:

(1) If $a \in B$ is a topological divisor of zero, but not a divisor of zero, then $aB \neq B$.

(2) Let D be a dense subset of B. Assume that for a certain sequence $\{x_n\} \subset B$, $\|x_n\| \geqq \delta > 0$ $(n = 1, 2, \dots)$, $\|x_n d\| + \|d x_n\| \to 0$, as $n \to \infty$, for every $d \in D$. Then every element of B is a topological divisor of zero in B.

In what follows we assume that the reader is familiar with the basic theory of compact Lie groups, as is presented for example in [1]. Let G be a compact, connected Lie group. Denote by $\hat{G}$ its dual. For $h \in L^p(G)$ $(p \geqq 1)$ we denote by $\|h\|_p$ the L^p-norm. For $\alpha \in \hat{G}$ and $T_\alpha \in \alpha$ the character function $\varphi_\alpha(g) = \mathrm{Tr}\, T_\alpha(g)$ is continuous on G.

Received January 27, 1983, and in revised form April 20, 1983.

Lemma. *Let G be a compact, connected, non-abelian Lie group. Then for every $h\in L^2(G)$ we have*

(i) $|h*\varphi_\alpha(g)|\leqq M_h,\ \forall\alpha\in\hat{G}$,

(ii) $h*\varphi_\alpha(g)=\varphi_\alpha*h(g)\to 0$ *as* $\alpha\to\{\infty\}$,

(iii) *there exists* $\delta>0$ *such that* $\|\varphi_\alpha\|\geqq\delta$ *for a certain* $\alpha\to\{\infty\}$.

Proof. (i) $|h*\varphi_\alpha(g)|\leqq\int|h(x)\cdot\varphi_\alpha(gx^{-1})|dx\leqq\|h\|_2\|\varphi_\alpha\|_2=\|h\|_2$.

(ii) Let $\hat{h}(\alpha)=\int h(x)\,T_\alpha(x)^*\,dx$; here $T_\alpha(x)^*$ denotes the adjoint of $T_\alpha(x)\in\in L(H_\alpha)$ $\big(L(H_\alpha)$ stands for all linear operators in $H_\alpha\big)$. Assume that $\dim H_\alpha=N_\alpha$. We have

$$\sum_{\alpha\in\hat{G}}N_\alpha\|\hat{h}(\alpha)\|_2^2=\|h\|_2^2,$$

where $\|\hat{h}(\alpha)\|_2^2=\operatorname{Tr}\hat{h}(\alpha)^*\hat{h}(\alpha)$.

Since

$$[\operatorname{Tr}\hat{h}(\alpha)^*\hat{h}(\alpha)\cdot N_\alpha]^{1/2}\to 0\quad\text{as}\quad\alpha\to\{\infty\}$$

and

$$h*\varphi_\alpha(g)=\int h(s^{-1}g)\operatorname{Tr}T_\alpha(s)\,ds=\int h(x)\operatorname{Tr}T_\alpha(g)\,T_\alpha(x)^*\,dx=\operatorname{Tr}T_\alpha(g)\,\hat{h}(\alpha),$$

therefore

$$|h*\varphi_\alpha(g)|=|\operatorname{Tr}T_\alpha(g)\hat{h}(\alpha)|\leqq[N_\alpha\operatorname{Tr}\hat{h}(\alpha)^*\hat{h}(\alpha)]^{1/2}\to 0\quad\text{as}\quad\alpha\to\{\infty\}.$$

(iii) Let T be a maximal torus in G. Since $\varphi_\alpha(g_1g_2)=\varphi_\alpha(g_2g_1)$, $\forall\alpha\in\hat{G}$, applying Weyl's theorem [1, Th. 6.1] we have

$$\int|\varphi_\alpha(g)|\,dg=\int_T|\varphi_\alpha(t)|\,u(t)\,dt,$$

where $u(t)=|p(t)|^2|W|^{-1}$, $|p(t)|^2=\prod_{j=1}^{m}4\sin^2\pi\theta_j(t)$, $|W|\in\mathbf{N}$ is a universal integer, and $\Theta_1,\dots,\Theta_m$ are distinct roots of G. But T is commutative, so

$$\varphi_\alpha(t)=\sum_{k=1}^{N_\alpha}\exp\big(2i\pi\lambda_k^{(\alpha)}(t)\big),$$

where $\lambda_k^{(\alpha)}(t)$ are real. Assume that $\dim T=n$. Then we have

$$\lambda_k^{(\alpha)}(t_1,\dots,t_n)=\sum_{p=1}^{n}a_{kp}^{(\alpha)}t_p,\quad a_{kp}^{(\alpha)}\in\mathbf{Z},\ \forall k,p.$$

Thus

$$|W|\int|\varphi_\alpha(t)|\,u(t)\,dt=\int_{I_n}\int_0^1|\exp(2\pi ia_{11}^{(\alpha)}t_1)A_1(\bar{t})+\dots+\exp(2\pi ia_{N_\alpha 1}^{(\alpha)}t_1)A_{N_\alpha}(\bar{t})|\,u(t)\,dt,$$

where $t=(t_1, \bar{t})$ and $|A_s(\bar{t})|=1$, $s=1, 2, \ldots, N_\alpha$, $I_n=[0, 1]^{n-1}$. Hence

$$|W| \int |\varphi_\alpha(t)| u(t)\, dt =$$

$$= \int_{I_n} \int_0^1 \big|1+\exp\big(2\pi i (a_{21}^{(\alpha)}-a_{11}^{(\alpha)}) t_1\big) A_2(\bar{t}) \bar{A}_1(\bar{t}) + \ldots$$

$$\ldots + \exp\big(2\pi i (a_{N_\alpha 1}^{(\alpha)}-a_{11}^{(\alpha)}) t_1\big) A_{N_\alpha}(\bar{t}) \bar{A}_1(\bar{t})\big| u(t_1, \bar{t})\, dt_1\, d\bar{t}.$$

Choose $\alpha\to\{\infty\}$ such that $a_{kp}^{(\alpha)}-a_{11}^{(\alpha)}\neq 0$ for every k, p. Applying Szegö's theorem we have

$$\int_{I_n} \int_0^1 \big|1+\ldots+\exp\big(2\pi i (a_{N_\alpha 1}^{(\alpha)}-a_{11}^{(\alpha)}) t_1\big) A_{N_\alpha}(\bar{t}) \bar{A}_1(\bar{t})\big| u(t_1, \bar{t})\, dt_1\, d\bar{t} \geqq$$

$$\geqq \int_{I_n} \exp \int_0^1 \log u(t_1, \bar{t})\, dt_1\, d\bar{t}.$$

Since

$$\int_0^1 \log \sin^2 r\, dr > -\infty,$$

so

$$\exp \int_0^1 \log u(t_1, \bar{t})\, dt_1 \geqq \delta(\bar{t}) > 0$$

and is a continuous function of $\bar{t}\in I_n$. Hence

$$\int_{I_n} \int_0^1 \big|1+\ldots+\exp\big(2\pi i (a_{N_\alpha 1}^{(\alpha)}-a_{11}^{(\alpha)}) t_1\big) A_{N_\alpha}(\bar{t}) \bar{A}_1(\bar{t})\big| u(t_1, \bar{t})\, dt_1\, d\bar{t} \geqq \delta,$$

for a certain $\delta>0$. Note also that the number

$$\int_{I_n} \exp \int_0^1 \log u(t_1, \bar{t})\, dt_1\, d\bar{t}$$

does not depend on α, and so

$$|W| \int |\varphi_\alpha(t)| u(t)\, dt \geqq \delta$$

for every $\alpha\in\hat{G}$. The proof is complete.

As is well known, no $h\in L^1(G)$ $(h\neq 0)$ is a divisor of zero in $L^1(G)$. Hence applying Lemma, (1), and (2) we get

Theorem. *Let G be a compact, connected Lie group. Then for every $h\in L^1(G)$ the mapping $L^1(G)\ni g\to h*g\in L^1(G)$ is not surjective.*

Proof. If G is abelian, the result holds by the theorem of Dieudonné. Hence we can assume that G is not abelian. By (i) and (ii) of Lemma and the Lebesgue dominated convergence theorem we have $\|h*\varphi_\alpha\|_1\to 0$ as $\alpha\to\{\infty\}$, for any $h\in L^2(G)$. Application of (1) and (2) ends the proof.

Remark 1. Since $L^p(G)$ is $L^1(G)$ module, for $p\geqq 1$, the above theorem can be easily extended to $L^p(G)$. Namely, for every $h\in L^1(G)$ the mapping $L^p(G)\ni g\to\to h*g\in L^p(G)$ is not surjective. The proof is the same as before (note that $\|\varphi_\alpha\|_p\geqq\geqq\|\varphi_\alpha\|_1$, $\forall\alpha\in\hat{G}$).

Remark 2. It seems that the above result is also true in the context of nilpotent Lie groups (this is true for the Heisenberg group of arbitrary dimension).

I would like to thank the referee for pointing out an error in the first version of this paper.

References

[1] J. F. Adams, *Lectures on Lie groups*, Benjamin, Inc. (New York—Amsterdam, 1970).
[2] J. Dieudonné, Sur le product de convolution II, *J. Math. Pures Appl.*, 39 (3) (1960), 275—292.
[3] J. T. Burnham, R. R. Goldberg, The convolution theorem of Dieudonné, *Acta Sci. Math.*, **36** (1974), 1—3.

INSTYTUT MATEMATYCZNY PAN
UL. SOLSKIEGO 30
31-027 KRAKÓW, POLAND

Acta Sci. Math. 47 (1984), 465—470

Über numerische Wertebereiche und Spektralwertabschätzungen

A. RHODIUS

0. Einleitung

Numerische Wertebereiche für lineare Operatoren in Hilberträumen werden seit den Arbeiten [4], [15] von F. HAUSDORFF und O. TOEPLITZ untersucht. G. LUMER [9] und F. L. BAUER [2] führten numerische Wertebereiche für Banachraum-Operatoren ein. Nach J. P. WILLIAMS [16] ist das Spektrum jedes stetigen Endomorphismus eines Banachraumes eine Teilmenge der abgeschlossen Hülle des Bauerschen numerischen Wertebereiches.

Die abgeschlossene Hülle des numerischen Wertebereiches von Lumer enthält im allgemeinen nur das approximative Punktspektrum. In der vorliegenden Note werden mit Hilfe von zu Halbnormen gehörenden Wertebereichen Einschließungsmengen für Teile des Spektrums angegeben. Diese Resultate können als Verallgemeinerungen der Sätze von Williams und Lumer auf in halbnormierten Räumen wirkende Operatoren aufgefaßt werden. Als Anwendungsmöglichkeiten ergeben sich Spektralwerteinschließungen für Hilbertraum-Operatoren, für Integraloperatoren mit stochastischen Kernen ebenso wie Ergebnisse für diskrete Markovprozesse bezüglich ihres asymptotischen Verhaltens.

1. Begriffe und Bezeichnungen

Es sei E ein Vektorraum über dem Körper C der komplexen Zahlen, p eine Halbnorm auf E und $T: E \to E$ ein Endomorphismus von E. Ferner bezeichne S_p die Einheitssphäre $\{x \in E: p(x)=1\}$ und $D_p(x)$ die Menge der Stützfunktionale:

$$D_p(x) = \{f \in E': f(x) = 1, \quad |f(y)| \leq p(y) \quad (y \in E)\} \ (x \in S_p).$$

Eingegangen am 22. Februar 1983.

Für die Abbildung $Q_p: S_p \to \mathfrak{P}(E')$ der Einheitssphäre S_p in die Potenzmenge $\mathfrak{P}(E')$ gelte $\emptyset \neq Q_p(x) \subseteqq D_p(x)$ $(x \in S_p)$.

Die Menge

$$V_{Q_p}(T) = \{f(Tx)\colon\ f \in Q_p(x),\quad x \in S_p\}$$

heißt numerischer Wertebereich von T bezüglich Q_p. (Vgl. [11].) Da für die zugelassenen Abbildungen Q_p die konvexe Hülle von $V_{Q_p}(T)$ mit der konvexen Hülle von $V_{D_p}(T)$ übereinstimmt, ist $\sup\{|\lambda|\colon \lambda \in V_{Q_p}(T)\}$ unabhängig von der speziellen Abbildung Q_p. Die Größe

$$v_p(T) = \sup\{|\lambda|\colon\ \lambda \in V_{D_p}(T)\}$$

heißt numerischer Radius des Endomorphismus T.

Unter dem Spektrum $\sigma(T)$ verstehen wir stets das algebraische Spektrum des Endomorphismus T, das heißt, die komplexe Zahl λ gehört genau dann zu $\sigma(T)$, wenn $T-\lambda I$ keine bijektive Abbildung von E ist. Im Falle stetiger Endomorphismen in Banachräumen ist das algebraische Spektrum bekanntlich genau das topologische Spektrum. Für eine Norm p bezeichnet man als approximatives Punktspektrum a.p.$\sigma(T)$ die Menge aller $\lambda \in C$, für die eine Folge (x_n) aus S_p mit $\lim_{n\to\infty} p((T-\lambda I)x_n)=0$ existiert.

Ist F eine invarianter Unterraum des Endomorphismus T, so bezeichne $T_{|F}$ die Einschränkung von T auf F. So bezeichnet zum Beispiel $T_{|F_p}$ die Einschränkung eines stetigen Endomorphismus T von (E, p) auf den Nullraum $F_p = \{x \in E\colon p(x)=0\}$.

2. Die Spektraleigenschaften numerischer Wertebereiche

Satz 1. *Es sei T ein stetiger Endomorphismus des vollständigen halbnormierten Raumes (E, p). Dann gilt*

$$\sigma(T) \setminus \sigma(T_{|F_p}) \subseteqq \overline{V_{D_p}(T)}.$$

Beweis. Mit E/F_p bezeichnen wir den Quotientenraum von E nach $F_p = \{x \in E\colon p(x)=0\}$ und mit $[x]$ die Restklasse $x+F_p$ modulo F_p. Durch die Beziehung $\|[x]\| = p(x)$ $(x \in E)$ ist eine Norm auf E/F_p definiert; $(E/F_p, \|\cdot\|)$ ist ein Banachraum. Da F_p bezüglich T invarianter Teilraum von E ist, wird durch T eine lineare Abbildung T_{F_p} (die sogenannte Quotientenabbildung) von E/F_p in sich induziert. $T_{F_p}[x]=[y]$ genau dann, wenn $Tx \in [y]$ gilt.

Da die stetigen Linearformen $f \in E'$ auf jeder Restklasse modulo F_p konstant sind, wird durch die Vorschrift $(jf)[x]=f(x)$ $(f \in E',\ x \in E)$ eine Abbildung j von E' in $(E/F_p)'$ definiert. Die Abbildung j ist eine eineindeutige, bezüglich der Supremumsnormen isometrische Abbildung von E' auf $(E/F_p)'$. Es gilt

$$V_{D_{\|\cdot\|}}(T_{F_p}) = \{f^*(T_{F_p}[x]) : f^* \in D_{\|\cdot\|}([x]),\ [x] \in S_{\|\cdot\|}\} =$$
$$= \{(jf)([Tx]) : f \in D_p(x),\ x \in S_p\} = \{f(Tx) : f \in D_p(x),\ x \in S_p\} = V_{D_p}(T).$$

Für den stetigen Endomorphismus T_{F_p} des Banachraumes $(E/F_p, \|\cdot\|)$ gilt nach einem Satz von WILLIAMS [16]

$$\sigma(T_{F_p}) \subseteqq \overline{V_{D_{\|\cdot\|}}(T_{F_p})} = \overline{V_{D_p}(T)}.$$

Andererseits ergibt sich leicht die für invariante Teilräume bekannte Beziehung $\sigma(T) \subseteqq \sigma(T_{|F_p}) \cup \sigma(T_{F_p})$ (siehe z. B. [7]), so daß die Behauptung folgt.

Satz 2. *Es sei T ein stetiger Endomorphismus der halbnormierten Raumes (E, p). Es sei L eine Menge komplexer Zahlen derart, daß für jedes $\lambda \in L$ eine Folge (x_n) existiert mit $\lim_{n\to\infty} p((T-\lambda I)x_n) = 0$ und nicht $\lim_{n\to\infty} p(x_n) = 0$. Dann gilt $L \subseteqq \overline{V_{Q_p}(T)}$.*

Beweis. Für $\lambda \in L$ existieren nach Voraussetzung eine Folge (x_n) aus E und ein $\varepsilon_0 > 0$ mit $\lim_{n\to\infty} p((T-\lambda I)x_n) = 0$ und $p(x_n) \geqq \varepsilon_0$ $(n \in N)$. Dann gilt mit $y_n = x_n/p(x_n)$ die Beziehung $p((T-\lambda I)y_n) \to 0$. Für jedes $f_n \in Q_p(y_n)$ folgt

$$f_n(Ty_n) = f_n((T-\lambda I)y_n) + \lambda f_n(y_n) \to \lambda,$$

also gilt $\lambda \in \overline{V_{Q_p}(T)}$.

Satz 3. *Es sei T ein stetiger Endormorphismus des normierten Raumes $(E, \|\cdot\|)$ in sich. F sei ein bezüglich T invarianter abgeschlossener Unterraum von $(E, \|\cdot\|)$ und $\tilde{p}(z) = \inf_{y \in F} \|y+z\|$ $(z \in E)$. Dann gilt*

$$a.p.\sigma(T) \setminus a.p.\sigma(T_{|F}) \subseteqq \overline{V_{Q_{\tilde{p}}}(T)}.$$

Beweis. Für $\lambda \in L := \text{a.p.}\sigma(T) \setminus \text{a.p.}\sigma(T_{|F})$ existiert eine Folge (x_n) mit $\|x_n\| = 1$, $\|(T-\lambda I)x_n\| \to 0$ und nicht $\tilde{p}(x_n) \to 0$. Denn aus $\tilde{p}(x_n) \to 0$ folgt die Existenz einer Folge (y_n) aus F mit $\|x_n - y_n\| \to 0$, damit ergäbe sich aus der Stetigkeit von T zusammen mit $\|(T-\lambda I)x_n\| \to 0$ die Beziehung $\|(T-\lambda I)y_n\| \to 0$; da $\|y_n\| \to 1$ gilt, würde sonst λ zu a.p.$\sigma(T_{|F})$ gehören. Da T auch bezüglich $\tilde{p}$ stetig ist, sind für $T, L, \tilde{p}$ alle Voraussetzungen des Satzes 2 erfüllt, womit die Behauptung folgt.

3. Anwendungen

3.1. Hilbertraum-Operatoren. Es sei T ein stetiger Endomorphismus des Hilbertraumes E; $\lambda_1, \lambda_2, \ldots, \lambda_l$ seien voneinander verschiedene Eigenwerte von T mit zugehörigen Eigenvektoren $x_1, x_2, \ldots, x_l$ $(Tx_i = \lambda_i x_i,\ i = 1, 2, \ldots, l)$. Wir setzen

$$\Sigma = \{f \in E:\ \|f\| = 1,\quad (x_i, f) = 0\quad (i = 1, 2, \ldots, l)\},$$

und

$$p(x) = \sup \{|(x,f)|: f \in \Sigma\} \quad (x \in E).$$

Dann gilt offensichtlich

$$V_{D_p}(T) = \{(Tx,f): f \in \Sigma,\ x \in E,\ (x,f) = p(x) = 1\}.$$

Hilfssatz 1. $V_{D_p}(T) = \{(Tx, x): \|x\| = 1,\ (x, x_i) = 0\ (i = 1, 2, \ldots, l).\}$

Beweis. Wir zeigen, daß zu jedem $(Tx,f) \in V_{D_p}(T)$ ein $z \in E$ existiert, für das $\|z\| = 1$, $z \perp \mathscr{L}(x_1, x_2, \ldots, x_l)$ und $(Tx, f) = (Tz, z)$ gelten. Da $F_p = \mathscr{L}(x_1, \ldots, x_l)$ als endlichdimensionaler Teilraum von E abgeschlossen ist, existiert zu x genau ein Paar (x_0, z) mit $x_0 \in F_p$, $z \perp F_p$ und $x = x_0 + z$. Da F_p bezüglich T invariant ist, gilt $(Tz, z) = (Tx - Tx_0, z) = (Tx, z)$.

Andererseits folgen aus $x - z \in F_p$ die Beziehungen $p(z) = (z, f) = 1$. Aus $z/\|z\| \in \sum$ ergibt sich $\|z\| = |(z, z/\|z\|)| \leqq p(z)$ und somit $\|z\| = p(z) = 1$. Wegen $1 = (z, f) \leqq \|z\| \|f\| = 1$ gilt $f = z$, was noch zu zeigen war.

Satz 4. *Es gilt*

$$\sigma(T) \setminus \{\lambda_1, \lambda_2, \ldots, \lambda_l\} \subseteqq \overline{\{(Tx, x): \|x\| = 1,\ (x, x_i) = 0,\quad i = 1, 2, \ldots, l\}}.$$

Beweis. Die Halbnorm p ist die kanonische Halbnorm von $(E, \|\cdot\|)$ bezüglich des Unterraumes $\mathscr{L}(\{x_1, \ldots, x_l\}) = F_p$. Damit ist (E, p) vollständig und T bezüglich der Halbnorm p stetig, so daß Satz 1 zusammen mit Hilfssatz 1 die Behauptung liefert.

3.2. Integraloperatoren mit stochastischen Kernen. Es sei $(X, \mathscr{B}, \mu)$ ein Maßraum mit dem positiven Maß μ und $B = B(X, \mathscr{B})$ die Menge der komplexwertigen $\mathscr{B}$-meßbaren beschränkten Funktionen auf X. Wir betrachten den Operator $T: B \to B$ mit

$$(Tx)(t) = \int_X H(t, s)\, x(s)\, d\mu(s) \quad (x \in B,\ t \in X).$$

Dabei sei H eine reellwertige $\mathscr{B} \times \mathscr{B}$-meßbare Funktion auf $X \times X$ und erfülle die Bedingungen

$$H(t, s) \geqq 0, \int_X H(t, s)\, d\mu(s) = 1 \quad (t, s \in X).$$

Bezüglich der Supremumsnorm $\|x\| = \sup_{s \in X} |x(s)|$ ist der Raum $(B, \|\cdot\|)$ vollständig, der Operator T ist beschränkt mit $\|T\| = 1$. Der mit der Oszillationshalbnorm $p(x) = \sup_{t, t' \in X} |x(t) - x(t')|$ versehene Raum (B, p) ist vollständig. Der Integraloperator T ist bezüglich p stetig mit

$$p(T) = \frac{1}{2} \sup_{t,t' \in X} \int_X |H(t, s) - H(t', s)| \, d\mu(s).$$

(Siehe [10], [13]).

Satz 5. *Für jede zur Oszillationshalbnorm p gehörende Daulitätsabbildung Q_p gilt*

$$a.p.\sigma(T) \setminus \{1\} \subseteqq \overline{V_{Q_p}(T)}.$$

Beweis. Wir benutzen Satz 2 und setzen $L = \text{a.p.}\sigma(T) \setminus \{1\}$. Für $\lambda \in L$ existiert eine Folge (x_n) mit $\|x_n\| = 1$ und $\|(T - \lambda I)x_n\| \to 0$. Es folgt $p((T - \lambda I)x_n) \to 0$. Andererseits gilt nicht $p(x_n) \to 0$; denn aus $p(x_n) \to 0$ und $\|x_n\| = 1$ folgt die Existenz einer konstanten Funktion c mit $\|x_n - c\| \to 0$ und somit aus der Stetigkeit des Operators T (bezüglich $\|\cdot\|$) die Gleichung $Tc = \lambda c$, also $\lambda = 1$.

Als Folgerung von Satz 5 ergibt sich für alle $\lambda \in \text{a.p.}\sigma(T) \setminus \{1\}$ die Abschätzung $|\lambda| \leqq v_p(T)$. Diese Ungleichung stellt eine Verschärfung der von E. Hopf [5], Bauer—Deutsch—Stoer [3], Anselone—Lee [1], Rhodius [10] angegebenen Abschätzungen für die von 1 verschiedenen Eigenwerte von T dar. In [12] ist eine Darstellung des numerischen Radius $v_p(T)$ in Abhängigkeit vom Kern H und dem Maß μ angegeben.

3.3. Homogene Markovketten mit allgemeinen Zustandsräumen. Jede homogene Markovkette $(X_n)_{n \in N}$ mit dem meßbaren Raum $(X, \mathscr{B})$ als Zustandsraum ist durch eine Übergangswahrscheinlichkeit P auf $(X, \mathscr{B})$ und eine Anfangsverteilung p auf $\mathscr{B}$ bestimmt. Es gelten $P(X_{n+1} \in A | X_n) = P(X_n, A)$ $(n \in N,\ A \in \mathscr{B})$ und $P(X_0 \in A) = p(A)$ $(A \in \mathscr{B})$. Die Markovkette heißt stark ergodisch, wenn eine Wahrscheinlichkeitsverteilung Q auf $\mathscr{B}$ derart existiert, daß

$$\lim_{m \to \infty} \sup_{t \in X,\, A \in \mathscr{B}} |P(X_m \in A | X_0 = t) - Q(A)| = 0.$$

Um die Eigenschaft der starken Ergodizität durch das Verhalten numerischer Wertebereiche zu charakterisieren, wird der Übergangswahrscheinlichkeit P ein Endomorphismus T des Raumes $B = B(X, \mathscr{B})$ der komplexwertigen $\mathscr{B}$-meßbaren beschränkten Funktionen auf X zugeordnet:

$$(Tx)(t) = \int_X x(s) P(t, ds) \quad (x \in B).$$

T ist bezüglich der Oszillationshalbnorm $p(x) = \sup_{t,t' \in X} |x(t) - x(t')|$ $(x \in B)$ stetig, und es gilt

$$p(T) = \sup_{t,t' \in X,\, A \in \mathscr{B}} |P(t, A) - P(t', A)|$$

(siehe [14]); $1 - p(T)$ ist also der zur Übergangswahrscheinlichkeit P gehörende Ergodizitätskoeffizient. Da (B, p) vollständig ist, ist Satz 1 anwendbar, und es gilt wegen $T1 = 1$ die Beziehung

$$\sigma(T) \setminus \{1\} \subseteq \overline{V_{D_p}(T)}.$$

Aufgrund der letzten Inklusion kann mit Sätzen über die Konvergenz von Potenzen linearer Operatoren (siehe z. B. [6], [8]) folgende Aussage bewiesen werden (siehe [14]).

Satz 6. *Die homogene Markovkette* $(X_n)_{n\in N}$ *ist genau dann stark ergodisch, wenn eine natürliche Zahl* m *existiert, so daß der numerische Radius* $v_p(T^m)$ *kleiner als* 1 *ist.*

Als Folgerung dieses Satzes erhält man unmittelbar die für homogene Markovketten bekannte Äquivalenz von starker und schwacher Ergodizität und eine Charakterisierung der starken Ergodizität durch den Ergodizitätskoeffizienten (siehe [14]).

Literatur

[1] P. M. Anselone and J. W. Lee, Spectral properties of integral operators with nonnegative kernels, *Linear Algebra Appl.*, **9** (1974), 67—87.

[2] F. L. Bauer, On the field of values subordinate to a norm, *Numer. Math.*, **4** (1962), 103—111.

[3] F. L. Bauer, E. Deutsch und J. Stoer: Abschätzungen für die Eigenwerte positiver linearer Operatoren, *Linear Algebra Appl.*, **2** (1969), 275—301.

[4] F. Hausdorff, Der Wertevorrat einer Bilinearform, *Math. Z.*, **3** (1919), 314—316.

[5] E. Hopf, An inequality for positive linear integral operators, *J. Math. Mech.*, **12** (1963), 683—692.

[6] J. J. Koliha, Power Convergence and Pseudoinverses of Operators in Banach Spaces, *J. Math., Anal. Appl.*, **48** (1974), 446—469.

[7] H. P. Lotz, Über das Spektrum positiver Operatoren, *Math. Z.*, **108** (1968), 15—32.

[8] G. R. Luecke, Norm convergence of T^n, *Canad. J. Math.*, **29** (1977), 1340—1344.

[9] G. Lumer, Semi-inner-product spaces, *Trans. Amer. Math. Soc.*, **100** (1961), 29—43.

[10] A. Rhodius, Eine Eigenwertabschätzung für Integraloperatoren mit stochastischen Kernen, *Comment. Math. Univ. Carolinae*, **18** (1977), 183—193.

[11] A. Rhodius, Über zu Halbnormen gehörende numerische Wertebereiche linearer Operatoren, *Math. Nachr.*, **86** (1978), 181—185.

[12] A. Rhodius, Numerische Wertebereiche linearer Operatoren und Einschließungsmengen für die nichttrivialen Eigenwerte von Integraloperatoren mit stochastischen Kernen, *Beiträge Anal.*, **14** (1979), 47—61.

[13] A. Rhodius, Ergodizität inhomogener Markovketten und numerische Wertebereiche, *Wiss. Z. PH 'K.F.W. Wander' Dresden*, **3** (1980), 47—60.

[14] A. Rhodius, A characterization of strong ergodicity by means of numerical ranges, in: *Proc. Conference Topology and Measure III* (GDR 1980), vol. 2 (Greifswald 1982), 229—233.

[15] O. Toeplitz, Das algebraische Analogon zu einem Satz von Fejér, *Math. Z.*, **2** (1918), 187—197.

[16] J. P. Williams, Spectra of products and numerical ranges, *J. Math. Anal. Appl.*, **17** (1967), 214—220.

SEKTION MATHEMATIK
PH „K. F. W. WANDER" DRESDEN
806 DRESDEN, PSF 365, DDR

Acta Sci. Math. 47 (1984), 471—485

Convergence of solutions of a nonlinear integrodifferential equation arising in compartmental systems

T. KRISZTIN

In honour of Professor Béla Szőkefalvi-Nagy on his 70th birthday

1. Introduction

The compartmental models play an important role in the mathematical description of biological processes, chemical reactions, economic and human interactions [1, 2, 8]. I. GYŐRI [3, 4] used nonlinear integrodifferential equations to describe compartmental systems with pipes and propounded the question whether the bounded solutions of the model equation have limits as $t\to\infty$. If the transit times of material flow between compartments are zero, then the model equations are ordinary differential equations. In this case there are known results on the existence of the limit of solutions [9, 12]. But these methods are not applicable if the transit times are not zero. The existence of the limit of solutions is also known in the case of nonzero transit times if the model equation is linear [5] or if the so-called transport functions are continuously differentiable [11]. But there occur compartmental systems in the applications such that the transport functions do not satisfy even the local Lipschitz continuity. For example, in hydrodynamical models, where the free outflow of water through a leak at the bottom of a container has a rate proportional to square root of the amount of water in the container. If the transport functions are continuous, monotone nondecreasing and the model equation has exactly one equilibrium state then the solutions tend to this one as $t\to\infty$ [4].

In this paper we examine stationary compartmental systems with pipes, which are described by nonlinear autonomous integrodifferential equations, and the transit times of material flow through pipes are characterized by distribution functions.

Received December 9, 1982.

We show that the model equations have equilibrium states if and only if their solutions are bounded (the equilibrium points need not be unique). The main result of this paper guarantees the existence of the limits of the bounded solutions if the transport functions are continuous, strictly increasing functions.

2. The model equation, notations and definitions

Consider a stationary n-compartmental system with pipes. It is well-known (see e.g. [3, 4]) that the state vector $x(t)$ is the solution of the following system of integrodifferential equations:

$$(1)\qquad \dot{x}_i(t) = -\sum_{j=0}^{n} h_{ji}(x_i(t)) + \sum_{j=1}^{n} \int_0^{\tau} h_{ij}(x_j(t-s))\, dF_{ij}(s) + I_i \quad (i = 1, \dots, n),$$

where

(a) $h_{ij}: R \to R$ is a continuous, monotone nondecreasing function, $h_{ij}(0)=0$ $(i=0, 1, \dots, n;\ j=1, \dots, n)$;

(b) $\tau > 0$;

(c) $F_{ij}: [0, \tau] \to [0, 1]$ is continuous from the left, monotone nondecreasing and $F_{ij}(0)=0$, $F_{ij}(\tau)=1$ $(i, j=1, \dots, n)$;

(d) $I_i \geqq 0$ $(i=1, \dots, n)$.

Denote by $C_1, \dots, C_n$ the compartments and by C_0 the environment of the compartmental system. In equation (1) the function h_{ij} is called the transport function, which is the rate of material outflow from C_j in the direction of C_i $(i=0, 1, \dots, n;$ $j=1, \dots, n)$. The nonnegative number I_i is the inflow rate of material flow from environment C_0 into compartment C_i $(i=1, \dots, n)$.

Since in equation (1) the components of the solution vector denote material amounts, it is a reasonable claim that solutions corresponding to nonnegative initial conditions should be nonnegative, and the model equation (1) should have a unique solution for any given initial condition. In Section 3 we prove that (1) has these properties.

Let R and R^n be the set of real numbers and the n-dimensional Euclidean space, respectively, and $|\cdot|$ denotes the norm in R^n. Denote by $C([a, b], R^n)$ the Banach space of continuous functions mapping the interval $[a, b]$ into R^n with the topology of uniform convergence.

It is natural to consider the space $C([-\tau, 0], R^n)$ for the state space of (1). Let $r=2n\tau$. Obviously, without loss of generality, $C=C([-r, 0], R^n)$ may also be regarded as a state space of (1). In this paper we use C for the phase space of (1).

Denote the norm of an element φ in C by $\|\varphi\| = \max_{-r \leqq s \leqq 0} |\varphi(s)|$. If $t_0 \in R$, $A>0$ and $x: [t_0-r, t_0+A) \to R^n$ is continuous, then for any $t \in [t_0, t_0+A)$ let $x_t \in C$ be defined by $x_t(s)=x(t+s)$, $-r \leqq s \leqq 0$.

A function $x: I \to R^n$ is said to be a solution of (1) on the interval I if x is continuous on I and $x(t)$ satisfies (1) for every $t \in I$ such that $t-r \in I$. For given $\varphi \in C$ we say that $x(\varphi)$ is a solution of (1) through $(0, \varphi)$ if there is an $A>0$ such that $x(\varphi)$ is a solution of (1) on $[-r, A)$ and $x_0(\varphi)=\varphi$.

It follows from conditions (a), (b), (c), (d) that for every $\varphi \in C$ there is a solution $x(\varphi)$ of (1) through $(0, \varphi)$ and if x is a noncontinuable, bounded solution of (1) on the interval $[-r, A)$, then $A=\infty$ [7, Theorems 2.2.1, 2.3.2].

We prove in Section 3 that if $\varphi \in C$, then equation (1) has at most one solution $x(\varphi)$ through $(0, \varphi)$.

Let $x(\varphi)$ be a solution of (1) on the interval $[-r, \infty)$, $\varphi \in C$. Define the ω-limit set $\Omega(\varphi)$ of the solution $x(\varphi)$ as follows: $\Omega(\varphi)=\{\psi \in C$: there is a sequence $\{t_n\}$ such that $t_n \to \infty$ and $\|\psi - x_{t_n}(\varphi)\| \to 0$ as $n \to \infty\}$. The set $M \subset C$ is said to be invariant if for every $\psi \in M$ equation (1) has a solution y on R such that $y_0=\psi$ and $y_t \in M$ for all $t \in R$. If $x(\varphi)$ is a bounded solution of (1) on $[-r, \infty)$, then $\Omega(\varphi)$ is nonempty, compact, connected, invariant and $x_t(\varphi) \to \Omega(\varphi)$ as $t \to \infty$ [7, Corollary 4.2.1].

Let $N \subset \{1, 2, \dots, n\}$ and define the directed graph $D_N=(V(D_N), A(D_N))$ to equation (1) as follows: $V(D_N)=\{v_i: i \in N\}$ is the set of vertices, $A(D_N)= =\{a_{ij}: h_{ij}(\cdot) \not\equiv 0,\ (i,j) \in N \times N\}$ is the set of arcs, where the arc a_{ij} is said to join v_j to v_i, v_j is the tail of a_{ij} and v_i is its head. A directed (v_j, v_i)-walk W from v_i to v_j is a finite non-null sequence $W=(a_{i_1 i_0}, a_{i_2 i_1}, \dots, a_{i_k i_{k-1}})$, where $a_{i_1 i_0}, a_{i_2 i_1}, \dots, a_{i_k i_{k-1}} \in A(D_N)$ and $i_0=i$, $i_k=j$. If $i_0, i_1, \dots, i_k$ are distinct, then the walk $W=(a_{i_1 i_0}, a_{i_2 i_1}, \dots, a_{i_k i_{k-1}})$ is called a directed (v_{i_k}, v_{i_0})-path. Two vertices v_i, v_j are diconnected in D_N if there are a directed (v_i, v_j)-path, and a directed (v_j, v_i)-path in D_N. The diconnection is an equivalence relation on set N. The directed subgraphs $D_{N_1}, D_{N_2}, \dots, D_{N_k}$ induced by the resulting partition $(N_1, N_2, \dots, N_k)$ of N are called the dicomponents of D_N. It is easy to see that there exists a dicomponent $D_{N_{i_0}}$ of D_N such that if $i \in N_{i_0}$ and $j \in N \setminus N_{i_0}$, then $a_{ij} \notin V(D_N)$.

3. Uniqueness, boundedness and some technical lemmas

In this section we prove some easy lemmas, which are necessary to the proof of the main result.

Define the functional $U: C \times C \to [0, \infty)$ as follows:

$$U(\varphi, \psi) = \sum_{i=1}^{n} \Big[|\varphi_i(0) - \psi_i(0)| + \sum_{j=1}^{n} \int_0^{\tau} \int_0^{s} \big| h_{ij}(\varphi_j(-u)) - h_{ij}(\psi_j(-u)) \big| \, du \, dF_{ij}(s) \Big],$$

$$\varphi = (\varphi_1, \dots, \varphi_n), \quad \psi = (\psi_1, \dots, \psi_n) \in C.$$

Lemma 1 claims the monotonicity of functional U along the solutions of (1).

Lemma 1. *If x and y are solutions of* (1) *on the interval* $[-r, A)$ *then* $U(x_t, y_t)$ *as a function of t is monotone nonincreasing on* $[0, A)$.

Proof. Let $u(t)=U(x_t, y_t)$, $t\in[0, A)$. Since x and y are solutions of (1) on $[-r, A)$, we have

$$\frac{d}{dt}[x_i(t)-y_i(t)] =$$

$$= -\sum_{j=0}^{n}[h_{ji}(x_i(t))-h_{ji}(y_i(t))]+\sum_{j=1}^{n}\int_0^{\tau}[h_{ij}(x_j(t-s))-h_{ij}(y_j(t-s))]\,dF_{ij}(s)$$

$$(t\in[0, A), \quad i = 1, \ldots, n).$$

Thus, from the monotonicity of functions h_{ij} it follows that

$$D^+|x_i(t)-y_i(t)| \leqq$$

$$\leqq -\sum_{j=0}^{n}|h_{ji}(x_i(t))-h_{ji}(y_i(t))|+\sum_{j=0}^{n}\int_0^{\tau}|h_{ij}(x_j(t-s))-h_{ij}(y_j(t-s))|\,dF_{ij}(s)$$

$$(t\in[0, A), \quad i = 1, \ldots, n).$$

Hence it is easy to see that

$$D^+u(t) \leqq$$

$$\leqq \sum_{i=1}^{n}\Big[-\sum_{j=0}^{n}|h_{ji}(x_i(t))-h_{ji}(y_i(t))|+\sum_{j=1}^{n}\int_0^{\tau}|h_{ij}(x_j(t-s))-h_{ij}(y_j(t-s))|\,dF_{ij}(s)+$$

$$+\sum_{j=1}^{n}|h_{ij}(x_j(t))-h_{ij}(y_j(t))|-\sum_{j=1}^{n}\int_0^{\tau}|h_{ij}(x_j(t-s))-h_{ij}(y_j(t-s))|\,dF_{ij}(s)\Big] =$$

$$= -\sum_{i=1}^{n}|h_{0i}(x_i(t))-h_{0i}(y_i(t))| \leqq 0 \quad (t\in[0, A)),$$

which, by using differential inequality [10, p. 15], completes the proof.

R. M. Lewis and B. D. O. Anderson [11] proved similar result provided that functions h_{ij} are continuously differentiable.

The uniqueness for the initial-value problem of (1) follows easily from Lemma 1.

Corollary 1. *For every $\varphi\in C$ equation* (1) *has a unique solution $x(\varphi)$ through* $(0, \varphi)$.

By using Lemma 1 and the properties of the ω-limit set one can readily verify that:

Corollary 2. *If* $\varphi \in C$ *and* $\psi \in \Omega(\varphi)$ *then there exists a sequence* $\{t_n\}$ *such that* $t_n \to \infty$ *and*

$$\sup_{u \geq 0} |x(\varphi)(t_n+u) - x(\psi)(u)| \to 0 \quad as \quad n \to \infty.$$

Define the functions $H_i: R^n \to R$ by

$$H_i(z_1, \ldots, z_n) = -\sum_{j=0}^{n} h_{ij}(z_i) + \sum_{j=1}^{n} h_{ij}(z_j) + I_i \quad (i = 1, \ldots, n),$$

where $(z_1, \ldots, z_n) \in R^n$. If $z^* \in R^n$ and $H_i(z^*) = 0$, $i = 1, \ldots, n$, then the constant function z^* is a solution of (1) on R, i.e. z^* is an equilibrium point of (1).

From Lemma 1 it is clear that the existence of an equilibrium point of (1) guarantees the solutions of (1) to be bounded.

Corollary 3. *If there exists* $z \in R^n$ *such that*

$$H_i(z) = 0 \quad (i = 1, \ldots, n), \tag{2}$$

then every solution of (1) *is bounded on* $[-r, \infty)$.

Corollary 3 is reversible in the following sense: if equation (1) has a bounded solution on $[-r, \infty)$ then equation (2) has a solution.

Lemma 2. *If* x *is a bounded solution of* (1) *on* $[-r, \infty)$ *and* $M_i = \overline{\lim_{t \to \infty}}\, x_i(t)$, $m_i = \underline{\lim_{t \to \infty}}\, x_i(t)$, $i = 1, \ldots, n$, *then*

(i) $H_i(M_1, \ldots, M_n) = 0 \quad (i = 1, \ldots, n)$;
(ii) $H_i(m_1, \ldots, m_n) = 0 \quad (i = 1, \ldots, n)$;
(iii) $h_{0i}(m_i) = h_{0i}(M_i) \quad (i = 1, \ldots, n)$.

Proof. We first prove that $H_i(M_1, \ldots, M_n) \geq 0$, $i = 1, \ldots, n$. Suppose this is not true. Then there is an $i_0 \in \{1, \ldots, n\}$ such that $H_{i_0}(M_1, \ldots, M_n) < 0$. Let $a = H_{i_0}(M_1, \ldots, M_n)$. Since functions h_{ij} are continuous, there exists $\varepsilon > 0$ such that

$$-\sum_{j=0}^{n} h_{ji_0}(M_{i_0} - \varepsilon) + \sum_{j=1}^{n} h_{i_0 j}(M_j + \varepsilon) + I_{i_0} < \frac{a}{2}. \tag{3}$$

Let T be chosen so that if $t \geq T$, then

$$\sup_{t \geq T - \tau} x_j(t) \leq M_j + \varepsilon \quad (j = 1, \ldots, n). \tag{4}$$

By using relations (3), (4) and the monotonicity of functions h_{ij} we have

$$\dot{x}_{i_0}(t) \leq -\sum_{j=0}^{n} h_{ji_0}(M_{i_0} - \varepsilon) + \sum_{j=1}^{n} h_{i_0 j}(M_j + \varepsilon) + I_{i_0} < \frac{a}{2} < 0$$

on the set $\{t \geqq T : x_{i_0}(t) \in [M_{i_0}-\varepsilon, M_{i_0}+\varepsilon]\}$. This contradicts the definition of M_{i_0}, proving the statement. By similar arguments we obtain $H_i(m_1, \ldots, m_n) \leqq 0$, $i=1, \ldots, n$. From the above and the equality $\sum_{i=1}^{n} H_i(z_1, \ldots, z_n) = \sum_{i=1}^{n} [I_i - h_{0i}(z_i)]$ it follows that

$$0 \leqq \sum_{i=1}^{n} H_i(M_1, \ldots, M_n) - \sum_{i=1}^{n} H_i(m_1, \ldots, m_n) = -\sum_{i=1}^{n} [h_{0i}(M_i) - h_{0i}(m_i)] \leqq 0,$$

which proves the lemma.

The proof of Lemma 2 is based on the idea of [4, Th. 3.2.1].

Corollary 4. *If equation* (2) *has exactly one solution, then for every solution* x *of* (1) *the limit* $\lim_{t\to\infty} x(t)$ *exists.*

Corollary 5. *If there exists* $i_0 \in \{1, \ldots, n\}$ *such that function* h_{0i_0} *is strictly monotone increasing, then for every bounded solution of* (1) *the limit* $\lim_{t\to\infty} x_{i_0}(t)$ *exists.*

Lemma 3. *If* $M_i, m_i, i=1, \ldots, n$, *are real numbers and*

(i) $M_i > m_i \quad (i=1, \ldots, n)$,

(ii) $H_i(M_1, \ldots, M_n) = 0 \quad (i=1, \ldots, n)$,

(iii) $H_i(m_1, \ldots, m_n) = 0 \quad (i=1, \ldots, n)$,

then for every $\varepsilon \in (0, \min_{i=1,\ldots,n}(M_i - m_i))$ *there exists* $z^*(\varepsilon) = (z_1^*, \ldots, z_n^*) \in R^n$ *such that*

(iv) $M_i - \varepsilon \leqq z_i^* \leqq M_i \quad (i=1, \ldots, n)$,

(v) *there is an* $i_0 \in \{1, \ldots, n\}$ *such that* $z_{i_0}^* = M_{i_0} - \varepsilon$,

(vi) $H_i(z_1^*, \ldots, z_n^*) = 0 \quad (i=1, \ldots, n)$.

Proof. From the equality $\sum_{i=1}^{n} H_i(z_1, \ldots, z_n) = \sum_{i=1}^{n} [I_i - h_{0i}(z_i)]$, the monotonicity of functions h_{0i} and (ii), (iii) it follows that $\sum_{i=1}^{n} H_i(z_1, \ldots, z_n) = 0$ for $z_i \in [m_i, M_i]$, $i=1, \ldots, n$. Let $\varepsilon \in (0, \min_{i=1,\ldots,n}(M_i - m_i))$ be given. Define the sequence $\{z_1^{(k)}, \ldots, z_n^{(k)}\}_{k=0}^{\infty}$ as follows:

(a) $z_i^{(0)} = M_i - \varepsilon \quad (i=1, \ldots, n)$,

(b) assume that $(z_1^{(k+1)}, \ldots, z_{j-1}^{(k+1)}, z_j^{(k)}, \ldots, z_n^{(k)})$ is defined such that $M_i - \varepsilon \leqq z_i^{(k+1)} \leqq M_i$, $i=1, \ldots, j-1$, $M_i - \varepsilon \leqq z_i^{(k)} \leqq M_i$, $i=j, \ldots, n$. Let $z_j^{(k+1)}$ be chosen according as $H_j(z_1^{(k+1)}, \ldots, z_{j-1}^{(k+1)}, z_j^{(k)}, \ldots, z_n^{(k)}) \leqq 0$ or >0. If $H_j(z_1^{(k+1)}, \ldots, z_{j-1}^{(k+1)}, z_j^{(k)}, \ldots, z_n^{(k)}) \leqq 0$ then let $z_j^{(k+1)} = z_j^{(k)}$. If $H_j(z_1^{(k+1)}, \ldots, z_{j-1}^{(k+1)}, z_j^{(k)}, \ldots, z_n^{(k)}) > 0$ then choose $z_j^{(k+1)}$ such that $z_j^{(k)} < z_j^{(k+1)} \leqq M_j$ and $H_j(z_1^{(k+1)}, \ldots, z_{j-1}^{(k+1)}, z_j^{(k+1)}, z_{j+1}^{(k)}, \ldots, z_n^{(k)}) = 0$. Since $H_j(z_1^{(k+1)}, \ldots, z_{j-1}^{(k+1)}, M_j, z_{j+1}^{(k)}, \ldots, z_n^{(k)}) \leqq H_j(M_1, \ldots, M_n) = 0$, the number $z_j^{(k+1)}$ exists.

Since the sequence $\{z_i^{(k)}\}_{k=0}^{\infty}$ is monotone nondecreasing and bounded, $z^*(\varepsilon)$

can be defined by $z_i^* = \lim_{k\to\infty} z_i^{(k)}$, $i=1, \dots, n$, $z^*(\varepsilon)=(z_1^*, \dots, z_n^*)$. We now prove that $z^*(\varepsilon)$ has properties (iv, (v), (vi). By the definition of $z^*(\varepsilon)$ (iv) is obviously satisfied. If (vi) is not true, then from $\sum_{i=1}^{n} H_i(z_1^*, \dots, z_n^*)=0$ it follows that there is an $i_0 \in \{1, \dots, n\}$ such that $H_{i_0}(z_1^*, \dots, z_n^*)>0$. Since H_{i_0} is continuous, one can find a number N such that $H_{i_0}(z_1^{(k+1)}, \dots, z_{i_0}^{(k+1)}, z_{i_0+1}^{(k)}, \dots, z_n^{(k)})>(1/2)H_{i_0}(z_1^*, \dots, z_n^*)>0$ for $k \geqq N$. But this contradicts the definition of $z_{i_0}^{(k+1)}$. If (v) is not true, then we can choose a number k_i for every $i \in \{1, \dots, n\}$ such that $z_i^{(k_i+1)} > z_i^{(k_i)} = M_i - \varepsilon$. Let $k_0 = \max\{k_1, \dots, k_n\}$ and $j = \max\{i \in \{1, \dots, n\} : k_i = k_0\}$. The definition of the sequence $\{z_i^{(k)}\}$ implies $H_i(z_1^{(k_i+1)}, \dots, z_{i-1}^{(k_i+1)}, z_i^{(k_i)}, \dots, z_n^{(k_i)})>0$, $i=1, \dots, n$. From the structure of H_i, the monotonicity of functions h_{ij} and the construction of $\{z_i^{(k)}\}$ it follows that $H_i(z_1^{(k_0+1)}, \dots, z_{j-1}^{(k_0+1)}, z_j^{(k_0)}, \dots, z_n^{(k_0)}) \geqq 0$ for $i \neq j$. Thus $\sum_{i=1}^{n} H_i(z_1^{(k_0+1)}, \dots, z_{j-1}^{(k_0+1)}, z_j^{(k_0)}, \dots, z_n^{(k_0)})>0$, which is a contradiction.

The following lemma includes the nonnegativity of trajectories and a comparison result.

Lemma 4 [13, Theorems 1,3]. *If* $\varphi, \psi \in C$, $\psi_i(s) \geqq \varphi_i(s) \geqq 0$ *for* $s \in [-r, 0]$, $i=1, \dots, n$, *and* $x(\varphi)(\cdot)$, $x(\psi)(\cdot)$ *are solutions of* (1) *on* $[-r, \infty)$ *through* $(0, \varphi)$, $(0, \psi)$, *then* $x_i(\psi)(t) \geqq x_i(\varphi)(t) \geqq 0$ *for* $t \in [0, \infty)$, $i=1, \dots, n$.

Lemma 5 [6, Theorem 3.1]. *Assume that* $\varphi \in C$ *and* $x(\varphi)$ *is a bounded solution of* (1) *on* $[-r, \infty)$. *If there exists a nonempty set* $H \subset (0, r]$ *such that*

(i) $\dot{x}_1(\psi)(0) \leqq 0$ *for all* $\psi \in \Omega(\varphi)$ *such that* $\psi_1(0) = \max_{-r \leqq s \leqq 0} \psi_1(s)$;

(ii) $\{\psi_1(-u) : u \in H\} = \{\psi_1(0)\}$ *for all* $\psi \in \Omega(\varphi)$ *such that* $\dot{x}_1(\psi)(0)=0$ *and* $\psi_1(0) = \max_{-r \leqq s \leqq 0} \psi_1(s)$;

(iii) *either there exist* $r_1, r_2 \in H$ *such that* r_1/r_2 *is irrational or these set* H *is infinite;*

then for any $\psi \in \Omega(\varphi)$ *the limit* $\lim_{t\to\infty} x_1(\psi)(t)$ *exists.*

Lemma 6. *Assume that* $\varphi \in C$, $x(\varphi)$ *is a bounded solution of* (1) *on the interval* $[-r, \infty)$ *and* $\psi \in \Omega(\varphi)$. *If the limit* $\lim_{t\to\infty} x(\psi)(t)$ *does not exist, then there are subsets* N_1, N_2 *of* $\{1, \dots, n\}$ *and real numbers* c_i, $i \in \{1, \dots, n\} \setminus N_1$, *such that*

(i) $N_2 \subseteqq N_1$;

(ii) $x_i(\psi)(\cdot) \equiv c_i$ *for* $i \in \{1, \dots, n\} \setminus N_1$;

(iii) *the limit* $\lim_{t\to\infty} x_i(\psi)(t)$ *does not exist for all* $i \in N_1$;

(iv) D_{N_2} *is a dicomponent of* D_{N_1};

(v) *for every* $i \in N_2$

$$\dot{x}_i(\psi)(t) = -\sum_{j \in N_2 \cup \{0\}} \tilde{h}_{ji}(x_i(\psi)(t)) + \sum_{j \in N_2} \int_0^{\tau} \tilde{h}_{ij}(x_j(\psi)(t-s))\, dF_{ij}(s) + \tilde{I}_i \qquad (t \in R), \tag{5}$$

where $\tilde{h}_{ij}(\cdot)=h_{ij}(\cdot)$, $i, j\in N_2$, $\tilde{h}_{0i}(\cdot)=h_{0i}(\cdot)+\sum_{j\in N\setminus N_2} h_{ji}(\cdot)$, $\tilde{I}_i=I_i+\sum_{j\in N\setminus N_1} h_{ij}(c_j)$, $i\in N_2$.

Proof. Let $N_0=\{i\in\{1, \dots, n\}$: the limit $\lim_{t\to\infty} x_i(\psi)(t)$ exists$\}$ and $c_i=\lim_{t\to\infty} x_i(\psi)(t)$, $i\in N_0$. From the definition of $\Omega(\varphi)$ and Corollary 2 it follows that $\lim_{t\to\infty} x_i(\varphi)(t)=c_i$ and $x_i(\psi)(\cdot)\equiv c_i$, $i\in N_0$. Let $N_1=\{1, \dots, n\}\setminus N_0$ and define the set N_2 as follows: D_{N_2} is a dicomponent of D_{N_1} such that if $i\in N_2$ and $j\in N_1\setminus N_2$ then $a_{ij}\notin V(D_{N_1})$. Clearly (iii), (iv), (v) are satisfied.

4. Convergence of the bounded solutions

In this section we give a sufficient condition for the existence of the limit of bounded solutions of (1).

Theorem. *If for every $i, j\in\{1, \dots, n\}$ either function $h_{ij}(\cdot)$ is strictly monotone increasing or $h_{ij}(\cdot)\equiv 0$, then, for each $\varphi\in C$ such that $x(\varphi)$ is a bounded solution of* (1) *on $[-r, \infty)$, the limit $\lim_{t\to\infty} x(\varphi)(t)$ exists.*

Proof. Assume that $\varphi\in C$, $x(\varphi)$ is a bounded solution of (1) on $[-r, \infty)$ and $\lim_{t\to\infty} x(\varphi)(t)$ does not exist. By Corollary 2 if $\psi\in\Omega(\varphi)$ then $\lim_{t\to\infty} x(\psi)(t)$ does not exist, either. Using Lemma 6 one can construct the equation (5), which has a bounded solution on $[-r, \infty)$ such that its components do not tend to constant as $t\to\infty$. Our aim is to show that equation (5) has not such a solution. This contradiction will prove Theorem.

Since (5) is a special case of (1), without loss of generality we can assume that $N_1=N_2=\{1, \dots, n\}$ in Lemma 6, i.e. $x(\varphi)$ is a solution of (1) on $[-r, \infty)$ such that for every $i\in\{1, \dots, n\}$ the limit $\lim_{t\to\infty} x_i(\varphi)(t)$ does not exist.

Let $M_i=\overline{\lim_{t\to\infty}}\, x_i(\varphi)(t)$ and $m_i=\underline{\lim_{t\to\infty}}\, x_i(\varphi)(t)$, $i=1, \dots, n$. By Corollary 2 and the definition of $\Omega(\varphi)$, for every $\psi\in\Omega(\varphi)$

(6) $$M_i=\overline{\lim_{t\to\infty}}\, x_i(\psi)(t),\quad m_i=\underline{\lim_{t\to\infty}}\, x_i(\psi)(t)\quad (i=1, \dots, n)$$

and

(7) $$m_i\leqq x_i(\psi)(t)\leqq M_i\quad (t\in R;\ i=1, \dots, n).$$

We now show that for every $\psi\in\Omega(\varphi)$

(8) $$\max_{-r\leqq s\leqq 0} \psi_i(s)=M_i\quad (i=1, \dots, n)$$

and

(9) $$\min_{-r\leqq s\leqq 0} \psi_i(s)=m_i\quad (i=1, \dots, n).$$

If (8) is not true for $\psi\in\Omega(\varphi)$, then without loss of generality one can assume that there exists $\varepsilon_0>0$ such that

(10) $$\max_{-r\leqq s\leqq 0}\psi_1(s)\leqq M_1-\varepsilon_0.$$

Let $i\in\{2, ..., n\}$ in the case $n>1$. Since v_1, v_i are diconnected in $D_{\{1,...,n\}}$, there exists a directed (v_i, v_1)-path $W=(a_{i_1 i_0}, a_{i_2 i_1}, ..., a_{i_m i_{m-1}})$ in $D_{\{1,...,n\}}$, where $i_0=1$, $i_m=i$. Suppose that for some $k\in\{0, 1, ..., m-1\}$ there is an $\varepsilon_k>0$ such that

(11) $$\max_{-r+k\tau\leqq s\leqq 0}\psi_{i_k}(s)\leqq M_{i_k}-\varepsilon_k.$$

From the strict monotonicity of function $h_{i_{k+1}i_k}(\cdot)$, (7), (11) and Lemma 2 it follows that on the set

$$S=\{t\in[-r+(k+1)\tau, 0]: \psi_{i_{k+1}}(t)=M_{i_{k+1}}\}$$

we have

$$\dot{x}_{i_{k+1}}(\psi)(t)\leqq-\sum_{j=0}^{n}h_{ji_{k+1}}(M_{i_{k+1}})+\sum_{\substack{j=1\\ j\neq i_k}}^{n}h_{i_{k+1}j}(M_j)+h_{i_{k+1}i_k}(M_{i_k}-\varepsilon_k)+I_{i_{k+1}}<$$

$$<H_{i_{k+1}}(M_1, ..., M_n)=0.$$

On the other hand, (7) and $x_{i_{k+1}}(t)=M_{i_{k+1}}$ imply $\dot{x}_{i_{k+1}}(\psi)(t)=0$, i.e. S is an empty set. Thus, there exists $\varepsilon_{k+1}>0$ such that

(12) $$\max_{-r+(k+1)\tau\leqq s\leqq 0}\psi_{i_{k+1}}(s)\leqq M_{i_{k+1}}-\varepsilon_{k+1}.$$

Since (10) is satisfied and (12) follows from (11), by using mathematical induction it can be seen that for some $\varepsilon_i>0$

$$\max_{-r+m\tau\leqq s\leqq 0}\psi_i(s)\leqq M_i-\varepsilon_i.$$

Since $i\in\{2, ..., n\}$ was arbitrary and W was a path, we have $m\leqq n-1$ and for some $\varepsilon>0$

(13) $$\max_{-\tau\leqq s\leqq 0}\psi_i(s)\leqq M_i-\varepsilon\quad(i=1, ..., n).$$

Apply Lemma 3: there exists $z^*=(z_1^*, ..., z_n^*)\in R^n$ such that $H_i(z_1^*, ..., z_n^*)=0$ and $M_i-\varepsilon\leqq z_i^*\leqq M_i$ for every $i\in\{1, ..., n\}$, $z_{i_0}^*=M_{i_0}-\varepsilon$ for some $i_0\in\{1, ..., n\}$. By (13) and Lemma 4 it follows that

$$x_{i_0}(\psi)(t)\leqq M_{i_0}-\varepsilon\quad(t\geqq 0),$$

which contradicts (6). Thus (8) is proved. By similar arguments one can show (9).

Let $T_{ij}\subset[0,\tau]$ denote the support of the Lebesgue—Stieltjes type measure induced by the distribution function F_{ij}, $i,j=1, ..., n$.

Define the set

$$H=(0,r]\cap\Big\{\bigoplus_{k=1}^{m} T_{i_k i_{k-1}}\colon (a_{i_1 i_0}, a_{i_2 i_1}, \dots, a_{i_m i_{m-1}}) \text{ is a directed } (v_1, v_1)\text{-walk in } D_{\{1,\dots,n\}}\Big\}.$$

See a special case at the end of this section. Since every two vertices are diconnected in $D_{\{1,\dots,n\}}$, for every $a_{ij}\in A(D_{\{1,\dots,n\}})$ there exists a directed (v_1, v_1)-walk W in $D_{\{1,\dots,n\}}$ such that $a_{ij}\in W$ and the length of W is at most $2n-1$. Thus, if H is empty, then $T_{ij}=\{0\}$ for every (i,j) such that $h_{ij}(\cdot)\not\equiv 0$, i.e. equation (1) is an ordinary differential equation. In this case τ can be an arbitrary small positive number. From this and (8), (9) it follows that $M_i=m_i$, $i=1,\dots,n$, which is a contradiction. Further on let us suppose that the set H is nonempty.

As regards the structure of set H we distinguish two cases:

Case. 1. Either there exist $r_1, r_2\in H$ such that r_1/r_2 is irrational or the set H is infinite.

Case 2. $H=\{p_1 r^*, p_2 r^*, \dots, p_K r^*\}$, where $r^*>0$, $0<p_1<\dots<p_K\leqq r/r^*$, p_i is an integer for each $i=1,\dots,K$ and $(p_1,\dots,p_K)=1$ (( ) denotes the greatest common divisor).

Case 1. Set H just satisfies condition (iii) of Lemma 5. (8) implies condition (i) of Lemma 5. To verify condition (ii) of Lemma 5 it will be sufficient to show that for each $\psi\in\Omega(\varphi)$, from $\dot{x}_1(\psi)(0)=0$, $\psi_1(0)=M_1$ it follows that $\psi_1(0)=\psi_1(-u)$ for all $u\in H$. Let $u=\sum_{k=1}^{m} t_{i_k i_{k-1}}\in H$, where $t_{i_k i_{k-1}}\in T_{i_k i_{k-1}}$, $k=1,\dots,m$. If $\dot{x}_1(\psi)(0)=0$ and $\psi_1(0)=M_1$ then by equation (1)

$$0=-\sum_{j=0}^{n} h_{j1}(M_1)+\sum_{j=1}^{n}\int_0^{\tau} h_{1j}\big(x_j(\psi)(-s)\big)\,dF_{1j}(s)+I_1. \tag{14}$$

From Lemma 2

$$0=-\sum_{j=0}^{n} h_{j1}(M_1)+\sum_{j=1}^{n}\int_0^{\tau} h_{1j}(M_j)\,dF_{1j}(s)+I_1. \tag{15}$$

From (8), (14), (15) and the monotonicity of functions $h_{1j}(\cdot)$ with the notation $i_m=1$

$$0=\int_0^{\tau}\big[h_{i_m i_{m-1}}(M_{i_{m-1}})-h_{i_m i_{m-1}}\big(x_{i_{m-1}}(\psi)(-s)\big)\big]\,dF_{i_m i_{m-1}}(s). \tag{16}$$

Since function $h_{i_m i_{m-1}}(\cdot)$ is strictly increasing and $t_{i_m i_{m-1}}\in T_{i_m i_{m-1}}$, (16) implies

$$\psi_{i_{m-1}}(-t_{i_m i_{m-1}})=M_{i_{m-1}}. \tag{17}$$

Using (8), (17) it is easy to see that $\dot{x}_{i_{m-1}}(\psi)(-t_{i_m i_{m-1}})=0$. Continuing this procedure we get

$$(18)\qquad \psi_{i_{k-1}}\Big(-\sum_{j=k}^{m} t_{i_j i_{j-1}}\Big) = M_{i_{k-1}} \quad (k = 1, \dots, m).$$

In the case $k=1$ relation (18) gives just $\psi_1(-u)=M_1$, which was to be proved. From (6) and Lemma 5 it follows $M_1=m_1$, which is a contradiction.

Case 2. Define the nonempty sets $A_0, A_1, \dots, A_m$ as follows:

(i) $\bigcup_{p=0}^{m} A_p=\{1, \dots, n\}$;

(ii) $A_0=\{1\}$;

(iii) $A_p=\{i\colon i\in\{1, \dots, n\}\setminus\bigcup_{k=0}^{p-1} A_k$ and there exists $j\in A_{p-1}$ such that $a_{ji}\in \in A(D_{\{1,\dots,n\}})\}$, $p=1, \dots, m$.

Let the function $S\colon \{2, 3, \dots, n\}\to\{1, 2, \dots, n\}$ eb defined in the following way: $S(i)\in A_{p-1}$ and $a_{S(i)i}\in A(D_{\{1,\dots,n\}})$ whenever $i\in A_p$, $p=1,\dots, m$. Let $\psi\in\Omega(\varphi)$, $y=x(\psi)$ and define the function

$$V(t) = \sum_{i=1}^{n} z_i(t),$$

where $z_1(t)=y_1(t)$ and

$$z_i(t) = \int_0^{\tau} \dots \int_0^{\tau} y_i\Big(t-\sum_{m=1}^{k} s_m\Big)\, dF_{i_k i_{k-1}}(s_k) \dots dF_{i_1 i_0}(s_1)$$

for $i=2, 3, \dots, n$, where $i_0, i_1, \dots, i_k$ are defined by $i_0=i$, $i_k=1$ and $S(i_m)=i_{m+1}$ for $m=0, 1, \dots, k-1$. Obviously $i_0, i_1, \dots, i_k$ may depend on i. Let $M_0=\sum_{i=1}^{n} M_i$ and $m_0=\sum_{i=1}^{n} m_i$. It is clear from (7) and the definition of V that

$$m_0 \leqq V(t) \leqq M_0 \quad (t\in R).$$

From the invariance of set $\Omega(\varphi)$ we have $y_t\in\Omega(\varphi)$ for all $t\in R$. By similar arguments as in Case 1, for every $t\in R$ from $y_1(t)=M_1$ it follows that $y_i(t-\sum_{m=1}^{k} s_m)=M_i$ whenever $s_m\in T_{i_m i_{m-1}}$, $m=1, \dots, k$; moreover $y_1(t-u)=M_1$ for each $u\in H$. Clearly, $V(t)=M_0$ implies $y_1(t)=M_1$. Thus, from $V(t)=M_0$ it follows that $V(t-u)=M_0$ for $u\in H$. Similarly, from $V(t)=m_0$ we obtain $V(t-u)=m_0$, $u\in H$. Hence and from (8), (9), (19) we have

$$(20)\qquad \max_{-r\leqq s\leqq 0} V(t+s) = M_0, \quad \min_{-r\leqq s\leqq 0} V(t+s) = m_0 \quad (t\in R).$$

Since $(p_1, \dots, p_K)=1$, from elementary number theory, there exist integers $n_1, \dots, n_K$ such that $\sum_{k=1}^{K} n_k p_k=1$. Let

$$N = \sum_{k=1}^{K} n_k^+ p_k - 1 \Big(= \sum_{k=1}^{K} n_k^- p_k\Big),$$

where n_k^+ and n_k^- are the positive and negative parts of n_k.

If $h=\sum_{k=1}^{K} a_k p_k$, where a_k is nonnegative integer, $k=1, \dots, K$, then hr^* is the sum of the elements of set H. Thus, from $V(t)=M_0$ and $V(t)=m_0$ it follows that $V(t-hr^*)=M_0$ and $V(t-hr^*)=m_0$, respectively. For every integer l, which is not less than N^2, the number lr^* is the sum of the elements of H. This is evident from the following:

$$l = N^2+k = N^2+aN+b = (N+a)N+b =$$

$$= (N+a)\sum_{k=1}^{K} n_k^- p_k + b\sum_{k=1}^{K} n_k p_k = \sum_{k=1}^{K} [(N+a)n_k^- + bn_k] p_k,$$

where k, a, b are nonnegative integers, $k=aN+b$, $b<N$.

Thus, from (20) it follows that there exist numbers $t_1, t_2 \in R$ such that

$$V(t_1 - ir^*) = M_0, \quad V(t_2 - ir^*) = m_0 \quad (i = 0, 1, 2, \dots). \tag{21}$$

From Lemma 2, (7) and the monotonicity of functions h_{0i} we have

$$\sum_{i=1}^{n} [-h_{0i}(y_i(t)) + I_i] = 0 \quad (t \in R).$$

Thus, by using that y is a solution of (1), one gets

$$\dot V(t) = \sum_{i=1}^{n} \dot z_i(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(t),$$

where

$$w_{11}(t) = \int_0^{\tau} h_{11}(y_1(t-s))\,dF_{11}(s) - h_{11}(y_1(t)),$$

$$w_{1i}(t) = \int_0^{\tau} h_{1i}(y_i(t-s))\,dF_{1i}(s) -$$

$$-\int_0^{\tau}\dots\int_0^{\tau} h_{1i}\Big(y_i\Big(t-\sum_{m=1}^{k} s_m\Big)\Big)\,dF_{i_k i_{k-1}}(s_k)\dots dF_{i_1 i_0}(s_1) \quad (i \geqq 2),$$

$$w_{j1}(t) = \int_0^{\tau}\dots\int_0^{\tau} h_{j1}\Big(y_1\Big(t-s-\sum_{m=1}^{l} s_m\Big)\Big)\,dF_{j1}(s)\,dF_{j_l j_{l-1}}(s_l)\dots dF_{j_1 j_0}(s_1) -$$

$$-h_{j1}(y_1(t)) \quad (j \geqq 2),$$

$$w_{ij}(t) = \int_0^{\tau}\dots\int_0^{\tau} h_{ij}\Big(y_j\Big(t-s-\sum_{m=1}^{k} s_m\Big)\Big)\,dF_{ij}(s)\,dF_{i_k i_{k-1}}(s_k)\dots dF_{i_1 i_0}(s_1) -$$

$$-\int_0^{\tau}\dots\int_0^{\tau} h_{ij}\Big(y_j\Big(t-\sum_{m=1}^{l} s_m\Big)\Big)\,dF_{j_l j_{l-1}}(s_l)\dots dF_{j_1 j_0}(s_1) \quad (i, j \geqq 2),$$

where $j_0, j_1, \ldots, j_l$ are defined by $j_0=j$, $j_l=1$ and $S(j_m)=j_{m+1}$ for $m=0, 1, \ldots, l-1$. Let W be a (v_j, v_1)-path in $D_{\{1,\ldots,n\}}$. Then $(W, a_{j_1 j_0}, a_{j_2 j_1}, \ldots, a_{j_l j_{l-1}})$ and $(W, a_{ij}, a_{i_1 i_0}, a_{i_2 i_1}, \ldots, a_{i_k i_{k-1}})$ are (a_1, a_1)-walks in $D_{\{1,\ldots,n\}}$ such that their lengths are at most $2n-1$. Thus, from the definition of H it follows that there exists a nonnegative u such that

$$u+s+\sum_{m=1}^{k} s_m = p_i r^* \quad \text{and} \quad u+\sum_{m=1}^{l} \tau_m = p_j r^*$$

for all $s\in T_{ij}$, $s_m\in T_{i_m i_{m-1}}$, $m=1, \ldots, k$, $\tau_m\in T_{j_m j_{m-1}}$, $m=1, \ldots, l$, for some nonnegative integers p_i, p_j, where p_i and p_j may depend on s_m, s, τ_m. That is, for $i, j\in \{1, \ldots, n\}$ functions $w_{ij}(t)$ have the following structure

$$w_{ij}(t) = \sum_{k=1}^{K_1} a_k v(t-b_k r^*+u) - \sum_{m=1}^{K_2} c_m v(t-d_m r^*+u),$$

where $\sum_{k=1}^{K_1} a_k = \sum_{m=1}^{K_2} c_m = 1$, $u\in R$, b_k and d_m are nonnegative integers for $k=1, \ldots, K_1$, $m=1, \ldots, K_2$, and the function $v\colon R\to R$ is bounded on R. See a special case at the end of this section.

If $|v(t)|\leq a$ for $t\in R$ and $b=\max_{k=1,\ldots,K_1, m=1,\ldots,K_2}\{b_k, d_m\}$, then

$$\frac{1}{L+1}\Big|\sum_{l=0}^{L}\Big(\sum_{k=1}^{K_1} a_k v(t-b_k r^*+u-lr^*) - \sum_{m=1}^{K_2} c_m v(t-d_m r^*+u-lr^*)\Big)\Big| =$$

$$= \frac{1}{L+1}\Big|\sum_{k=1}^{K_1} a_k \sum_{l=0}^{L} v(t+u-(b_k+l)r^*) - \sum_{m=1}^{K_2} c_m \sum_{l=0}^{L} v(t+u-(d_m+l)r^*)\Big| \leq$$

$$\leq \frac{1}{L+1}\Big|\sum_{k=1}^{K_1} a_k \sum_{s=b}^{L} v(t+u-sr^*) - \sum_{m=1}^{K_2} c_m \sum_{s=b}^{L} v(t+u-sr^*)\Big| + \frac{2ab}{L+1} =$$

$$= \frac{1}{L+1}\Big|\sum_{s=b}^{L} v(t+u-sr^*)\Big(\sum_{k=1}^{K_1} a_k - \sum_{m=1}^{K_2} c_m\Big)\Big| + \frac{2ab}{L+1} = \frac{2ab}{L+1} \to 0$$

as $L\to\infty$ uniformly in t on R. Hence we have

$$\frac{1}{L+1}\sum_{l=0}^{L} \dot V(t-lr^*) \to 0 \quad \text{as} \quad L\to\infty \tag{22}$$

uniformly in t on R.

On the other hand from (21) it follows that

$$\int_{t_2}^{t_1} \frac{1}{L+1}\sum_{l=0}^{L} \dot V(t-lr^*)\,dt = \frac{1}{L+1}\sum_{l=0}^{L}\int_{t_2-lr^*}^{t_1-lr^*} \dot V(t)\,dt = M_0-m_0 > 0$$

for all $L=0, 1, 2, \ldots$, which contradicts (22).

This completes the proof.

Remarks. The proof of Theorem for Case 2 is based on the idea of [6, Theorem 3.2].

We remark that the monotonicity conditions for functions h_{ij} cannot be omitted: if the functions are not monotone nondecreasing, then the equation (1) may have periodic solution [10].

We do not know whether the strict monotonicity conditions for h_{ij} is a necessary condition for the convergence of solutions of (1).

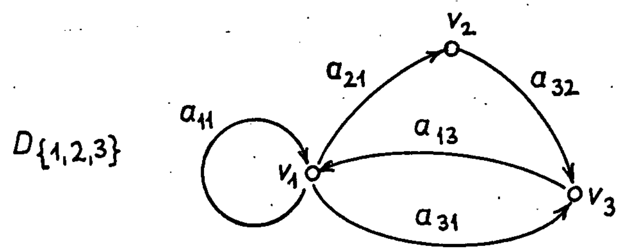
To illustrate the above proof we give a special case. Let us consider the system

$$\dot{x}_1(t) = -h_{11}(x_1(t))-h_{21}(x_1(t))-h_{31}(x_1(t))+h_{11}(x_1(t-1))+h_{13}(x_3(t-2))$$

$$\dot{x}_2(t) = -h_{32}(x_2(t))+h_{21}(x_1(t-1))$$

$$\dot{x}_3(t) = -h_{13}(x_3(t))+h_{31}(x_1(t-2))+\frac{1}{2}h_{32}(x_2(t))+\frac{1}{4}h_{32}(x_2(t-1))+\frac{1}{4}h_{32}(x_2(t-2)),$$

where functions $h_{11}, h_{21}, h_{31}, h_{13}, h_{32}$ are strictly increasing. Here directed graph $D_{\{1,2,3\}}, \tau, r, T_{ij}, H, A_p, V(t)$ and $\dot{V}(t)$ are the following:



$$\tau = 3; \quad r = 18;$$

$$T_{11} = \{1\}, \quad T_{13} = \{2\}, \quad T_{21} = \{1\}, \quad T_{31} = \{2\}, \quad T_{32} = \{0, 1, 2\};$$

$$H = \{1, 2, 3, \ldots, 18\};$$

$$A_0 = \{1\}, \quad A_1 = \{3\}, \quad A_2 = \{2\},$$

$$V(t) = y_1(t)+y_3(t-2)+y_2(t-2)/2+y_2(t-3)/4+y_2(t-4)/4;$$

$$\dot{V}(t) = [-h_{11}(y_1(t))+h_{11}(y_1(t-1))]+[-h_{31}(y_1(t))+h_{31}(y_1(t-4))]+$$
$$+[-h_{21}(y_1(t))+h_{21}(y_1(t-3))]/2+[-h_{21}(y_1(t))+h_{21}(y_1(t-4))]/4+$$
$$+[-h_{21}(y_1(t))+h_{21}(y_1(t-5))]/4.$$

Acknowledgements. The author would like to thank Professors I. Győri, L. Hatvani, L. Pintér and J. Terjéki for many discussions.

References

[1] G. L. ATKINS, *Multicompartment Models for Biological Systems,* Methuen (London, England, 1969).

[2] R. BELLMAN, Topics in pharmacokinetics II: Identification of time-lag processes, *Math. Biosci.,* **11** (1971), 1—11.

[3] I. GYŐRI, Delay differential and integro-differential equations in biological compartmental models, *Systems Sci.,* **8** (1982), 167—187.

[4] I. GYŐRI, Connections between compartmental systems with pipes and integro-differential equations, in: *Proceedings of the first conference of Rythms, Oscillations and Models in Biology* (1981), **8** (1982), 167—187.

[5] I. GYŐRI and J. ELLER, Compartmental systems with pipes, *Math. Biosci.,* **53** (1981), 223—247.

[6] J. R. HADDOCK and J. TERJÉKI, Liapunov—Razumikhin functions and an invariance principle for functional differential equations, *J. Differential Equations,* **48** (1983), 95—122.

[7] J. K. HALE, *Theory of Functional Differential Equations,* Springer-Verlag (New York, 1977).

[8] J. A. JACQUEZ, *Compartmental Analysis in Biology and Medicine,* Elsevier (New York, 1972).

[9] G. S. LADDE, Cellular systems. I, II, *Math. Biosci.,* **29** (1976), 309—330; **30** (1976), 1—21.

[10] V. LAKSHMIKANTHAM and S. LEELA, *Differential and Integral Inequalities,* Vol. I., Academic Press (New York, 1969).

[11] R. M. LEWIS and B. D. O. ANDERSON, Insensitivity of a class of nonlinear compartmental systems to the introduction of arbitrary time delays, *IEEE Trans. Circuits and Systems,* **27** (1980), 604—612.

[12] H. MAEDA, S. KODAMA and Y. OHTA, Asymptotic behaviour of nonlinear compartmental systems: nonoscillation and stability, *IEEE Trans. Circuits and Systems,* **25** (1978), 372—378.

[13] Y. OHTA, Qualitative analysis of nonlinear quasimonotone dynamical systems described by functional differential equations, *IEEE Trans. Circuits and Systems,* **28** (1981), 138—144.

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Acta Sci. Math. 47 (1984), 487—489

A formula for the solution of the difference equation $x_{n+1}=ax_n^2+bx_n+c$

DIETMAR DORNINGER and HELMUT LÄNGER

There are many papers dealing with the qualitative behaviour of the solution of the difference equation $x_{n+1}=ax_n^2+bx_n+c$, but up to now no explicit formula for the solution is known. (For a survey of results cf. [2].) In the following we deduce such a formula in a graph-theoretic context.

By a graph (V, E) with the vertex-set V and the set of edges E we mean an undirected graph without loops and without multiple edges. Thus the set E of edges of a graph (V, E) can be considered as a set of unordered pairs $\{v, w\}$, where v, w belong to the set V. A graph (V, E), wherein a certain vertex v_0 is distinguished as the "root" of the graph, will be called a rooted graph and will be denoted by (V, E, v_0). A rooted graph S which is a subgraph of a rooted graph G will be called a rooted subgraph of G, if the roots of S and G coincide.

Definition 1. For any non-negative integer n let T_n denote the rooted graph

$$(P(\{1, \ldots, n\}), \{\{M, M\setminus\{\max M\}\} | \emptyset \neq M \subseteq \{1, \ldots, n\}\}, \emptyset)$$

where $P(\{1, \ldots, n\})$ denotes the power set of $\{1, \ldots, n\}$ and $\max M$ the maximum number occurring within the subset M of $\{1, \ldots, n\}$.

Remark. T_n can be easily constructed inductively by observing $T_0=(\{\emptyset\}, \emptyset, \emptyset)$ and

$$T_{n+1} = (V(T_n)\cup\{M\cup\{n+1\} | M\in V(T_n)\}, E(T_n)\cup\{\{M, M\cup\{n+1\}\} | M\in V(T_n)\}, \emptyset)$$

for all $n \geqq 0$.

We say that a vertex M of T_n has cardinality k if the cardinality $|M|$ of the set M is k.

Lemma. *For any non-negative integer n, T_n is a rooted tree.*

Received March 1, 1983.

Proof. Let n be some fixed non-negative integer. Studying the definition of T_n one can see easily that there are no loops and that there always exists a path connecting an arbitrary vertex of T_n with the root $\emptyset$. Thus T_n is a connected graph (without loops). If T_n would contain a circle C, then C would have to have at least three vertices, since there are no loops and no double edges in T_n. Assume, M is a vertex of maximal cardinality of C. Then, by the definition of T_n, the vertices of C being adjacent to M would have to coincide, which is a contradiction. Hence T_n is a tree.

Definition 2. For a graph $G=(V(G), E(G))$ and for any subgraph $S=(V(S), E(S))$ of G let S_G denote the complete subgraph of G which has the vertex-set

$$V(S_G) = V(S)\cup\{x\in V(G)\,|\,\text{there exists some } y\in V(S) \text{ such that } \{x, y\}\in E(G)\}.$$

For a rooted graph G and for any rooted subgraph S of G the rooted subgraph S_G of G is defined analogously.

Theorem. *Let I be an arbitrary integral domain. Then the solution of the difference equation $x_{n+1}=ax_n^2+bx_n+c$ $(a, b, c\in I;\ n\geqq 0)$ is given by $x_n=x_0+nc$ if $(a, b)=(0, 1)$ and*

$$x_n = \bar{x}+\sum a^{|V(S)|-1}(f'(\bar{x}))^{|V(S_{T_n})\setminus V(S)|}(x_0-\bar{x})^{|V(S)|}$$

otherwise. Thereby $f(x)$ denotes the polynomial function ax^2+bx+c, $\bar{x}$ is an arbitrary fixed point of f (which in case $(a, b)\neq(0, 1)$ exists in a suitable extension field of I) and the sum is taken over all rooted subtrees S of T_n. (By definition $0^0:=1$.)

Proof. The solution in case $(a, b)=(0,1)$ is obvious. Therefore assume $(a, b)\neq(0, 1)$.

Then within the algebraic closure K of the quotient field of I there exists some fixed point of f, say $\bar{x}$. Performing the substitution $x_n=\bar{x}+y_1^{(n)}$ the difference equation $x_{n+1}=f(x_n)$ is transformed into the difference equation

$$y_1^{(n+1)} = y_1^{(n)}(ay_1^{(n)}+f'(\bar{x})). \tag{1}$$

Now consider the system

$$\begin{aligned} y_1^{(n+1)} &= y_1^{(n)}(ay_1^{(n)}+f'(\bar{x})\,y_2^{(n)}) \\ y_2^{(n+1)} &= y_2^{(n)}(0y_1^{(n)}+1y_2^{(n)}) \end{aligned} \tag{2}$$

of difference equations over K. As one can see easily, $y_1^{(n)}$ is a solution of (1) with the initial value $y_1^{(0)}$ if and only if $(y_1^{(n)}, 1)$ is a solution of (2) with the initial value $(y_1^{(0)}, 1)$. To solve the system (2) one can apply the formula

$$y_1^{(n)} = y_1^{(0)} \sum_{\substack{g:P(\{1,\dots,n\})\to\{1,2\} \\ g(\emptyset)=1}} \prod_{M:\emptyset\neq M\subseteq\{1,\dots,n\}} (a_{g(M\setminus\{\max M\}),g(M)}\, y_{g(M)}^{(0)})$$

(which was proved in [1]) where in our case $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix}=\begin{pmatrix} a & f'(\bar{x}) \\ 0 & 1\end{pmatrix}$.

Performing the index transformation $g \leftrightarrow V := g^{-1}(\{1\})$ we get

$$y_1^{(n)} = \sum a^{|V|-1}(f'(\bar{x}))^{|\{M\in P(\{1,\dots,n\})\setminus V \mid M\setminus\{\max M\}\in V\}|}(y_1^{(0)})^{|V|}$$

where the sum is taken over all subsets V of $P(\{1, \dots, n\})$ which contain the empty set as an element and have the property $\emptyset\neq M\in V\Rightarrow M\setminus\{\max M\}\in V$. We claim that the sets V are exactly the vertex-sets of the rooted subtrees of T_n. Given a set V one can see immediately that within the complete subgraph of T_n with vertex-set V there exists a path connecting each element of V with $\emptyset$. Thus the complete subgraph of T_n having V as its set of vertices is connected and hence is a rooted subtree of T_n. Conversely, let S be a rooted subtree of T_n. Then from each vertex M of S with $|M|\geqq 1$ to the root $\emptyset$ we can find a path $M=M_0, M_1, \dots, M_k=\emptyset$ $(k\geqq 1)$ within S. $|M_1|>|M_0|$ would imply $k>1$ and $|M_{m-1}|=|M_{m+1}|$ and hence $M_{m-1}=M_{m+1}$ for $m:=\min\{i \mid 1\leqq i<k,\ |M_{i+1}|<|M_i|\}$ contradicting the definition of a path. Therefore $|M_1|<|M_0|$ which implies $M_0\setminus\{\max M_0\}=M_1\in V(S)$. This shows that with every non-empty vertex M, S also contains the vertex $M\setminus\{\max M\}$ wherefrom we can conclude

$$y_1^{(n)} = \sum a^{|V(S)|-1}(f'(\bar{x}))^{|V(S_{T_n})\setminus V(S)|}(y_1^{(0)})^{|V(S)|},$$

the sum being taken over all rooted subtrees S of T_n. Replacing $y_1^{(n)}$ by $x_n-\bar{x}$ yields the result of the theorem.

Remark. If $a=f'(\bar{x})=1$, then $x_n=\bar{x}+\sum_{i=1}^{2^n} b_{ni}(x_0-\bar{x})^i$ where for all $n\geqq 0$ and for all i with $1\leqq i\leqq 2^n$, b_{ni} denotes the number of all rooted subtrees of T_n with exactly i vertices.

References

[1] H. LÄNGER, A formula for the solution of a system of difference equations over a commutative ring, in: *Contributions to General Algebra 2* (Proc. Klagenfurt Conf. 1982, ed. G. Eigenthaler, H. K. Kaiser, W. B. Müller and W. Nöbauer), Hölder—Pichler—Tempsky and Teubner (Wien and Stuttgart, 1983), 233—238.

[2] CH. PRESTON, *Iterates of maps on an interval*, Lecture Notes in Mathematics 999, Springer-Verlag (Berlin, 1983).

TECHNISCHE UNIVERSITÄT WIEN
INSTITUT FÜR ALGEBRA UND DISKRETE MATHEMATIK
WIEDNER HAUPTSTRASSE 8-10
1040 WIEN, AUSTRIA

Bibliographie

G. **Alexits, Approximation Theory** (Selected Papers), 298 pages, Akadémiai Kiadó, Budapest, 1983.

The volume is a collection of selected papers by George Alexits. It is my deep-seated conviction that this collection is of great value of mathematics. The thirthy-four articles included here cover a wide field of real analysis and show the characteristic mathematical style of Alexits, the admirably clear exposition of his profound mathematical ideas. More precisely the volume presents articles on approximation theory, the papers developing the theory of multiplicative function systems, and the recent items on function series. The earlier function-theoretic, set-theoretic and curve-theoretic papers of Alexits and his works on the history of mathematics have been left out together with those papers on the theory of function series, the results of which were incorporated in his monograph "Konvergenztheorie der Orthogonalreihen" (Akadémiai Kiadó, Budapest, 1960) published also in English and in Russian. The papers are reprinted in their original form, with the only exception being the English translation of an article originally published in Hungarian. In my view this article is one of the most significant papers of Alexits. In it he characterizes the Lipschitz class of order $\alpha=1$ by the order of approximation given by the Cesàro-means of the conjugate Fourier series. This paper was published in a Hungarian journal in 1941, presumably this was the reason that the result was reproved later in parts by A. Zygmund (1945) and M. Zamansky (1949). In addition to the papers, the volume contains a short description of the life and scientific activities of George Alexits and the full list of his scientific works. At the end of the book are some remarks and a list of errata. These remarks briefly discribe the effect of the presented papers and the further developments resulting from them, moreover they give references to later results, while the list of errors corrects some oversights and misprints in the originals.

The significance of Alexits' contributions to many areas of mathematics is nowadays well known. But, for the sake of correctness, it is necessary to mention in connection with the "Remarks" on p. 287 that the cited monograph of R. A. DeVore was not the first to give international recognition to the fact that Alexits proved already in 1941 both necessity and sufficiency of the characterization of the Lipschitz class $\alpha=1$ by (C, 1)-summation. The first monograph emphasizing this was that of P. L. Butzer and R. J. Nessel *Fourier Analysis and Approximation,* Birkhauser Verlag, Basel—Suttgart, 1971.

I am convinced that Professor Alexits had a wide international reputation by the time when his monograph on the convergence and summation problems of orthogonal series appeared in 1960 in three languages. This monograph has become one of the most cited works in the field of orthogonal series. Alexits was one of the most influential Hungarian mathematicians. He created a scientific school having numerous pupils in Hungary and all over the world.

Mathematicians working in approximation theory will surely find it very useful to have these selected papers of Alexits in one volume.

L. Leindler (Szeged)

V. I. Arnold, Geometrical Methods in the Theory of Ordinary Differential Equations (Grundlehren der Mathematischen Wissenschaften, 250), XI+334 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1983.

V. I. Arnold, Catastrophe Theory, 79 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

The title of the original edition of the first book is „Дополнительные главы теории обыкновенных дифференциальных уравнений" (Supplementary Chapters to the Theory of Ordinary Differential Equations). The translator (or the editor of the translation) chose the new title rightly because it characterizes both the topics and the treatment of the book. However, it is worth recording the original title, which shows that the present book is the continuation and supplement of the author's excellent introductory text-book crowned success, and that the book consists of *almost independent chapters.*

The first two chapters deal with special equations (differential equations invariant under groups of symmetries, implicit equations, the stationary Schrödinger equation, second order differential equations, first order partial differential equations) and present classical results, that can be found in most monographs. Nevertheless, after having read these chapters the reader feels as if he had been just acquainted with these results because their deep mysteries have become clear and understandable, setting the facts in their true light.

Chapter 3 is devoted to structural stability. In the real world there always exist small perturbations, which cannot be taken into consideration in the mathematical models. It is clear, that only those properties of the model may be viewed as the properties of the real process which are not very sensitive to a small change in the model. The investigation of these properties led to the notion of structural stability.

The organization of the chapter is typical Arnold. First he gives the naive definition of structural stability and illuminates it by examples. Then he gathers together the necessary tools and gives the final precise definition of structural stability. The definition is followed by a detailed analysis of the one-dimensional case, which helps the reader to intensify the new notion. Then he presents a survey on the differential equation on the torus, hyperbolic theory and Anosov systems.

Chapter 4 is concerned with perturbation theory. In the theory of differential equations there are some equations of special form (e.g. linear equations) which admit an exact analytic solution or a complete qualitative description. Perturbation theory gives methods for the study of equations close to one with known properties. One of the most important sections of this theory is the averaging method that has been used among others in the celestial mechanics since the time of Lagrange and Laplace. "Nevertheless, the problem of strict justification of the averaging method is still far from being solved" — writes the author, and the reviewer can recommend this part of the book as an excellent comprehensive introduction to this interesting and actual topic.

In Chapter 5 the reader finds Poincaré's theory of normal form, which is a very useful device in many topics such as in bifurcation theory, to which Chapter 6 is devoted. In the models of the real world, in general, there are some parameters. It may happen that arbitrarily small variations of the parameters at fixed values cause essential change of the pictures of the solutions. This phenomenon is called bifurcation. The author studies bifurcations of phase portraits of dynamical systems in the neighbourhood of equilibrium positions and closed trajectories.

The subject-matter of the second book (or booklet) can also be considered as a chapter of the geometrical theory of dynamical systems. The origins of catastrophe theory lie in Whitney's theory of singularities of smooth mappings and the bifurcation theory of dynamical systems. Interpreting — not always mathematically — the results of these theories, catastrophe theory tries to provide a universal method for the study of all jump transitions, discontinuities and sudden qualitative changes. It has aroused a great controversy not only among specialists but also in the popular press. This booklet explains what catastrophe theory is about and why it arouses such a controversy.

While the first book is of advanced level, the second one can be recommended also "to readers having minimal mathematical background but the reader is assumed to have an inquiring mind".

L. Hatvani (Szeged)

Bernard Aupetit, Propriétés Spectrales des Algebres de Banach (Lecture Notes in Mathematics, 735) X+192, Springer-Verlag, Berlin—Heidelberg—New York, 1979.

This nice book collects the results obtained till its publication in connection with the spectra of Banach algebra elements. It contains mainly the author's own results, but gives historical backgrounds and performs the classical results as well. The recent development of this area and the absence of a comprehensive work on this subject make this book very interesting and useful.

The author's interest started from two, apparently remote problems. These were: to generalize Newburgh's theorem on the continuity of the spectrum and to generalize the theorem of Hirschfeld and Želazko on the characterization of commutative algebras. Mixing in a suprising manner the methods of these areas, the author obtained a characterization of finite-dimensional algebras. The use of subharmonic functions and deep results of classical potential theory in functional analysis provides the essential new feature of his technique.

The text consists of five chapters. Continuity problems of the spectrum, characterizations of commutative, finite-dimensional, symmetric and C^* algebras, respectively, are systematically treated. Abundance of examples and counterexamples complete the discussions. Two appendices, one on Banach algebras and the other on potential theory, help the reader and make the text available for a wide audience. The book is recommended to everyone who is interested in this new field of functional analysis.

L. Kérchy (Szeged)

David Bleecker, Gauge Theory and Variational Principles (Global Analysis, Pure and Applied, Series A, No. 1), XX+179 pages, Addison—Wesley, London—Amsterdam—Don Mills—Ontario—Tokyo, 1981.

The present book is the first number of a new series on pure mathematics and applications of global analysis based on ideas of classical analysis and geometry. Series B will provide a collection of prerequisites for the reports of series A from the research frontiers.

The most successful models of the fundamental interactions of the matter as well as the most hopeful candidates for their unification are all gauge theories with local symmetries. The majority of developments of classical gauge field theory in the last 10 years is connected with the global aspects of the underlying fibre bundle theory. This book contains a detailed account of bundle theoretic foundations of gauge theory.

The author's point of view, that the particle fields are functions on the corresponding principal bundles, leads to very elegant formulation of the variational problems and Euler—Lagrange equations involved. This is done in Chapters 3—5 based on the geometric notations of the previous ones.

A short, clear explanation of the free Dirac's equation as Lagrange's equation for the Dirac spinor field on the spin bundle with Levi—Civita connection can be found in Chapter 6. In Chapter 7 a general framework is given for the unification of interactions, based on a construction to form a principal bundle with product group, a connection and Lagrangian on it from the principal bundles and their connections which are connected with the fields that are to be incorporated in a unified theory. The general scheme is applied to the Dirac electron field coupled to electromagnetic potential and to the original Yang—Mills nucleon model. In Chapter 8 the author treats the tensor calculus on a (pseudo-) Riemann manifold in the frame bundle picture. Chapter 9 is devoted to the unification of gravitation and Yang—Mills fields in the well-known Kaluza—Klein type way. The reality of the used canonical bundle metric is supported by calculation of its geodesics in Chapter 10, nicely intepreted as paths for the classical particle motion. Besides, Utiyama's theorem, the spontaneous symmetry breaking and the very basic notations of the characteristic classes in connection with the monopoles and instantons are treated within the additional topics of Chapter 10.

The book is very well organized, self-contained, concise and rigorous. In the preface and in the introduction to the chapters the intuitive ideas are also sketched by the author. It is highly recommended for everyone interested in gauge theory. Those working in the field as well as graduate students will find it useful without doubt.

L. Gy. Fehér (Szeged)

E. A. Coddington—H. S. V. de Snoo, Regular Boundary Value Problems Associated with Pairs of Ordinary Differential Expressions (Lecture Notes in Mathematics, 858), V+225 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1981.

This volume is devoted to the study of eigenvalue problems associated with pairs L, M of ordinary differential operators. The solutions f of $Lf=\lambda Mf$ subject to boundary conditions are considered. It is shown how these problems have a natural setting within the framework of subspaces in the direct sum of Hilbert spaces. A detailed discussion is worked out for the regular case, where the coefficients of the ordinary differential expressions L and M are sufficiently smooth and invertible functions on a closed bounded interval I, and M is positive in the sense that there exists a constant $c>0$ such that $(Mf,f)_2 \geqq c^2(f,f)_2$ for $f\in C_0^\infty(I)$. The key idea of the simultaneous diagonalization of two hermitian $n\times n$ matrices K, H, where $H>0$, is extended for the case where K, H are replaced by a pair of ordinary differential expressions L, M. The possible difficulties of the generalization are discussed in eleven chapters of this work. The authors say: "it is hoped that this detailed knowledge of the regular case will lead to a greater understanding of the more involved singular case".

The reader is assumed to have some familiarity with the main results proved in an earlier paper of the authors. We recommend these notes to everybody working in related fields of mathematics as well as to graduate students interested in the subject.

T. Krisztin (Szeged)

Combinatorial Mathematics X. Proceedings of the Conference held in Adelaide, Australia, August 23—27, 1982, edited by L. R. A. Casse (Lecture Notes in Mathematics, Vol. 1036), XI+419 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983.

These conference proceedings consist of seven invited papers and twenty-four contributed papers. According to the tradition of Australian conferences in combinatorial mathematics, a great part of the papers is concerned with finite geometries, Hadamard matrices, block designs and latin squares. Some papers investigate topics in combinatorial analysis, e.g. the Schröder—Etherington sequence, the solutions of $y^{(k)}(x)=y(x)$, and the method of generating combinatorial identities by stochastic processes.

The titles of invited papers are: C. C. Chen and N. Quimpo, Hamiltonian Cayley graphs of order pq; J. W. P. Hirschfeld, The Weil conjectures in finite geometry; D. A. Holton, Cycles in graphs;

A. D. Keedwell, Sequenceable groups, generalized complete mappings, neofields, and block designs; N. J. Pullmann, Unique coverings of graphs — A survey; D. Stinson, Room squares and subsquares; J. A. Thas, Geometries in finite projective spaces: recent results.

L. A. Székely (Szeged)

Complex Analysis and Spectral Theory (Seminar, Leningrad 1979/80), Edited by V. P. Havin and N. K. Nikol'skii (Lecture Notes in Mathematics, 864), IV+480 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1981.

This book may be considered as the third issue of selected works of the Seminar on Spectral Theory and Complex Analysis, organized by the Leningrad Branch of the Steklov Institute and the Leningrad University. It contains 9 papers written by the participants during the period 1979/80. The whole volume and most papers separately convincingly demonstrate how close the connection is between Spectral Theory and Complex Analysis both in their problems and methods.

The table of contents: 1. A. B. Aleksandrov, Essays on non Locally Convex Hardy Classes. — This paper contains a new approach to the problem of characterizing functions representable by Cauchy potential and at the same time, among others, gives a description of invariant subspaces of the shift operator. 2. E. M. Dyn'kin, The Rate of Polynomial Approximation in the Complex Domain. — This paper represents the classical Function Theory and provides a systematic exposition of the subject. 3. V. P. Havin, B. Jöricke, On a Class of Uniqueness Theorems for Convolutions. — This paper deals with a phenomenon of quasi-analicity exhibited by many operators commuting with translations. 4. S. V. Hruščev, S. A. Vinogradov, Free Interpolation in the Space of Uniformly Convergent Taylor Series. — The authors plan to publish in the future a survey of harmonic analysis in the space of the title and in the disc-algebra. The present paper collects some new results and some new approaches to the subject which appeared during their work. 5. S. V. Hruščev, N. K. Nikol'skii, B. S. Pavlov, Unconditional Basis of Exponentials and of Reproducing Kernels. — This nice paper contains a description of all subsets $\{\lambda_n\}_n$ of a half-plane $\{\lambda\in\mathbf{C}\colon \operatorname{Im}\lambda>\gamma\}$ such that the family $\{e^{i\lambda_n x}\}_n$ forms an unconditional basis in $L^2(I)$. (Here I is an interval of the real axis and the notion of unconditional basis is a slight generalization of the one of Riesz basis.) 6. S. V. Kisliakov, What is Needed for a 0-Absolutely Summing Operator to be Nuclear? — The results of this paper are concerned with the open problem: whether each continuous linear operator from the dual of the disc-algebra to a Hilbert space is 1-absolutely summing. 7. N. G. Makarov, V. I. Vasjunin, A Model for Non-contractions and Stability of the Continuous Spectrum. — The authors extend the Sz.-Nagy—Foiaş functional model from contractions to arbitrary bounded Hilbert space operators remaining in spaces with definite metrics and using auxiliary contractions. Applying this model they get nice results on the stability of the continuous spectrum in the case of "nearly unitary" operators. 8. N. A. Shirokov, Division and Multiplication by Inner Functions in Spaces of Analytic Functions Smooth up to the Boundary. — The results of this paper complete the list of basic classes X of "smooth analytic functions" with the property that for every function $f\in X$ and for every inner function I the relation $fI^{-1}\in X$ holds whenever fI^{-1} belongs to the Smirnov class. 9. A. L. Volberg, Thin and Thick Families of Rational Fractions. — A family of rational fractions $R_\Lambda=\{1/(z-\lambda)\colon \lambda\in\Lambda\}$, where $\Lambda\subset\{z\in\mathbf{C}\colon \operatorname{Im} z>0\}$, is called thick with respect to a Borel measure μ on the real line if R_Λ is dense in $L^2(\mu)$; R_Λ is called thin with respect to μ if e.g. the L^2-norms corresponding to μ and the Lebesgue measure are equivalent in the linear span of R_Λ. In this paper thick and thin families are described for measures with some properties.

L. Kérchy (Szeged)

Differential Equations Models (Edited by M. Braun, C. S. Coleman, D. A. Drew), XIX+380 pages;

Life Science Models (Edited by H. Marcus-Roberts, M. Thompson), XX+366 pages;
(Modules in Applied Mathematics, vol. 1, vol 4), Springer-Verlag, New York—Heidelberg—Berlin, 1983.

It is an old question even in the mathematical society "Why do people do mathematics?" There exist a great number of answers to this question from "We do mathematics because we enjoy doing mathematics" to "We do mathematics because it can be applied to the practice and other sciences". The first and last volume of the series "Modules in Applied Mathematics" convince us that good mathematics can be both enjoyable and applicable to the problems of the real world. These books show models which describe phenomena of nature or of the society and, simultaneously, they serve as a source of very interesting and very deep investigations in pure mathematics. For example, in population dynamics the co-existence of two interacting species is described by an autonomous system of two ordinary differential equations with polynomial right-hand sides. If the population shows periodical behaviour, then the system has a cycle as a trajectory. The following problem was posed by David Hilbert in 1900 and is still unsolved: what is the maximum number and position of the isolated cycles for a differential equations of this type?

Each chapter is concerned with a model. The construction of the chapters illustrates the steps of the method of the applied mathematics: the statement of the word problem; setting up to mathematical model; investigation of the model with the help of mathematical methods; the interpretation of the results.

The series has been written primarily for college teachers to be used in undergraduate programs. The independent chapters serve as the subject-matters of one-four lectures. Each chapter includes many exercises challenging the reader to further thinking, which are suitable to be posed for good students as well. Prerequisites for each chapter and suggestions for the teacher are provided.

The 23 chapters of the first volume are divided into six parts: I. Differential equations, models, and what to do with them; II. Growth and decay models: first order differential equations; III. Higher order linear models; IV. Traffic models; V. Interacting species: steady states of nonlinear systems; VI. Models leading to partial differential equations. Some of the most exciting problems: The Van Meegeren art forgeries; How long should a traffic light remain amber; Why the percentage of sharks caught in the Mediterranean Sea rose dramatically during World War I; The principle of competitive exclusion in population biology.

The fourth volume consists of three parts: I. Population models; II. Biomedicine: epidemics, genetics, and bioengineering; III. Ecology. The main mathematical devices used here are differential equations, probability theory, linear programming.

These excellent books will be very interesting and useful for both mathematicians interested in realistic applications of mathematics and those non-mathematicians wanting to know how modern mathematics is actually employed to solve relevant contemporary problems.

L. Hatvani (Szeged)

K. Donner, Extension of Positive Operators and Korovkin Theorems (Lecture Notes in Mathematics, 904) X+173 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982.

This book deals with positive and norm-preserving extensions of linear operators in Banach lattices. Imbedding of Banach lattices into cones with infinitely big elements (i.e. $R\cup(+\infty)$ is used instead of R) and using the tensor product method, a useful new technique is obtained for solving the problems mentioned above. The results lead to a simple description of Korovkin systems in L^p.

The text is divided into eight sections. The reader is supposed to be familiar with some basic knowledge in Banach lattice theory.

László Gehér (Szeged)

Dynamical Systems and Turbulence, Warwick 1980, Edited by D. A. Rand and L. S. Young (Lecture Notes in Mathematics, 898), VI+390 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1981.

The aim of the Organizing Committee was to bring together a wide variety of scientists from different backgrounds with a common interest in the problem of the dynamics of turbulence and related topics. The titles of some papers enumerated below show that this aim was fulfilled and so this volume is important and interesting for everyone who is interested in the general theory of dynamical systems.

There are two expository papers: D. Joseph: Lectures on bifurcation from periodic orbits; D. Schaeffer: General introduction to steady state bifurcation. Some of the contributed papers are: J. Guckenheimer: On a codimension two bifurcation; J. Hale: Stability and bifurcation in a parabolic equation; P. Holmes: Space- and time-periodic perturbations of the Sine-Gordon equation; I. P. Malta and J. Palis: Families of vector fields with finite modulus of stability; L. Markus: Controllability of multi-trajectories on Lie groups; W. de Melo, J. Palis and S. J. van Strien: Characterizing diffeomorphisms with modulus of stability one; S. J. van Strien: On the bifurcations creating horseshoes; F. Takens: Detecting strange attractors in turbulence.

L. Pintér (Szeged)

Emanuel Fischer, Intermediate Real Analysis, (Undergraduate Texts in Mathematics), XIV+770 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1983.

Today one finds a great number of books on introductory analysis, but sometimes the teacher cannot choose such a work from them which satisfies his students' background. The author — on the basis of his experience of many years — wrote a book for students who have completed a three-semester calculus course, possibly an introductory course in differential equations and one or two semesters of modern algebra. This determines the structure of the book and the spirit of the definitions and the proofs. Therefore, the author presents the material in "theorem — proof — theorem" fashion, interspersing definitions, examples and remarks.

The book is self-contained except for some theorems on finite sets.

At the end of Chapter XIV — having the title The Riemann Integral — we find Lebesgue's famous theorem: A function which is bounded on a bounded closed interval $[a, b]$ is Riemann-integrable if and only if the set of points in $[a, b]$ at which it is discontinuous has measure zero. We cited this theorem because in some sense it is characteristic for this book. The notions to understand this theorem are treated in the text, but the proof — which belongs to a next stage — is omitted. Nevertheless, the book concentrates on the specific and concrete by applying the theorems to obtain information about important functions of analysis.

Above all, this is a stylish book, well thought out and uses tested methods, which one could safely put into the hands of future users of mathematics. (There is an unexpected mistake in the Bibliography. Correctly the names of the authors of the world-famous book " Aufgaben und Lehrsätze aus der Analysis" are G. Pólya and G. Szegő.)

L. Pintér (Szeged)

G. B. Folland, Lectures on Partial Differential Equations (Tata Institute of Fundamental Research Lectures on Mathematics and Physics), VI+160 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1983.

This book consists of the notes for a course the author gave at the Tata Institute of Fundamental Research Centre in Bangalore in autumn 1981. The purpose of the course was the application of Fourier analysis (i.e. convolution operators as well as the Fourier transform itself) to partial differential equations. The book is divided into five chapters. In the first some basic results about convolutions and the Fourier transforms are given. In Chapter 2 the fundamental facts of partial differential operators with constant coefficients are studied. In the next one precise theory of L^2 differentiability is introduced to prove Hörmander's theorem on the hypoellipticity of constant coefficient differential operators. Chapter 4 comprises the basic theory of pseudo differential operators. The aim of the last chapter is to study how to measure the smoothing properties of pseudo differential operators of nonpositive order in terms of various important function spaces.

The reader is assumed to have familiarity with real analysis and to be acquainted with the basic facts about distributions. No specific knowledge of partial differential equations is assumed.

This book is directed to graduate students and mathematicians who are interested in the application of Fourier analysis.

T. Krisztin (Szeged)

F. Gécseg—M. Steinby, Tree Automata, 235 pages, Akadémiai Kiadó, Budapest, 1984.

The theory of tree automata is a relatively new field of theoretical computer science. More exactly, it is a new field of automata and formal language theory, though it has several aspects in common with flowchart theory, recursive program schemes, pattern recognition, theory of translations, mathematical logic, etc. The book of F. Gécseg and M. Steinby gives a systematic, mathematically rigorous summary of results on tree automata.

Every finite automaton, — more precisely, a finite-state recognizer — can be viewed as a finite universal algebra having unary operations only. This observation, though obvious, provides a way of generalization. Basically, a tree automaton is a finite universal algebra equipped with arbitrary finitary operations. However, problems investigated in the theory of tree automata essentially differ from that investigated in universal algebra. The introduction of tree automata as a new device was not only for the sake of generalizing automata theory. As explained in this book, the connection with context free grammars and languages, syntax directed translations, and other topics has been significant and vitally important.

The book consists of four chapters, a bibliography, and an index. Chapter I comprises an exposition on necessary universal algebra, lattice theory, finite automata and formal languages. Section 1 presents the terminology. Sections 2 and 3 recall some basic concepts of universal algebra, including terms, polynomials and free algebras. Section 4 deals with lattices, complete lattices, and a variant of Tarski's fixed-point theorem. Section 5 surveys finite-state recognizers and their relation to regular languages. Besides the various characterizations of regular languages, minimization and decidability results are also included. Section 6 is about Chomsky's hierarchy and, especially, context-free languages. Closure under operations, the pumping lemma, normal forms and decidability questions are treated. Section 7 reviews sequential machines. Almost all theorems on universal algebra and lattices appear with complete proofs. Automata and language theoretic proofs are mostly just outlined or omitted. Readers familiar with the topics of Chapter I may skim over it. Other readers will find enough material to understand the rest of the book, or, if needed, may consult the references given at the end of the chapter.

Chapter II is devoted to finite-state tree recognizers, i.e., tree automata without output. Section 1 explains the usage of the word tree for terms. Two kinds of tree recognizers are introduced in Section 2. Frontier-to-root recognizers read trees from the leaves toward the root, and root-to-frontier recognizers work in the opposite way. Both types have deterministic and nondeterministic versions. It is shown that all these recognizers accept the same class of tree languages — the so-called recognizable forests —, except for deterministic root-to-frontier recognizers. In Scetion 3 closure properties of recognizable forests are dealt with. Sections 4 and 5 give two different characterizations of recognizable forests through regular tree grammars and regular expressions. The latter is Kleene's theorem for recognizable forests. The minimization theory of deterministic frontier-to-root recognizers is developed in Section 6. Sections 7—9 provide four additional characterizations of recognizable forests: by means of congruences of the absolutely free term algebra, as fixed-points of forest equations, in terms of local forests, and by means of certain Medvedev-type operations. In Section 10 basic properties of recognizable forests are shown to be decidable. Section 11 treats deterministic root-to-frontier recognizers, their minimization, and characterizes forests accepted by these recognizers.

Chapter III provides a study of the connection of recognizable forests to context free grammars and languages. Section 1 exploites the yield function as a way of extracting a word from a tree and a language from a forest. In Section 2 the forest made up from the derivation trees of a context free grammar is shown to be recognizable. Hence, by the yield forming process, tree recognizers become acceptors for context free languages. Section 3 demonstrates some further properties of the yield function. The chapter ends with Section 4, where tree recognizers are used as acceptors for context free languages in an alternative way.

The last chapter, Chapter IV, treats tree automata with output, the so-called tree transducers. Two basic sorts of tree transducers are introduced in Section 1: frontier-to-root and root-to-frontier tree transducers. Many special cases and deterministic versions are investigated in the first two sections. These special cases give rise to the composition and decomposition theorems of tree transformations induced by tree transducers. This is the subject of Section 3. In Section 4, root-to-frontier tree transducers are generalized to transducers with regular look-ahead. Later this concept turns out to be a very useful tool in many ways. Section 6 provides a study of properties of surface forests, i.e. the images of regular forests under tree-transformations. Section 7 contains some auxiliary results in preparation for Section 8, where it is shown that an infinite hierarchy can be obtained by serial compositions of tree transformations. In the last section the equivalence problem of deterministic tree transducers is proven to be decidable.

Chapters II—IV also contain exercises and each of them ends with a historical and bibliographical overview reviewing some additional fields too. Applications of the theory are ignored, but interested readers may find enough orientation in the bibliographical notes.

The bibliography contains more than 250 entries. The index helps guide the reader in looking up notions and notations.

This well-written new book can be recommended as an important, systematic summary of the subject, as a reference book, and even for those who are familiar with some aspects of automata and formal language theory and want to increase their knowledge in this direction.

Zoltán Ésik (Szeged)

Geometric Dynamics. Proceedings, Rio de Janeiro, 1981. Edited by J. Palis Jr. (Lecture Notes in Mathematics, 1007), IX+827 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo 1983.

These are the Proceedings of an International Symposium on Dynamical Systems that took place at the Istituto de Matematica Pura e Aplicada, Rio de Janeiro, in July—August, 1981.

One hundred years before this conference H. Poincaré published his fundamental memoir "Sur les courbes définies par les équations différentielles", which was the origin of geometric or qualitative dynamics. Since that moment a great number of mathematicians have studied the properties of the trajectories of dynamical systems. New notions have come up, very interesting and deep problems have arisen.

The conference was participated by the most outstanding scholars in the West of this theory. They delivered 43 lectures on up-to-date topics. Some of them were: structural stability, entropy, local classification of vector fields, bifurcations, infinite dimensional dynamical systems (especially, functional differential equations), existence and nonexistence of periodic orbits, Lyapunov functions, Lyapunov exponents, strange attractors, random perturbations.

The Proceedings will be very useful for every scholar interested in the qualitative theory of differential equations.

L. Hatvani (Szeged)

Geometric Techniques in Gauge Theories. Proceedings of the Fifth Scheveningen Conference on Differential Equations, The Netherlands, August 23—28, 1981. Edited by R. Martini and E. M. de Jager (Lecture Notes in Mathematics, 926), IX+219 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982.

The volume contains 10 lectures delivered at the conference on gauge theory, one of the important subjects of contemporary mathematics and physics. The first two papers give an introduction to the geometry of gauge field theory (R. Hermann: Fiber spaces, connections and Yang—Mills fields; Th. Friedrich: A geometric introduction to Yang—Mills equations). Four lectures were devoted to physical phenomena occurring in gauge field theory, the majority of which is based on global properties of the fibre bundle underlying the field equations. These lectures (F. A. Bais: Symmetry as a clue to the physics of elementary particles; Topological excitations in gauge theories; an introduction from the physical point of view; P. J. M. Bongaarts: Particles, fields and quantum theory; E. F. Corrigan: Monopole solitons) of informative character provide a common language for mathematicians and theoretical physicists. A Trautman's report — Yang—Mills theory and gravitation: A comparison — summarizes the analogies and differences between gauge theories of internal symmetries and Einstein's theory of general relativity. Two articles deal with the twistor method which is promising for solving nonlinear partial differential equations of mathematical physics (M. G. Eastwood: The twistor description of linear fields; R. S. Ward: Twistor techniques in gauge theories). Prolongation theory is the concern of the final paper (P. Molino: Simple pseudopotentials for the *KdV*-equation).

This well arranged book with single lectures very clearly written provides a comprehensive survey of classical gauge theory and can be warmly recommended for all students and research workers interested in the subject.

L. Gy. Fehér (Szeged)

Geometries and Groups, Proceedings, Berlin 1981. Edited by M. Aigner and D. Jungnickel (Lecture Notes in Mathematics, 893), X+250 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1981.

This volume contains five invited and 11 contributed papers presented at the colloquium in honour of Professor Hanfried Lenz held at the Freie Universität Berlin in May 1981. The invited survey lectures given by F. Buekenhout, J. Doyen, D. R. Hughes, U. Ott and K. Strambach are devoted to combinatorial and group theoretical aspects of geometry. The contributed papers deal with various problems of combinatorics and finite geometry.

Péter T. Nagy (Szeged)

Allan Gut—Klaus D. Schmidt, Amarts and Set Function Processes (Lecture Notes in Mathematics, 1042), 258 pages. Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983.

These lecture notes are based on a series of talks on real-valued asymptotic martingales (amarts) held at Uppsala University, Sweden. The main purpose of them is to introduce the reader to the theory of asymptotic martingales, on whose part the notes require the knowledge of classical martingale theory.

The book is divided into three parts. In the first part Allan Gut gives an introduction to amarts. This introduction contains, for example, the history and basic properties of amarts, convergence and stability theorems, and the Riesz decomposition. The much longer second part was written by Klaus D. Schmidt and it deals with amarts from a measures theoretical point of view. We list only the chapter headings here: Introduction, Real amarts, Amarts in a Banach space, Amarts in a Banach lattice, Further aspects of amart theory. The book ends with a rich bibliography. The bibliopgrahy contains papers which deal with or were inspired by amarts as well as some papers concerning further generalizations of martingales.

The book gives a good introduction to this field and the rich, up-to-date bibliography helps to find a way in the literature of amarts.

Lajos Horváth (Szeged)

A. Haraux, Nonlinear Evolution Equations—Global Behavior of Solutions (Lecture Notes in Mathematics, 841), IX+313 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982.

It is common in the modern theory of partial differential equations that the original equation is rewritten into an ordinary differential equation in an infinite-dimensional Banach space of functions as a state space. This allows the application of certain methods of topological dynamics and the theory of finite dimensional ordinary differential equations to partial differential equations. If the original equation is non-linear (e.g. the Schrödinger equation arising from non-linear optics) then the associated infinite-dimensional equation is non-linear as well. These lecture notes contain the basic material of the two semester seminar course on equations of this type given by the author at Brown University during the academic year 1979—80.

The study is centered on semi-linear, quasi-autonomous systems.

Chapter A, which is of preparatory character, deals with the uniqueness of the solutions of the Cauchy problem. Then the basic notions and facts of the theory of monotone operators are given, which is the main tool of investigation in the book.

Chapter B is concerned with the existence of periodic solutions to quasi-autonomous systems with especial regard to linear and dissipative cases.

Chapters C and D are the most original parts of the book. Concerning autonomous dissipative and quasi-autonomous dissipative periodic systems, the author gives theorems on the asymptotic behaviour of the solutions as $t\to\infty$.

The knowledge of elementary Banach space theory and the introductory chapters on Cauchy problem in nonlinear partial differential equations are prerequisites to read the book.

These lecture notes, containing several results not published previously in the literature, will be very useful and interesting for mathematicians dealing with the theory and applications of nonlinear partial differential equations.

L. Hatvani (Szeged)

Harmonic Maps, Proceedings, New Orleans 1980, edited by R. J. Knill, M. Kalka and H. C. J. Sealey (Lecture Notes in Mathematics 949), 158 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982.

This volume contains papers contributed by participants of the N.S.F.—C.B.M.S. Regional Conference on Harmonic Maps at Tulane University in December 1980. The ten lectures given by James Eells and co-authored by Luc Lemaire at the conference are published separately in CBMS regional conference reports. The book gives a good survey on various topics connected with the theory of harmonic maps: singularities, deformation and stability theory, Cauchy—Riemann equations, Yang—Mills fields, foliations, and harmonic maps between classical spaces and surfaces.

Péter T. Nagy (Szeged)

Loo-Keng Hua, Selected Papers, Edited by H. Halberstam, XIV+889 pages, Springer-Verlag New York—Heidelberg—Berlin 1983.

Having edited recently some of Hua's books (*Introduction to Number Theory, Starting with the Unit Circle, Applications of Number Theory to Numerical Analysis,* the latter one written jointly with Wang Yuan) in English, it was just very timely to publish his Selected Papers. The Selected Papers consist of three main parts reflecting Hua's oeuvre in pure mathematics and a part classified miscellaneous, his biography and list of publications, and a sketch of his contributions to applied mathematics.

The first main part is, of course, number theory. It consists of 20 papers including his results on the estimation of exponential sums, on the generalized Waring's problem, on Goldbach's problem, on the Waring—Goldbach problem, on the Gauss circle problem, and on the number of partitions of a number into odd parts.

The second main part contains 18 papers on algebra and geometry, including Hua's results on the existence of pseudo-basis in p-groups, on semi-automorphisms of skew fields, on automorphisms of classical groups, and on the geometry of matrices.

The third main part is devoted to function theory in several variables (5 papers) in connection with partial differential equations and differential geometry.

We have to emphasize Hua's "offensive style" in solving mathematical problems what looms in his computations. Some of the present selected papers are the first English translations. This volume proves that those who know Loo-Keng Hua to be "only" number theorist are wrong.

L. A. Székely (Szeged)

Serge Lang, Undergraduate Analysis (Undergraduate Texts in Mathematics), 545 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1983.

This book is a revised and enlarged version of the author's "Analysis I", Addison—Wesley Publishing Company 1968. It is a logically self-contained first course in real analysis, "which presupposes the mathematical maturity acquired by students who ordinarily have had two years of calculus" (from the Foreword). The contents is as follows: Part 1: Review of Calculus (Sets and Mappings;

Real Numbers; Limits and Continuous Functions; Differentiation; Elementary Functions; The Elementary Real Integral); Part 2: Convergence (Normed Vector Spaces; Limits; Compactness; Series; The Integral in One Variable); Part 3: Applications of the Integral (Approximation with Convolutions; Fourier Series; Improper Integrals; The Fourier Integral); Part 4: Calculus in Vector Spaces (Functions on n-Space; Derivatives in Vector Spaces; Inverse Mapping Theroem; Ordinary Differential Equations); Part 5: Multiple Integration (Multiple Integrals; Differential Forms).

This survey shows how many topics are treated, more than in usual standard texts at this level. The emphasis is on the theoretical aspects, but the basic computational techniques are also demonstrated in detail. The central and deep concepts of analysis (convergence, limit, derivative, integral) are presented in a series of different forms, in ascending order of difficulty, and generality. There are many interesting technical and theoretical examples and problems, some easy, many hard; solutions to the problems are not included.

To conclude, this book is very well written and produced. Because of its flexible structure it is suitable for several advanced calculus and real analysis courses. It is not a book for the beginner, but it can be warmly recommended to all who want to learn the foundations of modern analysis.

Arnold Janz (Berlin)

Loren C. Larson, Problem-Solving Through Problems. Problem Books in Mathematics, XI + 344 pages with 104 illustrations, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1983.

This book is a volume of Springer-Verlag's new series Problem Books in Mathematics edited by P. Halmos. The reader can expect to find in this series collections of problems that have been discovered and gathered carefully together over years; interesting subjects not yet adequately treated elsewhere etc. As prototypes "Otto Dunkel Memorial Problem Book" and Pólya and Szegő's "Problems and Theorems in Analysis" are mentioned.

In another book of Pólya, in the world-famous "How to solve it" we find a chart of questions and answering some of them we have a good chance to obtain a solution of the problem. In answering the questions one of the crucial points is the knowledge of the various problem solving techniques. In this direction Larson's book will prove invaluable as a teaching aid. Chapter headings are: Heuristics; Two important principles: Induction and pigeonhole; Arithmetic; Algebra; Summation of series; Intermediate real analysis; Inequalities; Geometry.

One of the most interesting chapters is the first one, entitled Heuristics. The author focuses on the typically useful basic ideas such as: Search for a pattern; Draw a figure; Formulate an equivalent problem; Modify the problem; Choose effective notations; Exploit symmetry; Divide into cases; Consider extreme cases; Generalize. For example, in "Divide into cases" the problems can be divided into subproblems each of which can be handled separately in a case-by-case manner. The following three problems are solved: a) Prove that an angle inscribed in a circle is equal to one-half the central angle which subtends the same arc; b) A real valued function f, defined on the rational numbers, satisfies $f(x+y)=f(x)+f(y)$ for all rational x and y. Prove that $f(x)=f(1)x$ for all rational x; c) Prove that the area of a lattice triangle is equal to $I+(1/2)B-1$, where I and B denote respectively the number of interior and boundary lattice points of the triangle. Then some problems — from different branches of mathematics — for solution are listed and references to problems proposed in other chapters where this treated method may be useful. This is the structure of the other chapters too. At the end of the book one finds the sources of the more remarkable problems.

The style of the book is attractive, methods, problems and solutions are presented in a way which brings the printed page to life. No doubt, students and teachers will enjoy and use this book.

L. Pintér (Szeged)

George E. Martin, Transformation Geometry (Undergraduate Text in Mathematics) XII+237 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1982.

The main purpose of this book is to describe the Euclidean plane geometry by the study of its transformation groups. The text starts with a short introduction (Chapter 1). In Chapter 2 the concept of transformation groups is defined. Chapters 3—4 deal with translations, halfturns and reflections, using the method of analytic geometry. In Chapter 5 it is shown that any congruence can be represented as a product of at most three reflections. Chapter 6 investigates congruence transformations which can be represented as products of two reflections; it turns out that these are the translations and the rotations. Chapters 7—8 introduce the concept of the congruences of even and of odd types and a complete classification of congruences is given. Chapter 9 gives the equations of congruence transformations. Chapters 10—12 describe the discrete congruence groups; the seven discrete groups having translations in only one direction (called "Frieze Groups") and the seventeen discrete groups having two independent translations (called "Wallpaper Groups"). The periodic tesselations can be obtained as an application. Chapter 13 is devoted to similarity transformations. Chapter 14 contains the classical theorems of elementary geometry. In Chapter 15 the affine transformations are defined, and their linear operator representations are given. Chapter 16 gives a short indication as to how the classification of congruences in three-space can be obtained. In Chapter 17 the Euler polyhedron theorem is proved, the regular polyhedrons are constructed and their symmetry groups are given.

The text requires only elementary geometric knowledge. The reader will surely enjoy the book.

László Gehér (Szeged)

Mathematical Models as a Tool for the Social Sciences, edited by B. J. West, V+120 pages, Gordon and Breach, New York—London—Paris, 1980.

This book is a collection of the talks of a seminar at the University of Rochester. The eight lectures present themselves as interesting examples of mathematical model building in economic and natural history (R. W. Fogel: Historiography and retrospective econometrics; A. Budgor and B. J. West: Natural forces and extreme events — the latter is on floods and droughts in the Nile River Valley), the psychology of learning, selection making and speculation (A. O. Dick: A mathematical model of serial memory; J. Keilson and B. J. West: A simple algorithm of contract acceptance; B. J. West: The psychology of speculation: a simple model), politics (W. Riker: A mathematical theory of political coalitions), inpopulation growth (J. H. B. Kemperman: Systems of mating — in which the problem is how stable population patterns are formed in large populations under given mating systems), and for economic income distribution (W. W. Badger: An entropyutility model for the size distribution of income).

"There is no one way, and indeed no best way, to construct a mathematical model of a natural or social system" as the editor writes in his introduction, but he believes "that any problem which may be well formulated verbally, may be well formulated mathematically". All of the above models are interesting and novel enough. If you don't believe in them, construct your own and confront it with the already existing ones. The book is a very good reading.

Sándor Csörgő and Lajos Horváth (Szeged)

Mathematical Programming. The State of the Art, Bonn 1982, edited by A. Bachem, M. Grötschel and B. Korte, VIII+655 pages with 30 figures. Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983.

This book consists of 21 state-of-the-art tutorials of the 23 having constituted the main frame of the XI. International Symposium on Mathematical Programming held at the University of Bonn in 1982. These survey papers written by leading experts can introduce everyone to the recent and most important results in several areas of mathematical programming. The book contains a brief review about the Fulkerson Prize and Dantzig Prize won in the year 1982. Since it seems to be unjust to mention some papers and to neglect other ones, no matter how long their list is, we give the author and the title of all the papers. We hope that the reader will forgive us upon seeing the list: E. L. Allgower and K. Georg, Predictor-corrector and simplicial methods for approximating fixed points and zero points of nonlinear mappings; L. J. Billera, Polyhedral theory and commutative algebra; G. B. Dantzig, Reminescences about the origins of linear programming; R. Fletcher, Penalty functions; R. L. Graham, Applications of the FKG inequality and its relatives; S.-Å. Gustafson and K. D. Kortanek, Semi-infinite programming and applications; M. Iri, Applications of matroid theory; E. L. Lawler, Recent results in the theory of machine scheduling; L. Lovász, Submodular functions and convexity; J. J. Morè, Recent developments in algorithms and software for trust region methods; M. J. D. Powell, Variable metric methods for constrained optimization; W. R. Pulleyblank, Polyhedral combinatorics; Stephen M. Robinson, Generalized equations; R. T. Rockafellar, Generalized subgradients in mathematical programming; J. Rosenmüller; Nondegeneracy problems in cooperative game theory, R. B. Schnabel, Conic methods for unconstrained minimization and tensor methods for nonlinear equations; A. Schrijver, Min-max results in combinatorial optimization; N. Z. Shor, Generalized gradient methods of nondifferentiable optimization employing space dilatation operations; S. Smale, The problem of the average speed of the simplex method; J. Stoer, Solution of large linear systems of equations by conjugate gradient type methods; R. J.-B. Wets, Stochastic programming: solution techniques and approximation schemes.

L. A. Székely (Szeged)

Measure Theory, Oberwolfach 1981, Proceedings of the Conference Held at Oberwolfach, Germany, June 21—27, 1981, edited by D. Kölzow and D. Maharam-Stone (Lecture Notes in Mathematics, 945), XV+431 pages. Springer-Verlag, Berlin—Heidelberg—New York, 1982.

These conference proceedings consist of 36 papers on several fields of measure theory such as general measure theory, descriptive set theory and measurable selections, lifting and disintegration, differentiation of measures and integrals, measure theory and functional analysis, non-scalar-valued measures, measures on linear spaces, stochastic processes and ergodic theory.

Although I must not list here all the titles of papers, I have to mention some of them. R. J. Gardner in his paper 'The Regularity of Borel Measures' gives a detailed survey on regularity assumptions of Borel measures with 15 pages of references. H.-U. Hess 'A Kuratowski Approach to Wiener Measure' exhibits a procedure that may be considered an alternative way of constructing Wiener measure. J. R. Choksi and V. S. Prasadin 'Ergodic Theory on Homogeneous Measure Algebras' continues previous efforts to generalize ergodic theory.

The book contains open research problems discussed in the problem session of the conference.

L. A. Székely (Szeged)

G. H. Moore, Zermelo's Axiom of Choice: Its Origins, Development and Influence (Studies in the History of Mathematics and Physical Sciences 8), XIV+410 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1982.

This book of four chapters is the first full-lenght history of the Axiom of Choice. David Hilbert wrote in 1926 that Zermelo's Axiom of Choice was the axiom "most attacked up to the present in the mathematical literature...". Later Abraham Fraenkel added to this that "the axiom of choice is probably the most interesting and, in spite of its late appearance, the most discussed axiom of mathematics, second only to Euklid's axiom of parallels which was introduced more than two thousand years ago".

In Chapter 1, The Prehistory of the Axiom of Choice, the author indicates four major stages through which the use of arbitrary choices passed on the way to Zermelo's explicit formulation of the Axiom of Choice. The first stage — choosing an unspecified element from a single set or arbitrary choice of an element from each of finitely many sets — can be found in Euklid's Elements (if not earlier). The second stage was when Gauss and others made infinite number of choices by stating a rule. In the third stage mathematicians made infinite number of choices but left the rule unstated. This was the case, e.g., when Cauchy demonstrated a version of the Intermediate Value Theorem in 1821. The fourth stage, where mathematicians made infinitely many arbitrary choices for which, consequently, the Axiom of Choice was essential, began in 1871 by a paper of Heine on real analysis. Heine's proof, borrowed from Cantor, implicitly used the Axiom to show that his definition of continuity implies the earlier one introduced by Cauchy and Weierstrass.

The boundary between finite and infinite, the various definitions of finiteness (by Bolzano, Dedekind and Pierce) and the connections among them are also discussed in this chapter, as well as several implicit uses of the Axiom by Cantor.

At the end of this chapter two equivalent statements to the Axiom, the Well-Ordering Principle and the Trichotomy of Cardinals are mentioned which were stated by Cantor before Zermelo formulated the Axiom of Choice.

Chapter 2, Zermelo and His Critics (1904—1908) is an exploration of the debate started when in 1904 Zermelo published his proof that every set can be well-ordered. The major questions were: "What methods were permissible in mathematics? Must such methods be constructive? If so, what constituted a construction? What did it mean to say that a mathematical object existed?" From 1905 to 1908 eminent mathematicians in England, France, Germany, Holland, Hungary, Italy, and the United States debated the validity of his demostration. Never in modern times have mathematicians argued so publicly and so vehemently over a proof.

In Chapter 3 we can read Zermelo's reply to his critics and his axiomatization of set theory and the counteropinions of Poincaré and Russel among others. Some equivalent statements to the Axiom of Choice are also discussed,

Chapter 4, The Warsaw School, Widening Applications, Models of Set Theory (1918—1940) deals with the wide-spread applications and the modern independence results.

There are an Epilogue: After Gödel, and two appendices. The first one consists of five letters on set theory (written by Baire, Borel and Hadamard), and the second is "Deductive Relations Concerning the Axiom of Choice".

While the author brings out aspects of a history that will fascinate mathematical researchers and philosophers, this book is warmly recommended to everybody interested in set theory, in the philosophy of mathematics and in historical questions.

Lajos Klukovits (Szeged)

M. A. Naimark—A. I. Stern, Theory of Group Representations (Grundlehren der Mathematischen Wissenschaften, 246), IX+568 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1982.

This book is the second translation of the original edition of the book of M. A. Naimark, written in Russian, and in which M. A. Naimark describes his collaboration with A. I. Stern. The first translation was into French including a faithful transcription of misprints. The French translation already lists A. I. Stern as co-author. The book consists of 12 chapters. The text starts with a short algebraic foundation of representation theory. The next chapter summarizes the most important general results of the theory of representations of finite groups giving the representations of the symmetric group and of the group SL. Two chapters deal with topological groups, providing the general definition of a representation of a topological group, and especially the representation theory of compact groups in connection with the representations of the corresponding group algebra. In this part there are some mistakes that do not disturb the intelligibility of the text. Further chapters deal with the applications of the general theory of representations of compact groups. Two chapters investigate finite representations of the full linear group and of complex classical groups. The next one is devoted to covering spaces and simply connected groups. The last five chapters contain a detailed investigation of Lie groups and Lie algebras.

The reader is supposed to be familiar with linear algebra, elementary functional analysis and with the theory of analytic functions.

László Gehér (Szeged)

A. W. Naylor—G. R. Sell, Linear Operator Theory in Engineering and Science (Applied Mathematical Sciences, vol. 40), XV+624 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1982.

As a lecturer, sometimes I have had to teach parts of mathematical analysis to scientists or students in chemistry, biology, medicine. So I know that this task can be more difficult than to lecture the same subject to mathematicians. How difficult could it be then to write a book for engineers and scientists on functional analysis, which is one of the most abstract fields of mathematical analysis? Thus having read the exciting title of this book I was very curious about answers to some questions: How to introduce the concepts of linear operator theory to the readers not bringing enough experiences from the classical chapters of mathematical analysis that make the definitions natural and understandable? Which concepts, results and methods and how deeply are they to be included into a mathematically rigorous book if it is known that the readers are interested mostly in the applications of functional analysis to their own sciences?

Fortunately, the authors resolve these conflicts excellently and find the balance between the different points of view. In order to illuminate the abstract concepts they give lots of examples and exercises. As far as it is possible they use the geometry and finite-dimensional analogies for the heuristic preparation of the subject-matter. For example, Chapter 6, concerned with the spectral analysis of linear operators, is divided into three parts. The first one is the geometric analysis of linear combinations of orthogonal projections giving a resolution of the identity in a Hilbert space. In the second part the spectrum of general bounded and unbounded linear operators is introduced and illuminated by examples. The chapter is concluded with the spectral theorem for compact normal operators in a Hilbert space and its applications (matched filter, the Karhunen—Loève expansion for discrete random processes, ε-capacity of a linear channel). It has been a very good decision to deal with the spectral theory of compact operators separately because it is relatively simple but demonstrates the distinction between the finite- and infinite-dimensional cases, which is the big jump in spectral theory.

We recommend this excellent text-book to every engineer, scientist and applied mathematician making the first steps in functional analysis.

L. Hatvani (Szeged)

Donald J. Newman, A Problem Seminar. Problem Books in Mathematics, VIII+113 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1982.

This book contains some problems of D. J. Newman's problem seminar. The author says in the Preface: "There was once a bumper sticker that read "Remember the good old days when air was clean and sex was dirty?" Indeed, some of us are old enough to remember not only those good old days, but even the days when Math was fun (!), not the ponderous THEOREM, PROOF, THEOREM, PROOF, ..., but the whimsical, "I've got a good problem"."

This last sentence shows precisely what the reader can find on every page of this excellent book. The problems are interesting, natural, in general one cannot get away from them without having the solutions. This is not only the reviewer's personal impression but this was his experience after posing some problems of the text to his students.

The book consists of three parts: Problems, Hints and Solutions. Sometimes the solutions are not fully worked out, but the interested reader can fill the gaps. A great part of problems seems to be quite elementary, but in some cases the solution requires not only elementary notions. Therefore, the text forces the reader to do some more mathematics, to get acquainted with new notions. For illustration I tried to select a problem but I have so many favourites that I could not choose among them.

This problem seminar is warmly recommended to teachers, students and everyone who enjoy the fun and games of problem solving and have the opinion that asking and answering problems is what keeps a mathematician young in spirit.

L. Pintér (Szeged)

Ordinary Differential Equations and Operators, Proceedings, Dundee, 1982. Edited by W. N. Everitt and R. T. Lewis (Lecture Notes in Mathematics, 1032), XV+521 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983.

These are the Proceedings of the Symposium on Ordinary Differential Equations and Operators held in the Department of Mathematics at the University of Dundee, Scotland during the months of March, April, May, June and July 1982. They are dedicated to F. V. Atkinson by his many friends and colleagues in recognition of his mathematical contributions to the theory of differential equations.

The topics of the volume can be arranged in groups according to the many themes having been studied by F. V. Atkinson: boundary value problems, differential operators (Sturm—Liouville problems, spectral theory), second order oscillation theory, limit cycles, etc.

Some of the papers are surveys giving also the history of their topics, but the reader can find also articles including results not published before.

L. Hatvani (Szeged)

Ordinary and Partial Differential Equations. Proceedings, Dundee, Scotland 1980. Edited by W. N. Everitt and B. D. Sleeman (Lecture Notes in Mathematics, 846), XIV+384 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1981.

This volume contains lectures delivered at the sixth Conference on Ordinary and Partial Differential Equations held at the University of Dundee. As the name of the conference shows, the topics of lectures are taken from various branches of the theory of differential equations. To illustrate this assertion here are the titles of some lectures: Some unitarily equivalent differential operators with finite and infinite singularities; Nonlinear two-point boundary value problems; On the spectra of Schrödinger operators with a complex potential; Asymptotic distribution of eigenvalues of elliptic operators on unbounded domains; Some spectral gap results; Some topics in nonlinear wave propagation; Oscillation properties of weakly nonlinear differential equations; Norm inequalities for derivatives; Fixed point theorems; A bound for solutions of a fourth order dynamical system;

Convergence of solutions of infinite delay differential equations with an underlying space of continuous functions; Symmetry and bifurcation from multiple eigenvalues; Variational methods and almost solvability of semilinear equations.

The book is warmly recommended to everybody who works in differential equations and perhaps it will stimulate other readers to make research in this field.

L. Pintér (Szeged)

Ordinary and Partial Differential Equations, Proceedings, Dundee, Scotland, 1982. Edited by W. N. Everitt and B. D. Sleeman (Lecture Notes in Mathematics, 964), XVIII+726 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982.

These Proceedings include the lectures delivered at the seventh Conference on Ordinary and Partial Differential Equations which was held at the University of Dundee, Scotland, March 29—April 2, 1982.

Unfortunately, there is no room in this review to present the complete list of the 60 lectures, which shows a very wide spectrum. Some of the key words and phrases: boundary value problems, eigenvalue problems, eigenfunction expansions, oscillations, bifurcations, differential equations with delay, integrodifferential equations, stochastic functional differential equations, scattering theory, generalized Schrödinger operators, partial differential equations of infinite order, control theory, astronomy, thermodynamics.

Like the Proceedings of the earlier Dundee Conferences, this volume, which is dedicated to the University of Dundee on the occasion of its centenary celebrations, gives a good flavour of the actual problems of the theory of differential equations.

L. Hatvani (Szeged)

Radical Banach Algebras and Automatic Continuity (Proceedings, Long Beach 1981), Edited by J. M. Bachar, W. G. Bade, P. C. Curtis Jr., H. G. Dales and M. P. Thomas (Lecture Notes in Mathematics, 975), VII+470 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1983.

This collection contains 30 papers, the contributions to the conference indicated in the title, held at the California State University between July 13—17, 1981.

The editors write: "The basic problem of automatic continuity theory is to give algebraic conditions which ensure that a linear operator between, say, two Banach spaces is necessarily continuous. This problem is of particular interest in the case of a homomorphism between two Banach algebras. Other automatic continuity questions arise in the study of derivations from Banach algebras to suitable modules and in the study of translation invariant functionals on function spaces. There is a fundamental connection between questions of automatic continuity and the structure of radical algebras. ... The purpose of the conference was to present recent developments in these two areas and to explore the connections between them."

The volume is divided into five sections. Section I deals with the general theory of commutative radical Banach algebras and contains (together with a paper of F. Zouakia) two lengthy papers by J. Esterle. The first one gives a classification of these algebras, while the second one is devoted to the question of whether or not such algebras must contain non-trivial closed ideals. This latter problem is related to the invariant subspace problem for Banach spaces.

Papers in Section II (by H. G. Dales, Y. Domar, W. G. Bade, K. B. Laursen, M. P. Thomas, S. Grabiner, G. R. Allan, G. A. Willis, N. Gronbaek and G. F. Bachelis) are concerned with radical convolution algebras on $\mathbf{R}^+$ and $\mathbf{Z}^+$. The central problem here is to determine for which radical weights, ω, every closed ideal of $L^1(\omega)$ is a standard ideal, that is, an ideal consisting of those functions with support in an interval $[\alpha, \infty)$.

Section III contains papers by B. Aupetit, R. J. Loy, P. C. Curtis Jr., J. C. Tripp, P. G. Dixon, E. Albrecht, M. Neumann, H. G. Dales and G. A. Willis, and is devoted to the automatic continuity of homomorphisms (between semisimple, nonsemisimple, local and C^* algebras) and derivations.

The automatic continuity of (mostly translation invariant) linear functionals on Banach algebras is discussed in Section IV, which includes papers by G. H. Meisters, R. J. Loy and H. G. Dales.

Finally Section V contains a list of open problems, some well known and others posed at the conference.

L. Kérchy (Szeged)

D. M. Sandford, Using Sophisticated Methods in Resolution Theorem Proving (Lecture Notes in Computer Science, 90), VI+239 pages. Springer-Verlag, Berlin—Heidelberg—New York, 1980.

The motto of the volume "There are no solved problems; there are only problems that are more or less solved" indicates quite well the author's intention when choosing an area of research, the development of which — after a promising decade — has come to a sudden standstill. The author is right; the book convinces the reader that there remains a large room for further thinking on open problems in the theory of theorem proving, whose solutions can point ahead.

The main topic of the volume is a certain refinement of the familiar resolution principle, called Hereditary Lock Resolution (HLR, for short). HLR is an amalgamation of a modification of Boyer's Lock Resolution rule and an extension of the Model Strategy due to Luckham. The basic properties of HLR are presented in Chapter 2. Chapter 3 is devoted to completeness problems; in fact it is proved that HLR are a sound and complete inference rule. The last chapter deals with a general theory of model specification techniques. The results obtained are employed to show the flexibility and sophistication of models in pragmatic environments.

The book is not self-contained. Actually, its complete understanding requires a considerable amount of brackground knowledge in the "classical" theory of theorem proving. Accordingly, this volume can be useful for experts and graduate students.

P. Ecsedi-Tóth (Szeged)

Ryuzo Sato-Takayuki Nǒno, Invariance Principles and the Structure of Technology (Lecture Notes in Economics and Mathematical Systems 212), 94 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983.

This book is devoted to the study of the mathematical models of production theory in the period of technical progress. The production process can be described by an input-output function and the technical change can be considered as a 1-parameter transformation group acting on the manifold of input variables. Thus it is very natural to use the methods of Lie transformation groups in this theory.

The main results of this monograph are connected with invariance principles of production processes. The possible input-output functions are classified and the classical production functions are characterized by means of invariance properties.

Péter T. Nagy (Szeged)

J.-P. Serre, Linear Representations of Finite Groups (Graduate Texts in Mathematics, 42) Springer-Verlag, New York—Heidelberg—Berlin 1977, X+170 pages.

This book consists of three parts. One of them deals with the general theory and two are devoted to special questions of representation theory. The first part introduces the basic concepts of representation of finite groups, and describes the correspondence between representations and characters.

The proofs are elegant and as elementary as possible. A short indication shows how the preceding results carry over to compact groups. The general theory is applied for some known classical groups. The second part investigates degrees of representations and integrality properties of characters, induced representations, theorems of Artin and Brauer and their applications, rationality questions. The third part contains an introduction to the Brauer theory using the language of abelian categories. Several applications to the Artin representations are given. At the end of the text a short Appendix can be found on the definition of Artinian rings, the Grothendieck group, projective modules and discrete valuations.

László Gehér (Szeged)

J. Sesiano, Books IV to VII of Diophantus' Arithmetica in the Arabic translation attributed to Qustā ibn Lūqā, (Sources in the History of Mathematical and Physical Sciences 3), XII+502 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1982.

According to our present knowledge the Greek mathematician Diophantus of Alexandria (lived probably between 150 B. C. and A. D. 350, but it seems fairly probable that he flourished about A. D. 250) wrote at least two treatises: one of them dealing with problems in indeterminate equations and systems of equations, the Arithmetica, and another, a smaller tract, on polygonal numbers. Both are only partially extant today. We can read from the introduction of the Arithmetica that it originally consisted of thirteen Books. But only six of these have survived until now in Greek, and they have been edited and translated several times. The remaining seven were considered irretrievably lost until 1973, when Gerald Toomer learned of existence of a manuscript in A. Gulchin-i Ma'ānī's just-published catalogue of the mathematical manuscripts in the Mashhad Shrine Library. This manuscript, a codex, consists of four other, hiterto unknown Books in an Arabic translation which, since it is attributed to Qustā ibn Lūqā, must have been made around or after the middle of the ninth century.

This book, which has five parts, is based on the author's 1975 Ph. D. thesis at Brown University. Major changes, however, are found in the mathematical commentaries. The discussion of Greek and Arabic interpolations is entirely new, as is the reconstruction of the history of the Arithmetica from Diophantine to Arabic times.

In Part One the first chapter deals with historical questions: the authenticity of the Arabic Books, the placement of the Arabic Books among the presently known Books of the Arithmetica, Diophantus in islamic and Bysantine times. This analysis leads to the conclusion that the four Arabic Books are the IV—VII books of the Arithmetica. Three Greek Books precede the Arabic four and the other Greek Books follow them.

In Part Two we can find the English translation of the Arabic Books. Part Three, the largest one, contains the author's detailed mathematical commentaries on the material. Part Four is the complete Arabic text of the manuscript. Part Five is an extensive Arabic index.

There is an Appendix under the title Conspectus of the Problems of the Arithmetica.

"Readers — mathematicians and non-mathematicians alike — will gain new perspectives on the techniques of Greek algebra and will learn of the fate and modifications of a scientific classic in the time between its classical origin and its medieval Arabic translation."

Lajos Klukovits (Szeged)

D. J. Shoesmith—T. J. Smiley, Multiple-Conclusion Logic, pp. IX+396. Cambridge University Press, Cambridge—London—New York—Melbourne, 1978.

The volume is a systematic study of multiple-conclusion proofs which can have, as opposed to traditional proof theory, more than one conclusions, say $B_1, \ldots, B_n$. These are to be understood as the "field within which the truth must lie", provided, of course, that the premisses $A_1, \ldots, A_m$ are accepted. The subject goes back to the works of G. Gentzen, R. Carnap and W. Kneale.

The book is divided into four parts. In Part I the familiar logical notions are generalized for multiple-conclusion proof rules and the connections between conventional and multiple-conclusion logics are investigated. In particular, adequateness (completeness) of several multiple-conclusion proof rules is proved. Part II treats graph proofs. This concept has been introduced to give an explicit tool for describing interdependencies among the components of an argument; quoting the authors: "it is not enough that each of its component steps is valid in isolation: they must also relate to one another properly". Graph proofs enable one to visualize arguments (independently from any particular axiom system) and hence to investigate the connection between the "form of arguments" (i.e. their graphs) and the semantical notion of validity.

In the rest of the volume the authors apply the techniques developed in the first two parts. In Part III, a thorough study of a particular many-valued multiple-conclusion inference system can be found. It is proved, for example, that every finite-valued multiple-conclusion propositional calculus is finitely axiomatizable. The last part of the book is devoted to investigate how "natural deduction" can be replaced by direct multiple-conclusion proofs. In particular, cut-elimination-like theorems are proved for classical predicate and for intuitionistic propositional calculi.

The book is clearly written and easily comprehensible. It can be useful for proof theorists on expert and graduate levels.

P. Ecsedi-Tóth (Szeged)

Ya. G. Sinai, Theory of Phase Transitions: Rigorous Results. VIII+150 pages, Akadémiai Kiadó, Budapest and Pergamon Press, Oxford, 1982.

The concept of limit Gibbs distributions (LGD) is relatively new, it was introduced in 1968 by Dobrushin, Lanford and Ruelle. Their construction made possible the rigorous development of the theory of phase transition in a probabilistic language. However, the special mathematical structures related to statistical physics involve highly non-standard methods.

Sinai's outstanding book gives a systematic survey of the results obtained using the concept of LGD. A great deal of these results is due to the author himself and his school (Chapters II and IV).

The book is well constructed, each chapter is almost selfcontained. The presentation is clear, the author always finds the appropriate level of generality. Both mathematicians and physicists — if they are inclined to deal with statistical physics directly and seriously — can grasp the major problems of the theory of phase transitions and the necessary information to try to solve them.

Chapter I has an introductory character, the author defines the notion of LGD and elucidates it by the most important examples related to lattice systems (e.g. Ising model, Heisenberg's continuous spin model, Yang—Mills model). The existence of the LGD is proved for general lattice systems and for the lattice model of quantum field theory.

In Chapter II the existence of phase diagram for small $(r-1)$-parameter perturbations of a periodic Hamiltonian having r ground states is proved. The result is due to Sinai and Pirogov; the proof is based on a far-reaching generalization of the contour method proposed by Peierls for proving the existence of long range order in the Ising-model at low temperature.

In Chapter III continuous spin systems are considered. By the Dobrushin—Shlosman theorem there is no continuous symmetry breakdown in the two-dimensional Heisenberg model. On the other hand, in models of three or more dimensions at low temperature, as Fröhlich, Simon and Spencer have proved, a spontaneous breakdown of continuous symmetry is present.

Chapter IV is devoted to the exact mathematical foundation of the renormalization group method — due to Bleher and Sinai — in the theory of second-order phase transitions. Dyson's hierarchical model is studied in detail; this model is an instructive example, where all interesting phenomena arise. The most intriguing problem is to find non-Gaussian invariant distributions under the action of the renormalization group. A special kind of bifurcation theory is developed for solving the above problem.

The subject of this book is presented "in statu nascendi"; the deep mathematical tools treated by the author were further developed — a great deal even by the Moscow school of mathematical physics — since the book has been written.

András Krámli (Budapest)

Statistics and Probability, Proceedings of the 3rd Pannonian Symposium on Mathematical Statistics, Visegrád, Hungary, 13—18 September, 1982, edited by J. Mogyoródi, I. Vincze and W. Wertz, X+415 pages, Akadémiai Kiadó, Budapest and D. Reidel Publishing Company, Dordrecht—Boston—Lancaster, 1984.

The thirty-six papers included in this volume move on a very wide scale. This, of course, is no surprise if the major organizing principle of a conference is geographical. The authors are: G. Baróti, M. Bolla—G. Tusnády, E. Csáki, S. Csörgő—H. D. Keller, P. Deheuvels, I. Fazekas, L. Gerencsér, T. Gerstenkorn—T. Jarzebska, B. Gyires, L. Horváth, J. Hurt, P. Kosik—K. Sarkadi, A. Kováts, A. Krámli—D. Szász, M. Krutina, L. Lakatos, A. Lesanovsky, E. Lukacs, P. Lukács, Gy. Michaletzky, J. Mogyoródi, T. F. Mári, H. Neudecker—T. Wansbeek, H. Niederreiter, J. Pintér, W. Polasek, L. Rutkowski, F. Schipp, A. Somogyi, C. Stepniak, G. J. Székely, A. Vetier, I. Vincze, A. Wakolbinger—G. Eder, M. T. Weselowska—Janczarek and A. Zempléni. A subject index helps orientation.

Sándor Csörgő (Szeged)

Studies in Pure Mathematics. To the Memory of Paul Turán. Edited by P. Erdős, L. Alpár, G. Halász and A. Sárközy, 773 pages, Akadémiai Kiadó, Budapest and Birkhäusler Verlag, Basel—Boston—Stuttgart, 1983.

The volume, dedicated to the memory of Paul Turán includes 66 papers of 88 invited authors from 16 countries of the world. The subjects of the papers are in most cases near to Turán's researches, in many cases problems of Turán are solved or the works were initiated by his earlier results. Nearly half of the papers deal with number theory what was his favourite topic during his very successful mathematical activity.

The wide scope of topics which found place in this volume — number theory, theory of functions of a complex variable, approximation theory, Fourier series, differential equations, combinatorics, statistical group theory — reflects Turán's universality and his large influence in mathematics. His pioneering contribution to many branches of mathematics can never be forgotten. This volume gives also an impression of his endeavour of searching for new paths, since various flourishing fields represented here, as, e.g., his main achievement, the power sum method (to which topic he devoted two books already, the third appears in 1984 at J. Wiley Interscience Tracts Series under the title "On a new method in the analysis and its application"), furthermore extremal graph theory, probabilistic number theory, statistical group theory owe their birth or/and their main developments to ideas of Turán. The high level of the works has been ensured by the authors whose list is the following: H. L. Abbot, M. Ajtai, L. Alpár, J. M. Anderson, R. Askey, C. Belna, B. Bollobás, W. G. Brown, L. Carleson, F. R. K. Chung, J. Clunie, Á. Császár, J. Dénes, E. Dobrowolski, Á. Elbert, P. D. T. A. Elliott, P. Erdős, W. H. J. Fuchs, D. Gaier, T. Ganelius, R. L. Graham, K. Győry, G. Halász, F. Harary, B. Harris, I. Havas, W. K. Hayman, E. Heppner, E. Hlawka, L. Iliev, K.-H. Indlekofer, Mourad E.—H. Ismail, H. Jager, M. Jutila, J.—P. Kahane, I. Kátai, Y. Katznelson, K. H. Kim, B. Kjellberg, G. Kolesnik, J. Komlós, W. Lawton, L. Lorch, G. G. Lorentz, L. Lovász, A. Meir, Z. Miller, H. L. Montgomery, Y. Motohashi, W. Narkiewicz, D. J. Newman, H. Niederreiter, P. P. Pálfy, Z. Z. Papp, R. Pierre, J. Pintz, G. Piranian, Ch. Pommerenke, N. Purzitsky, Q. I. Rahman, F. W. Roush, I. Z. Ruzsa, H. Sachs, A. Sárközy, A. Schinzel, W. M. Schmidt, I. J. Schoenberg, W. Schwarz, S. M. Shah, A. B. Shidlovsky, H. Siebert, M. Simonovits, G. Somorjai, V. T. Sós, J. Spencer, C. L. Stewart, M. Stiebitz, E. G.. Straus, J. Surányi, H. P. F. Swinnerton-Dyer, M. Szalay, E. Szemerédi, P. Szüsz, R. Tijdeman, R. C. Vaughan, P. Vértesi, M. Waldschmidt, K. Wiertelak.

J. Pintz (Budapest)

The Mathematics and Physics of Disordered Media: Percolation, Random Walk, Modeling, and Simulation, Proceedings of a Workshop held at the IMA, University of Minnesota, Minneapolis, February 13—19, 1983, edited by B. D. Hughes and B. W. Ninham (Lecture Notes in Mathematics 1035), VIII+431 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983.

It is most appropriate to cite a few sentences from the charter of the Workshop: "One of the fundamental questions of the 1980's facing both mathematicians and scientists is the mathematical characterisation of disorder.... The last decade has seen the beginnings of a unity of methods and approaches in statistical mechanics, transport in amorphous and disordered materials, properties of heterogeneous polymers and composite materials, turbulent flow, phase nucleation, and interfacial science. All have an underlying structure characterised in some sense by chaos, self-avoiding irregular walks, percolation, and fractals. Some real progress has been made in understanding random walks and percolation processes on the one hand, and through mean field or effective medium approximation and simulation of liquids and porous media on the other. The subject is directly connected with the statistics of extreme events and important pragmatic areas like fracture of solids, comminution of particulate materials, and flow through porous media."

In this extremely carefully compiled workshop volume very well-known theoreticians and applied scientists present their views of the foundations of disordered media. Following a long introductory paper in two parts (by B. D. Hughes on random discrete models and by P. Prager on diffusions in disordered media), two papers emphasize the important role of stable distributions in various physical phenomena, nine papers discuss various aspects (theoretical and applied) of percolation theory, and the five further papers deal with probabilistic models of fluids, permeability, diffusion, waves and crack growth.

Among various other kind of specialists, this volume is certainly a must for the applied probabilist.

Sándor Csörgő (Szeged)

Twistor Geometry and Non-Linear Systems (Proceedings, Primorsko, 1980), edited by H. D' Doebner and T. D. Palev, (Lecture Notes in Mathematics, 970), V+216 pages, Springer-Verlag Berlin—Heidelberg—New York, 1982.

This book contains the review lectures given at the 4th Bulgarian Summer School on Mathematical Problems of Quantum Field Theory held in Primorsko, Bulgaria, in Septermber 1980. The list of the papers is as follows.

I. Twistor Geometry: 1. S. G. Gindikin; Integral geometry and twistors. — This is a new approach to twistor geometry using the methods of Gelfand' integral geometry. 2. Yu. I. Manin; Gauge fields and cohomology of analytic sheaves. — This gives a deep analysis of holomorphic Yang—Mills fields, the vacuum Yang—Mills equations and the full system of Yang—Mills—Dirac equations in the language of holomorphic vector bundles over analytic spaces. 3. Z. Perjés; Introduction to twistor particle theory. 4. N. J. Hitchin; Complex manifolds and Einstein's equations. — This is a generalization of Penrose's twistor theory based on the geometry of rational curves in complex manifolds.

II. Non-Linear Systems: 1. A. A. Kirillov; Infinite dimensional Lie groups: their orbits, invariants and representations. The geometry of moments. 2. A. S. Schwartz; A few remarks on the construction of solutions of non-linear equations. 3. A. K. Pogrebkov—M. C. Polivanov; Some topics in the theory of singular solutions of non-linear equations. 4. V. K. Melnikov; Symmetries and conservation laws of dynamical systems. — The infinite dimensional symmetry group and several infinite series of conservation laws are found for a nonlinear evolution equation. 5. M. A. Semonov—Tianshansky; Group-theoretical aspects of completely integrable systems. — This paper treats several applications of the so-called orbit method in representation theory. 6. A. V. Mikhailov; Relativistically invariant models of the field theory integrable by the inverse scattering method. 7. P. A. Nikolov—I. T. Todorov; Space-time versus phase space approach to relativistic particle dynamics.

The book gives a good account of the present stage of the subject. We recommend it to everybody working in related fields of mathematics or mathematical physics.

Péter T. Nagy (Szeged)

Frank W. Warner, Foundations of Differentiable Manifolds and Lie Groups, (Graduate Texts in Mathematics; 94) VI+271 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1983.

This Springer edition is a reproduction of the book originally published by Scott, Foresman and Co. in 1971. It is a very clear, detailed and carefully developed graduate-level textbook of analysis on manifolds. The reader must be familiar with the material by a good undergraduate course in algebra and analysis, some knowledge of point set topology, covering spaces and fundamental groups is also assumed. Chapters 1, 2 and 4 treat the fundamental methods of calculus on manifolds. These include differentiable manifolds, tangent vectors, submanifolds, implicit function theorems, vector fields, distributions and the Frobenius theorem, differential forms, integration, Stokes' theorem and the de Rham cohomology. Chapter 3 is devoted to the foundations of Lie group theory, including the relationship between Lie groups and Lie algebras, adjoint representation, properties of classical groups, the closed subgroup theorem and homogeneous spaces. The subject of Chapter 5 is the proof of a strong form of de Rham theorem. An axiomatic treatment of sheaf cohomology theory is given. The canonical isomorphism of all classical cohomology theories on manifolds is proved. In Chapter 6 the Hodge theorem and a complete description of the local theory of elliptic operators is presented, using Fourier series as the basic tool.

A lot of exercises are included, which constitute an integral part of the text. Some of them are routine, but in some cases they contain major theorems. Hints are provided for difficult exercises.

The book may be recommended to students and research workers interested in manifold theory.

Péter T. Nagy (Szeged)

Livres reçus par la rédaction

V. I. Arnold, Catastrophe theory, IX+79 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984. — DM 16,80.

Asymptotic Analysis II. Surveys and new trends. Edited by F. Verhults (Lecture Notes in Mathematics, Vol. 985), VI+497 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1983. — DM 62,—.

C. Berg—J. P. R. Christensen—P. Ressel, Harmonic analysis on semigroups. Theory of positive definite and related functions (Graduate Texts in Mathematics, Vol. 100), X+289 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984. — DM 118,—.

Combinatorial Mathematics X. Proceedings of the Conference held in Adelaide, Australia, August 23—27, 1982. Edited by L. R. A. Casse (Lecture Notes in Mathematics, Vol. 1036), XI+419 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983. — DM 49,—.

C. W. Curtis, Linear algebra. An introductory approach (Undergraduate Texts in Mathematics), X+337 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984. — DM 72,—.

J. Diestel, Sequences and series in Banach spaces (Graduate Texts in Mathematics, Vol. 92), XIII+261 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984. — DM 108,—.

B. A. Dubrovin—A. T. Fomenko—S. P. Novikov, Modern geometry — Methods and applications. Part 1: The geometry of surfaces, transformation groups, and fields (Graduate Texts in Mathematics, Vol. 93), Translated from the Russian by R. G. Burns, XV+464 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984. — DM 128,—.

H. D. Ebbinghaus—J. Flum—W. Thomas, Mathematical logic (Undergraduate Texts in Mathematics), IX+216 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1984. — DM 84,—.

H. M. Edwards, Galois theory (Graduate Texts in Mathematics, Vol. 101), XIII+152 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984. — DM 68,—.

Equadiff 82. Proceedings of the International Conference held in Würzburg, FRG, August 23—28, 1982. Edited by H. W. Knobloch and K. Schmitt (Lecture Notes in Mathematics, Vol. 1017), XXIII+666 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983. — DM 68,—.

P. Erdős—A. Hajnal—A. Máté—R. Radó, Combinatorial set theory: Partition relations for cardinals (Disquisitiones Mathematicae Hungaricae, 13), 347 pages, Akadémiai Kiadó, Budapest, 1984.

G. B. Folland, Lecture on partial differential equations (Tata Institute Lectures on Mathematics), VI+160 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1983. — DM 16,—.

F. Gécseg—M. Steinby, Tree automata, 235 pages, Akadémiai Kiadó, Budapest, 1984.

Geometric Dynamic. Proceedings of the International Symposium held at the Instituto de Matématica Pura e Aplicada, Rio de Janeiro, Brasil, July—August, 1981. Edited by J. Palis (Lecture Notes in Mathematics, Vol. 1007), IX+827 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983. — DM 84,—.

H. Gericke, Mathematik in Antike und Orient, XII+292 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984. — DM 98,—.

V. Guillemin—S. Sternberg, Symplectic techniques in physics, XI+468 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1984. — £ 32.50.

A. Gut—K. D. Schmidt, Amarts and set function processes (Lecture Notes in Mathematics, Vol. 1042), III+258 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983. — DM 33,50.

P. van den Heuvel, The stability of a macroeconomic system with quantity constraints (Lecture Notes in Economics and Mathematical Systems, Vol. 211), VII+169 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1983. — DM 30,—.

L. K. Hua, Selected papers. Edited by H. Halberstam, XIV+889 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1983. — DM 128,—.

M. H. Karwan—V. Lotfi—J. Telgen—S. Zionts, Redundancy in mathematical programming. A state-of-the-art survey (Lecture Notes in Economics and Mathematical Systems, Vol. 206), VII+286 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983. — DM 44,—.

L. C. Larson, Problem-solving through problems (Problem Books in Mathematics), XI+332 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983. — DM 88,—.

D. V. Lindley—W. F. Scott, New Cambridge elementary statistical tables, 80 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1984. — £ 2.00.

D. H. Luecking—L. A. Rubel, Complex analysis. A functional approach (Universitext), VII+176 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984. — DM 45,—.

The Mathematics and Physics of Disordered Media: Percolation, Random Walk, Modeling and Simulation. Proceedings of a Workshop held at the IMA, University of Minnesota, Minneapolis, February 13—19, 1983. Edited by B. D. Hughes and B. W. Ninham (Lecture Notes in Mathematics, Vol. 1035), VII+431 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983. — DM 49,—.

Measure Theory, Oberwolfach 1981. Proceedings of the Conference held at Oberwolfach, Germany, June 21—27, 1981. Edited by D. Kölzow and D. Maharam-Stone (Lecture Notes in Mathematics, Vol. 945), XV+431 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 49,—.

J. D. Murray, Asymptotic analysis (Applied Mathematical Sciences, Vol. 48), VII+164 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1984. — DM 65,—.

Y. Okuyama, Absolute summability of Fourier series and orthogonal series (Lecture Notes in Mathematics, Vol. 1067), VI+118 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984. — DM 21,50.

Ordinary and Partial Differential Equations. Proceedings of the Seventh Conference held at Dundee, Scotland, March 29—April 2, 1982. Edited by W. N. Everitt and B. D. Sleeman (Lecture Notes in Mathematics, Vol. 964), XVIII+726 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 76,—.

Ordinary Differential Equations and Operators. A Tribute to F. V. Atkinson. Proceedings of a Symposium held at Dundee, Scotland, March—July 1982. Edited by W. N. Everitt and R. T. Lewis (Lecture Notes in Mathematics, Vol. 1032), XV+521 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983. — DM 62,—.

L. C. Piccinini—G. Stampacchia—G. Vidossich, Ordinary differential equations in $\mathbf{R}^n$. Problems and methods. Translated from the Italian by A. LaBello, XII+385 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984. — DM 96,—.

Probability Measures on Groups VII. Proceedings of a Conference held in Oberwolfach, 24—30 April 1983. Edited by H. Heyer (Lecture Notes in Mathematics, Vol. 1064), X+588 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984. — DM 76,—.

R. A. Rosenbaum—P. Johnson, Calculus: Basic concepts and applications, XVI+422 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1984. — £ 12.50.

R. Sato—T. Nono, Invariance principles and the structure of technology (Lecture Notes in Economics and Mathematical Systems, Vol. 212), V+94 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983. — DM 24,—.

Séminaire de Probabilité XVIII, 1982/83. Proceedings. Edité par J. Azema, M. Yor (Lecture Notes in Mathematics, Vol. 1059), IV+518 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984. — DM 62,—.

I. M. Sigal, Scattering theory for many-body quantum mechanical systems. Rigorous results (Lecture Notes in Mathematics, Vol. 1011), IV+132 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983. — DM 19,80.

E. Solomon, Games programming, XI+257 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1984. — £ 7.50.

Statistics and Probability. Proceedings of the 3rd Pannonian Symposium on Mathematical Statistics, Visegrád, Hungary, 13—18 September, 1982. Edited by J. Mogyoródi, I. Vincze, and W. Wertz, X+415 pages, Akadémiai Kiadó, Budapest, 1984.

V. S. Varadarajan, Lie groups, Lie algebras, and their representations (Graduate Texts in Mathematics, Vol. 102), XIII+430 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984. — DM 88,—.

F. V. Warner, Foundations of differentiable manifolds and Lie groups (Graduate Texts in Mathematics, Vol. 94), IX+272 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983. — DM 58,—

TOMUS XLVII—1984—47. KÖTET

Bibliographie

R. E. Burkard—U. Derigs, Assignment and Matching Problems: Solutions Methods with FORTRAN-Programs. — Complex Analysis, Methods, Trends and Applications, — K. J. Devlin, Fundamentals of Contemporary Set Theory. — Differential Equations, Proceedings, Sao Paulo, 1981. — R. E. Edwards, Fourier Series, a Modern Introduction. — P. J. Federico, Descartes on Polyhedra. — J. E. Fenstad, General Recursion Theory. — J. Flum—M. Ziegler, Topological Model Theory. — Geometry and Analysis. — T. W. Hungerford, Algebra. — Iterative Solution of Nonlinear Systems of Equations, Proceedings, Oberwolfach, 1982. — U. Kastens—B. Hutt—E. Zimmermann, Gag: A Practical Compiler Generator. — T. Kato, A Short Introduction to Perturbation Theory for Linear Operators. — J. H. van Lint, Introduction to Coding Theory. — J. Lützen, The Prehistory of the Theory od Distributions. — Yu, I. Manin, A Course in Mathematical Logic. — Mathematical Modeling of the Hearing Process, Proceedings, Troy NY, 1980. — Numerical Integration of Differential Equations and Large Linear Systems, Proceedings, Bielefeld, 1980. — Probability in Banach Spaces IV, Proceedings, Oberwolfach, 1983. — Probability Theory and Mathematical Statistics, Proceedings, Tbilisi, 1982. — Séminaire de Probabilités XVII, 1981/82. — C. Sparrow, The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors. — Stability Problems for Stochastic Models. — The Mathematical Gardner. — J. Uhl—S. Drossopoulou—G. Porsch—G. Goss—M. Dausmann—G. Winterstein—W. Kirschgässner, An Attribute Grammar for Semantic Analysis of Ada. — M. I. Yadrenko, Special Theory of Random Fields. .... 249—262

G. Alexits, Approximation Theory (Selected Papers). — V. I. Arnold, Geometrical Methods in the Theory of Ordinary Differential Equations. — V. I. Arnold, Catastrophe Theory. — B. Aupetit, Propriétés Spectrales des Algèbres de Banach. — D. Bleecker, Gauge Theory and Variational Principles. — E. A. Coddington—H. S. V. Snoo, Regular Boundary Value Problems Associated with Pairs of Ordinary Differential Expressions. — Combinatorial Mathematics X, Proceedings, Adelaide 1982. — Complex Analysis and Spectral Theory (Seminar, Leningrad 1979/80). — Differential Equations Models. — Life Sciences Models. — K. Donner, Extension of Positive Operators and Korovkin Theorems. — Dynamical Systems and Turbolence, Warwick 1980. — E. Fischer, Intermediate Real Analysis. — G. B. Folland, Lectures on Partial Differential Equations. — F. Gécseg—M. Steinby, Tree Automata. — Geometric Dynamics, Proceedings, Rio de Janeiro 1981. — Geometric Techniques in Gauge Theories, Proceedings, Scheveningen 1981. — Geometries and Groups, Proceedings, Berlin 1981. — A.

INDEX—TARTALOM

ACTA SCIENTIARUM MATHEMATICARUM
SZEGED (HUNGARIA), ARADI VÉRTANÚK TERE 1

On peut s'abonner à l'entreprise de commerce des livres et journaux „Kultúra" (1061 Budapest, I., Fő utca 32)

ISSN 034-6523 Acta Univ. Szeged
ISSN 0001-6969 Acta Sci. Math.

INDEX: 26 024

84-2544 — Szegedi Nyomda — F. v.: Dobó József igazgató

Felelős szerkesztő és kiadó: Leindler László
A kézirat a nyomdába érkezett: 1984. június 6
Megjelenés: 1985. február

Példányszám: 1000. Terjedelem: 22,57 (A/5) ív
Készült monószedéssel, íves magasnyomással,
az MSZ 5601-24 és az MSZ 5602-55 szabvány szerint