ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

ADIUVANTIBUS

B. CSÁKÁNY
S. CSÖRGŐ
E. DURSZT
F. GÉCSEG
L. HATVANI
L. KÉRCHY
L. MEGYESI
F. MÓRICZ
P. T. NAGY
J. NÉMETH
L. PINTÉR
G. POLLÁK
L. L. STACHÓ
L. SZABÓ
I. SZALAY
Á. SZENDREI
B. SZ.-NAGY
K. TANDORI
J. TERJÉKI
V. TOTIK

REDIGIT

L. LEINDLER

TOMUS 51

FASC. 3—4

SZEGED, 1987

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

CSÁKÁNY BÉLA
CSÖRGŐ SÁNDOR
DURSZT ENDRE
GÉCSEG FERENC
HATVANI LÁSZLÓ
KÉRCHY LÁSZLÓ
MEGYESI LÁSZLÓ
MÓRICZ FERENC
NAGY PÉTER
NÉMETH JÓZSEF
PINTÉR LAJOS
POLLÁK GYÖRGY
STACHÓ LÁSZLÓ
SZABÓ LÁSZLÓ
SZALAY ISTVÁN
SZENDREI ÁGNES
SZŐKEFALVI-NAGY BÉLA
TANDORI KÁROLY
TERJÉKI JÓZSEF
TOTIK VILMOS

KÖZREMŰKÖDÉSÉVEL SZERKESZTI

LEINDLER LÁSZLÓ

51. KÖTET

FASC. 3—4

SZEGED, 1987

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

Pseudomodular lattices and continuous matroids

A. BJÖRNER*) and L. LOVÁSZ

Dedicated to the memory of András Huhn

0. Introduction

If $(E, \mathcal{M})$ is a linear matroid (i.e., represented by a subset of vectors in a linear space) then $(E, \mathcal{M})$ can be embedded in the full linear matroid (the matroid formed by all vectors in that linear space) in a natural way. The most significant property of the full linear matroids is that the lattice of their flats is modular. (In fact, apart from direct sums, loops and parallel elements and the non-desarguesian projective planes, this property characterizes full linear matroids.)

Other classes of matroids like graphic, algebraic and transversal matroids also have natural "full" members, which are, however, non-modular. For the case of full algebraic matroids, INGLETON and MAIN [7] proved that the following property (strictly weaker than modularity) still holds: any three lines such that any two are coplanar, but all three are not coplanar, have a point in common. LINDSTRÖM [11]—[13] observed that this fact is a basic property of full algebraic matroids, and used it to prove that several other geometric results on projective spaces, e.g. Desargues's Theorem, also carry over to full algebraic matroids.

DRESS and LOVÁSZ [4] proved various generalizations of the Ingleton—Main Lemma, and showed that one of them suffices to extend the minimax formula for matchings in linear matroids (LOVÁSZ [14]) to algebraic matroids. It was observed that full graphic matroids (or, equivalently, partition lattices) and full transversal matroids also have this property. Another related property, the existence of "pseudo-intersections", was established for the full algebraic matroids.

Modularity of the subspace lattice of linear spaces also plays a crucial role in a contruction, due to VON NEUMANN [16], of "continuous geometries". To obtain

Received September 1, 1986, and in revised form March 25, 1987.

*) Partially supported by the National Science Foundation.

these, one embeds the subspace lattice of the n-dimensional linear space over some field F in the subspace lattice of the nq-dimensional linear space over the same field, so that any flat x of rank $r(x)$ is mapped onto a flat of rank $q \cdot r(x)$. Such a "stretch embedding" which preserves meet and join can be constructed using the modularity of the lattices.

A construction of a "continuous partition lattice" based on stretch embeddings was given by Björner [2]. This construction depends on the fact that partition lattices have sufficiently many modular elements. A general scheme to obtain continuous analogues of sequences of geometric lattices was also outlined: the scheme depends on the existence of "stretch embeddings" between these lattices.

The main result of this paper is that the existence of "pseudointersections" in suitable sequences of geometric lattices can be used to construct stretch embeddings and thereby continuous analogues. In particular, we construct continuous transversal geometries, continuous algebraic geometries over any field, and obtain a new theoretical explanation for the existence of the continuous partition lattice.

Semimodular lattices with pseudointersections, which we call *pseudomodular*, seem to be worth studying even without an eye on continuous geometries. We shall show that such lattices arise in the study of antimatroids (abstract convexity spaces). In fact, an antimatroid with Caratheodory number 2 has a pseudomodular lattice of feasible sets.

1. Pseudomodular lattices

In this paper we shall assume some modest familiarity with lattices and matroids. For details concerning these notions see Birkhoff [1] and Welsh [17], respectively.

Let L be a semimodular lattice. We assume without further mention that all semimodular lattices considered have finite rank. Let $r(x)$ denote the rank function of L. For each $x, y \in L$, we set

$$P_{x,y} = \{z \leqq y\colon\ r(x \vee z) - r(z) = r(x \vee y) - r(y)\}.$$

Note that it would suffice to require that $r(x \vee z) - r(z) \leqq r(x \vee y) - r(y)$ in this definition, since the reverse inequality is always true by the submodularity of the rank function.

This set lies in the interval $[x \wedge y, y]$. To see this, let $z \in P_{x,y}$. Then $z \leqq (x \vee z) \wedge y$ and hence by the submodularity of the rank function,

$$\begin{aligned} r(z) &\leqq r\big((x \vee z) \wedge y\big) \leqq r(x \vee z) + r(y) - r\big((x \vee z) \vee y\big) = \\ &= r(x \vee z) + r(y) - r(x \vee y) = r(z). \end{aligned}$$

So $z = (x \vee z) \wedge y \geqq x \wedge y$.

Clearly, $P_{x,y}$ is a filter in the interval $[x\wedge y, y]$, i.e., if $z\in P_{x,y}$ and $z\leqq u\leqq y$ then $u\in P_{x,y}$.

If the set $P_{x,y}$ has a unique least element then we call this the *pseudointersection* of x and y and denote it by $x\urcorner y$. The lattice L is called *pseudomodular* if every pair of its elements have a pseudointersection.

Note that in general $x\urcorner y\neq y\urcorner x$ (cf. Lemma 1.1 below). Furthermore, the existence of a pseudointersection is not a symmetric relation. For an example of this, take three pairwise parallel lines in affine 3-space, not all in a plane. Let y denote one of these lines and x, the plane spanned by the other two. Then in the geometric lattice formed by the points of these lines, $x\urcorner y$ exists but $y\urcorner x$ does not.

The relationship between the pseudointersection and the (ordinary) intersection of two lattice elements is illuminated by the following lemma.

Lemma 1.1. *For any two elements x and y in a semimodular lattice L, the following are equivalent:*

(i) *x and y form a modular pair, i.e., $r(x\vee y)+r(x\wedge y)=r(x)+r(y)$.*
(ii) *$x\urcorner y$ exists and $x\urcorner y\leqq x$.*
(iii) *$x\urcorner y$ exists and $x\urcorner y=x\wedge y$.*
(iv) *$x\wedge y\in P_{x,y}$.*

Proof. All implications (i)→(iv)→(iii)→(ii)→(i) are straightforward.

The following lemma gives some means to verify the existence of pseudointersections.

Lemma 1.2. *For any two elements x and y of a semimodular lattice L, the following are equivalent:*

(i) *$x\urcorner y$ exists, i.e., $P_{x,y}$ has a unique least element.*
(ii) *$P_{x,y}$ is closed under meets.*
(iii) *If $u, v, z\in P_{x,y}$ and z covers u and v, then $u\wedge v\in P_{x,y}$.*

Proof. The only non-trivial implication is that (iii)→(i). To verify this, assume, by way of contradiction, that a and b are distinct minimal elements of $P_{x,y}$, and choose a and b so that $a\vee b$ is as low in the lattice as possible. Let u be an element in the interval $[a, a\vee b]$ covered by $a\vee b$ and let v be an element in the interval $[b, a\vee b]$ covered by $a\vee b$. Then by (iii), $u\wedge v\in P_{x,y}$. Let c be a minimal element of $P_{x,y}$ below $u\wedge v$, then $a\vee c\leqq u<a\vee b$ and $b\vee c\leqq v<a\vee b$. Since c is distinct from at least one of a and b, this contradicts the choice of a and b. (This lemma in fact holds for any filter in any interval of any lattice of finite length.)

It will be useful to remark that the assertion (iii) in Lemma 1.2 holds automatically if u, v is a modular pair, i.e., if u covers $u\wedge v$. For, by submodularity and the definition of $P_{x,y}$, we have the following:

$$1 = r(u) - r(u \wedge v) \geqq r(u \vee x) - r((u \wedge v) \vee x) \geqq$$
$$\geqq r(z \vee x) - r(v \vee x) = r(z) - r(v) = 1.$$

So equality must hold throughout, and equality in the first inequality means just that $u \wedge v \in P_{x,y}$.

Lemma 1.3. *Let L be a geometric lattice and $x, y \in L$. If $x \urcorner y$ exists then it is equal to the meet of all $z \in L$ such that y covers z and $x \vee y$ covers $x \vee z$.*

Proof. Let $t = \wedge \{z:\ y$ covers z and $y \vee x$ covers $z \vee x\}$. Note that the second condition on z is equivalent to $z \in P_{x,y}$. Hence $t \in P_{x,y}$ and thus $t \geqq x \urcorner y$. On the other hand, the interval $[x \urcorner y, y] = P_{x,y}$ is a geometric lattice and hence its bottom element is the meet of its coatoms. This proves that $t = x \urcorner y$.

The existence of pseudointersections can be characterized by the non-existence of certain configurations in the lattice. Such a result is stated in the following theorem.

Theorem 1.4. *Let L be any semimodular lattice. Then the following are equivalent:*

(i) *L is pseudomodular.*

(ii) *Let $a, b, c \in L$, and assume that $r(a \vee c) - r(a) = r(b \vee c) - r(b) = r(a \vee b \vee c) - r(a \vee b)$. Then $r((a \vee c) \wedge (b \vee c)) - r(a \wedge b) = r(a \vee c) - r(a)$.*

(iii) *Let $x, y, z \in L$ and assume that x covers $x \wedge z$ and y covers $y \wedge z$. Then $r(x \wedge y) - r(x \wedge y \wedge z) \leqq 1$.*

(iv) *Let $x, y, z, u \in L$, and assume that u covers x, y and z, and z covers $x \wedge z$ and $y \wedge z$. Then $r(x \wedge y) - r(x \wedge y \wedge z) \leqq 1$.*

(v) *Let $x, y, z, u \in L$, and assume that u covers x and y, $z \leqq u$, and z covers $x \wedge z$ and $y \wedge z$. Then $r(x \wedge y) - r(x \wedge y \wedge z) \leqq r(u) - r(z)$.*

Remark. Property (ii) has the following consequences. Since $c \leqq (a \vee c) \wedge (b \vee c)$, it implies that $r(c) - r(a \wedge b) \leqq r(a \vee c) - r(a)$. Also, it follows that a and $(a \vee c) \wedge (b \vee c)$ form a modular pair and $a \wedge (a \vee c) \wedge (b \vee c) = a \wedge b$. Hence, $a \wedge c = a \wedge (a \vee c) \wedge (b \vee c) \wedge \wedge c = a \wedge b \wedge c$. Similarly, $b \wedge c = a \wedge b \wedge c$. It also follows that $a \wedge c \leqq a \wedge b$.

Proof. (i)→(ii): Let $d = (a \vee c) \wedge (b \vee c)$. Then $a \in P_{d, a \vee b}$ since $a \vee c = a \vee d$ and $a \vee b \vee c = a \vee b \vee d$, and so $r(a \vee d) - r(a) = r(a \vee b \vee d) - r(a \vee b)$ by the hypothesis in (ii). Similarly $b \in P_{d, a \vee b}$ and hence by the pseudomodularity of L, $a \wedge b \in P_{d, a \vee b}$. This means that $r((a \wedge b) \vee d) - r(a \wedge b) = r(a \vee b \vee d) - r(a \vee b)$. Since $(a \wedge b) \vee d = d$ and $a \vee b \vee d = a \vee b \vee c$, this implies the assertion of (ii).

(ii)→(iii): We may assume that $z = (x \wedge z) \vee (y \wedge z)$; if this is not already the case we can just let $(x \wedge z) \vee (y \wedge z)$ play the role of z without changing the situation.

We may also assume that $x\wedge y\not\leqq z$ (since otherwise $x\wedge y=x\wedge y\wedge z$), and that $x\neq y$. It follows that $x\wedge z\neq y\wedge z$, and $x=(x\wedge z)\vee(x\wedge y)$, $y=(y\wedge z)\vee(x\wedge y)$. Also, $r(z)<r(z\vee(x\wedge y))\leqq r(z\vee x)\leqq r(z)+r(x)-r(z\wedge x)=r(z)+1$. Hence, $z\vee(x\wedge y)$ covers z. Now letting $a=x\wedge z$, $b=y\wedge z$ and $c=x\wedge y$ in (ii), assertion (iii) follows.

(iii)→(iv) An easy special case.

(iv)→(v): We prove this by induction on $r(u)-r(z)$; (iv) is just the special case of (v) when this difference is 1. We may assume that $x\wedge y\wedge z\neq x\wedge y$. Let p be an element of the interval $[x\wedge y\wedge z, x\wedge y]$ covering $x\wedge y\wedge z$. Then clearly $p\not\leqq z$ and so $p\not\leqq x\wedge z$ and $p\not\leqq y\wedge z$. Hence $v=z\vee p$ covers z by submodularity and similarly, $(z\wedge x)\vee p$ covers $z\wedge x$ and $(z\wedge y)\vee p$ covers $z\wedge y$. Clearly $(z\wedge x)\vee p\leqq v\wedge x<v$ and hence $v\wedge x=(z\wedge x)\vee p$. Hence $v\wedge x$ is covered by v and similarly, $v\wedge y$ is also covered by v. Applying (iv) with v, $v\wedge x$, $v\wedge y$ and z in place of u, x, y and z we obtain that

$$r(x\wedge y\wedge v)-r(x\wedge y\wedge z)\leqq 1.$$

Applying the induction hypothesis with u, x, y and v in place of u, x, y and z we obtain that

$$r(x\wedge y)-r(x\wedge y\wedge v)\leqq r(u)-r(v)=r(u)-r(z)-1.$$

This proves (v).

(v)→(i): We verify Lemma 1.2(iii). Let $u, v, z\in P_{x,y}$, where z covers both u and v. Then by the definition of $P_{x,y}$, $z\vee x$ covers $u\vee x$ and $v\vee x$, and $z\wedge(u\vee x)=u$, $z\wedge(v\vee x)=v$. So (v) can be applied with $u\vee x$, $v\vee x$, z and $z\vee x$ in place of x, y, z and u, and we obtain that

$$r((u\vee x)\wedge(v\vee x))-r((u\vee x)\wedge(v\vee x)\wedge z)\leqq r(z\vee x)-r(z).$$

Since, as remarked, $(u\vee x)\wedge(v\vee x)\wedge z=u\wedge v$, this implies that

$$r((u\wedge v)\vee x)-r(u\wedge v)\leqq r((u\vee x)\wedge(v\vee x))-r(u\wedge v)\leqq r(z\vee x)-r(z),$$

which proves that $u\wedge v\in P_{x,y}$.

Lindström [13] proved the following generalization of the Ingleton—Main Lemma for full algebraic matroids: if a, b and c are three flats such that $r(a)=r(b)=r(c)=n$, $r(a\vee b)=r(a\vee c)=r(b\vee c)=n+1$ and $r(a\vee b\vee c)=n+2$ then $a\wedge b=a\wedge c=b\wedge c=a\wedge b\wedge c$ and $r(a\wedge b\wedge c)=n-1$. This follows immediately from property (iv) in the above theorem. He conjectured that if a, b and c are three flats in an algebraic matroid such that $r(a)=r(b)=r(c)=n$, $r(a\vee b)=r(a\vee c)=r(b\vee c)=n+k$ and $r(a\vee b\vee c)=n+2k$ then $a\wedge b=a\wedge c=b\wedge c=a\wedge b\wedge c$ and $r(a\wedge b\wedge c)=n-k$. This conjecture follows from the inequality in (ii) easily.

2. Examples of pseudomodular lattices

In this section we discuss some classes of semimodular lattices which have pseudointersections. We start with two obvious examples:

Example 1. Modular lattices.

Example 2. Semimodular lattices of length at most 3.

Next we discuss three families of geometric lattices, (i.e., matroids) which have pseudointersections. These are "full" members in their own class (algebraic matroids, graphic matroids, transversal matroids) in a very natural way. The "full" linear matroids, i.e., linear spaces, have a modular subspace lattice and hence they are covered by Example 1. It would be important to understand the structure of those classes of matroids which have natural "full" members and why these full members tend to be pseudomodular.

Example 3. Full algebraic matroid lattices. These can be described as follows: let F and K be algebraically closed fields and $F\subset K$. Then the algebraically closed subfields of K containing F form a geometric lattice, which we denote by $\mathcal{L}(F, K)$.

The fact that $\mathcal{L}(F, K)$ has pseudointersections was shown by Dress and Lovász [4]. For the sake of completeness, we describe the simple construction of the operation $\urcorner$. So let X and Y be two algebraically closed fields with $F\subset X$, $Y\subset K$. Let $\{x_1, \ldots, x_m\}$ be a transcendence basis of X over F. Consider the ideal I of all polynomials over Y in m variables which are satisfied by $(x_1, \ldots, x_m)$, and a basis $q_1, \ldots, q_N$ of this ideal. We may assume that each q_i has at least one coefficient that is equal to 1. Then the algebraically closed subfield T of Y generated by the coefficients of $q_1, \ldots, q_N$ is the pseudointersection of X and Y.

Example 4. Partition lattices, i.e., circuit matroids of complete graphs. We show that the lattice of partitions of a set E has pseudointersections, using Lemma 1.2(iii). Assume that u, v and z are three partitions in $P_{x,y}$, and that z covers both u and v, i.e., both u and v arise from z by splitting a partition class into two. The fact that $u, v, z\in P_{x,y}$ implies that $r(x\vee z)-r(z)=r(x\vee u)-r(u)=r(x\vee v)-r(v)=r(x\vee y)-r(y)$, and hence $r(x\vee u)=r(x\vee v)=r(x\vee z)-1$. We want to show that $r(x\vee u)-r(x\vee(u\vee v))=r(u)-r(u\wedge v)$.

By the remark after Lemma 1.2, the only non-trivial case to consider is when u and v do not form a modular pair, i.e., when they arise from z by splitting the same class A in two different ways $A'_u\cup A''_u$ and $A'_v\cup A''_v$ so that the intersections $B_1=A'_u\cap A'_v$, $B_2=A'_u\cap A''_v$, $B_3=A''_u\cap A'_v$ and $B_4=A''_u\cap A''_v$ are all non-empty. So $r(u)-r(u\wedge v)=2$, and submodularity implies that $r(x\vee u)-r(x\vee(u\wedge v))\leqq 2$. Now if $r(x\vee u)-r(x\vee(u\wedge v))\leqq 1$ then the sets B_1, B_2, B_3 and B_4 cannot belong to different classes in $x\vee(u\wedge v)$ and hence there exists a sequence $x_1, \ldots, x_k$ of elements of E such that $x_1 \in B_i$, $x_n \in B_j$ $(i \neq j)$, no other member of the sequence belongs to A, and any two consecutive members of the sequence are either in one class of x or in one class of $u \wedge v$. Without loss of generality we may assume that $B_i \subset A'_u$ and $B_j \subset A''_u$. But then B_i and B_j must belong to the same class of $x \vee u$, which is a contradiction.

Example 5. Full transversal matroids. The full transversal matroid $\mathcal{TM}(r)$ of rank r is defined as follows. First we construct a bipartite graph G. Let S be a set with r elements; this will be one of the color classes. For each subset $S' \subseteqq S$, we take denumerably infinitely many new vertices and connect them by edges to the vertices in S'. The set T of these new vertices will be other color class. Now $\mathcal{TM}(r)$ is defined as the transversal matroid induced by G on T. So a set $T' \subseteqq T$ is independent iff G contains a matching covering T'.

Using König's Theorem, it is easy to show that the flats in $\mathcal{TM}(r)$ have the following structure: take a set $A \subseteqq S$, and also a set $B \subseteqq T$ such that every non-empty subset $B' \subseteqq B$ has at least $|B'|+1$ neighbors in $S-A$. Let $\Omega(A)$ denote the set of points x in T such that all neighbors of x are in A. Then $F(A, B) = \Omega(A) \cup B$ is a flat of rank $|A|+|B|$ in $\mathcal{TM}(r)$, and every flat is of this form.

The pseudomodularity of full transversal matroids (in fact, of a much larger class of transversal matroids) will follow from the results in the next section.

Example 6. Antimatroids with Caratheodory number 2. Antimatroids were introduced by Edelman [5] and Jamison-Waldner [8] as combinatorial abstractions of convex sets. For our purposes, the following definition will suffice. Let E be a finite set and $\mathcal{F}$, a family of subsets of E with the following properties:

a) if $X \in \mathcal{F}$ and $Y \in \mathcal{F}$ then $X \cup Y \in \mathcal{F}$;

b) if $X \in \mathcal{F}$, $X \neq \emptyset$ then there exists an element $x \in X$ such that $X - x \in \mathcal{F}$.

Then the pair $(E, \mathcal{F})$ is called an *antimatroid*. The members of $\mathcal{F}$ are called *feasible sets*, their complements are called *convex sets*. Since the family of convex sets is closed under intersection, we can define the *convex hull* of any subset X of E as the intersection of all convex supersets of X. These notions share many of the combinatorial properties of convex sets in the usual sense. We shall need the following two elementary facts: (I) if X and Y are feasible and $Y \not\subset X$, then there exists an element $y \in Y - X$ such that $X \cup y$ is feasible; (II) p is in the convex hull of G if and only if every feasible set containing p has a non-empty intersection with G.

We define the *Caratheodory number* of an antimatroid as the least integer k with the following property: whenever an element p is contained in the convex hull of a set G, it is also contained in the convex hull of some subset $G' \subset G$ with $|G'| \leqq k$. In the language of Korte and Lovász [9], the Caratheodory number is one less than the maximum size of a circuit of the antimatroid. For various properties of antimatroids, see also EDELMAN and JAMISON [6], KORTE and LOVÁSZ [10], BJÖRNER, KORTE and LOVÁSZ [3].

The feasible sets of an antimatroid of $\mathscr{F}$ form a semimodular lattice $\mathscr{L}(E, \mathscr{F})$ under ordinary inclusion. More strongly, $\mathscr{L}(E, \mathscr{F})$ is *locally free* (i.e., the elements covering any given element generate a Boolean subalgebra) and every locally free semimodular lattice has a unique representation as the feasible set lattice of an antimatroid (EDELMAN [5]). The rank of any element $X\in\mathscr{F}$ in this lattice is just its cardinality.

It was proved by KORTE and LOVÁSZ [9] that if the Caratheodory number of an antimatroid is 1 then its feasible sets are the ideals of a poset and hence the lattice is distributive (and therefore modular). Conversely, it is easy to see that for all other kinds of antimatroids, the lattice $\mathscr{L}(E, \mathscr{F})$ is non-modular.

We now prove that if $(E, \mathscr{F})$ has Caratheodory number at most two then $\mathscr{L}(E, \mathscr{F})$ is pseudomodular.

Let X and Y be any two feasible sets. Then by the definition of the rank and of $P_{X,Y}$, we have that

$$P_{X,Y} = \{Z\in\mathscr{F}:\ X\cap Y \subset Z \subset Y\}.$$

To show that X and Y have a pseudointersection, it suffices to verify the following (by Lemma 1.2(iii)): let Z, $Z-u$ and $Z-v$ be feasible sets with $X\cap Y\subset Z-u$, $Z-v$ and $Z\subset Y$, and let W be the largest feasible subset of $Z-u-v$; then $X\cap Y\subset W$. Suppose that this is not the case, then there exists an element $p\in(X\cap Y)-W$. Let $G=\{g\in E-W\colon W\cup\{g\}\in\mathscr{F}\}$. Then p is in the convex hull of G (this follows from properties (I) and (II) of antimatroids) and hence, by the definition of the Caratheodory number, we have a pair $\{q, r\}\subset G$ such that p is in the convex hull of $\{q, r\}$. Since p is an element in the feasible set $Z-u$, it follows that one of q and r must belong to $Z-u$. But none of q and r can belong to $Z-u-v$ since this would contradict the choice of W. Hence v must be one of q and r. Similarly, u must be one of q and r. But then X is a feasible set containing p but not q and r, which is a contradiction.

There are several important classes of antimatroids with Caratheodory number 2. We mention just a few:

Example 6a. Let E be any poset and let the convex sets be those sets which contain, along with any two comparable elements x, y, the whole interval $[x, y]$.

Example 6b. Let E be the vertex [edge] set of any tree T and let the convex sets be the vertex [edge] sets of subtrees.

Example 6c. Let E be any finite set in R^2 and let the convex sets be those subsets which contain, along with any two elements x and y, every point of E in the region of the plane bounded by the segment xy and by semilines pointing "upwards" from x and y.

3. Constructions preserving pseudomodularity

We show that some standard operations on semimodular lattices preserve pseudomodularity.

3.1. Direct product.

3.2. Truncation. For a semimodular lattice L and integer $k \geqq 1$, let $L_k = \{x \in L: r(x) < k \text{ or } x = 1\}$. Then the truncated lattice L_k is again semimodular, and it is easy to see that pseudomodularity is also preserved. The pseudointersection $x \urcorner_k y$ $(y \neq 1)$ in L_k is given by

$$x \urcorner_k y = \begin{cases} x \urcorner y, & \text{if } r(x \vee y) \leqq k \text{ in } L, \\ y, & \text{otherwise.} \end{cases}$$

A less trivial operation preserving pseudomodularity is the following:

3.3. Principal extension. Let L be a semimodular lattice and $w \in L - \{0\}$. The *principal extension* of L with respect to w is defined on the set

$$L' = L \cup \{y+p: y \in L,\ r(y \vee w) \geqq r(y)+2\}.$$

Here $y+p$ denotes a new element associated with the old lattice element y. The ordering is defined as before on the old elements, and by

$$\begin{aligned} x &\leqq y+p && \text{iff } x \leqq y, \\ x+p &\leqq y+p && \text{iff } x \leqq y, \\ x+p &\leqq y && \text{iff } x \vee w \leqq y \end{aligned}$$

for $x, y \in L$. In particular it follows that $0+p$, which we denote shortly by p, is an atom and more generally, x is covered by $x+p$ whenever the latter exists. One can verify that L', with this partial ordering, is a semimodular lattice, containing L as a sublattice.

(This construction is best known for a geometric lattice, i.e., the lattice of flats of a matroid. Then the principal extension of L with respect to w means creating a new point p of the matroid which is "in general position" on the flat w.)

Theorem 3.4. *A principal extension of a pseudomodular lattice is again pseudomodular.*

Proof. The proof is more-or-less straightforward; nevertheless, we include it here for completeness. Let L be a pseudomodular lattice and $w \in L - \{0\}$. Let L' be the principal extension of L with respect to w. Observe that the class of "new" elements is closed under intersection: $(x+p) \wedge (y+p) = (x \wedge y)+p$, and so is of course the class of "old" elements. Further, if x is "old" and $y+p$ is "new" then $x\wedge(y+p)=x\wedge y$ if $p\nleqq x$, and $x\wedge(y+p)=(x\wedge y)+p$ if $p\leqq x$.

We verify that condition (iv) of Theorem 1.4 holds for L'. Let x, y, z and u be elements of L' as in (iv). We may assume that they are distinct and that $x\wedge y\wedge z\neq$ $\neq x\wedge y$. The argument will be divided into several cases depending on the distribution of "new" elements among x, y, and z.

Case 1. x, y and z are "old". Then u also must be "old", and we know that (iv) is valid in L.

Case 2. $x=x_0+p$, $y=y_0+p$ and $z=z_0+p$ are "new" elements. Then condition (iii) applied to x_0, y_0 and z_0 within L, implies (iv) for x, y and z.

Case 3. $z=z_0+p$, and x, y are "old". Then we have the following subcases.

Subcase 3.1. $p\nleqq x, y$. Then $x\wedge z$ is an "old" element covered by z and hence, $x\wedge z=z_0$. Similarly, $y\wedge z=z_0$ and the assertion is obvious.

Subcase 3.2. $p\leqq x$ but $p\nleqq y$ (say). Then as before, $y\wedge z=z_0$ and hence $x\wedge y\wedge z=x\wedge z_0$. Since $x\wedge z=(x\wedge z_0)+p$, it follows that

$$r(x\wedge y)\leqq r(x)-1=r(x\wedge z)=r(x\wedge z_0)+1=r(x\wedge y\wedge z)+1,$$

which proves (iv).

Subcase 3.3. $p\leqq x, y$. Then $z\wedge x=(z_0\wedge x)+p$ is covered by $z=z_0+p$ by hypothesis, and hence $z_0\wedge x$ is covered by z_0. Similarly, $z_0\wedge y$ is covered by z_0. Since x is an "old" element above p, and u covers x, u must also be "old". Hence u, x, y and z_0 are elements of the old lattice L satisfying the conditions of Theorem 1.4 (v), and hence by the pseudomodularity of L, we obtain that

$$r(x\wedge y)-r(x\wedge y\wedge z_0)\leqq r(u)-r(z_0)=2.$$

Since $x\wedge y\wedge z=(x\wedge y\wedge z_0)+p$, again (iv) follows.

Case 4. $x=x_0+p$, and z is "old". By symmetry this also handles the case when y is "new" and z is "old".

Subcase 4.1. $p\nleqq z$. Then $x\wedge z$ is an "old" element covered by x and hence $x\wedge z=x_0$. So $x\wedge y\wedge z=x_0\wedge y$. Now $x\wedge y$ is either $x_0\wedge y$ or $(x_0\wedge y)+p$, which proves (iv).

Subcase 4.2. $p\leqq z$ and y is "old". Then $x\wedge z=(x_0\wedge z)+p$ is covered by $x=x_0+p$ by hypothesis, and hence $x_0\wedge z$ is covered by x_0. We can apply Theorem 1.4 (iii) to the "old" elements x_0, y and z and obtain that $r(x_0\wedge y)-r(x_0\wedge y\wedge z)\leqq 1$. Now, if $p\leqq y$ then $x\wedge y\wedge z=(x_0\wedge y\wedge z)+p$ and $x\wedge y=(x_0\wedge y)+p$; if $p\nleqq y$ then $x\wedge y\wedge$ $\wedge z=x_0\wedge y\wedge z$ and $x\wedge y=x_0\wedge y$. In either case, (iv) follows.

Subcase 4.3. $p \leqq z$ and $y = y_0 + p$. Then, as in the preceding case, $x_0 \wedge z$ is covered by x_0. Similarly, $y_0 \wedge z$ is covered by y_0. Apply (iii) to the elements x_0, y_0 and z, and obtain that $r(x_0 \wedge y_0) - r(x_0 \wedge y_0 \wedge z) \leqq 1$. Now, $x \wedge y \wedge z = (x_0 \wedge y_0 \wedge z) + p$ and $x \wedge y = (x_0 \wedge y_0) + p$, and we are done again.

Case 5. $x = x_0 + p$, $z = z_0 + p$, and y is "old". By symmetry this also handles the case when x is the only "old" element.

Subcase 5.1. $p \not\leqq y$. Then $y \wedge z$ is an "old" element covered by z, hence $y \wedge z = z_0$, and $x \wedge y \wedge z = x \wedge z_0 = x_0 \wedge z_0$. Since x covers $x \wedge z = (x_0 \wedge z_0) + p$, we get $r(x) - r(x_0 \wedge z_0) = 2$. So, $r(x \wedge y) \leqq r(x) - 1 = r(x \wedge y \wedge z) + 1$, and (iv) follows.

Subcase 5.2. $p \leqq y$. Since $x \wedge z = (x_0 \wedge z_0) + p$ and $y \wedge z = (y \wedge z_0) + p$, we have that $r(x) = r(x_0) + 1 = r(z_0) + 1 = r(x_0 \wedge z_0) + 2 = r(y \wedge z_0) + 2$. We may assume that $x_0 \wedge z_0 \not\leqq y$ (else $x \wedge z \leqq y$ and $x \wedge y \wedge z = x \wedge y$). Choose $t \in L$ so that $y \wedge z_0 < t < y$. Since $(x_0 \wedge z_0) \vee y = u$ covers y and $(x_0 \wedge z_0) \vee (y \wedge z_0) = z_0$ covers $y \wedge z_0$, it follows by semimodularity that $z' = (x_0 \wedge z_0) \vee t$ covers t. Clearly $z' \in L$, $r(z') = r(y) = r(t) + 1$, and $y \wedge z' = t$.

First, suppose that $x_0 \not\leqq z'$. Then $x_0 \wedge z' = x_0 \wedge z_0$, which is covered by x_0. Applying Theorem 1.4 (iii) to the "old" elements x_0, y, z', we obtain that $r(x_0 \wedge y) - 1 \leqq r(x_0 \wedge z' \wedge y) = r(x_0 \wedge z_0 \wedge y)$.

Second, suppose that $x_0 \leqq z'$. Since t covers $t \wedge z_0 = y \wedge z_0$, we may apply Theorem 1.4 (iii) to the elements x_0, t and z_0. This yields $r(x_0 \wedge t) - 1 \leqq r(x_0 \wedge z_0 \wedge t) = r(x_0 \wedge z_0 \wedge y)$. But $x_0 \wedge t = x_0 \wedge z' \wedge y = x_0 \wedge y$.

We have shown that in either case $r(x_0 \wedge y) - r(x_0 \wedge z_0 \wedge y) \leqq 1$. Since $x \wedge y = (x_0 \wedge y) + p$ and $x \wedge y \wedge z = (x_0 \wedge z_0 \wedge y) + p$, this proves (iv).

Observe that full transversal matroids, as defined in the previous section, can be obtained from Boolean algebras by principal extensions (infinitely often with respect to each flat). Hence the pseudomodularity of these matroids follows by an easy compactness argument. More generally, every matroid which can be obtained by principal extensions from Boolean algebras is pseudomodular. These matroids are all transversal, and can be represented as follows. Let G be a bipartite graph, and assume that one of its color classes S has r elements (the other may be finite of infinite). Also assume that for each $s \in S$, the other color class T contains an element which is connected only to s. Then the transversal matroid on T induced by G (in which a subset $T' \subseteqq T$ is independent iff G contains a matching covering T') is pseudomodular.

On the other hand, not every transversal matroid is pseudomodular: let $S = \{1, 2, 3, 4\}$, $T = \{a, b, c, d, e, f\}$, $V(G) = S \cup T$, and $E(G) = \{2a, 3b, 4c, 1d, 2d, 1e, 3e, 1f, 4f\}$. Then the transversal matroid induced by G on T is not pseudomodular (the flats *abde*, *acdf* and *bcef* violate condition (iv) of Theorem 1.4).

The fact that partition matroids are pseudomodular can be restated so that the Dilworth truncation of a Boolean algebra is pseudomodular. It is an interesting problem to find a broader class of lattices whose Dilworth truncations are pseudomodular.

4. Stretch embeddings and continuous matroids

We prove here the key theorem which will enable us to construct "stretch embeddings" and thereby continuous analogues of some classes of geometric lattices. This theorem generalizes a well-known result for modular lattices, see BIRKHOFF [1], pp. 73—74.

Theorem 4.1. *Let L be a pseudomodular lattice and $a_1, \dots, a_k$ elements of L such that $r(a_1)+\dots+r(a_k)=r(a_1\vee\dots\vee a_k)$. Then the sublattice generated by the intervals $[0, a_i]$ is isomorphic to the direct product of these intervals.*

Proof. Obviously, it suffices to consider the case $k=2$. Note that the submodularity of the lattice and the hypothesis that $r(a_1)+r(a_2)=r(a_1\vee a_2)$ imply that $r(x_1)+r(x_2)=r(x_1\vee x_2)$ for all $x_i\leqq a_i$.

Let L' be the sublattice generated by the intervals $[0, a_i]$. Define the mapping $\varphi(x_1, x_2)=x_1\vee x_2$. It is easy to see that this is an injection of $[0, a_1]\times[0, a_2]$ into L', and that this injection preserves joins. We will show that it also preserves meets. This will then also imply that the mapping is bijective.

Let $x_i, y_i\in[0, a_i]$ and set $p=(x_1\vee x_2)\wedge(y_1\vee y_2)$, $q=(x_1\wedge y_1)\vee(x_2\vee y_2)$. We want to show that $p=q$. It is obvious that $p\geqq q$. To show that equality holds, we show that p and q have the same rank. Clearly, $r(q)=r(x_1\wedge y_1)+r(x_2\wedge y_2)$.

To estimate $r(p)$, let $a=x_1$, $b=y_1$ and $c=x_2\vee y_2\vee p$ in Theorem 1.4 (ii). Then trivially $a\vee b\vee c=x_1\wedge y_1\vee x_2\vee y_2$ and hence

$$r(a\vee b\vee c) = r(x_1\vee y_1)+r(x_2\vee y_2).$$

Similarly we can compute that

$$r(a\vee b) = r(x_1\vee y_1),\quad r(a\vee c) = r(x_1)+r(x_2\vee y_2),\quad r(b\vee c) = r(y_1)+r(x_2\vee y_2).$$

This shows that a, b and c satisfy the conditions in Theorem 1.4 (ii), and hence by the pseudomodularity of L, we have

$$r(c) \leqq r(a\wedge b)+r(a\vee c)-r(a),$$

or, substituting,

$$r(x_2\vee y_2\vee p) \leqq r(x_1\wedge y_1)+r(x_2\vee y_2).$$

Interchanging the subscripts, we obtain

$$r(x_1\vee y_1\vee p) \leqq r(x_2\wedge y_2)+r(x_1\vee y_1).$$

Hence by submodularity,

$$r(p) \leqq r(p\vee x_1\vee y_1)+r(p\vee x_2\vee y_2)-r(x_1\vee y_1\vee x_2\vee y_2) \leqq$$
$$\leqq r(x_1\wedge y_1)+r(x_2\wedge y_2) = r(q).$$

This proves the theorem.

It takes a little time to see that this theorem does not hold automatically in every semimodular or geometric lattice. Let Σ_1 and Σ_2 be two disjoint planes in a rank 6 projective space, and let e_i and f_i be two lines in Σ_i. Construct a matroid by deleting the intersection point of e_1 and f_1 as well as the intersection point of e_2 and f_2 from the space. Then in the lattice of flats of this matroid, $e_1\wedge f_1=e_2\wedge f_2=0$, but $(e_1\vee e_2)\wedge(f_1\vee f_2)\neq 0$. This shows that at least the trivial mapping φ used in the proof above does not work. In fact, it is easy to see that this gives a counterexample.

The previous theorem enables us to construct "stretch embeddings" for various classes of matroids. Let $L_1, L_2, \ldots$ be a sequence of pseudomodular geometric lattices such that L_n has height n. Assume that for each $n, m\geqq 1$ such that $m|n$, there exist in L_n n/m elements $a_1, \ldots, a_{n/m}$ of rank m such that $a_1\vee\ldots\vee a_{n/m}=1$ and $[0, a_i]\cong L_m$. We call these elements the *representatives* of L_m in L_n.

It is now easy to define a *stretch embedding* of L_m in L_n, i.e., a lattice embedding $\varphi=\varphi^n_m$: $L_m\to L_n$ such that $r(\varphi(x))=(n/m)r(x)$ for each $x\in L_m$. For, let φ_i: $L_m\to[0, a_i]$ $(i=1, \ldots, n/m)$ be any isomorphism, and define $\varphi(x)=\varphi_1(x)\vee\ldots\vee\varphi_{n/m}(x)$. Theorem 4.1 implies that this is indeed a stretch embedding.

In the paper of BJÖRNER [2], a similar construction was described under the hypothesis that the elements $a_1, \ldots, a_{n/m}$ are modular. Since we assume the existence of pseudointersections for all pairs of elements, the construction in this paper is neither stronger nor weaker than that.

To construct the "continuous limit" of this sequence of geometric lattices, we have to assume that the mappings φ^n_m form a directed system, i.e., if $k|m$ and $m|n$ then $\varphi^n_m\circ\varphi^m_k=\varphi^n_k$. One may assure this by compatibly choosing the representatives. This was done for the partition lattices in BJÖRNER [2]; we describe below how such a choice can be made in the special families of matroids mentioned before.

Continuous algebraic matroids. Let F be an algebraically closed field. For each $n\geqq 1$, let K_n be an algebraically closed field extension of F of transcendence degree n, and let $L_n=\mathscr{L}(F, K_n)$. Let $\{x_1, \ldots, x_n\}$ be a transcendence basis. Let A_i be the algebraically closed subfield of K_n generated by $\{x_{(i-1)m+1}, \ldots, x_{im}\}$ $(i=1, \ldots, n/m)$. Then $A_1, \ldots, A_{n/m}$ are appropriate representatives of L_m in L_n, and it is easy to check that the induced mappings form a directed system.

Continuous transversal matroids. Let $L_n=\mathscr{TM}(n)$ be the full transversal matroid of rank n, constructed in Section 2. Let $S=\{x_1, \dots, x_n\}$. Assume that $m|n$ and let $S_i=\{x_{(i-1)m+1}, \dots, x_{im}\}$ for $i=1, \dots, n/m$. Then $\Omega(S_i)$ is a flat in L_n and these flats can be chosen as representatives of L_m in L_n. It is straighforward to check that the induced mappings form a directed system.

Now as in BJÖRNER [2], we can construct the direct limit $L_{(\infty)}$ of the system $\{L_k, \varphi_k^m\}$ and its completion L_∞, obtaining thereby continuous algebraic and transversal matroids. The study of these objects is, however, left to another paper.

References

[1] G. BIRKHOFF, *Lattice Theory*, 3rd ed., Amer. Math. Soc. (Providence, R. I., 1967).

[2] A. BJÖRNER, Continuous partition lattice, preprint, Dept. of Math., M.I.T., 1986.

[3] A. BJÖRNER, B. KORTE and L. LOVÁSZ, Homotopy properties of greedoids, *Adv. in Appl. Math.*, **6** (1985), 447—494.

[4] A. DRESS and L. LOVÁSZ, On some combinatorial properties of algebraic matroids, *Combinatorica*, **7** (1987), 39—48.

[5] P. H. EDELMAN, Meet-distributive lattices and the antiexchange closure, *Algebra Universalis*, **10** (1980), 290—299.

[6] P. H. EDELMAN and R. E. JAMISON, The theory of convex geometries, *Geom. Dedicata*, **19** (1985), 247—270.

[7] A. W. INGLETON and R. A. MAIN, Non-algebraic matroids exist, *Bull. London Math. Soc.*, **7** (1975), 144—146.

[8] R. E. JAMISON-WALDNER, A perspective of abstract convexity: classifying alignments by varieties, in: *Convexity and Related Combinatorial Geometry* (eds. D. C. Kay and M. Breem), Marcel Dekker (New York, 1982); pp. 113—150.

[9] B. KORTE and L. LOVÁSZ, Shelling structures, convexity, and a happy end, in: *Graph Theory and Combinatorics* (ed. B. Bollobás; Proc. Cambridge Combinatorial Conference in Honour of Paul Erdős), Academic Press (London, 1984); pp. 219—232.

[10] B. KORTE and L. LOVÁSZ, Polyhedral results for antimatroids, Rep. No. 85390—OR, Inst. f. Ökon. und Oper. Res., Univ. Bonn, 1985.

[11] B. LINDSTRÖM, A simple non-algebraic matroid of rank three, *Utilitas Math.*, **25** (1984), 95—97.

[12] B. LINDSTRÖM, A desarguesian theorem for algebraic combinatorial geometries, *Combinatorica*, **5** (1985), 237—239.

[13] B. LINDSTRÖM, A generalization of the Ingleton—Main Lemma and a class of non-algebraic matroids, *Combinatorica*, to appear.

[14] L. LOVÁSZ, Selecting independent lines from a family of lines in a projective space, *Acta Sci. Math.*, **42** (1980), 121—131.

[15] J. VON NEUMANN, Continuous Geometries, *Proc. Nat. Acad. Sci. USA*, **22** (1936), 92—100.

[16] J. VON NEUMANN, Examples of Continuous Geometries, *Proc. Nat. Acad. Sci. USA*, **22** (1936), 101—108.

[17] D. J. A. WELSH, *Matroid Theory*, Academic Press (London, 1976).

(A. B.)
DEPARTMENT OF MATHEMATICS
ROYAL INSTITUTE OF TECHNOLOGY
10044 STOCKHOLM, SWEDEN

(L. L.)
DEPARTMENT OF COMPUTER SCIENCE
EÖTVÖS LORÁND UNIVERSITY
1088 BUDAPEST, HUNGARY

Acta Sci. Math., 51 (1987), 309—318

Non-Arguesian configurations in a modular lattice

ALAN DAY[1] and BJARNI JÓNSSON[2]

To the memory of András Huhn

1. Introduction. In [1] we showed that if L is non-Arguesian, then there exist, in the ideal lattice of L, elements p_α, $\alpha\in 5^{[2]}$, and q_β, $\beta\in 5^{[3]}$, that are related to each other in a manner similar to the ten points and ten lines in a non-Arguesian configuration in a projective plane. In the lattice case, however, each p_α is a point in a plane P_α, and each q_β is a line in the plane Q_β, with all of these planes being intervals in the ideal lattice of L. Actually our construction yielded thirty two intervals $I_\mu=u_\mu/z_\mu$, $\mu\subseteq 5$, and it was shown that, with at most two exceptions, these intervals are non-degenerate projective planes. The exceptional intervals, $I_\emptyset$ and I_5, were shown to be projective geometries of dimension three or less.

Our present objective is to describe in greater detail how the various intervals I_μ fit together. The notation and terminology of [1] will be in effect. A non-Arguesian perspectivity configuration (or PC), $\mathbf{d}$, will be called *prime* if $\mathbf{d}$ covers $\mathbf{d}_*$ in $\mathbf{PC}(L)$. These PC's and their associated intervals $I_\mu=u_\mu/z_\mu$, $\mu\subseteq 5$, will be the primary objects of our investigation. To simplify the notation, we write I_i for $I_{\{i\}}$, I_{ij} for $I_{\{ij\}}$, $I_{\neg i}$ for $I_{5\setminus\{i\}}$, etc.

It is easy to see that if, $\emptyset\neq\mu\prec\nu\neq 5$ ($\prec$ means "is covered by"), then the planes I_μ and I_ν are either transposes of each other (possibly equal) or else they are connected by a two dimensional gluing (either loose or tight). Much less is known about the intervals $I_\emptyset$ and I_5. In the examples that have been constructed so far, these too are non-degenerate projective planes, but we do not know if this is always the case. We do however show that, if $I_\emptyset$ is either 2 or 3 dimensional, then it is non-degenerate. By duality, the same holds for I_5.

Received August 25, 1986.

[1]) Research supported by NSERC Operating Grant A-8190.

[2]) Research supported by NSF Grant DMS 860251.

There are two further technical conditions that apply to PC's. A PC, **d**, is called *stable* if, whenever two intervals of the form I_i and I_{ij} are transposes of each other, they are equal. This supposed restriction causes no real loss of generality since we will show that, for every non-stable **d**, there exists a stable prime **e** with $\mathbf{e}<\mathbf{d}$. A PC, **d**, is called *Boolean* if the two functions $\mu\to z_\mu$ and $\mu\to u_\mu$ of 2^5 into L are both lattice homomorphisms. Clearly, if **d** is Boolean, then $L':=\bigcup\{I_\mu\colon \mu\subseteqq 5\}$ is a sublattice of L of finite length. A fundamental result states that, if **d** is both Boolean and stable, $\{I_\mu\colon \mu\subseteqq 5\}$ consists of 2^r planes where $0\leqq r\leqq 3$. In this case the length of L' is at most 9, and each simple subdirect factor of L' has length 6 or less.

Much less is known about the case when **d** is stable but not Boolean. We do show however that in this case the twenty planes, I_μ, $2\leqq|\mu|\leqq 3$, are distinct from each other and from the planes of the form I_i or $I_{\neg i}$. Hopefully this case will be broken down eventually into subcases for which reasonable descriptions can be found.

Some examples of the above cases can be found in [3].

2. The gluings. Throughout this section we work with a fixed prime PC, **d**, in a modular lattice, L.

Lemma 2.1. *For distinct* $i,j\in 5$, $z_iz_j=z_\emptyset$.

Proof. By definition, $z_\emptyset$ is the meet of all the entries in the matrix **d**. Since each diagonal entry is the meet of all entries in its row (or column), it follows that $z_\emptyset$ is the meet of the diagonal entries in **d**. For distinct $i,j,k\in 5$, we have

$$z_iz_j=(d_{*ij}d_{*ik})(d_{*ij}d_{*jk})\leqq d_{*ik}d_{*jk}=z_k.$$

Consequently $z_iz_j=z_\emptyset$.

Lemma 2.2. *For all* $\mu,\nu\subseteqq 5$,

(1) $z_\mu+z_\nu=z_{\mu\cup\nu}$, *if* $\mu\cap\nu\neq\emptyset$;
(2) $z_\mu z_\nu=z_{\mu\cap\nu}$, *if* $\mu\cup\nu\neq 5$;
(3) $u_\mu+u_\nu=u_{\mu\cup\nu}$, *if* $\mu\cap\nu\neq\emptyset$;
(4) $u_\mu u_\nu=u_{\mu\cap\nu}$, *if* $\mu\cup\nu\neq 5$.

Proof. Statement (1) and its dual (4) are, respectively, parts (2) and (1) of [1; Lemma 5.2]. It therefore suffices to prove (2). Moreover we may assume that $|\mu|\leqq|\nu|$, $\mu\cap\nu\subset\nu$, and $|\mu\cup\nu|=4$. We consider four cases:

(A) $|\mu|=2$; $|\nu|=3$; $|\mu\cap\nu|=1$. We may assume $\mu=\{i,j\}$ and $\nu=\{i,k,m\}$. Then

$$z_\mu z_\nu=d_{*ij}(d_{*ik}+d_{*im})=d_{*ii}=z_i=z_{\mu\cap\nu}.$$

(B) $|\mu|=3$; $|\nu|=3$; $|\mu\cap\nu|=2$. We may assume $\mu=\{i,j,k\}$ and $\nu=\{i,j,m\}$.

Then:

$$z_\mu z_\nu = (d_{*ij}+d_{*jk})(d_{*ij}+d_{*jm}) = d_{*ij} = z_{ij} = z_{\mu\cap\nu}.$$

(C) $|\mu|=2$; $|\nu|=2$; $|\mu\cap\nu|=0$. We may assume $\mu=\{i, j\}$ and $\nu=\{k, m\}$. Then

$$z_\mu z_\nu = d_{*ij}d_{*km} \leqq d_{*ij}(d_{*ik}+d_{*im})d_{*km}(d_{*ik}+d_{*jk}) = d_{*ii}d_{*kk} = z_i z_k = z_\emptyset,$$

by 2.1.

(D) $|\mu|=1$; $|\nu|=3$; $|\mu\cap\nu|=0$. We may assume $\mu=\{i\}$ and $\nu=\{j, k, m\}$. Then

$$z_\mu z_\nu = z_i z_{ij} z_{jkm} = z_i z_j = z_\emptyset,$$

by (A) and 2.1.

Lemma 2.3. *For $i\in 5$, the four elements, $d_{ij}u_i$ with $j\neq i$, are four points in general position in the plane I_i.*

Proof. Let i, j, k, m, n be the distinct members of 5. Then, by computing with intervals,

$$d_{ij}u_i/z_i = d_{ij}u_{ikm}/d_{ij}u_{ikm}(d_{ik}+d_{im}) \cong (d_{ij}u_{ikm}+d_{ik}+d_{im})/(d_{ik}+d_{im}) =$$

(by transposition)

$$= (d_{ij}+d_{ik}+d_{im})u_{ikm}/(d_{ik}+d_{im}) = u_{ikm}/(d_{ik}+d_{im}).$$

Now $d_{ik}+d_{im}$ is a line in I_{ikm}, and is therefore covered by u_{ikm}. Thus $z_i \prec d_{ij}u_i$ for each $j\neq i$. To see that the four points are in general position, we compute

$$(d_{ij}u_i+d_{ik}u_i)d_{im}u_i \leqq (d_{ij}+d_{ik})d_{im} = d_{ii} = z_i.$$

Theorem 2.4. *If μ and ν are non-empty proper subsets of 5 with $\mu\prec\nu$, then either*

$$z_\mu \prec u_\mu z_\nu \quad \textit{and} \quad (u_\mu+z_\nu) \prec u_\nu,$$

or

$$z_\mu = u_\mu z_\nu \quad \textit{and} \quad (u_\mu+z_\nu) = u_\nu.$$

Proof. The intervals I_μ and I_ν are of the same length and have comparable upper and lower endpoints. Consequently, $z_\mu \prec u_\mu z_\nu$ if and only if $(u_\mu+z_\nu)\prec u_\nu$, and $z_\mu=u_\mu z_\nu$ holds just in case $(u_\mu+z_\nu)=u_\nu$ is true. Therefore we need only show that for each $\mu\prec\nu$, at least one of the four conditions holds. By duality, we need only consider $|\mu|=1$ or 2.

Assume that $\mu=\{i\}$ and $\nu=\{i,j\}$. By 2.3, $z_i\prec u_i d_{ij}$ whence $u_i z_{ij}$ must equal one of those two elements. Thus $z_\mu\prec u_\mu z_\nu$ or $z_\mu=u_\mu z_\nu$.

Assume now that $\mu=\{i,j\}$ and $\nu=\{i,j,k\}$. By the Main Theorem of [1], the element $q=d_{ij}+d_{ik}$ is a line in the plane I_ν, and qu_μ is a line on the point d_{ij} in I_μ. Now $z_\mu\leqq u_\mu z_\nu\leqq qu_\mu$. This last inequality must be strict since

$$d_{ij}z_{ijk} = d_{ij}(z_{ij}+z_{ik}) = z_{ij}+d_{ij}z_{ik} = z_{ij}+z_i = z_{ij} \prec d_{ij}.$$

Therefore the length of $u_\mu z_\nu/z_\mu$ is at most 2 and one of our relations must again hold.

Lemma 2.5. *For all* $i \in 5$, $u_\emptyset z_i$ *either covers or equals* $z_\emptyset$.

Proof. For distinct i, j, k, m in 5,

$$u_\emptyset z_i = z_i u_i u_{jkm} = z_i u_{jkm}, \quad \text{and} \quad z_\emptyset = z_i z_j = z_i d_{ij}(d_{jk}+d_{jm}) = z_i(d_{jk}+d_{jm}).$$

Since $(d_{jk}+d_{jm}) \prec u_{jkm}$, the conclusion follows.

Lemma 2.6. *Any four of the five elements,* z_i, $i \in 5$, *are independent over* $z_\emptyset$.

Proof. If $i, j, k, m \in 5$ are distinct, then

$$z_i(z_j+z_k+z_m) \leqq z_i z_{jkm} = z_\emptyset.$$

Theorem 2.7. *The following conditions are equivalent:*

(1) *The five elements,* $z_i u_\emptyset$, $i \in 5$, *are points in general position in* $I_\emptyset$*;*

(2) $I_\emptyset$ *is a non-degenerate 3-space;*

(3) length $(I_\emptyset)=4$;

(4) $z_\emptyset \prec z_i u_\emptyset$, *for all* $i \in 5$.

Proof. Now [1; Theorem 5.4] gives us that length $(I_\emptyset) \leqq 4$. Thus (1)⇒(2) and (2)⇒(3) are trivial. If $z_\emptyset = z_i u_\emptyset$, for some $i \in 5$, then $I_\emptyset \cong (z_i+u_\emptyset)/z_i$, a subinterval of a length 3 lattice. Therefore (3)⇒(4). Finally if (4) holds, then the $z_i u_\emptyset$ are five points in $I_\emptyset$ by 2.5. By 2.6, any four of these points are independent. From length $(I_\emptyset) \leqq 4$, we deduce (1).

Corollary 2.8. *If the conditions of the theorem hold, then for each* $i \in 5$, I_i *transposes down onto the interval* $u_\emptyset/z_i u_\emptyset$ *and* $u_\emptyset = \sum (z_s u_\emptyset\colon s \neq i)$.

Theorem 2.9. *If* length $(I_\emptyset)=3$, *then at least two of the intervals* I_i *transpose down onto* $I_\emptyset$. *Thus in this case as well,* $I_\emptyset$ *is a non-degenerate projective space (i.e. a plane).*

Proof. Since any four of the elements $z_i u_\emptyset$ are independent, at least one out of each four must be $z_\emptyset$. Therefore at least two of the five such elements must be $z_\emptyset$. But this forces, for these i, I_i to transpose down onto $I_\emptyset$ since both intervals are the same length.

Theorem 2.10. *The duals of* 2.7, 2.8, *and* 2.9 *also hold. In particular, if* I_5 *is of length* 3 *or* 4, *it is a non-degenerate projective space.*

3. Boolean configurations. The definition of a Boolean configuration in Section 1 contains redundancies. We already know, for instance, that whenever $\mu \cap \nu \neq \emptyset$, $z_\mu + z_\nu = z_{\mu \cap \nu}$ holds for any PC. In this section we will reduce the number of conditions needed to be checked in order to show that a PC, **d**, is Boolean. As before, we assume that **d** is a prime PC in a modular lattice, L.

Recall that a subset $U\subseteq L$ is called *distributive* if it generates a distributive sublattice of L. A 3-element subset $U=\{a, b, c\}$ is distributive if either $(a+b)c= =ac+bc$ or $(a+c)(b+c)=ab+c$.

Lemma 3.1. *The following conditions on a prime* PC *are equivalent:*

(1) $z_\mu+z_\nu=z_{\mu\cup\nu}$ *for all* $\mu, \nu\subseteq\mathbf{5}$;

(2) $z_\mu+z_\nu=z_{\mu\cup\nu}$ *for some* $\mu, \nu\neq\emptyset$ *with* $\mu\cap\nu=\emptyset$, *and* $\mu\cup\nu\neq\mathbf{5}$;

(3) $\{z_{ij}, z_{ik}, z_{jk}\}$ *is distributive for all pairwise distinct* $i, j, k\in\mathbf{5}$;

(4) $\{z_{ij}, z_{ik}, z_{jk}\}$ *is distributive for some pairwise distinct* $i, j, k\in\mathbf{5}$;

(5) $\{d_{23}, d_{24}, d_{34}\}$ *is distributive.*

Proof. By 2.2, (1) is equivalent to $z_\mu+z_\nu=z_{\mu\cup\nu}$ with the added condition that μ and ν are disjoint. By noting that $z_2=d_{23}d_{24}$, $z_3=d_{23}d_{34}$, and $z_{23}=d_{23}(d_{24}+d_{34})$, (5) is equivalent to $z_\mu+z_\nu=z_{\mu\cup\nu}$ with $\mu=\{2\}$ and $\nu=\{3\}$. By using the special automorphisms of $\mathbf{PC}(L)$, we get that (5) is equivalent to $z_\mu+z_\nu=z_{\mu\cup\nu}$ with the added constraint that μ and ν are disjoint singletons. This last property and 2.2 however easily imply that $z_\mu=\sum(z_i\colon i\in\mu)$ for all $\mu\subseteq\mathbf{5}$ and this implies (1). Therefore (1) is equivalent to (5).

A priori, (1) implies (2). Conversely, assume that (2) holds with μ or ν a non-singleton. If $\mu=\{i\}$ and $\nu\supseteq\{j, k\}$, then

$$z_{ij} = z_{ij}(z_\mu+z_\nu) = z_i+z_{ij}z_\nu = z_i+z_j.$$

If $\mu=\{i, j\}$ and $\nu=\{k, m\}$, then

$$z_{ijk} = z_{ijk}(z_\mu+z_\nu) = z_{ij}+z_{ijk}z_\nu = z_{ij}+z_k.$$

Thus this case reduces to the previous one. Therefore (1)$\Leftrightarrow$(2)$\Leftrightarrow$(5).

Now for distinct $i, j, k\in\mathbf{5}$, $\{z_{ij}, z_{ik}, z_{jk}\}$ is distributive if and only if

$$z_{ij}(z_{ik}+z_{jk}) = z_{ij}z_{ik}+z_{ij}z_{jk}.$$

The left side of this equation is z_{ij}, and the right side is z_i+z_j. Thus (4) implies (2) and (1) implies (3). This completes the proof.

Lemma 3.2. *For a* PC, **d**, *the following are equivalent:*

(1) $z_\mu z_\nu=z_{\mu\cap\nu}$ *for all* $\mu, \nu\subseteq\mathbf{5}$;

(2) $z_\mu z_\nu=z_{\mu\cap\nu}$ *for some* $\mu, \nu\neq\mathbf{5}$ *with* $\mu\cup\nu=\mathbf{5}$ *and* $\mu\cap\nu\neq\emptyset$;

(3) $\{z_{ijk}, z_{ijm}, z_{ijn}\}$ *is distributive for all distinct* $i, j, k, m\in\mathbf{5}$;

(4) $\{z_{ijk}, z_{ijm}, z_{ijn}\}$ *is distributive for some distinct* $i, j, k, m\in\mathbf{5}$.

Proof. In considering (2), we may assume that $|\mu|\leqq|\nu|$. The possible values for $s=|\mu|$ and $t=|\nu|$ are therefore

$$(s, t) = (4, 4), (3, 4), (3, 3), \text{ and } (2, 4).$$

For each of these four ordered pairs, (s, t), we define:

$(\forall_{st})$ $z_\mu z_\nu = z_{\mu\cap\nu}$ for all μ, ν with $\mu\cup\nu = \mathbf{5}$, $|\mu| = s$, and $|\nu| = t$;

$(\exists_{st})$ $z_\mu z_\nu = z_{\mu\cap\nu}$ for some μ, ν with $\mu\cup\nu = \mathbf{5}$, $|\mu| = s$, and $|\nu| = t$.

We claim that (3), (4), and each of the eight statements above are equivalent to each other.

Assume that i, j, k, m, n are all distinct in $\mathbf{5}$, and consider the equation

$$(*) \qquad z_{ijkm} z_{ijkn} = z_{ijk}.$$

This can be rewritten as

$$(z_{ijk} + z_{ijm})(z_{ijk} + z_{ijn}) = z_{ijk}$$

and since $z_{ijm} z_{ijn} = z_{ij} \leqq z_{ijk}$, this is equivalent to

$$(**) \qquad \{z_{ijk}, z_{ijm}, z_{ijn}\} \text{ is distributive.}$$

Since $(*)$ is $\{i, j, k\}$-symmetric and $(**)$ is $\{k, m, n\}$-symmetric, it follows that both conditions are invariant under all symmetries of the indices. Therefore (3), (4), $(\forall_{44})$, and $(\exists_{44})$ are equivalent.

If $(\forall_{44})$ holds, then

$$z_{ijk} z_{ijmn} = z_{ijkm} z_{ijkn} z_{ijmn} = z_{ijm} z_{ijn} = z_{ij},$$

and thus $(\forall_{34})$ holds. On the other hand if $(\exists_{34})$ holds, say $z_{ijk} z_{ijmn} = z_{ij}$, then

$$z_{ijkm} z_{ijmn} = z_{im} + z_{ijk} z_{ijmn} = z_{im} + z_{ij} = z_{ijm}$$

and $(\exists_{44})$ holds. Consequently, $(\forall_{44})$ is equivalent to both $(\exists_{34})$ and $(\forall_{34})$.

If $(\forall_{34})$ holds, then

$$z_{ijk} z_{imn} = z_{ijk} z_{ikmn} z_{ijmn} = z_{ij} z_{ik} = z_{ij},$$

and thus $(\forall_{33})$ holds. On the other hand if $(\exists_{33})$ holds, say $z_{ijk} z_{imn} = z_i$, then

$$z_{ijk} z_{ijmn} = z_{ij} + z_{ijk} z_{imn} = z_{ij} + z_i = z_{ij}$$

and $(\exists_{33})$ holds. Consequently, $(\forall_{44})$ is equivalent to both $(\exists_{33})$ and $(\forall_{33})$.

A similar argument shows that each of the statements $(\exists_{24})$ and $(\forall_{24})$ is equivalent to $(\forall_{44})$. Therefore (2), (3), and (4) are equivalent.

To obtain (2) implies (1) we need only consider complementary subsets of $\mathbf{5}$. Assuming (2), we obtain

$$z_i z_{jkmn} = z_{ij} z_{ik} z_{jkmn} = z_j z_k = z_\emptyset, \quad \text{and} \quad z_{ij} z_{kmn} = z_{ijk} z_{ijm} z_{kmn} = z_k z_m = z_\emptyset.$$

Thus (2) implies (1) and the proof is complete.

Corollary 3.3. *If, for some non-empty proper subsets,* $\mu \subset \nu \subseteqq \mathbf{5}$, $z_\mu = z_\nu$, *then* $\mathbf{d}$ *is Boolean.*

Proof. The inclusion $\mu\subset\nu$ implies that $u_\mu\leqq u_\nu$, and since I_μ and I_ν are both projective planes we get equality here as well. Now let $\varkappa=\nu\setminus\mu$ and $\lambda=\mathbf{5}\setminus\varkappa=(\mathbf{5}\setminus\nu)\cup\mu$. We compute

$$z_{\mu\cup\varkappa}=z_\nu=z_\nu+z_\varkappa=z_\mu+z_\varkappa,\quad\text{and}\quad z_{\nu\cap\lambda}=z_\mu=z_\mu z_\lambda=z_\nu z_\lambda.$$

By Lemmas 3.1, 3.2 and their duals, it follows that $\mathbf{d}$ is Boolean.

The above argument works in general to produce:

Theorem 3.4. *A* PC, $\mathbf{d}$, *is Boolean if and only if, for some distinct* $i,j\in\mathbf{5}$,

$$z_i+z_j=z_{ij},\quad \textit{and}\quad z_{\neg i}+z_{\neg j}=z_{\neg ij},\quad \textit{and}\quad u_i+u_j=u_{ij},\quad \textit{and}\quad u_{\neg i}+u_{\neg j}=u_{\neg ij}.$$

4. Stable configurations. We still assume that $\mathbf{d}$ is a prime PC in a modular lattice, L.

Theorem 4.1. *Let* $\mathbf{d}$ *be prime and stable. Then* $\mathbf{d}$ *is either Boolean or satisfies*

(***) *For all* $\mu,\nu\subseteqq\mathbf{5}$, *if* $\emptyset\subset\mu\prec\nu\subset\mathbf{5}$, *then* $z_\mu\prec u_\mu z_\nu$ *and* $u_\mu+z_\nu\prec u_\nu$.

Proof. Let $\mathbf{d}$ be stable, and take $\emptyset\subset\mu\prec\nu\subset\mathbf{5}$. By 2.4 we must have I_μ transposing up to I_ν, or $z_\mu\prec z_\nu u_\mu$ and $u_\mu+z_\nu\prec u_\nu$, If the first property holds, then let $\{i\}=\nu\setminus\mu$ and take $j\in\mu$. Now

$$u_j z_{ij}=u_j z_{ij} u_\mu z_\nu=u_j z_\mu z_{ij}=z_i.$$

Since $\mathbf{d}$ is stable, this implies $I_{ij}=I_i$, and hence $\mathbf{d}$ is Boolean by 3.3.

Lemma 4.2. *Let* $\mathbf{d}$ *be a prime* PC. *For any* $x\in d_{02}/z_0$, *there exists a unique* PC, $\mathbf{e}$, *such that*

$$e_{01}=d_{01}(x+d_{12}),\quad e_{02}=x,\quad e_{12}=d_{12}(x+d_{01}),$$

and for $\{i,j\}=\{0,1\}$ *and* $k\in\{3,4\}$,

$$e_{ik}=d_{ik}(d_{jk}+e_{01}).$$

Moreover if x *is not less than or equal to* z_{02}, *then* $\mathbf{e}$ *is non-Arguesian.*

Proof. The uniqueness of $\mathbf{e}$ is obvious for, by [1; Theorem 3.2], every PC in L is completely determined by the elements listed above. Thus we are left with showing the existence. This however also follows from [1; Lemma 2.4] and the quoted theorem. If $\mathbf{e}$ were Arguesian, then $\mathbf{e}=\mathbf{e}_*$, and

$$x=e_{02}=e_{*02}\leqq d_{*02}=z_{02}.$$

Lemma 4.3. *If* $\mathbf{d}$ *is not stable, then there exists a prime* PC, $\mathbf{e}<\mathbf{d}$ *that is both stable and Boolean.*

Proof. If **d** is a prime PC that is not stable then there exists $i, j \in \mathbf{5}$ such that I_i transposes up to I_{ij} but is not equal to I_{ij}. Thus for these i and j we have

$$z_{ij}u_j = z_j, \quad z_{ij}+u_i = u_j, \quad \text{and} \quad z_i < z_j.$$

By using the special automorphisms of $\mathbf{PC}(L)$, we may assume that $i=0$, and $j=2$. Using $x=d_{02}u_0$ in the previous lemma, we obtain a non-Arguesian PC, **e**, with $\mathbf{e}<\mathbf{d}$ and $e_{02}=d_{02}u_0$. To see that **e** is prime, we note that $z_0=z_0(\mathbf{d})\leqq z_{02}(\mathbf{e})<e_{02}$ (and that $z_0 \prec e_{02}$). Therefore $z_0=z_0(\mathbf{e})=z_{02}(\mathbf{e})\prec e_{02}$.

That **e** is Boolean follows from 3.3 and the fact that $z_0(\mathbf{e})=z_{02}(\mathbf{e})$, but **e** may not be stable. What this **e** has done is replace the transpose $I_0(\mathbf{d})$ up to $I_{02}(\mathbf{d})$ with the equality $I_0(\mathbf{e})=I_{02}(\mathbf{e})$. But for all $i\in\mathbf{5}$, direct calculations show that

$$z_i(\mathbf{d}) = z_i(\mathbf{e}) \leqq z_{ij}(\mathbf{e}) \leqq z_{ij}(\mathbf{d}).$$

Therefore this **e** preserves all equalities of the form $I_i(\mathbf{d})=I_{ij}(\mathbf{d})$. This means that after finitely many steps (at most 5^2) all transpositions are replaced by equalities and the resultant PC is both Boolean and stable.

Thus if L is a non-Arguesian modular lattice, we can find, in the lattice of ideals of L, a prime (non-Arguesian) PC, **d**. If **d** is stable, then **d** is either Boolean or satisfies $(***)$. If **d** is not stable, we can find a smaller PC, **e**, that is both stable and Boolean. Therefore every non-Arguesian variety of modular lattices contains a non-Arguesian lattice with a stable (non-Arguesian) PC. The Boolean case has a nice finite solution which we present in the next section. By [3], there exists infinitely many distinct stable PC's satisfying $(***)$, and these authors at least have found no classification of them. Our only general result is the following.

Theorem 4.4. *Let* **d** *be a stable non-Boolean* PC. *Then the twenty planes,* I_μ, $2\leqq|\mu|\leqq 3$, *are distinct from each other, and from the planes,* I_i *and* $I_{\neg i}$, $i\in\mathbf{5}$.

Proof. Let $\mu\neq\nu\subseteqq\mathbf{5}$ satisfy:

$$1 \leqq |\mu|, |\nu| \leqq 4, \quad \min\{|\mu|, |\nu|\} \leqq 3, \quad \text{and} \quad \max\{|\mu|, |\nu|\} \geqq 2.$$

We wish to show that the assumption, $z_\mu=z_\nu$, leads to a contradiction. We obtain this contradiction by producing a covering pair of subsets, $\varkappa\prec\lambda$, with $z_\varkappa=z_\lambda$, and invoking $(***)$.

If $\mu\cap\nu\neq\emptyset$, then $z_\mu+z_\nu=z_{\mu\cup\nu}$, and we may choose $\varkappa$ to be the set of smallest cardinality and λ to be any cover contained in $\mu\cup\nu$. This produces our contradiction on $(***)$. Therefore we may conclude that

$$[0] \quad \mu\cap\nu = \emptyset.$$

Therefore there exists $i\in\mu\setminus\nu$. But now we have for all $i\in\mu$

$$z_{\nu\cup\{i\}} = z_{\nu\cup\{i\}}+z_\nu = z_{\nu\cup\{i\}}+z_\mu = z_{\mu\cup\nu}.$$

To avoid conflict with $(***)$, we must have:

$$i\in\mu \text{ implies } [1] \quad \mu\cup\nu=5 \text{ and } 4\leqq \nu\cup|\{i\}|, \text{ or}$$
$$[2] \quad \mu\cup\nu=\nu\cup\{i\}.$$

We also have $j\in\nu\setminus\mu$ and the trick above can be applied again to produce

$$j\in\nu \text{ implies } [3] \quad \mu\cup\nu=5 \text{ and } 4\leqq|\mu\cup\{j\}|, \text{ or}$$
$$[4] \quad \mu\cup\nu=\mu\cup\{j\}.$$

Now [0] makes [2] equivalent to $\mu=\{i\}$, and [4] equivalent to $\nu=\{j\}$. Thus [1] and [4] are incompatible as well as [2] and [3]. Our initial assumptions deny the conjunction of [2] and [4], so we must have [1] and [3]. But this forces $3\leqq|\nu|$ and $|\mu|$ which contradicts [0]. This concludes the proof.

5. Stable Boolean configurations. Throughout this section, **d** will be a prime, Boolean, and stable PC in a modular lattice, L. The lattice homomorphisms,

$$z, u\colon \mathbf{2}^5 \to L,$$

produce Boolean congruences on $\mathbf{2}^5$ which are, of course, determined by their respective ideals, $\mathrm{Id}(z)$ and $\mathrm{Id}(u)$, of subsets congruent to $\emptyset$. Now $\{i\}\in\mathrm{Id}(z)\Leftrightarrow z_i=z_\emptyset\Leftrightarrow$for all $j\neq i$, $z_{ij}=z_j\Leftrightarrow$for all $j\neq i$, $u_{ij}=u_j\Leftrightarrow u_i=u_\emptyset\Leftrightarrow\{i\}\in\mathrm{Id}(u)$. Therefore $\mathrm{Id}(z)=\mathrm{Id}(u)$, and by factoring out this ideal we produce, for some r with $0\leqq r\leqq 5$, lattice embeddings

$$z', u'\colon \mathbf{2}^{\mathbf{r}} \to L.$$

Our first result shows that this r can be further restricted.

Lemma 5.1. *If* **d** *is Boolean and stable, then the set* $\{I_\mu\colon \mu\subseteqq 5\}$ *consists of* 2^r *planes for some* r, $0\leqq r\leqq 3$.

Proof. Let i, j, k, m, n be distinct members of $\mathbf{5}$, and assume that for all $s\neq n$, $z_{sn}>z_n$. From 2.3 and stability, this implies that for all $s\neq n$, $u_n z_{sn}=u_n d_{sn}$. 2.3 also says that $\{u_n d_{sn}\colon s\neq n\}$ are points in general position in I_n. But **d** is Boolean, and therefore

$$u_i d_{in} \leqq u_i z_{in}(u_j z_{jn}+u_k z_{kn}+u_m z_{mn}) \leqq z_{in} z_{jkmn} = z_n.$$

This is a contradiction.

Thus for every $n\in\mathbf{5}$, there exists an $s\neq n$ such that $z_{sn}=z_n$. Again since **d** is Boolean this implies that for every $n\in\mathbf{5}$, there exists an $s\neq n$ such that $z_s=z_\emptyset$. Elementary counting now produces two distinct $s\in\mathbf{5}$ with $z_s=z_\emptyset$.

We may therefore replace **5** by **r** for $0\leqq r\leqq 3$, and assume that we have lattice monomorphisms,

$$z, u\colon \mathbf{2}^{\mathbf{r}} \to L,$$

that satisfy:

(1) $I_\mu := u_\mu/z_\mu$ is a non-degenerate projective plane for all $\mu \subseteq \mathbf{r}$;

(2) For all $\emptyset \subseteq \mu \prec \nu \subseteq \mathbf{r}$, $z_\mu u_\nu \prec z_\nu$ and $u_\mu + z_\nu \prec u_\nu$.

We define $L' := \bigcup \{I_\mu : \mu \subseteq \mathbf{r}\}$. Clearly L' is a sublattice of L of finite length.

Lemma 5.2. *L' is simple if and only if for all $i \in \mathbf{r}$, $z_i \leqq u_{\neg i}$.*

Proof. If our condition fails, then, for the offending $i \in \mathbf{r}$, L' is the disjoint union of the filter, $\uparrow z_i$ and the ideal, $\downarrow u_{\neg i}$. Thus L' is not simple.

Conversely, assume the condition holds. We proceed by induction on r. If $r=0$, then L' is a non-degenerate projective plane and hence simple. If $0<r\leqq 3$, take a prime quotient q/p in L', and let θ be the congruence it generates. Since $L' = \uparrow z_i \cup \downarrow u_{\neg i}$ and $z_i \leqq u_{\neg i}$, we must have this quotient in $\uparrow z_i$ or in $\downarrow u_{\neg i}$. By induction, θ collapses either the filter or the ideal. Since $z_i < u_{\neg i}$, induction applies also to the other part and θ collapses all of L'. Therefore L' is simple.

Theorem 5.3. *Suppose $\mathbf{V}$ is a variety of modular lattices and assume that there exists a Boolean, prime* PC *in some member of $\mathbf{V}$. Then there exists in $\mathbf{V}$ a simple non-Arguesian lattice of length $3+r$, with $0 \leqq r \leqq 3$, and a Boolean, stable, prime* PC, $\mathbf{d}$, *in L with the following properties:*

(1) *L is generated by $\{d_{ij} : i \neq j \text{ in } \mathbf{5}\}$;*

(2) *The set $\{I_\mu : \mu \subseteq \mathbf{5}\}$ consists of precisely 2^r planes.*

Proof. By 4.3. there exists in some member L of $\mathbf{V}$ a PC, $\mathbf{d}$, that is prime, Boolean, and stable. By 5.1, the set $\{I_\mu : \mu \subseteq \mathbf{5}\}$ consists of 2^r distinct planes for some r, with $0 \leqq r \leqq 3$. We may assume without loss of generality that L is generated by the PC and is therefore the union of the planes I_μ.

Since L is obviously of finite length, we may assume that its length is as small as possible. We claim that in this case L is simple. To see this, we consider a homomorphism $\varphi : L \to S$, where S is simple and φ does not identify d_{01} and d_{*01}. Clearly $\varphi(\mathbf{d})$ is a (non-Arguesian) PC in S and in fact also prime and Boolean (since $\varphi(z_\mu(\mathbf{d})) = z_\mu(\varphi(\mathbf{d}))$ and similarly for the u's). The length of S therefore cannot be less than that of L. This makes φ an isomorphism and L simple.

References

[1] A. Day and B. Jónsson, A structural characterization of non-Arguesian lattices, *Order*, **2** (1986), 335—350.

[2] Ch. Herrmann, *S*-verklebte Summen von Verbänden, *Math. Z.*, **130** (1973), 255—274.

[3] D. Pickering, On minimal non-Arguesian lattice varieties, Ph. D. Thesis, University of Hawaii, 1985.

(A. D.)
DEPARTMENT OF MATHEMATICAL SCIENCES
LAKEHEAD UNIVERSITY
THUNDER BAY, ONT, P7B 5E1, CANADA

(B. J.)
DEPARTMENT OF MATHEMATICS
VANDERBILT UNIVERSITY
NASHVILLE, TENNESSEE 37235, U.S.A.

Acta Sci. Math., 51 (1987), 319—327

Complete congruence relations of concept lattices

KLAUS REUTER and RUDOLF WILLE

To the memory of András Huhn

1. Introduction. Although complete lattices have been a main subject of lattice theory for a long time, complete congruence relations of complete lattices have only rarely been studied. In this paper we describe a general approach to complete congruence relations generalizing ideas introduced in [3]. In our approach we understand complete lattices as concept lattices. This enables us to establish a one-to-one correspondence between complete congruence relations and compatible saturated subcontexts of suitable contexts. The use of this correspondence is demonstrated by proving that every distributive complete lattice in which each element is the supremum of $\vee$-irreducible elements is isomorphic to the lattice of all complete congruence relations of some complete lattice. The question remains open which complete lattices are isomorphic to such lattices of complete congruence relations. Examples are given that they need not be distributive.

2. Compatible and saturated subcontexts. A *subcontext* of a context (G, M, I) is understood as a triple $(H, N, I\cap(H\times N))$ with $H\subseteqq G$ and $N\subseteqq M$; we often write (H, N) instead of $(H, N, I\cap(H\times N))$. Throughout this section, (G, M, I) will be a context and (H, N) a subcontext of (G, M, I). For $g\in G$ and $m\in M$, g' and m' stands for $\{g\}'$ and $\{m\}'$, respectively. By $\pi(H, N)(A, B):=(A\cap H, B\cap N)$ for any concept (A, B) of (G, M, I), we define a map $\pi(H, N)$ from $\mathfrak{B}(G, M, I)$ into $\mathfrak{P}(H)\times\mathfrak{P}(N)$ where, in general, $\mathfrak{P}(S)$ is the complete lattice of all subsets of a set S. The subcontext (H, N) of (G, M, I) is said to be *compatible* if the following conditions are satisfied:

(1a) For all $h\in H$ and $m\in M\setminus h'$ there exists an $n\in N\setminus h'$ with $n'\supseteqq m'$;

(1b) for all $n\in N$ and $g\in G\setminus n'$ there exists an $h\in H\setminus n'$ with $h'\supseteqq g'$.

The notion of a compatible subcontext is the same as in [3] which follows from Proposition 1.

Received January 28, 1987, and in revised form April 17, 1987.

Proposition 1. *(H, N) is compatible if and only if $\pi(H, N)$ is a complete lattice homomorphism from $\mathfrak{B}(G, M, I)$ onto $\mathfrak{B}(H, N, I\cap(H\times N))$.*

Proof. Let (H, N) be compatible. By the basic theorem in [2], it must only be shown that $(A\cap H, B\cap N)$ is a concept of $(H, N, I\cap(H\times N))$ for $(A, B)\in\mathfrak{B}(G, M, I)$. Let $h\in H\setminus A$. Then there is an $m\in B$ with $(h, m)\notin I$, i.e. $m\in M\setminus h'$. By (1a), there exists an $n\in N\setminus h'$ with $n'\supseteq m'$. Hence $n\in B\cap N$ and so $h\notin(B\cap N)'$. It follows that $A\cap H= =(B\cap N)'\cap H$ and dually that $B\cap N=(A\cap H)'\cap N$. Thus, $(A\cap H, B\cap N)$ is a concept of $(H, N, I\cap H\times N))$. Conversely, let $(A\cap H, B\cap N)\in\mathfrak{B}(H, N, I\cap(H\times N))$ for all $(A, B)\in\mathfrak{B}(G, M, I)$. Now, let $h\in H$ and $m\in M\setminus h'$. Since $(m'\cap H, m''\cap N)$ is a concept of $(H, N, I\cap(H\times N))$ and $h\notin m'$, there exists an $n\in m''\cap N$ with $(h, n)\notin I$, i.e. $n\in N\setminus h'$ and $n'\supseteq m'$. Hence (H, N) satisfies (1a). Dually we obtain (1b). Thus, (H, N) is compatible.

Let $\Theta(H, N)$ be the set of all pairs of concepts (A, B) and (C, D) of (G, M, I) such that $\pi(H, N)(A, B)=\pi(H, N)(C, D)$, i.e. $\Theta(H, N)$ is the kernel of $\pi(H, N)$. If (H, N) is compatible, Proposition 1 yields that $\Theta(H, N)$ is a complete congruence relation on $\mathfrak{B}(G, M, I)$ and that $\mathfrak{B}(H, N, I\cap(H\times N))\cong\mathfrak{B}(G, M, I)/\Theta(H, N)$. Let us recall that a *complete congruence relation* of a complete lattice L is an equivalence relation Θ on L satisfying $(\bigwedge_{j\in J} x_j)\Theta(\bigwedge_{j\in J} y_j)$ and $(\bigvee_{j\in J} x_j)\Theta(\bigvee_{j\in J} y_j)$ if $x_j\Theta y_j$ for all $j\in J$.

The question arises how to reconstruct the compatible subcontext (H, N) from the complete congruence relation $\Theta(H, N)$. By the following definition, a complete congruence relation Θ of $\mathfrak{B}(G, M, I)$ is naturally transformed into a subcontext of (G, M, I):

$$G(\Theta):=\{g\in G\mid \gamma g:=(g'', g') \text{ is the smallest element of a } \Theta\text{-class}\},$$

$$M(\Theta):=\{m\in M\mid \mu m:=(m', m'') \text{ is the greatest element of a } \Theta\text{-class}\}.$$

To obtain (H, N) from $\Theta(H, N)$ via this definition, (H, N) has to be *saturated,* i.e. (H, N) must satisfy the following conditions:

(2a) For $g\in G$ and $X\subseteq H$, $g'=X'$ implies $g\in H$;

(2b) for $m\in M$ and $Y\subseteq N$, $m'=Y'$ implies $m\in N$.

Proposition 2. *Let (H, N) be compatible. Then (H, N) is saturated if and only if $H=G(\Theta(H, N))$ and $N=M(\Theta(H, N))$.*

Proof. First we assume that (H, N) is saturated. Let $h\in H$ and $(A, B)\in\mathfrak{B}(G, M, I)$. Then $h''\cap H=A\cap H$ implies $h\in A$ and so $h''\subseteq A$; hence $h\in G(\Theta(H, N))$. Now, let $g\in G(\Theta(H, N))$. Since γg is the smallest element of a $\Theta(H, N)$-class, it follows that $\gamma g=((H\cap g'')'', (H\cap g'')')$ and therefore $g'=(H\cap g'')'$. Hence $g\in H$ by (2a).

This proves that $H=G(\mathbf{\Theta}(H, N))$ and dually $N=M(\mathbf{\Theta}(H, N))$. Let us assume these equalities for the opposite direction of the proof. Now we use that for a complete congruence relation Θ of a complete lattice L the set of the smallest elements of the Θ-classes is closed under suprema and the set of the greatest elements of the Θ-classes is closed under infima. Let $g'=X'$ for $g\in G$ and $X\subseteq H=G(\mathbf{\Theta}(H, N))$. Then, by the basic theorem in [2], $\gamma g=\bigvee_{x\in X}\gamma x$ and so $g\in G(\mathbf{\Theta}(H, N))=H$. In this way we obtain (2a) and dually (2b). Thus, (H, N) is saturated.

The next proposition clarifies the nature of the complete congruence relation $\mathbf{\Theta}(H, N)$. In the formulation we use the notation $[z]\,\Theta$ for the equivalence class of Θ represented by z.

Proposition 3. *Let Θ be a complete congruence relation of $\mathfrak{B}(G, M, I)$. Then $(G(\Theta), M(\Theta))$ is a compatible and saturated subcontext of (G, M, I) satisfying $\Theta=\mathbf{\Theta}(G(\Theta), M(\Theta))$ if and only if $\{[\gamma h]\,\Theta \mid h\in G(\Theta)\}$ is a supremum-dense and $\{[\mu n]\,\Theta \mid n\in M(\Theta)\}$ is infimum-dense in $\mathfrak{B}(G, M, I)/\Theta$.*

Proof. Assume that $(G(\Theta), M(\Theta))$ is a compatible and saturated subcontext of (G, M, I) satisfying $\Theta=\mathbf{\Theta}(G(\Theta), M(\Theta))$. By Proposition 1, $\{(h''\cap G(\Theta), h'\cap M(\Theta)) \mid h\in G(\Theta)\}$ is supremum-dense in $\mathfrak{B}(G(\Theta), M(\Theta), I\cap(G(\Theta)\times M(\Theta)))$. Since Θ is the kernel of $\pi(G(\Theta), M(\Theta))$, it follows that $\{[\gamma h]\Theta \mid h\in G(\Theta)\}$ is supremum-dense in $\mathfrak{B}(G, M, I)/\Theta$ and dually that $\{[\mu n]\Theta \mid n\in M(\Theta)\}$ is infimum-dense in $\mathfrak{B}(G, M, I)/\Theta$. Let us assume these properties for the opposite direction of the proof. First we show that $(G(\Theta), M(\Theta))$ is compatible. Let $h\in G(\Theta)$ and $m\in M\setminus h'$. Then $[\gamma h]\Theta\nleqq[\mu m]\Theta$. Since $\{[\mu n]\Theta \mid n\in M(\Theta)\}$ is infimum-dense in $\mathfrak{B}(G, M, I)$, there exists an $n\in M(\Theta)\setminus h'$ with $\mu n\geqq\mu m$, i.e. $n'\supseteq m'$. This proves (1a) and dually (1b). From the fact that the smallest and greatest elements of the Θ-classes are closed under suprema and infima, respectively, it follows that $(G(\Theta), M(\Theta))$ is saturated. Now, let $(A, B)\in\mathfrak{B}(G, M, I)$ and let (A_-, B_-) and (A^-, B^-) be the smallest and greatest concept in the Θ-class containing (A, B). Then $\gamma g\leqq(A, B)$ for $g\in G(\Theta)$ implies $\gamma g\leqq(A_-, B_-)$. Therefore $A\cap G(\Theta)=$ $=A_-\cap G(\Theta)$ and dually $B\cap M(\Theta)=B^-\cap M(\Theta)$. Hence $(A, B)\,\Theta\,(C, D)$ implies $A\cap G(\Theta)=C\cap G(\Theta)$ and $B\cap M(\Theta)=D\cap M(\Theta)$, i.e. $(A, B)\,\mathbf{\Theta}(G(\Theta), M(\Theta))\,(C, D)$. Thus, we have $\Theta\subseteq\mathbf{\Theta}(G(\Theta), M(\Theta))$. The equality follows from $(A_-, B_-)=$ $=\gamma(A\cap G(\Theta))$ and $(A^-, B^-)=\mu(B\cap M(\Theta))$.

For subcontexts (H_1, N_1) and (H_2, N_2) of (G, M, I) we define $(H_1, N_1)\leqq$ $\leqq(H_2, N_2):\Leftrightarrow H_1\subseteq H_2$ and $N_1\subseteq N_2$. The set of all compatible and saturated subcontexts of (G, M, I) together with this order relation is denoted by $\mathfrak{S}(G, M, I)$. For the complete lattice of all complete congruence relations of a complete lattice L we use the notation $\mathfrak{C}(L)$. From Propositions 1, 2, and 3 we obtain the following theorem:

Theorem 4. *Let* (G, M, I) *be a context such that, for all complete congruence relations* Θ *of* $\mathfrak{B}(G, M, I)$, $\{[\gamma h]\Theta \mid h \in G(\Theta)\}$ *is supremum-dense and* $\{[\mu n]\Theta \mid n \in M(\Theta)\}$ *is infimum-dense in* $\mathfrak{B}(G, M, I)/\Theta$. *Then* $\Theta \mapsto (G(\Theta), M(\Theta))$ *describes an anti-isomorphism from* $\mathfrak{C}(\mathfrak{B}(G, M, I))$ *onto* $\mathfrak{S}(G, M, I)$.

For the study of $\mathfrak{C}(L)$ it is interesting to find suitable contexts (G, M, I) with $L \cong \mathfrak{B}(G, M, I)$ satisfying the assumption of Theorem 4. Obviously, $(L, L, \leqq)$ will do, but it would be better to find smaller contexts. The following lemma serves us with one method recognizing such contexts. Another method is given by Lemma 7.

Lemma 5. *Let* $\{\gamma g \mid g \in G\} \cup \{(M', M)\}$ *be an order ideal and let* $\{\mu m \mid m \in M\} \cup \cup \{(G, G')\}$ *be an order filter of* $\mathfrak{B}(G, M, I)$. *Then* $\{[\gamma h]\Theta \mid h \in G(\Theta)\}$ *is supremum-dense and* $\{[\mu n]\Theta \mid n \in M(\Theta)\}$ *is infimum-dense in* $\mathfrak{B}(G, M, I)/\Theta$ *for each complete congruence relation* Θ *of* $\mathfrak{B}(G, M, I)$.

Proof. For a complete congruence relation Θ of $\mathfrak{B}(G, M, I)$ let $(A, B)_\Theta$ be the smallest concept in the Θ-class containing the concept (A, B). It can be easily seen that $(A, B) \mapsto (A, B)_\Theta$ describes a $\bigvee$-preserving map from $\mathfrak{B}(G, M, I)$ into itself. From $(A, B) = \bigvee_{g \in A} \gamma g$ we obtain $(A, B)_\Theta = \bigvee_{g \in A} (\gamma g)_\Theta$ and so $[(A, B)]\Theta = = \bigvee_{g \in A} [(\gamma g)_\Theta]\Theta$. Since $(\gamma g)_\Theta = \gamma h$ for all $g \in G$ with $(\gamma g)_\Theta \neq (M', M)$ by assumption, the first assertion follows (and dually the second).

3. Closed subcontexts. After establishing the correspondence between complete congruence relations and compatible saturated subcontexts, the question arises how to construct compatible and saturated subcontexts. In general, this question seems difficult to answer. But there is a method which can be successfully applied in special cases. This method is based on the relations $\nearrow$ and $\swarrow$ of a context (G, M, I) which have been introduced in [3] as follows $(g \in G, m \in M)$:

$$g \nearrow m :\Leftrightarrow (g, m) \notin I \text{ and } m' \text{ is maximal in } \{n' \mid n \in M \setminus g'\},$$

$$g \swarrow m :\Leftrightarrow (g, m) \notin I \text{ and } g' \text{ is maximal in } \{h' \mid h \in G \setminus m'\}.$$

It has been useful to fill in the arrows in the cross-table describing the given context. An example is shown in Figure 1. A subcontext (H, N) of (G, M, I) is called *(arrow-) closed* if $h \nearrow m$ implies $m \in N$ for $h \in H$ and $m \in M$ and if $g \swarrow n$ implies $g \in H$ for $g \in G$ and $n \in N$. For example $(\{1, 4\}, \{a, d\})$ is a closed subcontext of the context described in Figure 1. A context (G, M, I) is called *doubly founded* if for all $(g, m) \in G \times M \setminus I$ there exists $h \in G$ and $n \in M$ with $g \nearrow n$, $n' \supseteqq m'$ and $h \swarrow m$, $h' \supseteqq g'$ (cf. [4]).

Lemma 6. *A compatible subcontext of a context* (G, M, I) *for which* $g_1' = g_2'$ *implies* $g_1 = g_2$ *for* $g_1, g_2 \in G$ *and* $m_1' = m_2'$ *implies* $m_1 = m_2$ *for* $m_1, m_2 \in M$, *is closed. Conversely, a closed subcontext of a doubly founded context is compatible.*

	a	b	c	d	e
1	$\swarrow\!\!\nearrow$	×	×	×	×
2	$\nearrow$	×	$\swarrow\!\!\nearrow$	×	$\swarrow$
3	×	$\swarrow\!\!\nearrow$	$\swarrow\!\!\nearrow$	×	$\swarrow$
4		$\swarrow$	×	$\swarrow\!\!\nearrow$	$\swarrow$

Figure 1

Lemma 6 is an immediate consequence of the definitions. Let us recall that a context (G, M, I) is said to be *reduced* if $g'=X'$ implies $g\in X$ for $g\in G$, $X\subseteq G$ and if $m'=Y'$ implies $m\in Y$ for $m\in M$, $Y\subseteq M$. Observe that each subcontext of a reduced context is saturated. A complete lattice L is called *doubly founded* if, for every pair $x<y$ in L, there exists a minimal element $s\in L$ with $s\leqq y$ and $s\nleqq x$ and a maximal element $t\in L$ with $x\leqq t$ and $y\nleqq t$. Such minimal and maximal elements are just the $\bigvee$-irreducible and $\bigwedge$-irreducible elements of L, respectively, and every element of L is the supremum of $\bigvee$-irreducible elements and the infimum of $\bigwedge$-irreducible elements of L. If $J(L)$ denotes the set of all $\bigvee$-irreducible elements of L and $M(L)$ the set of all $\bigwedge$-irreducible elements of L, then $L\cong\mathfrak{B}(J(L), M(L), \leqq)$ by the basic theorem in [2], and $(J(L), M(L), \leqq)$ is a reduced context. For a doubly founded lattice L, $(J(L), M(L), \leqq)$ is a doubly founded context; but the concept lattice of a doubly founded context need not be doubly founded (take $(\mathbf{N}, \mathbf{N}, \leqq)$).

Lemma 7. *Let $L:=\mathfrak{B}(G, M, I)$ be doubly founded and let Θ be a complete congruence relation of L. Then $\{[\gamma h]\Theta \mid h\in G(\Theta)\}$ is supremum-dense and $\{[\mu n]\Theta \mid n\in M(\Theta)\}$ is infimum-dense in L/Θ.*

Proof. Suppose there is a concept (A, B) with $[\bigvee\gamma(A\cap G(\Theta))]\Theta<[(A, B)]\Theta$. Let $(\bar{A}, \bar{B})$ and $(\bar{C}, \bar{D})$ be the greatest element in $[(A, B)]\Theta$ and $[\bigvee\gamma(A\cap G(\Theta))]\Theta$, respectively. Because of $(\bar{C}, \bar{D})<(\bar{A}, \bar{B})$, there exists a minimal concept (E, F) in L with $(E, F)\leqq(\bar{A}, \bar{B})$ and $(E, F)\nleqq(\bar{C}, \bar{D})$. Since (E, F) is $\bigvee$-irreducible in L, there is a $g\in G$ with $\gamma g=(E, F)$. Moreover, (E, F) must be the smallest element of $[(E, F)]\Theta$ and so $g\in A\cap G(\Theta)$. This contradicts $\gamma g\nleqq\bigvee\gamma(A\cap G(\Theta))$. Thus, the first assertion is proved and dually the second.

Lemmas 6 and 7 together with Theorem 4 yield the following theorem:

Theorem 8. *For a doubly founded complete lattice L, $\mathfrak{C}(L)$ is antiisomorphic to the complete lattice of all closed subcontexts of $(J(L), M(L), \leqq)$.*

Notice that the supremum and infimum of closed subcontexts (H_k, N_k) with $k\in K$ are just given by $\left(\bigcup_{k\in K} H_k, \bigcup_{k\in K} N_k\right)$ and $\left(\bigcap_{k\in K} H_k, \bigcap_{k\in K} N_k\right)$.

Corollary. *For a doubly founded complete lattice L, $\mathfrak{C}(L)$ is completely distributive.*

4. Lattices of complete congruence relations. It is a challenging problem to determine the class of all complete lattices which are isomorphic to some $\mathfrak{C}(L)$. Up to now, no complete lattice is known which does not belong to this class. As a positive result we prove that every distributive complete lattice with enough $\vee$-irreducible elements is isomorphic to some $\mathfrak{C}(L)$.

Theorem 9. *Let D be a distributive complete lattice in which each element is a supremum of $\vee$-irreducible elements. Then there exists a complete lattice L with $D \cong \mathfrak{C}(L)$.*

Proof. Let $J(D)$ be the set of all $\vee$-irreducible elements of D (notice that $0 \notin J(D)$). The following construction of a context (G, M, I) was stimulated by an (unpublished) idea of E. T. Schmidt:

$$G := J(D)\times\{1, 2, 3\}, \quad M := J(D)\times\{4, 6\}\cup D\times\{5\},$$

and

$$I := (J(D)\times\{1\})\times(J(D)\times\{4\})\cup(J(D)\times\{2\})\times(D\times\{5\})\cup(J(D)\times\{3\})\times(J(D)\times\{6\})\cup$$

$$\cup\{((s_1, i), (s_2, j)) \mid s_1, s_2\in J(D),\ s_1 \neq s_2,\ (i, j)\in\{(2, 4), (3, 4), (1, 6), (2, 6)\}\}\cup$$

$$\cup\{((s, i), (x, 5)) \mid s\in J(D),\ x\in D,\ s \leqq x,\ i\in\{1, 3\}\}.$$

For a concept (A, B) of the context $(J(D), D, \leqq)$ we define

$$\varrho(A, B) := (\bar{A}\times\{1, 2, 3\},\ \bar{A}\times\{4, 6\}\cup B\times\{5\})$$

where $\bar{A} := J(D)\setminus A$. We shall show that ϱ is an antiisomorphism from $\mathfrak{B}(J(D), D, \leqq)$ onto $\mathfrak{S}(G, M, I)$ which leads to $D \cong \mathfrak{C}(\mathfrak{B}(G, M, I))$ using Theorem 4 and the fact that $D \cong \mathfrak{B}(J(D), D, \leqq)$. Figure 2 visualizes the foregoing definitions.

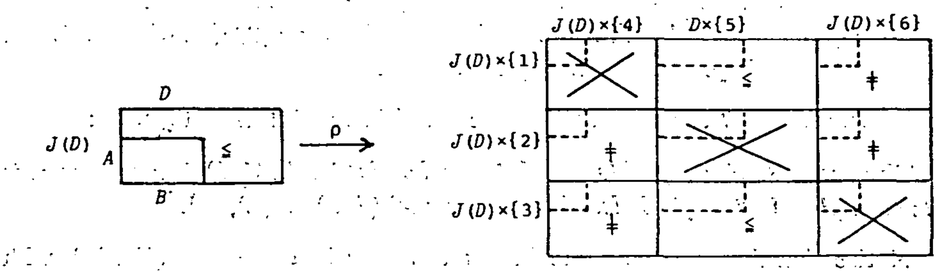


Figure 2

Let (A, B) be a concept of $(J(D), D, \leqq)$. First we show that the subcontext $(\bar{A}, B)$ of $(J(D), D, \leqq)$ satisfies the conditions (1a) and (2b). Let $h \in \bar{A}$ and $m \in D \setminus h'$. Then $\gamma h \not\leqq (A, B)$ and $\gamma h \not\leqq \mu m$. Since γh is $\vee$-irreducible and since $\mathfrak{B}(J(D), D, \leqq)$ is distributive, it follows that $\gamma h \not\leqq (A, B) \vee \mu m$. Hence there exists an $n \in B \setminus h'$ with $n' \supseteq m'$; this proves (1a). Let $m \in D$ and $Y \subseteq B$ with $m' = Y'$. As $A \subseteq Y'$ we get $m \in B$ and so (2b). Now we shall verify the conditions (1b), (2a), (1a), and (2b) for the subcontext $\varrho(A, B)$ of (G, M, I). Define $(H, N) := \varrho(A, B)$. For $g \in G$ and $m \in M$ we write g^I and m^I instead of g' and m', respectively, to avoid confusion; the prime symbol is used in this proof with respect to the context $(J(D), D, \leqq)$. Because of $G \setminus n^I \subseteq H$ for all $n \in N$, (H, N) satisfies (1b). (2a) follows from the fact that $g_1^I \supseteq g_2^I$ implies $g_1 = g_2$ for all $g_1, g_2 \in G$. For $h \in H$ we have $M \setminus h^I \subseteq N \cup D \times \{5\}$. Therefore (1a) holds because $(\bar{A}, B)$ satisfies (1a) in $(J(D), D, \leqq)$ as shown above. Since $m_1^I \supset m_2^I$ for $m_1, m_2 \in M$ and $m_1 \neq (1_D, 5)$ implies $m_1, m_2 \in D \times \{5\}$, (H, N) satisfies (2b) because this condition holds for $(\bar{A}, B)$ in $(J(D), D, \leqq)$. Thus, (H, N) is a compatible and saturated subcontext of (G, M, I).

Now we shall show that a compatible and saturated subcontext of (G, M, I) equals $\varrho(A, B)$ for some $(A, B) \in \mathfrak{B}(J(D), D, \leqq)$. It can be easily seen that $g \swarrow m$ for all $(g, m) \in G \times M \setminus I$ and $g \nearrow m$ for all $(g, m) \in G \times (J(D) \times \{4, 6\}) \setminus I$. By Lemma 6, a compatible subcontext of (G, M, I) must be of the form $(C \times \{1, 2, 3\}, C \times \{4, 6\} \cup B \times \{5\})$ with $C \subseteq J(D)$ and $B \subseteq D$; in addition, $s \leqq x$ has to be hold for all $s \in \bar{C} := J(D) \setminus C$ and $x \in B$. It remains to show that $(\bar{C}, B)$ is a concept of $(J(D), D, \leqq)$ if $(C \times \{1, 2, 3\}, C \times \{4, 6\} \cup B \times \{5\})$ is a compatible and saturated subcontext of (G, M, I). Suppose there is an $s \in C$ with $s \in B'$. Because of $s' \neq D$, we can choose an $x \in D \setminus s'$. By (1a), there exists a $(y, i) \in C \times \{4, 6\} \cup B \times \{5\} \setminus (s, 1)^I$ with $(y, i)^I \supseteq (x, 5)^I$. This implies $y \in B$ which contradicts $s \in B'$. Thus, $\bar{C} = B'$ is shown. Let $x \in D$ with $\bar{C} \subseteq x'$. For each $g \in G \setminus (x, 5)^I$ we have $g \in C \times \{1, 2, 3\}$ and $(x, 5) \in M \setminus g^I$. Hence, by (1a), there exists an $\alpha g \in (C \times \{4, 6\} \cup B \times \{5\}) \setminus g^I$ with $(\alpha g)^I \supseteq (x, 5)^I$. It follows that $\alpha g \in B \times \{5\}$ and $(x, 5)^I = (\alpha(G \setminus (x, 5)^I))^I$. Now (2b) yields $x \in B$ and therefore $B = \bar{C}'$.

Since $(A_1, B_1) \leqq (A_2, B_2) \Leftrightarrow \varrho(A_1, B_1) \geqq \varrho(A_2, B_2)$, it is shown that D is anti-isomorphic to $\mathfrak{S}(G, M, I)$. We apply Lemma 5 to see that (G, M, I) satisfies the assumption of Theorem 4. Obviously, $\{\gamma g \mid g \in G\} \cup \{(M', M)\}$ is an order ideal of $\mathfrak{B}(G, M, I)$. Let $\mu m < (A, B)$ for $m \in M$ and $(A, B) \in \mathfrak{B}(G, M, I) \setminus \{(G, G')\}$. Then $B = \tilde{B} \times \{5\}$ and so $\mu(\wedge \tilde{B}, 5) = (A, B)$. Hence $\{\mu m \mid m \in M\} \cup \{(G, G')\}$ is an order filter of $\mathfrak{B}(G, M, I)$. Finally, Theorem 4 yields $D \cong \mathfrak{C}(\mathfrak{B}(G, M, I))$.

The assumptions of Theorem 9 are fulfilled by distributive dually continuous lattices [1; p. 69] and, in particular, completely distributive complete lattices [1; p. 58]. Since the construction in the proof yields a finite context for a finite lattice D, the assumption of distributivity is unavoidable for this kind of construction. Nevertheless, there are non-distributive lattices $\mathfrak{C}(L)$ for certain infinite complete lattices L where $\mathfrak{C}(L)$ might even be finite. This we show by two examples.

Example 1. Let $\mathbf{Z}$ be the set of all integers and let $\mathbf{E}$ and $\mathbf{O}$ be the set of all even and odd integers, respectively. We define a context (G, M, I) as follows:

$$G := \mathbf{Z}\times\{1, 2, 3\}, \quad M := \mathbf{Z}\times\{4, 5, 6\},$$
$$I := \{((x, i), (y, j)) \mid x, y\in\mathbf{Z}, x \leqq y, (i, j)\in\{(1, 4), (2, 5), (3, 6)\}\}.$$

Now, we consider the following subcontexts:

$$(H_1, N_1) := (\mathbf{Z}\times\{1\}\cup\mathbf{E}\times\{2\}\cup\mathbf{O}\times\{3\},\ \mathbf{Z}\times\{4\}\cup\mathbf{O}\times\{5\}\cup\mathbf{E}\times\{6\}),$$
$$(H_2, N_2) := (\mathbf{O}\times\{1\}\cup\mathbf{Z}\times\{2\}\cup\mathbf{E}\times\{3\},\ \mathbf{E}\times\{4\}\cup\mathbf{Z}\times\{5\}\cup\mathbf{O}\times\{6\}),$$
$$(H_3, N_3) := (\mathbf{E}\times\{1\}\cup\mathbf{O}\times\{2\}\cup\mathbf{Z}\times\{3\},\ \mathbf{O}\times\{4\}\cup\mathbf{E}\times\{5\}\cup\mathbf{Z}\times\{6\}).$$

It can be easily checked that (H_i, N_i) is a compatible and saturated subcontext of (G, M, I) for $i=1, 2, 3$; furthermore, the subcontexts $(\emptyset, \emptyset)$, (H_1, N_1), (H_2, N_2), (H_3, N_3), and (G, M) form a sublattice of $\mathfrak{S}(G, M, I)$ isomorphic to M_3. This shows that $\mathfrak{C}(\mathfrak{B}(G, M, I))$ is not distributive.

Example 2. Let L_n be the complete lattice described by Figure 3.

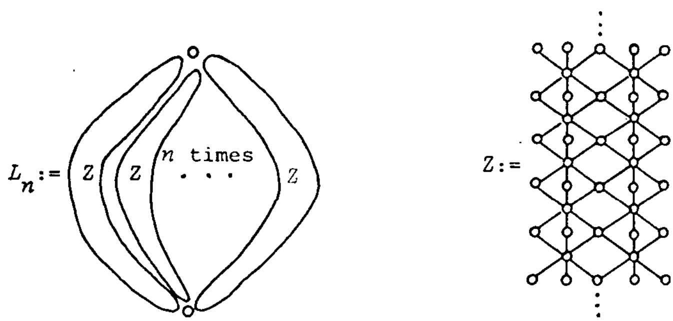
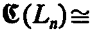


Figure 3

The lattice Z has only two non-trivial complete congruence relations. It follows that $\mathfrak{C}(L_n)\cong($$)^n\oplus 1$. For $n\geqq 2$, $\mathfrak{C}(L_n)$ is not distributive. Let us remark that $\mathfrak{C}(L_n)$ is antiisomorphic to the face lattice of an n-cube. The diagram of $\mathfrak{C}(L_2)$ is shown in Figure 4.

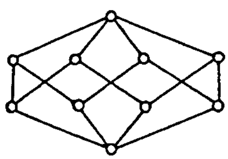


Figure 4

References

[1] G. Gierz, K. H. Hofmann, K. Keimel, J. D. Lawson, M. Mislove, D. S. Scott, *A compendium of continuous lattices*, Springer (Berlin—Heidelberg—New York, 1980).
[2] R. Wille, Restructuring lattice theory: an approach based on hierarchies of concepts, in: *Ordered sets* (ed. I. Rival), Reidel (Dordrecht—Boston, 1982); pp. 445—470.
[3] R. Wille, Subdirect decomposition of concept lattices, *Algebra Universalis*, **17** (1983), 275—287.
[4] R. Wille, Tensorial decomposition of concept lattices, *Order*, **2** (1985), 81—95.

TECHNISCHE HOCHSCHULE DARMSTADT
FACHBEREICH MATHEMATIK, ARBEITSGRUPPE ALLGEMEINE ALGEBRA
SCHLOSSGARTENSTR. 7
6100 DARMSTADT, BRD

Acta Sci. Math., 51 (1987), 329—334

Mal'cev conditions for varieties of subregular algebras

JAROMÍR DUDA

Although every congruence uniquely determines anyone of its blocks, the converse apparently does not hold in general. This trivial fact has given origin to various definitions of congruence "regularity" or "nice" congruences or "a good theory" of ideals etc. Recall from the literature that an algebra $\mathfrak{A}$ is called *regular* if every congruence on $\mathfrak{A}$ is uniquely determined by anyone of its blocks; an algebra $\mathfrak{A}$ with nullary operations $c_1, \ldots, c_n$ is called *weakly regular* (with respect to $c_1, \ldots, c_n$) whenever every congruence Ψ on $\mathfrak{A}$ is uniquely determined by its blocks $[c_1]\Psi, \ldots, [c_n]\Psi$. A natural continuation of these two concepts was introduced by J. TIMM, [12]. We write $A, B, \ldots$ for the universes of algebras $\mathfrak{A}, \mathfrak{B}, \ldots$.

Definition 1. An algebra $\mathfrak{A}$ is said to have *subregular congruences* (briefly: $\mathfrak{A}$ is *subregular*) if every congruence Ψ on $\mathfrak{A}$ is uniquely determined by its blocks $[b]\Psi$, $b \in B$, for any subalgebra $\mathfrak{B}$ of $\mathfrak{A}$.

It is already known that varieties of regular algebras and varieties of weakly regular algebras are definable by Mal'cev conditions, see [1, 2], [15] and [7] for the details. The objective of this note is to prove that also varieties of subregular algebras form Mal'cev class. We give here the explicit Mal'cev condition, see Theorem 1, since the characterizing identities enable us to prove that any variety of subregular algebras is congruence modular and n-permutable for some $n>1$. In addition we discuss the relationship between subregularity of tolerances and subregularity of congruences on algebras from a given variety. As a result of these considerations a simple Mal'cev condition for permutable varieties of subregular algebras is obtained.

Two lemmas will be needed in the sequel.

Lemma 1. *Let $\mathfrak{B}$ be a subalgebra of an algebra $\mathfrak{A}$, Ψ a congruence on $\mathfrak{A}$. The following conditions are equivalent:*

Received July 10, 1984.

(i) *Ψ is uniquely determined by its blocks* $[b]\Psi$, $b\in B$;

(ii) $\Psi=\Theta(\bigcup_{b\in B_0}(\{b\}\times A_b))$ *for some subsets* $B_0\subseteq B$ *and* $A_b\subseteq A$, $b\in B_0$.

Proof. (i)⇒(ii): If (i) holds then $\Psi=\Theta(\bigcup_{b\in B}([b]\Psi\times[b]\Psi))$. Using an evident fact that $\Theta(\bigcup_{b\in B}([b]\Psi\times[b]\Psi))=\Theta(\bigcup_{b\in B}(\{b\}\times[b]\Psi))$ the desired conclusion (ii) readily follows.

(ii)⇒(i): Conversely, suppose (ii). Then $\Psi\supseteq\bigcup_{b\in B_0}(\{b\}\times A_b)$ which gives that $[b]\Psi\times[b]\Psi\supseteq\{b\}\times A_b$ for every $b\in B_0$. By forming suitable set unions we obtain $\Psi=\bigcup_{b\in A}([b]\Psi\times[b]\Psi)\supseteq\bigcup_{b\in B}([b]\Psi\times[b]\Psi)\supseteq\bigcup_{b\in B_0}([b]\Psi\times[b]\Psi)\supseteq\bigcup_{b\in B_0}(\{b\}\times A_b)$. Hence $\Psi\supseteq\Theta(\bigcup_{b\in B}([b]\Psi\times[b]\Psi))\supseteq\Theta(\bigcup_{b\in B_0}(\{b\}\times A_b))=\Psi$, i.e. $\Psi=\Theta(\bigcup_{b\in B}([b]\Psi\times[b]\Psi))$, as required.

Remark 1. Evidently, the subsets B_0 and A_b, $b\in B_0$, from the previous lemma can be taken finite whenever Ψ is compact (=finitely generated).

H. A. Thurston has given a useful criterion for varieties of regular algebras in [13]. Lemma 2 shows that an analogue result holds for varieties of subregular algebras.

Lemma 2. *For a variety* $\mathbf{V}$ *the following conditions are equivalent:*

(i) *Every* $\mathfrak{A}\in\mathbf{V}$ *has subregular congruences;*

(ii) *a congruence on* $\mathfrak{A}\in\mathbf{V}$ *is trivial whenever it is trivial on a subalgebra* $\mathfrak{B}$ *of* $\mathfrak{A}$.

Proof. The implication (i)⇒(ii) is obvious.

(ii)⇒(i): Let $\mathfrak{B}$ be an arbitrary subalgebra of $\mathfrak{A}\in\mathbf{V}$. We have to prove that any congruence Ψ on $\mathfrak{A}$ is uniquely determined by blocks $[b]\Psi$, $b\in B$. To do this take the congruence $\Psi'=\Theta(\bigcup_{b\in B}([b]\Psi\times[b]\Psi))$ on $\mathfrak{A}$. Clearly, the subset $[B]\Psi=\bigcup_{b\in B}[b]\Psi$ is a subalgebra of $\mathfrak{A}$, moreover, the equality $[B]\mathfrak{A}=[B]\Psi'$ follows from the construction of Ψ'. Since $\Psi\supseteq\Psi'$ we can consider the congruence Ψ/Ψ' on the quotient algebra $\mathfrak{A}/\Psi'\in\mathbf{V}$. Apparently, Ψ/Ψ' is trivial on the subalgebra $[B]\Psi'/\Psi'\cap([B]\Psi'\times[B]\Psi')$ of $\mathfrak{A}/\Psi'$ hence, by hypothesis (ii), Ψ/Ψ' is trivial on the whole algebra $\mathfrak{A}/\Psi'$. In other words we have $\Psi=\Psi'$ which was to be proved.

Now we state the promised Mal'cev condition for varieties of subregular algebras (announced in [4] at first).

Theorem 1. *For a variety* $\mathbf{V}$ *the following conditions are equivalent:*

(1) *every* $\mathfrak{A}\in\mathbf{V}$ *has subregular congruences;*

(2) *there exist unary polynomials* $u_1, \ldots, u_n$, *ternary polynomials* $p_1, \ldots, p_n$ *and*

4-ary polynomials $s_1, \ldots, s_n$ *such that*

$$x = s_1(x, y, z, u_1(z))$$
$$s_i(x, y, z, p_i(x, y, z)) = s_{i+1}(x, y, z, u_{i+1}(z)), \quad 1 \leqq i < n,$$
$$y = s_n(x, y, z, p_n(x, y, z))$$
$$u_i(z) = p_i(x, x, z), \quad 1 \leqq i \leqq n,$$

hold in $\mathbf{V}$;

(3) *there exist unary polynomials* $u_1, \ldots, u_n$ *and ternary polynomials* $p_1, \ldots, p_n$ *such that*

$$(u_i(z) = p_i(x, y, z),\ 1 \leqq i \leqq n) \Leftrightarrow x = y$$

holds in $\mathbf{V}$.

Proof. (1)⇒(2): Take $\mathfrak{A}=\mathfrak{F}_{\mathbf{V}}(x, y, z)$, the free algebra in $\mathbf{V}$ on free generators x, y and z. Choose the subalgebra $\mathfrak{B}=\mathfrak{F}_{\mathbf{V}}(z)$ of $\mathfrak{A}$ and consider the principal congruence $\Theta(x, y)$ on $\mathfrak{A}$. Since $\mathfrak{A}$ is subregular, Lemma 1 (see also Remark 1) yields

$$(*) \qquad \Theta(x, y) = \Theta(\langle b_1, a_1\rangle, \ldots, \langle b_m, a_m\rangle)$$

for some elements $b_1, \ldots, b_m \in B$ and $a_1, \ldots, a_m \in A$. Applying the binary scheme, see [5, Thm. 1], to the congruence on the right hand side we get that

$$x = \sigma_1(u_1, p_1),$$
$$\sigma_i(p_i, u_i) = \sigma_{i+1}(u_{i+1}, p_{i+1}), \quad 1 \leqq i < n,$$
$$y = \sigma_n(p_n, u_n),$$

where $\sigma_1, \ldots, \sigma_n$ are binary algebraic functions over $\mathfrak{A}$ and

$$\langle u_1, p_1\rangle, \ldots, \langle u_n, p_n\rangle \in \{\langle b_1, a_1\rangle, \ldots, \langle b_m, a_m\rangle\}.$$

Using the fact that $\mathfrak{A}=\mathfrak{F}_{\mathbf{V}}(x, y, z)$ and $\mathfrak{B}=\mathfrak{F}_{\mathbf{V}}(z)$, the above equalities can be rewritten in the form

$$x = s_1(x, y, z, u_1(z), p_1(x, y, z)),$$
$$s_i(x, y, z, p_i(x, y, z), u_i(z)) = s_{i+1}(x, y, z, u_{i+1}(z), p_{i+1}(x, y, z)), \quad 1 \leqq i < n,$$
$$y = s_n(x, y, z, p_n(x, y, z), u_n(z)),$$

for some unary polynomials $u_1, \ldots, u_n$, ternary polynomials $p_1, \ldots, p_n$ and 5-ary polynomials $s_1, \ldots, s_n$ of $\mathbf{V}$. Moreover $(*)$ implies the identities $u_i(z)=p_i(x, x, z)$, $1 \leqq i \leqq n$. Now one can easily verify that the ternary polynomials $q_1, \ldots, q_n$ defined by $q_i(x, y, z)=s_i(x, z, z, p_i(y, z, z), p_i(x, y, z))$, $1 \leqq i \leqq n$, satisfy the identities

$$x = q_1(x, z, z),$$
$$q_i(x, x, z) = q_{i+1}(x, z, z), \quad 1 \leqq i < n,$$
$$z = q_n(x, x, z),$$

ensuring the $(n+1)$-permutability of $\mathbf{V}$, see [9, 10] or [8, p. 353]. Then, by [5, Thm. 2], unary scheme can be used to describe the congruence $\Theta(\langle b_1, a_1\rangle, \ldots, \langle b_m, a_m\rangle)$. In this way we obtain the same identities as above with an additional information that the polynomials $s_1, \ldots, s_n$ do not depend on the last variable. Hence we have

$$x = s_1(x, y, z, u_1(z)),$$
$$s_i(x, y, z, p_i(x, y, z)) = s_{i+1}(x, y, z, u_{i+1}(z)), \quad 1 \leqq i < n,$$
$$y = s_n(x, y, z, p_n(x, y, z)),$$
$$u_i(z) = p_i(x, x, z), \quad 1 \leqq i \leqq n,$$

as desired in (2).

The implication (2)⇒(3) is clear.

(3)⇒(1): Let $\mathfrak{B}$ be a subalgebra of $\mathfrak{A}\in\mathbf{V}$, Ψ a congruence on $\mathfrak{A}$. Assume that $[d]\Psi=\{d\}$ for every $d\in B$. Apparently $\langle u_i(c), p_i(a, b, c)\rangle=\langle p_i(a, a, c), p_i(a, b, c)\rangle\in\Psi$, $i=1, \ldots, n$, hold for any $\langle a, b\rangle\in\Psi$ and $c\in B$. Since $u_i(c)\in B$, $i=1, \ldots, n$, we have further $u_i(c)=p_i(a, b, c)$, $i=1, \ldots, n$. Then the hypothesis (3) gives $a=b$ proving the triviality of Ψ. Lemma 2 completes the proof.

Remark 2. Putting $u_1(z)=\ldots=u_n(z)=z$ $(u_1(z)=c_1, \ldots, u_n(z)=c_n$ for nullary operations $c_1, \ldots, c_n$ of $\mathbf{V})$ in Theorem 1 (2), (3) we immediately get the well-known Mal'cev conditions for varieties of regular (resp. weakly regular) algebras.

We have already proved

Corollary 1. *Any variety of subregular algebras is n-permutable for some* $n>1$.

Furthermore, the identities from Theorem 1 (2) yield

Corollary 2. *Any variety of subregular algebras is congruence modular.*

Proof. Define 4-ary polynomials $m_0, \ldots, m_{2n+1}$ by $m_0(x, y, z, w)=x$, $m_{2i-1}(x, y, z, w)=s_i(x, w, w, u_i(w))$ $(1\leqq i\leqq n)$, $m_{2i}(x, y, z, w)=s_i(x, w, w, p_i(y, z, w))$ $(1\leqq i\leqq n)$ and $m_{2n+1}(x, y, z, w)=w$. Then $m_{2i-1}(x, y, y, w)=s_i(x, w, w, u_i(w))=$ $=s_i(x, w, w, p_i(y, y, w))=m_{2i}(x, y, y, w)$ $(1\leqq i\leqq n)$, $m_0(x, x, w, w)=x=$ $=s_1(x, w, w, u_1(w))=m_1(x, x, w, w)$, $m_{2i}(x, x, w, w)=s_i(x, w, w, p_i(x, w, w))=$ $=s_{i+1}(x, w, w, u_{i+1}(w))=m_{2i+1}(x, x, w, w)$ $(1\leqq i<n)$, $m_{2n}(x, x, w, w)=$ $=s_n(x, w, w, p_n(x, w, w))=w=m_{2n+1}(x, x, w, w)$ and $m_j(x, y, y, x)=x$, $0\leqq j\leqq 2n+1$, since $x=s_1(x, x, x, u_1(x))=\ldots=s_n(x, x, x, u_n(x))$. Thus $m_0, \ldots, m_{2n+1}$ are the Day polynomials and the desired result follows from [3] (see also [8, p. 355]).

As we have already seen in Corollary 1, the subregularity of congruences implies the n-permutability of a given variety. Considering further the subregularity of tolerances or the subregularity of compatible reflexive relations something more can be stated. Simultaneously, an application of general compatible relations enable us to derive Mal'cev condition for permutable varieties of subregular algebras in a very simple form. First it will be convenient to make

Definition 2. An algebra $\mathfrak{A}$ is said to have *subregular tolerances (subregular compatible reflexive relations)* if every tolerance (compatible reflexive relation, respectively) T on $\mathfrak{A}$ is uniquely determined by subsets $\{x\in A\colon \langle b, x\rangle \in T\}$, $b\in B$, for any subalgebra $\mathfrak{B}$ of $\mathfrak{A}$.

It is a routine to paraphrase Lemma 1 for tolerance and for compatible reflexive relations. On the other hand Lemma 2 is redundant for the proof of our Theorem 2.

Theorem 2. *For a variety* $\mathbf{V}$ *the following conditions are equivalent:*

(1) *every* $\mathfrak{A}\in\mathbf{V}$ *has permutable and subregular congruences;*

(2) *every* $\mathfrak{A}\in\mathbf{V}$ *has subregular compatible reflexive relations;*

(3) *every* $\mathfrak{A}\in\mathbf{V}$ *has subregular tolerances;*

(4) *there exist unary polynomials* $u_1, \dots, u_n$, *ternary polynomials* $p_1, \dots, p_n$ *and* $(3+n)$*-ary polynomial* s *such that*

$$x = s(x, y, z, u_1(z), \dots, u_n(z))$$
$$y = s(x, y, z, p_1(x, y, z), \dots, p_n(x, y, z))$$
$$u_i(z) = p_i(x, x, z), \quad 1 \leqq i \leqq n,$$

hold in $\mathbf{V}$.

Proof. The implication (1)$\Rightarrow$(2) is a direct consequence of Werner's theorem [14]; (2)$\Rightarrow$(3) is trivial.

(3)$\Rightarrow$(4): Analogously as in the proof of Theorem 1 we take $\mathfrak{A}=\mathfrak{F}_{\mathbf{V}}(x, y, z)$, $\mathfrak{B}=\mathfrak{F}_{\mathbf{V}}(z)$ and $T(x, y)$ the smallest tolerance on $\mathfrak{A}$ containing the pair $\langle x, y\rangle$. Then the hypothesis (3) (here a modified version of Lemma 1 is used) gives

$$(**) \qquad T(x, y) = T(\langle u_1, p_1\rangle, \dots, \langle u_n, p_n\rangle)$$

for some elements $u_1, \dots, u_n\in B$ and $p_1, \dots, p_n\in A$. Consequently there is a $2n$-ary algebraic function σ over $\mathfrak{A}$ such that

$$x = \sigma(u_1, \dots, u_n, p_1, \dots, p_n),$$
$$y = \sigma(p_1, \dots, p_n, u_1, \dots, u_n).$$

Since $\mathfrak{A}=\mathfrak{F}_{\mathbf{V}}(x, y, z)$ and $\mathfrak{B}=\mathfrak{F}_{\mathbf{V}}(z)$, the above two equalities can be expressed as

$$x = s(x, y, z, u_1(z), \dots, u_n(z), p_1(x, y, z), \dots, p_n(x, y, z)),$$
$$y = s(x, y, z, p_1(x, y, z), \dots, p_n(x, y, z), u_1(z), \dots, u_n(z))$$

for some $(3+2n)$-ary polynomial s, unary polynomials $u_1, \dots, u_n$ and ternary polynomials $p_1, \dots, p_n$. Identities $u_i(z)=p_i(x, x, z)$, $1\leqq i\leqq n$, follow directly from $(**)$. From all the above identities one gets Mal'cev polynomial p by $p(x, y, z)= =s(x, z, z, p_1(y, z, z), \dots, p_n(y, z, z), p_1(x, y, z), \dots, p_n(x, y, z))$. Hence $\mathbf{V}$ is permutable and, again by [14], tolerances can be replaced by compatible reflexive relations in $(**)$. Then

$$x = s(x, y, z, u_1(z), \dots, u_n(z)),$$
$$y = s(x, y, z, p_1(x, y, z), \dots, p_n(x, y, z)),$$
$$u_i(z) = p_i(x, x, z), \quad 1 \leqq i \leqq n,$$

as required.

(4)⇒(1): Ternary polynomials $p_1, \dots, p_n$ satisfy condition (3) from Theorem 1, i.e. every algebra in **V** has subregular congruences. Permutability of **V** is ensured by Mal'cev polynomial $p(x, y, z) = s(x, z, z, p_1(y, z, z), \dots, p_n(y, z, z))$.

Remark 3. We have just proved that congruence permutability is implicit in subregularity of tolerances or in subregularity of compatible reflexive relations. The same phenomenon holds for regularity and weak regularity, see the earlier paper [6].

References

[1] B. Csákány, Characterization of regular varieties, *Acta Sci. Math.*, **31** (1970), 187—189.

[2] B. Csákány, Congruence and subalgebras, *Ann. Univ. Sci. Bqdapest. Eötvös Sect. Math.*, **18** (1975), 37—44.

[3] A. Day, A characterization of modularity for congruence lattices of algebras, *Canad. Math. Bull.*, **12** (1969), 167—173.

[4] J. Duda, Varieties of subregular algebras are definable by a Mal'cev condition (Abstract), *Comm. Math. Univ. Carolinae*, **22**, 3 (1981), 635.

[5] J. Duda, On two schemes applied to Mal'cev type theorems, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.*, **26** (1983), 39—45.

[6] J. Duda, Mal'cev conditions for regular and weakly regular subalgebras of the square, *Acta Sci. Math.*, **46** (1983), 29—34.

[7] K. Fichtner, Varieties of universal algebras with ideals, *Mat. Sbornik*, **75** (117) (1968), 445—453. (Russian)

[8] G. Grätzer, *Universal Algebra*, Second Expanded Edition, Springer-Verlag (Berlin—Heidelberg—New York, 1979).

[9] J. Hagemann, On regular and weakly regular congruences, Preprint Nr. 75, *Technische Hochschule Darmstadt* (1973).

[10] J. Hagemann and A. Mitschke, On n-permutable congruences, *Algebra Universalis*, 3 (1973), 8—12.

[11] A. I. Mal'cev, On the general theory of algebraic systems, *Mat. Sbornik*, **35** (77) (1954), 3—20. (Russian)

[12] J. Timm, On regular algebras, *Colloq. Math. Soc. János Bolyai 17. Contributions to universal algebra*, Szeged (1975), 503—514.

[13] H. A. Thurston, Derived operations and congruences, *Proc. London Math. Soc.*, (3) **8** (1958), 127—134.

[14] H. Werner, A Mal'cev condition for admissible relation, *Algebra Universalis*, 3 (1973), 263.

[15] R. Wille, *Kongruenzklassengeometrien, Lecture Notes in Mathematics*, 113, Springer-Verlag (Berlin, 1970).

KROFTOVA 21
616 00 BRNO 16
CZECHOSLOVAKIA

Acta Sci. Math., 51 (1987), 335—341

On some generalizations of Boolean algebras

J. PŁONKA

0. We shall consider only lattices and algebras of the type $\tau_0=(2, 2, 1)$ with fundamental operations $+, \cdot, '$, where $+$ and $\cdot$ are binary and $'$ is unary. Algebras of type τ_0 are often studied mainly as generalizations of Boolean algebras, e.g. pseudocomplemented lattices, Stone algebras (see [1], [3]—[6]).

In [4] we introduced the notion of a locally Boolean algebra as follows. An algebra $(A; +, \cdot, ')$ is called a locally Boolean algebra if $(A; +, \cdot)$ is a distributive lattice and there exists a congruence R of $(A; +, \cdot, ')$ such that any congruence class of R is a Boolean algebra with respect to the operations $+$, $\cdot$, and $'$ restricted to this class.

We use a similar idea in this paper. In Section 1 we introduce a special congruence $\sim$ in a lattice $\mathfrak{A}=(A; +, \cdot)$ and by means of it we construct a new algebra $\mathfrak{A}_\sim$ of type $(2, 2, 1)$. We show that all algebras $\mathfrak{A}_\sim$ form a variety (Theorem 1). In Section 2 we prove that if $\mathfrak{A}$ is distributive then it is isomorphic to a subdirect product of a Stone algebra and a distributive lattice with an additional constant operation $'$ whose value is the greatest element to this lattice.

1. Let $\mathfrak{A}=(A; +, \cdot)$ be a lattice. A congruence $\sim$ of $\mathfrak{A}$ will be called a b.u.-congruence of $\mathfrak{A}$ if it satisfies the following conditions (a)—(c):

(a) $\mathfrak{A}/\sim$ is a Boolean lattice;

(b) in any congruence class $[x]$ of $\sim$ there exists a greatest element $u([x])$;

(c) for any $x, y\in A$ we have:

$$u([x]+[y]) = u([x])+u([y]), \quad u([x]\cdot[y]) = u([x])\cdot u([y]).$$

Example 1. If $\mathfrak{A}$ is a finite chain then any congruence of it having two congruence classes is a b.u.-congruence. In fact a congruence class of a lattice must be convex.

If a lattice $\mathfrak{A}$ has a b.u.-congruence $\sim$ then we can define a new algebra $\mathfrak{A}_\sim$ of type τ_0 by putting $\mathfrak{A}_\sim=(A; +, \cdot, ')$ where the operations $+$ and $\cdot$ coincide in $\mathfrak{A}$

Received October 4, 1984.

and $\mathfrak{A}_\sim$ and the operation $'$ is defined by the formula $x'=u([x]^0)$ where $[x]^0$ is the complement of the congruence class $[x]$ in the lattice $\mathfrak{A}/\sim$.

We have

(i) any b.u.-congruence $\sim$ of a lattice $\mathfrak{A}$ is a congruence of $\mathfrak{A}_\sim$ such that $\mathfrak{A}_\sim/\sim$ is a Boolean algebra.

Lemma 1. *Any algebra* $\mathfrak{A}_\sim=(A;+,\cdot,')$ *satisfies the following system of identities:*

$$x+x=x, \quad x\cdot x=x, \tag{1}$$

$$x+y=y+x, \quad x\cdot y=y\cdot x, \tag{2}$$

$$(x+y)+z=x+(y+z), \quad (x\cdot y)\cdot z=x\cdot(y\cdot z), \tag{3}$$

$$x\cdot(x+y)=x=x+(x\cdot y), \tag{4}$$

$$((x')')'=x', \tag{5}$$

$$(x+y)'=x'\cdot y', \quad (x\cdot y)'=x'+y', \tag{6}$$

$$x+(x')'=(x')', \tag{7}$$

$$x'+(x')'=y'+(y')', \tag{8}$$

$$x'+(y'\cdot z')=(x'+y')\cdot(x'+z'). \tag{9}$$

Proof. The proof follows easily from (a)—(c). We prove for example (8). Let us denote $x''=(x')'$. Let $x\in A$. Then

$$x'+x''=u([x]^0)+u([x']^0)=u([x]^0+[x']^0)=u([x]^0+[u([x]^0)]^0)=$$
$$=u([x]^0+([x]^0)^0)=u([x]^0+[x]).$$

But the element $u([x]^0+[x])$ is the greatest element of the greatest class of $\mathfrak{A}$ so it is fixed and consequently (8) holds.

Lemma 2. *Let* $\mathfrak{A}=(A;+,\cdot,')$ *be an algebra satisfying* (1)—(9). *Then there exists a b.u.-congruence relation* $\sim$ *in the lattice* $(A;+,\cdot)$ *such that* $\mathfrak{A}$ *is identical with the algebra* $(A;+,\cdot)_\sim$.

Proof. Let us put for $x,y\in A$

$$x\sim y \Leftrightarrow x'=y'.$$

Obviously $\sim$ is an equivalence. If $a_1\sim a_2$ and $b_1\sim b_2$ then by (6) we have $(a_1+b_1)'=$ $=a_1'\cdot b_1'=a_2'\cdot b_2'=(a_2+b_2)'$, so $\sim$ satisfies the substitution law for $+$. Analogously $\sim$ satisfies the substitution law for $\cdot$ and $'$, so $\sim$ is a congruence in $\mathfrak{A}$ and consequently in $(A;+,\cdot)$.

To prove (a) it is enough to show that $\mathfrak{A}/\sim$ is a Boolean algebra. However by (5) we have $x''\sim x$ for any $x\in A$, so the identity $x''=x$ holds in $\mathfrak{A}/\sim$. By

(6), (5) and (8) we have

$$(x+x')' = (x'+x'')' = (y'+y'')' = (y+y')'$$

for any $x, y \in A$. So the identity $x+x'=y+y'$ holds in $\mathfrak{A}/\sim$. By (6) and (9) the distributive law

(10) $$x\cdot(y+z) = x\cdot y+x\cdot z$$

holds in $\mathfrak{A}/\sim$, so $\mathfrak{A}/\sim$ is a Boolean algebra.

To prove (b) we shall show that the element x'' is the greatest element in the class $[x]$. We have already shown above that $x''\sim x$ for any $x\in A$, so $x''\in[x]$. If $x\sim y$ then $x'=y'$ and $x''=y''$. Now by (7) $x''=u([x])$.

The condition (c) follows at once from (5) and (6).

Finally $u([x]^0)=u([x'])=(x')''=x'$, so the operations $'$ in $(A; +, \cdot)_\sim$ and $\mathfrak{A}$ coincide.

Let us denote by L^* the class of all algebras of the form $\mathfrak{A}_\sim$ for some lattice $\mathfrak{A}$ and a b.u.-congruence $\sim$ of $\mathfrak{A}$. By Lemmas 1 and 2 we have

Theorem 1. *The class L^* is a variety defined by the identities* (1)—(9).

Let us denote by D^* the class of all algebras $\mathfrak{A}_\sim$ where $\mathfrak{A}$ is a distributive lattice.

Corollary 1. *The class D^* is a variety defined by the identities* (1)—(8) *and* (10).

This follows from Lemmas 1 and 2.

2. Let us denote by L_1 the variety of algebras of type τ_0 satisfying (1)—(4) and the following two identities:

(11) $$x+y' = x',$$

(12) $$x' = y'.$$

We denote by D_1 the variety of algebras of type τ_0 defined by (1)—(4), (11), (12) and (10). Thus the algebras from L_1 and D_1 are lattices with unit defined by an additional operation $'$.

The construction of algebras $\mathfrak{A}_\sim$ can suggest that any algebra from L^* is isomorphic to a subdirect product of a Boolean algebra and an algebra from L_1. This however is not true even for the variety D^* as it is shown by the following example.

Example 2. Let us consider an algebra $\mathfrak{B}=(\{a, b, c\}; +, \cdot, ')$ where $(\{a, b, c\}; +, \cdot)$ is a lattice in which $a<b<c$ and $a'=c$, $b'=c'=a$. Then the equivalence relation $\sim$ with two classes $\{a\}$ and $\{b, c\}$ is a b.u.-congruence in the lattice $(\{a, b, c\}; +, \cdot)$ such that $\mathfrak{B}=(\{a, b, c\}; +, \cdot)_\sim$ (see the definition of $\sim$ in Lemma 2). However $\mathfrak{B}$ neither is a Boolean algebra nor belongs to D_1, and it is subdirectly irreducible since $\mathfrak{B}$ is a subdirectly irreducible Stone algebra (see [3]).

This example is not accidental. In fact, the next theorem shows that for algebras from D^* we always have a subdirect decomposition.

Let B_1 denote the variety of Stone algebras of type τ_0 (see [1]). We have that

(ii) the identities (1)—(8), (10) and

$$x \cdot x' = y \cdot y' \tag{13}$$

form an equational base for the variety of Stone algebras.

In fact the identity (13) together with the identities $x\cdot(x\cdot x')'=x$, $(x\cdot x')''= =x\cdot x'$, $x\cdot(x\cdot y)'=x\cdot y'$, $x'+x''=(x\cdot x')'$ form an equational base for B_1. Using subdirectly irreducible algebras from B_1 (see [3]) it is easy to check that these two systems of identities are equivalent.

For a variety V of algebras of type τ_0 we denote by $\mathrm{Id}\,(V)$ the set of all identities of type τ_0 satisfied in V. For two varieties V_1 and V_2 we denote by $V_1 \vee V_2$ the join of V_1 and V_2, and by $V_1 \otimes V_2$ the class of all algebras isomorphic to a subdirect product of two algebras $\mathfrak{A}_1$ and $\mathfrak{A}_2$ where $\mathfrak{A}_1 \in V_1$ and $\mathfrak{A}_2 \in V_2$.

Let $\mathfrak{A}=(A; +, \cdot, ')$ be an algebra of type τ_0.

Theorem 2. *The following four conditions are equivalent:*

(1°) $\mathfrak{A} \in D^*$,

(2°) $\mathfrak{A} \in B_1 \otimes D_1$,

(3°) $\mathfrak{A} \in B_1 \vee D_1$,

(4°) $\mathfrak{A}$ *satisfies the identities* (1)—(10).

To prove Theorem 2 we need some lemmas. In the next six lemmas we assume that the algebra $\mathfrak{A}=(A; +, ')$ belongs to D^*, so it satisfies (1)—(10) by Corollary 1.

Lemma 3. $\mathfrak{A}$ *satisfies the following identities:*

$$x' \cdot x'' = y' \cdot y'', \tag{14}$$

$$x \cdot x' = x \cdot x' \cdot x'', \tag{15}$$

$$(x+y)(x+y)' = xx' + yy', \tag{16}$$

$$(x \cdot y)(x \cdot y)' = xx' \cdot yy'. \tag{17}$$

Proof. By (6), (5), (3) and (8) we have $x'x''=(x''+x')'=(y''+y')'=y'y''$. By (7) and (2) we have $xx'=xx''x'=xx'x''$. By (14) we can denote by e the constant element of A with $e=x'x''$ for any $x \in A$. By (15) and (10) we have

$$(x+y)(x+y)' = (x+y)(x+y)'(x+y)'' = (x+y)\cdot e = x\cdot e + y \cdot e =$$
$$= xx'x'' + yy'y'' = xx' + yy'.$$

Finally

$$(x\cdot y)(x\cdot y)' = (x\cdot y)(x\cdot y)'(x\cdot y)'' = xye = xyee = xe\cdot ye = xx'\cdot yy'.$$

We define in $\mathfrak{A}$ two relations R_1 and R_2 by putting for $a, b\in A$

$$aR_1b \Leftrightarrow a+a'a'' = b+b'b'', \quad aR_2b \Leftrightarrow aa' = bb'.$$

Lemma 4. *The relation R_1 is a congruence in $\mathfrak{A}$.*

Proof. Obviously R_1 is an equivalence. If aR_1a_1 and bR_1b_1 then $(a+b)+(a+b)'(a+b)''=(a+b)+e=a+e+b+e=a_1+e+b_1+e=(a_1+b_1)+e=(a_1+b_1)+(a_1+b_1)'\cdot(a_1+b_1)''$. So R_1 satisfies the substitution law for $+$. To show the substitution law for $\cdot$ we use the distributivity of $+$ with respect of $\cdot$. If aR_1b then by (6) and (4) we have $(a+a'a'')'=(b+b'b'')'$, hence $a'(a'+a'')=b'(b'+b'')$, so $a'=b'$, and consequently $a'R_1b'$.

Lemma 5. *R_2 is a congruence of $\mathfrak{A}$.*

Proof. Obviously R_2 is an equivalence. The substitution law for $+$, for $\cdot$ and for $'$ follows at once from (16), (17) and (14), respectively.

Lemma 6. *$R_1\cap R_2=\omega$ where ω is the diagonal.*

Proof. If aR_1b and aR_2b then

$$\begin{aligned} a &= a+aa' = a+(a\cdot a'')\cdot a' = a+(aa')\cdot(a'a'') = a+(bb')(b'b'') = \\ &= a+bb'b'' = a+be = (a+b)(a+e) = (a+b)\cdot(a+e)(a+e) = \\ &= (a+b)(a+e)\cdot(b+e). \end{aligned}$$

Analogously, we can prove that $b=(b+a)\cdot(b+e)\cdot(a+e)$. So by (2) $a=b$.

Lemma 7. *$\mathfrak{A}/R_1$ is a Stone algebra.*

Proof. We shall show that for any $x, y\in A$ we have $(x\cdot x')R_1(y\cdot y')$. In fact

$$\begin{aligned} (xx')+(xx')'(xx')'' &= xx'+(x'+x'')x'x'' = xx'+x'x'' = \\ &= x'(x+x'') = x'x'' = e. \end{aligned}$$

Analogously $(yy')+(yy')'(yy')''=e$. So $xx'R_1yy'$. Thus the algebra $\mathfrak{A}/R_1$ satisfies (13) and by (ii) it is a Stone algebra.

Lemma 8. *$\mathfrak{A}/R_2$ belongs to D_1.*

Proof. By (11) and (12) we have to prove that for any $x\in A$ we have $(x+x')R_2x'$ and for any $x, y\in A$ we have $x'R_2y'$. In fact

$$(x+x')(x+x')' = (x+x')\cdot x'x'' = xx'x''+x'x'' = x'x'' = x'(x')'.$$

Further $x'(x')'=e=y'(y')'$.

Proof of Theorem 2. By Corollary 1 condition (1°) is equivalent to (4°). Obviously $B_1 \otimes D_1 \subset B_1 \vee D_1$, so (2°)⇒(3°). Further $B_1 \vee D_1 \subset D^*$ since each of the identities (1)—(10) belongs to $\mathrm{Id}(B_1)$ by (ii), and each of the identities (1)—(10) belongs to $\mathrm{Id}(D_1)$. So any of (1)—(10) belongs to $\mathrm{Id}(B_1) \cap \mathrm{Id}(D_1)$. Thus (3°)⇒(4°). To complete the proof it is enough to show that (1°)⇒(2°) i.e. any algebra $\mathfrak{A} \in D^*$ is isomorphic to a subdirect product of a Stone algebra and an algebra from D_1. However, this follows from Lemmas 4—8 and the decomposition theorem (see [2], Theorem 2, p. 123).

Remark 1. The distributive law (10) in Theorem 2 is an essential assumption, i.e. we cannot omit this identity in condition (4°) and substitute D^* by L^* and D_1 by L_1 in conditions (1°)—(3°).

In fact, we have the following:

(iii) the variety L^* is essentially larger than the variety $B_1 \vee L_1$.

Indeed, by Theorem 2 we have $B_1 \subset D^* \subset L^*$. Further $\mathrm{Id}(L^*) \subset \mathrm{Id}(L_1)$, as it is easy to check. So $L_1 \subset L^*$, and consequently $B_1 \vee L_1 \subset L^*$. Let us take the lattice $N_5 = (\{a, b, c, 0, 1\}; +, \cdot)$ where $0<a<b<1$, $0<c<1$, the elements a and c are incomparable and the elements b and c are incomparable. We consider in N_5 an equivalence $\sim$ with two equivalence classes $\{0, a, b\}$ and $\{c, 1\}$. Obviously $\sim$ is a b.u.-congruence in N_5 where $u(\{0, a, b\}) = b$ and $u(\{c, 1\}) = 1$. Hence the algebra $(N_5)_\sim$ belongs to L^*. However $(N_5)_\sim$ does not belong to $B_1 \vee L_1$, as the identity

(d) $$x + y \cdot y' = (x+y) \cdot (x+y')$$

belongs to $\mathrm{Id}(B_1) \cap \mathrm{Id}(L_1) = \mathrm{Id}(B_1 \vee L_1)$, while $(N_5)_\sim$ does not satisfy (d) since $a + c \cdot c' = a + c \cdot u(\{0, a, b\}) = a + c \cdot b = a$ and $(a+c) \cdot (a+c') = (a+c) \cdot (a+b) = b$.

Let B_0 denote the variety of Boolean algebras of type τ_0.

Corollary 2. *$B_0 \vee D_1 = B_0 \otimes D_1$ and $B_0 \vee D_1$ is defined by the identities* (1)—(10) *and*

(18) $$x + x' = y + y'.$$

Proof. Let us denote by K the variety of algebras of type τ_0 defined by (1)—(10) and (18). Obviously $B_0 \otimes D_1 \subset B_0 \vee D_1$ and $B_0 \vee D_1 \subset K$, since $\mathrm{Id}(K) \subset (\mathrm{Id}(B_0) \cap \mathrm{Id}(D_1))$. If $\mathfrak{A} \in K$ then $\mathfrak{A} \in D^*$, since $\mathrm{Id}(D^*) \subset \mathrm{Id}(K)$. So by Theorem 2 $\mathfrak{A}$ is isomorphic to a subdirect product of two algebras $\mathfrak{A}_1$ and $\mathfrak{A}_2$ with $\mathfrak{A}_1 \in B_1$ and $\mathfrak{A}_2 \in D_1$. But $\mathfrak{A}$ satisfies (18), so also $\mathfrak{A}_1$ does. Thus $\mathfrak{A}_1$ satisfies (1)—(10), (13) and (18), whence it is easy to show that $\mathfrak{A}_1$ is a Boolean algebra. This completes the proof.

Example 3. Let X be a set. Put $Y = \{\langle A, B \rangle : A, B \in 2^X, A \subset B\}$. We define an algebra $\mathfrak{B}_0$ of type τ_0 by putting $\mathfrak{B}_0 = (Y; +, \cdot, ')$ where $+ = \cup$, $\cdot = \cap$ and $(\langle A, B \rangle)' = \langle X \setminus A, X \rangle$.

By Corollary 2 and Theorem 2 $\mathfrak{A}$ belongs to D^* since it is a subdirect product of a Boolean algebra $\mathfrak{A}_1=(2^X; \cup, \cap, ')$ and an algebra $\mathfrak{A}_2=(2^X; \cup, \cap, ')$ where $Z'=X$ for any $Z\subset X$. If $|X|=1$ then $\mathfrak{B}_0$ has only 3 elements: $\langle\emptyset; \emptyset\rangle$, $\langle\emptyset, X\rangle$ and $\langle X, X\rangle$. But $\mathfrak{B}_0$ neither is a Stone algebra nor belongs to D_1. So $\mathfrak{B}_0$ is not a direct product of a Stone algebra and an algebra from D_1. This shows that Theorem 2 cannot be strengthed to direct product.

Remark 2. We can obtain results dual to those of this paper by assuming the existence of a least element $o([x])$ in (b), and by substituting u by o in (c). Then (7) must be substituted by $x+x''=x$, and so on.

References

[1] R. Balbes, P. Dwinger, *Distributive lattices*, Univ. Missouri Press (Columbia, 1974).

[2] G. Grätzer, *Universal Algebra*, Van Nostrand (Princeton, 1968).

[3] G. Grätzer, *General Lattice Theory*, Akademie-Verlag (Berlin, 1978).

[4] J. Płonka, On bounding congruences in some algebras having the lattice structure, in: Banach Center Publications, Vol. 9, Polish Scientific Publishers (Warsaw, 1982); pp. 203—207.

[5] H. Rasiowa, R. Sikorski, *The mathematics of metamathematics*, 3rd ed., Polish Scientific Publishers (Warsaw, 1970).

[6] T. Wesołowski, On some locally pseudocomplemented distributive lattices, *Demonstratio Mathematica*, 13 (4) (1980); pp. 907—918.

MATHEMATICAL INSTITUTE
OF THE POLISH ACADEMY OF SCIENCES
UL. KOPERNIKA 18
WROCŁAW, POLAND

Acta Sci. Math., 51 (1987), 343—348

Square subgroup of an abelian group

A. M. AGHDAM

Given an abelian group G, we call R a ring over G if the additive group $R^+=G$. In this situation we write $R=(G, *)$, where $*$ denotes the ring multiplication. The multiplication is not assumed to be associative. Every group G can be provided with a ring structure in a trivial way, by defining all products to be 0; such a ring is called a zero-ring. In general, we call a group G a *nil group* if there is no ring on G other than the zero-ring.

Suppose that H is a subgroup of G. G is called *nil modulo H* if $G*G\leqq H$ for every ring $(G, *)$ on G. It is clear that if G is nil modulo both H_1 and H_2 then G is nil modulo $H_1\cap H_2$, this suggests the following definition of the *square subgroup* $\square G$ of G:

$$\square G=\bigcap\{H\leqq G \mid G \text{ is nil modulo } H\}.$$

Clearly $\square G$ is the smallest subgroup with the property that G is nil modulo $\square G$. For the first time the square subgroup was studied in [1] by A. E. STRATTON and M. C. WEBB. The basic question about the square subgroup is whether $\dfrac{G}{\square G}$ is a nil group? If this is not true in general then under what conditions it is true and why it fails?

In this note we are investigating the square subgroup of an abelian group. We will show that the square subgroup of a torsion reduced group is equal to itself and we will prove that

$$\frac{G}{\square G}\cong\frac{D}{T}\oplus\frac{N}{\square N},\quad \square D\leqq T\leqq D,$$

where D is the maximal divisible subgroup of G and N is the reduced part of G; also, if G is a non-torsion group then

$$\frac{G}{\square G}\cong\frac{N}{\square N}.$$

Received August 29, 1983, and in revised form January 29, 1985.

By an example we will show that the square subgroup of a torsion-free group, in general, is not a direct summand of the group.

All groups considered in this paper are abelian, with addition as the group operation.

Proposition 1. *If G is cyclic (finite or infinite) then* $\square G=G$.

Proof. Let $\langle x\rangle$ be a cyclic group, define a ring on $\langle x\rangle$ by $(mx)(nx)=mnx$. In this ring x is the neutral element of $\langle x\rangle$, so, $\langle x\rangle=\langle x\rangle^2$ and hence

$$\square\langle x\rangle = \langle x\rangle.$$

Proposition 2. $A=B\oplus H$ *implies that* $\square B\leqq\square A$.

Proof. Suppose that there is a ring $(B, *)$ over B. We can define a ring $(A, \circ)$ by putting

$$(b+h)\circ(b'+h') = b*b',$$

this implies that $A\circ A=B*B$, hence $\square B\leqq\square A$.

Theorem 3 ([2], page 288). *A p-group G is a nil group if and only if it is divisible.*

Theorem 4 ([2], page 287). *A multiplication μ on a p-group A is completely determined by the values $\mu(a_i, a_j)$ with a_i, a_j running over a p-basis of A. Moreover, any choice of $\mu(a_i, a_j)\in A$ with a_i, a_j from a p-basis of A subject to the sole condition that*

$$o(\mu(a_i, a_j)) \leqq \min(o(a_i), o(a_j))$$

extends to a multiplication on A.

Lemma 5. *The reduced part of a p-group G has unbounded order if and only if any p-basic subgroup of G has unbounded order.*

Proof. Let $G=D\oplus N$, D is the maximal divisible subgroup of G. Let B be a p-basic subgroup of N. If B has bounded order then B is a direct summand of G, hence $G=D\oplus B\oplus N'$ and $\frac{G}{B}\cong D\oplus N'$; by the definition of B, N' should be divisible, a contradiction, that is $N'=0$. Consequently $N=B$ and is of bounded order. This concludes that N has unbounded order if and only if B has unbounded order.

Lemma 6. *Let G be a p-group. If the reduced part of G has unbounded order then* $\square G=G$.

Proof. Suppose that G is a p-group and the reduced part of G has unbounded order. Let $B=\bigoplus_{i\in I}\langle a_i\rangle$ be a p-basic subgroup of G. Let g be an arbitrary element of G with $o(g)=p^n$. By Lemma 5, B has unbounded order, hence there is a_K such that $o(a_K)>p^n$. In accordance with Theorem 4, a multiplication μ on G is uniquely determined if we put

$$\mu(a_i, a_j) = \begin{cases} 0 & \text{if either } i \neq K \text{ or } j \neq K, \\ g & \text{if } j = i = K; \end{cases}$$

hence $g\in \square G$, that is, $\square G=G$.

Lemma 7. *Let G be a reduced p-group, then* $\square G=G$.

Proof. If G has bounded order then $G=\bigoplus_{i\in A}\langle x_i\rangle$, and by Propositions 1, 2, $\square G=G$. If G has unbounded order then, by Lemma 6, $\square G=G$.

Theorem 8. *Let G be a reduced-torsion group, then* $\square G=G$.

Proof. $G=\bigoplus_p G_p$, G_p is a p-group. If G is reduced then G_p is reduced for all prime p. By Lemma 7 $\square G_p=G_p$. Therefore $\square G=G$.

Remark 1. Let G be a group. Let $R=(G, \eta)$ be a ring on G, then $\eta\in\mathrm{Hom}\,(G\otimes G, G)$ and $\eta(g_1\otimes g_2)=g_1g_2$, that is, $G^2=\mathrm{Im}\,\eta$, therefore

$$\square G = \langle \mathrm{Im}\,\eta \mid \eta\in\mathrm{Hom}\,(G\otimes G, G)\rangle.$$

Note. $A\otimes B$ means the tensor product of A and B.

Proposition 9. *Let G be a non-torsion group, then*

$$\langle \mathrm{Im}\,\theta \mid \theta\in\mathrm{Hom}\,(G, Z(p^\infty))\rangle = Z(p^\infty).$$

Proof. $Z(p^\infty)=\langle c_1, c_2, \dots, c_n, \dots \mid pc_1=0,\ pc_2=c_1, \dots, pc_n=c_{n-1}, \dots\rangle$. Let x be in G and the order of x be infinity, then the map $f(n)=nx$ ($n\in Z$, the set of integer numbers) defines a short exact sequence:

$$0 \to Z \xrightarrow{f} G \to M \to 0$$

which induces the short exact sequence:

$$0 \to \mathrm{Hom}\,(M, Z(p^\infty)) \to \mathrm{Hom}\,(G, Z(p^\infty)) \xrightarrow{f^*} \mathrm{Hom}\,(Z, Z(p^\infty)) \to 0,$$

the sequence being right exact because $\mathrm{Ext}\,(M, Z(p^\infty))=0$.

The definition of the map f^* is given by $f^*(\theta)=\theta f$ for all $\theta\in\mathrm{Hom}\,(G, Z(p^\infty))$. Now given $y\in Z(p^\infty)$ there is $\eta\in\mathrm{Hom}\,(Z, Z(p^\infty))$ such that $\eta(1)=y$. Since f^* is epic there is $\theta\in\mathrm{Hom}\,(G, Z(p^\infty))$ such that $f^*(\theta)=\eta$, hence $y=\eta(1)=(f^*(\theta))(1)=$ $=\theta(f(1))$, yielding the result.

Theorem 10. *Let G be a group, $G=D\oplus N$ where D is the maximal divisible subgroup of G. Then $\frac{G}{\square G}\cong\frac{D}{T}\oplus\frac{N}{\square N}$, where $\square D\leqq T\leqq D$. If G is a non-torsion group, then $\frac{G}{\square G}\cong\frac{N}{\square N}$.*

Proof. $G\otimes G\cong(D\otimes D)\oplus(D\otimes N)\oplus(N\otimes D)\oplus(N\otimes N)$. Since N is reduced and $D\otimes D$, $D\otimes N$, $N\otimes D$ are divisible, $\operatorname{Hom}(D\otimes N, N)=\operatorname{Hom}(N\otimes D, N)=$ $=\operatorname{Hom}(D\otimes D, N)=0$. Hence

$$\begin{aligned}\operatorname{Hom}(G\otimes G, G) \cong{} & \operatorname{Hom}(D\otimes D, D)\oplus\operatorname{Hom}(D\otimes N, D)\\ & \oplus\operatorname{Hom}(N\otimes D, D)\oplus\operatorname{Hom}(N\otimes N, N)\\ & \oplus\operatorname{Hom}(N\otimes N, D).\end{aligned}\tag{1}$$

So, by remark (1), $\square G=T\oplus\square N$ where $\square D\leqq T\leqq D$. This implies $\frac{G}{\square G}\cong\frac{D}{T}\oplus$ $\oplus\frac{N}{\square N}$.

Suppose that G is a non-torsion group. If the group of rational numbers is a subgroup of D, then $D=H\oplus K$, where H is a direct sum of the groups of rational numbers and K is a direct sum of quasicyclic groups. Hence $D\otimes D=H\otimes H$ is a direct sum of the groups of rational numbers.

$$\operatorname{Hom}(D\otimes D, D)\cong\operatorname{Hom}(H\otimes H, H)\oplus\operatorname{Hom}(H\otimes H, K),$$

because of $\langle\operatorname{Im}\theta|\theta\in\operatorname{Hom}(Q, Q)\rangle=Q$ (Q is the group of rational numbers), by Proposition 9 and Remark 1 $\square D=D$ and $\square G=D\oplus\square N$.

If D is a torsion group, then D is a direct sum of quasicyclic groups and N is a non-torsion group, hence $N\otimes N$ is non-torsion, too. By Proposition 9 $\langle\operatorname{Im}\eta|\eta\in\operatorname{Hom}(N\otimes N, D)\rangle=D$.

Consequently by (1) $\square G=D\oplus\square N$, this concludes that

$$\frac{G}{\square G}\cong\frac{N}{\square N}.\tag{2}$$

Remark 2. Let $G=Z(p^\infty)\oplus Z(p)$, then we have $G\otimes G\cong Z(p)\otimes Z(p)\cong Z(p)$. $\operatorname{Hom}(G\otimes G, G)\cong\operatorname{Hom}(Z(p)\otimes Z(p), Z(p^\infty))\oplus\operatorname{Hom}(Z(p)\otimes Z(p), Z(p))$. By remark (1) $\square G=\langle c_1\rangle\oplus Z(p)$. We deduce that $\square G$ is not a pure subgroup of G, consequently $\square G$ is not a direct summand of G. $\frac{G}{\square G}\cong\frac{Z(p^\infty)}{\langle c_1\rangle}$, that is (2) is not true in general when G is a torsion group. But $\frac{G}{\square G}$ is a nil group.

The following example shows that the square subgroup of a torsion-free group, in general, is not a direct summand.

Example. Let A be the subgroup of $Qx_1 \oplus Qx_2$ generated by the set

$$\left\{\frac{1}{p}x_1, \frac{1}{p^2}x_1+\frac{1}{p^5}x_2 \,\middle|\, p \text{ is running over } \pi\right\}$$

where π is the set of all prime numbers. $\frac{1}{p}x_1 \in A$ implies $x_1 = p\left(\frac{1}{p}x_1\right) \in A$, hence,

(3) $$h_p(x_1) \geqq 1 \quad \text{for all} \quad p \in \pi.$$

Suppose that $x_1 \in q^2 A$ for some prime q, then

$$x_1 = q^2 \sum_{p\in\psi}\left[\frac{\alpha_p}{p}x_1+\beta_p\left(\frac{1}{p^2}x_1+\frac{1}{p^5}x_2\right)\right],$$

where ψ is a finite set of prime numbers;

$$x_1 = q^2\left[\sum_{p\in\psi}\left(\frac{\alpha_p}{p}+\frac{\beta_p}{p^2}\right)x_1+\sum_{p\in\psi}\frac{\beta_p}{p^5}x_2\right].$$

Since $\{x_1, x_2\}$ is an independent set of A, $\sum_{p\in\psi}\frac{\beta_p}{p^5}=0$, this implies

(4) $$\beta_p \equiv 0 \pmod{p} \quad \text{for all} \quad p\in\psi.$$

We deduce $1=q^2\sum_{p\in\psi}\left(\frac{\alpha_p}{p}+\frac{\beta_p}{p^2}\right)$, this implies $q\in\psi$, so, by (4)

(5) $$\beta_q \equiv 0 \pmod{q}.$$

Let $\psi^0=\psi-\{q\}$, then

$$1 = q^2\left(\frac{\alpha_q}{q}+\frac{\beta_q}{q^2}\right)+q^2\sum_{p\in\psi^0}\left(\frac{\alpha_p}{p}+\frac{\beta_p}{p^2}\right) = q\alpha_q+\beta_q+q^2\sum_{p\in\psi^0}\left(\frac{\alpha_p}{p}+\frac{\beta_p}{p^2}\right),$$

this implies $\sum_{p\in\psi^0}\left(\frac{\alpha_p}{p}+\frac{\beta_p}{p^2}\right)$ is an integer, therefore $\beta_q\equiv 1 \pmod{q}$ a contradiction by (5). Consequently $x_1 \notin q^2A$. By (3) $h_p(x_1)=1$ for all $p\in\pi$. Hence, $t(x_1)=(1, 1, \ldots, 1, \ldots)$. Let $Z_p=\frac{1}{p^2}x_1+\frac{1}{p^5}x_2$ then $p^5Z_p=p^3x_1+x_2$, and since $h_p(x_1)=1$, $h_p(x_2)=4$ for all $p\in\pi$. Hence $t(x_2)=(4, 4, \ldots, 4, \ldots)$ and $t(x_2)>t(x_1)$. $t(x_1)$ and $t(x_2)$ are not idempotent, that is any ring $R=(A, *)$ over A satisfies $x_1^2=\alpha x_2$, $x_1x_2=x_2x_1=x_2^2=0$, α is a rational number.

Let $y=ux_1+vx_2$, $w=\beta x_1+\gamma x_2$ then, $yw=\alpha u\beta x_2$, this implies $A^2\leqq\langle x_2\rangle^*$ ($\langle x_2\rangle^*$ is a pure subgroup of A generated by x_2). Since $(A, *)$ was arbitrary,

$$\Box A \leqq \langle x_2\rangle^*. \tag{6}$$

Let $\alpha=1$, then $yw=\beta ux_2$, and by the structure of A, $\beta ux_2\in A$. Hence A is not a nil group. We claim that $A^2=\langle x_2\rangle^*$. For the proof it is enough to show that $\frac{1}{p^4}x_2\in A^2$.

Let $Z_p=\frac{1}{p^2}x_1+\frac{1}{p^5}x_2$, then $Z_p^2=\frac{1}{p^4}x_2$, so $\frac{1}{p^4}x_2\in A^2$ for all $p\in\pi$. By (6) $\Box A=\langle x_2\rangle^*$. Let $U=\{u\in Q\,|\,ux_1+vx_2\in A$ for some $v\in Q\}$. If $\Box A$ is a direct summand of A, then

$$A=\langle x_2\rangle^*\oplus B,\quad \frac{A}{\langle x_2\rangle^*}\cong B,\quad t(B)=t\left(\frac{A}{\langle x_2\rangle^*}\right)=t(U);$$

by the structure of A, $t(U)=(2, 2, \ldots, 2, \ldots)$. This implies $t(x_1)<t(B)<t(x_2)$ but this is impossible, since $r(A)=2$ ([2], page 112, Ex. 10).

Consequently $\Box A$ is not a direct summand of A.

Note. Since $\frac{A}{\Box A}$ is of rank one and its type is not idempotent, it follows that $\frac{A}{\Box A}$ is a nil group.

References

[1] A. E. Stratton, M. C. Webb, Abelian group, nil modulo a subgroup, need not have nil quotient group, *Publ. Math. Debrecen,* 27 (1980), 127—130.

[2] L. Fuchs, *Infinite Abelian Groups,* vol. 2, Academic Press (1973).

[3] A. E. Stratton, The type set of torsion-free rings of finite rank, *Comment. Math. Univ. St. Paul.,* 27 (1978), 199—211.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TABRIZ
TABRIZ, IRAN

Acta Sci. Math., 51 (1987), 349—363

Concrete characterization of partial endomorphism semigroups of graphs

V. A. MOLČANOV

Algebraic and elementary properties of partial endomorphism semigroups of graphs were studied by L. M. POPOVA [1, 2], JU. M. VAŽENIN [3] and A. M. KALMANOVIČ [4—6]. For these semigroups it is interesting to study the concrete characterization problem [7]: under which conditions is a partial transformation semigroup E equal to the partial endomorphism semigroup $E(G)$ of some graph G?

In the present paper we investigate this problem. The necessary and sufficient conditions for a partial transformation semigroup E to be equal to the partial endomorphism semigroup $E(G)$ of some graph G will be obtained in Theorem 2. We construct all kinds of such graphs in Theorem 1. At the end of the paper we apply our results to describe (in Theorem 3) graphs with equal partial endomorphisms and to investigate (in Theorem 4) the question: how are graphs determined by their partial endomorphism semigroups? Numerous other applications of Theorems 1 and 2 are briefly stated in [8].

1. Definitions, preliminary results

Let X be an arbitrary set with $|X|>1$, and let ϱ be a binary relation on X, $x\in X$, $A\subset X$. We put $X^2=X\times X$;

$$\varrho^{-1} = \{(x, y)\colon (y, x)\in\varrho\};\quad \operatorname{dom}\varrho = \{x\colon (\exists y)(x, y)\in\varrho\};$$

$$\varrho x = \{y\colon (x, y)\in\varrho\}\quad\text{and}\quad \varrho A = \{y\colon (\exists x\in A)(x, y)\in\varrho\}.$$

A one-valued binary relation $f\subset X^2$ is called a partial transformation of X (shortly p.transformation). If $x\in\operatorname{dom} f$, then fx denotes the image of x under f. A p.transformation f is called 3-bounded, if $|\operatorname{dom} f|<3$. We write $f=\begin{pmatrix} a & b \\ c & d\end{pmatrix}$, if $\operatorname{dom} f=\{a, b\}$ and $fa=c$, $fb=d$. The Cartesian power f^2 of f is the p.trans-

Received May 27, 1983.

formation of X^2 such that $f^2(x, y)=(fx, fy)$ for $x, y \in \operatorname{dom} f$. The identity transformation of a set A is denoted by Δ_A. We denote by $W(X)$ the symmetric semigroup of all p.transformations on a set X.

By a graph we mean a structure $G=(X, \varrho)$, where X is a non-empty set and $\varrho \subset X^2$. The elements of X and ϱ are called vertices and edges, respectively. The edge $(x, x) \in \varrho$ is called a loop. We denote $G^{-1}=(X, \varrho^{-1})$.

A p.transformation f of X is called a partial endomorphism (shortly p.endomorphism) of the graph G, if $f^2 \varrho \subset \varrho$, i.e. $(x, y) \in \varrho$ implies $(fx, fy) \in \varrho$ for any $x, y \in \operatorname{dom} f$. The p.endomorphisms of G form a semigroup $E(G)$ (under the composition), which is called the p.endomorphism semigroup of the graph G. $E_3(G)$ denotes the 3-bounded p.endomorphism semigroup of G.

Let E be a p.transformation semigroup on a set X. The canonical relations of E are defined by the formulas:

$$\tau \overset{\mathrm{df}}{=} \bigcup\{f\colon f \in E\}; \quad \beta \overset{\mathrm{df}}{=} \bigcup\{f^2\colon f \in E\};$$

$$A \overset{\mathrm{df}}{=} \{x \in X\colon X^2 \subset \beta^{-1}(x, x)\}; \quad B \overset{\mathrm{df}}{=} X \setminus A;$$

$$P \overset{\mathrm{df}}{=} \Delta_B \cup \{(x, y) \in X^2 \setminus \Delta_X\colon \tau x \times \tau y \subset \beta(x, y)\};$$

$$R \overset{\mathrm{df}}{=} \Delta_A \cup \{(x, y) \in X^2 \setminus \Delta_X\colon (\tau^{-1} x \times \tau^{-1} y) \setminus \Delta_X \subset \beta^{-1}(x, y)\};$$

$$Z \overset{\mathrm{df}}{=} P \cap R; \quad Q' \overset{\mathrm{df}}{=} X^2 \setminus (P \cup R)$$

and

$$Q \overset{\mathrm{df}}{=} ((\beta Q') \setminus R) \cup ((\beta^{-1} Q') \setminus P).$$

The intersections of any binary relation σ on X with the relations A^2, $(A \times B) \cup (B \times A)$ and B^2 are denoted by the same symbol but with indices 1, 2 and 3, respectively, i.e. $\sigma_1 = \sigma \cap A^2$, $\sigma_2 = \sigma \cap ((A \times B) \cup (B \times A))$ and $\sigma_3 = \sigma \cap B^2$.

We denote $\{f \in E\colon |\operatorname{dom} f| < 3\}$ by E_3.

In the following lemmas the canonical relations of p.transformation semigroups will be investigated.

Lemma 1. *If E contains all 3-bounded identity p.transformations of X, then* (i) *τ is a quasi-order*[1]) *on X,* (ii) *β is a quasi-order on X^2 and* (iii) *the canonical relations of E and E_3 are equal.*

Lemma 2. *If $Q_1 = Z_1 = \emptyset$ ($Q_3 = Z_3 = \emptyset$), then P_1 (R_3) is non-empty.*

Lemma 3. *If $\tau a = X$ for all $a \in B$ then the following conditions hold:*
(i) *if $B \neq \emptyset$ and $Z_1 \neq \emptyset$ then $R_1 = A^2$, $P_1 = A^2 \setminus \Delta_A$ and $P_2 \cup Q_2 \neq \emptyset$;*
(ii) *if $Z_2 \neq \emptyset$ then $P_2 = R_2 = (A \times B) \cup (B \times A)$;*
(iii) *if $A \neq \emptyset$ and $Z_3 \neq \emptyset$ then $P_3 = B^2$, $R_3 = B^2 \setminus \Delta_B$ and $R_2 \cup Q_2 \neq \emptyset$.*

[1]) A quasi-order on X is a reflexive and transitive binary relation on X.

The proofs of these lemmas easily follow from the definitions.

Now suppose that $E=E(G)$ for some graph $G=(X, \varrho)$. Then the canonical relations of E satisfy the following properties.

Lemma 4. *Let G be a graph such that $E(G)\neq W(X)$. Then*

(i) $X^2\setminus(\varrho\cup\varrho^{-1})\subset P$;

(ii) $\varrho\cap\varrho^{-1}\subset R$;

(iii) $\varrho\cap\Delta_X=\Delta_A$

and

(iv) $Q\subset(\varrho\setminus\varrho^{-1})\cup(\varrho^{-1}\setminus\varrho)$.

Proof. (i)—(iii) are obvious. It follows from the definition of Q' that $P\cap Q'= =R\cap Q'=\emptyset$. Hence $Q'\subset(\varrho\setminus\varrho^{-1})\cup(\varrho^{-1}\setminus\varrho)$. Let $(x, y)\in Q$. By definition there exists a $(u, v)\in Q'$ such that either $\begin{pmatrix} x & y \\ u & v \end{pmatrix}\in E$ and $(x, y)\notin P$ or $\begin{pmatrix} u & v \\ x & y \end{pmatrix}\in E$ and $(x, y)\notin R$. Then either $\begin{pmatrix} x & y \\ u & v \end{pmatrix}$ is a p.endomorphism of G and $(x, y)\in\varrho\cup\varrho^{-1}$ or $\begin{pmatrix} u & v \\ x & y \end{pmatrix}$ is a p.endomorphism of G and $(x, y)\notin\varrho\cap\varrho^{-1}$. Since $(u, v)\in(\varrho\setminus\varrho^{-1})\cup \cup(\varrho^{-1}\setminus\varrho)$, the vertices x, y are joined by one edge of G. Therefore $Q\subset(\varrho\setminus\varrho^{-1})\cup \cup(\varrho^{-1}\setminus\varrho)$, i.e. (iv) holds.

Lemma 5. *If the canonical relations of $E=E(G)$ satisfy $Z_1=\emptyset$ (resp. $Z_3=\emptyset$) then $R_1=(\varrho\cap\varrho^{-1})_1$ and $P=X^2\setminus(\varrho\cup\varrho^{-1})$ (resp. $R=\varrho\cap\varrho^{-1}$ and $P_3=(X^2\setminus(\varrho\cup\varrho^{-1}))_3$).*

Proof. Suppose that $Z_1=\emptyset$. By Lemma 2, $Q_1\neq\emptyset$ or $P_1\neq\emptyset$. If $(x, y)\in Q_1$ then, by Lemma 4, x and y are joined by one edge of G. If $(a, b)\in P_1$ then, by the definition, $\begin{pmatrix} a & b \\ b & a \end{pmatrix}\in E$. On the other hand, $(a, b)\notin R_1$, whence $\begin{pmatrix} u & v \\ a & b \end{pmatrix}\notin E$ for some $u, v\in X$. Then $(a, b)\notin\varrho\cup\varrho^{-1}$. Therefore $(x, y)\notin\varrho\cup\varrho^{-1}$, if $(x, y)\in P$, and $(x, y)\in\varrho\cap \cap\varrho^{-1}$, if $(x, y)\in R_1$. Using Lemma 4, we obtain that $R_1=(\varrho\cap\varrho^{-1})_1$ and $P=X^2\setminus(\varrho\cup\varrho^{-1})$. By the analogy the second statement of the lemma is proved.

Lemma 6. *If the canonical relations of $E=E(G)$ satisfy $B\neq\emptyset$ and $Z_2=Q=\emptyset$ then the following conditions hold:*

(i) if $Z_1\neq\emptyset$ then $R_1=(\varrho\cap\varrho^{-1})_1$ and either $P_2=(X^2\setminus(\varrho\cup\varrho^{-1}))_2$ or $P_2= =((\varrho\setminus\varrho^{-1})\cup(\varrho^{-1}\setminus\varrho))_2$;

(ii) if $Z_3\neq\emptyset$ then $P_3=(X^2\setminus(\varrho\cup\varrho^{-1}))_3$ and either $R_2=((\varrho\setminus\varrho^{-1})\cup(\varrho^{-1}\setminus\varrho))_2$ or $R_2=(\varrho\cap\varrho^{-1})_2$.

Proof. Suppose that $Q=Z_2=\emptyset$ and $B, Z_1\neq\emptyset$. Then, by Lemma 3, $P_2\neq\emptyset$. Let $(a, b)\in P\cap(A\times B)$. Hence $(a, b)\notin R$ and, by the definition, $\begin{pmatrix} x & y \\ a & b \end{pmatrix}\notin E$ for some $x\in\tau^{-1}a$ and $y\in\tau^{-1}b$. Then $(a, b)\notin\varrho\cap\varrho^{-1}$, whence $(a, b)\in(\varrho\setminus\varrho^{-1})\cup\cup(\varrho^{-1}\setminus\varrho)$ or $(a, b)\notin\varrho\cup\varrho^{-1}$. Since $\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} c & d \\ a & b \end{pmatrix}\in E$ for any $(c, d)\in P\cap(A\times B)$, we obtain by Lemma 4 that either $P_2=((\varrho\setminus\varrho^{-1})\cup(\varrho^{-1}\setminus\varrho))_2$ or $P_2=(X^2\setminus(\varrho\cup\varrho^{-1}))_2$. Moreover, $(x, y)\notin P$ and, by Lemma 4, $(x, y)\in\varrho\cup\varrho^{-1}$. Since $\begin{pmatrix} x & y \\ u & v \end{pmatrix}, \begin{pmatrix} x & y \\ v & u \end{pmatrix}\in E$ for any $(u, v)\in R_1$, we obtain that $R_1=(\varrho\cap\varrho^{-1})_1$. Thus (i) holds; (ii) can be proved in the same manner.

2. Main results

Let E be a p.transformation semigroup on a set X. Using the canonical relations of E we define the following conditions U_k and binary relations $D_{k,l}$ on E:

$$U_1 \overset{\text{df}}{=} (A = X \ \& \ Z \neq \emptyset); \quad U_2 \overset{\text{df}}{=} (\bigwedge_{i=1}^{3} Z_i \neq \emptyset);$$

$$U_3 \overset{\text{df}}{=} (Q_3 \neq \emptyset); \quad U_4 \overset{\text{df}}{=} (Q_3 = \emptyset \ \& \ Q_1 \neq \emptyset);$$

$$U_5 \overset{\text{df}}{=} (Q = Q_2 \ \& \ Q \neq \emptyset);$$

$$U_6 \overset{\text{df}}{=} (Q = Z_3 = \emptyset \ \& \ (Z_2 \neq \emptyset \vee Z_1 = \emptyset));$$

$$U_7 \overset{\text{df}}{=} (Q = Z_1 = \emptyset \ \& \ Z_2 \neq \emptyset);$$

$$U_8 \overset{\text{df}}{=} (Z_3 \neq \emptyset \ \& \ Q = Z_1 = Z_2 = \emptyset);$$

$$U_9 \overset{\text{df}}{=} (Z_1 \neq \emptyset \ \& \ A \neq X \ \& \ Q = Z_2 = Z_3 = \emptyset);$$

$$U_{10} \overset{\text{df}}{=} (Q = Z_2 = \emptyset \ \& \ Z_1 \neq \emptyset \ \& \ Z_3 \neq \emptyset);$$

$$D_{1,1} \overset{\text{df}}{=} \emptyset, \quad D_{1,2} \overset{\text{df}}{=} \Delta_X, \quad D_{1,3} \overset{\text{df}}{=} X^2;$$

$$D_{2,1} \overset{\text{df}}{=} \Delta_A; \quad D_{2,2} \overset{\text{df}}{=} A^2, \quad D_{2,3} \overset{\text{df}}{=} A\times X,$$

$$D_{2,4} \overset{\text{df}}{=} D_{2,3}^{-1}; \quad D_{2,5} \overset{\text{df}}{=} X^2\setminus B^2, \quad D_{2,6} \overset{\text{df}}{=} X^2\setminus\Delta_B;$$

$$D_{3,1} \overset{\text{df}}{=} \beta(a, b) \quad \text{and} \quad D_{3,2} \overset{\text{df}}{=} D_{3,1}^{-1} \quad \text{for} \quad (a, b)\in Q_3;$$

$$D_{4,1} \overset{\text{df}}{=} ((R\setminus Q)\cup\beta^{-1}(a, b))\setminus P \quad \text{and} \quad D_{4,2} \overset{\text{df}}{=} D_{4,1}^{-1} \quad \text{for} \quad (a, b)\in Q_1;$$

$$D_{5,1} \overset{\text{df}}{=} (R\setminus P)\cup\beta(a, b)\cup\sigma \quad \text{and} \quad D_{5,2} \overset{\text{df}}{=} D_{5,1}^{-1} \quad \text{for} \quad (a, b)\in Q\cap(A\times B)$$

and either $\sigma=\beta(d, c)$, if there exists $(c, d)\in(Q\cap(A\times B))\setminus\beta(a, b)$, or $\sigma=\emptyset$ otherwise;

$$D_{6,1} \stackrel{\mathrm{df}}{=} R;\quad D_{7,1} \stackrel{\mathrm{df}}{=} R_1;\quad D_{8,1} \stackrel{\mathrm{df}}{=} R\setminus B^2,\quad D_{8,2} \stackrel{\mathrm{df}}{=} R\setminus(B\times X),$$

$$D_{8,3} \stackrel{\mathrm{df}}{=} D_{8,2}^{-1};\quad D_{9,1} \stackrel{\mathrm{df}}{=} R,\quad D_{9,2} \stackrel{\mathrm{df}}{=} (A\times X)\cup R,$$

$$D_{9,3} \stackrel{\mathrm{df}}{=} D_{9,2}^{-1};\quad D_{10,1} \stackrel{\mathrm{df}}{=} R_1\cup R_2,\quad D_{10,2} \stackrel{\mathrm{df}}{=} (A\times X)\cup R_2,$$

$$D_{10,3} \stackrel{\mathrm{df}}{=} R_1\cup(R_2\cap(A\times B)),\quad D_{10,4} \stackrel{\mathrm{df}}{=} D_{10,2}^{-1},\quad D_{10,5} \stackrel{\mathrm{df}}{=} D_{10,3}^{-1}.$$

The sufficient conditions for a 3-bounded p.transformation semigroup to be equal to the 3-bounded p.endomorphism semigroup of some graph will be given in the following

Theorem 1. *Let E be a 3-bounded p.transformation semigroup on a set X and let the canonical relations of E satisfy the following conditions:*

(T$_1$) $\begin{pmatrix} a \\ x \end{pmatrix} \in E$ *for any* $a\in B$ *and* $x\in X$;

(T$_2$) $\begin{pmatrix} x & y \\ u & v \end{pmatrix}, \begin{pmatrix} x & y \\ v & u \end{pmatrix} \in E$ *imply* $(x, y)\in P$ *or* $(u, v)\in R$;

(T$_3$) $\left(\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \in E \Leftrightarrow \begin{pmatrix} x_1 & y_1 \\ x_3 & y_3 \end{pmatrix} \in E\right) \Rightarrow \begin{pmatrix} x_2 & y_2 \\ x_3 & y_3 \end{pmatrix} \in E$

and

$$\left(\begin{pmatrix} x_1 & y_1 \\ x_3 & y_3 \end{pmatrix} \in E \Leftrightarrow \begin{pmatrix} x_2 & y_2 \\ x_3 & y_3 \end{pmatrix} \in E\right) \Rightarrow \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \in E$$

for any $(x_i, y_i)\in Q$ $(i=\overline{1,3})$ *such that* $(x_k, x_{k+1}), (y_k, y_{k+1})\in\tau$ $(k=1, 2)$.

Then the following conditions hold:

(i) *E satisfies the one of the conditions* U_k $(1\leqq k\leqq 10)$;

(ii) *if E satisfies U_k, then the equality*

$$E = E_3(G) \tag{1}$$

holds if and only if $G=(X, D_{k,l})$ *for some number l.*

Using the canonical relations of p.transformation semigroups we obtain the following concrete characterization of the p.endomorphism semigroup of a graph.

Theorem 2. *Let E be a p.transformation semigroup on a set X. Then E is equal to the p.endomorphism semigroup of some graph if and only if E is the left idealizer*[2]) *of its subsemigroup E_3 in the symmetric semigroup $W(X)$ and the canonical relations of E satisfy the conditions* (T$_1$)—(T$_3$).

[2]) If S is a subsemigroup of a semigroup T, then the left idealizer of S in T is the largest subsemigroup L of T such that S is a left ideal of L.

Before drawing up our main results we verify five lemmas which make the proof of the theorems easier. Let E be a p.transformation semigroup on a set X satisfying (T_1)—(T_3). In the following propositions we collect some properties of the canonical relations of E.

Lemma 7. *E contains all 3-bounded identity p.transformations of X and the canonical relations of E satisfy $(x, y)\in\beta(x, y)$, $\tau a=A$, $\tau^{-1}b=B$ and $\tau^{-1}a=\tau b=X$ for any $x, y\in X$, $a\in A$, $b\in B$.*

Proof. By the definitions of A and (T_1), $\begin{pmatrix}x\\x\end{pmatrix}\in E$ for any $x\in X$. Consider distinct elements $x, y\in X$. Clearly, $(x, x), (y, y)\in\tau$. If $(x, y)\in P\cup R$, then, by the definition, $\begin{pmatrix}x & y\\x & y\end{pmatrix}\in E$. If $(x, y)\in Q$ then $\begin{pmatrix}x & y\\x & y\end{pmatrix}\in E$ by (T_3). We show that $\emptyset\in E$. Since $|X|>1$, there exist two distinct elements $x, y\in X$. Then the p.transformations $\begin{pmatrix}x\\x\end{pmatrix}$, $\begin{pmatrix}y\\y\end{pmatrix}$ and $\emptyset=\begin{pmatrix}x\\x\end{pmatrix}\circ\begin{pmatrix}y\\y\end{pmatrix}$ belong to E. So E contains all 3-bounded identity p.transformations of X. It follows from the definitions of β and A that $(x, y)\in\beta(x, y)$ and $\tau a=A$, $\tau^{-1}b=B$, $\tau^{-1}a=\tau b=X$ for any $x, y\in X$, $a\in A$, $b\in B$.

Lemma 8. *The canonical relations of E satisfy the following conditions:*

(i) $Q_1=Q_1'$,

(ii) $Q_3=Q_3'$,

(iii) $(\beta Q')\setminus R=\begin{cases}Q', & \text{if } Q_3'=\emptyset,\\ (\beta\{(a, b), (b, a)\})\setminus R, & \text{if there is } (a, b)\in Q_3';\end{cases}$

(iv) $(\beta^{-1}Q')\setminus P=\begin{cases}Q', & \text{if } Q_1'=\emptyset;\\ (\beta^{-1}\{(a, b), (b, a)\})\setminus P, & \text{if there is } (a, b)\in Q_1'.\end{cases}$

Proof. Suppose that $(x, y)\in Q_1\cup Q_3$. Then, by Lemma 7, $x, y\in\tau x=\tau y$ and, by the definition of Q, there exists a $(u, v)\in Q'$ such that either $(x, y)\notin R$ and $\begin{pmatrix}u & v\\x & y\end{pmatrix}\in E$ or $\begin{pmatrix}x & y\\u & v\end{pmatrix}\in E$ and $(x, y)\notin P$. We prove that $(x, y)\in Q'$. Suppose the contrary. Let $(x, y)\in P\cup R$. It follows that $\begin{pmatrix}x & y\\y & x\end{pmatrix}\in E$, by the definitions of P and R. Then, by (T_2), either $(u, v)\in P$ or $(u, v)\in R$, which both contradict the assumption. Consequently, $(x, y)\notin P\cup R$, $(x, y)\in Q'$ and $Q_1'\cup Q_3'=Q_1\cup Q_3$. Hence (i) and (ii) are satisfied.

Suppose now that $Q_3'=\emptyset$. By Lemma 7, $Q'\subset(\beta Q')\setminus R$. Conversely, let $(x, y)\in(\beta Q')\setminus R$. Clearly, $(x, y)\in Q$. If $(x, y)\in Q_1$ then $(x, y)\in Q_1'\subset Q'$. Suppose that $(x, y)\notin Q_1$. Then $(x, y)\in Q_2$, since $Q_3=Q_3'=\emptyset$. Let, for example, $(x, y)\in A\times B$. According to $(x, y)\in(\beta Q')\setminus R$, there exists an $(a, b)\in Q'$ such that $(x, y)\notin R$ and

$\begin{pmatrix} a & b \\ x & y \end{pmatrix} \in E$. Then $(a, b) \in Q_1' \cup Q_2'$ since $Q_3' = \emptyset$. By Lemma 7, $(a, b) \in A \times B$. If it were $(x, y) \in P$, then, by definitions, $(a, b) \in P$, which would contradict $P \cap Q' = \emptyset$. Hence $(x, y) \in X^2 \setminus (P \cup R) = Q'$. Thus $(\beta Q') \setminus R = Q'$, if $Q_3' = \emptyset$. Suppose now that there exists an $(a, b) \in Q_3'$. By Lemma 7, $\tau a = \tau b = X$. If $(x, y) \in (\beta Q') \setminus R \subset Q$ then, by (T_3), one of the p.transformations $\begin{pmatrix} a & b \\ x & y \end{pmatrix}$ and $\begin{pmatrix} a & b \\ y & x \end{pmatrix}$ belongs to E, i.e. $(x, y) \in (\beta \{(a, b), (b, a)\}) \setminus R$. On the other hand, it is easy to see that $(a, b) \in Q_3'$ implies $(\beta \{(a, b), (b, a)\}) \setminus R \subset (\beta Q') \setminus R$. Therefore Condition (iii) is satisfied. The latter condition can be proved in a similar way.

Lemma 9. *The canonical relations of E satisfy the following conditions:*

(i) *if* $(x, y) \in A^2 \cup B^2$, *then* $(x, y) \in Q$ *iff* $\begin{pmatrix} x & y \\ y & x \end{pmatrix} \notin E$;

(ii) *if* (x, y), $(u, v) \in Q$ *and* $x, y \in B$ *or* $u, v \in A$, *then one and only one of the p.transformations* $\begin{pmatrix} x & y \\ u & v \end{pmatrix}$ *and* $\begin{pmatrix} x & y \\ v & u \end{pmatrix}$ *belongs to E.*

Proof. Let $(x, y) \in A^2 \cup B^2$. By Lemma 7, $\begin{pmatrix} x & y \\ x & y \end{pmatrix} \in E$. Using (T_2) and the definitions of P and R, we obtain that the condition $(x, y) \in P \cup R$ is equivalent to $\begin{pmatrix} x & y \\ y & x \end{pmatrix} \in E$. Since $Q' = X^2 \setminus (P \cup R)$ and, by Lemma 8, $Q_1 \cup Q_3 = Q_1' \cup Q_3'$, the first statement of the lemma follows. Suppose now that $(x, y), (u, v) \in Q$ and $x, y \in B$ or $u, v \in A$. We may assume that $x, y \in B$. Then, by Lemma 7, $\tau x = \tau y = X$ and, by Lemma 8, $(x. y) \in Q_3 = Q_3'$, $(x, y) \notin P$. By (T_3), one of the p.transformations $\begin{pmatrix} x & y \\ u & v \end{pmatrix}$ and $\begin{pmatrix} x & y \\ v & u \end{pmatrix}$ belongs to E. Suppose that $\begin{pmatrix} x & y \\ u & v \end{pmatrix}, \begin{pmatrix} x & y \\ v & u \end{pmatrix} \in E$. Then, by (T_2), $(u, v) \in R$ since $(x, y) \notin P$. Hence, by the definition of Q, $(u, v) \in (\beta^{-1} Q') \setminus P$, i.e., by Lemma 8, there exists an $(a, b) \in Q_1'$ such that $(u, v) \notin P$ and $\begin{pmatrix} u & v \\ a & b \end{pmatrix} \in E$. This implies $\begin{pmatrix} x & y \\ a & b \end{pmatrix}, \begin{pmatrix} x & y \\ b & a \end{pmatrix} \in E$ and, by (T_2), $(a, b) \in R$ or $(x, y) \in P$, a contradiction. Consequently, one and only one of the p.transformations $\begin{pmatrix} x & y \\ u & v \end{pmatrix}$ and $\begin{pmatrix} x & y \\ v & u \end{pmatrix}$ belongs to E. For $u, v \in A$ the proof is analogous, therefore it is omitted.

Lemma 10. *If* $Q = \emptyset$ *and* $Z_2 \neq \emptyset$ *then one of the relations* Z_1, Z_3 *is nonempty.*

Proof. If $Q = \emptyset$, $Z_2 \neq \emptyset$ and $Z_1 = \emptyset$ then, by Lemmas 2 and 3, $P_2 = R_2 = (A \times B) \cup (B \times A)$, $P_1 \neq \emptyset$ and $R_3 \neq \emptyset$. Let $(x, y) \in R_3$ and $(u, v) \in P_1$. Then for any $(a, b) \in A \times B$ we have $\begin{pmatrix} x & y \\ a & b \end{pmatrix}, \begin{pmatrix} a & b \\ u & v \end{pmatrix} \in E$ and $\begin{pmatrix} x & y \\ u & v \end{pmatrix} \in E$. Since, by the definition of P, $\begin{pmatrix} u & v \\ v & u \end{pmatrix} \in E$, we obtain $\begin{pmatrix} x & y \\ v & u \end{pmatrix} \in E$ and, by (T_2), $(x, y \in P$. Consequently, $Z_3 \neq \emptyset$.

Later we will use the following

Lemma 11. *Let ϱ and σ be binary relations on X. Then $\varrho\subset\sigma$, $\varrho\cap\varrho^{-1}=\sigma\cap\sigma^{-1}$ and $\varrho\cup\varrho^{-1}=\sigma\cup\sigma^{-1}$ imply $\varrho=\sigma$.*

The proof follows from the relations:

$$\sigma = ((\sigma\cup\sigma^{-1})\setminus\sigma^{-1})\cup(\sigma\cap\sigma^{-1}) \subset ((\varrho\cup\varrho^{-1})\setminus\varrho^{-1})\cup(\varrho\cap\varrho^{-1}) = \varrho.$$

Now we turn to the main theorems of the paper.

Proof of Theorem 1. It is easy to see that (i) follows from Lemmas 2, 3 and 10.

Suppose now that E satisfies U_k $(1\leqq k\leqq 10)$. We prove that (1) holds for the graphs $G=(X, D_{k,l})$ and only for them. Note that it is sufficient to verify (1) for one of these graphs with mutually converse relations, since they have equal p.endomorphisms. Clearly, $\emptyset$ is a p.endomorphism of any graph and, by Lemma 7, $\emptyset\in E$. Thus for the proof of (1) we must show that a non-empty 3-bounded p.transformation $f=\begin{pmatrix} x & y \\ u & v\end{pmatrix}$ of X is a p.endomorphism of a graph G iff $f\in E$. We investigate the following ten cases concerning E.

Case 1. Let E satisfy U_1, i.e. $A=X$ and $Z\neq\emptyset$. Then, by Lemma 3, $R=X^2$ and E consists of all 3-bounded p.transformations of X. One can easily see that E is equal to $E_3(G)$ for all graphs $G=(X, D_{1,l})$ $(l=\overline{1,3})$. On the other hand, if a graph $G=(X,\varrho)$ satisfies (1), then either $\varrho=X^2=D_{1,3}$ (if $\varrho\setminus\Delta_X\neq\emptyset$) or $\varrho=\Delta_X=D_{1,2}$ (if $\emptyset\neq\varrho\subset\Delta_X$) or $\varrho=\emptyset=D_{1,1}$.

Case 2. Let E satisfy U_2, i.e. $Z_i\neq\emptyset$ $(i=\overline{1,3})$. Then, by Lemma 3, E consists of all 3-bounded p.transformations of X which map no elements of A into B. One can easily see that E is equal to $E_3(G)$ for all graphs $G=(X, D_{2,l})$ $(l=\overline{1,6})$. On the other hand, let a graph $G=(X,\varrho)$ satisfy (1). Then for any distinct elements $a,b\in B$, by Lemma 3 $\begin{pmatrix} a & b \\ b & a\end{pmatrix}\in E(G)$. Therefore either $(a,b)\in\varrho\cap\varrho^{-1}$ (and in this case $\varrho=X^2\setminus\Delta_B=D_{2,6}$) or $(a,b)\notin\varrho\cup\varrho^{-1}$ (and hence $\varrho\subset X^2\setminus B^2$). In the latter case $\begin{pmatrix} a & b \\ b & a\end{pmatrix}\notin E(G)$ for any $(a,b)\in A\times B$. Then either $(a,b)\in\varrho\cap\varrho^{-1}$ (and in this case $\varrho=X^2\setminus B^2=D_{2,5}$) or $(a,b)\in\varrho\setminus\varrho^{-1}$ (and hence $\varrho=A\times X=D_{2,3}$) or $(a,b)\in\varrho^{-1}\setminus\varrho$ (and hence $\varrho=X\times A=D_{2,4}$) or $(a,b)\notin\varrho\cup\varrho^{-1}$ (and hence $\varrho\subset A^2$). In the latter case, by Lemma 3, $\begin{pmatrix} a & b \\ b & a\end{pmatrix}\in E(G)$ for any distinct elements $a,b\in A$. Then either $(a,b)\in\varrho\cap\varrho^{-1}$ (and in this case $\varrho=A^2=D_{2,2}$) or $(a,b)\notin\varrho\cup\varrho^{-1}$ (and hence $\varrho=\Delta_A=D_{2,1}$).

Case 3. Let E satisfy U_3, i.e. $(a, b)\in Q_3$ for some $a, b\in B$. We prove (1) for the graph $G=(X, \varrho)$ with $\varrho=D_{3,1}=\beta(a, b)$. By Lemma 7, $(a, b)\in\varrho\setminus\varrho^{-1}$. If $L=\beta(a, b)\cup\beta(b, a)$ then, by the definitions of R and Lemma 8, $R\cup Q\subset L$ and $X^2=P\cup L$. Hence any pair $(x, y)\in X^2$ belongs to one of the relations $P\setminus L$ and L. If $(x, y)\in P\setminus L$ then, by the definition, $(x, y)\notin\varrho\cup\varrho^{-1}$ and, by Lemma 7 f belongs to both E and $E_3(G)$. Let $(x, y)\in L$, and for example, $(x, y)\in\beta(a, b)$. If $(x, y)\in\beta(b, a)$, then $(x, y)\in\varrho\cap\varrho^{-1}$ and $\begin{pmatrix} a & b \\ x & y \end{pmatrix}, \begin{pmatrix} a & b \\ y & x \end{pmatrix}\in E$ by the definition of β. Hence, from $f\in E$, it follows that $\begin{pmatrix} a & b \\ u & v \end{pmatrix}, \begin{pmatrix} a & b \\ v & u \end{pmatrix}\in E$ and $(u, v)\in\varrho\cap\varrho^{-1}$, whence f is a p.endomorphism of G. Conversely, if $f\in E_3(G)$, then $(u, v)\in\varrho\cap\varrho^{-1}$ and by the definition of G, $\begin{pmatrix} a & b \\ u & v \end{pmatrix}, \begin{pmatrix} a & b \\ v & u \end{pmatrix}\in E$. By Lemma 8, from the equality $Q_3=Q_3'$ and (T_2) it follows that $(u, v)\in R$, $f\in E$. Suppose now that $(x, y)\notin\beta(b, a)$. In this case $(x, y)\in Q$, $\begin{pmatrix} a & b \\ x & y \end{pmatrix}\in E$ and $(x, y)\in\varrho\setminus\varrho^{-1}$. If $f\in E$ then $\begin{pmatrix} a & b \\ u & v \end{pmatrix}\in E$ and $(u, v)\in\varrho$. Hence f is a p.endomorphism of G. Conversely, let $f\in E_3(G)$. Then $(u, v)\in\varrho=\beta(a, b)$. If $\begin{pmatrix} a & b \\ v & u \end{pmatrix}\in E$ then $(u, v)\in R$ and $f\in E$ by Lemma 8 and (T_2). If $\begin{pmatrix} a & b \\ v & u \end{pmatrix}\notin E$ then $(u, v)\notin R$ and $(u, v)\in Q$ by the definition of Q. In this case $\begin{pmatrix} x & y \\ v & u \end{pmatrix}\notin E$ and, by Lemma 9, $f\in E$. So (1) holds for the graph G.

Conversely, let a graph $G=(X, \varrho)$ satisfy (1). Then, by Lemma 4, $(a, b)\in(\varrho\setminus\varrho^{-1})\cup(\varrho^{-1}\setminus\varrho)$. If $(a, b)\in\varrho$ then $\beta(a, b)\subset\varrho$, and $\beta(a, b)\subset\varrho^{-1}$ otherwise. By Lemma 5, $R=\varrho\cap\varrho^{-1}$. If $(u, v)\in P\setminus(R\cup Q)$ then $\begin{pmatrix} a & b \\ u & v \end{pmatrix}, \begin{pmatrix} a & b \\ v & u \end{pmatrix}\notin E(G)$ and $(u, v)\notin\varrho\cup\varrho^{-1}$. Using Lemma 4 we obtain that $\varrho\cup\varrho^{-1}=R\cup Q$. From Lemmas 8, 9 it follows that the relation $\sigma=\beta(a, b)$ satisfies $\sigma\cap\sigma^{-1}=R$ and $\sigma\cup\sigma^{-1}=R\cup Q$. Then, by Lemma 11, either $\varrho=\beta(a, b)=D_{3,1}$ or $\varrho=\beta(b, a)=D_{3,2}$.

Case 4. Let E satisfy U_4, i.e. $Q_3=\emptyset$ and $(a, b)\in Q_1$ for some $a, b\in A$. We prove (1) for the graph $G=(X, \varrho)$ with $\varrho=D_{4,1}=((R\setminus Q)\cup\beta^{-1}(a, b))\setminus P$. Since $X^2=(R\setminus(P\cup Q))\cup P\cup Q$, any pair $(u, v)\in X^2$ belongs to one of the relations $R\setminus(P\cup Q)$, P and $Q\setminus P$. If $(u, v)\in R\setminus(P\cup Q)$ then $(u, v)\in\varrho\cap\varrho^{-1}$ and f belongs to both E and $E_3(G)$. Suppose that $(u, v)\in P$. Then $(u, v)\notin\varrho\cup\varrho^{-1}$ and, by the definition of P, either $u\in B$ (if $u=v$) or $\begin{pmatrix} u & v \\ a & b \end{pmatrix}, \begin{pmatrix} u & v \\ b & a \end{pmatrix}\in E$ (if $u\neq v$). Let $f\in E$. Then, by (T_2), $(x, y)\in P$ since $R\cap\Delta_B=R\cap Q'=\emptyset$. Consequently, $(x, y)\notin\varrho\cup\varrho^{-1}$ and f is a p.endomorphism of G. Conversely, if $f\in E_3(G)$, then $(x, y)\notin\varrho\cup\varrho^{-1}$, $(x, y)\in P$ and $f\in E$. Further, suppose that $(u, v)\in Q\setminus P$. By Lemma 9, there is a unique mapping of $\{u, v\}$ onto $\{a, b\}$. Let, for example, $\begin{pmatrix} u & v \\ a & b \end{pmatrix}\in E$. It follows that $(u, v)\in\varrho$. If $f\in E$, then $\begin{pmatrix} x & y \\ a & b \end{pmatrix}\in E$ and either $(x, y)\in P$ or $(x, y)\in Q$. Consequently, either

$(x, y)\notin \varrho\cup\varrho^{-1}$ or $(x, y)\in\varrho\setminus\varrho^{-1}$. Hence f is a p.endomorphism of G. Conversely, let $f\in E_3(G)$. Then either $(x, y)\notin\varrho\cup\varrho^{-1}$ or $(x, y)\in\varrho\setminus\varrho^{-1}$. From the definition of G it follows that either $(x, y)\in P$ or $(x, y)\in Q\setminus P$. In the former case $f\in E$ by the definition of P. In the latter case, by (T_2), $\begin{pmatrix} x & y \\ v & u \end{pmatrix}\notin E$ since $\begin{pmatrix} x & y \\ a & b \end{pmatrix}, \begin{pmatrix} u & v \\ a & b \end{pmatrix}\in E$. Then, by Lemma 9, $f\in E$. Thus (1) holds for the graph G.

Conversely, let a graph $G=(X, \varrho)$ satisfy (1). Then, by Lemma 4, $(a, b)\in(\varrho\setminus\varrho^{-1})\cup(\varrho^{-1}\setminus\varrho)$. It follows that $(\beta^{-1}(a, b))\setminus P\subset\varrho$, if $(a, b)\in\varrho$, and $(\beta^{-1}(a, b))\setminus P\subset\varrho^{-1}$ otherwise. By Lemma 5, $P=X^2\setminus(\varrho\cup\varrho^{-1})$. If $(x, y)\in R\setminus(P\cup Q)$ then $\begin{pmatrix} x & y \\ a & b \end{pmatrix}, \begin{pmatrix} x & y \\ b & a \end{pmatrix}\notin E$ and, by (1), $(x, y)\in\varrho\cap\varrho^{-1}$. Thus, by Lemma 4, $\varrho\cap\varrho^{-1}= =R\setminus(P\cup Q)$. From Lemmas 8 and 9 it follows that the relation $\sigma=((R\setminus Q)\cup \cup\beta^{-1}(a, b))\setminus P$ satisfies $\sigma\cap\sigma^{-1}=R\setminus(P\cup Q)$ and $\sigma\cup\sigma^{-1}=X^2\setminus P$. Hence, by Lemma 11, either $\varrho=\sigma=D_{4,1}$ or $\varrho=\sigma^{-1}=D_{4,2}$.

Case 5. Let E satisfy U_5, i.e. $Q=Q_2$ and $(a, b)\in Q$ for some $a\in A$, $b\in B$. We prove (1) for the graph $G=(X, \varrho)$ with $\varrho=D_{5,1}=(R\setminus P)\cup\beta(a, b)\cup\sigma$ where $\sigma=\beta(d, c)$, if there exists a $(c, d)\in(Q\cap(A\times B))\setminus\beta(a, b)$, and $\sigma=\emptyset$ otherwise. By U_5 and Lemma 8, $Q=Q'$. Since $X^2=(P\setminus(R_1\cup R_2))\cup(R\setminus P_3)\cup Q$, any pair $(x, y)\in X^2$ belongs to one of the relations $P\setminus(R_1\cup R_2)$, $R\setminus P_3$ and Q. If $(x, y)\in P\setminus(R_1\cup R_2)$ then, by the definition of G, $(x, y)\notin\varrho\cup\varrho^{-1}$ and f belongs to both E and $E_3(G)$. Suppose that $(x, y)\in R\setminus P_3$. This implies $(x, y)\in\varrho\cap\varrho^{-1}$. If $f\in E$ then $(u, v)\notin P_3$ since $(x, y)\in P_3$ otherwise. We show that $(u, v)\in R$. If $x, y\in B$ then by $Q=Q_2$ and, Lemma 9, $\begin{pmatrix} x & y \\ v & u \end{pmatrix}\in E$ and, by (T_2), $(u, v)\in R$. Now suppose that one of the elements x and y belongs to A, for example, $x\in A$. Since $f\in E$, $u\in A$. It follows that either $u, v\in A$ or $(u, v)\in A\times B$. In the former case, by $Q_2=Q'\neq\emptyset$, (T_2) and Lemmas 3, 9, $(u, v)\in R$. In the latter case $(x, y)\in R_2$ and $f\in E$ imply $(u, v)\in R_2$. So $(u, v)\in\varrho\cap\varrho^{-1}$ and f is a p.endomorphism of G. Conversely, if $f\in E_3(G)$ then $(u, v)\in\varrho\cap\varrho^{-1}$ and, by the definition of G, $(u, v)\in R\setminus P_3$. Thus $f\in E$. Now suppose that $(x, y)\in Q\cap(A\times B)$. Then either $\begin{pmatrix} a & b \\ x & y \end{pmatrix}\in E$ or $\begin{pmatrix} a & b \\ x & y \end{pmatrix}\notin E$. In the latter case, by (T_3), $\begin{pmatrix} c & d \\ x & y \end{pmatrix}\in E$. We may suppose that $\begin{pmatrix} a & b \\ x & y \end{pmatrix}\in E$. Then $(x, y)\in\varrho\setminus\varrho^{-1}$. Assume that $f\in E$. It follows $u\in A$ and $\begin{pmatrix} a & b \\ u & v \end{pmatrix}\in E$. If $v\in A$ then, by $Q=Q_2$ and Lemma 9, $\begin{pmatrix} u & v \\ v & u \end{pmatrix}\in E$. Consequently, $\begin{pmatrix} a & b \\ v & u \end{pmatrix}\in E$ and, by (T_2), $(u, v)\in R_1$. If $v\in B$ then $(u, v)\notin P_2$ since $(a, b)\in P_2$ otherwise. Hence (u, v) belongs to R_2 or Q_2. Moreover, in the latter case $(u, v)\in\beta(a, b)$. So $(u, v)\in\varrho\cap\varrho^{-1}$ or $(u, v)\in\varrho\setminus\varrho^{-1}$, whence $f\in E_3(G)$. Conversely, let $f\in E_3(G)$. If $\begin{pmatrix} x & y \\ v & u \end{pmatrix}\in E_3(G)$ then $(u, v)\in\varrho\cap\varrho^{-1}$. By the definition of G,

$(u, v)\in R\setminus P_3$. Then by the definition of R, $f\in E$. If $\begin{pmatrix} x & y \\ v & u \end{pmatrix}\notin E_3(G)$ then $(u, v)\in$ $\in(\varrho\setminus\varrho^{-1})\cup(\varrho^{-1}\setminus\varrho)$. It follows from the definition of G that $(u, v)\in Q\cap(A\times B)$ and $\begin{pmatrix} a & b \\ u & v \end{pmatrix}\in E$, since $\begin{pmatrix} a & b \\ x & y \end{pmatrix}\in E$. Then, by Lemma 9, $f\in E$. Thus (1) holds for the graph G.

Conversely, let a graph $G=(X, \varrho)$ satisfy (1). Then, by Lemma 4, $(a, b)\in(\varrho\setminus\varrho^{-1})\cup(\varrho^{-1}\setminus\varrho)$. If $(x, y)\in P_3$ then $\begin{pmatrix} x & y \\ a & b \end{pmatrix}$ and $\begin{pmatrix} x & y \\ b & a \end{pmatrix}$ are p.endomorphisms of G, and $(x, y)\notin\varrho\cup\varrho^{-1}$. If $(x, y)\in P_2\cap(A\times B)$ then $\begin{pmatrix} x & y \\ a & b \end{pmatrix}\in E$ and $\begin{pmatrix} a & b \\ x & y \end{pmatrix}$ $\notin E$, whence $(x, y)\notin\varrho\cup\varrho^{-1}$. Moreover $\begin{pmatrix} a & b \\ x & y \end{pmatrix}, \begin{pmatrix} a & b \\ y & x \end{pmatrix}\in E_3(G)$ for $(x, y)\in R_1$ and $\begin{pmatrix} a & b \\ x & y \end{pmatrix}$ $\in E$, $\begin{pmatrix} x & y \\ a & b \end{pmatrix}\notin E$ for $(x, y)\in R_2\cap(A\times B)$. In these cases $(x, y)\in\varrho\cap\varrho^{-1}$. If $(x, y)\in P_1\setminus R_1$ then $\begin{pmatrix} a & b \\ x & y \end{pmatrix}\notin E$ and, by Lemmas 8 and 9, $\begin{pmatrix} x & y \\ y & x \end{pmatrix}\in E$, whence $(x, y)\notin\varrho\cup\varrho^{-1}$. If $(x, y)\in R_3\setminus P_3$ then $\begin{pmatrix} x & y \\ a & b \end{pmatrix}\notin E$ and, by Lemmas 8 and 9 $\begin{pmatrix} x & y \\ y & x \end{pmatrix}\in E$, whence $(x, y)\in\varrho\cap$ $\cap\varrho^{-1}$. So $X^2\setminus(\varrho\cup\varrho^{-1})=P\setminus R_1$ and $\varrho\cap\varrho^{-1}=R\setminus P_3$. Moreover, $\beta(a, b)\subset\varrho$, if $(a, b)\in\varrho$, and $\beta(a, b)\subset\varrho^{-1}$ otherwise. Let there exist a $(c, d)\in(Q\cap(A\times B))\setminus\beta(a, b)$. Then, by Lemma 4, (c, d) belongs to $(\varrho\setminus\varrho^{-1})\cup(\varrho^{-1}\setminus\varrho)$. It follows that either $(d, c)\in\varrho$, if $(a, b)\in\varrho$, or $(d, c)\in\varrho^{-1}$ otherwise. Consequently, $\beta(d, c)\subset\varrho$, if $(a, b)\in\varrho$, and $\beta(d, c)\subset\varrho^{-1}$ otherwise. So the relation $\sigma=(R\setminus P)\cup\beta(a, b)\cup\beta(d, c)$ satisfies the conditions: $\sigma\cap\sigma^{-1}=R\setminus P_3=\varrho\cap\varrho^{-1}$, $\sigma\cup\sigma^{-1}=X^2\setminus(P\setminus R_1)=\varrho\cup\varrho^{-1}$ and either $\sigma\subset\varrho$ or $\sigma\subset\varrho^{-1}$. By Lemma 11, $\varrho=\sigma=D_{5,1}$ or $\varrho=\sigma^{-1}=D_{5,2}$.

Cases 6 and 7. If E satisfies U_6 (or U_7), then it is easy to verify that (1) holds for the graph $G=(X, \varrho)$ with $\varrho=R=D_{6,1}$ (or $\varrho=R_1=D_{7,1}$). On the other hand, if a graph $G=(X, \varrho)$ satisfies (1), then by Lemmas 4 and 7 $\varrho=R=D_{6,1}$ (or $\varrho=R_1=$ $=D_{7,1}$).

Case 8. Let E satisfy U_8, i.e. $Q=Z_1=Z_2=\emptyset$ and $Z_3\neq\emptyset$. We prove (1) for the graph $G=(X, \varrho)$ with $\varrho=D_{8,1}$ (and $\varrho=D_{8,2}$). If $(x, y)\in P$ then $(x, y)\notin\varrho\cup\varrho^{-1}$ and f belongs to both E and $E_3(G)$. Suppose that $(x, y)\in R\setminus P$. Then $(x, y)\notin B^2$ since U_8 and Lemma 3 imply $B^2\subset P$. If $(x, y)\in A^2$ then $(x, y)\in\varrho\cap\varrho^{-1}$ and, by Lemma 9, $\begin{pmatrix} x & y \\ y & x \end{pmatrix}\in E$. Hence $f\in E$ implies $\begin{pmatrix} x & y \\ v & u \end{pmatrix}\in E$ and, by (T_2), $(u, v)\in R$. It follows that $(u, v)\in\varrho\cap\varrho^{-1}$ and f is a p.endomorphism of G. Conversely, if $f\in E_3(G)$, then $(u, v)\in\varrho\cap\varrho^{-1}$ and, by the definition of G, $(u, v)\in R_1$. Thus $f\in E$. Let $(x, y)\notin A^2$. Without loss of generality, we can assume that $(x, y)\in A\times B$. For the graph G with the relation $D_{8,1}$ (or $D_{8,2}$), it follows $(x, y)\in\varrho\cap\varrho^{-1}$ (or $(x, y)\in\varrho\setminus\varrho^{-1}$). If $f\in E$ then, by the definition of R and (T_2), $(u, v)\in R_1\cup R_2$. Hence $(u, v)\in\varrho\cap\varrho^{-1}$

(resp. $(u,v)\in\varrho\setminus\varrho^{-1}$ or $(u,v)\in\varrho\cap\varrho^{-1}$) for the graph G with the relation $D_{8,1}$ (resp. $D_{8,2}$). It follows that f is a p.endomorphism of G. It is easy to verify that $f\in E_3(G)$ implies $f\in E$. So (1) holds for the graph G with any of the relations $D_{8,l}$ $(l=\overline{1,3})$. Conversely, let a graph $G=(X,\varrho)$ satisfy (1). Then, by Lemmas 5 and 6, $P=X^2\setminus(\varrho\cup\varrho^{-1})$, $R_1=(\varrho\cap\varrho^{-1})_1$ and either $R_2=(\varrho\cap\varrho^{-1})_2$ or $R_2=((\varrho\setminus\varrho^{-1})\cup\cup(\varrho^{-1}\setminus\varrho))_2$. Since $\begin{pmatrix}a & b\\ c & d\end{pmatrix}\in E$ for any $(a,b),(c,d)\in R_2\cap(A\times B)$, the relation ϱ equals one of the relations $D_{8,l}$ $(1\leqq l\leqq 3)$.

Case 9. Let E satisfy U_9, i.e. $A\neq X$, $Z_1\neq\emptyset$ and $Q=Z_2=Z_3=\emptyset$. We prove (1) for the graph $G=(X,\varrho)$ with $\varrho=D_{9,1}$ (and $\varrho=D_{9,2}$). If $(u,v)\in R\setminus(P_2\cup P_3)$ then $(u,v)\in\varrho\cap\varrho^{-1}$ and f belongs to both E and $E_3(G)$. If $(u,v)\in P_3$ then $(u,v)\notin\varrho\cup\cup\varrho^{-1}$ and $f\in E$ is equivalent to $(x,y)\in P_3$, i.e. $(x,y)\notin\varrho\cup\varrho^{-1}$. The latter is equivalent to $f\in E_3(G)$. Suppose that $(u,v)\in P_2\cap(A\times B)$. Then $(u,v)\in\varrho\setminus\varrho^{-1}$ (or $(u,v)\notin\varrho\cup\varrho^{-1}$) for the graph G with the relation $D_{9,2}$ (or $D_{9,1}$). Let $f\in E$. If $x,y\in B$ then, by $Q=\emptyset$ and Lemma 9, $\begin{pmatrix}x & y\\ y & x\end{pmatrix}\in E$. Hence $\begin{pmatrix}x & y\\ v & u\end{pmatrix}\in E$ and, by (T_2), $(x,y)\in P_3$. It follows that $(x,y)\notin\varrho\cup\varrho^{-1}$ and f is a p.endomorphism of G. If $(x,y)\notin B^2$ then $(u,v)\in P_2$ and $f\in E$ imply $(x,y)\in P_2\cap(A\times B)$. Thus $(x,y)\in\varrho\setminus\varrho^{-1}$ (or $(x,y)\notin\varrho\cup\cup\varrho^{-1}$) for the graph G with the relation $D_{9,2}$ (or $D_{9,1}$). Therefore f is a p.endomorphism of G. Conversely, if $f\in E_3(G)$ then $(x,y)\notin\varrho\cup\varrho^{-1}$. Hence, by the definition of G, either $(x,y)\in P_3$ or $(x,y)\in P_2$; moreover, in the latter case, $(x,y)\in\varrho\setminus\varrho^{-1}$ for the graph G with the relation $D_{9,1}$. It follows that (x,y) belongs to P_3 or $P_2\cap(A\times B)$, whence, by the definition of P, $f\in E$. Thus (1) holds for the graph G with each of the relations $D_{9,l}$ $(l=\overline{1,3})$.

Conversely, let a graph $G=(X,\varrho)$ satisfy (1). Then, by Lemmas 3 and 5, $R_1=A^2$, $R=\varrho\cap\varrho^{-1}$ and $P_3=(X^2\setminus(\varrho\cup\varrho^{-1}))_3$. Since $\begin{pmatrix}a & b\\ c & d\end{pmatrix}\in E$ for any $(a,b),(c,d)\in P_2\cap(A\times B)$, from Lemma 6 it follows that the relation ϱ equals one of the relations $D_{9,l}$ $(1\leqq l\leqq 3)$.

Case 10. Let E satisfy U_{10}, i.e. $Q=Z_2=\emptyset$ and $Z_1,Z_3\neq\emptyset$. We prove (1) for the graph $G=(X,\varrho)$ with $\varrho=D_{10,l}$ $(l=\overline{1,3})$. By U_{10} and Lemma 3, $P_3=B^2$ and $R_1=A^2$. If $(x,y)\in B^2$ or $(u,v)\in A^2$ then, by the definition of G, $(x,y)\notin\varrho\cup\varrho^{-1}$ or $(u,v)\in\varrho\cap\varrho^{-1}$, respectively. It follows that f belongs to both E and $E_3(G)$. If $(x,y)\in A^2$ and f belongs to E or $E_3(G)$, then by Lemmas 4 and 9, $(u,v)\in A^2$. Analogously, $(x,y)\in B^2$ if $(u,v)\in B^2$ and f belongs to E or $E_3(G)$. Suppose now that $(x,y)\notin B^2$ and $(u,v)\notin A^2$. Hence (x,y) and (u,v) belong to $(A\times B)\cup(B\times A)$. Let, for example, $(x,y)\in A\times B$. If $f\in E$ then $(u,v)\in A\times B$ and, by the definitions, the following condition holds:

$$(2)\qquad \begin{cases}\text{either } (x,y),\ (u,v)\in P_2, \text{ or } (x,y),\ (u,v)\in R_2,\\ \text{or } (x,y)\in P_2 \text{ and } (u,v)\in R_2.\end{cases}$$

It follows that f is a p.endomorphism of G for any $D_{10,l}$ $(l=\overline{1,3})$. Conversely, if $f\in E_3(G)$, then $(u,v)\in A\times B$ and, by the definition of G, (2) is satisfied. Hence, by the definition of P and R, $f\in E$. So (1) holds for the graph G with any of the relations $D_{10,l}$ $(l=\overline{1,5})$.

Conversely, let a graph $G=(X,\varrho)$ satisfy (1). Then, by Lemmas 3 and 6, $P_3=(X^2\setminus(\varrho\cup\varrho^{-1}))_3=B^2$ and $R_1=(\varrho\cap\varrho^{-1})_1=A^2$. Moreover, either $P_2=((\varrho\setminus\varrho^{-1})\cup(\varrho^{-1}\setminus\varrho))_2$ and $R_2=(\varrho\cap\varrho^{-1})_2$, or $P_2=(X^2\setminus(\varrho\cup\varrho^{-1}))_2$ and one of the following equalities holds: $R_2=(\varrho\cap\varrho^{-1})_2$ or $R_2=((\varrho\setminus\varrho^{-1})\cup(\varrho^{-1}\setminus\varrho))_2$. We show that other cases are impossible. By Lemma 3, there exist (a,b) and (c,d) in $A\times B$ such that $(a,b)\in P_2$ and $(c,d)\in R_2$. Using $Z_2=\emptyset$ and the definition of P and R, we obtain that $\begin{pmatrix} a & b\\ c & d\end{pmatrix}\in E$ and $\begin{pmatrix} c & d\\ a & b\end{pmatrix}\notin E$. Clearly, $\begin{pmatrix} x & y\\ u & v\end{pmatrix}\in E$ if (x,y) and (u,v) belong to $P_2\cap(A\times B)$ (or $R_2\cap(A\times B)$). It follows that the relation ϱ equals one of the relations $D_{10,l}$ $(1\leqq l\leqq 5)$. The proof is complete.

Proof of Theorem 2. Let $E=E(G)$ for some graph $G=(X,\varrho)$. From the definition of a p.endomorphism it follows that E is a left idealizer of its subsemigroup $E_3=E_3(G)$ in the symmetric semigroup $W(X)$. We prove (T_1)—(T_3). Clearly, these conditions hold for $E=W(X)$. Suppose that $E\neq W(X)$. Then, by Lemma 4, a vertex $a\in X$ has no loop iff $a\in B$. Thus $\begin{pmatrix} a\\ x\end{pmatrix}$ is a p.endomorphism of G for any $a\in B$ and $x\in X$, i.e. (T_1) holds. Consider now $(x_i,y_i)\in Q$ $(i=\overline{1,3})$ such that $(x_k,x_{k+1}),(y_k,y_{k+1})\in\tau$ $(k=1,2)$. Then, by Lemma 4, $(x_i,y_i)\in(\varrho\setminus\varrho^{-1})\cup(\varrho^{-1}\setminus\varrho)$ and (T_3) is satisfied. Suppose that $\begin{pmatrix} x & y\\ u & v\end{pmatrix},\begin{pmatrix} x & y\\ v & u\end{pmatrix}$ are p.endomorphisms of G. Then $(x,y)\notin\varrho\cup\varrho^{-1}$ or $(u,v)\in\varrho\cap\varrho^{-1}$, whence, by Lemma 4, $(x,y)\in P$ or $(u,v)\in R$. Thus (T_2) holds.

Conversely, let a p.transformation semigroup E satisfy the conditions of Theorem 2. Then the semigroups E and E_3 are determined by each other and, by Lemmas 1 and 7, their canonical relations are equal. Moreover, E_3 satisfies (T_1)—(T_3). By Theorem 1, $E_3=E_3(G)$ for some graph G. Then $E=E(G)$ since $E(G)$ is the left idealizer of $E_3(G)$ in $W(X)$. Theorem 2 is proved.

Remark. The conditions (T_1)—(T_3) are independent.

3. Applications

Two graphs, G and G' are called E-equivalent, if $E(G)=E(G')$.

As an application of Theorems 1 and 2 we describe E-equivalent graphs and graphs with isomorphic or elementarily equivalent [9] p.endomorphism semigroups.

Theorem 3. *The graphs $G=(X,\varrho)$ and $G'=(X,\varrho')$ are E-equivalent iff either*

$\varrho=\varrho'$ or $\varrho=\varrho'^{-1}$ or ϱ, ϱ' simultaneously belong to one of the following classes:

$$K_0 \stackrel{\mathrm{df}}{=} \{\emptyset, \Delta_X, X^2\};$$

$$K(A) \stackrel{\mathrm{df}}{=} \{\Delta_A, A^2, A\times X, X\times A, X^2\setminus B^2, X^2\setminus \Delta_B\};$$

$$K(A, \alpha, \gamma) \stackrel{\mathrm{df}}{=} \{\alpha\cup\gamma, \alpha\cup(\gamma\cap(A\times B)), \alpha\cup(\gamma\cap(B\times A))\};$$

$$K(A, \delta, \zeta) \stackrel{\mathrm{df}}{=} \{A^2\cup\delta\cup\zeta, (A\times X)\cup\delta\cup\zeta, (X\times A)\cup\delta\cup\zeta\};$$

$$K(A, \gamma) \stackrel{\mathrm{df}}{=} \{A^2\cup\gamma, A^2\cup(\gamma\cap(A\times B)), A^2\cup(\gamma\cap(B\times A)), (A\times X)\cup\gamma, (X\times A)\cup\gamma\},$$

for some proper subset $A\subset X$, $B=X\setminus A$ and symmetrical relations α, ζ, γ, δ such that $\Delta_A\subset\alpha\subsetneqq A^2$, $\emptyset\neq\zeta\subset B^2\setminus\Delta_B$, γ, $\delta\subset(A\times B)\cup(B\times A)$ and $\gamma\neq\emptyset$.

The proof follows from Theorems 1 and 2.

Denote by $\mathfrak{A}$ the class of all p.endomorphism semigroups of graphs. The signature Ω of $\mathfrak{A}$ consists of the single symbol $\cdot$ for the binary semigroup ooeration. A n-place predicate $\Phi(x_1, \ldots, x_n)$ is called formular in $\mathfrak{A}$ if there exists a formula $F(x_1, \ldots, x_n)$ of the signature Ω such that for every semigroup $S\in\mathfrak{A}$ and for every $x_1, \ldots, x_n\in S$, $F(x_1, \ldots, x_n)$ is true iff $\Phi(x_1, \ldots, x_n)$ is true. Consider the following formulas:

$$(3)\qquad O(x) \stackrel{\mathrm{df}}{=} (\forall y)(x\cdot y = y\cdot x = x);\quad I(x) \stackrel{\mathrm{df}}{=} (\neg O(x)\ \&\ x\cdot x = x);$$

$$J(x) \stackrel{\mathrm{df}}{=} (I(x)\ \&\ (\forall y)(I(y)\ \&\ x\cdot y = y \Rightarrow y\cdot x = x));$$

$$M(x) \stackrel{\mathrm{df}}{=} (J(x)\ \&\ (\forall y)(J(y)\ \&\ y\cdot x \neq x \Rightarrow O(x\cdot y)));$$

$$T(x, y) \stackrel{\mathrm{df}}{=} (M(x)\ \&\ M(y)\ \&\ (\exists z)(\neg O(y\cdot z\cdot x))).$$

Let $G=(X, \varrho)$ be a graph, and let $E=E(G)$. We write $\bar{x}=\binom{x}{x}$ for $x\in X$. Clearly, $x\mapsto\bar{x}$ $(x\in X)$ is a one-to-one mapping of X into E. Then any relation σ on X is mapped onto the relation $\bar{\sigma}$ on E, and any condition U on the relations σ_i on X is transformed into the condition $\bar{U}$ on the relations $\bar{\sigma}_i$ on E. Using (3) we can prove that the relations X, ϱ, $D_{k,l}$ and the canonical relations of E are mapped onto the relations $\bar{X}$, $\bar{\varrho}$, $\bar{D}_{k,l}$ and so on, which are determined by formular predicates in $\mathfrak{A}$. For example, $\bar{X}$ and $\bar{\tau}$ are determined by the formulas $M(x)$ and $T(x, y)$, resp. Denote by F_k and $R_{k,l}$ the formulas that determine the conditions $\bar{U}_k$ and the relations $\bar{D}_{k,l}$, resp. Clearly, F_k is a proposition and $R_{k,l}$ a two-place predicate.

Theorem 4. *Let G and G' be graphs, and let $E=E(G)$, $E'=E(G')$. Then the following holds:*

(i) *if E and E' are elementarily equivalent then G is elementarily equivalent to a graph that is E-equivalent to G';*

(ii) *E and E' are isomorphic iff G is isomorphic to a graph that is E-equivalent to G'.*

Proof. If E and E' are elementarily equivalent and E satisfies U_k then F_k is true for E and E'. Therefore, by Theorem 1, the relations of G and G' are equal to $D_{k,l}$ and $D_{k,m}$ for some l and m. Consider the graph G_1 with the vertex-set of G' and the relation $D_{k,l}$ for the semigroup E'. By Theorem 1, $E(G')=E(G_1)$. Thus the formulas $M(x)$ and $P_{k,l}(x, y)$ determine the graphs $\bar{G}$ (on E) and $\bar{G}_1$ (on E') such that $G\cong\bar{G}$ and $G_1\cong\bar{G}_1$. On the analogy of [3] we obtain that G is elementarily equivalent to one of the graphs $\bar{G}_1$ and $\bar{G}_1^{-1}$. So G is elementarily equivalent to a graph that is E-equivalent to G', i.e. (i) holds.

Now suppose that $E\cong E'$. Then the semigroups are elementarily equivalent. Using the previous reasoning, we can prove that an isomorphism of E onto E' determines the isomorphism of G onto one of the graphs G_1 and G_1^{-1}. Hence G is isomorphic to a graph that is E-equivalent to G'. Theorem 3 implies the converse assertion. Thus (ii) holds. This completes the proof.

Consider a graph G with a reflexive (or antireflexive) relation ϱ such that $\varrho\neq\emptyset$, Δ_X, X^2. One can easily see that the E-equivalence class of G consists only of the graphs G and G^{-1}. Hence Theorem 4 yields the results of [2, 3] on p.endomorphism semigroups of reflexive graphs.

References

[1] L. M. Popova, Generating relations of partial endomorphism semigroup of a finite linearly ordered set, *Uč. Zap. Leningrad. Gos. Ped. Inst.*, **238** (1962), 78—88.

[2] L. M. Popova, On one partial endomorphism semigroup of a set with a relation, *Uč. Zap. Leningrad. Gos. Ped. Inst.*, **238** (1962), 49—77.

[3] Ju. M. Važenin, Elementary properties of partial transformation semigroups of reflexive graphs, *Studies in Modern Algebra* (Sverdlovsk, 1977), 9—25.

[4] A. M. Kalmanovič, Partial endomorphism semigroups of graphs, *Dopovidi Acad. Nauk URSR*, 2 (1965), 147—150.

[5] A. M. Kalmanovič, An abstract characterization of the one-to-one partial endomorphism semigroup of a graph, *Dopovidi Acad. Nauk URSR*, **11** (1969), 983—985.

[6] A. M. Kalmanovič, Certain theorems on graphs and their endomorphism semigroups, *Dopovidi Acad. Nauk URSR*, **7** (1970), 586—588.

[7] B. Jónsson, *Topics in universal algebra*, Lecture notes in Math. 250, Springer-Verlag (Berlin—Heidelberg—New York, 1972).

[8] V. A. Molčanov, Endomorphism semigroups of graphs, *XVI-th All-Union Conf. General Algebra, Leningrad*, Vol. 2 (1981), 92—93.

[9] A. Robinson, *Introduction to model theory and to the metamathematics of algebra*, North-Holland (Amsterdam, 1963).

410028 SARATOV
RABOČAYA 19 KV. 48
S.S.S.R.

Acta Sci. Math., 51 (1987), 365—369

The analytic behavior of the holiday numbers

L. A. SZÉKELY

1. Introduction

Investigating Hilbert's fourth problem Z. I. SZABÓ [7] introduced the holiday numbers. In my previous paper [8] many combinatorial and algebraic properties of these numbers were treated. These properties are close to those of the Stirling numbers of the second kind. The aim of the present paper is to investigate the analytic behavior of the holiday numbers. We follow the main ideas of HARPER [1], who investigated the analytic behavior of the Stirling numbers of the second kind.

We recall from [8] two possible definitions of the holiday numbers. The holiday numbers of the first kind are $\psi(m, i)$ (of the second kind $\varphi(m, i)$), where

$$\sum_{m=0}^{\infty} (z^m/m!) \sum_{k=0}^{m} \psi(m, k) t^k = \left(1/\sqrt{1-2z}\right) \exp t\left(1/\sqrt{1-2z}-1\right), \tag{1}$$

$$\sum_{m=0}^{\infty} (z^m/m!) \sum_{k=0}^{m} \varphi(m, k) t^k = \left(1/(1-2z)\right) \exp\left(t\left(1/\sqrt{1-2z}-1\right)\right). \tag{1'}$$

The second definition is

$$\psi(m, k) = (2m+k-1)\psi(m-1, k)+\psi(m-1, k-1), \tag{2}$$

$$\psi(0, 0) = 1, \quad \psi(0, t) = 0 \quad \text{for} \quad t \neq 0,$$

$$\varphi(m, k) = (2m+k)\varphi(m-1, k)+\varphi(m-1, k-1), \tag{2'}$$

$$\varphi(0, 0) = 1, \quad \varphi(0, t) = 0 \quad \text{for} \quad t \neq 0.$$

We use the notations $\psi_m = \sum_k \psi(m, k)$ and $\varphi_m = \sum_k \varphi(m, k)$.

Received October 23, 1984.

2. Results

Statement 1. *The holiday numbers are strongly logconcave in the following sense: for* $0 \leqq k \leqq n$,

$$\psi(n, k)^2 > \psi(n, k-1)\psi(n, k+1), \quad \varphi(n, k)^2 > \varphi(n, k-1)\varphi(n, k+1).$$

The statement is a special case of KURTZ's theorem [2]. It follows that the holiday numbers are of unimodal distribution, for any n their maximum value is attained at most two times. The statement is important to get the corollaries of our theorems.

Theorem 2. *ψ_n and φ_n admit asymptotic expansions in the powers of $n^{1/3}$ in the following way:*

$$\psi_n \sim \left(n!2^n/e\sqrt{3\pi}\right)e^{2^{-2/3}\cdot 3\cdot n^{1/3}}(n^{-1/2}+a_1n^{-5/6}+a_2n^{-7/6}+\dots), \tag{3}$$

$$\varphi_n \sim \left(n!2^{n+1/3}/e\sqrt{3\pi}\right)e^{2^{-2/3}\cdot 3\cdot n^{1/3}}(n^{-1/6}+b_1n^{-1/2}+b_2n^{-5/6}+\dots). \tag{3'}$$

We have also

$$\psi_{n+1}/\psi_n = 2n+(2n)^{1/3}+O(1), \quad \varphi_{n+1}/\varphi_n = 2n+(2n)^{1/3}+O(1),$$

$$\psi_n^{-2}\{-\psi_{n+1}^2+\psi_n\psi_{n+2}-2\psi_n\psi_{n+1}-\psi_n^2\} \sim (2/3)(2n)^{1/3}, \tag{4}$$

$$\varphi_n^{-2}\{-\varphi_{n+1}^2+\varphi_n\varphi_{n+2}-2\varphi_n\varphi_{n+1}-\varphi_n^2\} \sim (2/3)(2n)^{1/3}. \tag{4'}$$

Theorem 3. *The holiday numbers are asymptotically normal in the following sense:*

$$\lim_{n\to\infty}(1/\psi_n)\sum_{j=0}^{x_n}\psi(n, j) = (1/\sqrt{2\pi})\int_{-\infty}^{x} e^{-t^2/2}\,dt, \tag{5}$$

$$\lim_{n\to\infty}(1/\varphi_n)\sum_{j=0}^{y_n}\varphi(n, j) = (1/\sqrt{2\pi})\int_{-\infty}^{y} e^{-t^2/2}\,dt, \tag{5'}$$

where

$$x_n = \psi_{n+1}/\psi_n-(2n+2)+(x/\psi_n)\{-\psi_{n+1}^2+\psi_n\psi_{n+2}-2\psi_n\psi_{n+1}-\psi_n^2\}^{1/2}, \tag{6}$$

$$y_n = \varphi_{n+1}/\varphi_n-(2n+3)+(y/\varphi_n)\{-\varphi_{n+1}^2+\varphi_n\varphi_{n+2}-2\varphi_n\varphi_{n+1}-\varphi_n^2\}^{1/2}; \tag{6'}$$

or

$$x_n = (2n)^{1/3}+x((2/3)(2n)^{1/3})^{1/2}, \quad y_n = (2n)^{1/3}+y((2/3)(2n)^{1/3})^{1/2}. \tag{6''}$$

Corollary 4. *Using the definitions of x_n, y_n in (6), (6') or (6'') we have*

$$\psi_n^{-2}\{-\psi_{n+1}^2+\psi_n\psi_{n+2}-2\psi_n\psi_{n+1}-\psi_n^2\}^{1/2}\psi(n, [x_n]) \to (1/\sqrt{2\pi})e^{-x^2/2},$$

$$\varphi_n^{-2}\{-\varphi_{n+1}^2+\varphi_n\varphi_{n+2}-2\varphi_n\varphi_{n+1}-\varphi_n^2\}^{1/2}\varphi(n, [y_n]) \to (1/\sqrt{2\pi})e^{-y^2/2},$$

i.e.

$$\psi(n, [x_n]) \sim (n!\,2^{n-7/6}/e\pi)\, e^{2^{-2/3}\cdot 3\cdot n^{1/3}} n^{-2/3} e^{-x^2/2},$$

$$\varphi(n, [y_n]) \sim (n!\,2^{n-5/6}/e\pi)\, e^{2^{-2/3}\cdot 3\cdot n^{1/3}} n^{-1/3} e^{-y^2/2}.$$

Corollary 5. *Suppose, for* $i=I_n$ *the maximum value of* $\psi(n,i)$ *(for* $i=J_n$ *the maximum value of* $\varphi(n,i)$*) is attained. Then for every* $\varepsilon>0$ *there exists* N *such that for* $n>N$

$$|I_n-(2n)^{1/3}| < \varepsilon n^{1/6}, \quad |J_n-(2n)^{1/3}| < \varepsilon n^{1/6}.$$

Corollary 6.

$$\max_j \psi(n,j) \sim (2\pi)^{-1/2} \psi_n^2 \{-\psi_{n+1}^2 + \psi_n \psi_{n+2} - 2\psi_n \psi_{n+1} - \psi_n^2\}^{-1/2},$$

$$\max_j \varphi(n,j) \sim (2\pi)^{-1/2} \varphi_n^2 \{-\varphi_{n+1}^2 + \varphi_n \varphi_{n+2} - 2\varphi_n \varphi_{n+1} - \varphi_n^2\}^{-1/2},$$

i.e.

$$\max_j \psi(n,j) \sim (n!\,2^{n-7/6}/e\pi)\, e^{2^{-2/3}\cdot 3\cdot n^{1/3}} n^{-2/3},$$

$$\max_j \varphi(n,j) \sim (n!\,2^{n-5/6}/e\pi)\, e^{2^{-2/3}\cdot 3\cdot n^{1/3}} n^{-1/3}.$$

The corollaries follow from the fact, that the convergence in Theorem 3 is uniform behind the integrals. Its reason is Statement 1, and the proof goes on the same way as in Harper's paper.

3. The proof of the theorems

In order to prove (3) and (3′) we have to give the asymptotic expansion of the coefficients of

$$(1/\sqrt{1-2z}) \exp(1/\sqrt{1-2z}-1) \quad \text{and} \quad (1/(1-2z)) \exp(1/\sqrt{1-2z}-1)$$

(cf. (1), (1′)). It is given in [6], in 25.3, in formula 25.35, in terms of Bessel—Wright functions. The asymptotic expansion of Bessel—Wright functions is given in [5], in 21.5, in formula 21.107. Comparing them we get (3) and (3′). By the theorem concerning the ratio of functions expanded in asymptotic power series ([4], 4.4, Thm. 5—6) we have the following expansions in the powers of $n^{1/3}$:

$$\psi_{n+1}/\psi_n \sim 2n + 2^{1/3} n^{1/3} + c_1 + \dots, \quad \varphi_{n+1}/\varphi_n \sim 2n + 2^{1/3} n^{1/3} + d_1 + \dots.$$

Now (4) and (4′) follow easily.

In order to prove Theorem 3 we recall a well-known theorem from probability theory and prove an easy lemma.

Lemma. *The polynomials*

$$P_m(t) = \sum_{k=0}^{m} \psi(m,k)t^k \quad and \quad Q_m(t) = \sum_{k=0}^{m} \varphi(m,k)t^k$$

have m distinct, real, negative roots.

Proof of the lemma. We prove the statement by mathematical induction. It holds for $P_0(t)=Q_0(t)=1$. By (2) and (2') we have

$$P_m(x) = (2m-1+x)P_{m-1}(x)+xP'_{m-1}(x), \tag{7}$$

$$Q_m(x) = (2m+x)Q_{m-1}(x)+xQ'_{m-1}(x). \tag{7'}$$

Let the roots of $P_{m-1}(x)$ be $z_1<z_2<\ldots<z_{m-1}<0$ by hypothesis. There are $m-2$ roots of P_m by Rolle's theorem in (z_1, z_{m-1}). There are two other roots by

$$P_m(z_{m-1}) = z_{m-1}P'_m(z_{m-1}) < 0 \quad \text{and} \quad P_m(0) = (2m-1)!! \quad \text{(see (2))},$$

and

$$\operatorname{sign} P_m(z_1) = \operatorname{sign} z_1 P'_{m-1}(z_1) = -\operatorname{sign} P'_{m-1}(-\infty) = -\operatorname{sign} P_m(-\infty).$$

A similar method applies for Q_m.

We continue the proof of Theorem 3. Let the roots of $P_n(x)$ be $\{-y_{nk}: k=1, \ldots, n\}$, the roots of $Q_n(x)$ be $\{-x_{nk}: k=1, \ldots, n\}$. We define the independent random variables Y^*_{nk} and X^*_{nk} by

$$P(Y^*_{nk}=0) = y_{nk}/(1+y_{nk}), \quad P(Y^*_{nk}=1) = 1/(1+y_{nk}),$$

$$P(X^*_{nk}=0) = x_{nk}/(1+x_{nk}), \quad P(X^*_{nk}=1) = 1/(1+x_{nk}).$$

Let $Z^*_n=\sum_k Y^*_{nk}$, $S^*_n=\sum_k X^*_{nk}$, F_{nk} and E_{nk} the distribution function of X^*_{nk}, Y^*_{nk}. Using (7), (7') we have

$$E(Z^*_n) = \sum_k \frac{1}{1+y_{nk}} = \left.\frac{P'_n(x)}{P_n(x)}\right|_{x=1} = \left.\frac{P_{n+1}(x)-(2n+1+x)P_n(x)}{xP_n(x)}\right|_{x=1} =$$

$$= \psi_{n+1}/\psi_n-(2n+2),$$

$$E(S^*_n) = \varphi_{n+1}/\varphi_n-(2n+3),$$

$$D^2(Z^*_n) = E\big(Z^*_n-E(Z^*_n)^2\big)^2 = \sum_k \left(\frac{1}{1+y_{nk}}-\frac{1}{(1+y_{nk})^2}\right) = \tag{8}$$

$$= \frac{P'_n(x)}{P_n(x)}+\left.\left(\frac{P'_n(x)}{P_n(x)}\right)'\right|_{x=1} = \frac{-\psi^2_{n+1}+\psi_n\psi_{n+2}-2\psi_n\psi_{n+1}-\psi^2_n}{\psi^2_n},$$

(we used (7) twice),

$$D^2(S^*_n) = E\big(S^*_n-E(S^*_n)^2\big)^2 = (-\varphi^2_{n+1}+\varphi_n\varphi_{n+2}-2\varphi_n\varphi_{n+1}-\varphi^2_n)/\varphi^2_n. \tag{8'}$$

Let us define

$$(9) \qquad Z_n = (1/D(Z_n^*))(Z_n^* - E(Z_n^*)) = \sum_k (1/D(Z_n^*))(X_{nk}^* - E(X_{nk}^*)),$$

$$(9') \qquad S_n = (1/D(S_n^*))(S_n^* - E(S_n^*)) = \sum_k (1/D(S_n^*))(Y_{nk}^* - E(Y_{nk}^*)).$$

From (4), (4′), (8), (8′) we get $D(Z_n^*) \to \infty$, $D(S_n^*) \to \infty$. We are in a position to apply the Lindeberg—Feller Theorem ([3], p. 295) for Z_n^* and S_n^*, since

$$|X_{nk}^* - E(X_{nk}^*)| \leqq 1, \quad |Y_{nk}^* - E(Y_{nk}^*)| \leqq 1,$$

and for a number n large enough

$$\sum_k \int_{|x| \geqq \varepsilon} x^2 \, dF_{nk}(x) = 0, \quad \sum_k \int_{|x| \geqq \varepsilon} x^2 \, dE_{nk}(x) = 0.$$

Since the generating function of a sum of independent random variables is the product of the generating functions,

$$\prod_{k=1}^{n} \frac{x + x_{nk}}{1 + x_{nk}} = \frac{P_n(x)}{P_n(1)}, \quad \prod_{k=1}^{n} \frac{x + y_{nk}}{1 + y_{nk}} = \frac{Q_n(x)}{Q_n(1)},$$

we have $P(Z_n^* = a) = \psi(n, a)/\psi_n$, $P(S_n^* = a) = \varphi(n, a)/\varphi_n$. Now the theorem is proved by (9), (9′).

References

[1] L. H. Harper, Stirling behavior is asymptotically normal, *Ann. Math. Stat.*, **38** (1967), 410—414.

[2] D. C. Kurtz, A note on concavity properties of triangular arrays of numbers, *J. Combin. Theory*, **13** (1972), 135—139.

[3] M. Loéve, *Probability Theory* (Second Edition), Van Nostrand (Princeton, 1960).

[4] E. J. Riekstins, *Asymptotic expansions of integrals. I*, Zinatne (Riga, 1974). (Russian)

[5] E. J. Riekstins, *Asymptotic expansions of integrals. II*, Zinatne (Riga, 1977). (Russian)

[6] E. J. Riekstins, *Asymptotic expansions of integrals. III*, Zinatne (Riga, 1981). (Russian).

[7] Z. I. Szabó, Hilbert's fourth problem. I, *Adv. in Math.*, **59** (1986), 185—301.

[8] L. A. Székely, Holiday numbers: sequences resembling to the Stirling numbers of second kind, *Acta Sci. Math.*, **48** (1985), 459—467.

EÖTVÖS LORÁND UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE
MÚZEUM KRT. 6—8
1088 BUDAPEST, HUNGARY

Acta Sci. Math., 51 (1987), 371—386

p-algebras with Stone congruence lattices

T. KATRIŇÁK and S. EL-ASSAR

1. Introduction. In [12] we have described by means of subdirect factorization congruence distributive algebras A whose congruence lattices $\mathrm{Con}(A)$ are atomic, Boolean, or Stonean. The purpose of this paper is to give an intrinsic characterization of those quasi-modular p-algebras whose congruence lattices are atomic, Stonean or relatively Stonean. To obtain this we use the representation of congruence relations of quasi-modular p-algebras in terms of congruence pairs. That means (see [11]), that every congruence relation $\alpha \in \mathrm{Con}(L)$ of a quasi-modular p-algebra L can be uniquely represented by a congruence pair (α_B, α_D), where α_B is a (Boolean) congruence relation of $B(L)$ and α_D a (lattice) congruence relation of $D(L)$.

We start with a description of congruence pairs corresponding to (relative) pseudocomplements in the lattice $\mathrm{Con}(L)$ (Theorem 1). By way of application, we characterize those quasi-modular p-algebras with atomic congruence lattices (Theorems 2, 3 and 4). As a second application we provide a characterization of (relative) Stone congruence lattices of quasi-modular p-algebras (Theorems 5, 6 and 11). Analogous, but deeper results, are obtained for distributive p-algebras (Theorems 7, 8 and 12).

2. Preliminaries. A (modular, distributive) *p-algebra* or *pseudocomplemented lattice* is an algebra $(L; \vee, \wedge, {}^*, 0, 1)$ in which the deletion of the unary operation * yields a bounded (modular, distributive) lattice and * is the operation of pseudocomplementation, that is, $x \leqq a^*$ if and only if $a \wedge x = 0$. A p-algebra is said to be *quasi-modular* if it satisfies the identity

$$[(x\wedge y)\vee z^{**}]\wedge x = (x\wedge y)\vee(z^{**}\wedge x).$$

The variety of quasi-modular p-algebras properly contains the class of modular p-algebras and is properly contained in the class of p-algebras satisfying the identity

$$x = x^{**}\wedge(x\vee x^*).$$

Receivde May 25, 1984.

If, for any p-algebra L, we write

$$B(L) = \{x \in L\colon\ x = x^{**}\} \quad \text{and} \quad D(L) = \{x \in L\colon\ x^* = 0\}$$

then $(B(L); \boldsymbol{\vee}, \wedge, {}^*, 0, 1)$ is a Boolean algebra (of closed elements) when $a \boldsymbol{\vee} b$ is defined to be $(a^* \wedge b^*)^*$, for any pair $a, b \in B(L)$, and $D(L)$ is a filter in L (of dense elements). By a congruence relation of a p-algebra we mean a lattice congruence of L preserving *. The relation γ of L defined by $a \equiv b(\gamma)$ if and only if $a^* = b^*$ is a congruence relation of L, called the *Glivenko congruence* of L, and $L/\gamma \cong B(L)$. The lattice $\mathrm{Con}(L)$ of all congruence relations of a p-algebra L is algebraic and distributive, which implies that $\mathrm{Con}(L)$ is a distributive p-algebra. The least and greatest elements of $\mathrm{Con}(L)$ will be denoted by Δ and ∇, respectively.

A distributive p-algebra L in which the identity

$$x^* \vee x^{**} = 1$$

holds is called a *Stone algebra (lattice)*. A *relative Stone algebra (lattice)* is a distributive lattice in which every interval $[a, b]$ is a Stone lattice.

A double p-algebra is an algebra $(L; \vee, \wedge, {}^*, {}^+, 0, 1)$ in which the deletion of $^+$ gives a p-algebra and the deletion of * gives a dual p-algebra, that is $a \vee x = 1$ if and only if $x \geqq a^+$. The relation Φ of L defined by

$$a \equiv b(\Phi) \quad \text{if and only if} \quad a^* = b^* \quad \text{and} \quad a^+ = b^+$$

is a congruence relation of L, called the *determination congruence*. It is known that a double p-algebra is *regular* (that is, any two congruence relations of L having a class in common are the same) if and only $\Phi = \Delta$ (see [16]).

A special class of distributive p-algebras is formed by the *Heyting algebras* $(L; \vee, \wedge, *, 0, 1)$, where $(L; \vee, \wedge, 0, 1)$ is a bounded lattice and $x \wedge y \leqq z$ if and only if $y \leqq x * z$. Then $x^* = x * 0$ plays the role of a pseudocomplement of x. It is easy to verify that $\mathrm{Con}(L)$ of a p-algebra L is even a Heyting algebra.

A lattice with 0 is called *atomic*, if for every $a \neq 0$ there exists an atom $p \leqq a$.

We refer to [1], [8] or [10] for the standard results about p-algebras and to [1], [9] or [16] for the standard results about double p-algebras. For general lattice-theoretic terminology, notation and results we follow G. Grätzer [6].

3. Congruence pairs. Let $(L; \vee, \wedge, {}^*, 0, 1)$, henceforth simply L, be a quasi-modular p-algebra. Let $\mathrm{Con}(L)$ denote the lattice of congruence relations of L. Since $\mathrm{Con}(L)$ is a Heyting algebra, there exists a complete Boolean algebra $B(\mathrm{Con}(L))$ of closed elements (congruences) and the filter of dense elements (congruences) $D(\mathrm{Con}(L))$. We shall also consider $\mathrm{Con}(B(L))$, the lattice of (Boolean) congruence relations of $B(L)$ and $\mathrm{Con}(D(L))$, the lattice of (lattice) congruence relations of $D(L)$.

Having $\Theta\in\mathrm{Con}(L)$, the restrictions $\Theta_B=\Theta|B(L)$ and $\Theta_D=\Theta|D(L)$ are congruence relations of $B(L)$ and $D(L)$, respectively. Hence, there exists an isotone map $\Theta\mapsto(\Theta_B, \Theta_D)$ from $\mathrm{Con}(L)$ into $\mathrm{Con}(B(L))\times\mathrm{Con}(D(L))$. The following definition is crucial (see also [11], [4]).

A pair $(\Theta_1, \Theta_2)\in\mathrm{Con}(B(L))\times\mathrm{Con}(D(L))$ is said to be a *congruence pair* of L if the following condition holds: $a\in B(L)$, $u\in D(L)$, $u\geqq a$ in L, and $a\equiv 1(\Theta_1)$ imply that $u\equiv 1(\Theta_2)$.

Theorem A (see [11, Theorem 1]). *Every congruence relation Θ of a quasi-modular p-algebra L determines a congruence pair (Θ_B, Θ_D) and, conversely, every congruence pair (Θ_1, Θ_2) of L determines a unique congruence relation Θ of L having the property that $\Theta_B=\Theta_1$ and $\Theta_D=\Theta_2$. Moreover, $x\equiv y(\Theta)$ if and only if $x^*\equiv y^*(\Theta_1)$ and $x\vee x^*=y\vee y^*(\Theta_2)$.*

In what follows we shall often identify $\Theta\in\mathrm{Con}(L)$ with the corresponding congruence pair (Θ_B, Θ_D). If there is no danger of confusion, we shall omit the subscripts in notation of some congruence pairs, e.g. $\Delta=(\Delta, \Delta)$, $\nabla=(\nabla, \nabla)$, (Δ, α).

Clearly, having $\alpha\in\mathrm{Con}(B(L))$, there exists $\mathrm{Ker}\,\alpha=J\in I(B(L))$ (=the lattice of all ideals of $B(L)$) such that $\alpha=\Theta[J]$. Similarly, for $\beta\in\mathrm{Con}(D(L))$, $\mathrm{Ker}\,\beta=\{x\in D(L)\colon x\equiv 1(\beta)\}$ is a filter of $D(L)$, i.e. $\mathrm{Ker}\,\beta\in F(D(L))$.

Given a quasi-modular p-algebra L, there is a map $\varphi(L)\colon B(L)\to F(D(L))$ defined as follows:

$$a\varphi(L)=\{x\in D(L)\colon x\geqq a^*\}=[a^*)\cap D(L).$$

This map proved instrumental in characterizing the quasi-modular p-algebras (see [13]). We shall need the following result.

Theorem B (see [13, Theorem 3]). *In a quasi-modular p-algebra L, the map $\varphi(L)\colon B(L)\to F(D(L))$ is a $\{0, 1, \vee\}$-homomorphism.*

Now, we can reformulate the definition of a congruence pair.

Lemma 1. *Let L be a quasi-modular p-algebra and let $(\Theta_1, \Theta_2)\in\mathrm{Con}(B(L))\times\mathrm{Con}(D(L))$. Then (Θ_1, Θ_2) is a congruence pair if and only if $J\varphi(L):=\bigcup(a\varphi(L)\colon a\in J)\subseteq\mathrm{Ker}\,\Theta_2$, where $J=\mathrm{Ker}\,\Theta_1$.*

Proof. Clearly, $a\in J=\mathrm{Ker}\,\Theta_1$ if and only if $a^*\equiv 1(\Theta_1)$. Therefore, $J\varphi(L)\subseteq\mathrm{Ker}\,\Theta_2$ if and only if (Θ_1, Θ_2) is a congruence pair.

From Lemma 1 we see that for every $\Theta_1\in\mathrm{Con}(B(L))$ with $J=\mathrm{Ker}\,\Theta_1$ there exists a smallest $\delta(\Theta_1)\in\mathrm{Con}(D(L))$ such that $J\varphi(L)\subseteq\mathrm{Ker}\,\delta(\Theta_1)$. That means, (Θ_1, Θ_2) is a congruence pair of L if and only if $\Theta_2\geqq\delta(\Theta_1)$. Dually, for every $\Theta_2\in\mathrm{Con}(D(L))$ there exists a largest ideal $J\in I(B(L))$ such that $J\varphi(L)\subseteq\mathrm{Ker}\,\Theta_2$, i.e. $(\Theta[J], \Theta_2)$ is a congruence pair. Notation: $\tau(\Theta_2)=\Theta[J]$. Evidently, (Θ_1, Θ_2) is a congruence pair of L if and only if $\tau(\Theta_2)\geqq\Theta_1$.

An abstract description of the lattice of all congruence pairs of quasi-modular p-algebras can be found in [4]. In the next theorem we give a description of (relative) pseudocomplements in $\mathrm{Con}(L)$ by means of congruence pairs.

Theorem 1. *Let L be a quasi-modular p-algebra and let $\alpha, \beta \in \mathrm{Con}(L)$. Then* $(\alpha_B \vee \beta_B, \alpha_D \vee \beta_D)$, $(\alpha_B \wedge \beta_B, \alpha_D \wedge \beta_D)$ *and* $(\alpha_B * \beta_B \wedge \tau(\alpha_D * \beta_D), \alpha_D * \beta_D)$ *are congruence pairs of* $\alpha \vee \beta$, $\alpha \wedge \beta$ *and* $\alpha * \beta$, *respectively. In particular,*

$$(\alpha_B^* \wedge \tau(\alpha_D), \alpha_D^*) \quad \textit{and} \quad ((\alpha_B^* \wedge \tau(\alpha_D^*))^* \wedge \tau(\alpha_D^{**}), \alpha_D^{**})$$

are congruence pairs of α^* *and* α^{**}, *respectively.*

Proof. Clearly, $(\alpha \vee \beta)_B \geqq \alpha_B \vee \beta_B$ and $(\alpha \vee \beta)_D \geqq \alpha_D \vee \beta_D$. Assume $a \equiv b(\alpha \vee \beta)$ for $a, b \in B(L)$. Then there exists a finite sequence $a = z_0, \ldots, z_n = b$ such that $z_{i-1} \equiv z_i(\alpha)$ or $z_{i-1} \equiv z_i(\beta)$ for every $i = 1, \ldots, n$. Therefore $z_{i-1}^{**} \equiv z_i^{**}(\alpha)$ or $z_{i-1}^{**} \equiv \equiv z_i^{**}(\beta)$, which implies $a \equiv b(\alpha_B \vee \beta_B)$. Hence $(\alpha \vee \beta)_B = \alpha_B \vee \beta_B$. A similar argument yields $(\alpha \vee \beta)_D = \alpha_D \vee \beta_D$, $(\alpha \wedge \beta)_B = \alpha_B \wedge \beta_B$ and $(\alpha \wedge \beta)_D = \alpha_D \wedge \beta_D$.

It is easy to verify that $(\alpha_B * \beta_B \wedge \tau(\alpha_D * \beta_D), \alpha_D * \beta_D)$ is a congruence pair of L. Clearly,

$$(\alpha_B, \alpha_D) \wedge (\alpha_B * \beta_B \wedge \tau(\alpha_D * \beta_D), \alpha_D * \beta_D) \leqq (\beta_B, \beta_D).$$

Assume $(\alpha_B, \alpha_D) \wedge (\eta_B, \eta_D) \leqq (\beta_B, \beta_D)$ in $\mathrm{Con}(L)$. Therefore, $\eta_B \leqq \alpha_B * \beta_B$ and $\eta_D \leqq \alpha_D * \beta_D$. Since (η_B, η_D) is a congruence pair, we have $\eta_B \leqq \tau(\eta_D) \leqq \tau(\alpha_D * \beta_D)$. Hence $(\eta_B, \eta_D) \leqq (\alpha_B * \beta_B \wedge \tau(\alpha_D * \beta_D), \alpha_D * \beta_D)$. The last part of Theorem can be established in the same way because $(\alpha_B, \alpha_D)^* = (\alpha_B, \alpha_D) * (\Delta, \Delta)$.

Corollary 1 (see [1, Theorem 2]). *Let L be a quasi-modular p-algebra. Then* $\mathrm{Con}(D(L)) \cong [\Delta, \gamma]$, *where γ is the Glivenko congruence.*

Proof. Consider the map $\alpha_2 \mapsto (\Delta, \alpha_2)$ from $\mathrm{Con}(D(L))$ into $\mathrm{Con}(L)$. Since $\gamma = (\Delta, \nabla)$, we see that this map is an isomorphism between $\mathrm{Con}(D(L))$ and $[\Delta, \gamma]$.

Corollary 2. *Let L be a quasi-modular p-algebra. Then* $\mathrm{Con}(B(L)) \cong [\gamma, \nabla]$.

Proof. Consider the map $\alpha_1 \mapsto (\alpha_1, \nabla)$ from $\mathrm{Con}(B(L))$ into $\mathrm{Con}(L)$. This map is an isomorphism between $\mathrm{Con}(B(L))$ and $[\gamma, \nabla]$.

4. Atomic congruence lattices. In [12] we have extended Tanaka's result [15, Theorem 1].

Theorem C. *Let A be a congruence distributive algebra. Then the following conditions are equivalent:*

(i) $\mathrm{Con}(A)$ *is atomic;*

(ii) $D(\mathrm{Con}(A))$ *is a principal filter;*

(iii) $B(\mathrm{Con}\,(A))$ *is atomic and every dual atom of* $B(\mathrm{Con}\,(A))$ *is completely meet-irreducible in* $\mathrm{Con}\,(A)$;

(iv) $\mathrm{Con}\,(A)$ *satisfies the (infinite) identity*

$$\wedge(x_i^{**}\colon i\in I) = (\wedge(x_i\colon i\in I))^{**}.$$

Lemma 2. *Let L be a quasi-modular p-algebra. Then* $\alpha=(\alpha_B, \alpha_D)$ *is an atom of* $\mathrm{Con}\,(L)$ *if and only if*

(i) $\alpha_B=\Delta$ *and* α_D *is an atom of* $\mathrm{Con}\,(D(L))$

or

(ii) $\alpha_D=\Delta$, $\alpha_B\leqq\tau(\Delta)$ *and* α_B *is an atom of* $\mathrm{Con}\,(B(L))$.

Proof. Suppose that (α_B, α_D) is an atom of $\mathrm{Con}\,(L)$. Two cases can arise: $\alpha_D\neq\Delta$ or $\alpha_D=\Delta$. In the first event $(\Delta, \alpha_D)\leqq\alpha$, whence $\alpha=(\Delta, \alpha_D)$ and α_D is an atom of $\mathrm{Con}\,(D(L))$. In the second case we obtain (ii). The converse is trivial.

Theorem 2. *Let L be a quasi-modular p-algebra. Then* $\mathrm{Con}\,(L)$ *is atomic if and only if*

(i) $\mathrm{Con}\,(D(L))$ *is atomic*

and

(ii) $\{a\in B(L)\colon a\varphi(L)=[1)\}$ *is an atomic ideal of* $B(L)$, *i.e. it is an atomic lattice.*

Proof. Assume that $\mathrm{Con}\,(L)$ is atomic. Therefore, $[\Delta, \gamma]$ is atomic as well. By Corollary 1 of Theorem 1 we obtain (i). Take $0\neq a\in B(L)$ with $a\varphi(L)=[1)$. By Lemma 1, $(\Theta[(a]], \Delta)\in\mathrm{Con}\,(L)$. There exists an atom $\alpha\in\mathrm{Con}\,(L)$ with $\alpha=(\alpha_B, \alpha_D)\leqq\leqq(\Theta[(a]], \Delta)$. Hence $\alpha_D=\Delta$, $\alpha_B\leqq\tau(\Delta)$ and α_B is an atom of $\mathrm{Con}\,(B(L))$ (Lemma 2). Thus $\mathrm{Ker}\,\alpha_B=(b]$ and b is an atom of $B(L)$ with $b\varphi(L)=[1)$.

Conversely, assume (i) and (ii). Take $\Delta\neq\alpha=(\alpha_B, \alpha_D)$ from $\mathrm{Con}\,(L)$. Two cases can occur: $\alpha_D\neq\Delta$ or $\alpha_D=\Delta$. In the first case, there is by (i) an atom $\beta\in\mathrm{Con}\,(D(L))$ with $\beta\leqq\alpha_D$. Hence (Δ, β) is by Corollary 1 to Theorem 1 an atom of $\mathrm{Con}\,(L)$ and $(\Delta, \beta)\leqq(\alpha_B, \alpha_D)$. In the second case, $\Delta\neq\alpha_B\leqq\tau(\Delta)$. There exists an atom $a\in J=\mathrm{Ker}\,\alpha_B$ by (ii). Hence $(\Theta[(a]], \Delta)$ is an atom of $\mathrm{Con}\,(L)$ (Lemma 2) and $(\Theta[(a]], \Delta)\leqq\alpha$.

Lemma 3. *Let K be an ideal of a Boolean algebra B and let K be an atomic sublattice of B. Let J be the ideal of B generated by all atoms of K. Then* $J^*=K^*$ *in the lattice $I(B)$ of all ideals of B.*

Proof. Clearly $J\subseteqq K$. Therefore, $J^*\supseteqq K^*$. Take $b\in J^*$. If $(b]\cap K\neq(0]$, then there exists an atom $a\in K$ such that $a\leqq b$. Hence $a\in J\cap J^*=(0]$, a contradiction. Thus, $K\cap J^*=(0]$, which implies $J^*\subseteqq K^*$. So, $J^*=K^*$.

Theorem 3. *Let L be a quasi-modular p-algebra. Then* $(\beta_1, \beta_2)\in\mathrm{Con}\,(L)$ *is the smallest element of* $D(\mathrm{Con}\,(L))$ *if and only if*

(i) β_2 *is the smallest element of* $D(\mathrm{Con}\,(D(L)))$
and

(ii) *the ideal* $K=\{a\in B(L)\colon a\varphi(L)=[1)\}$ *of* $B(L)$ *is atomic and* $\beta_1=\Theta[J]$, *where* J *is the ideal of* $B(L)$ *generated by all atoms of* K.

Proof. Let $(\beta_1,\beta_2)\in\mathrm{Con}\,(L)$ be the smallest element of $D(\mathrm{Con}\,(L))$. It is easy to verify that $(\tau(\Delta),\alpha)\in\mathrm{Con}\,(L)$ for every $\alpha\in\mathrm{Con}\,(D(L))$. Moreover, $(\tau(\Delta),\alpha)^*=\Delta$ if and only if $\alpha\in D(\mathrm{Con}\,(D(L)))$. Therefore, $(\beta_1,\beta_2)\leqq(\tau(\Delta),\alpha)$ for every $\alpha\in D(\mathrm{Con}\,(D(L)))$. Thus β_2 is the smallest element of $D(\mathrm{Con}\,(D(L)))$ and $\beta_1\leqq\tau(\Delta)$. Since $\Delta=(\beta_1,\beta_2)^*=(\beta_1^*\wedge\tau(\Delta),\Delta)$, we see that $\beta_1^*\leqq\tau(\Delta)^*$. But $\beta_1\leqq\tau(\Delta)$ implies $\beta_1^*\geqq\tau(\Delta)^*$. Hence $\beta_1^*=\tau(\Delta)^*$. Clearly, $\beta_1=\Theta[M]$ and $\tau(\Delta)=\Theta[K]$, where M is an ideal of $B(L)$ and $M\subseteq K$. According to Theorems C and 2, K is atomic. Without difficulties one can check that M contains all atoms of K, as $\beta_1^*=\tau(\Delta)^*$. Let J denote the ideal of $B(L)$ generated by all atoms of K. Lemma 3 yields $\beta_1^*=\tau(\Delta)^*=\Theta[J]^*$. Now, $(\Theta[J],\beta_2)^*=(\beta_1^*\wedge\tau(\Delta),\beta_2^*)=\Delta$ implies $\beta_1\leqq\Theta[J]$. Eventually, $\beta_1=\Theta[J]$.

Conversely, let L satisfy (i) and (ii). Take $\beta_1=\Theta[J]$ from $\mathrm{Con}\,(B(L))$ and $\beta_2\in\mathrm{Con}\,(D(L))$ as defined in (i) and (ii). Clearly, $(\beta_1,\beta_2)\in\mathrm{Con}\,(L)$, as $\beta_1\leqq\tau(\Delta)$. By Lemma 3, $\beta_1^*=\tau(\Delta)^*$. Therefore, $(\beta_1,\beta_2)^*=\Delta$; that means $(\beta_1,\beta_2)\in D(\mathrm{Con}\,(L))$. Consider $(\alpha_1,\alpha_2)\in D(\mathrm{Con}\,(L))$. Since $\alpha_2^*=\Delta$, we have $\beta_2\leqq\alpha_2$. In addition, $\alpha_1^*\wedge\wedge\tau(\Delta)=\Delta$. Hence $\alpha_1^*\leqq\tau(\Delta)^*=\beta_1^*$. Clearly $\alpha_1=\Theta[M]$ for some ideal M of $B(L)$. We claim that $M\supseteq J$. Really, if $J\nsubseteq M$, then there exists an atom $a\in J-M$ and $a\in M^*$. That means $\Theta[(a]]\leqq\alpha_1^*\wedge\beta_1=\Delta$, a contradiction. Therefore, $J\subseteq M$, as claimed. Hence $\beta_1\leqq\alpha_1$, and (β_1,β_2) is the smallest dense congruence relation of L. The proof is complete.

Lemma 4. *Let L be a quasi-modular p-algebra. Let* $a\in B(L)$ *with* $a\varphi(L)=[1)$. *Then* $(\Theta[(a]],\Delta)\in B(\mathrm{Con}\,(L))$.

Proof. Since $\Theta[(a]]\leqq\tau(\Delta)$, we see that $(\Theta[(a]],\Delta)\in\mathrm{Con}\,(L)$. By Theorem 1, $(\Theta[(a]],\Delta)^{**}=(\Theta[(a]]^{**}\wedge\tau(\Delta),\Delta)$. Since $\Theta[(a]]^{**}=\Theta[(a]]$, the proof is complete.

Theorem 4. *Let L be a quasi-modular p-algebra. Then* $B(\mathrm{Con}\,(L))$ *is atomic if and only if*

(i) $B(\mathrm{Con}\,(D(L)))$ *is atomic*
and

(ii) $\{a\in B(L)\colon a\varphi(L)=[1)\}$ *is an atomic ideal of* $B(L)$.

Proof. Assume that $B(\mathrm{Con}\,(L))$ is atomic. Let $\Delta\neq\alpha\in B(\mathrm{Con}\,(D(L)))$. Therefore, $(\Delta,\alpha)\in\mathrm{Con}\,(L)$. Clearly, $(\Delta,\alpha)^{**}=(\tau(\alpha^*)^*\wedge\tau(\alpha),\alpha)\neq\Delta$, by Theorem 1. By assumption there exists an atom (β_1,β_2) of $B(\mathrm{Con}\,(L))$ such that $(\beta_1,\beta_2)\leqq\leqq(\tau(\alpha^*)^*\wedge\tau(\alpha),\alpha)$. Evidently, $\beta_2^{**}=\beta_2$ in $\mathrm{Con}\,(D(L))$. Hence $\beta_2\leqq\alpha$. We claim that β_2 is an atom of $B(\mathrm{Con}\,(D(L)))$. First we show that $\beta_2 \neq \Delta$. Assume to the contrary that $\beta_2=\Delta$. Hence $\beta_1 \leqq \tau(\Delta)$. Since $\tau(\Delta) \leqq \tau(\alpha^*)$, we get $\beta_1 \leqq \tau(\alpha^*)^* \leqq \tau(\Delta)^*$. Therefore $\beta_1=\Delta$, a contradiction. Thus $\beta_2 \neq \Delta$. Take $\Delta \neq \eta \in B(\mathrm{Con}\,(D(L)))$ with $\eta \leqq \beta_2$. Therefore $\Delta \neq (\Delta, \eta) \leqq (\beta_1, \beta_2)$ implies $(\Delta, \eta)^{**}=(\beta_1, \beta_2)$, as (β_1, β_2) is an atom of $B(\mathrm{Con}\,(L))$. But $(\Delta, \eta)^{**}=(\tau(\eta^*)^* \wedge \tau(\eta), \eta)$. Hence $\eta=\beta_2$ and β_2 is an atom of $B(\mathrm{Con}\,(D(L)))$, as claimed. The second part of Theorem follows from Lemma 4.

Conversely, let L satisfy (i) and (ii). Consider $\Delta \neq (\alpha_1, \alpha_2) \in B(\mathrm{Con}\,(L))$. Clearly $\alpha_2=\alpha_2^{**}$ in $\mathrm{Con}\,(D(L))$. If $\alpha_2=\Delta$ then $\alpha_1 \leqq \tau(\Delta)$, and $\alpha_1=\Theta[J]$, where J is an ideal of $\{a \in B(L)\colon a\varphi(L)=[1)\}$. By (ii) there exists an atom $a \in J$. Put $\beta_1=\Theta[(a]]$ in $\mathrm{Con}\,(B(L))$. Clearly $(\beta_1, \Delta)^{**}=(\beta_1, \Delta) \leqq (\alpha_1, \Delta)$, using Lemma 4. Thus (β_1, Δ) is an atom of $B(\mathrm{Con}\,(L))$. Assume $\alpha_2 \neq \Delta$. Then there exists an atom $\beta_2 \leqq \alpha_2$ in $B(\mathrm{Con}\,(D(L)))$ by (i). Since $(\Delta, \beta_2) \leqq (\alpha_1, \alpha_2)$, we see that

$$(\Delta, \beta_2)^{**} = (\tau(\beta_2^*)^* \wedge \tau(\beta_2), \beta_2) \leqq (\alpha_1, \alpha_2) = (\alpha_1, \alpha_2)^{**}.$$

It remains to verify that $(\Delta, \beta_2)^{**}$ is an atom of $B(\mathrm{Con}\,(L))$. Really, suppose that there exists $\Delta \neq (\eta_1, \eta_2) \in B(\mathrm{Con}\,(L))$ with $(\eta_1, \eta_2) \leqq (\Delta, \beta_2)^{**}$. Two cases can arise: $\eta_2 \neq \Delta$ or $\eta_2=\Delta$. But $\eta_2 \neq \Delta$ implies $\beta_2=\eta_2$. Moreover, $(\Delta, \beta_2) \leqq (\eta_1, \eta_2) \leqq (\Delta, \beta_2)^{**}$ implies $(\eta_1, \eta_2)^{**}=(\eta_1, \eta_2)=(\Delta, \beta_2)^{**}$. Assume $\eta_2=\Delta$. Therefore, $\eta_1 \leqq \tau(\Delta) \leqq \tau(\beta_2^*)$. Similarly as above, $\eta_1 \leqq \tau(\beta_2^*)^* \leqq \tau(\Delta)^*$, which implies $\eta_1=\Delta$, a contradiction. Thus, $(\Delta, \beta_2)^{**}$ is an atom of $B(\mathrm{Con}\,(L))$ and the proof is complete.

5. Stonean congruence lattices.

Lemma 5. *Let L be a Stone lattice and $a \in L$. Then $[0, a]$ is also a Stone lattice.*

The proof is straightforward (see [8, 2.11]).

Theorem 5. *Let L be a quasi-modular p-algebra. Then $\mathrm{Con}\,(L)$ is a Stone lattice if and only if*

(i) $\mathrm{Con}\,(D(L))$ *is a Stone lattice,*

(ii) *if* $(\alpha_1, \alpha_2) \in \mathrm{Con}\,(L)$ *then* $\mathrm{Ker}\,(\alpha_1^* \wedge \tau(\alpha_2^*))=(a]$ *for some* $a \in B(L)$,

(iii) *if* $\alpha \in \mathrm{Con}\,(D(L))$ *then* $\tau(\alpha^{**}) \geqq (\tau(\alpha)^* \wedge \tau(\alpha^*))^*$.

Proof. Suppose that $\mathrm{Con}\,(L)$ is a Stone lattice. The condition (i) follows directly from Lemma 5 and Corollary 1 to Theorem 1. Take now $(\alpha_1, \alpha_2) \in \mathrm{Con}\,(L)$. By Theorem 1 and the hypothesis,

$$\nabla = (\alpha_1, \alpha_2)^* \vee (\alpha_1, \alpha_2)^{**} = (\alpha_1^* \wedge \tau(\alpha_2^*), \alpha_2^*) \vee ((\alpha_1^* \wedge \tau(\alpha_2^*))^* \wedge \tau(\alpha_2^{**}), \alpha_2^{**}).$$

Therefore,

$$(\alpha_1^* \wedge \tau(\alpha_2^*)) \vee [(\alpha_1^* \wedge \tau(\alpha_2^*))^* \wedge \tau(\alpha_2^{**})] = \nabla.$$

Consequently, $(\alpha_1^*\wedge\tau(\alpha_2^*))\vee(\alpha_1^*\wedge\tau(\alpha_2^*))^*=\nabla$. Hence, $(\alpha_1^*\wedge\tau(\alpha_2^*))^*$ is a complement of $\alpha_1^*\wedge\tau(\alpha_2^*)$ in $\mathrm{Con}\,(B(L))$, and $(\alpha_1^*\wedge\tau(\alpha_2^*))^*=(\alpha_1^*\wedge\tau(\alpha_2^*))^*\wedge\tau(\alpha_2^{**})$. Thus $\tau(\alpha_2^{**})\geqq\geqq(\alpha_1^*\wedge\tau(\alpha_2^*))^*$. As $(\tau(\alpha),\alpha)\in\mathrm{Con}\,(L)$ for every α from $\mathrm{Con}\,(D(L))$, this yields (iii). The condition (ii) follows from the fact that $\alpha_1^*\wedge\tau(\alpha_2^*)=\Theta[J]$ and $(\alpha_1^*\wedge\tau(\alpha_2^*))^*==\Theta[J^*]$ for some $J\in I(B(L))$. By the hypothesis, J^* is a complement of J in $I(B(L))$. It follows that $J=(a]$ and $J=(a^*]$ for some $a\in B(L)$ (see [5] or [7]).

Conversely, suppose that L satisfies (i)—(iii). Take $(\alpha_1,\alpha_2)\in\mathrm{Con}\,(L)$. Clearly $(\alpha_1,\alpha_2)\leqq(\tau(\alpha_2),\alpha_2)$. By Theorem 1 and the hypothesis,

$$(\alpha_1,\alpha_2)^*\vee(\alpha_1,\alpha_2)^{**}=(\alpha_1^*\wedge\tau(\alpha_2^*),\alpha_2^*)\vee(((\alpha_1^*\wedge\tau(\alpha_2^*))^*\wedge\tau(\alpha_2^{**}),\alpha_2^{**})=$$

$$=(\Theta[(a]],\alpha_2^*)\vee(\Theta[(a^*]],\alpha_2^{**})=\nabla,$$

because $(\alpha_1^*\wedge\tau(\alpha_2^*))^*\leqq(\tau(\alpha_2)^*\wedge\tau(\alpha_2^*))^*$. The proof is complete.

Corollary. *Let L be a quasi-modular p-algebra and let* $\mathrm{Con}\,(L)$ *be a Stone lattice. Then for* $(\alpha_1,\alpha_2)\in\mathrm{Con}\,(L)$ *we have*

(i) $(\alpha_1^{**},\alpha_2^{**})\in\mathrm{Con}\,(L)$,

(ii) $(\alpha_1,\alpha_2)^{**}=(a_1^{**},\alpha_2^{**})^{**}=((\alpha_1^*\wedge\tau(\alpha_2^*))^*,\alpha_2^{**})$

and

(iii) $\alpha\in\mathrm{Con}\,(D(L))$ *implies* $\tau(\alpha^{**})\geqq\tau(\alpha)^{**}\vee\tau(\alpha^*)^*$.

Proof. (i) Since, by Theorem 5, $\alpha_1^{**}\leqq\tau(\alpha_2^{**})$, we have $(\alpha_1^{**},\alpha_2^{**})\in\mathrm{Con}\,(L)$. (ii) and (iii) follows from Theorems 1 and 5.

In [12] we have also investigated algebras whose congruence lattices satisfy the (infinite) identity

$$\bigvee(x_i^{**}:\ i\in I)=(\bigvee(x_i:\ i\in I))^{**}. \tag{1}$$

Theorem 6. *Let L be a quasi-modular p-algebra. Then* $\mathrm{Con}\,(L)$ *satisfies the identity* (1) *if and only if*

(i) $\mathrm{Con}\,(L)$ *is a Stone lattice*

and

(ii) $B(\mathrm{Con}\,(L))$ *is finite.*

Proof. Assume that $\mathrm{Con}\,(L)$ satisfies the identity (1). Then by [12, Lemma 2], $\mathrm{Con}\,(L)$ is a Stone lattice and $B(\mathrm{Con}\,(L))$ is atomic. Moreover, [12, Theorem 9] says that L has an irredundant discrete subdirect factorization with finitely subdirectly irreducible factors. Let $\{\alpha_i:\ i\in I\}$ denote the set of all dual atoms of $B(\mathrm{Con}\,(L))$. Then by [12, Theorem 2], $(L/\alpha_i:\ i\in I)$ is the subdirect factorization of L in question. Therefore, every element $x\in L$ can be represented as $(x_i)_{i\in I}$, where $x_i\in L/\alpha_i$ for every $i\in I$. Take now the elements $v=0$ and $u=1$ from L, i.e. the smallest and the largest elements of L, respectively. Since the factorization of L is discrete, there exists a finite subset $I_1 \subseteq I$ such that $\{i \in I: u_i \neq v_i\} = I_1$. Moreover, $0 \leqq x \leqq 1$ implies $u_i = x_i = v_i$ for every $i \in I - I_1$. But the factorization $(L/\alpha_i: i \in I)$ is irredundant that means $\wedge(\alpha_j: j \in I, i \neq j) \neq \Delta$ for every $i \in I$. Hence $I = I_1$ is finite. $B(\mathrm{Con}(L))$ is an atomic and complete Boolean algebra. Therefore, $B(\mathrm{Con}(L))$ is finite.

Conversely, assume that L satisfies (i) and (ii). Therefore, $\mathrm{Con}(L)$ satisfies the identity $\vee(x_i^{**}: i \in I) = (\vee(x_i: i \in I))^{**}$ for every finite I. According to (ii), $\mathrm{Con}(L)$ enjoys the identity (1) for arbitrary I. The proof is complete.

Deeper results can be obtained for distributive p-algebras. First we recall two results.

Theorem D. *Let L be a distributive lattice with 0. Then L can be embedded in a generalized Boolean lattice B such that every congruence relation of L has one and only one extension to B, that means* $\mathrm{Con}(L) \cong \mathrm{Con}(B)$.

For the proof see [6, Lemma II.4.5].

Theorem E ([10, Theorem 2]). *Every distributive p-algebra can be embedded in a Heyting algebra H of order 3 (i.e. $D(H)$ is relatively complemented) such that*

(i) *every congruence relation of L has one and only one extension to H, i.e.* $\mathrm{Con}(L) \cong \mathrm{Con}(H)$,

(ii) $B(L) = B(H)$

and

(iii) *$D(H)$ is an extension of $D(L)$ such that* $\mathrm{Con}(D(L)) \cong \mathrm{Con}(D(H))$.

For the proof of (ii) and (iii) see the proof of [10, Theorem 2].

Theorem 7. *Let L be a distributive p-algebra. Then* $\mathrm{Con}(L)$ *is a Stone lattice if and only if*

(i) *$D(L)$ is relatively complemented,*

(ii) *the dual lattice $\breve{L}$ is a Stone lattice*

and

(iii) *$B(\breve{L})$ is a complete Boolean algebra.*

Proof. Let $\mathrm{Con}(L)$ be a Stone lattice. Then there exists a Heyting algebra H of order 3 such that L is a subalgebra of the p-algebra H and $\mathrm{Con}(L) \cong \mathrm{Con}(H)$ (Theorem E). It is well known that $\mathrm{Con}(H) \cong F(H)$, that means, every congruence relation of H is uniquely determined by a filter of H. Hence $F(H)$ is a Stone lattice. Now we can apply [7, Satz 9]. Therefore, (a) the dual lattice $\breve{H}$ is a Stone lattice and (b) $B(\breve{H})$ is a complete Boolean algebra. Evidently, $B(\breve{H}) \subseteq B(H) = B(L)$ (Theorem E). Take $a \in L$. There exists a dual pseudocomplement $a^+ \in B(\breve{H})$ of a in H, i.e. $a \vee x = 1$ if and only if $x \geqq a^+$. As $B(\breve{H}) \subseteq B(L)$, $a^+ \in L$ and L is dual pseudocomplemented, that means $\check{L}$ is pseudocomplemented. Moreover, $\check{L}$ is a Stone lattice by (a). Again, $B(\check{L}) \subseteq B(\check{H}) \subseteq B(L)$ implies $B(\check{L}) = B(\check{H})$. Hence, $B(\check{L})$ is a complete Boolean algebra by (b). We have established (ii) and (iii).

Now we shall prove (i). Using (a) we see that H is a double p-algebra. Moreover, by the hypothesis, $D(H)$ is relatively complemented. Therefore, H is a regular double p-algebra (see [9, Theorem 2]). Above we have shown that L is a subalgebra of the double p-algebra H (see also Theorem E). But the regular double p-algebras form a variety (see [9, Theorem 2] or [16]). Hence, again by [9, Theorem 2], L is also regular and this implies that $D(L)$ is relatively complemented.

Conversely, let L satisfy (i)—(iii). Then L is a distributive double p-algebra. By [9, Theorem 2], L is a regular double p-algebra, because $D(L)$ is relatively complemented. According to [9, Theorem 1], L forms a (double) Heyting algebra H. But every congruence relation of L is also a Heyting algebra congruence relation, that means $\mathrm{Con}\,(L) = \mathrm{Con}\,(H)$ (see [10, Lemma 1]). Therefore, $\mathrm{Con}\,(L) \cong F(H) = F(L)$. Now, conditions (ii) and (iii) imply by [7, Satz 9] that $F(L)$ is a Stone lattice. Thus, $\mathrm{Con}\,(L)$ is a Stone lattice and the proof is complete.

For the next Theorem we need the following

Lemma 6. *Let L be a distributive lattice with* 0. *Then* $B(\mathrm{Con}\,(L))$ *is finite if and only if L is finite.*

Proof. By Theorem D there is an extension K of L such that K is a generalized Boolean lattice and $\mathrm{Con}\,(L) \cong \mathrm{Con}\,(K)$. Every congruence relation of K is uniquely determined by its kernel, that means $\mathrm{Con}\,(K) \cong I(K)$. By assumption, $B(I(K))$ is finite. Take $a \in K$. We claim that $(a] \in B(I(K))$. Really, if $J \in I(K)$, then $J^* = \{x \in K \colon x \wedge y = 0 \text{ for every } y \in J\}$. Consider $(a]^*$ and $(a]^{**}$. It suffices to show that $(a]^{**} = (a]$. Clearly $(a] \subseteq (a]^{**}$. Choose $b \in (a]^{**}$. Take $c = a \vee b$ and observe $(c] \in I(K)$. $(c]$ is a Boolean lattice. Since $(a] \vee ((a]^* \wedge (c]) = (c]$, we see that $b \leqq a$ and $(a] = (a]^{**} \in B(I(K))$, as claimed. Hence K is finite, as $B(I(K))$ is finite. Consequently, L is finite. The converse implication is trivial.

Theorem 8. *Let L be a distributive p-algebra. Then* $\mathrm{Con}\,(L)$ *satisfies the identity* (1) *if and only if*

(i) $\mathrm{Con}\,(L)$ *is a Stone lattice,*

(ii) $D(L)$ *is finite*

and

(iii) $\{a \in B(L) \colon a\varphi(L) = [1)\}$ *is finite.*

Proof. Let $\mathrm{Con}\,(L)$ satisfy the identity (1). The condition (i) follows from Theorem 6. Again from Theorem 6 we know that $B(\mathrm{Con}\,(L))$ is finite. Hence, $B(\mathrm{Con}\,(L))$ is atomic. With regard to Theorem 4, $B(\mathrm{Con}\,(D(L)))$ is also atomic and the set of all atoms of $B(\mathrm{Con}\,(L))$ comprises

$$\{(\Delta, \alpha)\in \mathrm{Con}\,(L)\colon\ \alpha \text{ is an atom of } B(\mathrm{Con}\,(D(L)))\}$$

and

$$\{(\Theta[(a]], \Delta)\in \mathrm{Con}\,(L)\colon\ a\varphi(L) = [1) \text{ and } a \text{ is an atom of } B(L)\}.$$

This and Lemma 6 imply (ii), because $B(\mathrm{Con}\,(L))$ is finite. Now we shall establish (iii). Again by the hypothesis the set of atoms $a\in B(L)$ such that $a\varphi(L)=[1)$ is finite. Observe $b\in B(L)$ with $b\varphi(L)=[1)$. We claim that $b=a_1\vee\ldots\vee a_n$, where $a_i\varphi(L)=[1)$ and a_i is an atom of $B(L)$ for every $i=1, \ldots, n$. Let $a\in B(L)$ be a join of atoms a_i of $B(L)$ with $a_i\leqq b$, i.e. $a=a_1\vee\ldots\vee a_n$. Therefore, $a\leqq b$. Then there exists $c\in B(L)$ such that $a\wedge c=0$ and $b=a\vee c$. Hence $b\varphi(L)=a\varphi(L)\vee \vee c\varphi(L)=[1)$ (see Theorem B). Consequently, $c\varphi(L)=[1)$. If $t\leqq c$ and t is an atom of $B(L)$ then by assumption $t\leqq a$. This implies $c=0$. Thus $b=a$, as claimed. Now it is easy to show that $\{a\in B(L)\colon\ a\varphi(L)=[1)\}$ is finite.

Conversely, let L satisfy (i)—(iii). By Theorem 4, $B(\mathrm{Con}\,(L))$ is atomic and the set of all atoms of $B(\mathrm{Con}\,(L))$ is finite. Therefore, the Boolean algebra $B(\mathrm{Con}\,(L))$ is finite. The rest follows from Theorem 6.

Before closing this section we shall generalize Beazer's [1, Theorem 6] (see also [2]). We shall characterize those finite p-algebras, which have the same congruence lattices as the finite distributive p-algebras.

Having an (arbitrary) finite p-algebra L, then $\mathrm{Con}\,(L)$ is a finite distributive lattice, and thus, $\mathrm{Con}\,(L)$ can be considered as a finite double p-algebra. In this case we introduce the ideal $\bar{D}(\mathrm{Con}\,(L))$ of dual dense elements from $\mathrm{Con}\,(L)$, that means, $\alpha\in\bar{D}(\mathrm{Con}\,(L))$ if and only if $\alpha^+=\nabla$.

Theorem 9. *Let L be a finite p-algebra. Then the following statements are equivalent:*

(i) *there exists a finite distributive p-algebra L' such that* $\mathrm{Con}\,(L)\cong\mathrm{Con}\,(L')$*;*
(ii) $D(\mathrm{Con}\,(L))$ *is a Boolean lattice;*
(iii) $\bar{D}(\mathrm{Con}\,(L))$ *is a Boolean lattice;*
(iv) $\mathrm{Con}\,(L)$ *is a regular double p-algebra.*

Proof. By assumption, $\mathrm{Con}\,(L)$ is finite and distributive. Now the equivalence between (ii)—(iv) follows from [9, Theorem 2]. Assume (i). Then there exists a finite Heyting algebra H of order 3 with $\mathrm{Con}\,(H)\cong\mathrm{Con}\,(L)$. Since H is finite, we see that H is a double p-algebra. Eventually, H is regular, because H is of order 3. The same is also true for the dual lattice $\breve{H}$. But $\mathrm{Con}\,(H)\cong F(H)\cong\breve{H}$. Hence $\breve{H}\cong\mathrm{Con}\,(L)$, and (iv) is true. Conversely, assume (iv). Let H denote the dual lattice of $\mathrm{Con}\,(L)$. Clearly, H is also a regular double p-algebra. By [9, Theorem 2] H is in fact a Heyting algebra of order 3. Let L' be H considered as a p-algebra. Then

$\mathrm{Con}(L')\cong F(H)$, by [10, Lemma 1]. Since H is finite, we see that $F(H)\cong \breve{H}\cong$ $\cong \mathrm{Con}(L)$, and (i) is established.

Lemma 7. *Let L be a finite quasi-modular p-algebra. Then* $\bar{D}(\mathrm{Con}(L))=$ $=[\Delta, \gamma]$ *(that means that the Glivenko congruence is the largest dual dense element of* $\mathrm{Con}(L)$*).*

Proof. We know that $\gamma=(\Delta, \nabla)$. Take $\alpha=(\alpha_1, \alpha_2)\in \mathrm{Con}(L)$ with $\nabla=\gamma\vee\alpha$. Therefore, $\alpha_1=\nabla$. As (α_1, α_2) is a congruence pair, we see that $\alpha_2=\nabla$. Now, assume $\alpha\geqq\gamma$ for some $\alpha\in\bar{D}(\mathrm{Con}(L))$. Corollary 2 to Theorem 1 says that $\nabla\neq\alpha\in[\gamma, \nabla]\cong \mathrm{Con}(B(L))$. But $\mathrm{Con}(B(L))\cong B(L)$, as L is finite. Take the complement α' of α in $[\gamma, \nabla]$. But $\alpha'\neq\nabla$ is impossible, because $\alpha\in\bar{D}(\mathrm{Con}(L))$. Hence $\alpha'=\nabla$, which implies $\alpha=\gamma$.

Theorem 10. *Let L be a finite quasi-modular p-algebra. Then there exists a finite distributive p-algebra L' such that* $\mathrm{Con}(L)\cong \mathrm{Con}(L')$ *if and only if* $\mathrm{Con}(D(L))$ *is a Boolean lattice.*

Proof. Corollary 1 to Theorem 1 and Lemma 7 imply that $\bar{D}(\mathrm{Con}(L))=$ $=[\Delta, \gamma]\cong \mathrm{Con}(D(L))$. Hence, by Theorem 9, $\mathrm{Con}(D(L))$ is a Boolean lattice if and only if there exists a finite distributive lattice L' such that $\mathrm{Con}(L)\cong \mathrm{Con}(L')$.

Corollary (see [1, Theorem 6]). *Let L be a finite modular p-algebra. Then there exists a finite distributive p-algebra L' such that* $\mathrm{Con}(L)\cong \mathrm{Con}(L')$.

Proof. $D(L)$ is a finite modular lattice. It is well known that the congruence lattice of a finite modular lattice is Boolean. Hence $\mathrm{Con}(D(L))$ is a Boolean lattice. The rest follows from Theorem 10.

6. Relative Stone congruence lattices. We start with general results.

Lemma 8. *Let L be a distributive lattice with 1. The following statements are equivalent:*

(i) *L is relative Stone;*

(ii) *for every $a\in L$, $[a, 1]$ is a Stone lattice;*

(iii) *for every $a\leqq b$ in L, $[a, b]$ is a relative Stone lattice;*

(iv) *L is a Brouwerian lattice (i.e. relatively pseudocomplemented) satisfying the identity $x*y\vee y*x=1$.*

Proof. The equivalences between (i), (ii) and (iii) follow from Lemma 5. The equivalence between (i) and (iv) can be found in [8, 2.10].

Lemma 9. *Let L be a Heyting algebra. Then L is a relative Stone lattice if and only if*

(i) *L is a Stone lattice*

and

(ii) $D(L)$ *is relative Stone.*

For the proof see [8, 2.13].

Lemma 10. *Let L be a quasi-modular p-algebra and let $B(L)$ be finite. Then* $(\alpha_1, \alpha_2) \in D(\mathrm{Con}\,(L))$ *if and only if*

(i) $\alpha_2^* = \Delta$, *i.e.* $\alpha_2 \in D(\mathrm{Con}\,(D(L)))$

and

(ii) $\alpha_1 \geqq \tau(\Delta)$.

Proof. Assume $(\alpha_1, \alpha_2) \in D(\mathrm{Con}\,(L))$. Then, by Theorem 1, $\Delta = (\alpha_1, \alpha_2)^* = (\alpha_1^* \wedge \tau(\Delta), \Delta)$. So, $\alpha_2^* = \Delta$. Moreover, $\Delta = \alpha_1^* \wedge \tau(\Delta)$ in $\mathrm{Con}\,(B(L))$. Since $B(L)$ is finite, we have $B(L) \cong \mathrm{Con}\,(B(L))$. Hence $\alpha = \alpha^{**}$ for every $\alpha \in \mathrm{Con}\,(B(L))$. Now, $\Delta = \alpha_1^* \wedge \tau(\Delta)$ implies $\alpha_1^{**} = \alpha_1 \geqq \tau(\Delta)$ proving (ii). Conversely, (i) and (ii) imply $(\alpha_1, \alpha_2)^* = (\alpha_1^* \wedge \tau(\Delta), \Delta) = \Delta$, as $\alpha_1^* \leqq \tau(\Delta)^*$.

Lemma 11. *Let B be a Boolean algebra. Then* $\mathrm{Con}\,(B)$ *is a relative Stone lattice if and only if B is finite.*

Proof. Let $\mathrm{Con}\,(B)$ be a relative Stone lattice. This is equivalent to the fact that $I(B/J)$ is a Stone lattice for every $J \in I(B)$ (Lemma 8). But $I(B)$ is a Stone lattice if and only if B is complete (see [5] or [7, Satz 9]). By [3, Theorem 4.3] every infinite complete Boolean algebra contains an ideal J such that B/J is not complete. That means $I(B/J)$ is not a Stone lattice. Hence B is finite. The converse is trivially true.

Theorem 11. *Let L be a quasi-modular p-algebra. Then* $\mathrm{Con}\,(L)$ *is a relative Stone lattice if and only if*

(i) $\mathrm{Con}\,(L)$ *is a Stone lattice,*

(ii) $B(L)$ *is finite,*

(iii) $\mathrm{Con}\,(D(L))$ *is a relative Stone lattice*

and

(iv) *for any* $\alpha, \beta \in \mathrm{Con}\,(D(L))$ *with* $\alpha \geqq \beta$, $\beta \in D(\mathrm{Con}\,(D(L)))$ *and* $\tau(\beta) \geqq \tau(\Delta)$ *it is true that* $\tau(\alpha * \beta)^* \leqq \tau((\alpha * \beta) * \beta)$.

Proof. Let $\mathrm{Con}\,(L)$ be relative Stone. (i) follows from Lemma 9. Corollary 2 to Theorem 1 says that $\mathrm{Con}\,(B(L)) \cong [\gamma, \nabla]$. Using Lemma 8 we see that $[\gamma, \nabla]$ is also relative Stone. Hence $\mathrm{Con}\,(B(L))$ is relative Stone. By Lemma 11, $B(L)$ is finite and (ii) is established. The condition (iii) follows from the hypothesis and Corollary 1 to Theorem 1. Eventually we shall prove (iv). Lemma 9 and the hypothesis imply that $D(\mathrm{Con}\,(L))$ is relative Stone. Take $\alpha_2 = \alpha$ and $\beta_2 = \beta$ from $\mathrm{Con}\,(D(L))$ with $\alpha \geqq \beta$, $\beta \in D(\mathrm{Con}\,(D(L)))$ and $\tau(\beta) \geqq \tau(\Delta)$. Since $(\tau(\Delta), \alpha_2), (\tau(\Delta), \beta_2) \in D(\mathrm{Con}\,(L))$ (see Lemma 10), there exist $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in D(\mathrm{Con}\,(L))$ with $(\alpha_1, \alpha_2) \geqq (\beta_1, \beta_2)$.

By the hypothesis, $[(\beta_1, \beta_2), \nabla]$ is a Stone lattice (Lemma 8). The pseudocomplements of elements in this interval can be calculated using Theorem 1. Therefore,

$$(\alpha_1, \alpha_2) * (\beta_1, \beta_2) = (\alpha_1 * \beta_1 \wedge \tau(\alpha_2 * \beta_2), \alpha_2 * \beta_2)$$

and

$$((\alpha_1, \alpha_2) * (\beta_1, \beta_2)) * (\beta_1, \beta_2) = (\alpha_1 * \beta_1 \wedge \tau(\alpha_2 * \beta_2), \alpha_2 * \beta_2) * (\beta_1, \beta_2) =$$
$$= (((\alpha_1 * \beta_1) \wedge \tau(\alpha_2 * \beta_2)) * \beta_1 \wedge \tau((\alpha_2 * \beta_2) * \beta_2), (\beta_1 * \beta_2) * \beta_2).$$

By the hypothesis

$$(\alpha_1, \alpha_2) * (\beta_1, \beta_2) \vee ((\alpha_1, \alpha_2) * (\beta_1, \beta_2)) * (\beta_1, \beta_2) = \nabla.$$

Since $B(L)$ is finite, we have $\mathrm{Con}(B(L)) \cong B(L)$. This implies that in $\mathrm{Con}(B(L))$ pseudocomplements are complements (i.e. $\alpha^* = \alpha'$) and $\alpha * \beta = \alpha' \vee \beta$. Bearing this in mind we see that $((\alpha_1 * \beta_1) \wedge \tau(\alpha_2 * \beta_2)) * \beta_1$ is the complement of $\alpha_1 * \beta_1 \wedge \tau(\alpha_2 * \beta_2)$ in $[\beta_1, \nabla]$. Therefore,

$$((\alpha_1 * \beta_1) \wedge \tau(\alpha_2 * \beta_2)) * \beta_1 = (\alpha_1 * \beta_1)^* \vee \tau(\alpha_2 * \beta_2)^* \vee \beta_1 \leqq \tau((\alpha_2 * \beta_2) * \beta_2)$$

and consequently, $\tau(\alpha_2 * \beta_2)^* \leqq \tau((\alpha_2 * \beta_2) * \beta_2)$.

Conversely, suppose that L satisfies (i)—(iv). With regard to (i) and Lemma 9 it suffices to show that $D(\mathrm{Con}(L))$ is a relative Stone lattice. Take $(\beta_1, \beta_2) \in D(\mathrm{Con}(L))$. By Lemma 10, $\beta_2^* = \Delta$ and $\tau(\beta_2) \geqq \tau(\Delta)$. We want to show that $[(\beta_1, \beta_2), \nabla]$ is a Stone lattice (Lemma 8). Take $(\alpha_1, \alpha_2) \geqq (\beta_1, \beta_2)$ in $\mathrm{Con}(L)$. Evidently, $(\alpha_1, \alpha_2) * (\beta_1, \beta_2)$ and $((\alpha_1, \alpha_2) * (\beta_1, \beta_2)) * (\beta_1, \beta_2)$ is a pseudocomplement of (α_1, α_2) and $(\alpha_1, \alpha_2) * (\beta_1, \beta_2)$, respectively, in $[(\beta_1, \beta_2), \nabla]$. By Theorem 1,

$$(\delta_1, \delta_2) = (\alpha_1, \alpha_2) * (\beta_1, \beta_2) \vee ((\alpha_1, \alpha_2) * (\beta_1, \beta_2)) * (\beta_1, \beta_2) =$$
$$= ((\alpha_1 * \beta_1 \wedge \tau(\alpha_2 * \beta_2))) \vee (((\alpha_1 * \beta_1 \wedge \tau(\alpha_2 * \beta_2)) * \beta_1 \wedge \tau((\alpha_2 * \beta_2) * \beta_2)), \alpha_2 * \beta_2 \vee$$
$$\vee (\alpha_2 * \beta_2) * \beta_2).$$

Condition (iii) implies $\alpha_2 * \beta_2 \vee (\alpha_2 * \beta_2) * \beta_2 = \nabla$. Clearly, $\alpha_2 \leqq (\alpha_2 * \beta_2) * \beta_2$ yields $\beta_1 \leqq \alpha_1 \leqq \tau((\alpha_2 * \beta_2) * \beta_2)$. The last condition, (ii) and (iv) imply

$$(\alpha_1 * \beta_1 \wedge \tau(\alpha_2 * \beta_2)) * \beta_1 = (\alpha_1 \wedge \beta_1^*) \vee \tau(\alpha_2 * \beta_2)^* \vee \beta_1 \leqq \tau((\alpha_2 * \beta_2) * \beta_2).$$

Now, it is easy to see that $(\delta_1, \delta_2) = \nabla$. Thus $\mathrm{Con}(L)$ is relative Stone and the proof is complete.

Before establishing the last theorem we need a concept. A lattice L is said to be *locally finite* if all intervals in L are finite.

Lemma 12. *Let L be a Stone lattice. Assume that $B(L)$ is finite. Let $J \in I(L)$. Then $J \in D(I(L))$, i.e. $J^* = (0]$, if and only if $J \cap D(L) \neq \emptyset$.*

Proof. Assume that $J \in D(I(L))$. Assume to the contrary that $J \cap D(L) = \emptyset$. It is well known that there exists a prime ideal $P \in I(L)$ such that $J \subseteq P$ and $P \cap D(L) = \emptyset$. Note that $a \in P$ implies $a^{**} \in P$, as $a^* \wedge a^{**} = 0$. Let $a \in B(L)$ be the join of all elements from $P \cap B(L)$. Since $B(L)$ is finite and L is a Stone lattice, we have $(a] = P$. Evidently $a \neq 1$. Hence $(a^*] \leqq J^*$, a contradiction. Thus $J \cap D(L) \neq 0$. The converse statement is trivially true.

Theorem 12. *Let L be a distributive p-algebra. Then* $\mathrm{Con}(L)$ *is a relative Stone lattice if and only if*

(i) *$B(L)$ is finite,*

(ii) *$D(L)$ is locally finite and relatively complemented,*

(iii) *the dual lattice $\breve{L}$ is a Stone lattice*

and

(iv) *the ideal of dual dense elements $\bar{D}(L)$ (i.e. $\bar{D}(L) = D(\breve{L})$) is locally finite and relatively complemented.*

Proof. Suppose that $\mathrm{Con}(L)$ is a relative Stone lattice. Combining Lemma 9, Theorem 7 and Theorem 11 we get (i), (iii) and that $D(L)$ is relatively complemented. In other words, L is a regular double p-algebra (see [9, Theorem 2]). Again by this theorem we get that $\bar{D}(L)$ is also relatively complemented. By Theorem 11 $\mathrm{Con}(D(L))$ is a relative Stone lattice. But $\mathrm{Con}(D(L)) \cong F(D(L))$. Take $a \in D(L)$. Then $[[1), [a)]$ is an interval in the lattice of all filters $F(D(L))$. Since $[a)$ is a Boolean lattice and $[[1), [a)] = F([a))$, we see that $[a)$ is finite, as $[[1), [a)]$ is a relative Stone lattice (see Lemma 11). Thus $D(L)$ is locally finite and (ii) is completely established. It remains to prove the locally finiteness of $\bar{D}(L)$. Since every congruence relation $\Theta \in \mathrm{Con}(L)$ is also a Heyting algebra congruence relation of L ([10, Lemma 1]), we see that $\mathrm{Con}(L) \cong F(L)$. Take $b \in \bar{D}(L)$, i.e. $b^+ = 1$. Evidently, $(b]$ is a Boolean lattice and $F((b]) \cong [[b), [0)]$. By assumption $[[b), [0)]$ is a relative Stone lattice. Therefore, by Lemma 11, $(b]$ is a finite Boolean lattice. Thus $\bar{D}(L)$ is locally finite, and of course, relatively complemented.

Conversely, suppose that L satisfies (i)—(iv). Theorem 7 says that $\mathrm{Con}(L)$ is a Stone lattice. According to Lemma 10 it suffices to prove that $D(\mathrm{Con}(L))$ is a relative Stone lattice. This follows from the fact (Lemma 9) that for every $\alpha \in D(\mathrm{Con}(L))$, $[\alpha, \nabla]$ is a Stone lattice. Again [9, Theorem 2] and [10, Lemma 1] imply that $\mathrm{Con}(L) \cong F(L)$. Let $\mathrm{Ker}\,\alpha = K \in F(L)$. With regard to Lemma 12, $K \cap \bar{D}(L) \neq \emptyset$. Take $b \in K \cap \bar{D}(L)$. By (iv), $(b]$ is a finite Boolean lattice. Thus $\mathrm{Con}(L)$ is a relative Stone lattice and the proof is complete.

References

[1] R. BEAZER, On congruence lattices of some p-algebras and double p-algebras, *Algebra Universalis*, **13** (1981), 379—388.

[2] J. BERMAN, Congruence relations of pseudocomplemented distributive lattices, *Algebra Universalis*, 3 (1973), 288—293.

[3] PH. DWINGER, On the completeness of the quotient algebras of a complete Boolean algebra. I, *Indag. Math.*, **20** (1958), 448—456.

[4] S. EL-ASSAR, Two notes on the congruence lattice of the p-algebras, *Acta Math. Univ. Comenian*, XLVI—XLVII (1985), 13—20.

[5] O. FRINK, Pseudo-complements in semi-lattices, *Duke Math. J.*, **29** (1962), 505—514.

[6] G. GRÄTZER, *General Lattice Theory*, Birkhäuser Verlag (Basel, 1978).

[7] T. KATRIŇÁK, Pseudokomplementäre Halbverbände, *Mat. Časop.*, **18** (1968), 121—143.

[8] T. KATRIŇÁK, Die Kennzeichnung der distributiven pseudokomplementären Halbverbände, *J. Reine Angew. Math.*, **241** (1970), 160—179.

[9] T. KATRIŇÁK, The structure of distributive double p-algebras. Regularity and congruences, *Algebra Universalis*, 3 (1973), 238—246.

[10] T. KATRIŇÁK, The congruence lattice of distributive p-algebras, *Algebra Universalis*, 7 (1977), 265—271.

[11] T. KATRIŇÁK, Essential and strong extensions of p-algebras, *Bull. Soc. Roy. Sci. Liège*, **49** (1980), 119—124.

[12] T. KATRIŇÁK and S. EL-ASSAR, Algebras with Boolean and Stonean congruence lattices, *Acta Math. Hung.*, **48** (1986), 301—316.

[13] T. KATRIŇÁK and P. MEDERLY, Constructions of p-algebras, *Algebra Universalis*, **17** (1983), 288—316.

[14] W. C. NEMITZ, Implicative semi-lattices, *Trans. Amer. Math. Soc.*, **117** (1965), 128—142.

[15] T. TANAKA, Canonical subdirect factorization of lattices, *J. Sci. Hiroshima Univ., Ser. A*, **16** (1952), 239—246.

[16] J. VARLET, A regular variety of type $\langle 2, 2, 1, 1, 0, 0\rangle$, *Algebra Universalis*, 2 (1972), 218—223.

(T. K.)
KATEDRA ALGEBRY A TEÓRIE ČÍSEL
MATEMATICKO-FYZIKÁLNEJ FAKULTY
UNIVERZITY KOMENSKÉHO
84215 BRATISLAVA, MLYNSKÁ DOLINA, CZECHOSLOVAKIA

Acta Sci. Math., 51 (1987), 387—402

Structure-filters in equality-free model theory

P. ECSEDI-TÓTH

Using a natural definition of (finite) meets of structures and that of the lattice ordering induced by the meet, we introduce the concept of filters on the similarity class of structures. Our main problem here is to answer the question whether the theories of such filters are characterizable by purely syntactical means. Restricting our considerations mostly to equality-free first order languages, we provide an affirmative solution to this problem.

1. Finite meet of structures has been introduced as a simple set theoretic construction in [3], where we have proved the following

Theorem 1.1 ([3], Theorem 2.14). *Let T be an equality-free first order theory. Then the two assertions below are equivalent:*

(i) *T has a set of universal equality-free Horn axioms;*

(ii) *T is preserved under finite meets (cf. Definition 3.6, below).*

It was shown, too, that this theorem fails for theories containing equality; more precisely, (ii) does not entail (i) if the equality is present, while the converse implication (i)$\Rightarrow$(ii) holds in general.

Our starting point in the present work is that, disregarding some set theoretic difficulties, the class of all similar structures forms a weak partial meet-semilattice. It is well-known, that the lattice ordering is uniquely determined in weak partial meet-semilattices. By means of the lattice ordering, filters are definable in the traditional way, and so the following natural questions arise:

(1) Which sentences (theories) are preserved under the lattice ordering induced by the meet?

(2) Which sentences (theories) have a class of models that forms a filter in the weak partial meet-semilattice of structures?

Received April 19, 1984.

We shall give here a complete answer to question (1) (cf. Corollary 4.4, Theorems 4.5, 4.6, 4.7), and a partial one to question (2) (cf. Theorems 5.3 and 5.5), in the sense, that we restrict our attention to equality-free languages, only.

It would be natural, too, to introduce and investigate the duals of these concepts; i.e. the join of structures and ideals of structures. These notions, however cannot be treated analogously to the meets and filters. For example, the meet of structures can be defined without any restrictions on the universes of structures (cf. Definition 3.1, below), nevertheless, a similar definition of the join would involve either the assumption that the universes of all structures in question are the same, or the permission for partial structures (in which functions may be partial). Beyond this definitional difficulty, some of our results do not have analogous dual forms. Thus, it seemed better to deal with these dual question in a separate paper.

2. Some of our assertions refer explicitly to proper classes, and so, in order to avoid set theoretic difficulties, the choice of the underlying set theory is important; in fact, our considerations could be carried out e.g. in the Bernays—Gödel set theory. We shall, however, present the material informally; the formal set theoretic development would be rather tedious.

By a similarity type t we shall mean an ordered quintuple $t=\langle\mathscr{R}, \mathscr{F}, \mathscr{C}, t_{\mathscr{R}}, t_{\mathscr{F}}\rangle$, where $\mathscr{R}$, $\mathscr{F}$, $\mathscr{C}$ are pairwise disjoint sets, $\mathscr{C}\neq\emptyset$, $t_{\mathscr{R}}\colon \mathscr{R}\to\omega-\{0\}$, $t_{\mathscr{F}}\colon \mathscr{F}\to\omega-\{0\}$.

By a structure of type t, we mean an ordered quadruplet

$$\mathfrak{A}=\langle|\mathfrak{A}|, \langle R_r^{(\mathfrak{A})}\rangle_{r\in\mathscr{R}}, \langle F_f^{(\mathfrak{A})}\rangle_{f\in\mathscr{F}}, \langle C_c^{(\mathfrak{A})}\rangle_{c\in\mathscr{C}}\rangle,$$

where $|\mathfrak{A}|$ is a nonvoid set, the universe of $\mathfrak{A}$; for each $r\in\mathscr{R}, f\in\mathscr{F}$ and $c\in\mathscr{C}$, $R_r^{(\mathfrak{A})}$ is a $t_{\mathscr{R}}(r)$-ary relation, $F_f^{(\mathfrak{A})}$ is a $t_{\mathscr{F}}(f)$-ary function and $C_c^{(\mathfrak{A})}$ is a constant on the set $|\mathfrak{A}|$, respectively.

From now on, we shall keep an arbitrary similarity type t be fixed. The class of all structures of type t will be denoted by $\mathfrak{M}^t$; we shall denote the elements of $\mathfrak{M}^t$ by German capitals, $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$, $\mathfrak{D}$, maybe with indices.

We shall use the standard notions and notations of [2]. Additionally, we need some supplementary facts, collected together in the rest of this section.

First, we mention the equality-free version of the well-known Łoś—Tarski preservation theorem (cf. [2], Theorem 3.2.2, p. 124).

Theorem 2.1 ([3], Lemma 2.10). *Let T be an equality-free first order theory. Then, the following two assertions are equivalent:*

(i) *T is preserved under substructures;*

(ii) *T has a set of universal equality-free axioms.*

Analogously, one can prove without major difficulty the dual form of this theorem.

Theorem 2.2. *Let T be an arbitrary equality-free first order theory. Then, the two assertions below are equivalent:*

(i) *T is preserved under extensions;*

(ii) *T has a set of existential equality-free axioms.*

The following concept has been introduced also in [3].

Definition 2.3 ([3], 2.4). Let X be an arbitrary set and consider the absolutely free algebra $\mathfrak{Fr}(X\cup\mathscr{C})$ of type t generated by the set $X\cup\mathscr{C}$ (cf. [5], Definition 0.4.19(i), Remarks 0.4.20, pp. 130—131). Let $\mathfrak{A}\in\mathfrak{M}^t$. It is well-known, that for arbitrary $h\colon X\cup\mathscr{C}\to|\mathfrak{A}|$, such that for all $c\in\mathscr{C}$, $h(c)=C_c^{(\mathfrak{A})}$ holds, there exists a unique homomorphism $\bar{h}$ from $\mathfrak{Fr}(X\cup\mathscr{C})$ into $\mathfrak{A}$ for which $h\subseteq\bar{h}$ (cf. [5], Definition 0.4.23, Theorem 0.4.24, Theorem 0.4.27(i), pp. 131—132). We define the *free structure* $\mathfrak{Fr}_h\mathfrak{A}$ induced by h over $\mathfrak{A}$ as follows:

(i) let $|\mathfrak{Fr}_h\mathfrak{A}|=|\mathfrak{Fr}(X\cup\mathscr{C})|$;

(ii) for every $r\in\mathscr{R}$, $t_{\mathscr{R}}(r)=n+1$ and for arbitrary elements $a_0,\dots,a_n\in|\mathfrak{Fr}_h\mathfrak{A}|$, let

$$\langle a_0,\dots,a_n\rangle\in R_r^{(\mathfrak{Fr}_h\mathfrak{A})}\Leftrightarrow\langle\bar{h}(a_0),\dots,\bar{h}(a_n)\rangle\in R_r^{(\mathfrak{A})},$$

where $\bar{h}$ is the unique extension of h to a homomorphism from $\mathfrak{Fr}(X\cup\mathscr{C})$ into $\mathfrak{A}$;

(iii) for every $f\in\mathscr{F}$, such that $t_{\mathscr{F}}(f)=n+1$ and for arbitrary $a_0,\dots,a_n\in|\mathfrak{Fr}_h\mathfrak{A}|$, let

$$F_f^{(\mathfrak{Fr}_h\mathfrak{A})}(a_0,\dots,a_n)=F_f^{(\mathfrak{Fr}(X\cup\mathscr{C}))}(a_0,\dots,a_n);$$

(iv) finally, for all $c\in\mathscr{C}$, let

$$C_c^{(\mathfrak{Fr}_h\mathfrak{A})}=C_c^{(\mathfrak{Fr}(X\cup\mathscr{C}))}.$$

It was shown in [3], that $\mathfrak{Fr}_h\mathfrak{A}$ is correctly defined and is of type t, provided $\mathfrak{A}\in\mathfrak{M}^t$. We shall need the following

Theorem 2.4 ([3], Lemma 2.5). *Let $\mathfrak{A}\in\mathfrak{M}^t$ and X be a set, $h\colon X\cup\mathscr{C}\to|\mathfrak{A}|$ such that $h(c)=C_c^{(\mathfrak{A})}$ for all $c\in\mathscr{C}$. If h is onto, then $\mathfrak{A}$ and $\mathfrak{Fr}_h\mathfrak{A}$ are elementarily equivalent for equality-free sentences.*

The next assertion is, on the one hand, a particular case of a well-known result of Shoenfield (cf. [2], Theorem 3.1.16, p. 118) in two respects: firstly, it concerns equality-free languages only, and secondly, it is restricted to the lowest levels of the quantifier hierarchy. On the other hand, however, it is a generalization of the mentioned result, since it is about theories instead of single sentences. Our proof, presented here, is purely model theoretic in character and differs from the one given in [2], p. 118.

Theorem 2.5. *Let T be an equality-free first order theory. Then, the following assertions are equivalent:*

(i) *T has both a set of universal equality-free and a set of existential equality-free axioms;*

(ii) *T is preserved under both substructures and extensions;*

(iii) *T has a set of quantifier-free (i.e. $\Pi_0 = \Sigma_0 = \Delta_0$) equality-free axioms.*

Proof. (i) and (ii) are equivalent by Theorem 2.1, and Theorem 2.2. Also, (iii) implies (i) trivially, since every quantifier-free (and equality-free) sentence can be considered as a universal, as well as an existential (equality-free) sentence. To complete the proof, we show that (ii) entails (iii).

First, we prove the following fact.

(3) Let $\mathfrak{A}, \mathfrak{B} \in \mathfrak{M}^t$, and assume that for any quantifier-free and equality-free sentence ψ, $\mathfrak{A} \models \psi \Leftrightarrow \mathfrak{B} \models \psi$. Then $\mathfrak{A} \models T \Leftrightarrow \mathfrak{B} \models T$.

Let X be an arbitrary set with cardinality large enough such that the onto mappings $h\colon X \cup \mathscr{C} \to |\mathfrak{A}|$ and $g\colon X \cup \mathscr{C} \to |\mathfrak{B}|$ exist. Consider the free structures $\mathfrak{Fr}_h \mathfrak{A}$ and $\mathfrak{Fr}_g \mathfrak{B}$ and let us denote by $\mathfrak{A}'$ and $\mathfrak{B}'$ those substructures of $\mathfrak{Fr}_h \mathfrak{A}$ and $\mathfrak{Fr}_g \mathfrak{B}$ which are generated by the set of constants, respectively. (By assumption, there exist constants in $\mathfrak{Fr}_h \mathfrak{A}$ and $\mathfrak{Fr}_g \mathfrak{B}$, so $\mathfrak{A}'$ and $\mathfrak{B}'$ exist.)

We claim that $\mathfrak{A}' = \mathfrak{B}'$.

Indeed, by Definition 2.3, we see that

$$|\mathfrak{Fr}_h \mathfrak{A}| = |\mathfrak{Fr}_g \mathfrak{B}|, \tag{4}$$

and for every $c \in \mathscr{C}$ and $f \in \mathscr{F}$,

$$C_c^{(\mathfrak{Fr}_h \mathfrak{A})} = C_c^{(\mathfrak{Fr}_g \mathfrak{B})}, \tag{5}$$

$$F_f^{(\mathfrak{Fr}_h \mathfrak{A})} = F_f^{(\mathfrak{Fr}_g \mathfrak{B})}. \tag{6}$$

From (4), (5) and (6), it follows that $|\mathfrak{A}'| = |\mathfrak{B}'|$ and $C_c^{(\mathfrak{A}')} = C_c^{(\mathfrak{B}')}$, $F_f^{(\mathfrak{A}')} = F_f^{(\mathfrak{B}')}$, for every $c \in \mathscr{C}$, $f \in \mathscr{F}$.

Finally, let $a_0, \ldots, a_n \in |\mathfrak{A}'|$ and $r \in \mathscr{R}$, such that $t_{\mathfrak{a}}(r) = n+1$. By the definition of $\mathfrak{A}'$, there are closed terms $\tau_0, \ldots, \tau_n$ such that

$$\tau_0^{(\mathfrak{A}')} = a_0, \ldots, \tau_n^{(\mathfrak{A}')} = a_n,$$

(where $\tau_i^{(\mathfrak{A}')}$ denotes the "value of τ_i in $\mathfrak{A}'$", cf. [2], 1.3.13, p. 27). Hence

$$\langle a_0, \ldots, a_n \rangle \in R_r^{(\mathfrak{A}')} \Leftrightarrow \langle \tau_0^{(\mathfrak{A}')}, \ldots, \tau_n^{(\mathfrak{A}')} \rangle \in R_r^{(\mathfrak{A}')} \Leftrightarrow \mathfrak{A}' \models r(\tau_0, \ldots, \tau_n).$$

Since $\tau_0, \ldots, \tau_n$ are closed, $\mathfrak{A}' \models r(\tau_0, \ldots, \tau_n)$ implies that $\mathfrak{Fr}_h \mathfrak{A} \models r(\tau_0, \ldots, \tau_n)$. By Theorem 2.4, $\mathfrak{A} \models r(\tau_0, \ldots, \tau_n)$. According to the assumption of (3), $\mathfrak{B} \models r(\tau_0, \ldots, \tau_n)$, from which $\mathfrak{Fr}_g \mathfrak{B} \models r(\tau_0, \ldots, \tau_n)$ and $\mathfrak{B}' \models r(\tau_0, \ldots, \tau_n)$ follow, again by Theorem 2.4 and by the closedness of $\tau_0, \ldots, \tau_n$.

This, however, means that $\langle \tau_0^{(\mathfrak{B}')}, \ldots, \tau_n^{(\mathfrak{B}')}\rangle \in R_r^{(\mathfrak{B}')}$ and so, using the fact that for all i $(0 \leqq i \leqq n)$, $\tau_i^{(\mathfrak{B}')} = a_i$, which follows from (5) and (6), we obtain: $\langle a_0, \ldots, a_n\rangle \in R_r^{(\mathfrak{B}')}$. Hence, $R_r^{(\mathfrak{A}')} \subset R_r^{(\mathfrak{B}')}$. The converse implication $R_r^{(\mathfrak{B}')} \subset R_r^{(\mathfrak{A}')}$ can be established similarly. Thus $R_r^{(\mathfrak{A}')} = R_r^{(\mathfrak{B}')}$ and, r being chosen arbitrarily, we have $\mathfrak{A}' = \mathfrak{B}'$.

If $\mathfrak{A} \models T$, then by Theorem 2.4, $\mathfrak{Fr}_h \mathfrak{A} \models T$, and since T is preserved under substructures, $\mathfrak{A}' \models T$. So, $\mathfrak{B}' \models T$. But T is preserved under extensions, too, hence $\mathfrak{Fr}_g \mathfrak{B} \models T$, whence we obtain $\mathfrak{B} \models T$; by Theorem 2.4. The converse implication $\mathfrak{B} \models T \Rightarrow \mathfrak{A} \models T$ can be seen in an analogous way.

Thus (3) is proved.

If T is inconsistent, then the set $\{r(c, c, \ldots, c), \neg r(c, c, \ldots, c)\}$, where $r \in \mathscr{R}$, $t_{\mathscr{R}}(r) = n+1$ and $c \in \mathscr{C}$ are arbitrary, is an axiom system for T in the required form. (In fact, speaking on equality-free languages, we may assume that $\mathscr{R} \neq \emptyset$, for otherwise no formula exists; on the other hand, $\mathscr{C} \neq \emptyset$ by assumption.)

Let us suppose that T is consistent and set $T_0 = \{\varphi \mid \varphi$ is a quantifier-free equality-free sentence and $T \models \varphi\}$. Then, $T \models T_0$ and so T_0 is consistent.

Let $\mathfrak{C} \models T_0$ be arbitrary. We claim that there is a structure $\mathfrak{D}$, such that $\mathfrak{D} \models T$, and for every quantifier-free equality-free sentence ψ, $\mathfrak{C} \models \psi \Leftrightarrow \mathfrak{D} \models \psi$.

Indeed, let $\Sigma = \{\varphi \mid \varphi$ is a quantifier-free equality-free sentence and $\mathfrak{C} \models \varphi\}$. Then $\Sigma \cup T$ is consistent. For if $\Sigma \cup T$ were inconsistent, then we could find a finite subset $\{\sigma_0, \ldots, \sigma_m\} \subset \Sigma$ such that $T \models \neg(\sigma_0 \wedge \ldots \wedge \sigma_m)$. But the sentence $\neg(\sigma_0 \wedge \ldots \wedge \sigma_m)$ is itself a quantifier-free equality-free sentence and so it is in T_0, hence $\mathfrak{C} \models \neg(\sigma_0 \wedge \ldots \wedge \sigma_m)$. Nevertheless, $\mathfrak{C} \models \sigma_0 \wedge \ldots \wedge \sigma_m$, by the definition of Σ. This contradiction indicates that $\Sigma \cup T$ is consistent.

Let $\mathfrak{D}$ be a model of $\Sigma \cup T$ and let ψ be an arbitrary equality-free quantifier-free sentence. If $\mathfrak{C} \models \psi$, then $\psi \in \Sigma$ and so $\mathfrak{D} \models \psi$. If $\mathfrak{C} \not\models \psi$, then $\mathfrak{C} \models \neg\psi$ and so $\neg\psi \in \Sigma$, hence $\mathfrak{D} \models \neg\psi$, i.e. $\mathfrak{D} \not\models \psi$.

Thus, $\mathfrak{C}$ and $\mathfrak{D}$ satisfy the condition of (3), and $\mathfrak{C} \models T$ follows from $\mathfrak{D} \models T$, by (3). □

3. Definition 3.1 ([3], 1.2). Let $t = \langle \mathscr{R}, \mathscr{F}, \mathscr{C}; t_{\mathscr{R}}, t_{\mathscr{F}}\rangle$ be a similarity type and let

$$\mathfrak{A}_i = \langle |\mathfrak{A}_i|, \langle R_r^{(\mathfrak{A}_i)}\rangle_{r \in \mathscr{R}}, \langle F_f^{(\mathfrak{A}_i)}\rangle_{f \in \mathscr{F}}, \langle C_c^{(\mathfrak{A}_i)}\rangle_{c \in \mathscr{C}}\rangle$$

be structures of type t for $i < n+1$, where $n \in \omega$. We define the set theoretic meet of $\mathfrak{A}_i$, $i < n+1$ as follows:

$$\bigcap_{i<n+1} \mathfrak{A}_i = \langle \bigcap_{i<n+1} |\mathfrak{A}_i|, \langle \bigcap_{i<n+1} R_r^{(\mathfrak{A}_i)}\rangle_{r \in \mathscr{R}}, \langle \bigcap_{i<n+1} F_f^{(\mathfrak{A}_i)}\rangle_{f \in \mathscr{F}}, \bigcap_{i<n+1} \langle C_c^{(\mathfrak{A}_i)}\rangle_{c \in \mathscr{C}}\rangle$$

where the meets on the right hand side of the equation are meant in the sense of set theory (i.e. the meet of functions is taken as the meet of sets of pairs representing those functions; the meet of sequences of constants is defined again as the meet of ordered sets).

If $\bigcap_{i<n+1} \mathfrak{A}_i \in \mathfrak{M}^t$, then it is called the *model theoretic meet* (from now on, simply, the meet) of $\mathfrak{A}_i$, $i<n+1$. We shall use the infix notation $\mathfrak{A}_0 \cap \mathfrak{A}_1 \cap \ldots \cap \mathfrak{A}_n$ for the meet of $\mathfrak{A}_i$, $i<n+1$.

Clearly, $\bigcap_{i<n+1} \mathfrak{A}_i$ always exists as a tuple. The meet of the structures $\mathfrak{A}_i$, $i<n+1$, however, is a partial operation: it may well happen, that the meet of $\mathfrak{A}_i$, $i<n+1$ does not exist even if $\bigcap_{i<n+1} |\mathfrak{A}_i| \neq \emptyset$. We shall use synonymously the following two expressions:

$$\text{“}\bigcap_{i<n+1} \mathfrak{A}_i \in \mathfrak{M}^t\text{”} \quad \text{and} \quad \text{“}\mathfrak{A}_0 \cap \mathfrak{A}_1 \cap \ldots \cap \mathfrak{A}_n \text{ exists”}.$$

The meet, if exists, is very close to the set theoretic meet. In particular, it possesses the following familiar properties.

Lemma 3.2. *Let* $\mathfrak{A}, \mathfrak{B}, \mathfrak{C} \in \mathfrak{M}^t$ *be arbitrary.*

(i) $\mathfrak{A} \cap \mathfrak{A} = \mathfrak{A}$, *hence* $\mathfrak{A} \cap \mathfrak{A} \in \mathfrak{M}^t$.

(ii) *If* $\mathfrak{A} \cap \mathfrak{B} \in \mathfrak{M}^t$, *then* $\mathfrak{B} \cap \mathfrak{A} \in \mathfrak{M}^t$, *and* $\mathfrak{A} \cap \mathfrak{B} = \mathfrak{B} \cap \mathfrak{A}$.

(iii) *If* $\mathfrak{A} \cap \mathfrak{B} \in \mathfrak{M}^t$ *and* $\mathfrak{B} \cap \mathfrak{C} \in \mathfrak{M}^t$, *then* ($\alpha$) *and* ($\beta$) *below are equivalent and any of them implies* $(\mathfrak{A} \cap \mathfrak{B}) \cap \mathfrak{C} = \mathfrak{A} \cap (\mathfrak{B} \cap \mathfrak{C})$:

(α) $(\mathfrak{A} \cap \mathfrak{B}) \cap \mathfrak{L} \in \mathfrak{M}^t$,

(β) $\mathfrak{A} \cap (\mathfrak{B} \cap \mathfrak{L}) \in \mathfrak{M}^t$.

Proof. (i) and (ii) are trivial.

(iii): Assume that $\mathfrak{A} \cap \mathfrak{B} \in \mathfrak{M}^t$, $\mathfrak{B} \cap \mathfrak{C} \in \mathfrak{M}^t$. If ($\alpha$) is true, i.e. $(\mathfrak{A} \cap \mathfrak{B}) \cap \mathfrak{C} \in \mathfrak{M}^t$, then consider $\mathfrak{A} \cap (\mathfrak{B} \cap \mathfrak{C})$. By the associativity of the set theoretic meet, which is immediate by Definition 3.1, we have $(\mathfrak{A} \cap \mathfrak{B}) \cap \mathfrak{C} = \mathfrak{A} \cap (\mathfrak{B} \cap \mathfrak{C})$, hence ($\beta$) is true and $(\mathfrak{A} \cap \mathfrak{B}) \cap \mathfrak{C} = \mathfrak{A} \cap (\mathfrak{B} \cap \mathfrak{C})$ holds. The converse can be established similarly. □

An immediate consequence of this lemma is the following

Theorem 3.3. *Let t be a fixed similarity type. Then, the class of all structures of type t forms a weak partial meet-semilattice.*

Definition 3.4. Let us define the binary relation $\leq$ on $\mathfrak{M}^t$ by the item: for any $\mathfrak{A}, \mathfrak{B} \in \mathfrak{M}^t$, $\mathfrak{A} \leq \mathfrak{B}$ iff $\mathfrak{A} \cap \mathfrak{B}$ exists and $\mathfrak{A} \cap \mathfrak{B} = \mathfrak{A}$.

If $\mathfrak{A} \leq \mathfrak{B}$, then we say that “$\mathfrak{A}$ is a *weak substructure* of $\mathfrak{B}$”, or equivalently, that “$\mathfrak{B}$ is a *weak extension* of $\mathfrak{A}$”.

The next assertion collects some useful facts about the relation $\leq$. The proof is an easy verification or can be readily obtained from the general theory of lattices [4].

Lemma 3.5. *Let* $\mathfrak{A}, \mathfrak{B}, \mathfrak{C} \in \mathfrak{M}^t$.

(i) $\leq$ *is a partial ordering on* $\mathfrak{M}^t$.

(ii) *If* $\mathfrak{A} \cap \mathfrak{B}$ *exists, then*

(α) $\mathfrak{A} \sqcap \mathfrak{B} \leqq \mathfrak{A}$ *and* $\mathfrak{A} \sqcap \mathfrak{B} \leqq \mathfrak{B}$;

(β) $\mathfrak{C} \leqq \mathfrak{A}$ *and* $\mathfrak{C} \leqq \mathfrak{B}$ *imply that* $\mathfrak{C} \leqq \mathfrak{A} \sqcap \mathfrak{B}$.

(iii) *If* $\mathfrak{A} \subset \mathfrak{B}$ *then* $\mathfrak{A} \leqq \mathfrak{B}$ *(where* $\subset$ *stands for the traditional concept of substructures). The converse implication is not true in general.*

(iv) $\mathfrak{A} \leqq \mathfrak{B}$ *iff* $|\mathfrak{A}| \subset |\mathfrak{B}|$ *and the identity mapping* $\imath: |\mathfrak{A}| \to |\mathfrak{B}|$, *defined by* $\imath(a)=a$, *is a homomorphism in the model theoretic sense.*

The clause (iii) of this lemma justifies the adjective "weak" in the naming of weak substructures.

Definition 3.6. Let T be an arbitrary first order theory. We say that

(i) T is *preserved under weak substructures* (resp. *under weak extensions*) iff for all $\mathfrak{A}, \mathfrak{B} \in \mathfrak{M}^t$, if $\mathfrak{A} \models T$ and $\mathfrak{B} \leqq \mathfrak{A}$ (resp. $\mathfrak{A} \leqq \mathfrak{B}$), then $\mathfrak{B} \models T$;

(ii) T is *preserved under finite meets* iff for all $\mathfrak{A}_0, \mathfrak{A}_1, \ldots, \mathfrak{A}_n \in \mathfrak{M}^t$, if $\mathfrak{A}_0 \models T$, $\mathfrak{A}_1 \models T, \ldots, \mathfrak{A}_n \models T$ and $\mathfrak{A}_0 \sqcap \mathfrak{A}_1 \sqcap \ldots \sqcap \mathfrak{A}_n$ exists, then $\mathfrak{A}_0 \sqcap \mathfrak{A}_1 \sqcap \ldots \sqcap \mathfrak{A}_n \models T$.

The next assertion is a slight strengthening of Lemma 3.5(ii), (iii), and is true for arbitrary first order languages.

Theorem 3.7. *Let T be an arbitrary first order theory.*

(i) *If T is preserved under weak substructures, then T is preserved under finite meets.*

(ii) *If T is preserved under finite meets, then T is preserved under traditional substructures.*

(iii) *None of these implications in* (i) *and* (ii) *can be reversed in general.*

Proof. (i): Let us suppose that T is preserved under weak substructures; let $\mathfrak{A}_0, \mathfrak{A}_1, \ldots, \mathfrak{A}_n \in \mathfrak{M}^t$, and assume that for all $i<n+1$, $\mathfrak{A}_i \models T$ and the meet $\mathfrak{A}_0 \sqcap \mathfrak{A}_1 \sqcap \ldots \sqcap \mathfrak{A}_n$ exists. By Lemma 3.5(ii) it is easily seen that $\mathfrak{A}_0 \sqcap \mathfrak{A}_1 \sqcap \ldots \sqcap \mathfrak{A}_n \leqq \mathfrak{A}_0$ and so, $\mathfrak{A}_0 \sqcap \mathfrak{A}_1 \sqcap \ldots \sqcap \mathfrak{A}_n \models T$, because T is preserved under weak substructures.

(ii): Let T be such that T is preserved under finite meets. Let $\mathfrak{A} \models T$, $\mathfrak{B} \subset \mathfrak{A}$. We define the structure $\mathfrak{A}'$ as follows. First set $|\mathfrak{A}'| = |\mathfrak{B}| \cup ((|\mathfrak{A}| - |\mathfrak{B}|) \times \{|\mathfrak{A}|\})$; then define $h\colon |\mathfrak{A}| \to |\mathfrak{A}'|$ by the item

$$h(a) = \begin{cases} a & \text{iff } a \in |\mathfrak{B}| \\ \langle a, |\mathfrak{A}| \rangle & \text{iff } a \in |\mathfrak{A}| - |\mathfrak{B}|. \end{cases}$$

Obviously, h is one-to-one and is onto. For all $c \in \mathscr{C}$, let $C_c^{(\mathfrak{A}')} = h(C_c^{(\mathfrak{A})})$. For every $f \in \mathscr{F}$, $t_{\mathscr{F}}(f) = n+1$ and for arbitrary elements $a'_0, \ldots, a'_n \in |\mathfrak{A}'|$, let

$$F_f^{(\mathfrak{A}')}(a'_0, \ldots, a'_n) = h\big(F_f^{(\mathfrak{A})}(h^{-1}(a'_0), \ldots, h^{-1}(a'_n))\big).$$

Finally, for every $r \in \mathscr{R}$, $t_{\mathscr{R}}(r) = n+1$ and elements $a'_0, \ldots, a'_n \in |\mathfrak{A}'|$; let

$$\langle a'_0, \ldots, a'_n \rangle \in R_r^{(\mathfrak{A}')} \Leftrightarrow \langle h^{-1}(a'_0), \ldots, h^{-1}(a'_n) \rangle \in R_r^{(\mathfrak{A})}.$$

Then $\mathfrak{A}'$ is correctly defined and $\mathfrak{A}'\in\mathfrak{M}^t$, provided $\mathfrak{A}\in\mathfrak{M}^t$. Moreover, $\mathfrak{A}'$ is isomorphic to $\mathfrak{A}$ by h. Thus $\mathfrak{A}'\models T$. By the construction, $\mathfrak{A}\cap\mathfrak{A}'=\mathfrak{B}$ and so, T being preserved under finite meets, $\mathfrak{B}\models T$.

(iii): Let us consider the (equality-free) theories

$$T_1 = \{(\forall x)r(x)\}, \quad T_2 = \{(\forall x)(r(x)\vee q(x))\},$$

where r and q are distinct unary relation symbols of an appropriate particular similarity type t.

By Theorem 2.1, T_2 is preserved under traditional substructures but, according to Theorem 1.1, is not preserved under finite meets.

Similarly, Theorem 1.1 shows, that T_1 is preserved under finite meets. Nevertheless, T_1 is not preserved under weak substructures as the following counterexample indicates. (This follows also from Theorem 4.7, below.)

Obviously, T_1 is consistent; let $\mathfrak{A}$ be a model of T_1. Let us define the structure $\mathfrak{C}$ as follows. First set $|\mathfrak{C}|=|\mathfrak{A}|$. Then, for every $f\in\mathcal{F}$, and $c\in\mathcal{C}$, put $F_f^{(\mathfrak{C})}=F_f^{(\mathfrak{A})}$ and $C_c^{(\mathfrak{C})}=C_c^{(\mathfrak{A})}$. Finally, for every $r\in\mathcal{R}$, let $R_r^{(\mathfrak{C})}=\emptyset$.

Trivially, $\mathfrak{C}\not\models T_1$ and $\mathfrak{C}\cap\mathfrak{A}=\mathfrak{C}$, i.e. $\mathfrak{C}<\mathfrak{A}$. $\square$

The "dual" of this theorem is simply a reformulation of Lemma 3.5(iii) in terms of preservation properties.

Theorem 3.8. *Let T be an arbitrary first order theory. If T is preserved under weak extensions, then T is preserved under (traditional) extensions. The converse fails in general.*

Proof. Trivial by Lemma 3.5(iii). $\square$

Corollary 3.9. *Let T be an arbitrary first order theory.*

(i) *If T is preserved under weak substructures or under finite meets, then T has a set of universal axioms. If, in addition, T is equality-free, then it has a universal axiom system which is equality-free.*

(ii) *If T is preserved under weak extensions, then T has a set of existential axioms, which are equality-free, provided T is such.*

Proof. (i): In contrary to the assertion, let us suppose that T has no universal axioms. Then T is consistent. By the well-known Łoś—Tarski preservation theorem ([2], Theorem 3.2.2, p. 124), we can find a model $\mathfrak{A}$ of T and a substructure $\mathfrak{B}$ of $\mathfrak{A}$, such that $\mathfrak{B}\not\models T$. By Theorem 3.7, T is preserved under neither weak substructures nor finite meets. If T is equality-free, then using Theorem 2.1 in place of the Łoś—Tarski theorem, the same argument applies.

(ii): Similar. $\square$

4. This section is devoted to answering the question (1).

Definition 4.1. Let us suppose that φ is an arbitrary first order formula. By predicate logic, φ is equivalent to a formula ψ of the form

(7)

$$\psi = (Q_1x_1\ldots Q_nx_n)\bigwedge_{i=1}^{m}(\neg\varphi_{i1}\vee\ldots\vee\neg\varphi_{ij_i}\vee\psi_{i1}\vee\ldots\vee\psi_{is_i}\vee\neg\varepsilon_{i1}\vee\ldots\vee\neg\varepsilon_{ik_i}\vee\eta_{i1}\vee\ldots\vee\eta_{il_i}),$$

where $n, m\in\omega$; for all i, $1\leqq i\leqq m$, $j_i, s_i, k_i, l_i\in\omega$, $\varphi_{i1}, \ldots, \varphi_{ij_i}, \psi_{i1}, \ldots, \psi_{is_i}$ are proper atomic formulae of the form $r(\tau_0, \ldots, \tau_v)$ for some $r\in\mathscr{R}$, $t_{\mathscr{R}}(r)=v+1$ and terms $\tau_0, \ldots, \tau_v$; and $\varepsilon_{i1}, \ldots, \varepsilon_{ik_i}, \eta_{i1}, \ldots, \eta_{il_i}$ are equations of the form $\tau_0\equiv\tau_1$, for some terms τ_0, τ_1; and finally, for all z, $1\leqq z\leqq n$, $Q_z\in\{\forall, \exists\}$.

We say that ψ (of the form (7)) is an *equationally-augmented negative* (resp. *positive*) formula, an EAN-formula (resp. EAP-formula), for short, iff for all i, $1\leqq i\leqq m$, $s_i=0$ (resp. $j_i=0$).

Lemma 4.2. *Let T be an arbitrary first order theory. If T has a set of existential* EAP-*axioms, then T is preserved under weak extensions.*

Proof. It will suffice to prove, that every existential EAP-sentence φ is preserved under weak extensions. We shall proceed by induction.

First we observe some trivial facts. Let $\mathfrak{A}, \mathfrak{B}\in\mathfrak{M}^t$, and $\mathfrak{B}\leqq\mathfrak{A}$. We shall denote the set of variables by V.

(8) If $k\colon V\to|\mathfrak{B}|$, then $k\colon V\to|\mathfrak{A}|$; that is, every assignment relative to $\mathfrak{B}$ can as well be regarded as an assignment relative to $\mathfrak{A}$.

(9) For every $r\in\mathscr{R}$, $R_r^{(\mathfrak{B})}\subset R_r^{(\mathfrak{A})}$, by Lemma 3.5(iv).

(10) If τ is a term in the variables $x_1, \ldots, x_n$, then for all $k\colon V\to|\mathfrak{B}|$, $\tau^{(\mathfrak{B})}[k]=$ $=\tau^{(\mathfrak{A})}[k]$, by (8) and by Lemma 3.5(iv). (Here $\tau^{(\mathfrak{B})}[k]$ (resp. $\tau^{(\mathfrak{A})}[k]$) stands for "the value of τ in $\mathfrak{B}$ (resp. in $\mathfrak{A}$) at k"; cf. [2], 1.3.13, p. 27).

Now, let us suppose, that $\varphi=(\exists x_1\ldots\exists x_n)\psi$, where ψ is an atomic formula in the variables $x_1, \ldots, x_n$, and let $\mathfrak{B}\models\varphi$. Then there is an assignment $k\colon V\to|\mathfrak{B}|$, such that

$$\mathfrak{B}\models\psi[k]. \tag{11}$$

Recalling that ψ is in one of the following three forms: $\tau_0\equiv\tau_1$, $\neg(\tau_0\equiv\tau_1)$ and $r(\tau_0, \ldots, \tau_v)$, we see that, in any case, $\mathfrak{A}\models\psi[k]$ is immediate from (11) by (8), (9) and (10).

The induction trivially passes over all the remaining cases, hence the assertion is proved. □

The converse of this lemma holds, too.

Theorem 4.3. *Let T be an arbitrary first order theory. If T is preserved under weak extensions, then T has a set of existential* EAP *axioms.*

Proof. If T is inconsistent, then the set $\{(\exists x)\neg(x\equiv x)\}$ is an axiom system for T in the required form. Hence we may assume, that T is consistent. Let $\Gamma=\{\varphi|\varphi$ is an existential EAP sentence and $T\models\varphi\}$. Then, obviously, $T\models\Gamma$ and Γ is consistent. We shall prove that $\Gamma\models T$.

Let $\mathfrak{A}\models\Gamma$. First we show, that there is a structure $\mathfrak{B}$ such that $\mathfrak{B}\models T$, and every existential EAP sentence holding in $\mathfrak{B}$ holds in $\mathfrak{A}$. To see this, let $\Sigma=\{\neg\varphi|\varphi$ is an existential EAP sentence and $\mathfrak{A}\models\neg\varphi\}$. We claim that $\Sigma\cup T$ is consistent. Indeed, if $\Sigma\cup T$ were inconsistent, then we could find a finite subset $\{\neg\sigma_0, \ldots, \neg\sigma_m\}\subset\Sigma$ such that $T\models\neg(\neg\sigma_0\wedge\ldots\wedge\neg\sigma_m)$. But $\neg(\neg\sigma_0\wedge\ldots\wedge\neg\sigma_m)$ is equivalent to an existential EAP sentence, say σ, and thus $T\models\sigma$ implies that $\sigma\in\Gamma$, hence $\mathfrak{A}\models\sigma$, that is $\mathfrak{A}\models\neg(\neg\sigma_0\wedge\ldots\wedge\neg\sigma_m)$. This, however, contradicts to the assumption that $\mathfrak{A}\models\neg\sigma_0, \ldots, \mathfrak{A}\models\neg\sigma_m$. So $\Sigma\cup T$ is consistent. Let $\mathfrak{B}$ be an arbitrary model of $\Sigma\cup T$ and suppose that χ is an existential EAP sentence which is true in $\mathfrak{B}$. Assume that $\mathfrak{A}\not\models\chi$, i.e. $\mathfrak{A}\models\neg\chi$. Then $\neg\chi\in\Sigma$ which entails that $\mathfrak{B}\models\neg\chi$, a contradiction. Thus $\mathfrak{A}\models\chi$.

Next we show that if $\mathfrak{B}$ is such that $\mathfrak{B}\models T$ and every existential EAP sentence holding in $\mathfrak{B}$ holds in $\mathfrak{A}$, then there are structures $\mathfrak{A}'$, $\mathfrak{B}'$ for which we have $\mathfrak{A}\prec\mathfrak{A}'$, $\mathfrak{B}'\leqq\mathfrak{A}'$ and $\mathfrak{B}'$ is isomorphic to $\mathfrak{B}$. (Here $\prec$ stands for the traditionally defined concept "elementary submodel", cf. [2], p. 107.)

Let c_a and d_b be new constant symbols for every $a\in|\mathfrak{A}|$ and $b\in|\mathfrak{B}|$, respectively, thus forming the diagram languages of $\mathfrak{A}$ and $\mathfrak{B}$ (cf. [2], p. 108). Make sure that $\{c_a|a\in|\mathfrak{A}|\}\cap\{d_b|b\in|\mathfrak{B}|\}=\emptyset$. Let $\Gamma_{\mathfrak{A}}$ be the elementary diagram of $\mathfrak{A}$ (cf. [2], p. 108). Let $\Delta_{\mathfrak{B}}^{+\mathrm{ea}}$ be the set of all positive atomic sentences and all negated equations in the diagram language of $\mathfrak{B}$ which hold in the diagram expansion $(\mathfrak{B}, b)_{b\in|\mathfrak{B}|}$ (cf. [2], p. 108). (That is, $\Delta_{\mathfrak{B}}^{+\mathrm{ea}}$ is a proper subset of the diagram $\Delta_{\mathfrak{B}}$ of $\mathfrak{B}$, cf. [2], p. 68, obtained from $\Delta_{\mathfrak{B}}$ by omitting all elements of the form $\neg r(\tau_0, \ldots, \tau_m)$.)

We claim that $\Gamma_{\mathfrak{A}}\cup\Delta_{\mathfrak{B}}^{+\mathrm{ea}}$ is consistent. Let us suppose the contrary: $\Gamma_{\mathfrak{A}}\cup\Delta_{\mathfrak{B}}^{+\mathrm{ea}}$ is inconsistent. Then we can find a finite subset $\{\delta_0, \ldots, \delta_m\}\subset\Delta_{\mathfrak{B}}^{+\mathrm{ea}}$, such that $\Gamma_{\mathfrak{A}}\models\neg(\delta_0\wedge\ldots\wedge\delta_m)$. Since the elements of $\{d_b|b\in|\mathfrak{B}|\}$ do not appear in $\Gamma_{\mathfrak{A}}$, we can treat them in $\neg(\delta_0\wedge\ldots\wedge\delta_m)$ as free variables. It follows from the universal Closure Theorem of predicate logic, that for an appropriately large $n\in\omega$,

$$\Gamma_{\mathfrak{A}}\models(\forall x_1\ldots\forall x_n)\neg(\delta_0(x_1, \ldots, x_n)\wedge\ldots\wedge\delta_m(x_1, \ldots, x_n)).$$

In particular, $(\mathfrak{A}, a)_{a\in|\mathfrak{A}|}\models(\forall x_1\ldots\forall x_n)\neg(\delta_0(x_1, \ldots, x_n)\wedge\ldots\wedge\delta_m(x_1, \ldots, x_n))$, and so

$$\mathfrak{A}\models(\forall x_1\ldots\forall x_n)\neg(\delta_0(x_1, \ldots, x_n)\wedge\ldots\wedge\delta_m(x_1, \ldots, x_n)) \tag{12}$$

because no elements of $\{c_a|a\in|\mathfrak{A}|\}$ appear in the sentence

$$\chi=(\forall x_1\ldots\forall x_n)\neg(\delta_0(x_1, \ldots, x_n)\wedge\ldots\wedge\delta_m(x_1, \ldots, x_n)).$$

On the other hand, however, $(\mathfrak{B}, b)_{b\in|\mathfrak{B}|}\models\delta_0\wedge\ldots\wedge\delta_m$, and so

$$\mathfrak{B}\models(\exists x_1\ldots\exists x_n)(\delta_0(x_1, \ldots, x_n)\wedge\ldots\wedge\delta_m(x_1, \ldots, x_n)).$$

But the sentence $(\exists x_1\ldots\exists x_n)(\delta_0(x_1, \ldots, x_n)\wedge\ldots\wedge\delta_m(x_1, \ldots, x_n))$ is an existential EAP sentence, hence, by assumption

$$\mathfrak{A}\models(\exists x_1\ldots\exists x_n)(\delta_0(x_1, \ldots, x_n)\wedge\ldots\wedge\delta_m(x_1, \ldots, x_n)),$$

which contradicts to (12). Thus $\Gamma_{\mathfrak{A}}\cup\Delta_{\mathfrak{B}}^{+\mathrm{ea}}$ is consistent, indeed.

Let $(\mathfrak{A}', a', b')_{a\in|\mathfrak{A}|, b\in|\mathfrak{B}|}$ be a model of $\Gamma_{\mathfrak{A}}\cup\Delta_{\mathfrak{B}}^{+\mathrm{ea}}$ (where a' and b' denote the interpretations of the new constant symbols c_a and d_b for every $a\in|\mathfrak{A}|$ and $b\in|\mathfrak{B}|$, respectively). We may assume that for all $a\in|\mathfrak{A}|$, $a'=a$; i.e. $|\mathfrak{A}|\subset|\mathfrak{A}'|$. Then $\mathfrak{A}\prec\mathfrak{A}'$, because $(\mathfrak{A}', a, b')_{a\in|\mathfrak{A}|, b\in|\mathfrak{B}|}\models\Gamma_{\mathfrak{A}}$. Let us define the mapping $g\colon|\mathfrak{B}|\to|\mathfrak{A}'|$ by the equation $g(b)=b'$. Since $(\mathfrak{A}', a, b')_{a\in|\mathfrak{A}|, b\in|\mathfrak{B}|}\models\Delta_{\mathfrak{B}}^{+\mathrm{ea}}$, it is easily seen that g is an isomorphism in the algebraic sense (leaving relations out of consideration) and that g is a model theoretic homomorphism (when relations are considered, too). By Lemma 3.5(iv), there is a weak substructure $\mathfrak{B}'$ of $\mathfrak{A}'$, such that $\mathfrak{B}'$ and $\mathfrak{B}$ are isomorphic by g.

Now, $\mathfrak{B}\models T$ implies $\mathfrak{B}'\models T$. T is preserved under weak extensions, hence $\mathfrak{A}'\models T$. By $\mathfrak{A}\prec\mathfrak{A}'$, we have $\mathfrak{A}\models T$, which was to be proved. □

Corollary 4.4. *Let T be an arbitrary first order theory. Then, the two assertions below are equivalent:*

(i) *T is preserved under weak extensions;*

(ii) *T has a set of existential* EAP *axioms.*

Proof. Immediate by Lemma 4.2 and Theorem 4.3. □

The dual of Corollary 4.4 has a somewhat simpler proof; in fact, we need the compactness property only, and we shall not use elementary submodels.

Theorem 4.5. *Let T be an arbitrary first order theory. Then the two assertions below are equivalent:*

(i) *T is preserved under weak substructures;*

(ii) *T has a set of universal* EAN *axioms.*

Proof. (i)$\Rightarrow$(ii): We may assume that T is consistent for otherwise the set $\{(\forall x)\neg(x\equiv x)\}$ shows that (ii) is true.

Let $\Gamma=\{\varphi|\varphi$ is a universal EAN sentence and $T\models\varphi\}$. Then $T\models\Gamma$ and Γ is consistent.

Let $\mathfrak{A}\models\Gamma$ and consider the set $\Delta_{\mathfrak{A}}^{+\mathrm{ea}}$ (defined in the very same way as $\Delta_{\mathfrak{B}}^{+\mathrm{ea}}$ was defined in the proof of Theorem 4.3, but, of course, $\mathfrak{B}$ replaced everywhere by $\mathfrak{A}$).

We claim that $\Delta_{\mathfrak{A}}^{+ea}\cup T$ is consistent. To see this let $\{\delta_0, ..., \delta_m\}\subset\Delta_{\mathfrak{A}}^{+ea}$. Then, the sentence

$$\chi = (\exists x_1 ... \exists x_n)(\delta_0(x_1, ..., x_n)\wedge ... \wedge\delta_m(x_1, ..., x_n))$$

is true in $\mathfrak{A}$ for an appropriate $n\in\omega$. But χ must hold in some model of T, since otherwise (when χ is false in every model of T), we would have $\neg\chi\in\Gamma$, because $\neg\chi$ is a universal EAN sentence, and so, we would arrive to the contradiction $\mathfrak{A}\models\neg\chi$. Thus, $\{\delta_0, ..., \delta_m\}$ is consistent with T and, by compactness, $\Delta_{\mathfrak{A}}^{+ea}\cup T$ is consistent.

Let $(\mathfrak{B}, a')_{a\in|\mathfrak{A}|}$ be a model of $\Delta_{\mathfrak{A}}^{+ea}\cup T$ (where a' stands for the interpretation of the newly added constant symbol c_a for each $a\in|\mathfrak{A}|$). Let $g: |\mathfrak{A}|\to|\mathfrak{B}|$ be defined by the item $g(a)=a'$. Since $(\mathfrak{B}, a')_{a\in|\mathfrak{A}|}\models\Delta_{\mathfrak{A}}^{+ea}$, it is easy to see that g is an isomorphism in the algebraic sense (relations dropped) and is a homomorphism if we consider relations, too. It follows from Lemma 3.5(iv) that there is a weak substructure $\mathfrak{B}'$ of $\mathfrak{B}$ such that $\mathfrak{A}$ is isomorphic to $\mathfrak{B}'$.

T is preserved under weak substructures, hence $\mathfrak{B}'\models T$ follows from $\mathfrak{B}\models T$ and $\mathfrak{A}'\leqq\mathfrak{B}$. Thus, $\mathfrak{A}\models T$, i.e. Γ is an axiom system for T.

(ii)⇒(i): It suffices to prove that every universal EAN sentence φ is preserved under weak substructures. This can be done by a simple argument; details are omitted. □

The statement of Theorem 4.5 is a slight strengthening of a result due to H. Andréka, I. Németi and I. Sain (cf. [1], § 6. Theorem 1; [6], Theorem 1, Theorem 3). Their proof, however, is purely category theoretic in character and works only if T is assumed to be universal. By Theorem 3.7, the assumption that T is universal, does not mean the loss of generality; nevertheless, this is not clear from the category theoretical framework.

For equality-free languages we prove

Theorem 4.6. *Let T be an equality-free consistent first order theory. Then, the following two assertions are equivalent:*

(i) *T is preserved under weak extensions;*

(ii) *T has a set of existential positive equality-free axioms.*

Proof. (i)⇒(ii): Let $\Gamma=\{\varphi|\varphi$ is an existential positive equality-free sentence and $T\models\varphi\}$. T is assumed to be consistent, hence Γ is consistent, because $T\models\Gamma$. We shall prove that $\Gamma\models T$.

Let $\mathfrak{A}\models\Gamma$. Just as in the proof of Theorem 4.3, we see that there is a structure $\mathfrak{B}$, such that $\mathfrak{B}\models T$, and every existential positive equality-free sentence holding in $\mathfrak{B}$ holds in $\mathfrak{A}$. Let $\mathfrak{A}$ and $\mathfrak{B}$ be fixed in the rest of this proof.

For every $b\in|\mathfrak{B}|$, let d_b be a new constant symbol and form the diagram language of $\mathfrak{B}$ (the language constructed from the non-logical symbols of t and the new set of constant symbols $\{d_b | b \in |\mathfrak{B}|\}$). Let $\Delta_{\mathfrak{B}}^{+\mathrm{ef}}$ be the set of all (positive) atomic sentences of the form $r(\tau_0, \ldots, \tau_n)$, where $r \in \mathscr{R}$, $t_{\mathscr{R}}(r) = n+1$ and $\tau_0, \ldots, \tau_n$ are terms in the diagram language of $\mathfrak{B}$, which are true in $(\mathfrak{B}, b)_{b \in |\mathfrak{B}|}$. Let Σ be the set of all equality-free sentences (of the original language) which hold in $\mathfrak{A}$.

Following closely the way the consistency of $\Gamma_{\mathfrak{A}} \cup \Delta_{\mathfrak{B}}^{+\mathrm{ea}}$ is established in the proof of Theorem 4.3, one proves that $\Sigma \cup \Delta_{\mathfrak{B}}^{+\mathrm{ef}}$ is consistent.

Let $(\mathfrak{C}, b')_{b \in |\mathfrak{B}|}$ be a model of $\Sigma \cup \Delta_{\mathfrak{B}}^{+\mathrm{ef}}$ (where, as usual, b' denotes the interpretation of d_b for each $b \in |\mathfrak{B}|$). First we show the following statement is true:

(13) For every equality-free first order sentence φ,

$$\mathfrak{A} \models \varphi \Leftrightarrow \mathfrak{C} \models \varphi.$$

Indeed, if $\mathfrak{A} \models \varphi$, then $\varphi \in \Sigma$ and thus $(\mathfrak{C}, b')_{b \in |\mathfrak{B}|} \models \varphi$, from which $\mathfrak{C} \models \varphi$ follows, because the elements of the set $\{d_b | b \in |\mathfrak{B}|\}$ cannot appear in φ. On the other hand, if $\mathfrak{A} \not\models \varphi$, i.e. $\mathfrak{A} \models \neg\varphi$, then $\neg\varphi \in \Sigma$, and so $\mathfrak{C} \models \neg\varphi$ is obtained. Thus (13) is proved.

Let X be an arbitrary set such that $\operatorname{card} X \geqq \operatorname{card} |\mathfrak{C}|$. Let h: $X \cup \{d_b | b \in |\mathfrak{B}|\} \to$ $\to |\mathfrak{C}|$ and g: $\{d_b | b \in |\mathfrak{B}|\} \to |\mathfrak{B}|$ be two onto mappings, such that for all $b \in |\mathfrak{B}|$, $h(d_b) = b'$ and $g(d_b) = b$. Such mappings h and g exist. Let us form the free structures $\mathfrak{C}' = \mathfrak{Fr}_h (\mathfrak{C}, b')_{b \in |\mathfrak{B}|}$ and $\mathfrak{B}' = \mathfrak{Fr}_g (\mathfrak{B}, b)_{b \in |\mathfrak{B}|}$. By Theorem 2.4, $\mathfrak{C}' \models \Sigma \cup \Delta_{\mathfrak{B}}^{+\mathrm{ef}}$ and $\mathfrak{B}' \models T \cup \Delta_{\mathfrak{B}}^{+\mathrm{ef}}$. We shall show that $\mathfrak{B}' \leqq \mathfrak{C}'$. Obviously, $|\mathfrak{B}'| \subset |\mathfrak{C}'|$, and for all $b \in |\mathfrak{B}|$

$$C_{c_b}^{(\mathfrak{B}')} = C_{c_b}^{(\mathfrak{C}')} \tag{14}$$

is immediate by Definition 2.3. Similarly, for every $f \in \mathscr{F}$, $t_{\mathscr{F}}(f) = n+1$, and $b_0, \ldots, b_n \in |\mathfrak{B}'|$, we have

$$F_f^{(\mathfrak{B}')}(b_0, \ldots, b_n) = F_f^{(\mathfrak{C}')}(b_0, \ldots, b_n). \tag{15}$$

It follows from (14) and (15), that for any closed term τ in the diagram language of $\mathfrak{B}$, the equation

$$\tau^{(\mathfrak{B}')} = \tau^{(\mathfrak{C}')} \tag{16}$$

holds.

Let $r \in \mathscr{R}$, $t_{\mathscr{R}}(r) = n+1$, $b_0, \ldots, b_n \in |\mathfrak{B}'|$. By the definition of $\mathfrak{B}'$, we can find closed terms $\tau_0, \ldots, \tau_n$ of the diagram language of $\mathfrak{B}$, such that $b_0 = \tau_0^{(\mathfrak{B}')}, \ldots, b_n = \tau_n^{(\mathfrak{B}')}$. Hence, the following chain of implications is obtained:

$$\langle b_0, \ldots, b_n \rangle \in R_r^{(\mathfrak{B}')} \Rightarrow \langle \tau_0^{(\mathfrak{B}')}, \ldots, \tau_n^{(\mathfrak{B}')} \rangle \in R_r^{(\mathfrak{B}')} \Rightarrow \mathfrak{B}' \models r(\tau_0, \ldots, \tau_n) \Rightarrow$$
$$\Rightarrow r(\tau_0, \ldots, \tau_n) \in \Delta_{\mathfrak{B}}^{+\mathrm{ef}} \Rightarrow (\mathfrak{C}, b')_{b \in |\mathfrak{B}|} \models r(\tau_0, \ldots, \tau_n).$$

Using Theorem 2.4 again, we can continue:

$$(\mathfrak{C}, b')_{b \in |\mathfrak{B}|} \models r(\tau_0, \ldots, \tau_n) \Rightarrow \mathfrak{C}' \models r(\tau_0, \ldots, \tau_n) \Rightarrow \langle \tau_0^{(\mathfrak{C}')}, \ldots, \tau_n^{(\mathfrak{C}')} \rangle \in R_r^{(\mathfrak{C}')}$$

from which $\langle b_0, \ldots, b_n\rangle \in R_r^{(\mathfrak{C}')}$ follows.

By Lemma 3.5(iv), we see that $\mathfrak{B}' \leqq \mathfrak{C}'$.

Since $\mathfrak{B}' \models T$ and T is preserved under weak extensions, we have $\mathfrak{C}' \models T$, and by Theorem 2.4, $\mathfrak{C} \models T$. By (13), $\mathfrak{A} \models T$, which was to be proved.

(ii)⇒(i): Immediate by Lemma 4.2. □

Using a similar (but somewhat simpler) argument, one can prove the dual of this theorem

Theorem 4.7. *Let T be an equality-free consistent first order theory. Then, the following assertions are equivalent:*

(i) *T is preserved under weak substructures;*

(ii) *T has a set of universal negative equality-free axioms.*

5. This section is devoted to answering question (2) in the particular case when equality is excluded from the language.

Definition 5.1. Let $K \subset \mathfrak{M}^t$.

(i) K is said to be *closed under finite meets* iff for arbitrary $\mathfrak{A}_0, \ldots, \mathfrak{A}_n \in K$, if $\mathfrak{A}_0 \cap \ldots \cap \mathfrak{A}_n$ exists, then $\mathfrak{A}_0 \cap \ldots \cap \mathfrak{A}_n \in K$.

(ii) K is *closed under extensions (weak extensions)* iff for arbitrary $\mathfrak{A} \in K$ and $\mathfrak{B} \in \mathfrak{M}^t$, $\mathfrak{A} \subset \mathfrak{B}$ $(\mathfrak{A} \leqq \mathfrak{B})$ entails $\mathfrak{B} \in K$.

Obviously, if T is an arbitrary first order theory and "OPERATION" stands for one of the following items: "finite meets", "extensions", and "weak extensions", then the assertion "T is preserved under OPERATION" is equivalent to the assertion "K is closed under OPERATION where $K = \{\mathfrak{A} | \mathfrak{A} \models T\}$".

Definition 5.2. By a *filter of structures* we shall mean a nonvoid class $K \subset \mathfrak{M}^t$ such that K is closed under both finite meets and weak extensions.

The following assertion characterizes filters of structures from a model theoretical point of view.

Theorem 5.3. *Let T be an arbitrary equality-free first order theory and let K be the class of all models of T. Then the following two assertions are equivalent:*

(i) *T has a set of quantifier-free atomic equality-free axioms;*

(ii) *K is a filter of structures.*

Proof. First we note that both (i) and (ii) imply that T is consistent.

(i)⇒(ii): It is obvious that every equality-free quantifier-free atomic sentence can be considered as an existential positive equality-free sentence and as a universal equality-free Horn sentence, simultaneously. Thus, T is preserved under both weak extensions and finite meets by Lemma 4.2 and Theorem 1.1, respectively; whence

K is closed under both weak extensions and finite meets; i.e. K is a filter of structures (for $K \neq \emptyset$).

(ii)$\Rightarrow$(i): Let us suppose, that K is a filter of structures, i.e. that K is closed under finite meets and weak extensions. It follows that T is preserved under finite meets and weak extensions.

Let $\Gamma = \{\varphi | \varphi$ is an equality-free, quantifier-free atomic sentence, $T \models \varphi\}$. Obviously, $T \models \Gamma$. We shall prove that $\Gamma \models T$.

Let $\mathfrak{C} \models \Gamma$ be arbitrary and set $\Sigma = \{\neg\sigma | \sigma$ is an equality-free, quantifier-free atomic sentence such that $\mathfrak{C} \models \neg\sigma\}$.

Let $\neg\sigma \in \Sigma$ be arbitrary. Then $\{\neg\sigma\} \cup T$ is consistent, for otherwise we would have $T \models \neg(\neg\sigma)$, i.e. $T \models \sigma$, and so $\sigma \in \Gamma$; from which the contradiction $\mathfrak{C} \models \sigma$ would follow.

Let $\{\neg\sigma_0, \ldots, \neg\sigma_m\} \subset \Sigma$, and for every i, $0 \leqq i \leqq m$, let $\mathfrak{B}_i$ be a model of $\{\neg\sigma_i\} \cup T$. Let X be any set such that $\operatorname{card} X \geqq \operatorname{card} |\mathfrak{B}_0| \cup \ldots \cup \operatorname{card} |\mathfrak{B}_m|$, and let g_i: $X \cup \mathscr{C} \to |\mathfrak{B}_i|$ be an onto mapping for each i, $0 \leqq i \leqq m$. Let us consider the free structures $\mathfrak{Fr}_{g_i} \mathfrak{B}_i$, $0 \leqq i \leqq m$. It is immediate by Definition 2.3, that $\mathfrak{B} = \mathfrak{Fr}_{g_0} \mathfrak{B}_0 \sqcap \ldots \sqcap \mathfrak{Fr}_{g_m} \mathfrak{B}_m$ exists; moreover, for any i, $0 \leqq i \leqq m$, $\mathfrak{Fr}_{g_i} \mathfrak{B}_i \models \{\neg\sigma_i\} \cup T$, by Theorem 2.4. Since T is preserved under finite meets, and σ_i is atomic, we have for every i, $0 \leqq i \leqq m$ that $\mathfrak{B} \models \{\neg\sigma_i\} \cup T$, i.e. $\mathfrak{B} \models \{\neg\sigma_0, \ldots, \neg\sigma_m\} \cup T$. By compactness, $\Sigma \cup T$ is consistent.

Let $\mathfrak{D}$ be a model of $\Sigma \cup T$. If ψ is an arbitrary equality-free, quantifier-free atomic sentence such that $\mathfrak{C} \not\models \psi$, then $\neg\psi \in \Sigma$, hence $\mathfrak{D} \not\models \psi$. It follows that for any equality-free, quantifier-free atomic sentence ψ, $\mathfrak{D} \models \psi$ implies $\mathfrak{C} \models \psi$.

Let Y be an arbitrary set such that $\operatorname{card} Y \geqq \operatorname{card} |\mathfrak{C}| \cup \operatorname{card} |\mathfrak{D}|$ and let h_1: $Y \cup \mathscr{C} \to |\mathfrak{C}|$, h_2: $Y \cup \mathscr{C} \to |\mathfrak{D}|$ be two onto mappings for which $h_1(c) = C_c^{(\mathfrak{C})}$, and $h_2(c) = C_c^{(\mathfrak{D})}$, for any $c \in \mathscr{C}$. Considering the free structures $\mathfrak{Fr}_{h_1} \mathfrak{C}$ and $\mathfrak{Fr}_{h_2} \mathfrak{D}$ we still have for any equality-free, quantifier-free atomic sentence ψ, that $\mathfrak{Fr}_{h_2} \mathfrak{D} \models \psi$ entails $\mathfrak{Fr}_{h_1} \mathfrak{C} \models \psi$. By Definition 2.3 and Lemma 3.5(iv), $\mathfrak{Fr}_{h_2} \mathfrak{D} \leqq \mathfrak{Fr}_{h_1} \mathfrak{C}$. But $\mathfrak{Fr}_{h_2} \mathfrak{D} \models T$ (by Theorem 2.4) and T is preserved under weak extensions, hence $\mathfrak{Fr}_{h_1} \mathfrak{C} \models T$. By Theorem 2.4, $\mathfrak{C} \models T$. $\square$

From a purely formalist point of view one may adopt the following notion:

Definition 5.4. By a *quasi-filter of structures* we mean a class $K \subset \mathfrak{M}^t$ such that K is closed under finite meets and ordinary extensions.

The analogue of Theorem 5.3 for this concept reads as follows.

Theorem 5.5. *Let T be an arbitrary equality-free first order theory and let K be the class of all models of T. Then, the following two assertions are equivalent:*

(i) *T has a set of quantifier-free equality-free Horn axioms;*

(ii) *K is a quasi-filter of structures.*

Proof. Similar to the proof of Theorem 5.3. □

We note that none of Theorems 5.3 and 5.5 generalize for theories with equality. Let us consider for example the theory $T=\{c_1=d_1 \vee c_2=d_2\}$, where c_1, c_2, d_1, d_2 are constant symbols. It is trivial that T is preserved under finite meets and weak extensions, by definition. Hence, K, the class of all models of T, is a filter of structures. T, however, has neither an atomic nor a Horn set of axioms in general, thus Theorem 5.3 is not true for this theory. Since every filter of structures is a quasi-filter of structures, Theorem 5.5 is false for T, too.

References

[1] H. Andréka, I. Németi, Generalization of the concept of variety and quasivariety to partial algebras through category theory, *Dissertationes Math.* (Rozprawy Matematyczne) CCIV (1983).

[2] C. C. Chang, H. J. Keisler, *Model Theory*, North-Holland (Amsterdam—London, 1973).

[3] P. Ecsedi-Tóth, A characterization of quasi-varieties in equality-free languages, *Acta Sci. Math.*, 47 (1984), 41—54.

[4] G. Gratzer, *General Lattice Theory*. Akademie-Verlag (Berlin, 1978).

[5] L. Henkin, J. D. Monk, A. Tarski, *Cylindric Algebras. Part I*, North-Holland (Amsterdam—London, 1971).

[6] I. Németi, I. Sain, Cone-implicational subcategories and some Birkhoff-type theorems, *Coll. Math. Soc. J. Bolyai 29*, North-Holland (Amsterdam—London, 1981), 535—578.

SZKI
DONÁTI u. 35—45
H—1015 BUDAPEST, HUNGARY

Acta Sci. Math., 51 (1987), 403—411

A classification of the set of linear functions in prime-valued logic

IVAN STOJMENOVIĆ

1. Introduction

Let $P_k = \bigcup_{n \in \omega} \{f \mid f\colon E_k^n \to E_k\}$, where $E_k = \{0, 1, \ldots, k-1\}$; i.e. P_k denotes the set of all k-valued logical functions. A subset G of P_k is said to be *closed* if it is closed under superposition (e.g. see [4]).

Let $H \subset P_k$ be a fixed closed set. If $F \subseteqq H$ then we say that

(i) F is *complete in* $H \Leftrightarrow$ every element of H is obtained from F by superposition;

(ii) F is *H-maximal* $\Leftrightarrow$ F is closed and no G exists such that $F \subset G \subset H$ (proper inclusion) and G is closed;

(iii) F is a *base in* $H \Leftrightarrow F$ is finite and complete in H and no complete subset of F exists;

(iv) F is a *pivotal set in* $H \Leftrightarrow F$ is finite and for every $f \in F$ there is an H-maximal F' such that $f \notin F'$ but $F - \{f\} \subseteqq F'$.

From these definitions it follows that a base is a complete pivotal set of functions.

The *rank* of a base (pivotal set) is the number of elements of the base (pivotal set).

Let m be the cardinality of the set of all H-maximal sets and suppose that this set is well-ordered. There exist closed sets H for which m is not finite ([5]). If m is finite then a subset F of H is complete in H iff F is not contained in any H-maximal set ([4]).

If $f \in H$, then the class a_f determined by f is an element of $\{0, 1\}^m$ such that $a_i = 0$ iff $f \in H_i$, where a_i is the i-th component of a_f and H_i is the i-th H-maximal set $(1 \leqq i \leqq m)$ in the well-ordering mentioned above. For $F \subseteqq H$, one can define the class a_F determined by F as the union of classes determined by the elements of F. Therefore, if $F = \{f_1, \ldots, f_s\}$ then $a_F = \{a_{f_1}, \ldots, a_{f_s}\}$. This set a_F can be represented as an element a'_F of $\{0, 1\}^m$ such that $a'_F = \vee(a_{f_1}, \ldots, a_{f_s})$, where bitwise OR operation $\vee$ is defined in the following way: the i-th component $a_F^{(i)}$ of a'_F is equal to 0 iff the i-th component of all classes a_{f_j} $(1 \leqq j \leqq s)$ is equal to 0.

Received March 20, 1984.

From this definition it follows that the set F is complete iff $a'_F = 1^m$. Also, we infer that F is a pivotal set if $a'_F \neq a'_{F \setminus \{f_j\}}$ for all j, $1 \leq j \leq s$. From these considerations one can remark that if F is complete (pivotal set, base), $f, f' \in F$ and $a_f = a_{f'}$ (i.e. f and f' are functions of the same class) then $F \cup \{f'\} \setminus \{f\}$ is complete (pivotal set, base) and $a_F = a_{F \cup \{f'\} \setminus \{f\}}$.

All P_2-maximal sets and maximal sets of P_2-maximal sets are described in [10]. P_3-maximal sets are determined in [4], and maximal sets of P_3-maximal sets are exhibited in [7] and other papers.

All different classes a_f for the set P_2 are investigated in [6], and for P_3 in [8], [9] and [11].

Let us recall some well-known closed sets in P_k.

The set L_k of *linear* functions is defined in the following way:

$$L_k = \Big\{a_0 + \sum_{i=1}^{n} a_i x_i \pmod{k} \,\Big|\, a_0 \in E_k,\ a_i \in E'_k,\ 1 \leq i \leq n,\ n \in \omega, \text{ where } E'_k = E_k \setminus \{0\}\Big\}.$$

Let $a = \sum_{i=1}^{n} a_i$. It is well-known that L_k is a P_k-maximal set iff k is a prime number ([4]).

The set S_k of *selfdual* functions is defined as follows:

$$S_k = \{f \mid f(x_1+1, \ldots, x_n+1) = f(x_1, \ldots, x_n)+1 \pmod{k},\ n = 1, 2, \ldots\}.$$

$T_k^j = \{f \mid f(j, \ldots, j) = j\}$ is the set of functions preserving j $(0 \leq j \leq k-1)$.

Let $\overline{X} = L_k \setminus X$ for each $X \subset L_k$. The intersection of the sets $X_1, \ldots, X_i \subset L_k$ will be denoted by $X_1 \ldots X_i$.

From the results in papers [1], [2], [3] it follows

Theorem 1. *Let $p \in \omega$ be an arbitrary prime. Then there are $p+2$ L-maximal sets. These are:*

(i) $L^j = L_p T_p^j$, *for every* j, $j = 0, 1, \ldots, p-1$,

(ii) $L^p = L_p S_p = \{a_0 + \sum_{i=1}^{n} a_i x_i \mid a = 1 \pmod{p}\}$, *the set of linear selfdual functions,*

(iii) $L^{(1)} = \{a_0 + a_1 x \mid a_0, a_1 \in E_p\}$, *the set of unary linear functions.*

Let 0^t denote the sequence $\underbrace{00\ldots0}_{t}$, and 1^t denote $\underbrace{11\ldots1}_{t}$.

In this paper we prove that there exist $2p+4$ different classes determined by functions of L_p. The number of different classes determined by bases in L_p is $4\binom{p+1}{2}$, and the number of different classes determined by pivotal noncomplete sets of L_p is $\binom{p+4}{2} - 2$.

2. Classification of L_p

Theorem 2. *Let $p\in\omega$ be an arbitrary prime. Then there are $2p+4$ different classes (denoted by $c_1, c_2, \dots, c_{2p+4}$) of functions in L_p. These classes and the corresponding sets of functions are:*

$$L^0L^1\dots L^{p-1}L^pL^{(1)}, \quad c_1 = 0^{p+2};$$

$$L^0L^1\dots L^{p-1}L^p\bar{L}^{(1)}, \quad c_2 = 0^{p+1}1;$$

$$\bar{L}^0\bar{L}^1\dots\bar{L}^{i-4}L^{i-3}\bar{L}^{i-2}\dots\bar{L}^pL^{(1)}, \quad c_i = 1^{i-3}0^{p+3-i}0, \quad \textit{where} \quad 3 \leqq i \leqq p+3;$$

$$\bar{L}^0\bar{L}^1\dots\bar{L}^{j-p-5}L^{j-p-4}\bar{L}^{j-p-3}\dots\bar{L}^p\bar{L}^{(1)}, \quad c_j = 1^{j-p-4}01^{2p+5-j},$$

$$\textit{where} \quad p+4 \leqq j \leqq 2p+4.$$

Proof. Let $f(x_1, \dots, x_n)=a_0+\sum_{i=1}^{n} a_ix_i \pmod p$ and $\sum_{i=1}^{n} a_i=a$. Consider the equation $a_0+ay=y$.

Case a) Let $a_0=0$, $a=1$. Then the equation is $y=y$ which is satisfied by every y. This implies that $f\in L^0L^1\dots L^p$. The function $f(x)=x$ is in the set $L^{(1)}$, and it is a function of the class c_1. The function $a_1x_1+\dots+a_nx_n$ where $a=1$ and $n\geqq 2$ is in the set $\bar{L}^{(1)}$, and so it is a function of the class c_2.

Case b) $a_0\neq 0$, $a=1$. Then we obtain $a_0=0$, so it has no solution. Hence, the function f is in the set $\bar{L}^0\bar{L}^1\dots\bar{L}^{p-1}L^p$. The function a_0+x for $a_0\neq 0$ is in the set $L^{(1)}$ and it is a function of the class c_{p+3}. The function $a_0+a_1x_1+\dots+a_nx_n \pmod p$ for $a_0\neq 0$ and $a=1$, $n\geqq 2$ is in the set $\bar{L}^{(1)}$, and it is a function of the class c_{2p+4}.

Case c) $a\neq 1$. $y_1\neq y_2$ implies $(a-1)y_1\neq(a-1)y_2$. From this it follows that $(a-1)y$ takes on p different values, when y ranges from 0 to $p-1$. It follows that there exists exactly one y_0 such that $(a-1)y_0=-a_0$, i.e. $a_0+ay_0=y_0$. This implies that the function f is in the set L^{y_0}, and it is not in the sets L^i for $i\neq y_0$, $1\leqq i\leqq p-1$. Since $a\neq 1$, f is not in the set L^p. The function $f=i$ (constant) is in the set $L^{(1)}$, and it is a function of the class c_{i+3}. The function $f=i+ax_1+(p-a)x_2$ $(a\neq 0)$ is in the set $\bar{L}^{(1)}$ and it is a function of the class c_{p+4+i}.

Theorem is proved, because all possible cases have been considered.

3. Classes determined by bases of L_p

Theorem 3. *Let $p\in\omega$ be an arbitrary prime. Then the number of different classes determined by bases in L_p and the number of different classes determined by pivotal noncomplete sets in L_p for each rank are shown in the following table:*

rank	*bases*	*pivotal noncomplete sets*
1	0	$2p+3$
2	$3\binom{p+1}{2}$	$\binom{p+1}{2}+p+1$
3	$\binom{p+1}{2}$	0
$\geqq 4$	0	0

Proof. From the definitions it is easy to see that the class $c_1=0^{p+2}$ is not included in any pivotal set, and there is no base of rank 1. The classes $c_2, c_3, \ldots, c_{2p+4}$ are different from 0^{p+2} and 1^{p+2}. Hence, these classes define the classes determined by pivotal noncomplete sets of rank 1 of L_p.

We begin the investigation of bases and pivotal noncomplete sets of rank $\geqq 2$ by the following remarks:

$\vee(c_i, c_j) = 1^{p+1}0$ for $3 \leqq i, j \leqq p+3$;

$\vee(c_i, c_j) = 1^{p+2}$ for $p+4 \leqq i, j \leqq 2p+4$;

$\vee(c_2, c_i) \notin \{c_2, c_i, 1^{p+2}\}$ for $3 \leqq i \leqq p+3$;

$\vee(c_2, c_i) = c_i$ for $p+4 \leqq i \leqq 2p+4$;

$\vee(c_i, c_j) = 1^{p+2}$ for $3 \leqq i \leqq p+3$, $p+4 \leqq j \leqq 2p+4$ and $j \neq i+p+1$;

$\vee(c_i, c_{i+p+1}) = c_{i+p+1}$ for $3 \leqq i \leqq p+3$;

$\vee(c_2, c_i, c_j) = 1^{p+2}$ for $3 \leqq i < j \leqq p+3$.

From these remarks it follows that bases of rank 2 may contain any two functions of classes c_i and c_j, where i and j satisfy the condition $p+4\leqq i<j\leqq 2p+4$, or the conditions $3\leqq i\leqq p+3$, $p+4\leqq j\leqq 2p+4$ and $j\neq i+p+1$.

Also, one can infer that pivotal incomplete sets of rank 2 consist either of two functions of classes c_i and c_j, $3\leqq i<j\leqq p+3$, or a function of class c_2 and a function of class c_i, $3\leqq i\leqq p+3$.

From the remarks above it follows that no pivotal set of rank $\geqq 3$ exists which contains a function of class c_i for $p+4\leqq i\leqq 2p+4$. Hence, pivotal sets of rank $\geqq 3$ may contain only functions of the class c_2 and classes c_i for $3\leqq i\leqq p+3$. But, from the first remark we conclude that $\vee(c_{i_1}, c_{i_2}, \ldots, c_{i_s})=\vee(c_{i_1}, c_{i_2})=1^{p+1}0$ for

$3 \leqq i_1, \ldots, i_s \leqq p+3$. Hence, a pivotal set cannot contain functions from more than two classes c_i for $3 \leqq i \leqq p+3$. Therefore, no base or pivotal set of rank $\geqq 4$ exists. From $\vee(c_2, c_i, c_j) = 1^{p+2}$ $(3 \leqq i < j \leqq p+3)$ we conclude that pivotal sets of rank 3 are complete. Thus, no pivotal noncomplete set of rank 3 exists and a base of rank 3 consists of a function of class c_2 and two functions of the classes c_i and c_j, where $3 \leqq i < j \leqq p+3$.

From the above considerations the theorem follows.

Corollary 1. *The maximal rank of a base of the set L_p is 3, and the maximal rank of a pivotal noncomplete set is 2.*

Corollary 2. *There is no base of rank 1 (i.e. Sheffer function) in the set L_p.*

Corollary 3. *The number of different classes determined by bases in L_p (p prime) is $4\binom{p+1}{2}$.*

Corollary 4. *The number of different classes determined by pivotal noncomplete sets of L_p (p prime) is $\binom{p+1}{2}+3p+4=\binom{p+4}{2}-2$.*

The number of n-ary linear functions of class c_i $(1 \leqq i \leqq 2p+4)$ will be denoted by $t_n(i)$.

Theorem 4. $t_0(i)=1$ *for* $3 \leqq i \leqq p+2$, $t_0(i)=0$ *otherwise;* $t_1(1)=1$, $t_1(p+3)=p-1$, $t_1(i)=p-2$ *for* $3 \leqq i \leqq p+2$, $t_1(i)=0$ *otherwise;* $t_n(2)=((p-1)^n-(-1)^n)/p$, $t_n(2p+4)=(p-1)\,t_n(2)$; $t_n(i)=((p-1)^{n+1}+(-1)^n)/p$ *for* $p+4 \leqq i \leqq 2p+3$, $t_n(i)=0$ *otherwise* $(n \geqq 2)$.

Proof. The statement follows easily from considerations in the proof of Theorem 2. For $n=0$ and $n=1$ the assertion is obvious. For $n>1$ $t_n(2)$ is equal to the number of sequences $a_1, \ldots, a_n$ which satisfy the condition $a_1+\ldots+a_n=1 \pmod p$. If $a_1+\ldots+a_{n-1}=1 \pmod p$, then no solution of the equation $a_1+\ldots+a_n=1 \pmod p$ exists (since $a_i \neq 0$, $1 \leqq i \leqq n$). If $a_1+\ldots+a_{n-1} \neq 1 \pmod p$, then there exists exactly one solution of the equation $a_1+\ldots+a_n=1 \pmod p$. It follows that $t_n(2)=(p-1)^{n-1}-t_{n-1}(2)$, $t_2(2)=p-2$. By induction on n it is easy to prove that $t_n(2)=((p-1)^n-(-1)^n)/p$. If $p+4 \leqq i \leqq 2p+3$, then from $t_n(i)=(p-1)^n-t_n(2)$ we obtain $t_n(i)=((p-1)^{n+1}+(-1)^n)/p$.

The number of functions of the class c_i which depend on at most n variables is denoted by $t_{\leqq n}(i)$.

From Theorem 4 the following theorem is easily derived.

Theorem 5. $t_{\leqq 0}(i)=t_0(i)$; $t_{\leqq 1}(i)=t_0(i)+t_1(i)$;

$$t_{\leqq n}(1)=1, \quad t_{\leqq n}(i)=p-1 \quad \textit{for} \quad 3\leqq i\leqq p+3;$$

$$t_{\leqq n}(2)=((p-1)^{n+1}-(p-1)^2)/(p-2)-((-1)^n+1)/2)/p;$$

$$t_{\leqq n}(2p+4)=(p-1)t_{\leqq n}(2);$$

$$t_{\leqq n}(i)=((p-1)^2((p-1)^n-1)/(p-2)+(1+(-1)^n)/2)/p \quad \textit{for} \quad p+4\leqq i\leqq 2p+3.$$

Let B_i^n and P_i^n denote the number of bases and the number of pivotal incomplete sets of rank i which consist of functions depending on at most n variables.

From Theorems 2, 3 and 5 it is easy to prove the following

Theorem 6.

$$B_2^n=pt_{\leqq n}(2p+4)t_{\leqq n}(p+4)+\binom{p}{2}t^2_{\leqq n}(p+4)+pt_{\leqq n}(p+3)t_{\leqq n}(p+4)+$$

$$+pt_{\leqq n}(2p+4)t_{\leqq n}(3)+t_{\leqq n}(p+3)t_{\leqq n}(2p+4)+p^2t_{\leqq n}(3)t_{\leqq n}(p+4);$$

$$B_3^n=t_{\leqq n}(2)\left(pt_{\leqq n}(p+3)t_{\leqq n}(3)+\binom{p}{2}t^2_{\leqq n}(3)\right);$$

$$P_1^n=t_{\leqq n}(2)+t_{\leqq n}(p+3)+t_{\leqq n}(2p+4)+pt_{\leqq n}(3)+pt_{\leqq n}(p+4);$$

$$P_2^n=t_{\leqq n}(2)(t_{\leqq n}(p+3)+pt_{\leqq n}(3))+pt_{\leqq n}(p+3)t_{\leqq n}(3)+\binom{p}{2}t_{\leqq n}(3);$$

$$B_1^n=B_4^n=B_5^n=\ldots=P_3^n=P_4^n=\ldots=0.$$

Analogously one can obtain the numbers of bases and pivotal noncomplete sets which contain functions depending on exactly n variables.

4. Classification of L_p-maximal sets

We may assume further that $p\geqq 3$ (prime number). The properties of L_2-maximal sets follow immediately from Post's lattice ([10]).

Let us define some familiar closed sets in L_p: $L^{(0)}=\{0, 1, \ldots, p-1\}$, $L_{s0}=L^pL^0=\{a_1x_1+\ldots+a_nx_n|a=1,\ n=1, 2, \ldots\}$, $L_i^{(1)}=L^{(1)}L^i=\{a_0+a_1x\,|\,a_0+a_1i=i,\ a_0, a_1\in E_p\}$ for $0\leqq i\leqq p-1$, $L_p^{(1)}=L^{(1)}L^p=\{x, x+1, \ldots, x+k-1\}$.

We shall mean by the multiplicative order of $a\in E_p'$ the least integer $r(a)=r\geqq 1$ for which $a^r=1$ holds. If $p-1$ is divisible by j, then E_p' has $\varphi(j)$ elements of order j ($\varphi(j)$ denotes Euler's φ-function). Let $a_{i0}+a_ix\in L^{(1)}\setminus L^{(0)}$, $i\geqq 1$, $r(a_i)=r_i$. Let us denote by lcm $(r_1, r_2, \ldots)$ the least common multiple of the numbers $r_1, r_2, \ldots$.

Let the number $p-1$ have the decomposition to powers of primes $p-1=q_1^{\alpha_1}q_2^{\alpha_2}\ldots q_u^{\alpha_u}$ with all $q_1=2<q_2<\ldots<q_u$ primes, $\alpha_i\geqq 1, p_i=(p-1)/q_i$ and $L^{(1,i)}=\{a_0+ax|r(a)(\geqq 1)$ divides $p_i\}$, $i=1, 2, \ldots, u$.

The maximal sets for all L_p-maximal sets are determined by Demetrovics and Bagyinszki ([2]).

Theorem 7 ([2]). *There are exactly two L^i-maximal sets for* $0 \leqq i \leqq p$ *(p is a prime number):* L_{s0} *and* $L_i^{(1)}$.

Theorem 8. *If* $1 \leqq i \leqq p$ *then there are exactly four different classes determined by functions in* L^i, *two different classes determined by bases of* L^i *(one for both of ranks* 1 *and* 2*) and two different classes determined by pivotal noncomplete sets in* L^i *(both of them are of rank* 1*).*

Proof. The function x belongs to the set $L_{s0}L_i^{(1)}$ and the function $2x+(p-1)y$ is in the set $L_{s0}\bar{L}_i^{(1)}$. The function $x+1$ is an element of the set $\bar{L}_{s0}L_p^{(1)}$ and the function $2x-i$ is in the set $\bar{L}_{s0}L_i^{(1)}$ for $1 \leqq i \leqq p-1$. Base functions $x+y+(p-i)$ and $2x+(p-1)y+1$ ([2]) belong to the sets $\bar{L}_{s0}\bar{L}_i^{(1)}$ for $0 \leqq i \leqq p-1$ and $\bar{L}_{s0}\bar{L}_p^{(1)}$ respectively. Thus all four possible classes determined by the functions in L^i are nonempty. The other parts of the theorem follow immediately.

We are going to investigate classes determined by functions in $L^{(1)}$.

Theorem 9 ([2]). *The following* $u+p+1$ *sets are* $L^{(1)}$*-maximal:*

$$L^{(1,i)} \cup L^{(0)}, \quad i = 1, 2, \ldots, u,$$
$$L_i^{(1)} \cup L^{(0)}, \quad i = 0, 1, \ldots, p-1,$$
$$L^{(1)} \setminus L^{(0)}.$$

The next three lemmas are useful for the classification of $L^{(1)}$.

Lemma 1. *For the elements of* $L^{(1)}$ *we have:*

(a) $a_0+x \in L_i^{(1)}$ *iff* $a_0=0$, *for* $i=0, 1, \ldots, p-1$;

(b) *If* $a>1$ *then for each* i $(0 \leqq i \leqq p-1)$ *there exists exactly one* a_0 *for which* $a_0+ax \in L_i^{(1)}$;

(c) $a_0 \in L_i^{(1)}$ *iff* $a_0=i$.

The proof is omitted.

Lemma 2. $L_i^{(1)}L_j^{(1)}=\{x\}$ *for* $0 \leqq i<j \leqq p-1$.

Proof. From $a_0+a_1 i=i$ and $a_0+a_1 j=j$ it follows that $a_1=1$ and $a_0=0$.

Lemma 3. *Let* t_i *be a sequence such that* $t_i=q_i$ *or* $t_i=1$ *for each* $i=1, 2, \ldots, u$, $t=(p-1)/(t_1 \ldots t_u)$ *and* a *is a number for which* $r(a)=t$. *If we define the sets* A_i $(1 \leqq i \leqq u)$ *such that* $A_i=\bar{L}^{(1,i)}$ *for* $t_i=1$ *and* $A_i=L^{(1,i)}$ *for* $t_i=q_i$ *then the function* $f=a_0+ax$ *is in the set* $A_1A_2 \ldots A_u$.

Proof. If $t_i=1$ then p_i is not divisible by $r(a)$. Hence $a_0+ax \in \bar{L}^{(1,i)}=A_i$. If $t_i=q_i$ then $r(a)$ divides p_i. Thus $a_0+ax \in L^{(1,i)}=A_i$.

Theorem 10. *The number of different classes determined by functions in* $L^{(1)}$ *is* $p2^u+3$ *if* $p-1\neq q_1q_2\dots q_u$ *and* $p(2^u-1)+3$ *otherwise.*

Proof. Suppose that the $L^{(1)}$-maximal sets are ordered as in Theorem 9. $0^{u+p}1$ is the class determined by the functions in the set $L^{(1)}$. The $(u+p+1)$-component of all other classes is 0. From Lemmas 1—3 we infer that the class 0^{u+p+1} is determined only by the function x and the class 0^u1^p0 is determined by functions a_0+x for $a_0\neq 0$. We may assume further that $f=a_0+ax$ and $a>1$. From Lemma 2 it follows that exactly one component among the components $u+1, u+2, \dots, u+p$ is equal to 0. We derive from Lemma 3 that all the 2^u possible classes with respect to the first u $L^{(1)}$-maximal sets are nonempty. But, if $p-1=q_1\dots q_u$ for $t_i=q_i$ $(1\leqq i\leqq u)$ we get $t=a=1$ in Lemma 3. It follows from Lemma 1 (b) and Lemma 2 that each of these classes with respect to the first u $L^{(1)}$-maximal sets can be supplemented to a class determined by functions in $L^{(1)}$ in p different ways.

The proof is complete.

Corollary 5. *Each base of* $L^{(1)}$ *contains a constant.*

Corollary 6. *For* $p=3$ *there are exactly* 6 *classes determined by functions in* $L^{(1)}$: $0^5, 0^41, 01^30, 10110, 11010, 1^300$.

Theorem 11 ([2]). *The cardinality of the bases of* $L^{(1)}$ *is* $\geqq 3$.

Theorem 12. *The maximal rank of classes determined by bases in* $L^{(1)}$ *is* $u+2$.

Proof. Each base of $L^{(1)}$ contains a function of the class $0^{u+p}1$. There is a subset of the base containing no more than u functions for which bitwise OR gives the value 1^u with respect to the first u components. From Lemmas 1 and 3 we obtain that no more than one component among components $u+1, \dots, u+p$ has the value 0. Hence, except the $u+1$ functions considered above, this base may contain at most one function. Thus, each base of $L^{(1)}$ consists of at most $u+2$ functions.

Theorem 13. *If* $p-1=q_1^{\alpha_1}$ *(for example, if* $p=3$ *or* $p=5$*) then each base in* $L^{(1)}$ *contains exactly three functions.*

Proof. In the case $p-1=q_1^{\alpha_1}$ we have $u=1$ and so this theorem is proved by using Theorems 11 and 12.

Acknowledgement. The author is thankful for the comments given by the referee which have certainly improved the readibility of the paper.

References

[1] J. BAGYINSZKI, J. DEMETROVICS, The structure of the maximal linear classes in prime-valued logics, *C.R. Math. Rep. Acad. Sci. Canada*, **2** (1980), 209—213.

[2] J. BAGYINSZKI, J. DEMETROVICS, The lattice of linear classes in prime-valued logics, *Banach Center Publ.*, PWN, 8/1979.

[3] J. BAGYINSZKI, J. DEMETROVICS, Lineáris osztályok szerkezete prímszám értékű logikában, *MTA SzTAKI Közlemények*, 16/1976, 25—52.

[4] S. V. JABLONSKIĬ, Functional constructions in k-valued logics, *Trudy Mat. Inst. Steklov*, **51** (1958), 5—142. (Russian)

[5] YU. I. YANOV, A. A. MUČNIK, Existence of k-valued closed classes without a finite basis, *Dokl. Akad. Nauk. SSSR*, **127** (1959), 44—46. (Russian)

[6] L. KRNIĆ, Classes of bases of propositional logic, *Glas. Mat.*, **20** (1965), 23—32. (Russian)

[7] D. LAU, Submaximale Klasses von P_3, *Elektron. Informationsverarb. Kybernet.*, **18** (1982), 227—243.

[8] M. MIYAKAWA, Functional completeness and structure of three-valued logic I. Classification of P_3, *Researches of Electrotechnical Laboratory*, 717, (Tokyo, 1971) 1—85.

[9] M. MIYAKAWA, Enumeration of bases of three-valued logical functions, in: *Finite Algebra and Multiple-Valued Logic* (Proc. Conf. Szeged, 1979), Colloq. Math. Soc. János Bolyai, 28, North-Holland (Amsterdam, 1981), 469—487.

[10] E. POST, *Two-valued Iterative Systems of Mathematical Logic*, Ann. of Math. Stud., 5, Princeton Univ. Press (1941).

[11] I. STOJMENOVIĆ, Classification of P_3 and the enumeration of bases of P_3, *Rev. of Res. Fac. of Sci. Math. Ser.*, **14** (1984), 73—80.

INSTITUTE OF MATHEMATICS
UNIVERSITY OF NOVI SAD
21000 NOVI SAD, DR. ILIJE DJURIČIĆA 4
YUGOSLAVIA

Acta Sci. Math., 51 (1987), 413—421

On superalgebras of the polydisc algebra

RAUL E. CURTO*, PAUL S. MUHLY*, TAKAHIKO NAKAZI** and T. YAMAMOTO**

Let $\mathbf{T}$ be the unit circle and, for $n \geqq 1$, let A_n be the uniform closure in $C(\mathbf{T}^2)$ of the algebra of polynomials in $z^k w^l$, where k and l are integers, $l \geqq 0$, and $k \geqq 0$ whenever $0 \leqq l \leqq n-1$. Each A_n contains the polydisc algebra and the intersection of the A_n is the polydisc algebra. In this paper we give a characterization of the subspaces of $L^2(\mathbf{T}^2)$ which are invariant under multiplication by the functions in A_n. The characterization is somewhat complicated, as one would expect, since for $n > 1$, A_n is not a Dirichlet algebra. In fact, for $n > 1$, the point in the maximal ideal space of A_n represented by Lebesgue measure on $\mathbf{T}^2$ has an infinite dimensional set of representing measures. Nevertheless, as a result of our analysis, we find that each simply invariant subspace of $L^2(\mathbf{T}^2)$ for A_n is finitely generated and the number of generators required is $\leqq n$. Examples can be constructed where n generators are necessary. Our analysis enables us to extend results of the third author and to parametrize the weak-$*$ closed superalgebras of A_n.

1. Introduction

Let X be a compact Hausdorff space, let $C(X)$ be the space of complex-valued continuous functions on X, and let A be a uniform algebra on X. For $\varphi \in M_A$, the maximal ideal space of A, set $A_0 = \{f \in A \colon \varphi(f) = 0\}$.

Definition 1.1. Let $\varphi \in M_A$, let σ be a representing measure (on X) for φ, and let $\mathfrak{M}$ be a (closed) subspace of $L^2(X, \sigma)$. Then $\mathfrak{M}$ is said to be *simply invariant* (for A) if $A\mathfrak{M} \subset \mathfrak{M}$, but $[A_0\mathfrak{M}]_2 \neq \mathfrak{M}$ (where $[\ ]_2$ denotes L^2-closure).

Let ∂_A denote the Shilov boundary of A and N_φ denote the set of representing measures for $\varphi \in M_A$ whose support is contained in ∂_A. Note that N_φ is a weak-$*$

* Supported in part by a National Science Foundation grant (U.S.A.).
** Supported in part by Kakenhi (Japan).
Received July 10, 1984; revised May 20, 1985.

compact convex set of probability measures on ∂_A. The general theory of simply invariant subspaces is known only in the case when $N_\varphi \cap L^1(X, \sigma)$ is finite dimensional. For instance, if A is a Dirichlet algebra then $N_\varphi \cap L^1(X, \sigma) = \{\sigma\}$, and the simply invariant subspaces of $L^2(X, \sigma)$ have been characterized (cf. [2, p. 132]). In particular, Beurling's theorem can be derived from that characterization (the disc algebra, after all, is a Dirichlet algebra on the unit circle $\mathbf{T}$).

In this note we focus our attention on the following class of function algebras, A_n, $n \geqq 1$, contained in $C(\mathbf{T}^2)$. The general theory of invariant subspaces does not apply to these algebras. Nevertheless, as we shall show, it is possible to give a fairly complete and concrete description of their invariant subspaces.

Definition 1.2. Let $\mathbf{T}^2$ be the 2-torus and let n be an integer, $n \geqq 1$. By A_n we shall denote the uniform algebra on $\mathbf{T}^2$ of all continuous functions on $\mathbf{T}^2$ that can be uniformly approximated by polynomials in $z^k w^l$, where $l \geqq 0$, and $k \geqq 0$ when $0 \leqq l \leqq n-1$.

Equivalently, A_n may be described as the set of all functions f in $C(\mathbf{T}^2)$ such that $\hat{f}$ is supported in the upper half-plane and, in the second quadrant, $\hat{f}$ is supported on or above the line $y=n$. We have $A_1 \supsetneqq A_2 \supsetneqq \dots$ and $\bigcap_{n=1}^{\infty} A_n = A_\infty$, the polidisc algebra. Observe that A_n is a Dirichlet algebra precisely when $n=1$. Let σ be the Haar measure on $\mathbf{T}^2$ and define

$$\varphi_n(f) = \int_{\mathbf{T}^2} f \, d\sigma \quad (f \in A_n).$$

Clearly, $\varphi_n \in M_{A_n}$ and $\sigma \in N_{\varphi_n}$ for all n. Also note that $\partial A_n = \mathbf{T}^2$ for all n. However, $N_{\varphi_n} \cap L^1(\mathbf{T}^2, \sigma)$ is not finite dimensional for $n \geqq 2$, as may be seen quite easily.

Our hope is that an understanding of the A_n's will help us understand better the polydisc algebra A_∞. After all, in one obvious sense, A_∞ is the limit of the A_n. In another somewhat more vague sense, as we shall see, it appears that the lattice of invariant subspaces of A_∞ is approximated by the invariant subspace lattices of the A_n. The following proposition, however, shows that in still another sense all the A_n, $n < \infty$, are similar to A_1. Observe that for $n < \infty$, $w^{n-1} A_1 \subset A_n$ and therefore, $|A_n| = |A_1|$, where $|A_n| = \{|f| : f \in A_n\}$. However, $|A_\infty| \subsetneqq |A_1|$. Since $A_n \subset A_1$, there is a natural embedding ϱ_n of M_{A_1} into M_{A_n}, given by restriction. Similarly, $\varrho_\infty \colon M_{A_1} \to M_{A_\infty}$ is an embedding.

Proposition 1.3. *For each finite n, $\varrho_n \colon M_{A_1} \to M_{A_n}$ is surjective, while $\varrho_\infty \colon M_{A_1} \to M_{A_\infty}$ is not surjective.*

Proof. Let $\varphi \in M_{A_n}$. There are two possibilities: $|\varphi(z)| = 1$ or $|\varphi(z)| < 1$. In the first case, define $\tilde{\varphi}(\bar{z}^k w^l) = \overline{\varphi(z)}^k \varphi(w)^l$ $(k \geqq 0, l \geqq 1)$. Then $\tilde{\varphi} \in M_{A_1}$ and

$\tilde{\varphi}|_{A_n}=\varphi$. If $|\varphi(z)|<1$ then $\varphi(w^n)=\varphi(z^k)\varphi(\bar{z}^k w^n)$ for all $k\geqq 0$, so that $|\varphi(w)|^n\leqq\leqq|\varphi(z)|^k$ (all k), which implies that $\varphi(w)=0$. By [1, Theorem 5], φ has a unique extension $\tilde{\varphi}$ to A_1, and $\tilde{\varphi}\in M_{A_1}$. Therefore, $\varrho_n(\tilde{\varphi})=\varphi$. For the second assertion, observe that the proof just given shows that if $\varphi\in M_{A_1}$ and $|\varphi(z)|<1$ then $\varphi(w)=0$. It clearly follows that ϱ_∞ cannot be onto because M_{A_∞} can be identified with the bidisc $\mathbf{D}\times\mathbf{D}$.

Definition 1.4. We shall let $\mathscr{A}$, $\mathscr{B}$, and $\mathscr{C}$ denote the following subalgebras of $C(\mathbf{T}^2)$:

i) $\mathscr{A}$ is the uniform closure of the polynomials in the first variable z;

ii) $\mathscr{B}$ is the uniform closure of the polynomials in z, $\bar{z}$, and w;

and

iii) $\mathscr{C}$ is the uniform closure of the polynomials in z and $\bar{z}$.

Observe that:

i) $\mathscr{A}$ is isomorphic to the disc algebra;

ii) $\mathscr{C}$ is isomorphic to $C(\mathbf{T})$;

iii) $\mathscr{B}$ is isomorphic to the tensor product of the disc algebra and $C(\mathbf{T})$;

iv) $\mathscr{B}$ is also the uniform closure of $\bigcup_{k=0}^{\infty}\bar{z}^k A_n$ (all n);

v) $A_n\subsetneqq\mathscr{B}$ (all n);

vi) $\bigcap_{k=0}^{\infty} z^k A_n=w^n\mathscr{B}$ (all n);

vii) $\mathscr{B}=(\sum_{l=0}^{n-1}\oplus w^l\mathscr{C})\oplus w^n\mathscr{B}$ (all n); and

viii) $A_n=(\sum_{l=0}^{n-1}\oplus w^l\mathscr{A})\oplus w^n\mathscr{B}$ (all n).

Definition 1.5. The closure in $L^2(\sigma)$ of A_n, $[A_n]_2$, will be denoted H_n^2 and the closure of $\mathscr{B}$ in $L^2(\sigma)$ will be denoted $\mathbf{H}^2$. Likewise, we define $H_n^\infty=[A_n]_*$ and $\mathbf{H}^\infty=[\mathscr{B}]_*$, where $[\ ]_*$ denotes weak-$*$ closure in $L^\infty(\sigma)$. For $p=2,\infty$, we set $H_{n,0}^p=\{f\in H_n^p|\int f\,d\sigma=0\}$. Finally, we define $\mathscr{L}^2=[\mathscr{C}]_2$, $\mathscr{L}^\infty=[\mathscr{C}]_*$, $\mathscr{H}^2=[\mathscr{A}]_2$, and $\mathscr{H}^\infty=[\mathscr{A}]_*$.

Observe that for $p=2,\infty$, $\mathscr{L}^p$ and $\mathscr{H}^p$ are spaces of functions in the first variable, z, only, while the splittings described above yield the decompositions

$$\mathbf{H}^p=(\sum_{l=0}^{n-1}\oplus w^l\mathscr{L}^p)\oplus w^n\mathbf{H}^p\quad(\text{all } n),$$

and

$$H_n^p=(\sum_{l=0}^{n-1}\oplus w^l\mathscr{H}^p)\oplus w^n\mathbf{H}^p.$$

These decompositions are crucial to our analysis. In Section 2 we use them to describe completely the non-simply invariant subspaces of $L^2(\sigma)$, and in Section 3 we use them to describe the simply invariant subspaces of $L^2(\sigma)$. Finally, in Section 4, we use them to determine the structure of the weak-$*$ closed superalgebras of H_n^∞.

2. Non-simply invariant subspaces

For $n<\infty$, H_n^2 is a simply invariant subspace (for A_n) while $\mathbf{H}^2$ is not. The following proposition gives an easy criterion to determine when an invariant subspace is simply invariant. First, we list some important properties of the algebras $A_{n,0}$:

(i) $A_{1,0}=zA_1$, and

(ii) $A_{n,0}=zA_n+[w, w^2, \ldots, w^{n-1}]$, where [ ] denotes linear span.

Proposition 2.1. *Let $\mathfrak{M}$ be an invariant subspace of $L^2(\sigma)$. Then $\mathfrak{M}$ is simply invariant for A_n if and only if $z\mathfrak{M}\subsetneqq\mathfrak{M}$.*

Proof. If $n=1$, $[A_{1,0}\mathfrak{M}]_2=[zA_1\mathfrak{M}]_2$, so that if $[A_{1,0}\mathfrak{M}]_2=\mathfrak{M}$, then $z\mathfrak{M}=\mathfrak{M}$. Conversely, if $z\mathfrak{M}=\mathfrak{M}$, then $[A_1z\mathfrak{M}]_2=[A_1\mathfrak{M}]_2=\mathfrak{M}$. If $n\neq 1$, $[A_{n,0}\mathfrak{M}]_2= =[z\mathfrak{M}+w\mathfrak{M}+\ldots+w^{n-1}\mathfrak{M}]_2$, by (ii) above, and therefore $[A_{n,0}\mathfrak{M}]_2=[z\mathfrak{M}+w\mathfrak{M}]_2$. Hence if $\mathfrak{M}$ is simply invariant, then $z\mathfrak{M}\subsetneqq\mathfrak{M}$. Assume now that $[A_{n,0}\mathfrak{M}]_2=\mathfrak{M}$. Then from what we have just seen, $[z\mathfrak{M}+w\mathfrak{M}]_2=\mathfrak{M}$. Consequently, $[z\mathfrak{M}+w^n\mathfrak{M}]_2= =[z(\mathfrak{M}+w^{n-1}\mathfrak{M})+w^n\mathfrak{M}]_2=[z\mathfrak{M}+w^{n-1}(z\mathfrak{M}+w\mathfrak{M})]_2=[z\mathfrak{M}+w^{n-1}\mathfrak{M}]_2$.

By repeating this argument, we find that $[z\mathfrak{M}+w^n\mathfrak{M}]_2=[z\mathfrak{M}+w\mathfrak{M}]_2=\mathfrak{M}$. But $\bar{z}w^n\in A_n$, so $\mathfrak{M}=[z\mathfrak{M}+w^n\mathfrak{M}]_2=z[\mathfrak{M}+\bar{z}w^n\mathfrak{M}]_2=z\mathfrak{M}$, as desired.

Corollary 2.2. *Let $\mathfrak{M}$ be an invariant subspace of $L^2(\sigma)$. Then $\mathfrak{M}$ is not simply invariant if and only if*

$$\mathfrak{M} = \chi_{E_1}q\mathbf{H}^2\oplus\chi_{E_2}L^2(\sigma),$$

where χ_{E_1} and χ_{E_2} denote the characteristic functions of two measurable sets E_1 and E_2, $\chi_{E_1}\in\mathscr{L}^\infty$, $\chi_{E_1}+\chi_{E_2}\leqq 1$, and $|q|=1$ a.e. (σ).

Proof. The sufficiency is clear. If $z\mathfrak{M}=\mathfrak{M}$, then $\mathfrak{M}$ is invariant under $\mathscr{B}$. Since $\mathscr{B}$ contains the Dirichlet algebra A_1 on which σ is multiplicative, we may apply [6, First example, p. 165] to conclude that $\mathfrak{M}$ is of the form $\mathfrak{M}=\chi_E q[D]_2$, where $D=\{f\in L^\infty\colon f\mathfrak{M}\subset\mathfrak{M}\}$, q is unimodular, and $\chi_E\in D$. By [5, Example 3.(1)], D has the form $D=\chi_F\mathbf{H}^\infty+(1-\chi_F)L^\infty$, where $\chi_F\in\mathbf{H}^\infty$. Letting $\chi_{E_1}=\chi_E\chi_F$ and $\chi_{E_2}=\chi_E(1-\chi_F)$, we see [5] that $\chi_{E_1}\in\mathscr{L}^\infty$ and $\mathfrak{M}$ has the desired representation.

An alternate proof of this result may be based on [4] as follows. Since $z\mathfrak{M}=\mathfrak{M}$, $\mathfrak{M}$ is invariant under $\mathbf{H}^\infty$. But $\mathbf{H}^\infty$ may be viewed as the non-self-adjoint crossed product determined by the identity automorphism of $L^\infty(\mathbf{T})$. Hence the result follows from the analysis in Section 3 of [4] (see in particular Theorem 3.3 and Proposition 3.4).

Before we proceed, we need a definition.

Definition 2.3. Let $\mathfrak{M}$ be an invariant subspace of $L^2(\sigma)$. Then we define $\mathfrak{M}_{-\infty}$ to be $[\bigcup_{k\geq 0} \bar{z}^k\mathfrak{M}]_2$ and $\mathfrak{M}_\infty$ to be $[\bigcap_{k\geq 0} z^k\mathfrak{M}]_2$.

Clearly $\mathfrak{M}_\infty\subset\mathfrak{M}\subset\mathfrak{M}_{-\infty}$. Moreover, both $\mathfrak{M}_\infty$ and $\mathfrak{M}_{-\infty}$ are non-simply invariant. By Corollary 2.2 we can describe both $\mathfrak{M}_\infty$ and $\mathfrak{M}_{-\infty}$. However, if $\mathfrak{M}$ is simply invariant, more can be said.

Proposition 2.4. *Let $\mathfrak{M}$ be a simply invariant subspace. Then $\mathfrak{M}_{-\infty}=q_1\mathbf{H}^2$ and $\mathfrak{M}_\infty=q_2\mathbf{H}^2$, where q_1 and q_2 are unimodular.*

Proof. By Proposition 2.1, $z\mathfrak{M}\subsetneqq\mathfrak{M}$, so $\mathfrak{M}_\infty\subsetneqq\mathfrak{M}\subsetneqq\mathfrak{M}_{-\infty}$. By Corollary 2.2,

$$\mathfrak{M}_{-\infty}=\chi_{E_1}q_1\mathbf{H}^2\oplus\chi_{E_2}L^2, \quad \text{with} \quad \chi_{E_1}+\chi_{E_2}\leqq 1, \quad |q_1|=1,$$

and

$$\mathfrak{M}_\infty=\chi_{F_1}q_2\mathbf{H}^2\oplus\chi_{F_2}L^2, \quad \text{with} \quad \chi_{F_1}+\chi_{F_2}\leqq 1, \quad |q_2|=1.$$

Since $\mathfrak{M}_\infty\subset\mathfrak{M}\subset\mathfrak{M}_{-\infty}$, it follows that $\chi_{F_1}+\chi_{F_2}\leqq\chi_{E_1}+\chi_{E_2}$ and $\chi_{F_2}\leqq\chi_{E_2}$. Since $\bar{z}^kw^n\in A_n$ for all $k\geqq 0$ and $A_n\mathfrak{M}\subset\mathfrak{M}$, we see that $\bar{z}^kw^n\mathfrak{M}\subset\mathfrak{M}$ for all $k\geqq 0$. Therefore, $w^n\mathfrak{M}_{-\infty}\subset\mathfrak{M}$, thus $w^n\mathfrak{M}_{-\infty}=w^nz^k\mathfrak{M}_{-\infty}\subset z^k\mathfrak{M}$ for all $k\geqq 0$, so $w^n\mathfrak{M}_{-\infty}\subset\subset\bigcap_{k\geq 0} z^k\mathfrak{M}=\mathfrak{M}_\infty$. Consequently, $w^n\chi_{E_2}L^2\subset\chi_{F_2}L^2$, and so $\chi_{E_2}=\chi_{F_2}$. Likewise, $\chi_{E_1}=\chi_{F_1}$, because $w^n\chi_{E_1}q_1\mathbf{H}^2\subset\chi_{F_1}q_2\mathbf{H}^2$. Thus we find that $\mathfrak{M}_{-\infty}\ominus\mathfrak{M}_\infty==\chi_{E_1}(q_1\mathbf{H}^2\ominus q_2\mathbf{H}^2)$ which, in turn, is contained in $\chi_{E_1}q_1(\mathbf{H}^2\ominus w^n\mathbf{H}^2)$, since $w^n\mathfrak{M}_{-\infty}\subset\mathfrak{M}_\infty$. Set $\mathfrak{M}_0=\mathfrak{M}\ominus\mathfrak{M}_\infty$. Then since $z\mathfrak{M}_\infty=\mathfrak{M}_\infty$, but $z\mathfrak{M}\subsetneqq\mathfrak{M}$, it follows that $z\mathfrak{M}_0\subsetneqq\mathfrak{M}_0$. If f is a nonzero function in $\mathfrak{M}_0\ominus z\mathfrak{M}_0$, then for all $k>0$, we have $0=(f,z^kf)=\iint_{T^2}|f(e^{i\theta},e^{i\varphi})|^2e^{-ik\theta}\,d\theta\,d\varphi$. Since $|f|$ is real, this implies that $\int_T|f(e^{i\theta},e^{i\varphi})|^2\,d\varphi$ is constant, a.e., in θ. Since f is nonzero and χ_{E_1} is a function of θ alone, we conclude that $\chi_{E_1}=1$. Thus $\mathfrak{M}_{-\infty}=q_1\mathbf{H}^2$ and $\mathfrak{M}_\infty=q_2\mathbf{H}^2$, as promised.

Remark 2.5. When $\mathfrak{M}=H_n^2$, we see that $\mathfrak{M}_{-\infty}=\mathbf{H}^2$ while $\mathfrak{M}_\infty=w^n\mathbf{H}^2$.

3. Simply invariant subspaces

Suppose that $\mathfrak{N}$ is a simply invariant subspace such that $w^l\mathbf{H}^2=\mathfrak{N}_\infty\subset\mathfrak{N}\subset\subset\mathfrak{N}_{-\infty}=\mathbf{H}^2$ where $1\leqq l\leqq n$. Then applying Lax's generalization of Beurling's theorem, we find that $\mathfrak{N}$ has a very special form. Specifically, using [3, VI.3, p. 60], we see that there is a $j\leqq l$ and there are functions $f_{ik}\in\mathscr{L}^2$, $1\leqq i\leqq j$, $0\leqq k\leqq l-1$, such that

a) $\sum_{k=0}^{l-1} f_{ij}\overline{f_{mk}} = \delta_{im}$, $1 \leqq i, m \leqq j$, and

b) $\mathfrak{N} = [z;\ f_1, \ldots, f_j]_2 \oplus w^l\mathbf{H}^2$

where $f_i = \sum_{k=0}^{l-1} f_{ik} w^k$, $1 \leqq i \leqq j$, and where $[z; f_1, \ldots, f_j]_2$ denotes the smallest subspace containing $f_1, f_2, \ldots, f_j$ that is invariant under multiplication by z. For instance, it is clear that

$$H_l^2 = [z;\ 1, w, \ldots, w^{l-1}]_2 \oplus w^l\mathbf{H}^2 \quad \text{and} \quad H_{l,0}^2 = [z;\ z, w, \ldots, w^{l-1}]_2 \oplus w^l\mathbf{H}^2.$$

If, now, F is a unimodular function and if $\mathfrak{M} = F\mathfrak{N}$, where $\mathfrak{N}$ is of the above form, then $\mathfrak{M}$ is easily seen to be simply invariant, but of course, $\mathfrak{M}$ need no longer be nestled between some $w^l\mathbf{H}^2$ and $\mathbf{H}^2$. Our goal, Theorem 3.2, is to show that every simply invariant subspace can be expressed in this way as $F\mathfrak{N}$.

Proposition 3.1. *Let $\mathfrak{M}$ be a simply invariant subspace for A_n and (for $n \geqq 2$) assume that $A_{n-1}\mathfrak{M} \not\subset \mathfrak{M}$. Then $\mathfrak{M} = F\mathfrak{N}$ where F is a unimodular function on $\mathbf{T}^2$ and $\mathfrak{N}$ is a simply invariant subspace such that $\mathfrak{N}_\infty = w^n\mathbf{H}^2$ and $\mathfrak{N}_{-\infty} = \mathbf{H}^2$.*

Proof. By Proposition 2.4, $\mathfrak{M}_{-\infty} = q_1\mathbf{H}^2$ and $\mathfrak{M}_\infty = q_2\mathbf{H}^2$, where $|q_1| = |q_2| = 1$. Since $q_2\mathbf{H}^2 \subset q_1\mathbf{H}^2$, we must have $\bar{q}_1 q_2 \in \mathbf{H}^2$ and $q_1\bar{q}_2 w^n \in \mathbf{H}^2$ (recall that $w^n\mathfrak{M}_{-\infty} \subset \subset \mathfrak{M}_\infty$). Set $q = \bar{q}_1 q_2$, so that $q \in \mathbf{H}^2$ and $w^n\bar{q} \in \mathbf{H}^2$. Therefore $q = \sum_{k=0}^{n} c_k w^k$, where $c_k \in \mathscr{L}^2$. Since $|q| = 1$, we have $q = \sum_{k=0}^{n} a_k \chi_{E_k} w^k$, where each a_k is a function of z alone, $|a_k| = 1$ a.e. on E_k, $0 \leqq k \leqq n$, and $\sum_{k=0}^{n} \chi_{E_k} = 1$. Since $q_1 q\mathbf{H}^2 \subset \mathfrak{M} \subset q_1\mathbf{H}^2$, we see that $\chi_{E_0} q_1\mathbf{H}^2 = \chi_{E_0} q_1 q\mathbf{H}^2 \subset \chi_{E_0}\mathfrak{M} \subset \chi_{E_0} q_1\mathbf{H}^2$, and therefore, $\chi_{E_0}\mathfrak{M} = = \chi_{E_0} q_1 q\mathbf{H}^2 \subset q_1 q\mathbf{H}^2 = \mathfrak{M}_\infty$. Now we may assume that $\chi_{E_0} \not\equiv 1$, for otherwise $\mathfrak{M} = \mathfrak{M}_\infty$ and so $\mathfrak{M}$ is not simply invariant. Moreover, $\chi_{E_0}\mathfrak{M} \subset \mathfrak{M}$ and, if $\chi_{E_0} \not\equiv 0$, then it is easy to see that $\bar{z}\mathfrak{M} \subset \mathfrak{M}$, so that $\mathfrak{M}$ is not simply invariant. (Indeed, on the basis of the Wold decomposition for an isometry, it is straightforward to show that if a subspace $\mathfrak{M}$ is invariant for a unitary operator U and if $\mathfrak{M}$ is also invariant for some nontrivial spectral projection of U, then $\mathfrak{M}$ reduces U. In our special situation, χ_{E_0} is a spectral projection for multiplication by z since χ_{E_0} is a function of z alone.) Thus $\chi_{E_0} \equiv 0$. Put $D = \{f \in L^\infty \colon f\mathfrak{M} \subset \mathfrak{M}\}$. Then $H_n^\infty \subset D$ and $q\mathbf{H}^\infty \subset D$, since $q\mathbf{H}^\infty\mathfrak{M} \subset q_1 q\mathbf{H}^2 = q_2\mathbf{H}^2 \subset \mathfrak{M}$. Hence $w\mathscr{H}^\infty$ and (since $\bar{a}_k\chi_{E_1} \in \mathbf{H}^\infty$) $w\chi_{E_1}\mathscr{L}^\infty$ are both contained in D. Since $\mathscr{L}^\infty$ is isomorphic to $L^\infty(\mathbf{T})$ with $\mathscr{H}^\infty$ corresponding to $H^\infty(\mathbf{T})$, it follows that if $\chi_{E_1} \not\equiv 0$, then $[\mathscr{H}^\infty + \chi_{E_1}\mathscr{L}^\infty]_* = \mathscr{L}^\infty$. But then $w\mathscr{L}^\infty \subset D$ and $H_1^\infty \subset D$. Thus $A_{n-1}\mathfrak{M} \subset \mathfrak{M}$, a contradiction. Thus, $\chi_{E_1} \equiv 0$. One shows similarly that $\chi_{E_2} = \ldots = \chi_{E_{n-1}} \equiv 0$ and $\chi_{E_n} \equiv 1$. Therefore, $q = w^n a_n$. Set $F = q_1$ and $\mathfrak{N} = \bar{q}_1\mathfrak{M}$ to complete the proof.

Theorem 3.2. *Let $\mathfrak{M}$ be a simply invariant subspace for A_n. Then $\mathfrak{M}=F\mathfrak{N}$ for some unimodular function F and a simply invariant subspace $\mathfrak{N}$ such that $\mathfrak{N}_{-\infty}=\mathbf{H}^2$ and $\mathfrak{N}_{\infty}=w^l\mathbf{H}^2$ for some l, $1\leqq l\leqq n$. Moreover, $\mathfrak{M}\cap Fw^{l-1}\mathbf{H}^2=Fw^{l-1}qH_1^2$, where q is a unimodular function in $\mathscr{L}^\infty$.*

Proof. Let l, $1\leqq l\leqq n$, be the smallest integer such that $A_l\mathfrak{M}\subset\mathfrak{M}$. Proposition 3.1 then establishes the first part of the theorem. Now, $\mathfrak{M}\cap Fw^{l-1}\mathbf{H}^2=F(\mathfrak{N}\cap w^{l-1}\mathbf{H}^2)$, and $\mathfrak{N}\cap w^{l-1}\mathbf{H}^2=q\mathscr{H}^2w^{l-1}\oplus w^l\mathbf{H}^2$, because $\bar{w}^{l-1}((\mathfrak{N}\cap w^{l-1}\mathbf{H}^2)\ominus w^l\mathbf{H}^2))$ is a simply invariant subspace of $\mathscr{L}^2$ under multiplication by z. Therefore the second part of the theorem follows.

The following corollary is of course well known since A_1 is a Dirichlet algebra. However, our methods provide an alternate proof.

Corollary 3.3. *If $n=1$ and $\mathfrak{M}$ is a simply invariant subspace, then $\mathfrak{M}=FH_1^2$ for some unimodular function F.*

Proof. Obviously, l must be 1 in this case, so that $\mathfrak{N}=\mathfrak{N}\cap\mathbf{H}^2=qH_1^2$, which implies that $\mathfrak{M}=F\mathfrak{N}=FH_1^2$.

Corollary 3.4. *Let $\mathfrak{M}$ be a simply invariant subspace for A_n. Then* $\dim(\mathfrak{M}\ominus z\mathfrak{M})=1$ *if and only if $\mathfrak{M}=FH_1^2$ for some unimodular function F.*

Proof. The sufficiency is clear. By Theorem 3.2, $\mathfrak{M}=\tilde{F}\mathfrak{N}$ for some unimodular function $\tilde{F}$ and a simply invariant subspace $\mathfrak{N}$ such that $\mathfrak{N}_{-\infty}=\mathbf{H}^2$ and $\mathfrak{N}_\infty=w^l\mathbf{H}^2$ for some l, $1\leqq l\leqq n$. We claim that $l=1$. This will give the desired result, as in the proof of the previous corollary. Since $\dim(\mathfrak{M}\ominus z\mathfrak{M})=1$, we also have $\dim(\mathfrak{N}\ominus z\mathfrak{N})=1$, so that $\mathfrak{N}\ominus z\mathfrak{N}=[\mathbf{C}f]_2$ for some function $f=\sum_{k=0}^{l}f_kw^k$, where $f_k\in\mathscr{L}^2$ $(0\leqq k\leqq l)$. Since $\mathfrak{N}\supset\mathfrak{N}_\infty=w^l\mathbf{H}^2$, f must be orthogonal to w^l and therefore $f_l=0$. Moreover, $fw^{l-1}=f_0w^{l-1}+w^lg$, where $g\in\mathbf{H}^2$, so that $f_0w^{l-1}\in\mathfrak{N}\ominus\mathfrak{N}_\infty$. Now, $\mathfrak{N}\ominus\mathfrak{N}_\infty=[\bigcup_{i\geqq0}z^if]_2=[z;f]_2$, and there exists a sequence $\{g_m\}\subset\mathscr{H}^\infty$ such that $g_mf\to f_0w^{l-1}$ in L^2. By projecting onto $w^{l-1}\mathscr{L}^2$ we get: $g_mf_{l-1}\to f_0$. Assume that $l>1$. Then $g_m\sum_{k=0}^{l-2}f_kw^k=g_m(f-f_{l-1}w^{l-1})\to0$, and in particular, $g_mf_0\to0$. However, by the second part of Theorem 3.2 we must have $f_0w^{l-1}\in w^{l-1}qH_1^2$, or $f_0w^{l-1}=w^{l-1}qh$, where $|q|=1$, $q\in\mathscr{L}^\infty$ and $h\in\mathscr{H}^2$. Therefore $|f_0|=|h|$ a.e. If $f_0=0$ a.e. then $\mathfrak{N}_{-\infty}\subset w\mathbf{H}^2$, so that $|f_0|>0$ on a set of positive measure. That forces $|h|>0$ a.e. and then $|f_0|>0$ a.e. If $\{g_{m_i}\}$ is a subsequence such that $g_{m_i}f_0\to0$ a.e., the previous observation implies that $g_{m_i}\to0$ a.e., so that $g_{m_i}f_{l-1}\to0$ a.e., or $f_0=0$ a.e.. This contradiction establishes the original claim and completes the proof.

4. Weak-$*$ closed superalgebras

The following theorem generalizes [5, Theorem 4] (see [5, Example 3.(1)]).

Theorem 4.1. *Let B be a weak-$*$ closed subalgebra of L^∞ containing H_n^∞. Then either $B \subset \mathbf{H}^\infty$, or $B = \chi_E \mathbf{H}^\infty + (1-\chi_E)L^\infty$, for some measurable set E with $\chi_E \in \mathscr{L}^\infty$. If $B \subset \mathbf{H}^\infty$ then $\bigcap_{k \geq 0} z^k B = w^l \mathbf{H}^\infty$ for some l, $1 \leq l \leq n$.*

Proof. Put $B_{-\infty} = [\bigcup_{k\geq 0} \bar{z}^k B]_*$ and $B_\infty = \bigcap_{k \geq 0} z^k B$. Then $B_\infty \subset B \subset B_{-\infty}$. By [5, Lemma 1] and Corollary 2.2, $B_\infty = \chi_{E_1} q_1 \mathbf{H}^\infty + \chi_{E_2} L^\infty$, where $\chi_{E_1} \in \mathscr{L}^\infty$, $\chi_{E_1} + \chi_{E_2} = 1$, and $B_{-\infty} = \chi_{F_1} q_2 \mathbf{H}^\infty + \chi_{F_2} L^\infty$, with $\chi_{F_1} \in \mathscr{L}^\infty$ and $\chi_{F_1} + \chi_{F_2} = 1$. As in the case of invariant subspaces of L^2, $w^n B_{-\infty} \subset B_\infty$. Thus $w^n \chi_{F_2} L^\infty \subset \chi_{E_2} L^\infty$, and this implies $\chi_{E_2} = \chi_{F_2}$, because $\chi_{E_2} L^\infty \subset \chi_{F_2} L^\infty$. Since $B_{-\infty}$ is also an algebra and $q_2 \in B_{-\infty}$, we get $q_2 B_{-\infty} \subset B_{-\infty}$. Thus $B_{-\infty} \subset \bar{q}_2 B_{-\infty}$. This implies that $\chi_{E_1} B_{-\infty} \subset \subset \bar{q}_2 \chi_{E_1} q_2 \mathbf{H}^\infty = \chi_{E_1} \mathbf{H}^\infty$. In particular, $\chi_{E_1} B \subset \chi_{E_1} \mathbf{H}^\infty$. Put $D = \chi_{E_1} B + \chi_{E_2} \mathbf{H}^\infty$. Then D is a weak-$*$ closed superalgebra of H_n^∞ and $D \subset \mathbf{H}^\infty$. We shall consider two cases:

Case 1: $B \not\subset \mathbf{H}^\infty$. In this case $\chi_{E_2} \not\equiv 0$. Consequently (as in the proof of Proposition 3.1) $[\mathscr{H}^\infty + \chi_{E_2} \mathscr{L}^\infty]_* = \mathscr{L}^\infty$. We have $\mathscr{H}^\infty \subset B$, hence $D \supset \mathscr{H}^\infty + \chi_{E_2} \mathscr{L}^\infty$, and so $D \supset \mathscr{L}^\infty$. This implies $D \supset \mathbf{H}^\infty$, which yields $D = \mathbf{H}^\infty$. Now $\chi_{E_1} B = \chi_{E_1} D = \chi_{E_1} \mathbf{H}^\infty$. On the other hand, $\chi_{E_2} L^\infty = \chi_{E_2} B_\infty \subset \chi_{E_2} B \subset \chi_{E_2} B_{-\infty} \subset \chi_{E_2} L^\infty$. Consequently $\chi_{E_2} B = \chi_{E_2} L^\infty$, and so we can conclude $B = \chi_{E_1} \mathbf{H}^\infty + (1 - \chi_{E_1}) L^\infty$.

Case 2: $B \subset \mathbf{H}^\infty$. In this case $\chi_{E_2} \equiv 0$. Since $w^n \mathbf{H}^\infty \subset B \subset \mathbf{H}^\infty$ and $B_\infty = q_1 \mathbf{H}^\infty$, $q_1 = \sum_{j=0}^n \chi_{S_j} w^j$, where $\chi_{S_j} \in \mathscr{L}^\infty$, $0 \leq j \leq n$, and $\sum_{j=0}^n \chi_{S_j} = 1$. If $\chi_{S_0} \not\equiv 0$, then $B = \mathbf{H}^\infty$ because $\chi_{S_0} = q_1 \chi_{S_0} \in B$ and $zB \subset B$. If k is the first integer such that $\chi_{S_k} \not\equiv 0$ then $B \supset w^k H^\infty$ and $B_\infty = w^k \mathbf{H}^\infty$. For, if $\chi_{S_k} \equiv 1$ then $B \supset w^k \mathbf{H}^\infty$ trivially. If $\chi_{S_k} \not\equiv 1$ then $B \supset w^k \mathbf{H}^\infty$ because $w^k \chi_{S_k} \in B$ and $zB \subset B$. By the hypothesis on k, $q_1 = w^k$ and therefore $B_\infty = w^k \mathbf{H}^\infty$.

When $n=2$ in the above theorem, more can be said about B.

Theorem 4.2. *Let B be a weak-$*$ closed subalgebra of $\mathbf{H}^\infty$ containing H_2^∞, and assume that $\bigcap_{k \geq 0} z^k B = w^2 \mathbf{H}^\infty$. Then $B = \mathscr{H}^\infty \oplus w\bar{q}\mathscr{H}^\infty \oplus w^2 \mathbf{H}^\infty$, where q is an inner function.*

Proof. Consider $B_0 = B \cap \mathscr{L}^\infty$. B_0 is a weak-$*$ closed subalgebra of $\mathscr{L}^\infty$ containing $\mathscr{H}^\infty$; moreover, if $B_0 = \mathscr{L}^\infty$ then $\mathscr{L}^\infty \subset \bigcap_{k \geq 0} z^k B$, a contradiction. Therefore $B_0 = \mathscr{H}^\infty$, i.e., $\mathscr{H}^\infty \subset B$. Let P_1 be the orthogonal projection from $\mathbf{H}^2$ onto $w\mathscr{H}^2$. Since $\bigcap_{k \geq 0} z^k B = w^2 \mathbf{H}^\infty$ and $\mathscr{H}^\infty \subset B$, it follows that $P_1 B := \{P_1 f \colon f \in B\} \subset B$, and that $B=\mathcal{H}^\infty \oplus P_1 B \oplus w^2 \mathrm{H}^\infty$. Moreover, $P_1 B = w\mathfrak{M}_1$, where $\mathfrak{M}_1 := \{f \in \mathcal{L}^\infty : wf \in B\}$ is an $\mathcal{H}^\infty$-submodule of $\mathcal{L}^\infty$; $\mathfrak{M}_1$ is, therefore, of the form $\mathfrak{M}_1 = \bar{q}\mathcal{H}^\infty$, for some unimodular function $q \in \mathcal{L}^\infty$. Since $\mathcal{H}^\infty \subset \mathfrak{M}_1$, we easily get that q is inner. Thus, $P_1 B = w\bar{q}\mathcal{H}^\infty$.

References

[1] R. Arens, A Banach algebra generalization of conformal mappings of the disc, *Trans. Amer. Math. Soc.*, **81** (1956), 501—513.

[2] T. Gamelin, *Uniform algebras*, Prentice-Hall, Englewood Cliffs, N.J. (1969).

[3] H. Helson, *Lectures on invariant subspaces*, Academic Press (New York, 1964).

[4] M. McAsey, P. Muhly, K.-S. Saito, Nonselfadjoint crossed products. II, *J. Math. Soc. Japan*, 33 (1981), 485—495.

[5] T. Nakazi, Nonmaximal weak-* Dirichlet algebras, *Hokkaido Math. J.*, **5** (1976), 88—96.

[6] T. Nakazi, Invariant subspaces of weak-* Dirichlet algebras, *Pacific J. Math.*, **69** (1977), 151—167.

(R. C. AND P. M.)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF IOWA
IOWA CITY, IOWA 52242 U.S.A.

(T. N. AND T. Y.)
DEPARTMENT OF APPLIED MATHEMATICS
RESEARCH INSTITUTE OF APPLIED ELECTRICITY
HOKKAIDO UNIVERSITY
SAPPORO 060, JAPAN

Acta Sci. Math., 51 (1987), 423—433

An extension of the Lindeberg—Trotter operator-theoretic approach to limit theorems for dependent random variables

II. Approximation theorems with O-rates, applications to martingale difference arrays

PAUL L. BUTZER, HERIBERT KIRSCHFINK and DIETMAR SCHULZ

This is Part II of the paper [10]. The contents of Part I, particularly the notations and preliminary results, are assumed to be known. References are in alphabetical order in each part, a few of the basic papers of Part I being recalled here. The sections are numbered consecutively.

Whereas Part I is concerned with convergence assertions without as well as with o-rates for dependent r.v.'s, all established with the help of the conditional Trotter operator first defined there, the purpose of Part II is to deal with O-error estimates, not only for convergence in distribution but also for the uniform convergence of distribution functions. The specializations to martingales carried out in Section 8 enable one to compare the results with those of other authors. Firstly some modifications and corrections are made to Part I.

3. A generalization of the Trotter-operator for dependent r.v.'s. — A revisit. Let us recall the definition of the generalized Trotter operator in terms of the conditional expectation given in Section 3.

Definition 1. Let $X\in\mathfrak{L}(\Omega, \mathfrak{A}, P)$ and $\mathfrak{G}$ be an arbitrary sub-σ-algebra of $\mathfrak{A}$. The conditional Trotter operator $V_{X|\mathfrak{G}}\colon C_B\to C_B\times(\mathfrak{Z}(\Omega, \mathfrak{G}))$ of X relative to $\mathfrak{G}$ is defined for $f\in C_B$ by

$$V_{X|\mathfrak{G}}f(y):=\inf_{x\in A_\alpha(y,f)} E[f(X+x)|\mathfrak{G}]\quad (y\in\mathbf{R})$$

for an $\alpha>0$ with $\alpha\in Q$ (=set of rationals), where $A_\alpha(y,f):=\{x\in Q; f(x)>f(y), y\in B_{\alpha x}\}$, $B_{\alpha x}:=\{y\in\mathbf{R}^1,\ |x-y|<\alpha\}$. Again, $(V_{X|\mathfrak{G}}f)(y,\omega):=(V_{X|\mathfrak{G}}f(y))(\omega)$.

Received July 18, 1984, and in revised form December 23, 1986.

In comparison with Definition 1 of [10] the present version has been modified by the introduction of the infimum. This will assure that assertions to be derived with this operator theoretical approach are valid almost surely in $\omega\in\Omega$ not only for each fixed $y\in\mathbf{R}$ but uniformly in $y\in\mathbf{R}$. The space $\mathbf{R}^1$ endowed with the usual topology has a countable base, namely $\mathfrak{B}:=\{B_{\alpha x};\ \alpha, x\in Q,\ x>0\}$. Such spaces, namely complete, separable metric spaces are called "Polish spaces"; they are in particular Borel spaces. Now it is well known that each finite Borel-measure μ on a Polish space is a regular measure (see e.g. [16, p. 373]). This ensures the existence of a regular distribution $P_{X|\mathfrak{G}}$ which is in particular $\mathfrak{G}$-measurable for each fixed $B\in\mathfrak{B}$ as well as a measure on $\mathfrak{B}$ for each fixed $\omega\in\Omega$. Therefore the integral representation of the conditional expectation (2.12) of [10] holds. In view of these considerations, the above infimum is taken only countably often for all $y\in\mathbf{R}$, so uniformly for all $y\in\mathbf{R}$. The condition "$f(x)>f(y)$" assures that the conditional Trotter operator will coincide with the classical one in case $\mathfrak{A}(X)$ is independent of $\mathfrak{G}$.

Under this modification Lemma 2 and Corollary 1 of [10] is readily seen to be valid, Lemma 3 is superfluous, and Lemmas 4 (the case $n=2$ of La. 5) and 5 are to be replaced by

Lemma 5. *Let $(X_n)_{n\in\mathbf{N}}$ be a sequence of r.v.'s from $\mathfrak{L}(\Omega, \mathfrak{A}, P)$, $(\mathfrak{G}_n)_{n\in\mathbf{N}}$ a sequence of sub-σ-algebras from $\mathfrak{A}$ for which $\mathfrak{G}_0:=\{\Omega, \emptyset\}\subset\mathfrak{G}_1\subset\mathfrak{G}_2\subset\ldots\subset\mathfrak{G}_n\subset\ldots$.*

a) *For each $f\in C_B$ one has*

$$V_{X_1|\mathfrak{G}_1}(V_{X_2|\mathfrak{G}_2}(\ldots V_{X_n|\mathfrak{G}_n}f(\ )\ldots))(y,\omega)=$$

$$=(V_{X_1|\mathfrak{G}_1}V_{X_2|\mathfrak{G}_2}\ldots V_{X_n|\mathfrak{G}_n}f)(y,\omega)=(V_{S_n|\mathfrak{G}_1}f)(y,\omega)\quad \text{a.s.}\quad (y\in\mathbf{R};\ n\in\mathbf{N}).$$

If, in particular $\mathfrak{G}_1=\mathfrak{G}_0$, then

$$(V_{X_1|\mathfrak{G}_1}\ldots V_{X_n|\mathfrak{G}_n}f)(y,\omega)=V_{S_n}f(y,\omega)\quad a.s.\quad (y\in\mathbf{R};\ n\in\mathbf{N}).$$

b) *If $(Z_n)_{n\in\mathbf{N}}$ is a further sequence from $\mathfrak{L}(\Omega, \mathfrak{A}, P)$, it being assumed that the Z_n are independent amongst themselves as well as of the X_n, then for each $f\in C_B$,*

$$\|V_{S_n|\mathfrak{G}_1}f-V_{\sum_{k=1}^{n}Z_k}f\|\leqq\sum_{k=1}^{n}\|V_{X_k|\mathfrak{G}_k}f-V_{Z_k}f\|\quad (n\in\mathbf{N}).$$

If, in particular, $\mathfrak{G}_k=\mathfrak{G}_0$, all $k\in\mathbf{N}$, then

$$\|V_{S_n}f-V_{\sum_{k=1}^{n}Z_k}f\|\leqq\sum_{k=1}^{n}\|V_{X_k}f-V_{Z_k}f\|\quad (n\in\mathbf{N}).$$

Proof. a) First take $n=2$. Noting (2.7) and (2.11) of [10], the latter being valid only for $\mathfrak{G}\subset\mathfrak{G}'$, one readily has

$$(V_{X_1|\mathfrak{G}_1}V_{X_2|\mathfrak{G}_2}f)(y,\omega)=\big(V_{X_1|\mathfrak{G}_1}\{\inf_{\bar{x}\in A_\alpha(\cdot,f)}E[f(X_2+\bar{x}(\cdot))|\mathfrak{G}_2]\}\big)(y,\omega)=$$
$$=\inf_{x\in A_\alpha(y,V_{X_2|\mathfrak{G}_2}f)}\{E[f(X_1+X_2+x)|\mathfrak{G}_1](\omega)\}=V_{X_1+X_2|\mathfrak{G}_1}f(y,\omega).$$

The general result follows by induction, the particular case by Lemma 2e) of [10].

b) The proof follows immediately by Corollary 1 and Lemmas 1 and 2e) of [10], as well as by the following result: Let $U_1,\dots,U_n,\dots$ and $V_1,\dots,V_n,\dots$ be contraction endomorphisms of C_B such that U_iU_j is only defined for $i\leqq j$, but the V_i may commute amongst themselves, and $U_iV_j=V_jU_i$ for any i,j. Then for each $f\in C_B$

$$\|U_1\dots U_nf-V_1\dots V_nf\|\leqq\sum_{k=1}^{n}\|U_kf-V_kf\|\quad(n\in\mathrm{N}).$$

6. O-approximation theorems for convergence in distribution

6.1. General theorems. In the proofs of the O-estimates of Section 6 the hypotheses of the corresponding o-convergence theorems of Section 5 may either be weakened or partially dropped entirely. Thus Lindeberg conditions are not needed either for the sequences $(X_k)_{k\in\mathrm{N}}$ or $(Z_k)_{k\in\mathrm{N}}$; the conditional moments of the r.v.'s X_k relative to $\mathfrak{G}_k$ need only coincide with the moments of Z_k up to the order $r-1$ (compare (6.2)).

Theorem 7. *Let $(X_k)_{k\in\mathrm{N}}$ be a sequence of (possibly) dependent r.v.'s from $\mathfrak{L}(\Omega,\mathfrak{A},P)$, let $(\mathfrak{G}_k)_{k\in\mathrm{N}}$ be a sequence of sub-σ-algebras of $\mathfrak{A}$ with $\mathfrak{G}_0:=\{\Omega,\emptyset\}\subset\subset\mathfrak{G}_1\subset\mathfrak{G}_2\subset\dots\subset\mathfrak{G}_n\subset\dots$, and Z be a φ-decomposable r.v. with decomposition components Z_k, $k\in\mathrm{N}$. Assume that for an $r\in\mathrm{N}\setminus\{1\}$*

$$E[|X_k|^r|\mathfrak{G}_k]\leqq M_{k,r}\quad a.s.\quad(k\in\mathrm{N})\tag{6.1}$$

for some constant $M_{k,r}>0$ as well as $E[|Z_k|^r]<\infty$. If furthermore

$$E[X_k^j|\mathfrak{G}_k]=E[Z_k^j]\quad a.s.\quad(k,j\in\mathrm{N};\ 1\leqq j\leqq r-1),\tag{6.2}$$

then there holds for $f\in C_B$

$$\|V_{\varphi(n)S_n|\mathfrak{G}_1}f-V_Zf\|\leqq 2c_{2,r}\omega_r\Big(\Big[\frac{\varphi(n)^r}{(r-1)!}M(n)\Big]^{1/r};f;C_B\Big),\tag{6.3}$$

where

$$M(n):=\sum_{k=1}^{n}(M_{k,r}+E[|Z_k|^r]),\tag{6.4}$$

$c_{2,r}$ being the constant of (2.1) in [10].

Proof. In view of (2.7) and (2.8) one has for $f\in C_B$ and any $g\in C_B^r$,

$$(6.5)\qquad \left|\inf_{x\in A_\alpha(y,f)} E[f(\varphi(n)S_n+x)\,|\,\mathfrak{G}_1]-E[f(Z+y)]\right| \leqq$$
$$\leqq 2\|f-g\|+\left|\inf_{x\in A_\alpha(y,g)} E[g(\varphi(n)S_n+x)\,|\,\mathfrak{G}_1]-E[g(Z+y)]\right|.$$

Further, in view of Lemma 5b,

$$(6.6)\qquad \|V_{\varphi(n)S_n|\mathfrak{G}_1}g-V_Zg\| \leqq \sum_{k=1}^{n}\|V_{\varphi(n)X_k|\mathfrak{G}_k}g-V_{\varphi(n)Z_k}g\|.$$

Furthermore, there holds the estimate

$$(6.7)\qquad \left|\inf_{x\in A_\alpha(y,g)}\{E[g(\varphi(n)X_k+x)\,|\,\mathfrak{G}_k]\}-E[g(\varphi(n)Z_k+y)]\right| \leqq$$
$$\leqq \sup_{x\in A_\alpha(y,g)}\left\{\left|E[g(\varphi(n)X_k+x)\,|\,\mathfrak{G}_k]-E[g(\varphi(n)Z_k+x)]\right|\right\}.$$

On account of the integral representation (2.12), and Taylor's formula applied to $g(u+x)$ twice,

$$\left|E[g(\varphi(n)X_k+x)\,|\,\mathfrak{G}_k]-E[g(\varphi(n)Z_k+x)]\right|=$$
$$=\left|\int_{\mathbf{R}} g(u+x)\,dF_{\varphi(n)X_k}(u\,|\,\mathfrak{G}_k)(\omega)-\int_{\mathbf{R}} g(u+x)\,dF_{\varphi(n)Z_k}(u)\right| \leqq$$
$$\leqq \left|\int_{\mathbf{R}}\left\{\sum_{j=0}^{r-1}\frac{u^j}{j!}g^{(j)}(x)\right\}d\left(F_{\varphi(n)X_k}(u\,|\,\mathfrak{G}_k)-F_{\varphi(n)Z_k}(u)\right)\right|+$$
$$+\left|\int_{\mathbf{R}}\frac{1}{(r-2)!}\left[\int_0^1(1-t)^{r-2}\{g^{(r-1)}(x+tu)-g^{(r-1)}(x)\}u^{r-1}\,dt\right]dF_{\varphi(n)X_k}(u|\mathfrak{G}_k)\right|+$$
$$+\left|\int_{\mathbf{R}}\frac{1}{(r-2)!}\left[\int_0^1(1-t)^{r-2}\{g^{(r-1)}(x+tu)-g^{(r-1)}(x)\}u^{r-1}\,dt\right]dF_{\varphi(n)Z_k}(u)\right|.$$

Since $g\in C_B^r$, $g^{(r-1)}\in \mathrm{Lip}\,(1;1;C_B)$ with Lipschitz constant $L_g=\|g^{(r)}\|$. So for $0<t\leqq 1$, $\left|\{g^{(r-1)}(x+tu)-g^{(r-1)}(x)\}u^{r-1}\right|\leqq\|g^{(r)}\|\,|u|^r$. In view of (6.2) and (6.1) there holds

$$(6.8)\qquad \sum_{k=1}^{n}\left|E[g(\varphi(n)X_k+x)\,|\,\mathfrak{G}_k]-E[g(\varphi(n)Z_k+x)]\right|\leqq$$
$$\leqq \sum_{k=1}^{n}\sum_{j=0}^{r-1}\left|\frac{1}{j!}g^{(j)}(x)\left\{\int_{\mathbf{R}}u^j\,d[F_{\varphi(n)X_k}(u\,|\,\mathfrak{G}_k)-F_{\varphi(n)Z_k}(u)]\right\}\right|+$$
$$+\left|\frac{\|g^{(r)}\|}{(r-1)!}\sum_{k=1}^{n}\int_{\mathbf{R}}|u|^r\,d[F_{\varphi(n)X_k}(u\,|\,\mathfrak{G}_k)+F_{\varphi(n)Z_k}(u)]\right|\leqq$$
$$\leqq \|g^{(r)}\|\left|\frac{\varphi(n)^r}{(r-1)!}\sum_{k=1}^{n}(M_{k,r}+E[|Z_k|^r])\right|=\|g^{(r)}\|\,\frac{\varphi(n)^r}{(r-1)!}M(n).$$

All in all, noting (6.5), (6.6) and (6.7),

$$\|V_{\varphi(n)S_n}f(y)-V_Zf(y)\| \leqq 2 \inf_{g\in C_B^r}\left\{\|f-g\|+\|g^{(r)}\|_{C_B}\frac{\varphi(n)^r}{(r-1)!}\,M(n)\right\} \leqq$$

$$\leqq 2K\left(\frac{\varphi(n)^r}{(r-1)!}\,M(n);\, f;\, C_B;\, C_B^r\right).$$

This gives (6.3) in view of (2.1).

The proof of the following result is immediate, noting La. 2e.

Theorem 8. a) *If the hypotheses in Theorem 7 are satisfied, and if in particular* $\mathfrak{G}_1=\mathfrak{G}_0$, *then*

$$\|V_{\varphi(n)S_n}f-V_Zf\| \leqq 2c_{2,r}\omega_r\left(\left[\frac{\varphi(n)^r}{(r-1)!}\,M(n)\right]^{1/r};\, f;\, C_B\right). \tag{6.9}$$

b) *If further* $f\in \mathrm{Lip}\,(\alpha; r, C_B)$, $0<\alpha\leqq r$, *then the left side of* (6.9) *has the bound*

$$2c_{2,r}\,L_f\,\varphi(n)^\alpha M(n)^{\alpha/r}. \tag{6.10}$$

Remark 1. The basic condition (6.2) of Thm. 7, which together with the assumed monotonicity of the $\mathfrak{G}_k$ is the only condition upon which the dependency structure of the r.v.'s X_k in question is subjected, could be replaced by the much weaker order condition

$$\sum_{=1}^{n}|E[X_k^j|\mathfrak{G}_k]-E[Z_k^j]| = O\left(\frac{\varphi(n)^r}{r!}\,M(n)\right) \quad \text{a.s.} \quad (1\leqq j\leqq r-1;\; n\to\infty). \tag{6.2*}$$

This will also insure the estimate (6.8) as does condition (6.2).

A comparable weaker version is given in [9] or [5] in the case of a weak invariance principle for dependent random functions. A further paper [6] deals in more detail with conditions like (6.2)*, called pseudo-moment conditions (with orders).

Remark 2. Concerning the proofs of Theorem 1, and analogously of Thms. 2—6 of [10], it should be mentioned that they have to be modified and corrected by taking the definition of the conditional Trotter operator in the form given here and by using likewise the arguments involving inequalities (6.6) and (6.7) of the proof of Thm. 7. This will assure results comparable to Thms. 7 and 8 for "little-o-rates" when assuming Lindeberg conditions provided $(\mathfrak{G}_n)_{n\in\mathbf{N}}$ is additionally assumed to be a monotone non-decreasing sequence. In fact, assertion (2.11) needed here is only valid if $\mathfrak{G}\subset\mathfrak{G}'$. In regard to GOVINDARAJULU, cited in [10], the authors cannot follow the proof of his main Theorem 3.1, in particular the step involving the norms on p. 1016, since the conditional expectations occurring there only hold for each fixed $y\in\mathbf{R}$ a.s. in $\omega\in\Omega$.

6.2. The CLT and WLLN with O-rates. The following statements dealing with the CLT are applications of Theorems 7 and 8, the usual specialisations being carried out.

Theorem 9. *Let $(X_k)_{k\in\mathbf{N}}$, $(\mathfrak{G}_k)_{k\in\mathbf{N}}$ be given as in Theorem 7, and let X^* be a standard normally distributed r.v. Set $\sigma_k^2=\mathrm{Var}\,[X_k]$, $k\in\mathbf{N}$ and $s_n^2=\Sigma_{k=1}^n\sigma_k^2$, and assume that $E[|X_k|^r|\mathfrak{G}_k]\leqq M_{k,r}$ a.s. for some constant $M_{k,r}>0$ as well as that*

$$(6.11)\qquad E[X_k^j\,|\,\mathfrak{G}_k]=\sigma_k^jE[X^{*j}]\quad a.s.\quad (k,j\in\mathbf{N},\ 1\leqq j\leqq r).$$

a) *Under these hypotheses one has for any $f\in C_B$*

$$(6.12)\qquad \|V_{s_n^{-1}S_n|\mathfrak{G}_1}f-V_{X^*}f\|_{C_B}\leqq 2c_{2,r}\omega_r\left(\left[\frac{s_n^{-r}}{(r-1)!}\,\overline{M}(n)\right]^{1/r};\ f;\ C_B\right),$$

where $\overline{M}(n):=\Sigma_{k=1}^n(M_{k,r}+\sigma_k^rE[|X^*|^r])$.

b) *If $f\in\mathrm{Lip}\,(\alpha;r;C_B)$, $0<\alpha\leqq r$, and $\mathfrak{G}_1=\mathfrak{G}_0$, then*

$$(6.13)\qquad \|V_{s_n^{-1}S_n}f-V_{X^*}f\|_{C_B}\leqq 2c_{2,r}L_fs_n^{-\alpha}\overline{M}(n)^{\alpha/r}.$$

Concerning the proof, just as in that of Theorem 2 condition (2.4) is satisfied with $Z_k=\sigma_kX^*$, $\varphi(n)=s_n^{-1}$. Since $[|X^*|^r]<\infty$, $r\in\mathbf{N}$, and so assertion (6.12) follows from (6.3), assertion (6.13) follows from (6.10).

In the case of the following WLLN with "O"-rates the basic moment condition, in this case (6.2) for $r>2$, must, for the same reasons as in Theorem 6, be weakened (see (6.15)), whereas for $r=2$ (6.10) reduces to the non-trivial requirement (6.17).

Theorem 10. *Let $(X_k)_{k\in\mathbf{N}}$, $(\mathfrak{G}_k)_{k\in\mathbf{N}}$, Z_k with $P(Z_k=0)=1$, $k\in\mathbf{N}$ be defined as in Theorem 7.*

a) *If for some $r\in\mathbf{N}\setminus\{1\}$*

$$(6.14)\qquad E[|X_k|^r\,|\,\mathfrak{G}_k]\leqq M_{k,r}\quad a.s.\quad (k\in\mathbf{N})$$

and if there exist constants c_j such that

$$(6.15)\quad \varphi(n)^j\sum_{k=1}^n|E[X_k^j\,|\,\mathfrak{G}_k]|\leqq c_j\varphi(n)^r\sum_{k=1}^nM_{k,r}\quad a.s.\quad (0\leqq j\leqq r-1;\ n\in\mathbf{N}),$$

then for $f\in\mathrm{Lip}\,(r;r;C_B)$,

$$(6.16)\qquad \|V_{\varphi(n)S_n|\mathfrak{G}_1}f-f\|_{C_B}\leqq\left(c_f+\frac{L_f}{(r-1)!}\right)\varphi(n)^r\sum_{k=1}^nM_{k,r}$$

with $c_f:=\Sigma_{j=0}^{r-1}c_j\|f^{(j)}\|_{C_B}/j!$.

b) *If* $r=2$, *one has for* $f\in C_B^2$, *provided* $E[X_K|\mathfrak{G}_k]=0$ *a.s. and* $E[X_k^2|\mathfrak{G}_k]<M_{k,r}$ *a.s.*, $k\in\mathbf{N}$,

$$\|V_{n^{-1}S_n}f-f\|_{C_B}\leqq 2c_{2,2}L_f n^{-2}\sum_{k=1}^{n} M_{k,2}. \tag{6.17}$$

Proof. Condition (2.4) is satisfied for independent r.v.'s Z_k with $P_{Z_k}=P_{X_0}$ for all $k\in\mathbf{N}$. Since $E[|X_0|^j]=0$ for any $j\geqq 1$, a Taylor expansion up to the order $r-1$ yields, similarly as in the proof of Theorem 7,

$$|V_{\varphi(n)X_k|\mathfrak{G}_k}f(y,\omega)-V_{\varphi(n)Z_k}f(y)|\leqq \tag{6.18}$$

$$\leqq\sum_{j=1}^{r-1}\frac{\varphi(n)^j}{(r-1)!}\|f^{(j)}\|_{C_B}E[|X_k|^j|\mathfrak{G}_k]+L_f\frac{\varphi(n)^r}{(r-1)!}M_{k,r}.$$

Assertion (6.16) now follows by using condition (6.15) in formula (6.18). Part b) is the particular case of Theorem 8b) for $r=\alpha=2$ and $\varphi(n)=n^{-1}$, noting that $P_Z=P_{X_0}$.

7. O-approximation theorems for convergence in distribution. Just as in the case of martingales (cf. [5], [6]) it is possible to transfer our results concerned with rates for the weak convergence of the distributions $P_{\varphi(n)S_n}$ to P_Z to those for strong convergence. This is possible by applying a result contained implicitly in ZOLOTAREV [18], formulated explicitly in e.g. [5]. Using this result one can deduce from Theorem 7 the following theorem, noting that conditions (7.1) and (7.2) yield, for $f\in C_B^r$,

$$\|V_{\varphi(n)S_n}f-V_Zf\|_{C_B}=O(n\varphi(n)^r)\quad(n\to\infty).$$

Theorem 11. *Let* $(X_k)_{k\in\mathbf{N}}$, $(\mathfrak{G}_k)_{k\in\mathbf{P}}$, Z_k, $k\in\mathbf{N}$, *be defined as in Theorem* 7, *let the limiting r.v.* $Z\in\mathfrak{L}(\Omega,\mathfrak{A},P)$, *with distribution function* F_Z, *satisfy condition*

$$|F_Z(x_1)-F_Z(x_2)|\leqq M_Z|x_1-x_2|\quad(x_1,x_2\in\mathbf{R})$$

for some constant $M_Z>0$, *and assume that for* $r\in\mathbf{N}\setminus\{1\}$

$$E[|X_k|^r|\mathfrak{G}_k]<M_r\quad a.s.,\quad E[|Z_k|^r]<M_r^*\quad(k\in\mathbf{N}), \tag{7.1, 2}$$

M_r, M_r^* *being positive constants, independent of* k. *If further* (6.2) *holds, then*

$$\sup_{x\in\mathbf{R}}|F_{\varphi(n)S_n}(x)-F_Z(x)|=O\big(\varphi(n)^{r/(r+1)}n^{1/(r+1)}\big)\quad(n\to\infty). \tag{7.3}$$

If one applies Theorem 11 to the r.v. $Z:=X^*$, one obtains the following Berry—Esséen type estimates for dependent r.v.'s.

Theorem 12. *Let the assumptions of Theorem* 9 *be satisfied. If there exist two positive constants m, M such that* $m<\sigma_k^2<M$, *one obtains*

$$\sup_{x\in\mathbf{R}}|F_{s_n^{-1}S_n}(x)-F_{X^*}(x)|=O(s_n^{r/(r+1)}n^{1/(r+1)})\quad(n\to\infty). \tag{7.4}$$

If the r.v.'s X_k and Z_k, $k\in\mathbf{N}$ are identically distributed, and $\sigma_k^2=1$, all $k\in\mathbf{N}$, then for $r=3$

$$\sup_{x\in\mathbf{R}} |F_{n^{-1/2}s_n}(x)-F_{X^*}(x)| = O(n^{-1/8}) \quad (n\to\infty). \tag{7.5}$$

Setting $\varphi(n)=s_n^{-1}$ one can show, just as in the proof of Theorem 9b, that (7.4) follows from (7.3). Estimate (7.5) is a result of (7.4) since $s_n^{-1}=n^{-1/2}$ for $\sigma_k^2=1$.

8. Applications to martingale difference arrays. Whereas the dependency structure of the r.v.'s in question has so far been very general, it will be concretized in this section. The particular type of dependency to be considered will be that defined by a martingale difference array (MDA). A MDA is a double indexed array $(X_{nk})_{1\le k\le k_n}$, $n\in\mathbf{N}$ of r.v.'s from $\mathfrak{L}(\Omega, \mathfrak{A}, P)$ that is connected with a scheme $(\mathfrak{F}_{nk})_{0\le k\le k_n}$, $n\in\mathbf{N}$ of sub-σ-algebras of $\mathfrak{A}$ in such a form that the following three conditions are satisfied:

i) the sequence $(\mathfrak{F}_{nk})_{0\le k\le k_n}$ is monotone non-decreasing in k for each $n\in\mathbf{N}$,
ii) X_{nk} is measurable with respect to $\mathfrak{F}_{nk}$ for $1\le k\le k_n$,
iii) $E[X_{nk}|\mathfrak{F}_{n,k-1}]=0$ a.s. for $1\le k\le k_n$, $n\in\mathbf{N}$.

The general convergence theorem of this paper, Theorem 1, may be applied to MDA, as well as that supplied with o-rates, namely Theorem 4. But in order to avoid repetitions in the formulations we shall just consider the applications of Theorem 7 and 12 to yield

Theorem 13. *Let $(X_{nk})_{1\le k\le k_n}$, $n\in\mathbf{N}$, be a MDA, $(\mathfrak{F}_{nk})_{0\le k\le k_n}$, $n\in\mathbf{N}$, the associated array of sub-σ-algebras of $\mathfrak{A}$ (non-decreasing in k per definition) with $\mathfrak{F}_{n0}=\{\emptyset, \Omega\}$ for all $n\in\mathbf{N}$, and let Z be a φ-decomposable r.v. with decomposition components Z_{nk}, $1\le k\le k_n$.*

Assume further that for an $r\in\mathbf{N}\setminus\{1\}$

$$E[|X_{nk}|^r\,|\,\mathfrak{F}_{n,k-1}] < M_{nk,r} \quad a.s. \quad (1\le k\le k_n;\ n\in\mathbf{N}) \tag{8.1}$$

for some constant $M_{nk,r}>0$, as well as

$$E[|Z_{nk}|^r] < \infty \quad (1\le k\le k_n;\ n\in\mathbf{N}) \tag{8.2}$$

together with

$$E[X_{nk}^j\,|\,\mathfrak{F}_{n,k-1}] = E[Z_{nk}^j] \quad a.s. \quad (1\le k\le k_n;\ n\in\mathbf{N};\ 1\le j\le r-1;\ j\in\mathbf{N}). \tag{8.3}$$

Then for any $f\in \mathrm{Lip}(\alpha; r; C_B)$, $0<\alpha\le r$, there holds for $T_{nk_n}:=\varphi(k_n)\Sigma_{k=1}^{k_n}X_{nk}$ the estimate

$$\|V_{T_{nk_n}}f-V_Zf\| \le 2c_{2,r}L_f\varphi(n_k)^\alpha \sum_{k=1}^{k_n}(M_{nk,r}+E[|Z_{nk}|^r]) \tag{8.4}$$

with $c_{2,r}$ and L_f from Theorem 8.

If for each $n \in \mathbf{N}$ *the r.v.'s* X_{nk} *and* Z_{nk} *are in particular identically distributed for all* $1 \leqq k \leqq k_n$, *and if there holds*

$$E[|X_{nk}|^r | \mathfrak{F}_{n,k-1}] < M_{n,r} \quad a.s. \quad (1 \leqq k \leqq k_n,\ n \in \mathbf{N})$$

where $M_{n,r}$ *is a positive constant independent of* k, *as well as* (8.2) *and* (8.3), *then for* $f \in C_B^r$

$$\|V_{T_{nk_n}} f - V_Z f\| \leqq L_f \frac{\varphi(k_n)^r}{(r-1)!} k_n (M_{n,r} + E[|Z_{nk}|^r]). \tag{8.5}$$

Proof. Assertion (8.4) is a direct application of Theorem 8b), replacing the X_k by X_{nk} and the $\mathfrak{G}_k$ by $\mathfrak{F}_{n,k-1}$, noting that $\mathfrak{G}_1 = \mathfrak{F}_{n,0} = \{\emptyset, \Omega\}$, and that the distribution P_Z of the limit r.v. Z can, for each natural k_n, be representated as $P_Z = P_{\varphi(k_n)\Sigma_{k=1}^{k_n} Z_{nk}}$, whereby the independent decomposition components Z_k of (2.4) have here been written in the preciser form Z_{nk}. Inequality (8.5) follows by (6.8) in the proof of Theorem 7.

Now to the application of Theorem 12 to MDA; it is the CLT with rates for MDA.

Theorem 14. *Let* $(X_{nk})_{1 \leqq k \leqq k_n}$, $n \in \mathbf{N}$ *and* $(\mathfrak{F}_{nk})_{0 \leqq k \leqq k_n}$, $n \in \mathbf{N}$ *be defined as in Theorem* 13. *Let* $m_n' < \sigma_{nk}^2 := \mathrm{Var}\,(X_{nk}) < M_n'$, $1 \leqq k \leqq k_n$, $n \in \mathbf{N}$. *Assume further that for* $r \in \mathbf{N} \setminus \{1\}$

$$E[|X_{nk}|^r | \mathfrak{F}_{n,k-1}] < M_{n,r} \quad a.s. \quad (1 \leqq k \leqq k_n,\ n \in \mathbf{N}),$$

$M_{n,r}$ *being positive constants, independent of* k, *as well as*

$$E[X_{nk}^j | \mathfrak{F}_{n,k-1}] = \sigma_{nk}^j E[X^{*j}] \quad a.s. \quad (1 \leqq k \leqq k_n,\ n \in \mathbf{N};\ 1 \leqq j \leqq r-1;\ j \in \mathbf{N}).$$

Then one has for $s_{n,k_n} := (\Sigma_{k=1}^{k_n} \sigma_{nk}^2)^{1/2}$,

$$\sup_{x \in \mathbf{R}} |F_{s_{n,k_n}^{-1} \sum_{k=1}^{k_n} X_{nk}}(x) - F_{X^*}(x)| = O(s_{n,k_n}^{r/(r+1)} k_n^{1/(r+1)}) \quad (n \to \infty). \tag{8.6}$$

If, in addition, for each $n \in \mathbf{N}$ *the r.v.'s* X_{nk} *are identically distributed for all* $1 \leqq k \leqq k_n$, *and* $\sigma_{nk}^2 = 1$ *for* $1 \leqq k \leqq k_n$, $n \in \mathbf{N}$, *then for* $r = 3$

$$\sup_{x \in \mathbf{R}} |F_{k_n^{-1/2} \sum_{k=1}^{k_n} X_{nk}}(x) - F_{X^*}(x)| = O(k_n^{-1/8}) \quad (n \to \infty). \tag{8.7}$$

The proof of this theorem consists in a consequent application of Theorem 12, using the special case of MDA with $\mathfrak{G}_k = \mathfrak{F}_{n,k-1}$.

If one would take $k_n = n$ in Theorem 14, then the rate in (8.7) reduces to $O(n^{-1/8})$, one which was also attained by HEYDE and BROWN [14], CHOW and TEICHER [11 p. 314] as well as by ERICKSON, QUINE and WEBER [12]. Improvements of this rate were achieved by HALL and HEYDE [13 p. 84], namely with $O(n^{-1/4} \log n)$, by MUKERJEE [17] with $O(n^{-1/4})$, KATO [15] with $O(n^{-1/2}(\log n)^3)$ as well as by BOLT-

HAUSEN [3] with $O(n^{-1/2}\log n)$, whereby the better rates of convergence by Kato and Bolthausen are restricted to uniformly bounded r.v.'s. If one just assumes the boundedness of the third absolute moments of the r.v.'s X_{nk} as well as the "near constancy" of the partial sums of the conditional variances, here expressed in the form

$$(8.8) \qquad s_{n,k_n}^{-1} \sum_{k=1}^{k_n} E[X_{nk}^2 | \mathfrak{F}_{n,k-1}] \to 1 \quad \text{in probability} \quad (n \to \infty),$$

then the rate $O(n^{-1/4})$ is the best that has been obtained so far. It should be noted that condition (8.6) for $j=2$ implies (8.8); however, an assertion comparable to Theorem 14 could also be deduced by means of the conditional Trotter operator under the weaker assumption (8.8).

It must further be mentioned that the rates of Theorem 13, deduced from Theorem 7, dealing with rates for dependent r.v.'s, are just as good as those obtained in [4], [1], [2], [8], [9] and [5] for independent r.v.'s, MDS or MDA by means of the strongly modified Dvoretzky-method of proof mentioned in the introduction of [10]. But the conditional Trotter operator introduced in Section 3 allows one to prove the fundamental limit theorems equipped with rates in a *unified* way not only for various types of dependent r.v.'s but also for independent r.v.'s. It should be added that the definition and proofs involving the conditional Trotter operator and its properties also make use of set functions. So this operator theoretic approach stands in contrast to the more intricate "measure-theoretic" approach dealt with in most papers concerned with stochastic processes, in particular Markov processes.

It may be observed that the conditional Lindeberg—Trotter operator approach even makes it possible to deal with general limit theorems for Markov processes equiped with rates, see [7]. Similar results would be possible for inverse martingales or other dependency structure types.

The research of the third named author was supported by DFG grant Bu 166/37—4. The authors would like to thank a DFG-referee for pointing out an error in Part I of this paper, corrected in this part. They are also indebted to Dr. Dietmar Pfeifer, Heisenberg Professor, Aachen, for suggesting the use of "Polish spaces" to overcome the difficulty as well as for his critical reading of the manuscript.

References

[1] A. K. BASU, On the rate of approximation in the central limit theorem for dependent random variables and random vectors, *J. Multivariate Anal.*, **10** (1980), 565—578.

[2] A. K. BASU, On rates of convergence to infinitely divisible laws for dependent random variables, *Canad. J. Statist.*, **8** (1980), 235—247.

[3] E. Bolthausen, Exact convergence rates in some martingale central limit theorems, *Ann. Probab.*, **10** (1982), 672—688.

[4] P. L. Butzer and L. Hahn, General theorems on the rate of convergence in distribution of random variables I. General limit theorems; II. Applications to the stable limit laws and weak law of large numbers, *J. Multivariate Anal.*, **8** (1978), 181—201; 202—221.

[5] P. L. Butzer and H. Kirschfink, Donsker's weak invariance principle with rates for $C[0, 1]$-valued, dependent random-functions, *Approximation Theory and its Applications*, **2** (1986), 55—77.

[6] P. L. Butzer and H. Kirschfink, General limit theorems with o-rates and Markov-processes under pseudo-moment conditions, *Z. Anal. Anwendungen* (1988) (in print).

[7] P. L. Butzer and H. Kirschfink, The conditional Lindeberg—Trotter operator in the resolution of limit theorems with rates for dependent random variables; Applications to Markovian processes, *Z. Anal. Anwendungen* (1988) (in print).

[8] P. L. Butzer and D. Schulz, General random sum limit theorems for martingales with O-rates, *Z. Anal. Anwendungen*, **2** (2) (1983), 97—109.

[9] P. L. Butzer and D. Schulz, Approximation theorems for martingale difference arrays with applications to randomly stopped sums, *Mathematical Structures — Computational Mathematics — Mathematical Modelling*, **2**, pp. 121—130. (Papers dedicated to Academian L. Iliev's 70th Anniversary; Ed. Bl. Sendov, Sofia, 1984.)

[10] P. L. Butzer and D. Schulz, An extension of the Lindeberg—Trotter-operator theoretic approach to limit theorems for dependent random variables. I. General convergence theorems, approximation theorems with o-rates, *Acta Sci. Math.*, **48** (1985), 37—54.

[11] Y. S. Chow and H. Teicher, *Probability Theory*, Springer-Verlag (New York, 1978).

[12] R. V. Erickson, M. P. Quine and N. C. Weber, Explicit bounds for the departure from normality of sums of dependent random variables, *Acta Math. Acad. Sci. Hungar.*, **34** (1979), 27—32.

[13] P. Hall and C. C. Heyde, *Martingale Limit Theory and its Application*, Academic Press (New York, 1980).

[14] C. C. Heyde and B. M. Brown, On the departure from normality of a certain class of martingales, *Ann. Math. Statist.*, **41** (1970), 2161—2165.

[15] Y. Kato, Convergence rates in central limit theorems for martingale differences, *Bull. Math. Statist.*, **18** (1979), 1—9.

[16] R. G. Laha and V. K. Rohatgi, *Probability Theory*, John Wiley & Sons (New York, 1979).

[17] H. G. Mukerjee, On an improved rate of convergence to normality for sums of dependent random variables with applications to stochastic approximation, *Acta Math. Acad. Sci. Hungar.*, **40** (1982), 229—236.

[18] V. M. Zolotarev, Approximation of the distributions of sums of independent random variables with values in infinite dimensional spaces, *Teor. Verojatnost. i. Primenen*, **21** (1976), 741—758. (Russian; English translation: *Theor. Probab. Appl.*, **21** (1976), 721—737.)

LEHRSTUHL A FÜR MATHEMATIK
AACHEN UNIVERSITY OF TECHNOLOGY
5100 AACHEN, FEDERAL REPUBLIC OF GERMANY

Acta Sci. Math., 51 (1987), 435—440

On asymptotic Toeplitz operators

JOSÉ BARRÍA

A symbol map is constructed for the C^*-algebra generated by the Toeplitz and compact operators on the space $H^2(\Sigma)$ associated with a semigroup Σ of a locally compact abelian group. As a consequence it follows that the essential range of the symbol is contained in the essential spectrum of the corresponding Toeplitz operator.

Let G be a locally compact abelian group with dual group $\hat{G}$, and let Σ denote a fixed sub-semigroup of $\hat{G}$ which is a Borel subset of $\hat{G}$. Let μ and $\hat{\mu}$ be the normalized Haar measures on G and $\hat{G}$, respectively. Let $L^2(G)$ and $L^2(\hat{G})$ be the corresponding Hilbert spaces of square-integrable functions. The Fourier transform $\mathcal{F}$ is an isometry from $L^2(G)$ onto $L^2(\hat{G})$. We denote by $H^2(\Sigma)$ the subspace of $L^2(G)$ consisting of the functions f for which $\mathcal{F}f$ is in $L^2(\Sigma)$, that is

$$H^2(\Sigma) = \{f\in L^2(G)\colon \mathcal{F}f \text{ is supported on } \Sigma\}.$$

Let P denote the orthogonal projection of $L^2(G)$ onto $H^2(\Sigma)$. If φ is a bounded measurable function on G, write M_φ for the multiplication operator defined on $L^2(G)$ by

$$M_\varphi f = \varphi f$$

and T_φ for the compression of M_φ defined on $H^2(\Sigma)$ by

$$T_\varphi f = PM_\varphi f = P(\varphi f).$$

The operator T_φ is called a *Toeplitz operator* with symbol φ.

The semigroup Σ induces a partial order $\leqq$ on $\hat{G}$ in which $\alpha\leqq\beta$ if $\beta\alpha^{-1}$ is in Σ. With this partial order, $\hat{G}$ and Σ are directed sets. Furthermore, if $\alpha\in\Sigma$ and $\alpha\leqq\beta$, then $\beta\in\Sigma$.

An operator T on $H^2(\Sigma)$ is called an *asymptotic Toeplitz operator* if the net $\{T_\alpha^* TT_\alpha\colon \alpha\in\Sigma\}$ converges strongly. The class of all asymptotic Toeplitz operators on $H^2(\Sigma)$ will be denoted by (AT).

Received September 13, 1984.

In case G is the unit circle group $\mathbf{T}$, $\hat{G}$ is the integers $\mathbf{Z}$, and Σ is the semigroup $\mathbf{Z}^+$ of non-negative integers, then T is an asymptotic Toeplitz operator if the sequence $\{U^{*n}TU^n\}_{n=1}^{\infty}$ converges strongly (here U is the unilateral shift on H^2). In [1] this asymptotic notion was defined and used to assign a symbol to any operator in the C^*-algebra generated by the Toeplitz and Hankel operators. In this paper the construction of the symbol map is carried out in the more general setting of locally compact abelian groups. As a consequence, it follows that the spectrum of M_φ is contained in the essential spectrum of T_φ.

Toeplitz operators on locally compact abelian groups were first studied by L. A. Coburn and R. G. Douglas [2]. One of their concerns was the C^*-algebra generated by the Toeplitz operators with symbol in the algebra of almost periodic functions. They proved that this algebra of operators modulo its commutator ideal is $*$-isomorphic to the algebra of almost periodic functions. Our results show that the C^*-algebra generated by all Toeplitz and compact operators is $*$-homomorphic to $L^\infty(G)$, and the kernel of this homomorphism is the ideal of operators T (in the algebra) such that $T_\alpha^* T T_\alpha \to 0$ $(\alpha \in \Sigma)$ strongly.

The following elementary facts about Toeplitz operators will be useful:

$$T_{\varphi+\psi} = T_\varphi + T_\psi, \quad T_\varphi^* = T_{\bar{\varphi}},$$
$$T_\alpha f = \alpha f \quad \text{if } \alpha \text{ is in } \Sigma.$$

For the rest of the paper we make the following assumptions: μ and $\hat{\mu}$ are σ-finite measures, $\hat{\mu}(\Sigma) > 0$, $\hat{G}$ is generated by Σ (i.e. $\hat{G} = \Sigma\Sigma^{-1}$), Σ is not dense in $\hat{G}$.

From [2] we have that $\sigma(M_\varphi)$, the spectrum of M_φ, is contained in $\sigma(T_\varphi)$, and $\|T_\varphi\| = \|\varphi\|_\infty$ for φ in $L^\infty(G)$.

Lemma 1. *If K is a compact operator on $H^2(\Sigma)$, then $K \in (\mathrm{AT})$ and $KT_\alpha \to 0$ $(\alpha \in \Sigma)$ strongly.*

Proof. It is enough to prove the last assertion. For this, let f be fixed in $H^2(\Sigma)$. Since K is a compact operator, and $KT_\alpha f = K(\alpha f)$, we only need to show that the net $\{\alpha f\colon \alpha \in \Sigma\}$ converges weakly to zero. For g in $H^2(\Sigma)$ we have $\bar{f}g \in L^1(G)$, and therefore

$$(g, \alpha f) = \int_G g\bar{\alpha}\bar{f}\, d\mu = \mathscr{F}(\bar{f}g)(\alpha)$$

for all α in Σ. By the Riemann—Lebesgue's Lemma ([4], Remark 28.42), given $\varepsilon > 0$ there exists a compact set F in $\hat{G}$ such that

$$|\mathscr{F}(\bar{f}g)(\sigma)| < \varepsilon \quad \text{for all } \sigma \text{ in } \hat{G} \setminus F.$$

Next we show that $\alpha\Sigma \subseteqq \hat{G} \setminus F$ for some α in Σ. If this is not the case, then $\alpha\Sigma \cap F \neq \Box$ for all α in Σ. Since $\alpha_1\alpha_2\Sigma \subseteqq \alpha_i\Sigma$ for α_1, α_2 in Σ, it follows that the family $\{\alpha\Sigma \cap F\colon \alpha \in \Sigma\}$

has the finite intersection property. Hence there exists α_0 in the closure Σ^- of Σ such that $\alpha_0 \in \alpha\Sigma^-$ for all α in Σ, therefore $\alpha_0\Sigma^{-1} \subseteqq \Sigma^-$ and so $\alpha_0\hat{G} \subseteqq \Sigma^-$. This is a contradiction because $\alpha_0\hat{G}=\hat{G}$ and Σ is not dense in $\hat{G}$.

So far we have proved that there exists α in Σ such that $|\mathscr{F}(\bar{f}g)(\sigma)|<\varepsilon$ for all σ in $\alpha\Sigma$. This shows that $|(g, \beta f)|<\varepsilon$ for all β in Σ such that $\beta \geqq \alpha$.

Lemma 2. [2] *If E is a compact subset of $\hat{G}$, then there exists α in Σ such that $\alpha E \subseteqq \Sigma$.*

Proof. See [2], § 2.

Lemma 3. *Let A be an operator on $H^2(\Sigma)$ such that $T_\alpha^* A T_\alpha = A$ for all α in Σ. Then $M_\alpha^* APM_\alpha \to M_\varphi$ $(\alpha\in\Sigma)$ weakly, and $A=T_\varphi$ for some φ in $L^\infty(G)$. Furthermore, $M_\alpha^* PM \to I$ $(\alpha\in\Sigma)$ strongly.*

Proof. Let f be in $L^2(G)$ such that $\mathscr{F}f$ has compact support E in $\hat{G}$. From Lemma 2 there exists α_0 in Σ such that $\alpha E\subseteqq\Sigma$ for $\alpha\geqq\alpha_0$. If $\sigma\in\hat{G}\setminus\Sigma$ and $\alpha\geqq\alpha_0$, then $\alpha^{-1}\sigma\in\hat{G}\setminus E$, and so $\mathscr{F}(\alpha f)(\sigma)=(\mathscr{F}f)(\alpha^{-1}\sigma)=0$. Hence $\alpha f\in H^2(\Sigma)$ and $M_\alpha^* APM_\alpha f=M_\alpha^* A(\alpha f)$. If A is the identity, we conclude that $M_\alpha^* PM_\alpha f=f$ for $\alpha\geqq\alpha_0$. This completes the proof of the last assertion of the lemma.

Let $B_\alpha=M_\alpha^* APM_\alpha$ for α in Σ. Let g be in $L^2(G)$ such that $\mathscr{F}g$ has support contained in E. If $\alpha\geqq\alpha_0$, from above we have

$$(B_\alpha f, g) = \big(A(\alpha f), \alpha g\big) = \big(AT_{\alpha\alpha_0^{-1}}(\alpha_0 f), T_{\alpha\alpha_0^{-1}}(\alpha_0 g)\big) =$$
$$= \big(T^*_{\alpha\alpha_0^{-1}} AT_{\alpha\alpha_0^{-1}}(\alpha_0 f), \alpha_0 g\big) = \big(A(\alpha_0 f), \alpha_0 g\big)$$

because $\alpha\alpha_0^{-1}\in\Sigma$. Hence there exists an operator B on $L^2(G)$ such that $B_\alpha\to B$ $(\alpha\in\Sigma)$ weakly.

For σ in Σ we have

$$(M_\sigma^* Bf, g) = \lim_\alpha (M_\sigma^* M_\alpha^* APM_\alpha f, g) = \lim_\alpha (M_{\alpha\sigma}^* APM_{\alpha\sigma} M_\sigma^* f, g) = (BM_\sigma^* f, g).$$

Therefore $M_\sigma B=BM_\sigma$ for all σ in $\hat{G}$. Since the subspace spanned by $\hat{G}$ is weak* dense in $L^\infty(G)$ ([4], Lemma 31.4), it follows that $M_\psi B=BM_\psi$ for all ψ in $L^\infty(G)$. Since the algebra of multiplication operators is maximal abelian, then $B=M_\varphi$ for some φ in $L^\infty(G)$. Finally, it is easy to see that $(T_\varphi f, g)=(Af, g)$ for f, g in $H^2(\Sigma)$.

Corollary 4. *If T is an asymptotic Toeplitz operator on $H^2(\Sigma)$, then $T_\alpha^* TT_\alpha\to T_\varphi$ $(\alpha\in\Sigma)$ strongly, for some φ in $L^\infty(G)$.*

Proof. If $T_\alpha^* TT_\alpha\to A$ $(\alpha\in\Sigma)$ strongly, then $T_\alpha^* AT_\alpha=A$ for all α in Σ. Now the result follows from Lemma 3.

We define the map Φ: (AT)$\to L^\infty(G)$ by $\Phi(T)=\varphi$ where $T_\alpha^* T T_\alpha \to T_\varphi$ $(\alpha\in\Sigma)$ strongly. The function φ is called the symbol of T. Lemma 1 shows that $\Phi(K)=0$ if K is a compact operator.

Corollary 5. *The class of asymptotic Toeplitz operators is a norm closed subspace. The map Φ is a linear contraction.*

Proof. Let $T\in$(AT) and $\Phi(T)=\varphi\in L^\infty(G)$. Then

$$\|T_\varphi f\| = \lim_\alpha \|T_\alpha^* T T_\alpha f\| \leqq \|T\|\,\|f\|.$$

Hence

$$\|\Phi(T)\| = \|\varphi\|_\infty = \|T_\varphi\| \leqq \|T\|.$$

Let $T_n\in$(AT) be such that $\|T-T_n\|\to 0$. Let $\varphi_n\in L^\infty(G)$ be such that $\Phi(T_n)=\varphi_n$. Since

$$\|\varphi_n-\varphi_m\|_\infty = \|\Phi(T_n-T_m)\| \leqq \|T_n-T_m\|,$$

there exists φ in $L^\infty(G)$ such that $\|\varphi_n-\varphi\|_\infty\to 0$. Now we have

$$\|T_\alpha^* T T_\alpha f - T_\varphi f\| \leqq \|T-T_n\|\,\|f\| + \|T_\alpha^* T_n T_\alpha f - T_{\varphi_n} f\| + \|\varphi_n-\varphi\|_\infty \|f\|,$$

therefore $T\in$(AT) and $\Phi(T)=\varphi$.

Corollary 6. *If K is a compact operator on $H^2(\Sigma)$, then $\|T_\varphi\|\leqq\|T_\varphi+K\|$ for φ in $L^\infty(G)$. Therefore the subspace $\{T_\varphi+K$: $\varphi\in L^\infty(G)$, K compact$\}$ is norm closed.*

Proof. By Lemma 1, $\Phi(T_\varphi+K)=\varphi$. Since Φ is a contraction,

$$\|T_\varphi\| = \|\varphi\|_\infty = \|\Phi(T_\varphi+K)\| \leqq \|T_\varphi+K\|.$$

Corollary 7. *If φ is in $L^\infty(G)$ and $H_\varphi=P^\perp M_\varphi|H^2(\Sigma)$, then $H_\varphi T_\alpha\to 0$ $(\alpha\in\Sigma)$ strongly.*

Proof. For f in $H^2(\Sigma)$ and α in Σ we have

$$H_\varphi T_\alpha f = P^\perp(\alpha\varphi f) = P^\perp M_\alpha(\varphi f).$$

Therefore $\|H_\varphi T_\alpha f\|=\|M_\alpha^* P^\perp M_\alpha(\varphi f)\|\to 0$ $(\alpha\in\Sigma)$ by Lemma 3.

Lemma 8. *Let $T=T_{\varphi_1}T_{\varphi_2}\dots T_{\varphi_n}$ with $\varphi_i\in L^\infty(G)$. Then T is an asymptotic Toeplitz operator and $\Phi(T)=\varphi_1\varphi_2\dots\varphi_n$.*

Proof. For φ in $L^\infty(G)$ and H_φ as defined in Corollary 7 we have

$$M_\varphi = \begin{pmatrix} T_\varphi & * \\ H_\varphi & * \end{pmatrix}.$$

with respect to the decomposition $L^2(G)=H^2(\Sigma)\oplus H^2(\Sigma)^\perp$. If φ and ψ are in $L^\infty(G)$, then $M_{\psi\varphi}=M_\psi M_\varphi$ and therefore (multiply matrices and compare upper left corners) $T_{\psi\varphi}-T_\psi T_\varphi=AH_\varphi$ for some operator A. Applying this last equality to the telescoping sum

$$T-T_{\varphi_1\varphi_2\dots\varphi_n}=T_{\varphi_1}T_{\varphi_2\dots\varphi_n}-T_{\varphi_1(\varphi_2\dots\varphi_n)}+T_{\varphi_1}(T_{\varphi_2}T_{\varphi_3\dots\varphi_n}-T_{\varphi_2(\varphi_3\dots\varphi_n)})+$$
$$+T_{\varphi_1}T_{\varphi_2}(T_{\varphi_3}T_{\varphi_4\dots\varphi_n}-T_{\varphi_3(\varphi_4\dots\varphi_n)})+\dots+T_{\varphi_1}T_{\varphi_2}\dots T_{\varphi_{n-2}}(T_{\varphi_{n-1}}T_{\varphi_n}-T_{\varphi_{n-1}\varphi_n})$$

we conclude that each of the $n-1$ summands on the right can be written as BH_φ for some operator B and some φ in $L^\infty(G)$. From Corollary 7 we have that $BH_\varphi T_\alpha\to 0$ $(\alpha\in\Sigma)$ strongly. Therefore $(T-T_{\varphi_1\varphi_2\dots\varphi_n})T_\alpha\to 0$ $(\alpha\in\Sigma)$ strongly. From this it follows that $T\in(\mathrm{AT})$ and $\Phi(T)=\varphi_1\varphi_2\dots\varphi_n$.

Theorem 9. *Let $\mathscr{A}$ be the C^*-algebra generated by the Toeplitz and compact operators on $H^2(\Sigma)$. Then $\mathscr{A}$ is contained in the class of asymptotic Toeplitz operators. Furthermore, the restriction of Φ to $\mathscr{A}$ is a $*$-homomorphism.*

Proof. Let $\mathscr{A}_0$ be the linear manifold generated by the compact operators and all the finite products of Toeplitz operators. Clearly $\mathscr{A}_0$ is an algebra which is closed under the operation of taking adjoint, and the norm closure of $\mathscr{A}_0$ is equal to $\mathscr{A}$. Since (AT) is a subspace, from Lemmas 1 and 8 it follows that $\mathscr{A}_0$ is contained in (AT), and the restriction of Φ to $\mathscr{A}_0$ is clearly a $*$-homomorphism. Since (AT) is norm closed, then $\mathscr{A}\subseteq(\mathrm{AT})$, and the proof is complete.

Remark. In general, (AT) is not an algebra, it is not even closed under adjointion (cf. [1]).

Corollary 10. *If φ is in $L^\infty(G)$, then the spectrum of M_φ is contained in the essential spectrum of T_φ.*

Proof. Since the spectrum of M_φ is the essential range of φ, it will be enough to show that if T_φ is a Fredholm operator, then φ has an inverse in $L^\infty(G)$. Let $\mathscr{A}$ be the C^*-algebra defined in Theorem 9. If $\mathbf{K}$ is the closed ideal of compact operators on $H^2(\Sigma)$, then $\mathscr{A}/\mathbf{K}$ is a C^*-algebra. If T_φ is Fredholm, then $[T_\psi]$ is invertible in $\mathscr{A}/\mathbf{K}$, so there exists S in $\mathscr{A}$ such that $T_\varphi S-I$ is compact. Therefore $\Phi(T_\varphi S-I)=0$. Since Φ is a homomorphism on $\mathscr{A}$, $\varphi\cdot\Phi(S)\equiv 1$ a.e. $[\mu]$. Since $\Phi(S)$ is in $L^\infty(G)$, then φ is invertible in $L^\infty(G)$.

Remark. In Corollary 10 it is actually proved that the spectrum of M_φ is contained in the intersection of the left essential spectrum and the right essential spectrum of T_φ.

Remark. From Theorem 9 we have that $TS-ST$ is in ker Φ for any S and T in $\mathscr{A}$. Therefore the commutator ideal of $\mathscr{A}$ is contained in ker Φ. For Toeplitz operators on the unit circle this inclusion is an equality [1]. Is this true in general?

References

[1] J. Barría, P. R. Halmos, Asymptotic Toeplitz operators, *Trans. Amer. Math. Soc.*, **273** (1982), 621—630.

[2] L. A. Coburn, R. G. Douglas, C^*-algebras of operators on a half-space. I, *Publ. Math. I.H.E.S.*, **40** (1971), 59—68.

[3] R. G. Douglas, R. Howe, On the C^*-algebra of Toeplitz operators on the quarter-plane, *Trans. Amer. Math. Soc.*, **158** (1971), 203—217.

[4] E. Hewitt, R. Ross, *Abstract harmonic analysis. II*, Springer-Verlag (New York, 1970).

INSTITUTO VENEZOLANO DE INVESTIGACIONES CIENTÍFICAS,
APARTADO 1827
CARACAS 1010—A, VENEZUELA

Current address:
DEPARTMENT OF MATHEMATICS
SANTA CLARA UNIVERSITY
SANTA CLARA, CALIFORNIA 95053, USA

Acta Sci. Math., 51 (1987), 441—456

Ideals and Lie ideals of operators

C. K. FONG and G. J. MURPHY

1. Introduction

Let $\mathfrak{H}$ denote an infinite dimensional (complex) Hilbert space and $\mathscr{B}(\mathfrak{H})$ the algebra of all (bounded, linear) operators on $\mathfrak{H}$. We say a linear manifold $\mathscr{L}$ in $\mathscr{B}(\mathfrak{H})$ is *unitarily invariant* if $U^*\mathscr{L}U\subseteq\mathscr{L}$ for all unitaries U in $\mathscr{B}(\mathfrak{H})$. If $\mathscr{L}$ is such a manifold and $\mathfrak{K}$ is another Hilbert space of the same dimension as $\mathfrak{H}$, then we can "transport" $\mathscr{L}$ to a unitarily invariant manifold of operators acting on $\mathfrak{K}$ by taking any unitary transformation W from $\mathfrak{H}$ onto $\mathfrak{K}$ and setting $\mathscr{L}_{\mathfrak{K}}=W\mathscr{L}W^*$. That $\mathscr{L}_{\mathfrak{K}}$ is unitarily invariant, and that its definition is independent of the choice of W, follow from the fact that $\mathscr{L}$ is unitarily invariant. In particular, if we consider the case when $\mathfrak{K}=\mathfrak{H}\oplus\mathfrak{H}$, then $\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{H}}$ is a unitarily invariant manifold of operators which can be expressed as 2×2 operator matrices with entries in $\mathscr{B}(\mathfrak{H})$. Thus we can define the following two manifolds in $\mathscr{B}(\mathfrak{H})$:

(*)
$$\mathscr{L}^c=\{T|_{\mathfrak{H}\oplus 0}\colon T\in\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{H}}\}=\left\{A\in\mathscr{B}(\mathfrak{H})\colon \begin{pmatrix}A & B\\ C & D\end{pmatrix}\in\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{H}} \text{ for some } B, C, D\in\mathscr{B}(\mathfrak{H})\right\},$$

(* *)
$$\mathscr{I}_{\mathscr{L}}=\left\{B\in\mathscr{B}(\mathfrak{H})\colon \begin{pmatrix}0 & B\\ 0 & 0\end{pmatrix}\in\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{H}}\right\}.$$

It was shown in [5] that $\mathscr{I}_{\mathscr{L}}$ is an ideal in $\mathscr{B}(\mathfrak{H})$, and that $[\mathscr{B}(\mathfrak{H}),\mathscr{I}_{\mathscr{L}}]\subseteq\mathscr{L}\subseteq$ $\subseteq\mathscr{I}_{\mathscr{L}}+\mathbf{C}I$. This fact covers a part of the following theorem, also proved in the same paper.

Theorem 1 ([5]). *Let $\mathscr{L}$ be a linear manifold in $\mathscr{B}(\mathfrak{H})$. Then the following conditions are equivalent:*

(1) *$\mathscr{L}$ is unitarily invariant;*

(2) *$\mathscr{L}$ is a Lie ideal in $\mathscr{B}(\mathfrak{H})$;*

Received January 29, 1984.

(3) *there exists an ideal* $\mathscr{I}$ *in* $\mathscr{B}(\mathfrak{H})$ *such that*

$$[\mathscr{B}(\mathfrak{H}), \mathscr{I}] \subseteq \mathscr{L} \subset \mathscr{I} + \mathbf{C}I.$$

(The above results are shown in [5] only in the case where $\mathfrak{H}$ is separable; but, in fact, everything works in the non-separable case too. See remarks following Theorem 2.)

For a unitarily invariant manifold $\mathscr{L}$ in $\mathscr{B}(\mathfrak{H})$, the ideal $\mathscr{I}$ of the condition (3) above is uniquely determined by $\mathscr{L}$ (shown in Section 2), and will be called the *associate ideal* of $\mathscr{L}$. Among other results in Section 2, we show that $\mathscr{L}^c$ (defined by ($*$)) is either $\mathscr{I}_{\mathscr{L}}$ or $\mathscr{I}_{\mathscr{L}} + \mathbf{C}I$. A consequence, shown in Section 3, is the following useful characterization of ideals in $\mathscr{B}(\mathfrak{H})$: a linear manifold $\mathscr{L}$ in $\mathscr{B}(\mathfrak{H})$ is a proper ideal if and only if $\mathscr{L}$ is unitarily invariant, $I \notin \mathscr{L}$ and $\mathscr{L}^c \subseteq \mathscr{L}$. Several applications of this result (or its variant) are given in Section 3.

In Section 4 we give some characterizations of ideals in C^*-algebras satisfying a certain condition, viz., we show that the ideals are precisely the linear manifolds $\mathscr{L}$ for which $P\mathscr{L}P \subseteq \mathscr{L}$ for all projections in the algebra.

The proof of Theorem 1 previously mentioned in [5] uses the following weaker form of a theorem of Fillmore:

Theorem 2 ([3]). *Every operator in* $\mathscr{B}(\mathfrak{H})$ *is a linear combination of projections.*

The original proof of this result is quite complicated. We include an appendix to this paper in which we prove Theorem 1 in such a way that, not only do we obtain it without Theorem 2, but the latter theorem actually drops out as a bonus in the process. To generalize Theorem 1 for Hilbert spaces not necessarily separable, we need to extend a theorem of Calkin [1] to the non-separable case. Since this extension is by no means straightforward, we also include its proof in the appendix.

We use standard notation: $\mathscr{C}_2$ denotes the Hilbert—Schmidt class and $\mathscr{C}_p$ ($p > 0$) the ideal of operators such that $(T^*T)^{p/4} \in \mathscr{C}_2$. If $\mathscr{S}$, $\mathscr{T}$ are linear manifolds in $\mathscr{B}(\mathfrak{H})$, we write $\mathscr{S}\mathscr{T}$ (resp. $[\mathscr{S}, \mathscr{T}]$) for the linear span of all operators of the form ST (resp. $[S, T] = ST - TS$) where $S \in \mathscr{S}$ and $T \in \mathscr{T}$. All Hilbert spaces are assumed to be infinite dimensional, but they are not required to be separable unless otherwise stated.

2. The associate ideal of a Lie ideal in $\mathscr{B}(\mathfrak{H})$

Let $\mathscr{L}$ be a unitarily invariant manifold in $\mathscr{B}(\mathfrak{H})$. As mentioned in the introduction, the set

$$\mathscr{I}_{\mathscr{L}} = \left\{ B \in \mathscr{B}(\mathfrak{H}) \colon \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix} \in \mathscr{L}_{\mathfrak{H} \oplus \mathfrak{H}} \right\}$$

forms an ideal. We call $\mathscr{I}_{\mathscr{L}}$ *the associate ideal* of $\mathscr{L}$ or *the ideal associated with* $\mathscr{L}$. There are several ways to describe this ideal, as the following proposition shows.

Proposition 3. *Let $\mathscr{L}$ be a unitarily invariant manifold and $\mathscr{I}$ be an ideal in $\mathscr{B}(\mathfrak{H})$. Then the following conditions are equivalent:*

(1) *$\mathscr{I}$ is the associate ideal of $\mathscr{L}$;*
(2) *$[\mathscr{B}(\mathfrak{H}), \mathscr{I}] \subseteqq \mathscr{L} \subseteqq \mathscr{I} + \mathbf{C}I$;*
(3) *$[\mathscr{B}(\mathfrak{H}), \mathscr{I}] \subseteqq \mathscr{L}$ and $[\mathscr{B}(\mathfrak{H}), \mathscr{L}] \subseteqq \mathscr{I}$;*
(4) *$\mathscr{I}$ is the largest ideal among those ideals $\mathscr{J}$ satisfying $[\mathscr{B}(\mathfrak{H}), \mathscr{J}] \subseteqq \mathscr{L}$;*
(5) *$\mathscr{I}$ is the smallest ideal among those ideals $\mathscr{J}$ satisfying $\mathscr{L} \subseteqq \mathscr{J} + \mathbf{C}I$;*
(6) *$\mathscr{I}$ is the ideal generated by $[\mathscr{B}(\mathfrak{H}), \mathscr{L}]$;*
(7) *$\mathscr{I} + \mathbf{C}I = \{T \in \mathscr{B}(\mathfrak{H}) : [\mathscr{B}(\mathfrak{H}), T] \subseteqq \mathscr{L}\}$.*

For the proof of the above proposition, we need the following lemma.

Lemma 4. *Let $\mathscr{A}$ be an algebra with identity I and $\mathscr{B} = \mathscr{M}_2(\mathscr{A})$ be the algebra of all 2×2 matrices with entries in $\mathscr{A}$. Then, for two ideals $\mathscr{I}_1$ and $\mathscr{I}_2$ in $\mathscr{B}$, $[\mathscr{B}, [\mathscr{B}, \mathscr{I}_1]] \subseteqq \mathscr{I}_2$ implies $\mathscr{I}_1 \subseteqq \mathscr{I}_2$.*

Proof. Let $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathscr{I}_1$. Then

$$\left[\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \left[\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} A & B \\ C & D \end{pmatrix}\right]\right] = \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix} \in \mathscr{I}_1 \cap \mathscr{I}_2.$$

Hence $\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \in \mathscr{I}_1$ and it suffices to show that $\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \in \mathscr{I}_2$. Now $\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} = \begin{pmatrix} 0 & A \\ D & 0 \end{pmatrix} \in \mathscr{I}_1$ and, by using the same computation as above, we obtain $\begin{pmatrix} 0 & A \\ D & 0 \end{pmatrix} \in \mathscr{I}_2$. Hence

$$\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} = \begin{pmatrix} 0 & A \\ D & 0 \end{pmatrix}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \in \mathscr{I}_2. \quad \square$$

Proof of Proposition 3. (1)⇒(2) follows from Theorem 1. (2)⇒(3) is obvious. To show (3)⇒(6), let $\mathscr{J}$ be the ideal generated by $[\mathscr{B}(\mathfrak{H}), \mathscr{L}]$. Then we have $[\mathscr{B}(\mathfrak{H}), \mathscr{L}] \subseteqq \mathscr{J} \subseteqq \mathscr{I}$ and hence $[\mathscr{B}(\mathfrak{H}), [\mathscr{B}(\mathfrak{H}), \mathscr{I}]] \subseteqq [\mathscr{B}(\mathfrak{H}), \mathscr{L}] \subseteqq \mathscr{J}$. It follows from Lemma 4 (since $\mathscr{B}(\mathfrak{H})$ and $\mathscr{M}_2(\mathscr{B}(\mathfrak{H}))$ are isomorphic algebras) that $\mathscr{I} \subseteqq \mathscr{J}$. Therefore $\mathscr{I} = \mathscr{J}$. Similarly we can show that (3)⇒(4) and (2)⇒(5). Since the ideal $\mathscr{J}$ described by either (4), (5) or (6) is unique and since the associate ideal fits into each of these descriptions, we have (4)⇒(1), (5)⇒(1) and (6)⇒(1). Thus conditions (1) to (6) are equivalent.

Finally, let $\mathscr{I}_0$ be the ideal associated with $\mathscr{L}$ and

$$\mathscr{S} = \{T \in \mathscr{B}(\mathfrak{H}) : [\mathscr{B}(\mathfrak{H}), T] \subseteqq \mathscr{L}\}.$$

Then it is easy to see that $\mathscr{S}$ is unitarily invariant and $\mathscr{I}_0 \subseteqq \mathscr{S}$. Let $\mathscr{J}$ be the associate ideal of $\mathscr{S}$. Since

$$[\mathscr{B}(\mathfrak{H}), [\mathscr{B}(\mathfrak{H}), \mathscr{J}]] \subseteqq [\mathscr{B}(\mathfrak{H}), \mathscr{S}] \subseteqq \mathscr{L} \subseteqq \mathscr{I}_0 + \mathbf{C}I$$

it follows from Lemma 4 that $\mathscr{J}\subseteqq\mathscr{I}_0$. Therefore we have $\mathscr{J}\subseteqq\mathscr{I}_0\subseteqq\mathscr{S}\subseteqq\mathscr{J}+\mathbf{C}I$. Hence $\mathscr{J}=\mathscr{I}_0$ and $\mathscr{S}=\mathscr{J}+\mathbf{C}I=\mathscr{I}_0+\mathbf{C}I$. We have proved (1)⇒(7). Conversely, if $\mathscr{S}=\mathscr{I}+\mathbf{C}I$, then $\mathscr{I}+\mathbf{C}I=\mathscr{I}_0+\mathbf{C}I$ and hence $\mathscr{I}=\mathscr{I}_0$. Therefore (7)⇒(1) follows. □

For brevity, in the rest of this section, we replace the term "unitarily invariant linear manifold in $\mathscr{B}(\mathfrak{H})$" by its synonym "Lie ideal in $\mathscr{B}(\mathfrak{H})$".

By definition, the associate ideal $\mathscr{I}$ of a Lie ideal $\mathscr{L}$ in $\mathscr{B}(\mathfrak{H})$ is obtained by taking the upper right corners of 2×2 matrices in $\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{H}}$. The next result says, if we take the upper left corners instead, then either $\mathscr{I}$ or $\mathscr{I}+\mathbf{C}I$ is produced.

Proposition 5. *If $\mathscr{L}$ is a Lie ideal and $\mathscr{I}$ is its associate ideal, then either $\mathscr{L}^c=\mathscr{I}$ or $\mathscr{L}^c=\mathscr{I}+\mathbf{C}I$.*

Proof. From Theorem 1, we have $\mathscr{L}\subseteqq\mathscr{I}+\mathbf{C}I$. It is elementary that if $\mathscr{J}$ is an ideal in $\mathscr{B}(\mathfrak{H})$, then $\mathscr{J}^c=\mathscr{J}$. Hence we have $\mathscr{L}^c\subseteqq\mathscr{I}^c+\mathbf{C}I=\mathscr{I}+\mathbf{C}I$.

Now we show $\mathscr{I}\subseteqq\mathscr{L}^c$. Let $T\in\mathscr{I}$. Then $\begin{pmatrix}0 & T\\ 0 & 0\end{pmatrix}\in\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{H}}$. Let W be the unitary operator on $\mathfrak{H}\oplus\mathfrak{H}$ given by the matrix $\frac{1}{\sqrt{2}}\begin{pmatrix}I & I\\ -I & I\end{pmatrix}$. Then, since $\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{H}}$ is unitarily invariant, we have

$$2W^*\begin{pmatrix}0 & T\\ 0 & 0\end{pmatrix}W=\begin{pmatrix}T & T\\ -T & -T\end{pmatrix}\in\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{H}}$$

and hence $T\in\mathscr{L}^c$.

We have shown that $\mathscr{I}\subseteqq\mathscr{L}^c\subseteqq\mathscr{I}+\mathbf{C}I$ from which it follows that either $\mathscr{L}^c=\mathscr{I}$ or $\mathscr{L}^c=\mathscr{I}+\mathbf{C}I$. □

Now we consider some "permanence properties" of Lie ideals and their associate ideals. First we state the following obvious fact without proof in order to put it into record.

Proposition 6. *If $\{\mathscr{L}_j\}$ is a family of Lie ideals in $\mathscr{B}(\mathfrak{H})$, then the intersection $\bigcap_j\mathscr{L}_j$ and the sum $\sum_j\mathscr{L}_j$ are Lie ideals also. If, furthermore, $\mathscr{I}_j$ is the associate ideal of $\mathscr{L}_j$ for each j, then the associate ideals of $\bigcap_j\mathscr{L}_j$ and $\sum_j\mathscr{L}_j$ are $\bigcap_j\mathscr{I}_j$ and $\sum_j\mathscr{I}_j$ respectively.*

The next "permanence property" is less obvious and more interesting.

Proposition 7. *If $\mathscr{L}_1$, $\mathscr{L}_2$ are Lie ideals in $\mathscr{B}(\mathfrak{H})$ with $\mathscr{I}_1$, $\mathscr{I}_2$ as their associate ideals, then $[\mathscr{L}_1,\mathscr{L}_2]$ is a Lie ideal with $\mathscr{I}_1\mathscr{I}_2$ as its associate ideal.*

Remark. It is easy to check that $\mathscr{I}_1\mathscr{I}_2$ is an ideal in $\mathscr{B}(\mathfrak{H})$. Since every ideal in $\mathscr{B}(\mathfrak{H})$ is self-adjoint, we have $\mathscr{I}_1\mathscr{I}_2=\mathscr{I}_2\mathscr{I}_1$.

Proof. It is easy to check that $[\mathscr{L}_1, \mathscr{L}_2]$ is a Lie ideal. Let $\mathscr{I}$ be the ideal associated with $[\mathscr{L}_1, \mathscr{L}_2]$. From $\mathscr{L}_j \subseteqq \mathscr{I}_j + \mathbf{C}I\,(j=1,2)$ we have $[\mathscr{L}_1, \mathscr{L}_2] \subseteqq \mathscr{I}_1\mathscr{I}_2 + \mathbf{C}I$. Hence, by (1)$\Leftrightarrow$(5) in Proposition 3, we have $\mathscr{I} \subseteqq \mathscr{I}_1\mathscr{I}_2$. Next, suppose that $A_j \in \mathscr{I}_j$ $(j=1,2)$. Then we have

$$\begin{pmatrix} 0 & A_1 \\ 0 & 0 \end{pmatrix} \in (\mathscr{L}_1)_{\mathfrak{H}\oplus\mathfrak{H}}, \quad \begin{pmatrix} 0 & 0 \\ A_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\begin{pmatrix} 0 & A_2 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \in (\mathscr{L}_2)_{\mathfrak{H}\oplus\mathfrak{H}}$$

and hence

$$\begin{pmatrix} A_1A_2 & 0 \\ 0 & -A_2A_1 \end{pmatrix} = \left[\begin{pmatrix} 0 & A_1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ A_2 & 0 \end{pmatrix}\right] \in [\mathscr{L}_1, \mathscr{L}_2]_{\mathfrak{H}\oplus\mathfrak{H}}.$$

Therefore it follows from Proposition 5 that $A_1A_2 \in \mathscr{I}$. □

Remark. In case $\mathscr{L}_1$ and $\mathscr{L}_2$ actually are ideals, i.e. $\mathscr{L}_1 = \mathscr{I}_1$ and $\mathscr{L}_2 = \mathscr{I}_2$, we have an easier proof as follows. We have to show $[\mathscr{B}(\mathfrak{H}), \mathscr{I}_1\mathscr{I}_2] \subseteqq [\mathscr{I}_1, \mathscr{I}_2] \subseteqq \subseteqq \mathscr{I}_1\mathscr{I}_2 + \mathbf{C}I$. The second inclusion is obvious. The first follows from the identity $[T, AB] = [TA, B] + [BT, A]$.

Example. Let $\mathscr{C}_1^0$ denote the set of trace class operators on $\mathfrak{H}$ of trace zero. Then, by using some properties of the trace function, we have $[\mathscr{C}_2, \mathscr{C}_2] \subseteqq \mathscr{C}_1^0$. G. Weiss [17] has shown that $[\mathscr{C}_2, \mathscr{C}_2] \neq \mathscr{C}_1^0$. Using this result, it is observed in [5, Remark 1] that $[\mathscr{B}(\mathfrak{H}), \mathscr{C}_1] \neq \mathscr{C}_1^0$. Now we claim that for no Lie ideal $\mathscr{L}$ do we have $\mathscr{C}_1^0 = [\mathscr{B}(\mathfrak{H}), \mathscr{L}]$ or $\mathscr{C}_1^0 = [\mathscr{L}, \mathscr{L}]$. Since, if the associate ideal of $\mathscr{L}$ is $\mathscr{I}$, the associate ideals of $\mathscr{C}_1^0$, $[\mathscr{B}(\mathfrak{H}), \mathscr{L}]$ and $[\mathscr{L}, \mathscr{L}]$ are $\mathscr{C}_1$, $\mathscr{I}$ and $\mathscr{I}^2$ respectively, $\mathscr{C}_1^0 = [\mathscr{B}(\mathfrak{H}), \mathscr{L}]$ would imply that $\mathscr{I} = \mathscr{C}_1$ and $\mathscr{C}_1^0 = [\mathscr{L}, \mathscr{L}]$ would imply $\mathscr{I}^2 = \mathscr{C}_1$, i.e. $\mathscr{I} = \mathscr{C}_2$, both contradictory to the results previously mentioned. We do not know whether we can have $\mathscr{C}_1^0 = [\mathscr{I}, \mathscr{J}]$ for distinct ideals $\mathscr{I}$ and $\mathscr{J}$.

3. A characterization of operator ideals and its applications

We now turn our attention to the main theme of this paper: characterizations of operator ideals.

Proposition 8. *A linear manifold $\mathscr{L}$ in $\mathscr{B}(\mathfrak{H})$ is either an ideal or $\mathscr{I} + \mathbf{C}I$ for some ideal $\mathscr{I}$ if and only if $\mathscr{L}$ is unitarily invariant and $\mathscr{L}^c \subseteqq \mathscr{L}$.*

Proof. Suppose that $\mathscr{L}$ is a unitarily invariant manifold in $\mathscr{B}(\mathfrak{H})$ and $\mathscr{L}^c \subseteqq \mathscr{L}$. It follows from Theorem 1 that $\mathscr{L}$ is a Lie ideal and $\mathscr{L} \subseteqq \mathscr{I} + \mathbf{C}I$ where $\mathscr{I}$ is its associate ideal. From Proposition 5, we have $\mathscr{I} \subseteqq \mathscr{L}^c$. Now we have $\mathscr{I} \subseteqq \mathscr{L}^c \subseteqq \subseteqq \mathscr{L} \subseteqq \mathscr{I} + \mathbf{C}I$. Hence either $\mathscr{L} = \mathscr{I}$ or $\mathscr{L} = \mathscr{I} + \mathbf{C}I$.

For the proof of the "only if" part, it suffices to note that, if $\mathscr{I}$ is an ideal in $\mathscr{B}(\mathfrak{H})$, then $\mathscr{I}_{\mathfrak{H}\oplus\mathfrak{H}}$ consists of all 2×2 operator matrices with entries in $\mathscr{I}$. □

The following immediate consequence of the above proposition is a useful characterization of ideals in $\mathscr{B}(\mathfrak{H})$.

Proposition 9. *A linear manifold $\mathscr{L}$ in $\mathscr{B}(\mathfrak{H})$ is a proper ideal if and only if it is unitarily invariant, $\mathscr{L}^c\subseteq\mathscr{L}$ and $I\notin\mathscr{L}$.*

Remark. The "if" part of Proposition 9, under the additional assumption that $\mathscr{L}$ contains all Hilbert—Schmidt operators, was obtained by Sourour [16].

Next we give a few applications, labelled as examples, of the above two propositions. In many cases, it is convenient to think of $\mathscr{L}^c$ in the following way. Take any subspace $\mathfrak{M}$ in $\mathfrak{H}$ with $\dim\mathfrak{M}=\dim\mathfrak{M}^\perp(=\dim\mathfrak{H})$ and let

$$\mathscr{L}^{\mathfrak{M}}=\{\text{compression of } T \text{ to } \mathfrak{M}\colon\ T\in\mathscr{L}\}.$$

Then $\mathscr{L}^{\mathfrak{M}}$ is a unitarily invariant manifold and $(\mathscr{L}^{\mathfrak{M}})_{\mathfrak{H}}=\mathscr{L}^c$. Thus, roughly speaking, $\mathscr{L}^c$ can be obtained by taking the compression of $\mathscr{L}$ to a subspace $\mathfrak{M}$ with $\dim\mathfrak{M}=$ $=\dim\mathfrak{M}^\perp$ and then transporting it back to $\mathfrak{H}$.

Example 1. Let $\mathscr{S}$ be a linear manifold of numerical sequences converging to zero. We consider the set $\mathscr{I}$ of those operators T such that, for each orthonormal sequence $\{e_n\}$ in $\mathfrak{H}$, the sequence $\{(Te_n, e_n)\}_{n=1}^\infty$ is in $\mathscr{S}$. Then it is easy to see that $\mathscr{I}$ is a unitarily invariant manifold which does not contain I. By the obvious fact that an orthonormal sequence in a subspace is also an orthonormal sequence in the whole Hilbert space, one can see the validity of the inclusion $\mathscr{I}^c\subseteq\mathscr{I}$. By Proposition 9, it follows that $\mathscr{I}$ is an ideal.

If we take $\mathscr{S}=l^p$, the set of all numerical sequences $\{\lambda_j\}$ such that $\sum_{j=1}^\infty|\lambda_j|^p<\infty$, then the corresponding ideal $\mathscr{I}$ turns out to be the $\mathscr{C}^p$-class of operators. If π_j is a sequence of positive numbers decreasing to zero such that $\sum_{n=1}^\infty\pi_j=\infty$, and if $\mathscr{S}$ is the set of numerical sequences $\{\lambda_j\}$ satisfying $\sum_{j=1}^\infty\pi_j|\lambda_j|<\infty$, then the corresponding ideal $\mathscr{I}$ is σ_π which is defined in [7].

Example 2. An operator T in $\mathscr{B}(\mathfrak{H})$ (here $\mathfrak{H}$ is assumed to be separable) is said to be universally absolutely bounded if, for every orthonormal basis $\{e_n\}$ in $\mathfrak{H}$, the matrix

$$\begin{pmatrix} |(Te_1, e_1)| & |(Te_2, e_1)| & \dots \\ |(Te_1, e_2)| & |(Te_2, e_2)| & \dots \\ \dots & \dots & \dots \end{pmatrix}$$

represents a bounded operator on l^2. Let $\mathscr{U}$ be the set of all universally bounded operators. Clearly $\mathscr{U}$ is a unitarily invariant manifold, $I\in\mathscr{U}$ and, for an infinite dimensional subspace $\mathfrak{M}$ of $\mathfrak{H}$, each operator in $\mathscr{U}^{\mathfrak{M}}$ is also universally absolutely bounded. Hence it follows from Proposition 8 that $\mathscr{U}=\mathscr{I}+\mathbf{C}I$ for some ideal $\mathscr{I}$. In fact, HALMOS and SUNDER [10] showed that $\mathscr{U}=\mathscr{C}_2+\mathbf{C}I$. Our discussion here can be used to shorten their proof.

Example 3. For $p>0$, let $\mathscr{U}_p$ be the set of all those operators $T\in\mathscr{B}(\mathfrak{H})$ ($\mathfrak{H}$ is separable) satisfying the condition that the matrix

$$\begin{pmatrix} |(Te_1, e_1)|^p & |(Te_2, e_1)|^p \dots \\ |(Te_1, e_2)|^p & |(Te_2, e_2)|^p \dots \\ \dots & \dots \quad \dots \end{pmatrix}$$

represents a bounded operator on l^2 for every orthonormal basis $\{e_n\}$ of $\mathfrak{H}$. From the inequality $(a+b)^p\leqq 2^p(a^p+b^p)$ $(a, b\geqq 0)$ we see that $\mathscr{U}_p$ is a linear manifold. It is easy to check that $\mathscr{U}_p$ is unitarily invariant, $I\in\mathscr{U}_p$ and $\mathscr{U}_p^c\subseteqq\mathscr{U}_p$. Hence, by Proposition 8, $\mathscr{U}_p=\mathscr{I}_p+\mathbf{C}I$ for some ideal $\mathscr{I}_p$. For $p\geqq 2$, it follows from a classical result of Schur (which says, for two $n\times n$ matrices (a_{kj}) and (b_{kj}), $\|(a_{kj}b_{kj})\|\leqq$ $\leqq\|(a_{kj})\|\,\|(b_{kj})\|$; see [15]) that $\mathscr{U}_p=\mathscr{B}(\mathfrak{H})$. As we have mentioned in Example 2, $\mathscr{U}_1=\mathscr{C}_2+\mathbf{C}I$. We do not know how to describe $\mathscr{I}_p$ in an explicit way when $1<p<2$ or $0<p<1$.

Example 4. Let $\mathscr{T}$ be the set of all those operators in $\mathscr{B}(\mathfrak{H})$ ($\mathfrak{H}$ is separable) which, in any matrix representation, allow triangular truncation. More precisely, $T\in\mathscr{T}$ if and only if, for an arbitrary orthonornal basis $\{e_n\}$ in $\mathfrak{H}$, the triangular matrix

$$\begin{pmatrix} (Te_1, e_1) & (Te_2, e_1) & (Te_3, e_1)\dots \\ 0 & (Te_2, e_2) & (Te_3, e_2)\dots \\ 0 & 0 & (Te_3, e_3)\dots \\ & & \ddots \end{pmatrix}$$

represents a bounded operator on l^2. Then it is easy to see that $\mathscr{T}$ is a unitarily invariant manifold and $I\in\mathscr{T}$. A little reflexion on forming submatrices reveals that $\mathscr{T}^c\subseteqq\mathscr{T}$. Hence it follows from Proposition 8 that $\mathscr{T}=\mathscr{I}+\mathbf{C}I$ for some ideal $\mathscr{I}$. It follows from a result of MACAEV (see [7]) that $\mathscr{I}$ contains all those operator T with their s-numbers $\{s_n(\tau)\}$ satisfying $\sum_{n=1}^{\infty} n^{-1}s_n(T)<\infty$.

Example 5. Let (X, m) be a separable σ-finite measure space which is not purely atomic. We say that an operator T on $\mathfrak{H}=\mathscr{L}^2(X, m)$ is an integral operator if $Tx(s)=\int k(s, t)x(t)\,dm(t)$ a.e. $(x\in\mathfrak{H})$ for some measurable function k on $X\times X$. Proposition 8 can be used to give a simplified proof of a result due to KOROTKOV: if U^*TU is an integral operator for every unitary U, then $T\in\mathscr{C}_2+\mathbf{C}I$. For details, we refer to SOUROUR [16].

4. Characterization of ideals in certain classes of C^*-algebras

In the present section, we give some characterizations of ideals in certain general C^*-algebras which share some "noncommutative" features with $\mathscr{B}(\mathfrak{H})$.

For the next two results, we consider those unital C^*-algebras $\mathscr{A}$ which satisfy the following condition:

(C) Every unitary element in $\mathscr{A}$ can be expressed as a product of a scalar and several symmetries (i.e. hermitian unitaries) in $\mathscr{A}$.

That $\mathscr{B}(\mathfrak{H})$ satisfies condition (C) is a consequence of the following result of HALMOS and KAKUTANI [9]:

Theorem 10. *Each operator on an infinite dimensional Hilbert space is a product of four symmetries.*

This result was generalized by FILLMORE [3] to properly infinite von Neumann algebras. Note that if $\mathscr{A}$ is a commutative C^*-algebra and $\dim \mathscr{A} \geqq 2$, then condition (C) fails.

The notion of unitarily invariant manifolds in $\mathscr{B}(\mathfrak{H})$ can be extended to general C^*-algebras in a straightforward manner: in a C^*-algebra $\mathscr{A}$, a linear manifold $\mathscr{L}$ is unitarily invariant if and only if $U^*\mathscr{L}U \subseteqq \mathscr{L}$ for all unitary elements U in $\mathscr{A}$. The following result characterizes unitarily invariant manifolds in a C^*-algebra satisfying condition (C).

Proposition 11. *A linear manifold $\mathscr{L}$ in a C^*-algebra $\mathscr{A}$ satisfying* (C) *is unitarily invariant if and only if $(I-P)\mathscr{L}P \subseteqq \mathscr{L}$ for all projections P in $\mathscr{A}$.*

Proof. Suppose that $\mathscr{L}$ is unitarily invariant. Let P be a projection in $\mathscr{A}$ and $T \in \mathscr{L}$. Then both $U = I-2P$ and $V = P + i(I-P)$ are unitary and hence $T_1 \equiv \frac{1}{2}(T - U^*TU) \in \mathscr{L}$ and

$$(I-P)TP = \frac{1}{2}(T_1 - iV^*T_1V) \in \mathscr{L}.$$

Conversely, suppose that $(I-P)\mathscr{L}P \subseteqq \mathscr{L}$ for all projections P. Let S be a symmetry in $\mathscr{A}$. Then $S = 2P-I$ for some projection P. Hence, for $T \in \mathscr{L}$,

$$STS = T - 2\big(PT(I-P) + (I-P)TP\big) \in \mathscr{L}.$$

By condition (C) we see that $\mathscr{L}$ is unitarily invariant, since $\mathscr{A}$ is linearly spanned by unitaries. □

Proposition 12. *A linear manifold $\mathscr{L}$ in a C^*-algebra $\mathscr{A}$ satisfying* (C) *is an ideal if and only if $P\mathscr{L}P \subseteqq \mathscr{L}$ for all projections P in $\mathscr{A}$.*

Proof. Suppose that $P\mathscr{L}P\subseteqq\mathscr{L}$ for all projections P. Let $T\in\mathscr{L}$ and S be a symmetry so that $S=2P-I$ for some projection P. Then

$$STS=2(PTP+(I-P)T(I-P))-T\in\mathscr{L}.$$

Hence, by condition (C), $\mathscr{L}$ is unitarily invariant. Therefore, by Proposition 11, $(I-P)\mathscr{L}P\subseteqq\mathscr{L}$ for all projections P in $\mathscr{A}$. Now, for a symmetry $S=2P-I$ and $T\in\mathscr{L}$, we have

$$TS=2(PTP+(I-P)TP)-T\in\mathscr{L}.$$

By condition (C) again, we have $TU\in\mathscr{L}$ for each $T\in\mathscr{L}$ and each unitary U. Since unitary elements span $\mathscr{A}$ linearly, we have $\mathscr{L}\mathscr{A}\subseteqq\mathscr{L}$. In the same way we can show that $\mathscr{A}\mathscr{L}\subseteqq\mathscr{L}$. Hence $\mathscr{L}$ is an ideal of $\mathscr{A}$. □

A linear manifold $\mathscr{L}$ in a C^*-algebra $\mathscr{A}$ is said to be a *Jordan ideal* if $AX+XA\in\mathscr{L}$ for all $A\in\mathscr{L}$ and $X\in\mathscr{A}$. It is shown in [5, Theorem 3] that Jordan ideals in $\mathscr{B}(\mathfrak{H})$ are just associative ideals. This result can be generalized for a class of C^*-algebras wider than $\mathscr{B}(\mathfrak{H})$:

Corollary 13. *If $\mathscr{L}$ is a Jordan ideal in a C^*-algebra $\mathscr{A}$ which satisfies condition* (C), *then $\mathscr{L}$ is an associative ideal.*

Proof. Let P be a projection in $\mathscr{A}$ and $T\in\mathscr{L}$. Then

$$P(PT+TP)+(PT+TP)P=2PTP+(PT+TP)\in\mathscr{L}$$

and hence $PTP\in\mathscr{L}$. Now the corollary follows from Proposition 12. □

Sourour has informed the authors that, in case $\mathscr{A}=\mathscr{B}(\mathfrak{H})$, Proposition 12 can be deduced in the following way. Assume that $\mathscr{L}$ is a linear manifold such that $P\mathscr{L}P\subseteqq\mathscr{L}$ for all projections P. For $T\in\mathscr{L}$ and a projection P we have $TP+PT=T+PTP-(I-P)T(I-P)\in\mathscr{L}$. By the fact that projections span $\mathscr{B}(\mathfrak{H})$ linearly (Theorem 2), we see that $\mathscr{L}$ is a Jordan ideal. Now it follows from [5, Theorem 3] that $\mathscr{L}$ is an associative ideal.

The condition (C) in Proposition 12 is essential. For example, if $\mathscr{A}=C[0, 1]$, then there is no proper projection in $\mathscr{A}$ and hence the inclusion $P\mathscr{L}P\subseteqq\mathscr{L}$ is automatically satisfied for every linear manifold $\mathscr{L}$ in $\mathscr{A}$; but of course there are linear manifolds in $\mathscr{A}$ which are not ideals.

In the next result, we let $\mathscr{B}$ be a C^*-algebra with the identity I, $\mathscr{A}=\mathscr{M}_2(\mathscr{B})$ and P_0 be the projection in $\mathscr{A}$ given by the matrix $\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$.

Proposition 14. *A linear manifold $\mathscr{L}$ in $\mathscr{A}$ is an ideal if and only if $\mathscr{L}$ is unitarily invariant and $P_0\mathscr{L}P_0\subseteqq\mathscr{L}$.*

Proof. Suppose $\mathscr{L}$ is unitarily invariant and $P_0\mathscr{L}P_0 \subseteqq \mathscr{L}$. Let

$$\mathscr{I} = \left\{B \in \mathscr{B} : \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix} \in \mathscr{L}\right\}.$$

If U is a unitary element in $\mathscr{B}$ and $B \in \mathscr{I}$, then

$$\begin{pmatrix} I & 0 \\ 0 & U \end{pmatrix}^* \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & U \end{pmatrix} = \begin{pmatrix} 0 & BU \\ 0 & 0 \end{pmatrix} \in \mathscr{L},$$
$$\begin{pmatrix} U & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U & 0 \\ 0 & I \end{pmatrix}^* = \begin{pmatrix} 0 & UB \\ 0 & 0 \end{pmatrix} \in \mathscr{L},$$

and hence BU and UB are in $\mathscr{I}$. Since unitary elements in $\mathscr{B}$ span the whole algebra $\mathscr{B}$, we see that $\mathscr{I}$ is an ideal in $\mathscr{B}$. Now let $\mathscr{J}$ be the set of all 2×2 matrices with entries in $\mathscr{I}$. Then $\mathscr{J}$ is an ideal in $\mathscr{A}$.

Let $T = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ be an element in $\mathscr{L}$. We are going to show that $T \in \mathscr{J}$. For this purpose, we introduce the following unitary elements in $\mathscr{A}$:

$$U = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad V = \begin{pmatrix} I & 0 \\ 0 & iI \end{pmatrix}, \quad J = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \quad W = \frac{1}{\sqrt{2}} \begin{pmatrix} I & I \\ -I & I \end{pmatrix}.$$

Then we have

$$S \equiv \frac{1}{2}(T - U^*TU) = \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix} \in \mathscr{L}, \quad \frac{1}{2}(S - iV^*SV) = \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix} \in \mathscr{L}$$

and

$$\frac{1}{2} J(S + iV^*SV)J = \begin{pmatrix} 0 & C \\ 0 & 0 \end{pmatrix} \in \mathscr{L}.$$

Hence, $B, C \in \mathscr{I}$. We also have

$$2W^*P_0TP_0W = \begin{pmatrix} A & A \\ A & A \end{pmatrix} \in \mathscr{L}, \quad 2W^*P_0(JTJ)P_0W = \begin{pmatrix} D & D \\ D & D \end{pmatrix} \in \mathscr{L}.$$

By the previous argument, we have $A, D \in \mathscr{I}$.

Next we show that $\mathscr{J} \subseteqq \mathscr{L}$ and let $T = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathscr{I}$. From the definition of $\mathscr{J}$ we know A, B, C and D are in $\mathscr{I}$, or, in other words,

$$S_1 = \begin{pmatrix} 0 & A \\ 0 & 0 \end{pmatrix}, \quad S_2 = \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}, \quad S_3 = \begin{pmatrix} 0 & C \\ 0 & 0 \end{pmatrix}, \quad S_4 = \begin{pmatrix} 0 & D \\ 0 & 0 \end{pmatrix}$$

are in $\mathscr{L}$. We have to show that

$$T_1 = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \quad T_2 = \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}, \quad T_3 = \begin{pmatrix} 0 & 0 \\ C & 0 \end{pmatrix}, \quad T_4 = \begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix}$$

are in $\mathscr{L}$. This can be seen from the following identities:

$$T_1 = 2P_0(WS_1W^*)P_0, \quad T_2 = S_2, \quad T_3 = JS_3J, \quad T_4 = 2J(P_0WS_4W^*P_0)J,$$

where J and W are the unitary operators previously defined. □

Ideas similar to those in the above proof appear in [12].

Corollary 15. *If P_0 is a projection in $\mathscr{B}(\mathfrak{H})$ with $\dim P_0\mathfrak{H}=\dim(I-P_0)\mathfrak{H}$ and if $\mathscr{L}$ is a unitarily invariant manifold in $\mathscr{B}(\mathfrak{H})$ satisfying $P_0\mathscr{L}P_0\subseteq\mathscr{L}$, then $\mathscr{L}$ is an ideal in $\mathscr{B}(\mathfrak{H})$.*

Proof. This follows from the previous proposition and the fact that $\mathscr{B}(\mathfrak{H})$ and $\mathscr{M}_2(\mathscr{B}(\mathfrak{H}))$ are isomorphic C^*-algebras. □

Example. Let $\mathscr{L}\subseteq\mathscr{K}(\mathfrak{H})$ be an ideal in $\mathscr{K}(\mathfrak{H})$, i.e., for $X\in\mathscr{K}(\mathfrak{H})$ and $A\in\mathscr{L}$, we have $XA\in\mathscr{L}$ and $AX\in\mathscr{L}$. In general, $\mathscr{L}$ is not necessarily an ideal in $\mathscr{B}(\mathfrak{H})$. Among other things, it was shown in [6] that if $\mathscr{L}$ is also a Lie ideal and $\mathscr{L}$ is countably generated as an ideal of $\mathscr{K}(\mathfrak{H})$, then $\mathscr{L}$ is also an ideal of $\mathscr{B}(\mathfrak{H})$. This result can be proved in the following alternative way.

Let $\mathscr{I}$ be the linear span of operators of the form XAY, where $A\in\mathscr{S}$ and $X, Y\in\mathscr{K}(\mathfrak{H})$. It is easy to see that $\mathscr{I}$ is an ideal in $\mathscr{B}(\mathfrak{H})$ and $\mathscr{I}\subseteq\mathscr{L}$. On the other hand, it follows from a lemma in [6] that there is a projection P_0 in $\mathscr{B}(\mathfrak{H})$ such that $\dim P_0\mathfrak{H}=\dim(I-P_0)\mathfrak{H}$ and P_0S, $SP_0\in\mathscr{I}$ for all S in $\mathscr{S}$. Notice that each element S in $\mathscr{L}$ can be expressed as a finite sum:

$$S = D+\Sigma_j(\alpha_jA_j+B_jX_j+Y_jC_j)$$

where $\alpha_j\in\mathbf{C}$; $A_j, B_j, C_j\in\mathscr{S}$; $X_j, Y_j\in\mathscr{K}(\mathfrak{H})$ and $D\in\mathscr{I}$. For such a sum, we have

$$P_0SP_0 = P_0DP_0+\Sigma_j(\alpha_jP_0(A_jP_0)+(P_0B_j)XP_0+P_0Y_j(C_jP_0))\in\mathscr{I}$$

since $\mathscr{I}$ is an ideal in $\mathscr{B}(\mathfrak{H})$ and the operators D, A_jP_0, P_0B_j, C_jP_0 are all in $\mathscr{I}$. Now it follows from Corollary 15 and $\mathscr{I}\subseteq\mathscr{L}$ that $\mathscr{L}$ is an ideal in $\mathscr{B}(\mathfrak{H})$.

Finally, we have the following characterization of ideals in $\mathscr{B}(\mathfrak{H})$.

Proposition 16. *If $\mathscr{L}$ is a unitarily invariant manifold in $\mathscr{B}(\mathfrak{H})$ consisting of compact operators and if $T\in\mathscr{L}$ implies $|T|=(T^*T)^{1/2}\in\mathscr{L}$, then $\mathscr{L}$ is an ideal.*

Proof. Again, let $\mathscr{I}$ be the ideal of those operators B such that $\begin{pmatrix}0 & B\\ 0 & 0\end{pmatrix}\in\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{H}}$. Then, from Theorem 1, we have $\mathscr{L}\subseteq\mathscr{I}$. Next we show that $\mathscr{I}\subseteq\mathscr{L}$. Since every ideal in $\mathscr{B}(\mathfrak{H})$ is linearly spanned by its positive elements, it suffices to show that positive elements in $\mathscr{I}_{\mathfrak{H}\oplus\mathfrak{H}}$ are in $\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{H}}$. So let $T=\begin{pmatrix}A & B\\ C & D\end{pmatrix}$ be a positive element in $\mathscr{I}_{\mathfrak{H}\oplus\mathfrak{H}}$. Then A, B, C and D are in $\mathscr{I}$. Hence

$$\begin{pmatrix}0 & B\\ C & 0\end{pmatrix} = \begin{pmatrix}0 & B\\ 0 & 0\end{pmatrix}+\begin{pmatrix}0 & I\\ I & 0\end{pmatrix}\begin{pmatrix}0 & C\\ 0 & 0\end{pmatrix}\begin{pmatrix}0 & I\\ I & 0\end{pmatrix}\in\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{H}},$$

$$\begin{pmatrix}0 & 0\\ 0 & D\end{pmatrix} = \left|\begin{pmatrix}0 & D\\ 0 & 0\end{pmatrix}\right|\in\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{H}},\quad \begin{pmatrix}A & 0\\ 0 & 0\end{pmatrix} = \left|\begin{pmatrix}0 & I\\ I & 0\end{pmatrix}\begin{pmatrix}0 & A\\ 0 & 0\end{pmatrix}\begin{pmatrix}0 & I\\ I & 0\end{pmatrix}\right|\in\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{H}}.$$

Hence we obtain $T\in\mathscr{L}_{\mathfrak{H}\oplus\mathfrak{S}}$. □

Appendix

In this appendix we give a transparent proof of Theorem 1 and Theorem 2 based on an idea in [4]. The main tool we use in this proof is Halmos—Kakutani's Theorem (Theorem 10): every unitary operator can be expressed as a product of not more than four symmetries. This theorem can be deduced constructively by the following three short steps: first express it as a direct sum of countably many blocks such that each block has the same dimension as the Hilbert space; then, using this expression, write the operator as a product of two bilateral shifts (of infinite rank); finally, write each bilateral shift as a product of two symmetries. (For details, we refer to [9].) From this argument we see that the symmetries involved can be chosen in such a way that their eigen-subspaces have the same dimension as the underlying Hilbert space.

In order to reveal the essential part of our argument in proving Theorem 1, we consider a more general situation. We let $\mathscr{B}$ be a unital C^*-algebra, $\mathscr{A}=\mathscr{M}_2(\mathscr{B})$ and $\mathscr{S}$ be the set of all those symmetries of the form $U^*\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}U$, where I is the identity in $\mathscr{B}$ and U is a unitary element in $\mathscr{A}$. We consider the following condition:

(C′) each unitary element in $\mathscr{A}$ is a product of finitely many elements in $\mathscr{S}$ and a scalar.

It follows from our previous remark that for $\mathscr{A}=\mathscr{B}(\mathfrak{H})$, condition (C′) is satisfied.

In the following three lemmas, we always assume that $\mathscr{A}$ is the C^*-algebra described in the previous paragraph and condition (C′) is satisfied. Furthermore, we assume that $\mathscr{L}$ is a unitarily invariant manifold in $\mathscr{A}$,

$$\mathscr{J}=\left\{B\in\mathscr{B}\colon \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}\in\mathscr{L}\right\}$$

and

$$\mathscr{I}=\left\{\begin{pmatrix} A & B \\ C & D \end{pmatrix}\in\mathscr{A}\colon A, B, C \text{ and } D \text{ are in } \mathscr{J}\right\}.$$

By using the same argument as that in the proof of Proposition 14, we see that $\mathscr{J}$ is an ideal in $\mathscr{B}$ and $\mathscr{I}$ is an ideal in $\mathscr{A}$.

Lemma A. *With the above assumption, we have* $[\mathscr{I},\mathscr{A}]\subseteq\mathscr{L}$.

Proof. It suffices to show that $[\mathscr{I}, U]\subseteq\mathscr{L}$ for all unitary elements U in $\mathscr{A}$. First we note that if $\begin{pmatrix} A & B \\ C & D \end{pmatrix}\in\mathscr{I}$, then

$$\left[\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\right]=2\begin{pmatrix} 0 & -B \\ C & 0 \end{pmatrix}=2\left(\begin{pmatrix} 0 & -B \\ 0 & 0 \end{pmatrix}+\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\begin{pmatrix} 0 & C \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\right)\in\mathscr{L}$$

since $\mathscr{L}$ is unitary and $B, C \in \mathscr{J}$. Now if $T \in \mathscr{I}$ and $S = W^* \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} W \in \mathscr{S}$ (W is unitary), then

$$[T, S] = W^* \left[WTW^*, \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \right] W \in \mathscr{L}.$$

Now we consider an arbitrary unitary element U in $\mathscr{A}$ and show that $[T, U] \in \mathscr{L}$. By condition (C′), U can be written as a product $\lambda S_1 S_2 \dots S_n$ where $S_1, \dots, S_n \in \mathscr{S}$ and $\lambda \in \mathbf{C}$. We proceed by induction on n. Let $V = S_2 S_3 \dots S_n$. Then

$$[T, U] = \lambda [TS_1, V] + \lambda [VT, S_1].$$

Since $VT \in \mathscr{I}$ and $TS_1 \in \mathscr{I}$, we have $[VT, S_1] \in \mathscr{L}$ by our previous argument and $[TS_1, V] \in \mathscr{L}$ by our induction assumption. Therefore $[T, U] \in \mathscr{L}$. □

Lemma B. *The linear span of $\mathscr{S}$ includes $[\mathscr{A}, \mathscr{A}]$.*

Proof. Let $\mathscr{L}_0$ be the linear span of $\mathscr{S}$. Then $\mathscr{L}_0$ is unitarily invariant. Let $\mathscr{J}_0$ and $\mathscr{I}_0$ be the ideals defined from $\mathscr{L}_0$ in the same way as $\mathscr{J}, \mathscr{I}$ defined from $\mathscr{L}$. Since $\frac{1}{2}\begin{pmatrix} I & I \\ I & I \end{pmatrix} \in \mathscr{L}_0$, we see that $I \in \mathscr{J}_0$ and hence $\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \in \mathscr{I}_0$. Therefore $\mathscr{I}_0 = \mathscr{A}$. Thus, by Lemma A, $[\mathscr{A}, \mathscr{A}] = [\mathscr{A}, \mathscr{I}_0] \subset \mathscr{L}_0$. □

Lemma C. $[\mathscr{L}, [\mathscr{A}, \mathscr{A}]] \subseteq \mathscr{I} \cap \mathscr{L}$.

Proof. It follows from Lemma B that it suffices to show $[\mathscr{L}, \mathscr{S}] \subseteq \mathscr{I} \cap \mathscr{L}$. If $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathscr{L}$, then, by an argument similar to that in the proof of Proposition 14, we can show that $\begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ C & 0 \end{pmatrix}$ are in $\mathscr{I} \cap \mathscr{L}$ and hence

$$\left[\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \right] = 2 \begin{pmatrix} 0 & -B \\ C & 0 \end{pmatrix} \in \mathscr{I} \cap \mathscr{L}.$$

If $S = W^* \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} W \in \mathscr{S}$, where W is unitary, and if $T \in \mathscr{L}$, then

$$[T, S] = W^* \left[WTW^*, \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \right] W \in \mathscr{I} \cap \mathscr{L}. \quad \square$$

Proof of Theorem 2. Apply Lemma B to the case $\mathscr{A} = \mathscr{B}(\mathfrak{H})$ and note that $[\mathscr{B}(\mathfrak{H}), \mathscr{B}(\mathfrak{H})] = \mathscr{B}(\mathfrak{H})$, (see [8]). □

Proof of Theorem 1. By Halmos—Kakutani's Theorem and Theorem 2, we can easily deduce the equivalence of (1) and (2). (For details, see [5].) That (3) implies (1) is obvious. It remains to show (1)⇒(3). By Lemma A, Lemma C and the fact that $[\mathscr{B}(\mathfrak{H}), \mathscr{B}(\mathfrak{H})] = \mathscr{B}(\mathfrak{H})$, we have $[\mathscr{B}(\mathfrak{H}), \mathscr{I}] \subseteq \mathscr{L}$ and $[\mathscr{B}(\mathfrak{H}), \mathscr{L}] \subseteq \mathscr{I}$.

Now Theorem 1 follows from the following theorem of Calkin [1]:

Theorem D. *If $\mathcal{I}$ is a proper ideal in $\mathcal{B}(\mathfrak{H})$, $T\in\mathcal{B}(\mathfrak{H})$ and if $[T, \mathcal{B}(\mathfrak{H})]\subseteq\mathcal{I}$, then $T\in\mathcal{I}+\mathbf{C}I$.*

Calkin only showed this theorem for the case when $\mathfrak{H}$ is separable. Now we prove this theorem under the assumption that $\mathfrak{H}$ is nonseparable.

Let $\mathcal{M}(\mathfrak{H})$ be the unique maximal ideal in $\mathcal{B}(\mathfrak{H})$. (Thus, for an operator S on $\mathfrak{H}$, $S\in\mathcal{M}(\mathfrak{H})$ if and only if there exists a projection E in $\mathcal{B}(\mathfrak{H})$ such that $ESE=S$ and $\dim E\mathfrak{H}<\dim\mathfrak{H}$.) Let $\mathcal{C}(\mathfrak{H})$ be the "Calkin algebra" $\mathcal{B}(\mathfrak{H})/\mathcal{M}(\mathfrak{H})$. Let t be the canonicali mage of T in $\mathcal{C}(\mathfrak{H})$.

Since $\mathcal{I}$ is self-adjoint, with no loss of generality, we may assume $T=T^*$. Let $T=\int\lambda\, dE_\lambda$ be the spectral decomposition of T. Note that $\lambda\in\sigma(t)$ if and only if, for all $\varepsilon>0$, $\dim E(\lambda-\varepsilon, \lambda+\varepsilon)\mathfrak{H}=\dim\mathfrak{H}$.

First we demonstrate that $\sigma(t)$ is a singleton. Assume the contrary: we have $\lambda_1, \lambda_2\in\sigma(t)$ with $\lambda_1\neq\lambda_2$. Choose $\varepsilon>0$ such that the intervals $[\lambda_1-\varepsilon, \lambda_1+\varepsilon]$ and $[\lambda_2-\varepsilon, \lambda_2+\varepsilon]$ are disjoint. Let $\mathfrak{H}_j=E[\lambda_j-\varepsilon, \lambda_j+\varepsilon]\mathfrak{H}$ $(j=1, 2)$ and $\mathfrak{K}=\mathfrak{H}\ominus(\mathfrak{H}_1\oplus\mathfrak{H}_2)$. Then $\dim\mathfrak{H}_1=\dim\mathfrak{H}_2=\dim\mathfrak{H}$. Let U be a unitary transformation from $\mathfrak{H}$ onto $\mathfrak{H}\oplus\mathfrak{H}\oplus\mathfrak{K}$ such that $U\mathfrak{H}_1=\mathfrak{H}\oplus 0\oplus 0$, $U\mathfrak{H}_2=0\oplus\mathfrak{H}\oplus 0$ and $U\mathfrak{K}=0\oplus 0\oplus\mathfrak{K}$. Then

$$U^{-1}TU=\left(\begin{array}{cc|c} T_1 & 0 & \cdot \\ 0 & T_2 & \\ \hline \cdot & & * \end{array}\right)$$

for some hermitian operators T_1 and T_2 in $\mathcal{B}(\mathfrak{H})$ with disjoint spectra. By a well-known result of ROSENBLUM [14], there exists an operator A in $\mathcal{B}(\mathfrak{H})$ such that $T_1A-AT_2=I$. Let

$$X=U\left(\begin{array}{cc|c} 0 & A & \\ 0 & 0 & \\ \hline & & 0 \end{array}\right)U^{-1}.$$

Then

$$TX-XT=U\left(\begin{array}{cc|c} 0 & I & \\ 0 & 0 & \\ \hline & & 0 \end{array}\right)U^{-1}\in\mathcal{I}$$

and hence $I\in\mathcal{I}$. Therefore $\mathcal{I}=\mathcal{B}(\mathfrak{H})$, a contradiction to our assumption that $\mathcal{I}$ is proper.

We have $T=S+\lambda I$ for some $\lambda\in\mathbf{C}$ and self-adjoint operator S in $\mathcal{M}(\mathfrak{H})$. Choose a projection E in $\mathcal{B}(\mathfrak{H})$ such that $ESE=S$ and $\dim E\mathfrak{H}=\dim(I-E)\mathfrak{H}$. Let W be a unitary transformation from $\mathfrak{H}$ onto $\mathfrak{H}\oplus\mathfrak{H}$ such that $W(E\mathfrak{H})=\mathfrak{H}\oplus 0$ and $W((I-E)\mathfrak{H})=0\oplus\mathfrak{H}$. Then $WSW^{-1}=\begin{pmatrix} S_0 & 0 \\ 0 & 0\end{pmatrix}$ for some $S_0\in\mathcal{B}(\mathfrak{H})$. Let

$V=W^{-1}\begin{pmatrix}0 & I\\ 0 & 0\end{pmatrix}W$. Then we have

$$SV-VS=W^{-1}\begin{pmatrix}0 & S_0\\ 0 & 0\end{pmatrix}W\in\mathscr{I}$$

and hence $S_0\in\mathscr{I}$. Therefore $S\in\mathscr{I}$. □

By using the above three lemmas, Fillmore's extension [3] of Halmos and Kakutani's Theorem and the fact that $[\mathscr{A}, \mathscr{A}]=\mathscr{A}$ for a properly infinite von Neumann algebra [13], we can show the following two results.

Theorem 1′. *Let $\mathscr{L}$ be a linear manifold in a properly infinite von Neumann algebra. Then the following conditions are equivalent:*

(1) *$\mathscr{L}$ is unitarily invariant;*
(2) *$\mathscr{L}$ is a Lie ideal in $\mathscr{A}$, i.e., $[\mathscr{L}, \mathscr{A}]\subseteqq\mathscr{L}$;*
(3) *there is an ideal $\mathscr{I}$ in $\mathscr{A}$ such that $[\mathscr{A}, \mathscr{I}]\subseteqq\mathscr{L}$ and $[\mathscr{A}, \mathscr{L}]\subseteqq\mathscr{I}$.*

Theorem 2′ [13]. *Every element in a properly infinite von Neumann algebra is a linear combination of projections.*

As in Section 2, in a properly infinite von Neumann algebra, we can define the associate ideals of Lie ideals. Also we can show that conditions (1), (3), (4), (6) in Proposition 3 are equivalent in this general situation.

References

[1] J. W. Calkin, Two-sided ideals and congruences in the ring of bounded operators in Hilbert space, *Ann. of Math.*, (2) **42** (1941), 839—873.
[2] P. Fillmore, On products of symmetries, *Canad. J. Math.*, **18** (1966), 897—900.
[3] P. Fillmore, Sums of operators with square zero, *Acta Sci. Math.*, **28** (1967), 285—288.
[4] C. K. Fong, Equivalence of unitarily invariant norms, submitted.
[5] C. K. Fong, C. R. Miers and A. R. Sourour, Lie and Jordan ideals of operators on Hilbert space, *Proc. Amer. Math. Soc.*, **84** (1982), 516—520.
[6] C. K. Fong and H. Radjavi, On ideals and Lie ideals of compact operators, submitted.
[7] I. C. Gohberg and M. G. Krein, *Theory and Applications of Volterra Operators in Hilbert space*, Trans. Math. Monographs, Vol. 24, Amer. Math. Soc. (Providence, 1970).
[8] P. R. Halmos, Commutators of operators, *Amer. J. Math.*, **74** (1952), 237—240.
[9] P. R. Halmos and S. Kakutani, Products of symmetries, *Bull. Amer. Math. Soc.*, **64** (1958), 77—78.
[10] P. R. Halmos and V. S. Sunder, *Bounded integral operators on L^2*, Ergeb. Math. Grenzgeb. **96**, Springer (New York, 1978).
[11] C. R. Miers, Closed Lie ideals in operator algebras, preprint.
[12] G. J. Murphy, Lie ideals in associative algebras, submitted.

[13] C. Pearcy and D. M. Topping, Sums of small numbers of idempotents, *Mich. J. Math.*, **14** (1967), 453—465.

[14] M. Rosenblum, On the operator equation $BX-XA=Q$, *Duke Math. J.*, **23** (1956), 263—269.

[15] I. Schur, Bemerkungen zur Theorie der Beschrankten bilinearformen mit unendlich vielen veränder, *J. Reine Angew. Math.*, **140** (1911), 1—28.

[16] A. R. Sourour, A note on integral operators, *Acta Sci. Math.*, **41** (1979), 375—379.

[17] G. Weiss, Commutators of Hilbert—Schmidt operators, preprint.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TORONTO
TORONTO, CANADA

C. K. Fong's current address:
DEPARTMENT OF MATHEMATICS
AND STATISTICS
CARLETON UNIVERSITY
OTTAWA, ONTARIO K1S 5B6
CANADA

Acta Sci. Math., 51 (1987), 457—460

A note on Schmüdgen's classes $\mathfrak{N}_1$ and $\mathfrak{N}_\infty^\infty$ of pairs generated by Toeplitz operators

V. VASYUNIN*)

1. K. SCHMÜDGEN [1] introduced the following class of pairs of (unbounded) self-adjoint operators.

Definition 1. Let A, B be self-adjoint operators on a Hilbert space $\mathscr{H}$. The pair $\{A, B\}$ belongs to the class $\mathfrak{N}_1$ if there exists a dense linear manifold $\mathscr{D}$ in $\mathscr{H}$ such that

(i) $\mathscr{D} \subseteq \operatorname{Dom}(AB) \cap \operatorname{Dom}(BA)$ and $ABf = BAf$ for all $f \in \mathscr{D}$,

(ii) $A|\mathscr{D}$ and $B|\mathscr{D}$ are essentially self-adjoint.

Schmüdgen gives the following criterion for a pair $\{A, B\}$ to be in $\mathfrak{N}_1$. (In what follows $\mathscr{R}(\cdot)$ means "range of".)

Theorem 0 (Theorem 1.7 in [1]). *Suppose $\{A, B\} \in \mathfrak{N}_1$, $\alpha \in \mathbf{R} \setminus \sigma(A)$ and $\beta \in \mathbf{R} \setminus \sigma(B)$. Then the operators $X \stackrel{\text{def}}{=} (A-\alpha)^{-1}$ and $Y \stackrel{\text{def}}{=} (B-\beta)^{-1}$ satisfy the following conditions:*

$$\operatorname{Ker} X = \operatorname{Ker} Y = \{0\}, \tag{1}$$

$$\overline{\mathscr{R}([X, Y])} \cap \mathscr{R}(X) = \overline{\mathscr{R}([X, Y])} \cap \mathscr{R}(Y) = \{0\}. \tag{2}$$

Conversely, if X and Y are bounded self-adjoint operators satisfying (1) *and* (2), *then $\{X^{-1}+\alpha, Y^{-1}+\beta\} \in \mathfrak{N}_1$ for all $\alpha, \beta \in \mathbf{R}$.*

The main method in [1] to construct pairs belonging to $\mathfrak{N}_1$ is to consider pairs of the form $\{(\operatorname{Re} T)^{-1}, (\operatorname{Im} T)^{-1}\}$ for certain operators T. Among others Toeplitz operators with analytic symbols have been investigated in [1]. It was shown that Toeplitz operators with symbols which are cyclic for the backward shift do not generate a pair in $\mathfrak{N}_1$ ([1], Proposition 3.3). Moreover, the polynomials φ for which $\{(\operatorname{Re} T_\varphi)^{-1}, (\operatorname{Im} T_\varphi)^{-1}\} \in \mathfrak{N}_1$ are characterized in [1].

*) Research supported by Naturwissenschaftlich-Theoretisches Zentrum (Karl-Marx-Universität, Leipzig).

Received October 16, 1984.

The aim of this note is to show that Schmüdgen's method works in fact for Toeplitz operators with arbitrary analytic (or antianalytic) symbols.

Suppose $\varphi\in H^\infty$. Let T_φ be the multiplication by φ on H^2. Let $X \overset{\text{def}}{=} \operatorname{Re} T_\varphi$ and $Y \overset{\text{def}}{=} \operatorname{Im} T_\varphi$. As usual, S^* is the backward shift, P_+ is the orthogonal projection of L^2 onto H^2 and $P_-=I-P_+$ is the projection onto H^2_- and $\bigvee\{\dots\}$ denotes the closed linear span of $\{\dots\}$.

Lemma 1. $\mathscr{R}([X, Y])=\bigvee\{(S^*)^n\varphi\colon n\geqq 1\}$.

Proof. First note that, for any $h\in H^2$, $[T_\varphi^*, T_\varphi]h=(P_+\bar\varphi\varphi-\varphi P_+\bar\varphi)h=$ $=P_+\varphi P_-\bar\varphi h=H_{\bar\varphi}^* H_{\bar\varphi} h$, where $H_{\bar\varphi}\colon H^2\to H^2_-$ $(H_{\bar\varphi}h=P_-\bar\varphi h)$ is the Hankel operator with symbol $\bar\varphi$. Hence we have $\overline{\mathscr{R}([X,Y])}=\overline{\mathscr{R}([T_\varphi^*, T_\varphi])}=\overline{\mathscr{R}(H_{\bar\varphi}^*H_{\bar\varphi})}=\overline{\mathscr{R}(H_{\bar\varphi}^*)}=$ $=\overline{P_+\varphi H^2_-}=\bigvee\{P_+\bar z^n\varphi\colon n\geqq 1\}$. Now the assertion follows.

According to Beurling's theorem, the S^*-invariant subspace $\bigvee\{(S^*)^n\varphi\colon n\geqq 1\}$ has the form $H^2\ominus\Theta H^2$ with a certain inner function Θ or $\Theta=0$. We introduce the bounded analytic functions φ_+ and φ_- by

$$\varphi_\pm(z)=\frac{1}{2}\Theta(z)\big(\varphi(z)\pm\overline{\varphi(z)}\big)\quad\text{for}\quad |z|=1.$$

$\Theta\bar\varphi$ is indeed analytic, because $(\bar\varphi\Theta, \bar z^n)=(\Theta, (S^*)^n\varphi)=0$ for $n\geqq 1$.

Theorem 1. $\{(\operatorname{Re} T_\varphi)^{-1}, (\operatorname{Im} T_\varphi)^{-1}\}\in\mathfrak{N}_1$ *if and only if φ_+ and φ_- are non-zero outer functions.*

Proof. Let us note at first that for $\Theta=0$ we have by Lemma 1 $\overline{\mathscr{R}([X, Y])}=H^2$, i.e. condition (2) in Theorem 0 is not fulfilled. Hence we may assume that Θ is a non-zero function. Since the only bounded self-adjoint Toeplitz operator with non-trivial kernel is the zero operator, the conditions $X\neq 0$ and $Y\neq 0$ imply $\operatorname{Ker} X=$ $=\operatorname{Ker} Y=\{0\}$, i.e. condition (1) in Theorem 0.

We show that $(H^2\ominus\Theta H^2)\cap\mathscr{R}(X)=\{0\}$ iff φ_+ is outer. Since $Xf=$ $=\frac{1}{2}P_+(\varphi+\bar\varphi)f=P_+\Theta\varphi_+f$, we have

$$P_+\bar\Theta Xf=P_+\bar\Theta P_+\Theta\varphi_+f=P_+\bar\varphi_+f=T_{\varphi_+}^*f.$$

Therefore,

$$(H^2\ominus\Theta H^2)\cap\mathscr{R}(X)=\{Xf\colon Xf\perp\Theta H^2\}=\{Xf\colon P_+\bar\Theta Xf=0\}=$$
$$=X\operatorname{Ker} T_{\varphi_+}^*=X\big(H^2\ominus\overline{\mathscr{R}(T_{\varphi_+})}\big)=X(H^2\ominus\varphi_+^i H^2),$$

where φ_+^i is the inner part of φ_+. Since $\operatorname{Ker} X=\{0\}$, $(H^2\ominus\Theta H^2)\cap\mathscr{R}(X)=\{0\}$ if and only if φ_+ is outer. Similarly it follows that $(H^2\ominus\Theta H^2)\cap\mathscr{R}(Y)=\{0\}$ if and only if φ_- is outer. By Theorem 0, this completes the proof of Theorem 1.

Corollary 1. *If $\varphi\in H^\infty$ is S^*-cyclic, then* $\{(\operatorname{Re} T_\varphi)^{-1}, (\operatorname{Im} T_\varphi)^{-1}\}\notin\mathfrak{N}_1$.

Proof. Note that φ and $S^*\varphi$ are S^*-cyclic simultaneously. In this case $\Theta=0$ and $\varphi_+=\varphi_-=0$.

Lemma 2. *If* $\bigvee\{(S^*)^n\varphi\colon n\geqq 1\}=H^2\ominus\Theta H^2$, *then* $\Theta\bar{\varphi}$ *and* Θ *have no common inner divisor.*

Proof. Let ϑ be a common inner divisor of $\Theta\bar{\varphi}$ and Θ and let $\Theta'\overset{\text{def}}{=}\Theta\bar{\vartheta}$. Then $\Theta'\bar{\varphi}\in H^2$ and $((S^*)^n\varphi, \Theta'f)=(\bar{z}^n\bar{f}, \Theta'\bar{\varphi})=0$ for $n\geqq 1$ and $f\in H^2$. Therefore, $\Theta' H^2\subseteqq \subseteqq \Theta H^2$, i.e., $\bar{\vartheta}=\bar{\Theta}\Theta'\in H^2$ and ϑ is a constant function.

If Θ is a finite Blaschke product, then φ is meromorphic in $\bar{\mathbf{C}}$, the function $\bar{\bar{\varphi}}$ defined by $\bar{\bar{\varphi}}(z)=\overline{\varphi(1/\bar{z})}$ is meromorphic too, and $\varphi_\pm(z)=\frac{1}{2}\,\Theta(z)\big(\varphi(z)\pm\bar{\bar{\varphi}}(z)\big)$ for $|z|\leqq 1$.

Corollary 2. *Let Θ be a finite Blaschke product. Then,* $\{(\operatorname{Re} T_\varphi)^{-1}, (\operatorname{Im} T_\varphi)^{-1}\}\in\mathfrak{N}_1$ *if and only if* $\varphi^2(z)\neq\bar{\bar{\varphi}}^2(z)$ *for every* $z\in\mathbf{C}$, $|z|\neq 1$.

Proof. Suppose that $\varphi^2(z)=\bar{\bar{\varphi}}^2(z)$ for some $z\in\mathbf{C}$, $|z|\neq 1$. Since $\varphi^2(1/\bar{z})= =\overline{\bar{\bar{\varphi}}^2(z)}=\overline{\varphi^2(z)}=\bar{\bar{\varphi}}^2(1/\bar{z})$, we can assume without loss of generality that $|z|<1$. Hence $\varphi_+\varphi_-$ has a zero inside the unit circle. Therefore it is not outer.

Suppose now that φ_+ (or φ_-) is not outer. Then it has a zero, say z_0, inside the unit circle (see the remark just before Corollary 2). According to Lemma 2, $\Theta(z_0)\neq 0$ and therefore $\varphi(z_0)+\bar{\bar{\varphi}}(z_0)=0$ (or $\varphi(z_0)-\bar{\bar{\varphi}}(z_0)=0$, resp.), i.e. $\varphi^2(z_0)=\bar{\bar{\varphi}}^2(z_0)$.

2. In [2] the study of commuting unbounded self-adjoint operators was continued. The more general classes $\mathfrak{N}_{rs}$ are introduced in [2]. Here we only need the class $\mathfrak{N}_\infty^\infty$.

Definition 2. Let A, B be self-adjoint operators on a Hilbert space $\mathscr{H}$. The pair $\{A, B\}$ is in the class $\mathfrak{N}_\infty^\infty$ if there exists a dense linear manifold $\mathscr{D}$ in $\mathscr{H}$ such that

(i)′ $\mathscr{D}\subseteqq\operatorname{Dom}(A^jB^k)\cap\operatorname{Dom}(B^kA^j)$ and $A^jB^kf=B^kA^jf$ for all $f\in\mathscr{D}$ and all $j, k=0, 1, \dots;$

(ii)′ $A^k|\mathscr{D}$ and $B^k|\mathscr{D}$ are essentially self-adjoint for all $k\geqq 1$.

For polynomial symbols it was shown in [2, Theorem 4.1] that all pairs $\{(\operatorname{Re} T_\varphi)^{-1}, (\operatorname{Im} T_\varphi)^{-1}\}\in\mathfrak{N}_1$ are in fact in the class $\mathfrak{N}_\infty^\infty$. Using the same method as in [2] we prove this assertion for arbitrary analytic symbols.

Theorem 2. *For arbitrary* $\varphi\in H^\infty$ *the following are equivalent:*

$$\{(\operatorname{Re} T_\varphi)^{-1}, (\operatorname{Im} T_\varphi)^{-1}\}\in\mathfrak{N}_1, \tag{3}$$

$$\{\operatorname{Re} T_\varphi)^{-1}, (\operatorname{Im} T_\varphi)^{-1}\}\in\mathfrak{N}_\infty^\infty. \tag{4}$$

Lemma 3. $Q_{rs} \overset{\text{def}}{=} \vee\{\mathscr{R}(X^j Y^k[X, Y]): j<r,\ k<s\} = H^2 \ominus \Theta^{r+s-1} H^2$.

Proof. Since the subspace $H^2 \ominus \Theta H^2$ is invariant under the operator T_φ^*, it is sufficient to show that

$$\vee\{\mathscr{R}(T_\varphi^k[X, Y]):\ k < n\} = H^2 \ominus \Theta^n H^2.$$

We prove this assertion by induction. By Lemma 1 this is true in case $n=1$. Suppose that

$$\vee\{\mathscr{R}(T_\varphi^k[X, Y]):\ k < n\} = H^2 \ominus \Theta^n H^2 \overset{\text{def}}{=} K_n.$$

Then

$$\vee\{\mathscr{R}(T_\varphi^k[X, Y]):\ k < n+1\} = \vee\{T_\varphi K_n, K_n\}.$$

If $f \perp \vee\{T_\varphi K_n, K_n\}$, then $f=\Theta^n g$ and $P_+\bar\varphi f=\Theta^n h$, for some $g \in H^2$, $h \in H^2$. Hence $\Theta h=\bar\Theta^{n-1} P_+\bar\varphi \Theta^n g=\bar\Theta^{n-1}(\Theta\bar\varphi)\Theta^{n-1} g=(\Theta\bar\varphi)g$. According to Lemma 2, $\Theta\bar\varphi$ and Θ have no common inner divisor. Thus $g \in \Theta H^2$ and $f \in \Theta^{n+1} H^2$. Therefore, $K_{n+1} \subseteq \vee\{T_\varphi K_n, K_n\}$. On the other hand, $(\varphi K_n, \Theta^{n+1} H^2)=(K_n, \Theta^n(\Theta\bar\varphi) H^2)=0$. Hence $\vee\{T_\varphi K_n, K_n\}=K_{n+1}$ which completes the induction proof.

Proof of Theorem 2. Since (4)⇒(3) is obvious by definition we only have to prove the implication (3)⇒(4). Suppose that (3) is fulfilled. Then, by Theorem 1, φ_+ and φ_- are outer. To prove (4), we apply Corollary 1.9 in [2]. By this Corollary, it is sufficient to verify the following two conditions:

(x) $\quad Xf \in Q_{r+1,s} \Rightarrow f \in Q_{rs}$ for all $r \geqq 0$, $s \geqq 0$ and all $f \in H^2$,

(y) $\quad Yf \in Q_{r,s+1} \Rightarrow f \in Q_{r,s}$ for all $r \geqq 0$, $s \geqq 0$ and all $f \in H^2$.

Let $Xf \in Q_{r+1,s}=H^2 \ominus \Theta^n H^2$, $n=r+s$, i.e., $Xf=P_+\Theta\bar\varphi_+ f \perp \Theta^n H^2$. Hence $0=(P_+\Theta\bar\varphi_+ f, \Theta^n H^2)=(f, \varphi_+\Theta^{n-1} H^2)$. Therefore, since φ_+ is outer, $f \in H^2 \ominus \Theta^{n-1} H^2=Q_{rs}$. In a similar way we see that (y) is satisfied if φ_- is outer. This completes the proof.

References

[1] K. Schmüdgen, On commuting unbounded self-adjoint operators. I, *Acta Sci. Math.*, **47** (1984), 131—146.

[2] K. Schmüdgen, J. Friedrich, On commuting unbounded self-adjoint operators. II, *Integral Equations Operator Theory*, **7** (1984), 815—867.

FONTANKA 27, LOMI
191011, LENINGRAD, USSR

Acta Sci. Math., 51 (1987), 461—465

Non-atomic measure spaces and Fredholm composition operators

R. K. SINGH and T. VELUCHAMY

1. Introduction. Let $(X, \mathcal{S}, \lambda)$ be a sigma-finite measure space and let T be a measurable nonsingular ($\lambda T^{-1}(E)=0$ whenever $\lambda(E)=0$ for $E\in\mathcal{S}$) transformation from X into itself. Then the composition transformation C_T on $L^2(\lambda)$ is defined as $C_T f = f\circ T$ for every $f\in L^2(\lambda)$. If the range of C_T is in $L^2(\lambda)$ and C_T is bounded, then we call C_T the composition operator induced by T. It has been proved that a nonsingular measurable transformation T induces a composition operator C_T if and only if there exists a constant $M>0$ such that $\lambda T^{-1}(E)\leqq M\lambda(E)$ for every $E\in\mathcal{S}$. Hence the induced measure λT^{-1} is absolutely continuous with respect to the measure λ. Let f_0 denote the Radon—Nikodym derivative of the measure λT^{-1} with respect to λ.

The main purpose of this paper is to study Fredholm, essentially unitary and essentially normal composition operators on $L^2(\lambda)$ when the underlying measure space is non atomic. In case X is the unit interval of the real line and λ is the Lebesgue measure on the Borel subsets X it turns out that the composition operator C_T on $L^2(\lambda)$ is Fredholm if and only if C_T is invertible [2]. We prove here that the above result is true for a general non-atomic measure space. We also prove that the set of essentially unitary composition operators on $L^2(\lambda)$ coincides with the set of unitary composition operators on $L^2(\lambda)$ and the set of essentially isometric composition operators coincide with the set of isometric composition operators on $L^2(\lambda)$. It is also proved that when C_T has dense range, C_T is essentially normal if and only if C_T is normal. Note that a measure space $(X, \mathcal{S}, \lambda)$ is said to be non-atomic if for every nonnull $E\in\mathcal{S}$, there exists a nonnull $F\in\mathcal{S}$ such that $F\subset E$ and $\lambda(F)<\lambda(E)$.

Definitions. Let $B(H)$ denote the Banach algebra of all operators on a Hilbert space H and $C(H)$ denote the ideal of compact operators on H. Let π be the natural homomorphism from $B(H)$ to the Calkin algebra $B(H)/C(H)$. An

Received September 10, 1984.

operator $A\in B(H)$ is said to be Fredholm, essentially unitary, essentially normal, an essential isometry or an essential coisometry according as $\pi(A)$ is invertible, unitary, normal, an isometry or a coisometry, respectively. It has been proved that A is Fredholm if and only if A has closed range, and the kernel of A and the kernel of A^* are finite dimensional. A is called quasiunitary if A^*A-I and AA^*-I are finite rank operators [2].

2. Fredholm composition operators. If $C_T\in B(L^2(\lambda))$, then we know that $C_T^*C_T=M_{f_0}$ [3]. So, $\ker C_T=\ker C_T^*C_T=\ker M_{f_0}=L^2(X_0)$, where $X_0=\{x: f_0(x)=0\}$. The following theorem computes the kernel of C_T^* which is useful in proving the main theorem of this section.

Theorem 2.1. *Let* $C_T\in B(L^2(\lambda))$. *Then*

$$\ker C_T^* = \Big\{f\colon f\in L^2(\lambda) \text{ and } \int_{T^{-1}(E)} f\,d\lambda = 0 \text{ for all } E\in\mathcal{S}\Big\}.$$

Proof. Let $f\in L^2(\lambda)$. Then $f\in\ker C_T^*$ if and only if the inner product $\langle f, g\rangle=0$ for every $g\in(\ker C_T^*)^\perp=\overline{\operatorname{Ran} C_T}$. Since the span of the characteristic function $\{X_{T^{-1}(E)}\colon \lambda T^{-1}(E)<\infty \text{ and } E\in\mathcal{S}\}$ is dense in $\overline{\operatorname{Ran} C_T}$, we conclude that $f\in\ker C_T^*$ if and only if $\int_{T^{-1}(E)} f\,d\lambda=0$ for every $E\in\mathcal{S}$. Hence the proof is completed.

Definition. If $(X, \mathcal{S}, \lambda)$ is a measure space, then the sigma-algebra $T^{-1}(\mathcal{S})= =\{T^{-1}(E)\colon E\in\mathcal{S}\}$ is said to be essentially all of $\mathcal{S}$ if for every $E\in\mathcal{S}$ there exists $T^{-1}(F)\in T^{-1}(\mathcal{S})$ such that $\lambda(E\triangle T^{-1}(F))=\lambda\{(E\setminus T^{-1}(F))\cup(T^{-1}(F)\setminus E)\}=0$.

It has been proved by WHITLEY [6] and SINGH and KUMAR [4] that C_T has dense range if and only if $T^{-1}(\mathcal{S})$ is essentially all of $\mathcal{S}$. This we can conclude from the above theorem also.

Corollary 2.1. *No characteristic function belongs to* $\ker C_T^*$. *In fact, no positive function belongs to* $\ker C_T^*$.

Corollary 2.3. *If* $C_T\in B(L^2(\lambda))$, *then* $\ker C_T\subset\ker C_T^*$ *implies that* C_T *is an injection.*

Theorem 2.4. *Let* C_T *be a normal composition operator on* $L^2(\lambda)$. *Then* C_T *is Fredholm if and only if* C_T *is invertible.*

Proof. Since every normal composition operator on $L^2(\lambda)$ is an injection [4], the result follows.

The above theorem is not true in general as evident from the following example:

Example 2.5. Let l^2 denote the Hilbert space of all square summable sequences of complex numbers. Define the operator $A\colon l^2 \to l^2$ by

$$(Ax)(n) = \begin{cases} 0 & \text{if } n = 1 \\ x_n & \text{if } n > 1 \end{cases}$$

for $x = \{x_n\colon n \in \mathbf{N}\}$ in l^2. Then $A = A^*$ and hence A is normal. Also $\dim\ker A = \dim\ker A^* = 1$ and the range of A is closed. Hence A is a normal Fredholm operator. But clearly, A is not invertible.

From now on we assume that the measure space $(X, \mathcal{S}, \lambda)$ is non-atomic. The following theorem shows that the set of Fredholm composition operators and the set of invertible composition operators on $L^2(\lambda)$ coincide.

Theorem 2.6. *Let $C_T \in B(L^2(\lambda))$. Then C_T is Fredholm if and only if C_T is invertible.*

Proof. Suppose C_T is Fredholm. Then $\ker C_T$ and $\ker C_T^*$ are finite dimensional and C_T has closed range. But $\ker C_T = L^2(X_0)$, where $X_0 = \{x\colon f_0(x) = 0\}$ and λ is non-atomic implies that $\ker C_T$ is $\{0\}$. Hence to prove that C_T is invertible it is enough to prove that C_T has dense range. Suppose C_T does not have dense range. Then there exists a measurable set G in $\mathcal{S}$ such that G is not in $T^{-1}(\mathcal{S})$. We can find a measurable set E such that $T^{-1}(E) \supset G$. Let $T^{-1}(E) = G \cup F$. Then F is a nonnull measurable set and F does not belong to $T^{-1}(\mathcal{S})$. If we partition E into countable disjoint measurable sets, then at least one set among those partitions, say E^1, will be such that $T^{-1}(E^1)$ contains nonnull measurable subsets G^1 of G and F^1 of F where G^1 and F^1 are not in $T^{-1}(\mathcal{S})$ and $\lambda(E^1) < 1$. Again partition E^1. Then we get at least one E^2 such that $\lambda(E^2) < 1/2$ and $T^{-1}(E^2)$ containing nonnull parts of G and F which are not in $T^{-1}(\mathcal{S})$. Repeat this process. If at each stage of partition, there is exactly one measurable set E^n such that $T^{-1}(E^n)$ contains nonnull parts G^n of G and F^n of F such that G^n and F^n are not in $T^{-1}(E)$, then E^n can be made to approach a null set, since $\lambda(E^n) < 1/n$. This will imply that G and F are in $T^{-1}(\mathcal{S})$ which is a contradiction. Hence we can get a disjoint sequence $\{E_n\colon n \in \mathbf{N}\}$ of measurable subsets of E such that for every n, $T^{-1}(E_n) = G_n \cup F_n$ where $G_n \subset G$ and $F_n \subset F$ and G_n and F_n are not in $T^{-1}(\mathcal{S})$. Now consider the sequence $\{G_n\colon n \in \mathbf{N}\}$. This is a disjoint sequence and G_n does not belong to $T^{-1}(\mathcal{S})$ for every $n \in \mathbf{N}$. Also $G_1 \cup G_2 \cup \ldots \cup G_k$ does not belong to $T^{-1}(\mathcal{S})$ for every $k \in \mathbf{N}$. For, if $G_1 \cup G_2 \in T^{-1}(\mathcal{S})$, then there is a measurable set $K \subset E_1 \cup E_2$ such that $T^{-1}(K) = G_1 \cup G_2$. Hence $T^{-1}(K \cap E_1) = G_1$ which implies that $G_1 \in T^{-1}(\mathcal{S})$ which is a contradiction.

Now, since G_1 does not belong to $T^{-1}(\mathcal{S})$, there is a function f_1 in $L^2(\lambda)$ such that $\int_{T^{-1}(E)} f_1\, d\lambda = 0$ for every $E \in \mathcal{S}$ and $\int_{G_1} f_1\, d\lambda \neq 0$. For, if not, then $\ker C_T^* \subset (X_{G_1})^\perp$ and hence $(X_{G_1})^{\perp\perp} \subset \operatorname{Ran} C_T$. This implies that $X_{G_1} \in \operatorname{Ran} C_T$ which is a contradiction since G_1 does not belong to $T^{-1}(\mathscr{S})$. Again, there is a function f_2 in $L^2(\lambda)$ such that $\int_{T^{-1}(E)} f_2 d\lambda = 0$ for every $E \in \mathscr{S}$ and $\int_{G_1} f_2 d\lambda = 0$ but $\int_{G_2} f_2 d\lambda \neq 0$. For, if not, then $(\ker C_T^*) \cap (X_{G_1})^\perp \subset (X_{G_2})^\perp$. Hence $(X_{G_2})^{\perp\perp} \subset \operatorname{span}\{\operatorname{Ran} C_T, X_{G_1}\}$. This implies that $X_{G_2} = f + \alpha x_{G_1}$ for some f in $\operatorname{Ran} C_T$ and $\alpha \in \mathbf{C}$. Hence $f = X_{G_2} - \alpha X_{G_1}$ and this will imply that f is not measurable with respect to the sigma algebra $T^{-1}(\mathscr{S})$ which is a contradiction. Proceeding like this we will get a sequence $\{f_n \colon n \in \mathbf{N}\} \subset \ker C_T^*$ such that $\int_{G_k} f_n d\lambda$ is not equal to zero for $k = n$ and is zero for $k < n$. Hence, no two functions in $\{f_n \colon n \in \mathbf{N}\}$ are linearly dependent and hence $\dim \ker C_T^* = \infty$ which is a contradiction. Hence the theorem is proved.

3. Essentially unitary and essentially normal composition operators. First we shall characterise essentially isometric and essentially coisometric composition operators on $L^2(\lambda)$.

Theorem 3.1. *Let $C_T \in B(L^2(\lambda))$. Then C_T is an essential isometry if and only if C_T is an isometry.*

Proof. Let C_T be an essential isometry. Then $\pi(C_T)^* \pi(C_T) = \pi(I)$ which implies that $C_T^* C_T - I$ is compact. But $C_T^* C_T - I = M_{f_0} - I$ is compact on $L^2(\lambda)$ if and only if $f_0 = 1$ a.e. [5]. This implies that C_T is an isometry and hence the proof is completed.

Theorem 3.2. *Let $C_T \in B(L^2(\lambda))$. Then C_T is an essential coisometry if and only if it is a coisometry.*

Proof. Let C_T be an essential coisometry. Now, it is clear that C_T is an essential coisometry if and only if $C_T C_T^* - I$ is compact. But

$$C_T C_T^* - I = \begin{cases} M_{f_0 \circ T^{-1}} & \text{on } \overline{\operatorname{Ran} C_T} \\ -I & \text{on } \ker C_T^*. \end{cases}$$

Since $\overline{\operatorname{ran} C_T}$ and $\ker C_T^*$ are invariant under $C_T C_T^* - I$, $C_T C_T^* - I$ is compact if and only if $M_{f_0 \circ T^{-1}}$ is compact on $\overline{\operatorname{ran} C_T}$ and $-I$ is compact on $\ker C_T^*$. But $-I$ is compact on $\ker C_T^*$ if and only if $\ker C_T^*$ is finite dimensional which further implies that $\ker C_T^* = \{0\}$. Hence C_T has dense range and $M_{f_0 \circ T^{-1}}$ is compact on $\overline{\operatorname{ran} C_T} = L^2(\lambda)$ if and only if $f_0 \circ T = 1$ a.e. This implies that C_T is a coisometry and hence the theorem is proved.

Theorem 3.3. *Let $C_T \in B(L^2(\lambda))$. Then C_T is essentially unitary if and only if C_T is unitary.*

Proof. C_T is essentially unitary if and only if C_T is essentially an isometry and essentially a coisometry. Hence the theorem follows from Theorems 3.1 and 3.2.

Corollary 3.4. *Let* $C_T \in B(L^2(\lambda))$. *Then* C_T *is quasiunitary if and only if* C_T *is unitary.*

Theorem 3.5. *Let* C_T *be a composition operator on* $L^2(\lambda)$ *with dense range. Then* C_T *is essentially normal if and only if* C_T *is normal.*

Proof. C_T is essentially normal if and only if $C_T^* C_T - C_T C_T^*$ is compact. But when C_T has dense range, $C_T^* C_T - C_T C_T^* = M_{f_0 - f_0 \circ T}$ and hence $M_{f_0 - f_0 \circ T}$ is compact on $L^2(\lambda)$ implies that $f_0 = f_0 \circ T$ a.e. and this further implies that C_T is normal. Thus the theorem is proved.

References

[1] R. G. Douglas, *Banach Algebra Techniques in Operator Theory,* Academic Press (New York—London, 1972).

[2] Ashok Kumar, Fredholm composition operators, *Proc. Amer. Math. Soc.*, **79** (1980), 233—236.

[3] Eric A. Nordgren, Composition operators on Hilbert spaces, in: *Hilbert space operators* (Proc. Conf. Calif., Calif. State Univ., Long Beach, Calif., 1977), Lecture Notes in Mathematics, No. 693, Springer-Verlag (Berlin, 1978); pp. 37—63.

[4] R. K. Singh and Ashok Kumar, Characterization of invertible, unitary and normal composition operators, *Bull. Austral. Math. Soc.*, **19** (1978), 81—95.

[5] R. K. Singh and Ashok Kumar, Compact composition operators, *J. Austral. Math. Soc. Ser. A*, **28** (1979), 309—314.

[6] Robert Whitley, Normal and quasi-normal composition operators, *Proc. Amer. Math. Soc.*, **70** (1978), 114—118.

(R. K. S.)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF JAMMU
JAMMU — 180 001, INDIA

(T. V.)
DEPARTMENT OF MATHEMATICS
GOVT. ARTS COLLEGE
OOTACAMUND — 643 002, TAMIL NADU, INDIA.

Acta Sci. Math., 51 (1987), 467—474

Best approximation of a normal operator in the trace norm

RICHARD BOULDIN

1. Introduction. A problem that has received considerable attention is the classification of operators that have a unique best approximation among the nonnegative operators (a unique positive approximant) in one norm or another. For the operator norm this was done in [4] and, consequently, it solved a problem posed in [8]. Those results were generalized in [5], [9], [2] and other papers. The problem of approximation in trace norm was specifically excluded in [2], and it was noted how the methods given there failed in the case of the trace norm. This paper gives a characterization of those normal operators with a unique positive approximant in the trace norm. The result is a striking contrast to the characterizations given previously for other norms.

We are concerned throughout this paper with (bounded linear) operators on a separable Hilbert space $\mathfrak{H}$. For any operator T we use the associated operator $|T|=(T^*T)^{1/2}$ and the Caratesian decomposition $T=B+iC$ with $B=(1/2)(T+T^*)$ and $C=(1/2i)(T-T^*)$. We refer to B as $\operatorname{re} T$ and to C as $\operatorname{im} T$. For a compact operator T we let $s_1(T), s_2(T), \ldots$ denote the eigenvalues of $|T|$ in nonincreasing order repeated according to multiplicity. If we have

$$\sum_{j=1}^{\infty} s_j(T) < \infty$$

then we say that T is trace class and the preceding sum is the trace norm, denoted $\|T\|_1$. If T is not trace class then $\|T\|_1$ is defined to be infinity.

For a self-adjoint operator B we define B^+ to be $\frac{1}{2}(|B|+B)$ and B^- to be $\frac{1}{2}(|B|-B)$; we note that $B=B^+-B^-$ and $|B|=B^++B^-$. If $E(\cdot)$ is the spectral measure for B then it follows from the usual operational calculus that $B^+=BE([0,\infty))$ and $B^-=BE((-\infty,0])$. If T is a given operator and P is a nonnegative operator

Received October 17, 1984.

such that $\infty > \|T-P\|_1$ and $\|T-R\|_1 \geqq \|T-P\|_1$ for every nonnegative operator R then we say that P is a *trace class positive approximant of T*.

We shall frequently use the following inequality for the trace class operator T where $\{e_j\}$ is some orthonormal set:

$$\|T\|_1 \geqq \sum_j |\langle Te_j, e_j\rangle|.$$

This follows from the Corollary on p. 40 of [10].

2. Preliminary results. Of course, not all operators can be approximated by a nonnegative operator using the trace norm. The next theorem gives convenient conditions for recognizing when a given operator can be approximated.

Theorem 1. *For a given operator $T=B+iC$, $B^*=B$, $C^*=C$, the following conditions are equivalent:*

(i) *there exists a nonnegative operator P such that $T-P$ is trace class;*

(ii) *the operator C is trace class and the spectrum of B, denoted $\sigma(B)$, not in the interval $[0, \infty)$ consists of isolated eigenvalues, say $\{\lambda_j\}$ repeated according to multiplicity, such that $\sum_j |\lambda_j| < \infty$;*

(iii) *the operator $(T-B^+)$ is trace class.*

Proof. (i) *implies* (ii): Let D be the trace class operator $T-P$ and note that $B=P+\operatorname{re} D$, $C=\operatorname{im} D$. According to Weyl's Theorem B and P have the same Weyl spectrum. (See [1], for example.) For any normal operator A the Weyl spectrum coincides with the points of $\sigma(A)$ that are not isolated eigenvalues with finite multiplicity. (See [3, Theorem 3] or [1, Theorem 5.1].) It is elementary that $\operatorname{re} D$ and $\operatorname{im} D$ are trace class operators.

Let $\{\lambda_j\}$ be an enumeration of the negative eigenvalues of B, repeated according to multiplicity, and let $\{e_j\}$ be an orthonormal sequence of eigenvectors with e_j corresponding to λ_j. Note that $\|\operatorname{re} D\|_1 = \|P-B\|_1 \geqq \sum_j |\langle (P-B)e_j, e_j\rangle| = \sum_j (\langle Pe_j, e_j\rangle - \lambda_j) \geqq \sum_j -\lambda_j = \sum_j |\lambda_j|$.

(ii) *implies* (iii): Let $\{\lambda_j\}$ and $\{e_j\}$ have the same meaning as given in the first part. If D is defined by $D=\sum_j \langle\cdot, e_j\rangle \lambda_j e_j$ then $\|D\|_1 = \sum_j |\lambda_j|$. Note that $B=B^+ + D$, since $B^- = BE((-\infty, 0])$ where $E(\cdot)$ is the spectral measure for B. We note that $T-B^+ = D+iC$; which proves (iii).

(iii) *implies* (i): This is obvious.

Next we show that if an operator can be approximated in trace norm by a nonnegative operator then it has a trace class positive approximant.

Theorem 2. *If the operator T satisfies one of the conditions in Theorem 1 then T has a trace class positive approximant.*

Proof. Recall that the conjugate space for the Banach space of compact operators on the underlying Hilbert space $\mathfrak{H}$ is the space of trace class operators on $\mathfrak{H}$. (See [10, p. 48], for example.) Recall that any closed sphere in the conjugate space is compact in the weak star topology. (See [6, p. 424], for example.) Let R be a nonnegative operator such that $(T-R)$ is trace class and let $\mathscr{B}$ denote the set of operators

$$\{T-P\colon\ P \geqq 0,\ \|T-P\|_1 \leqq \|T-R\|_1\}.$$

In order to show that $\mathscr{B}$ is weak star compact it suffices to show that $\mathscr{B}$ is weak star closed.

Let $\{T-R_\alpha\colon\ \alpha\in A\}$ be a net from $\mathscr{B}$ that converges to $T-P$ in the weak star topology; thus, $\lim_\alpha \operatorname{tr}(T-R_\alpha)X=\operatorname{tr}(T-P)X$ for every compact operator X. It suffices to show that $(T-P)$ belongs to $\mathscr{B}$. Let positive ε and compact operator X be given. Note that

$$\begin{aligned}|\operatorname{tr}(T-P)X| &= |\operatorname{tr}[(T-P)-(T-R_\alpha)+(T-R_\alpha)]X| \leqq \\ &\leqq |\operatorname{tr}(T-P)X-\operatorname{tr}(T-R_\alpha)X|+|\operatorname{tr}(T-R_\alpha)X| < \varepsilon+\|T-R_\alpha\|_1\|X\| \leqq \\ &\leqq \varepsilon+\|T-R\|_1\|X\|\end{aligned}$$

provided $\alpha>\beta$ where β belongs to A and depends on ε and X. It follows from the preceding inequalities that

$$\|T-P\|_1 \leqq \varepsilon+\|T-R\|_1$$

for the arbitrarily chosen ε. Hence, $(T-P)$ belongs to $\mathscr{B}$ and, thus, $\mathscr{B}$ is weak star compact.

From elementary topology we know that any lower semicontinuous function defined on a compact set assumes its infimum. Thus, it suffices to show that $f(A)=\|A\|_1$ is lower semicontinuous on the space of trace class operators. Note that

$$\|A\|_1 = \sup\{|\operatorname{tr}(AX)|\colon\ X \text{ is a compact contraction}\}.$$

Since the supremum of any collection of lower semicontinuous function is lower semicontinuous, we conclude that $f(A)$ is lower semicontinuous on the compact set $\mathscr{B}$. This completes the proof.

Theorems 1 and 2 might lead the reader to conjecture that B^+ is always a trace class positive approximant for $T=B+iC$, $B^*=B$, $C^*=C$. Such a conjecture is false, as we demonstrate. Define T by

$$T = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}+i\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$

and note that $B^+=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$. It is routine to determine that the spectrum of $|T-B^+|$ is $\{((3+\sqrt{5})/2)^{1/2}, ((3-\sqrt{5})/2)^{1/2}\}$ and, consequently, $\|T-B^+\|_1=\operatorname{tr}|T-B^+|=\sqrt{5}$. Since $\|T\|_1=2$, we see that the zero operator is closer to T than B^+ is.

3. Main results. Note that T in the counterexample at the end of the preceding section is not a normal operator. If T is normal then there is a simple trace class positive approximant.

Theorem 3. *If $A=B+iC$, $B^*=B$, $C^*=C$, is a normal operator satisfying one of the conditions in Theorem 1 then B^+ is a trace class positive approximant for A.*

Proof. It is clear that $-B^-+iC=A-B^+$ is a normal operator, and by Theorem 1 it is trace class. Clearly B^- and C are commuting self-adjoint trace class operators and there is an orthonormal basis, say $\{e_j\}$, that diagonalizes both operators. Let z_j be the eigenvalue of $-B^-+iC$ corresponding to the eigenvector e_j for each j. Since $B^+=BE([0,\infty))$ where $E(\cdot)$ is the spectral measure for B, it is routine to see that $B^+e_j=0$ for every j. Thus, we have $\langle Ae_j, e_j\rangle = \langle(-B^-+iC)e_j, e_j\rangle = z_j$.

For any nonnegative operator R we note that

$$\|A-R\|_1 \geqq \sum_j |\langle (A-R)e_j, e_j\rangle| = \sum_j |\langle -Re_j e_j\rangle + z_j| \geqq \sum |z_j| = \|A-B^+\|_1.$$

The preceding inequality proves that B^+ is a trace class positive approximant of A.

It follows from the main theorem in [2] that B^+ is the unique positive approximant in the Schatten p-norm $\|\cdot\|_p$, with $p\geqq 2$, for the normal operator $A=B+iC$, $B^*=B$, $C^*=C$. The next lemma shows that no statement like the preceding is true when the norm used is $\|\cdot\|_1$.

Lemma 4. *Let α, β, γ and δ be positive numbers and define A by*

$$A=\begin{bmatrix}\alpha+i\gamma & 0\\ 0 & \beta-i\delta\end{bmatrix}.$$

Two trace class positive approximants of A are

$$(\operatorname{re} A)^+ = \begin{bmatrix}\alpha & 0\\ 0 & \beta\end{bmatrix} \quad \textit{and} \quad R=\begin{bmatrix}\alpha & \varepsilon\\ \varepsilon & \beta\end{bmatrix}$$

where ε is chosen to satisfy $\gamma\delta\geqq\varepsilon^2>0$ and $\alpha\beta\geqq\varepsilon^2$.

Proof. By Theorem 3 we know that $(\operatorname{re} A)^+$ is a trace class positive approximant; thus, the $\|\cdot\|_1$-distance between A and the nonnegative 2×2 matrices is

$$\|A-(\operatorname{re} A)^+\|_1 = \left\|\begin{bmatrix}i\gamma & 0\\ 0 & -i\delta\end{bmatrix}\right\|_1 = \gamma+\delta.$$

Thus, it suffices to show that $\|A-R\|_1\leqq\gamma+\delta$. Straightforward computations show that the spectrum of $|A-R|=[(A-R)^*(A-R)]^{1/2}$ is

$$\{2^{-1/2}(\delta^2+\gamma^2+2\varepsilon^2\pm((\delta^2+\gamma^2+2\varepsilon^2)^2-4[(\gamma^2+\varepsilon^2)(\delta^2+\varepsilon^2)-\varepsilon^2(\gamma+\delta)^2])^{1/2})^{1/2}\}.$$

It follows that

$$\|A-R\|_1^2=(\operatorname{tr}|A-R|)^2=\delta^2+\gamma^2+2\varepsilon^2+\sqrt{4[(\gamma^2+\varepsilon^2)(\delta^2+\varepsilon^2)-\varepsilon^2(\gamma+\delta)^2]}=$$
$$=\delta^2+\gamma^2+2\gamma\delta=(\delta+\gamma)^2.$$

This proves the lemma.

Theorem 5. *Let A be a normal operator satisfying one of the conditions in Theorem 1. If the eigenvalues of A include $z=\alpha+i\gamma$ and $w=\beta-i\delta$ with $\alpha,\gamma,\beta,\delta>0$ then A does not have a unique trace class positive approximant.*

Proof. Write A as an orthogonal direct sum $A_0\oplus A_1$ such that the spectrum of A_0 is the set $\{z, w\}$. Clearly the direct sum of trace class positive approximants of A_0 and A_1, respectively, is a trace class positive approximant for A. It follows from Lemma 4 that we can construct multiple approximants for A_0 and, hence, for A.

Before we are done we shall prove the converse of the preceding theorem. First, we must accumulate some appropriate basic results. The next lemma gives another circumstance in which $(\operatorname{re} T)^+$ is a trace class positive approximant of T.

Lemma 6. *Let $T=B+iC$, $B^*=B$, $C^*=C$, be an operator satisfying one of the conditions in Theorem 1. If $B\geqq 0$ then B is a trace class positive approximant for T.*

Proof. Let $\{e_j\}$ be an orthonormal basis of eigenvectors for C and let λ_j be the eigenvalue corresponding to e_j for each j. If R is any nonnegative operator then we have

$$\|T-R\|_1\geqq\sum_j|\langle(T-R)e_j,e_j\rangle|=\sum_j[(\langle(B-R)e_j,e_j\rangle)^2+\lambda_j^2]^{1/2}\geqq$$
$$\geqq\sum_j|\lambda_j|=\|C\|_1=\|T-B\|_1.$$

This proves the lemma.

By strengthening the hypothesis of the preceding lemma we get a uniqueness result.

Theorem 7. *Let $T=B+iC$, $B^*=B$, $C^*=C$, be an operator satisfying one of the conditions of Theorem 1. If $B\geqq 0$ and $C\geqq 0$ then B is the unique trace class positive approximant for T.*

Proof. Choose $\{e_j\}$ and λ_j as in the proof of Lemma 6, and note that B is a trace class positive approximant of T according to that lemma. For R any trace class positive approximant of T we have

$$\|T-R\|_1\geqq\sum_j|\langle(T-R)e_j,e_j\rangle|\geqq\Big|\sum_j\langle(T-R)e_j,e_j\rangle\Big|=$$
$$=\Big|\sum_j\langle(B-R)e_j,e_j\rangle+i\sum_j\lambda_j\Big|\geqq\sum_j\lambda_j=\|C\|_1=\|T-R\|_1.$$

Since equality must hold throughout the preceding inequalities, we have

$$\|T-R\|_1 = \Big|\sum_j \langle (T-R)e_j, e_j\rangle\Big| = |\operatorname{tr}(T-R)|.$$

By the last part of Theorem 8.6 of [7, pp. 104—105], we conclude that $e^{-i\theta}(T-R)$ is a nonnegative operator for $\theta = \arg \operatorname{tr}(T-R)$. The equality of the third and fourth lines in the earlier inequalities shows that $\operatorname{tr}(T-R) = i\sum_j \lambda_j$. Thus, we know that $-i(T-R) = -i(B-R)+C$ is a nonnegative operator. This implies that $B-R=0$, which is the desired conclusion.

The next theorem gives another situation where $(\operatorname{re} T)^+$ is the unique trace class positive approximant for T.

Theorem 8. *Let $A=B+iC$, $B^*=B$, $C^*=C$, be a normal operator satisfying one of the conditions in Theorem 1. If $B \leqq 0$ then the zero operator 0 is the unique trace class positive approximant of A.*

Proof. Let $\{e_j\}$ be an orthonormal basis consisting of eigenvectors of A and let z_j be the eigenvalue corresponding to e_j for each j. Note that $\operatorname{re} z_j \leqq 0$ for each j. If R is any nonnegative operator then we have

$$\begin{aligned}\|A-R\| &\geqq \sum_j |\langle (A-R)e_j, e_j\rangle| = \sum_j |\langle z_j - (Re_j, e_j\rangle| = \\ &= \sum_j [(\langle Re_j, e_j\rangle - \operatorname{re} z_j)^2 + (\operatorname{im} z_j)^2]^{1/2} \geqq \\ &\geqq \sum_j [(\operatorname{re} z_j)^2 + (\operatorname{im} z_j)^2]^{1/2} = \sum_j |z_j| = \|A\|_1.\end{aligned}$$

This proves that 0 is a trace class positive approximant of A.

Furthermore, if R is any trace class positive approximant of A then equality holds in each of the preceding inequalities. It follows that $\langle Re_j, e_j\rangle = 0$ for each j and hence, R must be 0. The uniqueness is proved.

Before we can exploit Theorems 7 and 8 we need an elementary observation about matrices of operators.

Lemma 9. *If $R = \begin{pmatrix} 0 & D \\ D^* & B \end{pmatrix}$ is a nonnegative operator on $\mathfrak{H}_0 \oplus \mathfrak{H}_1$ then $B \geqq 0$ and $D=0$.*

Proof. Assume that there exists some f in $\mathfrak{H}_1$ such that $Df \neq 0$, and define e by $e = (-\gamma/\|Df\|^2)Df$ where γ is an arbitrary positive number. Note that

$$\left\langle \begin{bmatrix} 0 & D \\ D^* & B \end{bmatrix} \begin{bmatrix} e \\ f \end{bmatrix}, \begin{bmatrix} e \\ f \end{bmatrix} \right\rangle = -2\gamma + \langle Bf, f\rangle.$$

This contradicts the nonnegativity of R and, thus, it shows that $D=0$.

Since

$$\left\langle R\begin{bmatrix}0\\f\end{bmatrix}, \begin{bmatrix}0\\f\end{bmatrix}\right\rangle \geqq 0$$

for any f in $\mathfrak{H}_1$, it is clear that $B \geqq 0$.

Using the results of 7, 8 and 9 we can prove a partial converse for Theorem 5.

Theorem 10. *Let $A=B+iC$, $B^*=B$, $C^*=C$, be a normal operator satisfying one of the conditions in Theorem* 1. *If the spectrum of A, denoted $\sigma(A)$, is contained in $\{z$: either $\operatorname{im} z \geqq 0$ or $\operatorname{re} z \leqq 0\}$ then B^+ is the unique trace class positive approximant for A.*

Proof. Let $E(\cdot)$ be the spectral measure for A and define E_0, E_1, A_0 and A_1 by $E_0=E(\{z\colon \operatorname{re} z \leqq 0\})$, $E_1=E(\{z\colon \operatorname{re} z \geqq 0,\ \operatorname{im} z \geqq 0\})$, $A_0=AE_0$, $A_1=AE_1$. The hypothesis concerning $\sigma(A)$ shows that $A=A_0 \oplus A_1$. According to Theorem 8, 0 is the unique trace class positive approximant of A_0; according to Theorem 7, the unique trace class positive approximant of A_1 is $(\operatorname{re} A_1)$. It suffices to show that $B^+=0 \oplus \operatorname{re} A_1$ is the unique trace class positive approximant for A.

We use Theorem 8.7 of [7, pp. 105—106] in the first inequality below. If R is a nonnegative operator then we have

$$\begin{aligned} \|A-R\|_1 &\geqq \|E_0(A-R)E_0\|_1+\|E_1(A-R)E_1\|_1 = \\ &= \|A_0-E_0RE_0\|_1+\|A_1-E_1RE_1\|_1 \geqq \\ &\geqq \|A_0\|_1+\|A_1-\operatorname{re} A_1\|_1 = \|A-B^+\|_1. \end{aligned}$$

The preceding computation shows that B^+ is a trace class positive approximant for A. Furthermore, if R is any trace class positive approximant for A then $E_0RE_0=0$, $E_1RE_1=\operatorname{re} A_1$ by the uniqueness of the approximants of A_0 and A_1. It now follows from Lemma 9 that $R=0 \oplus \operatorname{re} A_1=B^+$, which proves the theorem.

Using Theorems 5 and 10 we characterize the normal operators that have a unique trace class positive approximant.

Theorem 11. *Let $A=B+iC$, $B^*=B$, $C^*=C$, be a normal operator that satisfies one of the conditions in Theorem* 1. *There is a unique trace class positive approximant for A if and only if $\sigma(A)$ is contained in one or the other of the two sets $\{z$: either $\operatorname{im} z \geqq 0$ or $\operatorname{re} z \leqq 0\}$, $\{z$: either $\operatorname{im} z \leqq 0$ or $\operatorname{re} z \leqq 0\}$.*

Proof. If $\sigma(A)$ is contained in the first set then it is immediate from Theorem 10 that B^+ is the unique trace class positive approximant of A. If $\sigma(A)$ is contained in the second set then $\sigma(A^*)$ is contained in the first set and B^+ is the unique trace class positive approximant of A^*. For any nonnegative operator R we have $\|A^*-R\|_1=\|A-R\|_1$ (by Lemma 8 of [10, p. 39], for example). It follows that B^+ is the unique trace class positive approximant for A.

If A has a unique trace class positive approximant then Theorem 5 shows that A does not have eigenvalues in each of the sets $\{z\colon \operatorname{im} z<0\}$ and $\{z\colon \operatorname{im} z>0\}$. Thus, the eigenvalues of A are contained in one or the other of the two sets $\{z\colon \text{either } \operatorname{im} z\geqq 0 \text{ or } \operatorname{re} z\leqq 0\}$, $\{z\colon \text{either } \operatorname{im} z\leqq 0 \text{ or } \operatorname{re} z\leqq 0\}$. According to Theorem 1, $A-B^+$ is trace class and so A is a compact perturbation of B^+, that is $A=B^++(A-B^+)$. By Weyl's theorem A and B^+ have the same Weyl spectrum. For each of these normal operators the Weyl spectrum consists of the points that are not isolated eigenvalues with finite multiplicity. Clearly the Weyl spectrum of B^+ (and, hence, the Weyl spectrum of A) is contained in the interval $[0, \infty)$. Since both the Weyl spectrum of A and the eigenvalues of A are contained in one of the desired sets, we conclude that $\sigma(A)$ is contained in one or the other of the sets indicated in the statement of the theorem.

References

[1] S. K. Berberian, The Weyl spectrum of an operator, *Indiana Univ. Math. J.*, **20** (1970), 529—544.

[2] R. H. Bouldin, Best approximation of a normal operator in the Schatten p-norm, *Proc. Amer. Math. Soc.*, **80** (1980), 277—282.

[3] R. H. Bouldin, Essential spectrum for a Hilbert space operator, *Trans. Amer. Math. Soc.*, **163** (1972), 437—445.

[4] R. H. Bouldin, Positive approximants, *Trans. Amer. Math. Soc.*, **177** (1973), 391—403.

[5] C. K. Chui, P. W. Smith and J. D. Ward, Approximation with restricted spectra, *Math. Japan*, **144** (1975), 289—297.

[6] N. Dunford and J. T. Schwartz, *Linear Operators. II*, Interscience (New York, 1963).

[7] I. C. Gohberg and M. G. Krein, *Introduction to the Theory of Linear Nonselfadjoint Operators*, Transl. Math. Monographs, vol. 18, Amer. Math. Soc. (Providence, RI., 1969).

[8] P. R. Halmos, Positive approximants of operators, *Indiana Univ. Math. J.*, **21** (1971/72), 951—960.

[9] P. R. Halmos, Spectral approximants of normal operators, *Proc. Edinburgh Math Soc.*, **19** (1974), 51—58.

[10] R Schatten, *Norm Ideals of Completely Continuous Operators*, Springer-Verlag (New York/Berlin, 1960).

DEPARTMENT OF MATHEMATICS
THE UNIVERSITY OF GEORGIA
ATHENS, GEORGIA 30602, U.S.A.

Acta Sci. Math., 51 (1987), 475—489

Contractions weakly similar to unitaries. II

LÁSZLÓ KÉRCHY

In this paper we continue the study of contractions, weakly similar to unitaries, begun in [9]. Here we consider the case, when the characteristic function is not isometric a.e. on the unit circle, and prove the reflexivity of such contractions under a general assumption. Our paper is organized as follows. After giving the necessary definitions and notations in Section 0, we introduce the notion of weak similarity in Section 1. Our main result is proved in Section 2, while in Section 3 we make some concluding remarks. The theory of contractions, elaborated by B. SZ.-NAGY and C. FOIAŞ will be applied, the main reference is their monograph [12].

0. Definitions and notations

If $\mathfrak{H}$ is a (complex, separable) Hilbert space, then $\mathscr{L}(\mathfrak{H})$ denotes the set of all (bounded, linear) operators acting on $\mathfrak{H}$. For an arbitrary subset $\mathscr{A}\subset\mathscr{L}(\mathfrak{H})$, Lat $\mathscr{A}$ stands for the lattice of invariant subspaces of $\mathscr{A}$, while for an arbitrary set S of (closed) subspaces of $\mathfrak{H}$, $\operatorname{Alg} S$ is the algebra of operators which leave invariant each element of S. A subalgebra $\mathscr{A}\subset\mathscr{L}(\mathfrak{H})$ is called reflexive, if $\operatorname{Alg}\operatorname{Lat}\mathscr{A}=\mathscr{A}$ (cf. [5]).

For an operator $T\in\mathscr{L}(\mathfrak{H})$, $\operatorname{Alg} T$ denotes the weakly closed algebra generated by T and the identity. It is clear that $\operatorname{Lat} T=\operatorname{Lat}\operatorname{Alg} T$. T is called reflexive, if $\operatorname{Alg} T$ is reflexive, i.e. $\operatorname{Alg}\operatorname{Lat} T=\operatorname{Alg} T$. $\{T\}'$ and $\{T\}''$ denote the commutant and bicommutant of T, respectively, and $\operatorname{Lat}'' T:=\operatorname{Lat}\{T\}''$, $\operatorname{Hyplat} T:=\operatorname{Lat}\{T\}'$. If T is a completely non-unitary (c.n.u.) contraction, then

$$H^\infty(T):=\{w(T):\ w\in H^\infty\},$$

where H^∞ denotes the Hardy class of bounded analytic functions, and the Sz.-Nagy, Foiaş functional calculus is applied for T.

Received November 19, 1984.

The contraction $T \in \mathscr{L}(\mathfrak{H})$ belongs to the class C_{11} or C_{10}, if for every non-zero vector $h \in \mathfrak{H}$ we have

$$\lim_{n\to\infty} \|T^n h\| \neq 0 \neq \lim_{n\to\infty} \|T^{*n} h\|,$$

or

$$\lim_{n\to\infty} \|T^n h\| \neq 0 = \lim_{n\to\infty} \|T^{*n} h\|,$$

respectively. If T is a C_{11}-contraction, then

$$\mathrm{Lat}_1 T := \{\mathfrak{M} \in \mathrm{Lat}\, T \colon T|\mathfrak{M} \in C_{11}\}$$

is a lattice under set-inclusion as partial ordering, in which the greatest lower bound "$\overset{(1)}{\cap}$" is generally different from the intersection "$\cap$". $\mathrm{Hyplat}_1 T := \mathrm{Lat}_1 T \cap \cap \mathrm{Hyplat}\, T$ is a sublattice of $\mathrm{Lat}_1 T$ (cf. [8]).

$\mathbf{D}$ will denote the open unit disc of the complex plane, C its boundary, and m the normalized Lebesgue measure on C. For a contraction $T \in \mathscr{L}(\mathfrak{H})$, $\mathfrak{D}_T := ((I-T^*T)\mathfrak{H})^-$ and $\mathfrak{D}_{T^*} := ((I-TT^*)\mathfrak{H})^-$ denote its defect spaces, and $\{\Theta_T(\lambda), \mathfrak{D}_T, \mathfrak{D}_{T^*}\}$ its characteristic function in the sense of Sz.-Nagy and Foiaş, i.e. λ alters in $\mathbf{D}$ and $\Theta_T(\lambda) \in \mathscr{L}(\mathfrak{D}_T, \mathfrak{D}_{T^*})$ is defined by

$$\Theta_T(\lambda) = [-T + \lambda(I-TT^*)^{1/2}(I-\lambda T^*)^{-1}(I-T^*T)^{1/2}]\,|\,\mathfrak{D}_T.$$

Moreover, Δ_T stands for the operator-valued function defined on C by the formula

$$\Delta_T(e^{it}) = [I - \Theta_T(e^{it})^* \Theta_T(e^{it})]^{1/2}.$$

(Θ_T has radial limit a.e. on C.)

If $T \in \mathscr{L}(\mathfrak{H})$, $S \in \mathscr{L}(\mathfrak{K})$, then $\mathscr{I}(T, S)$ denotes the set of intertwining operators:

$$\mathscr{I}(T, S) = \{X \in \mathscr{L}(\mathfrak{H}, \mathfrak{K}) \colon XT = SX\}.$$

We say that T can be injected into S, and write $T \overset{i}{\prec} S$, if $\mathscr{I}(T, S)$ contains an injection; T is a quasi-affine transform of S, if $\mathscr{I}(T, S)$ contains a quasi-affinity, i.e. an injection with dense range; and T, S are quasi-similar, if they are quasi-affine transforms of each other.

A system $\{\mathfrak{H}_n\}_n$ of subspaces of $\mathfrak{H}$ is called to be basic, if $\mathfrak{H}_n \dot{+} (\bigvee_{k\neq n} \mathfrak{H}_k) = \mathfrak{H}$, for every n, and $\bigcap_n (\bigvee_{k\geq n} \mathfrak{H}_k) = \{0\}$ (cf. [1]).

1. Weak similarity

We begin by introducing the notion of contractions, weakly similar to unitaries, in a bit more general setting than in [9]. Namely, we give the following

Definition 1. The operators $T\in\mathcal{L}(\mathfrak{H})$ and $S\in\mathcal{L}(\mathfrak{K})$ are called *weakly similar,* if there exist basic systems $\{\mathfrak{H}_n\}_n$ and $\{\mathfrak{K}_n\}_n$ in $\mathfrak{H}$ and $\mathfrak{K}$, respectively, such that $\mathfrak{H}_n\in\operatorname{Hyplat} T$, $\mathfrak{K}_n\in\operatorname{Hyplat} S$, and $T|\mathfrak{H}_n$ is similar to $S|\mathfrak{K}_n$, for every n.

$T\in\mathcal{L}(\mathfrak{H})$ is *weakly similar to unitary,* if T is weakly similar to a unitary operator.

Remark 2. Weak similarity is clearly a weaker relation than similarity, but stronger than quasi-similarity. In fact, let P_n and Q_n denote the projections onto the subspaces $\mathfrak{H}_n$ and $\mathfrak{K}_n$ with respect to the decompositions $\mathfrak{H}=\mathfrak{H}_n\dot{+}(\bigvee_{k\neq n}\mathfrak{H}_k)$ and $\mathfrak{K}=\mathfrak{K}_n\dot{+}(\bigvee_{k\neq n}\mathfrak{K}_k)$, respectively. Now, choosing intertwining affinities $A_n\in\mathcal{I}(T|\mathfrak{H}_n, S|\mathfrak{K}_n)$, for every n, and sequences $\{\alpha_n\}_n$ and $\{\beta_n\}_n$ of positive numbers such that

$$\sum_n \alpha_n\|A_n\|\,\|P_n\|<\infty \quad\text{and}\quad \sum_n \beta_n\|A_n^{-1}\|\,\|Q_n\|<\infty,$$

we can define intertwining quasi-affinities $X\in\mathcal{I}(T,S)$ and $Y\in\mathcal{I}(S,T)$ by the equations

$$Xf=\sum_n \alpha_n A_n P_n f \quad (f\in\mathfrak{H}) \quad\text{and}\quad Yg=\sum_n \beta_n A_n^{-1}Q_n g \quad (g\in\mathfrak{K}).$$

The operator occurring in [9, Proposition 2] provides an example for a C_{11}-contraction which is not weakly similar to unitary. Since every C_{11}-contraction is quasi-similar to a unitary operator, we obtain that weak similarity is an actually stronger relation than quasi-similarity in the class of C_{11}-contractions. (Quasi-similarity was characterized in the class C_{11} in terms of decomposibility by C. Apostol [1].)

Remark 3. In [9] a contraction $T\in\mathcal{L}(\mathfrak{H})$ is called weakly similar to unitary, if there exists a basic system $\{\mathfrak{H}_n\}_n$ consisting of hyperinvariant subspaces of T such that $T|\mathfrak{H}_n$ is similar to a unitary operator $U_n\in\mathcal{L}(\mathfrak{K}_n)$, for every n. However, we can define a unitary operator U acting on the space $\mathfrak{K}=\bigoplus_n\mathfrak{K}_n$ as the orthogonal sum $U=\bigoplus_n U_n$. Constructing an intertwining quasi-affinity $X\in\mathcal{I}(T,U)$ as in the preceding remark, an application of [8, Proposition 6] shows that $\mathfrak{K}_n=(X\mathfrak{H}_n)^-\in\operatorname{Hyplat} U$, for every n. Therefore, T is weakly similar to U, i.e. T is weakly similar to unitary in the sense of our present definition too. Hence the two definitions coincide.

We recall that by [9, Theorem 4] a contraction T is weakly similar to unitary if and only if T is of class C_{11} and its characteristic function Θ_T is (boundedly) invertible a.e. on the unit circle C.

We finish the discussion of weak similarity by the following

Proposition 4. *Weak similarity is an equivalence relation in the class of C_{11}-contractions.*

Proof. We have to verify only transitivity. So let us assume that $T\in\mathcal{L}(\mathfrak{H})$, $S\in\mathcal{L}(\mathfrak{K})$ and $R\in\mathcal{L}(\mathfrak{L})$ are C_{11}-contractions such that T is weakly similar to S and S is weakly similar to R. Then, there exist basic systems $\{\mathfrak{H}_n\}_n$ and $\{\mathfrak{K}_n\}_n$ consisting of hyperinvariant subspaces of T and S, respectively, such that $\mathcal{I}(T|\mathfrak{H}_n, S|\mathfrak{K}_n)$ contains an affinity A_n, for every n. Similarly, we can find basic systems $\{\mathfrak{K}'_n\}_n$ and $\{\mathfrak{L}_n\}_n$ formed by hyperinvariant subspaces of S and R, respectively, such that $\mathcal{I}(S|\mathfrak{K}'_n, R|\mathfrak{L}_n)$ contains an affinity B_n, for every n. Since each of the above subspaces is C_{11}-invariant, and the C_{11}-hyperinvariant subspace lattice of any C_{11}-contraction is countably distributive (cf. [8, Proposition 2]), we can easily verify that the subspaces

$$\mathfrak{K}''_n = \bigvee_{i=1}^{n} (\mathfrak{K}_i \overset{(1)}{\bigcap} \mathfrak{K}'_{n+1-i})\in \mathrm{Hyplat}_1 S, \quad n = 1, 2, \ldots,$$

form a basic system in $\mathfrak{K}$. (Cf. also [9, Lemma 7].)

It follows immediately that the system

$$\{\mathfrak{H}'_n = \bigvee_{i=1}^{n} A_i^{-1}(\mathfrak{K}_i \overset{(1)}{\bigcap} \mathfrak{K}'_{n+1-i})\}_n$$

will be basic in $\mathfrak{H}$. Taking into account that the commutant $\{S\}'$ splits into the direct sum $\{S\}'=\{S|\mathfrak{K}_i\}'+\{S|\bigvee_{j\neq i}\mathfrak{K}_j\}'$ we infer that $\mathfrak{K}_i\overset{(1)}{\bigcap}\mathfrak{K}'_{n+1-i}\in \mathrm{Hyplat}_1(S|\mathfrak{K}_i)$. This implies that $A_i^{-1}(\mathfrak{K}_i\overset{(1)}{\bigcap}\mathfrak{K}'_{n+1-i})\in \mathrm{Hyplat}_1(T|\mathfrak{H}_i)$, and in virtue of the splitting $\{T\}'= =\{T|\mathfrak{H}_i\}'+\{T|\bigvee_{j\neq i}\mathfrak{H}_j\}'$ we conclude $A_i^{-1}(\mathfrak{K}_i\overset{(1)}{\bigcap}\mathfrak{K}'_{n+1-i})\in \mathrm{Hyplat}_1 T$. Since this holds for every $1\leq i\leq n$, we obtain that $\mathfrak{H}'_n\in \mathrm{Hyplat}_1 T$, for every n.

Similarly, we can prove that the subspaces $\mathfrak{L}'_n=\bigvee_{i=1}^{n} B_i(\mathfrak{K}_i\overset{(1)}{\bigcap}\mathfrak{K}'_{n+1-i})\in \mathrm{Hyplat}_1 R$, $n=1, 2, \ldots$, form a basic system in $\mathfrak{L}$. Since $T|\mathfrak{H}'_n$ is obviously similar to $R|\mathfrak{L}'_n$, for every n, we get that T and R are weakly similar.

2. Reflexivity of contractions, weakly similar to unitaries

Our main result is the following

Theorem 5. *Let T be a c.n.u. contraction which is weakly similar to unitary. If there exists a function $f\in(\Delta_T L^2(\mathfrak{D}_T))^-$ such that*

$$\int_C \log \|\Theta_T(e^{it})f(e^{it})\|\, dm(t) > -\infty, \tag{1}$$

then

(i) $H^\infty(T)=\text{Alg}\,T\neq\{T\}''$,

(ii) $\text{Lat}\,T\neq\text{Lat}_1\,T$, *and*

(iii) *T is reflexive.*

This theorem is a generalization of Wu's results (cf. [15, Theorem 3] and [16, Theorem 3.8]), who considered c.n.u. C_{11}-contractions with finite defect indeces, and is a counterpart of [9, Theorem 9], which is connected with contractions whose characteristic function is isometric on a subset of positive measure of the unit circle.

The assumption $\int_C \log\|\Theta_T(e^{it})f(e^{it})\|\,dm(t)>-\infty$ $(f\in(\Delta_T L^2(\mathfrak{D}_T))^-)$, occurring in our theorem, implies that $f(e^{it})\neq 0$ a.e. on C. Hence $\text{rank}\,\Delta_T(e^{it})\geqq 1$, i.e. $\Theta_T(e^{it})$ is not isometric a.e. on C.

Conversely, let us assume that, for the c.n.u. C_{11}-contraction T, $\Theta_T(e^{it})$ is not isometric a.e. on C. It follows that $\text{rank}\,\Delta_T(e^{it})\geqq 1$ a.e., and so the operator R of multiplication by e^{it} on the space $(\Delta_T L^2(\mathfrak{D}_T))^-$ is unitarily equivalent to an operator of the form $\bigoplus_n M_{\alpha_n}$, where $C=\alpha_1\supset\alpha_2\supset\dots$ are Borel subsets of C and M_{α_n} denotes the multiplication operator by e^{it} on the space $L^2(\alpha_n, m)$. (Cf. [7, Lemma 1].) This implies that we can find a vector $f\in(\Delta_T L^2(\mathfrak{D}_T))^-$ such that $R|\bigvee_{n\geqq 0} R^n f\in C_{10}$. Then we infer by Lemma 9 to be proved later that

$$\int_C \log\|f(e^{it})\|\,dm(t)>-\infty.$$

Let us assume in addition that Θ_T has a scalar multiple. On account of [12, Proposition V.7.1] this happens exactly when

$$\int_C \log\|\Theta_T(e^{it})^{-1}\|\,dm(t)<\infty.$$

Hence, taking into account that

$$\|\Theta_T(e^{it})f(e^{it})\|\geqq\|\Theta_T(e^{it})^{-1}\|^{-1}\|f(e^{it})\|,$$

we obtain

$$\int_C \log\|\Theta_T(e^{it})f(e^{it})\|\,dm(t)>-\infty.$$

Since in virtue of [9, Remark 5] T is in particular weakly similar to unitary, the assumptions of our theorem are fulfilled.

Therefore, taking also into consideration [9, Theorem 9 and Corollary 12] and that the question of reflexivity can be reduced to the case of c.n.u. contractions (cf. the proof of [2, Theorem 5]), we obtain the following

Corollary 6. *If T is a C_{11}-contraction whose characteristic function Θ_T has a scalar multiple, then T is reflexive. If we assume in addition that T is c.n.u. and*

$\Theta_T(e^{it})$ is not isometric a.e. on C, then

$$H^\infty(T) = \operatorname{Alg} T \neq \{T\}'' \quad \text{and} \quad \operatorname{Lat} T \neq \operatorname{Lat}_1 T,$$

while if $\Theta_T(e^{it})$ is isometric on a set of positive measure, then

$$H^\infty(T) \neq \operatorname{Alg} T = \{T\}'' \quad \text{and} \quad \operatorname{Lat} T = \operatorname{Lat}_1 T.$$

The proof of Theorem 5 follows the general outline of Wu's proof in [15]. The framework is the functional model of c.n.u. contractions. So we are starting by recalling some basic facts on model-operators. Since we are interested only in C_{11}-contractions we may restrict attention to contractive analytic functions whose values are operators acting in one Hilbert space.

So let us given a purely contractive analytic function $\{\Theta(\lambda), \mathfrak{E}, \mathfrak{E}\}$, where $\mathfrak{E}$ is a separable Hilbert space and $\Theta(\lambda) \in \mathscr{L}(\mathfrak{E})$ for every $\lambda \in \mathbf{D}$. The model-operator associated with Θ is defined in the following way. Let Δ denote the measurable operator-valued function defined by $\Delta(e^{it}) = [I - \Theta(e^{it})^* \Theta(e^{it})]^{1/2}$, and let us consider the Hilbert space

$$\mathfrak{K}_+ = H^2(\mathfrak{E}) \oplus (\Delta L^2(\mathfrak{E}))^-$$

of vector-valued functions. The operator $V \in \mathscr{L}(H^2(\mathfrak{E}), \mathfrak{K}_+)$, $Vw = \Theta w \oplus \Delta w$ $(w \in H^2(\mathfrak{E}))$ will be an isometry, and the subspace $VH^2(\mathfrak{E})$ of $\mathfrak{K}_+$ will be invariant under the operator U_+ of multiplication by e^{it} in $\mathfrak{K}_+$. Then the model-space is by definition

$$\mathfrak{H} = \mathfrak{K}_+ \ominus VH^2(\mathfrak{E}),$$

and the model-operator $T = S(\Theta)$ is the compression of U_+ onto $\mathfrak{H}$:

$$T = P_{\mathfrak{H}} U_+ | \mathfrak{H},$$

where $P_{\mathfrak{H}}$ denotes the orthogonal projection of $\mathfrak{K}_+$ onto $\mathfrak{H}$.

U_+ will be the minimal isometric dilation of T. The subspace $\mathfrak{R} = (\Delta L^2(\mathfrak{E}))^-$ reduces U_+ to a unitary operator $R = U_+ | \mathfrak{R}$, called the residual part of T. Since $VH^2(\mathfrak{E}) \in \operatorname{Lat} U_+$, it follows that $P_{\mathfrak{H}} U_+ = TP_{\mathfrak{H}}$, and so the operator

$$Y = P_{\mathfrak{H}} | \mathfrak{R} (= (P_{\mathfrak{R}} | \mathfrak{H})^*) \tag{2}$$

intertwines R and T: $Y \in \mathscr{I}(R, T)$. Moreover, on account of [12, Proposition II.3.5] Y is a quasi-affinity if T is a C_{11}-contraction.

By the Lifting Theorem there is a close connection between the commutants of U_+ and T (cf. [12, Theorem II.2.3] and [13]). Namely, let us denote by $\{U_+\}'_0$ the set of those operators in the commutant of U_+ which leave invariant the subspace $VH^2(\mathfrak{E})$: $\{U_+\}'_0 = \{\hat{Q} \in \{U_+\}' : \hat{Q} VH^2(\mathfrak{E}) \subset VH^2(\mathfrak{E})\}$. Then the Lifting Theorem says that the mapping

$$\pi : \{U_+\}'_0 \to \{T\}', \quad \pi\hat{Q} = P_{\mathfrak{H}} \hat{Q} | \mathfrak{H} \quad (\hat{Q} \in \{U_+\}'_0)$$

will be a well-defined, contractive, surjective, algebra-homomorphism.

Let us consider the matrix of an arbitrary operator $\hat{Q}\in\{U_+\}_0'$ with respect to the decomposition $\mathfrak{K}_+=H^2(\mathfrak{E})\oplus\mathfrak{R}$:

$$\hat{Q}=\begin{bmatrix} A & D \\ B & C \end{bmatrix}.$$

Since $\hat{Q}$ commutes with U_+, it follows that $D\in\mathcal{I}(R, U_+|H^2(\mathfrak{E}))$. Taking into account that R is unitary and $\bigcap_{n\geq 0} U_+^n H^2(\mathfrak{E})=\{0\}$, we deduce that $D=0$, i.e. $\mathfrak{R}\in\operatorname{Lat}\hat{Q}$. The relation $\hat{Q}VH^2(\mathfrak{E})\subset VH^2(\mathfrak{E})$ implies that $P_{\mathfrak{H}}\hat{Q}=QP_{\mathfrak{H}}$, where $Q=\pi\hat{Q}$. Hence, considering restrictions onto the subspace $\mathfrak{R}$, we obtain

$$YC=QY, \tag{3}$$

where C commutes with R.

Now we prove a lemma on the model-operator T introduced above.

Lemma 7. *Let us given* $\hat{Q}=\begin{bmatrix} A & 0 \\ B & C \end{bmatrix}\in\{U_+\}_0'$ *and* $Q=\pi\hat{Q}\in\{T\}'$. *If T is a C_{11}-contraction, then $Q=0$ is equivalent to $C=0$, and $C\in\{R\}''$ implies $Q\in\{T\}''$. Moreover, if T is weakly similar to unitary, then $C\in\{R\}''$ and $Q\in\{T\}''$ are equivalent.*

Proof. If T is a C_{11}-contraction, then the operator Y defined in (2) is a quasi-affinity. Hence the intertwining relation (3) yields that Q and C are equal to zero simultaneously.

Let us assume that $C\in\{R\}''$, and let us consider an arbitrary operator $Q'\in\{T\}'$. Since the mapping π is surjective, we can find an operator $\hat{Q}'=\begin{bmatrix} A' & 0 \\ B' & C' \end{bmatrix}\in\{U_+\}_0'$ such that $\pi\hat{Q}'=Q'$. In virtue of our assumption $C\in\{R\}''$ it follows that the operator $\hat{Q}'':=\hat{Q}\hat{Q}'-\hat{Q}'\hat{Q}\in\{U_+\}_0'$ has a matrix of the form

$$\begin{bmatrix} A'' & 0 \\ B'' & CC'-C'C \end{bmatrix}=\begin{bmatrix} A'' & 0 \\ B'' & 0 \end{bmatrix}.$$

Therefore, on account of the first part of our lemma, proved before, we conclude that $\pi\hat{Q}''=0$. However, π being an algebra-homomorphism this yields that $0=\pi\hat{Q}''=$ $=QQ'-Q'Q$, i.e. Q commutes with Q'.

Let us assume now that T is weakly similar to unitary, $Q\in\{T\}''$, and let us consider an arbitrary operator $C'\in\{R\}'$. On account of [9, Theorem 4] T belongs to C_{11} and $\Theta_T(e^{it})$ is boundedly invertible a.e. on C. Let $\alpha_n\subset C$ be the measurable set $\alpha_n=\{e^{it}\colon \|\Theta(e^{it})^{-1}\|\leq n\}$, for every n. Then $\{\alpha_n\}_n$ forms an increasing sequence such that $m(C\setminus(\bigcup_n\alpha_n))=0$. Consequently, if χ_{α_n} denotes also the operator of multiplication by the characteristic function χ_{α_n} of α_n, then the sequence $\{\chi_{\alpha_n}\}_n\subset\{R\}''$ tends to the identity operator $I_{\mathfrak{R}}$ in the strong operator topology.

For every n, let $\hat{Q}_n \in \{U_+\}'$ denote the operator whose matrix in the decomposition $\mathfrak{K}_+ = H^2(\mathfrak{E}) \oplus \mathfrak{R}$ is the following

$$\hat{Q}_n = \begin{bmatrix} 0 & 0 \\ -C'(\chi_{\alpha_n} \Delta \Theta^{-1}) & \chi_{\alpha_n} C' \end{bmatrix}.$$

Here $\chi_{\alpha_n} \Delta \Theta^{-1}$ stands for the operator of multiplication by the bounded, measurable, operator-valued function $\chi_{\alpha_n} \Delta \Theta^{-1}$. Since for every $w \in H^2(\mathfrak{E})$ we have

$$\hat{Q}_n V w = \hat{Q}_n(\Theta w \oplus \Delta w) = 0 \oplus (-C' \chi_{\alpha_n} \Delta \Theta^{-1} \Theta w + \chi_{\alpha_n} C' \Delta w) = 0 \oplus 0,$$

it follows that $\hat{Q}_n \in \{U_+\}'_0$. Hence $Q_n = \pi \hat{Q}_n \in \{T\}'$, and so $Q_n Q = Q Q_n$, for every n. In virtue of the first part of our lemma we conclude that

$$\chi_{\alpha_n}(C'C) = (\chi_{\alpha_n} C')C = C(\chi_{\alpha_n} C') = \chi_{\alpha_n}(CC')$$

holds, for every n. Taking into account that $\{\chi_{\alpha_n}\}_n$ converges to the identity, we obtain that

$$C'C = CC'.$$

Therefore, C belongs to $\{R\}''$, and so the proof is completed.

In order to formulate our second lemma on the model-operator T we introduce the operator-valued function $\Delta_*(e^{it}) = [I - \Theta(e^{it})\Theta(e^{it})^*]^{1/2}$. Then the operator R_*, called the $*$-residual part of T, is defined as the multiplication by e^{it} on the Hilbert space $\mathfrak{R}_* = (\Delta_* L^2(\mathfrak{E}))^-$. The following lemma, which is a generalization of [16, Lemma 3.4] (cf. also [11]), is proved in [10].

Lemma 8. *If T is a C_{11}-contraction, then the mapping*

$$X\colon \mathfrak{H} \to \mathfrak{R}_*, \quad X(u \oplus v) = -\Delta_* u + \Theta v \quad (u \oplus v \in \mathfrak{H}), \tag{4}$$

is a (well-defined) quasi-affinity, belonging to $\mathscr{I}(T, R_*)$. Moreover, its product $Z = XY \in \mathscr{I}(R, R_*)$ with the operator Y, defined in (2), acts as a multiplication by Θ, i.e.

$$(Zv)(e^{it}) = \Theta(e^{it})v(e^{it})$$

holds a.e. on C, for every $v \in \mathfrak{R}$.

Finally, we need two lemmas concerning absolutely continuous unitary operators.

Lemma 9. *Let U be the operator of multiplication by e^{it} on the space $\mathfrak{K} = L^2(\mathfrak{F})$, where $\mathfrak{F}$ is a Hilbert space, and for any non-zero vector $h \in \mathfrak{K}$ let $\mathfrak{K}_h$ denote the invariant subspace $\mathfrak{K}_h = \bigvee_{n \geq 0} U^n h$. Then the restriction $U|\mathfrak{K}_h$ belongs to the class C_{10} if and only if*

$$\int_C \log \|h(e^{it})\| \, dm(t) > -\infty.$$

Proof. Let $\mathfrak{K}_{h,0}$ denote the linear manifold $\mathfrak{K}_{h,0}=\{p(U)h: p(\lambda)$ is a complex polynomial$\}$, and let us define the mapping $V_0\colon \mathfrak{K}_{h,0}\to L^2(C, m)$ by $(V_0(p(U)h))(e^{it})=$ $=p(e^{it})\|h(e^{it})\|_{\mathfrak{F}}$. It is immediate that V_0 is a (linear) isometry (hence well-defined), and so it can be extended to an isometry $V\in\mathcal{L}(\mathfrak{K}_h, L^2(C, m))$. Since we evidently have $V_0(U|\mathfrak{K}_{h,0})=MV_0$, where M denotes the operator of multiplication by e^{it} in $L^2(C, m)$, it follows that

$$V(U|\mathfrak{K}_h) = MV.$$

This yields that $\operatorname{ran} V\in\operatorname{Lat} M$ and $U|\mathfrak{K}_h$ is unitarily equivalent to $M|\operatorname{ran} V$. Therefore $U|\mathfrak{K}_h$ belongs to the class C_{10} if and only if so does the operator $M|\operatorname{ran} V$. However, taking into account that

$$\operatorname{ran} V = \bigvee_{n\geqq 0} M^n\|h\|,$$

we conclude that $M|\operatorname{ran} V\in C_{10}$ holds exactly when

$$\int_C \log\|h(e^{it})\|\, dm(t) > -\infty.$$

(Cf. the Szegő—Kolmogoroff—Krein theorem in [6].)

Lemma 10. *Let $U\in\mathcal{L}(\mathfrak{K})$ be an absolutely continuous unitary operator, and let us consider an operator $C\in\{U\}''$. If C leaves invariant a non-zero subspace $\mathfrak{M}\in\operatorname{Lat} U$ such that $U|\mathfrak{M}\in C_{10}$, then C is of the form $C=\delta(U)$, where δ is a function from H^∞.*

Proof. Since $C\in\{U\}''$, we infer by the spectral theorem (cf. [4]) that C has the form $C=\delta(U)$ with an appropriate function $\delta\in L^\infty(m)$.

The assumption $U|\mathfrak{M}\in C_{10}$ implies that $U|\mathfrak{M}$ is a unilateral shift. Consequently, the subspace $\mathfrak{L}=\mathfrak{M}\ominus U\mathfrak{M}$ is wandering for U, i.e. the sequence $\{U^n\mathfrak{L}\}_{n=-\infty}^{\infty}$ consists of pairwise orthogonal subspaces. Let us consider the subspace

$$\hat{\mathfrak{M}} = \bigoplus_{n=-\infty}^{\infty} U^n\mathfrak{L},$$

which clearly reduces U. Taking into account that $C\in\{U\}''$ we conclude that $\hat{\mathfrak{M}}\in\operatorname{Lat} C$ and

$$C|\hat{\mathfrak{M}} = \delta(U)|\hat{\mathfrak{M}} = \delta(U|\hat{\mathfrak{M}}).$$

Hence, we obtain that

$$\delta(U|\hat{\mathfrak{M}})\mathfrak{M}\subset\mathfrak{M}. \tag{5}$$

Let us consider now the Fourier-representation of $\hat{\mathfrak{M}}$, i.e. the unitary map

$$\Phi\colon \hat{\mathfrak{M}}\to L^2(\mathfrak{L}),\quad \Big(\Phi\Big(\bigoplus_{n=-\infty}^{\infty} U^n h_n\Big)\Big)(e^{it}) = \sum_{n=-\infty}^{\infty} e^{itn} h_n$$

($h_n \in \mathfrak{L}$, for every n). Φ intertwines $U|\hat{\mathfrak{M}}$ with the operator $M \in \mathscr{L}(L^2(\mathfrak{L}))$ of multiplication by e^{it}: $\Phi(U|\hat{\mathfrak{M}}) = M\Phi$. This yields the relation

$$\Phi\delta(U|\hat{\mathfrak{M}}) = \delta(M)\Phi. \tag{6}$$

Consequently, on account of (5) and (6) we infer

$$\delta(M)H^2(\mathfrak{L}) \subset H^2(\mathfrak{L}),$$

which implies that $\delta \in H^\infty$, and the proof is finished.

Now we are ready to prove our main theorem.

Proof of Theorem 5. It is enough to show that $\operatorname{Alg}\operatorname{Lat} T \subset H^\infty(T)$. Indeed, then on account of the relations $H^\infty(T) \subset \operatorname{Alg} T \subset \operatorname{Alg}\operatorname{Lat} T$ it follows that

$$H^\infty(T) = \operatorname{Alg} T = \operatorname{Alg}\operatorname{Lat} T,$$

hence T is reflexive. Moreover, in virtue of [9, Corollary 12] we obtain $\operatorname{Alg} T \neq \{T\}''$, and taking into consideration that $\operatorname{Alg}\operatorname{Lat}_1 T = \{T\}''$ (cf. the proof of [9, Proposition 13]) we conclude $\operatorname{Lat} T \neq \operatorname{Lat}_1 T$.

So let $Q \in \operatorname{Alg}\operatorname{Lat} T$ be an arbitrary operator. We shall show that $Q \in H^\infty(T)$. On account of [12, Theorem VI.2.3] we may assume that T is a model-operator $T = S(\Theta)$, where $\{\Theta(\lambda), \mathfrak{E}, \mathfrak{E}\}$ is a purely contractive, analytic function, outer from both sides.

Since Q clearly belongs to $\operatorname{Alg}\operatorname{Lat}'' T$, we infer by the reflexivity of $\{T\}''$ (cf. [14]) that

$$Q \in \{T\}''.$$

On account of the Lifting Theorem there is an operator $\hat{Q} = \begin{bmatrix} A & 0 \\ B & C \end{bmatrix} \in \{U_+\}'_0$ such that $Q = \pi\hat{Q}$. An application of Lemma 7 gives that

$$C \in \{R\}''.$$

In order to be able to apply Lemma 10 we have to show that $C\mathfrak{M} \subset \mathfrak{M}$ for a non-zero subspace $\mathfrak{M} \in \operatorname{Lat} R$ such that $R|\mathfrak{M} \in C_{10}$.

By the assumption there exists a vector $f \in \mathfrak{R}$ such that

$$\int_C \log \|g(e^{it})\|\, dm(t) > -\infty$$

for the function $g = \Theta f$. Now, on account of Lemma 8 we know that g is contained in $\mathfrak{R}_*$. Moreover, applying Lemma 9 we obtain that

$$R_* | \mathfrak{R}_{*,g} \in C_{10}, \tag{7}$$

for the subspace $\mathfrak{R}_{*,g} = \bigvee_{n \geqq 0} R_*^n g \in \operatorname{Lat} R_*$. Then the intertwining relation $R_* X = XT$, where X is the operator defined in (4), implies that the subspace $\mathfrak{L} = X^{-1}\mathfrak{R}_{*,g}$ is invariant for T. Since, by Lemma 8, X is injective, we see that $T|\mathfrak{L}$ can be injected into $R_*|\mathfrak{R}_{*,g}$:

(8) $$T|\mathfrak{L} \overset{i}{\prec} R_*|\mathfrak{R}_{*,g}.$$

We conclude by (7) and (8) that the operator $T|\mathfrak{L}$ is also of class C_{10}.

An analogous argumentation yields that the subspace

$$\mathfrak{M} = Y^{-1}\mathfrak{L},$$

where Y is defined by (2), is invariant for R and

$$R|\mathfrak{M} \in C_{10}.$$

Since the non-zero vector f clearly belongs to $\mathfrak{M}$, it follows that $\mathfrak{M}$ is non-zero. On the other hand, $\mathfrak{L}\in \operatorname{Lat} T$ and $Q\in \operatorname{Alg}\operatorname{Lat} T$ imply

$$\mathfrak{L}\in \operatorname{Lat} Q.$$

Hence, the intertwining relation (3) yields that

$$\mathfrak{M}\in \operatorname{Lat} C.$$

Now, we can apply Lemma 10 to obtain that C has the form $C=\delta(R)$, with a suitable function $\delta\in H^\infty$.

Since the operator Y intertwines R and T too, we infer

$$\delta(T)Y = Y\delta(R) = YC.$$

Comparing this equality with (3) we conclude that

$$\delta(T)Y = QY.$$

Consequently, taking into account that Y is a quasi-affinity we obtain

$$Q = \delta(T).$$

The theorem is proved.

3. Concluding remarks

Under more general assumptions we are able to prove the following weaker version of part (i) of Theorem 5.

Proposition 11. *If T is a c.n.u. C_{11}-contraction such that $\Theta_T(e^{it})$ is not isometric a.e. on C, then*

$$H^\infty(T) = \operatorname{Alg}_* T,$$

where $\operatorname{Alg}_* T$ denotes the algebra generated by T and the identity, and closed in the ultraweak operator topology.

Proof. First of all we note the elementary fact that if an operator S is similar to a normal operator N, then $\|S\| \geqq \|N\|$. Indeed, similarity preserves the spectrum, so if r_S, r_N denote the spectral radii of S and N, respectively, then we can write $\|N\| = r_N = r_S \leqq \|S\|$.

Let us assume now that $T \in \mathcal{L}(\mathfrak{H})$ is a c.n.u. C_{11}-contraction such that $\Theta_T(e^{it})$ is not isometric a.e. on C. We can find an absolutely continuous unitary operator $U \in \mathcal{L}(\mathfrak{K})$, which is quasi-similar to T (cf. [12, Proposition II.3.5 and Theorem II.6.4]). By a result of Apostol (cf. [1]) there are basic systems $\{\mathfrak{H}_n\}_n$ and $\{\mathfrak{K}_n\}_n$ in $\mathfrak{H}$ and $\mathfrak{K}$, respectively, such that $\mathfrak{H}_n \in \operatorname{Lat} T$, $\mathfrak{K}_n \in \operatorname{Lat} U$ and $T|\mathfrak{H}_n$ is similar to $U|\mathfrak{K}_n$, for every n. Moreover, it can be achieved that the subspaces $\{\mathfrak{K}_n\}_n$ are pairwise orthogonal, i.e. the decomposition $\mathfrak{K} = \bigoplus_n \mathfrak{K}_n$ reduces U.

Let us given an arbitrary function $w \in H^\infty$. Since $w(T|\mathfrak{H}_n)$ is similar to the normal operator $w(U|\mathfrak{K}_n)$, we infer that

$$\|w(T)\| \geqq \|w(T)|\mathfrak{H}_n\| = \|w(T|\mathfrak{H}_n)\| \geqq \|w(U|\mathfrak{K}_n)\| = \|w(U)|\mathfrak{K}_n\|,$$

for every n, hence

$$\|w(T)\| \geqq \sup_n \|w(U)|\mathfrak{K}_n\| = \|w(U)\|.$$

However, $\Theta_T(e^{it})$ being not isometric a.e. on C, it follows by [7, Corollary 1] and [12, Proposition II.3.4] that $\sigma(U) = C$, and so $\|w(U)\| = \|w\|_\infty$. Therefore, we conclude that $\|w(T)\| \geqq \|w\|_\infty$. Since the opposite direction always holds (cf. [12, Theorem III.2.1]), we obtain that

$$\|w(T)\| = \|w\|_\infty,$$

for every $w \in H^\infty$, i.e. the Sz.-Nagy, Foiaş functional calculus is an isometry. But then on account of [3, Theorem 3.2] we get that

$$H^\infty(T) = \operatorname{Alg}_* T,$$

and the proof is finished.

It is left open whether the statements of Theorem 5 remain true under the assumption of Proposition 11, even in the case when T is weakly similar to unitary. The following example illuminates where difficulties arise.

Example 12. Let $\{\alpha_n\}_{n=1}^\infty$ be a sequence of pairwise disjoint Borel subsets of the unit circle C such that $m(\alpha_n) > 0$, for every n, and $\sum_n m(\alpha_n) = 1$. Let us choose an arbitrary sequence $\{c_n\}_{n=1}^\infty$ of positive numbers, where $c_n < 1$ for every n, and for each n let us define a (scalar-valued) outer function ϑ_n by the boundary condition

$$|\vartheta_n(e^{it})| = c_n \chi_{\alpha_n}(e^{it}) + \chi_{C \setminus \alpha_n}(e^{it}) \quad \text{a.e.}$$

Let us consider a separable, infinite dimensional Hilbert space $\mathfrak{E}$, and an orthonormal basis $\{e_n\}_n$ in $\mathfrak{E}$. Then $\{\Theta(\lambda), \mathfrak{E}, \mathfrak{E}\}$ will stand for the contractive, operator-valued, analytic function, whose matrix is

$$[\Theta(\lambda)] = \operatorname{diag}(\vartheta_n(\lambda)), \quad \lambda\in \mathbf{D},$$

in this basis. We shall examine the model-operator

$$T = S(\Theta),$$

which, of course, depends on the choice of sequences $\{\alpha_n\}_n$ and $\{c_n\}_n$.

Since Θ is outer from both sides, it follows that T is of class C_{11}. Moreover, the identity

$$\|\Theta(e^{it})^{-1}\| = \sum_n \chi_{\alpha_n}(e^{it})c_n^{-1},$$

being valid a.e. on C, implies that T is weakly similar to unitary (cf. [9, Theorem 4]).

Since $(\Delta L^2(\mathfrak{E}))^-$ splits into the orthogonal sum $(\Delta L^2(\mathfrak{E}))^- = \bigoplus_n (\Delta_n L^2(\mathfrak{E}_n))^-$, where $\mathfrak{E}_n$ is the one-dimensional subspace of $\mathfrak{E}$ spanned by e_n, $\Delta_n(e^{it})$ acts on $\mathfrak{E}_n$, and $\Delta_n(e^{it})e_n = (1-|\vartheta_n(e^{it})|^2)^{1/2}e_n$, it follows easily that relation (1) in Theorem 5 is satisfied with a vector $f\in(\Delta L^2(\mathfrak{E}))^-$ if and only if Θ has a scalar multiple, i.e. if

$$\infty > \int_C \log\|\Theta(e^{it})^{-1}\|\,dm(t) = \sum_n (\log c_n^{-1})\,m(\alpha_n)$$

holds. Hence, Theorem 5 can be applied exactly when

$$\sum_n m(\alpha_n)\log c_n^{-1} < \infty. \tag{9}$$

Let us examine now what the spectrum $\sigma(T)$ of T is like. We know by [12, Theorem VI.4.1] that a point μ of the unit disc $\mathbf{D}$ belongs to the spectrum if and only if $\Theta(\mu)$ is not invertible, which is equivalent to the condition

$$\sup_n |\vartheta_n(\mu)|^{-1} = \infty.$$

Taking into account that

$$\exp\left[\frac{1-|\mu|}{1+|\mu|}\log c_n^{-1}\cdot m(\alpha_n)\right] \leqq |\vartheta_n(\mu)|^{-1} = \exp\Big[\int_{\alpha_n} P_r(\varphi-t)\log c_n^{-1}\,dm(t)\Big] \leqq$$

$$\leqq \exp\left[\frac{1+|\mu|}{1-|\mu|}\log c_n^{-1}\cdot m(\alpha_n)\right],$$

where $\mu = re^{i\varphi}$, and P_r denotes the Poisson-kernel, we infer that $\mu\in\sigma(T)$ exactly when the equality

$$\sup_n [m(\alpha_n) \log c_n^{-1}] = \infty, \tag{10}$$

independent of μ, holds. Since $\sigma(T)$ always includes the whole unit circle C (cf. [12, Theorem VI.4.1]), we obtain that $\sigma(T)=\mathbf{D}^-$ or $\sigma(T)=C$ according to the case when (10) is fulfilled or not.

Taking into consideration that the essential spectrum $\sigma_e(T)$ of the C_{11}-contraction T coincides with its spectrum, we conclude that $\sigma_e(T)$ is dominating in $\mathbf{D}$, i.e. T is a (BCP)-operator (cf. [2]) if and only if (10) holds. But then [2, Theorem 1] also yields that the statements of Theorem 5 are true. (Cf. also the beginning of the proof of Theorem 5.)

Summerizing, we have obtained that the statements (i)—(iii) of Theorem 5 are valid if (9) or (10) are fulfilled, i.e. either if the sequence $\{m(\alpha_n) \log c_n^{-1}\}_n$ tends to zero fast enough or if it is unbounded. The intermediate case remains open.

Added in proof (December 10, 1987). In a subsequent paper, appearing in *Acta Math. Hung.* **50** (1987), further developing the methods of this work we succeeded in answering the question raised above.

References

[1] C. Apostol, Operators quasi-similar to a normal operator, *Proc. Amer. Math. Soc.*, **53** (1975), 104—106.

[2] H. Bercovici, C. Foiaş, J. Langsam and C. Pearcy, (BCP)-operators are reflexive, *Michigan Math. J.*, **29** (1982), 371—379.

[3] S. Brown, B. Chevreau and C. Pearcy, Contractions with rich spectrum have invariant subspaces, *J. Operator Theory*, **1** (1979), 123—136.

[4] N. Dunford and J. Schwartz, *Linear Operators*, Part II, Interscience Publishers (New York—London, 1963).

[5] P. R. Halmos, Ten problems in Hilbert space, *Bull. Amer. Math. Soc.*, **76** (1970), 887—933.

[6] K. Hoffman, *Banach Spaces of Analytic Functions*, Prentice-Hall (Englewood Cliffs, N. J., 1962).

[7] L. Kérchy, On the commutant of C_{11}-contractions, *Acta Sci. Math.*, **43** (1981), 15—26.

[8] L. Kérchy, Subspace lattices connected with C_{11}-contractions, *Anniversary Volume on Approximation Theory and Functional Analysis* (eds. P. L. Butzer, R. L. Stens, B. Sz.-Nagy), Birkhäuser Verlag (Basel—Boston—Stuttgart, 1984), pp. 89—98.

[9] L. Kérchy, Contractions being weakly similar to unitaries, *Operator Theory: Advances and Applications*, Vol. **17**, Birkhäuser Verlag (Basel, 1986), 187—200.

[10] L. Kérchy, A description of invariant subspaces of C_{11}-contractions, *J. Operator Theory*, **15** (1986), 327—344.

[11] S. O. Sickler, The invariant subspaces of almost unitary operators, *Indiana Univ. Math. J.*, **24** (1975), 635—650.

[12] B. Sz.-Nagy and C. Foiaş, *Harmonic Analysis of Operators on Hilbert Space*, North Holland—Akadémiai Kiadó (Amsterdam—Budapest, 1970).

[13] B. SZ.-NAGY and C. FOIAŞ, On the structure of intertwining operators, *Acta Sci. Math.*, **35** (1973), 225—254.
[14] K. TAKAHASHI, Double commutants of operators quasi-similar to normal operators, *Proc. Amer. Math. Soc.*, **92** (1984), 404—406.
[15] P. Y. WU, C_{11}-contractions are reflexive, *Proc. Amer. Math. Soc.*, **77** (1979), 68—72.
[16] P. Y. WU, Bi-invariant subspaces of weak contractions, *J. Operator Theory*, **1** (1979), 261—272.

BOLYAI INSTITUTE
UNIVERSITY SZEGED
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Acta Sci. Math., 51 (1987), 491—500

Integral manifolds, stability and decomposition of singularly perturbed systems in Banach space

V. A. SOBOLEV*)

1. Introduction. This paper is dealing with the study of infinite dimensional singularly perturbed systems near an integral manifold.

Consider the system

$$\dot{x} = f(t, x, y, \varepsilon) \tag{1.1}$$
$$\varepsilon\dot{y} = Ay+\varepsilon g(t, x, y, \varepsilon)$$

where x and y are elements of Banach spaces X and Y with norms $\|\cdot\|$, A is a constant linear bounded operator in Y, and

$$f\colon R\times X\times Y\times[0, \varepsilon_0] \to X, \quad g\colon R\times X\times Y\times[0, \varepsilon_0] \to Y$$

are continuous nonlinear operator functions. Using the method of integral manifolds [1, 2] we shall study the stability problem for (1.1) and the problem of decomposition of (1.1) by transforming it to the form

$$\dot{u} = F(t, u, \varepsilon), \tag{1.2}$$

$$\varepsilon\dot{v} = Av+\varepsilon G(t, u, v, \varepsilon). \tag{1.3}$$

Then we shall apply this method for investigation of linear singularly perturbed systems.

2. Slow manifold. We first recall the definition of an integral manifold for the equation $\dot{x}=X(t, x)$ where x is an element of a Banach space. A set S is said to be an integral manifold if for $(t_0, x_0)\in S$, the solution $(t, x(t))$, $x(t_0)=x_0$ is in S for $t\in R$. If $(t, x(t))\in S$ only at a finite interval, then we shall say that S is a local integral manifold.

*) This research was completed while the author was visiting the Department of Mathematics at the Budapest University of Technology.

Received August 2, 1984, and in revised form August 4, 1986.

Let $B_r=\{y\in Y,\ \|y\|\leqq r\}$, $I_{\varepsilon_0}=[0,\varepsilon_0]$, $\Omega=R\times X\times B_r\times I_{\varepsilon_0}$. Assume that f and g are bounded and satisfy the Lipschitz condition in x, y on Ω:

$$\|f(t,x,y,\varepsilon)\|\leqq M,\quad \|g(t,x,y,\varepsilon)\|\leqq M, \tag{2.1}$$

$$\begin{aligned}\|f(t,x,y,\varepsilon)-f(t,\bar{x},\bar{y},\varepsilon)\|&\leqq l(\|x-\bar{x}\|+\|y-\bar{y}\|),\\ \|g(t,x,y,\varepsilon)-g(t,\bar{x},\bar{y},\varepsilon)\|&\leqq l(\|x-\bar{x}\|+\|y-\bar{y}\|),\end{aligned} \tag{2.2}$$

where M and l are positive constants.

Assume that the spectrum $\sigma(A)$ of the linear bounded operator A satisfies the inequality $\operatorname{Re}\sigma(A)\leqq-2\alpha<0$. Then there exists a positive number K such that

$$\|e^{At}\|\leqq Ke^{-\alpha t},\quad t\geqq 0. \tag{2.3}$$

We shall say that the integral manifold of system (1.1) is a slow manifold if it can be represented of form $y=h(t,x,\varepsilon)$, where h is a continuous operator-function. If ε_0 is sufficiently small then for each $\varepsilon\in(0,\varepsilon_0)$ the system (1.1) has an integral manifold (slow manifold) represented of form $y=\varepsilon h(t,x,\varepsilon)$ (see [1], p. 438). Here h is a continuous and bounded operator-function defined on $\Omega_1=R\times X\times I_{\varepsilon_0}$ and satisfies the Lipschitz condition in x:

$$\|h(t,x,\varepsilon)-h(t,\bar{x},\varepsilon)\|\leqq\Delta\|x-\bar{x}\|,\quad \Delta>0. \tag{2.4}$$

Moreover, if f and g are continuously differentiable on Ω to k order and their derivatives are bounded and Lipschitzian in x, y then h is continuously differentiable on Ω_1 to k and its derivatives are bounded and Lipschitzian in x. In this case the operator-function h can be represented as asymptotic expansion $\varepsilon h=\varepsilon h_1(t,x)+\dots$ $\dots+\varepsilon^k h_k(t,x)+h_{k+1}(t,x,\varepsilon)$ where $h_{k+1}=O(\varepsilon^{k+1})$. The coefficients h_i of this expansion can be found from the equation

$$\varepsilon\frac{\partial h}{\partial t}+\varepsilon\frac{\partial h}{\partial x}f(t,x,\varepsilon h,\varepsilon)=Ah+g(t,x,\varepsilon h,\varepsilon). \tag{2.5}$$

For finite dimensional systems this method of approximating slow manifolds was essentially used in [3]. The method of approximation used in [4] can be generalized to infinite dimensional problems in an obvious way.

The flow on a slow manifold is governed by the reduced equation (1.2), where $F(t,u,\varepsilon)=f(t,u,\varepsilon h(t,u,\varepsilon),\varepsilon)$.

It is well-known for finite dimensional spaces X that the condition $f(t,0,0,\varepsilon)=0$, $g(t,0,0,\varepsilon)=0$ implies $h(t,0,\varepsilon)=0$ and if the zero solution of (1.2) is stable (asymptotically stable, unstable) then the zero solution of (1.1) is stable (asymptotically stable, unstable). We shall prove below this statement for infinite dimensional X.

3. Integral manifold for auxiliary system. Let us suppose that f and g are continuously differentiable on Ω and their derivatives are bounded and Lipschitzian in x, y and introduce new variables u, z and x_1 by the formulae $z=y-\varepsilon h(t, x, \varepsilon)$, $x_1=x-u$ where u satisfies (1.2). Consider the following auxiliary differential system

$$\begin{aligned} \dot{u} &= F(t, u, \varepsilon) \\ \dot{x}_1 &= f_1(t, u, x_1, z, \varepsilon) \qquad (3.1) \\ \varepsilon \dot{z} &= Az+\varepsilon Z(t, u, x_1, z, \varepsilon), \end{aligned}$$

where $f_1=f(t, u+x_1, z+\varepsilon h(t, u+x_1, \varepsilon), \varepsilon)-F(t, u, \varepsilon)$,

$$Z = g(t, u+x_1, z+\varepsilon h(t, u+x_1, \varepsilon), \varepsilon)-g(t, u+x_1, \varepsilon h(t, u+x_1, \varepsilon), \varepsilon)-$$

$$-\varepsilon \frac{\partial h}{\partial x}(t, u+x_1, \varepsilon)[f(t, u+x_1, z+\varepsilon h(t, u+x_1, \varepsilon), \varepsilon)-f(t, u+x_1, \varepsilon h(t, u+x_1, \varepsilon), \varepsilon)].$$

By means of our assumptions it is easy to show that there exists a constant $N>0$ such that f_1 and Z satisfy the following inequalities

$$\|f_1(t, u, x_1, z, \varepsilon)\| \leqq N(\|x_1\|+\|z\|), \tag{3.2}$$

$$\|Z(t, u, x_1, z, \varepsilon)\| \leqq N\|z\|, \tag{3.3}$$

$$\|f_1(t, u, x_1, z, \varepsilon)-f_1(t, u, \bar{x}_1, \bar{z}, \varepsilon)\| \leqq N(\|x_1-\bar{x}_1\|+\|z-\bar{z}\|), \tag{3.4}$$

$$\|Z(t, u, x_1, z, \varepsilon)-Z(t, u, \bar{x}_1, \bar{z}, \varepsilon)\| \leqq N(\|x_1-\bar{x}_1\|+\|z-\bar{z}\|), \tag{3.5}$$

$$\begin{aligned} &\|f_1(t, u, x_1, z, \varepsilon)-f_1(t, \bar{u}, \bar{x}_1, \bar{z}, \varepsilon)\| \leqq \\ &\leqq N[1+\|x_1\|+\|z\|][\|z-\bar{z}\|+(1+\|x_1\|)\|x_1-\bar{x}_1\|+(\|x_1\|+\|z\|)\|u-\bar{u}\|], \end{aligned} \tag{3.6}$$

$$\|Z(t, u, x_1, z, \varepsilon)-Z(t, \bar{u}, \bar{x}_1, \bar{z}, \varepsilon)\| \leqq N[\|z-\bar{z}\|+\|z\|(\|u-\bar{u}\|+\|x_1-\bar{x}_1\|)], \tag{3.7}$$

where $t\in R$, $u, \bar{u}\in X$, $x_1, \bar{x}_1\in X$, $z, \bar{z}\in B_{r_1}$, $0<r_1\leqq r$.

We shall show that the system (3.1) has an integral manifold represented of form $x_1=\varepsilon H(t, u, z, \varepsilon)$, where H is an operator-function defined and continuous on $\Omega_2=R\times X\times B_\varrho\times I_{\varepsilon_1}$, $0<\varrho<\frac{r_1}{K}$, $0<\varepsilon\leqq\varepsilon_0$, and H satisfies the inequalities:

$$\|H(t, u, z, \varepsilon)\| \leqq a\|z\|, \tag{3.8}$$

$$\|H(t, u, z, \varepsilon)-H(t, u, \bar{z}, \varepsilon)\| \leqq b\|z-\bar{z}\|, \tag{3.9}$$

$$\|H(t, u, z, \varepsilon)-H(t, \bar{u}, z, \varepsilon)\| \leqq c\|z\|\cdot\|u-\bar{u}\|, \tag{3.10}$$

with $a, b, c>0$ for $t\in R$, $u, \bar{u}\in X$, $z, \bar{z}\in B_\varrho$, $\varepsilon\in I_{\varepsilon_1}$.

The flow on this manifold is governed by reduced equations (1.2), (1.3), where $F=f(t, u, \varepsilon h(t, u, \varepsilon), \varepsilon)$, $G=Z(t, u, \varepsilon H(t, u, v, \varepsilon), v, \varepsilon)$.

Moreover, every solution of (1.1) with $\|y(t_0)-\varepsilon h(t_0, x(t_0), \varepsilon)\| \leqq \varrho$ can be represented of form

$$x = u+\varepsilon H(t, u, v, \varepsilon),$$

$$(3.11) \qquad y = v+\varepsilon h(t, x, \varepsilon) = v+\varepsilon h\big(t, u+\varepsilon H(t, u, v, \varepsilon), \varepsilon\big),$$

where u, v is the corresponding solution of (1.2), (1.3).

Our proof of this statements is modelled on KELLEY [5].

Let S be the set of operator-functions εH: $\Omega_2 \to X$ such that H is continuous and satisfies (3.8)—(3.10). Let d be a metric on S defined by

$$d(\varepsilon H, \varepsilon \bar{H}) = \sup\left\{\frac{1}{\|z\|}\varepsilon\|H(t, u, z, \varepsilon)-\bar{H}(t, u, z, \varepsilon)\|,\ t\in R,\ u\in X,\ z\in B_\varrho\right\}$$

for each $\varepsilon\in(0, \varepsilon_1]$, εH, $\varepsilon\bar{H}\in S$ and note that S is a complete metric space with respect to d.

For each $\varepsilon H\in S$, we consider the system

$$(3.12) \qquad \dot{u} = F(t, u, \varepsilon),$$

$$(3.13) \qquad \varepsilon\dot{z} = Az+\varepsilon Z\big(t, u, \varepsilon H(t, u, z, \varepsilon), z, \varepsilon\big),$$

with solutions denoted by $u=\Phi(t, t_0, u_0, \varepsilon)$, $z=\Psi(t, t_0, u_0, z_0, \varepsilon|H)$ where $\Phi(t_0, t_0, u_0, \varepsilon)=u_0$, $\Psi(t_0, t_0, u_0, z_0, \varepsilon|H)=z_0$. The operator-functions $F(t, u, \varepsilon)$, $Z(t, u, \varepsilon H(t, u, z, \varepsilon), z, \varepsilon)$ are uniformly bounded on their domains, hence, any solution of (3.12), (3.13) is defined for all t.

As usually, (see [1, 2, 5]) the equality $x_1=\varepsilon H(t, u, z, \varepsilon)$ describes an integral manifold for (3.1) if and only if the operator-function εH is a solution of the equation

$$(3.14) \qquad \varepsilon H(\tau, u, z, \varepsilon) = -\int_\tau^\infty f_1\big(t, \Phi(t, \tau, u, \varepsilon), \varepsilon H(t, \Phi(t, \tau, u, \varepsilon), \Psi(t, \tau, u, z, \varepsilon|H), \varepsilon\big), \Psi(t, \tau, u, z, \varepsilon|H), \varepsilon\big)\,dt.$$

Let $\varphi(t)=\Phi(t, \tau, u, \varepsilon)$, $\psi(t)=\Psi(t, \tau, u, z, \varepsilon|H)$ then by the "variation of constants" formula

$$\psi(t) = e^{(1/\varepsilon)A(t-\tau)}z+\int_\tau^t e^{(1/\varepsilon)A(t-s)}Z\big(s, \varphi(s), \varepsilon H\big(s, \varphi(s), \psi(s), \varepsilon\big)\,ds.$$

By (2.3), (3.3) and (3.8) there holds for all $-\infty<\tau\leqq t<\infty$, $\|z\|\leqq\varrho$, $\varepsilon\in(0, \varepsilon_1]$:

$$\|\psi(t)\| \leqq Ke^{-(\alpha/\varepsilon)(t-\tau)}\|z\|+\int_\tau^t Ke^{-(\alpha/\varepsilon)(t-s)}N\|\psi(s)\|\,ds.$$

Therefore, by Gronwall's Lemma, we obtain

$$\|\psi(t)\| \leqq Ke^{-(\alpha_1/\varepsilon)(t-\tau)}\|z\|, \quad -\infty<\tau\leqq t<\infty, \tag{3.15}$$

where $\alpha_1=\alpha-\varepsilon KN>\gamma>0$ for sufficiently small ε_1.

Now define an operator T on S by setting

$$T(H)(\tau, u, z, \varepsilon) = -\int_\tau^\infty f_1(t, \varphi(t), \varepsilon H(t, \varphi(t), \psi(t), \varepsilon), \psi(t), \varepsilon)\, dt. \tag{3.16}$$

The improper integral here converges by virtue of (3.2), (3.8) and (3.15). It is clear that $T(H)$ as defined in (3.16) is continuous on Ω_2. Also, by (3.2), (3.8) and (3.15) we obtain

$$\|T(H)(\tau, u, z, \varepsilon)\| \leqq \int_\tau^\infty N(1+\varepsilon a)Ke^{-(\alpha_1/\varepsilon)(t-\tau)}\|z\|\, dt = \varepsilon\frac{NK}{\alpha_1}(1+\varepsilon a)\|z\|,$$

and therefore $T(H)$ satisfies the boundedness condition required by (3.8) if $\varepsilon_1\dfrac{NK}{\alpha_1}<1$ and $a\geqq\dfrac{NK}{\alpha_1}\Big/\left(1-\varepsilon\dfrac{NK}{\alpha_1}\right)$.

To prove that $T(H)$ satisfies the conditions, required by (3.9), (3.10) we reason as follows. Let $u\in X$, $z, \bar{z}\in B_\varrho$, $\psi_1=\Psi(t, \tau, u, \bar{z}, \varepsilon|H)$. Then, by (3.5), (3.9), (2.3) and by the "variations of constants" formula we have

$$\|\psi(t)-\psi_1(t)\| \leqq Ke^{-(\alpha/\varepsilon)(t-\tau)}\|z-\bar{z}\| + \int_\tau^t Ke^{-(\alpha/\varepsilon)(t-s)}N(1+\varepsilon b)\|\psi(s)-\psi_1(s)\|\, ds.$$

Therefore, by Gronwall's Lemma, we obtain

$$\|\psi(t)-\psi_1(t)\| \leqq Ke^{-(\alpha_2/\varepsilon)(t-\tau)}\|z-\bar{z}\|, \quad -\infty<\tau\leqq t<\infty, \tag{3.17}$$

$$\alpha_2 = \alpha-\varepsilon KN(1+\varepsilon b).$$

Then, by (3.4), (3.9) and (3.17)

$$\|T(H)(\tau, u, z, \varepsilon)-T(H)(\tau, u, \bar{z}, \varepsilon)\| \leqq \int_\tau^\infty N(1+\varepsilon b)\|\psi(t)-\psi_1(t)\|\, dt \leqq$$

$$\leqq \varepsilon\frac{KN}{\alpha_2}(1+\varepsilon b)\|z-\bar{z}\|$$

It is clear that for sufficiently small ε_1 a constant b can be choosen such that $\alpha_2>\gamma$ and $\dfrac{K^2N}{\alpha_2}(1+\varepsilon_1 b)\leqq b$. From this inequality it follows that $T(H)$ satisfies the Lipschitz condition required by (3.9).

In exactly the same way by the inequality

$$\|\Phi(t, \tau, u, \varepsilon)-\Phi(t, \tau, \bar{u}, \varepsilon)\| \leqq e^{l(1+\varepsilon\Delta)(t-\tau)}\|u-\bar{u}\|, \quad -\infty<\tau\leqq t<\infty$$

and (3.10), (3.7) and (3.6) it is easy to show that for some $c>0$ and sufficiently small ε_1 the operator-function $T(H)$ satisfies the condition (3.10). Now, let $\varepsilon H, \varepsilon\bar{H}\in S$, $\psi_2(t)=\Psi(t, \tau, u, z, \varepsilon|\bar{H})$. Then by (3.4) and (3.9)

$$\|T(H)(\tau, u, z, \varepsilon)-T(\bar{H})(\tau, u, z, \varepsilon)\| \leqq \tag{3.18}$$

$$\leqq \int_\tau^\infty N[(1+\varepsilon b)\|\psi(t)-\psi_2(t)\|+\varepsilon\|H(t, \varphi(t), \psi_2(t), \varepsilon)-\bar{H}(t, \varphi(t), \psi_2(t), \varepsilon)\|]\,dt \leqq$$

$$\leqq \int_\tau^\infty N[(1+\varepsilon b)\|\psi(t)-\psi_2(t)\|+Ke^{-(\gamma/\varepsilon)(t-\tau)}\|z\|\,d(\varepsilon H, \varepsilon\bar{H})]\,dt.$$

Using (3.5) and (3.9) we find that

$$\|\psi(t)-\psi_2(t)\| \leqq$$

$$\leqq \int_\tau^t Ke^{-(\alpha/\varepsilon)(t-\tau)}N[(1+\varepsilon b)\|\psi(s)-\psi_2(s)\|+Ke^{-(\gamma/\varepsilon)(t-s)}\|z\|\,d(\varepsilon H, \varepsilon\bar{H})]\,dt.$$

Substitution of this into (3.18) yields

$$\frac{1}{\|z\|}\|T(H)(\tau, u, z, \varepsilon)-T(\bar{H})(\tau, u, z, \varepsilon)\| \leqq \varepsilon\frac{KN}{\gamma}\left[(1+\varepsilon b)\frac{KN}{\alpha_2-\gamma}+1\right]d(\varepsilon H, \varepsilon\bar{H}).$$

From this last inequality it easily follows that T is a contraction mapping if ε_1 is sufficiently small.

Thus, T is a contraction mapping of S into itself and so, by the known Banach Contraction Principle, T must have a unique fixed point $\varepsilon H\in S$. The operator-function εH is a solution of (3.14) and, therefore, the equality $x_1=\varepsilon H(t, u, z, \varepsilon)$ represents an integral manifold for (3.1). The flow on this manifold is governed by (1.2), (1.3) where

$$F=f(t, u, \varepsilon h(t, u, \varepsilon), \varepsilon), \quad G=Z(t, u, \varepsilon H(t, u, v, \varepsilon), v, \varepsilon).$$

4. Decomposition and stability. Our next object is to obtain the representation (3.11). Let $x=x(t)$, $y=y(t)$ be a solution of (1.1) with $x(t_0)=x_0$, $y(t_0)=y_0$, $\|y_0-\varepsilon h(t_0, x_0, \varepsilon)\|\leqq\varrho$. We shall show that there exists a solution $u=u(t)$, $u(t_0)=u_0$, $v=v(t)$, $v(t_0)=v_0$ of (1.2), (1.3) such that

$$\begin{aligned} x(t) &= u(t)+\varepsilon H(t, u(t), v(t), \varepsilon),\\ y(t) &= v(t)+\varepsilon h(t, x(t), \varepsilon). \end{aligned} \tag{4.1}$$

It is sufficient to show that (4.1) holds for $t=t_0$. Substitution $t=t_0$ into (4.1) yields

$$x_0 = u_0+\varepsilon H(t_0, u_0, v_0, \varepsilon), \quad y_0 = v_0+\varepsilon h(t_0, x_0, \varepsilon)$$

and, therefore, $v_0=y_0-\varepsilon h(t_0, x_0, \varepsilon)$. For u_0 we obtain the equation

$$x_0 = u_0+\varepsilon H\big(t_0, u_0, y_0-\varepsilon h(t_0, x_0, \varepsilon), \varepsilon\big). \tag{4.2}$$

This last equation can be represented of form

$$u_0 = P(u_0, \varepsilon) = x_0-\varepsilon H\big(t_0, u_0, y_0-\varepsilon h(t_0, x_0, \varepsilon), \varepsilon\big).$$

From (3.10) it is easy to obtain that for each $\varepsilon\in(0, \varepsilon_1]$ and fixed x_0, y_0 such that $\|y_0-\varepsilon h(t_0, x_0, \varepsilon)\|\leqq\varrho<\dfrac{1}{\varepsilon_1 c}$, P is a contraction mapping of X into itself and so, by the Banach Contraction Mapping Principle, P must have a unique fixed point $u_0\in X$ which is the required solution of (4.2).

Now, we consider the stability problem for (1.1). Using (4.1) we obtain that every solution $x=x(t)$, $y=y(t)$ with $\|y_0-\varepsilon h(t_0, x_0, \varepsilon)\|\leqq\varrho$ can be represented as

$$\begin{aligned} x(t) &= u(t)+\varphi_1(t), \\ y(t) &= \varepsilon h\big(t, u(t), \varepsilon\big)+\varphi_2(t), \end{aligned} \tag{4.3}$$

where $(u(t), \varepsilon h(t, u(t), \varepsilon))$ is a solution lying in the manifold $y=\varepsilon h(t, x, \varepsilon)$; $\varphi_1= =\varepsilon H(t, u(t), v(t), \varepsilon)$, $\varphi_2=v(t)+\varepsilon h\big(t, u(t)+\varepsilon H(t, u(t), v(t), \varepsilon), \varepsilon\big)-\varepsilon h(t, u(t), \varepsilon)$. This and (2.4), (3.8) and (3.15) allow us to write

$$\begin{aligned} \|\varphi_1(t)\| &\leqq \varepsilon a K e^{-(\gamma/\varepsilon)(t-t_0)}\|v_0\|, \\ \|\varphi_2(t)\| &\leqq (1+\varepsilon^2 a\Delta)K e^{-(\gamma/\varepsilon)(t-t_0)}\|v_0\|, \\ \varepsilon\in(0, \varepsilon_1], \quad t &\geqq t_0, \quad v_0 = y_0-\varepsilon h(t_0, x_0, \varepsilon). \end{aligned} \tag{4.4}$$

Assume that $f(t, 0, 0, \varepsilon)=0$, $g(t, 0, 0, \varepsilon)=0$; then $h(t, 0, \varepsilon)=0$ and $F(t, 0, \varepsilon)=0$. By (4.3) and (4.4) we obtain

$$\|x(t)\| \leqq \|u(t)\|+\varepsilon a K e^{-(\gamma/\varepsilon)(t-t_0)}\|v_0\|.$$

$$\|y(t)\| \leqq \varepsilon\Delta\|u(t)\|+(1+\varepsilon^2 a\Delta)K e^{-(\gamma/\varepsilon)(t-t_0)}\|v_0\|, \quad t\geqq t_0.$$

From this last inequalities it easily follows that if the zero solution of (1.2) is stable (asymptotically stable) then the zero solution of (1.1) is stable (asymptotically stable). It is obvious that the instability of the zero solution of (1.2) implies the instability of the zero solution of (1.1).

Now, we can summarize our results in the following

Theorem 4.1. *Let f and g in* (1.1) *be continuous, bounded and satisfy* (2.1), (2.2) *on $R\times X\times B_r\times I_{\varepsilon_0}$; let us assume that the spectrum of the linear bounded operator A satisfies* $\operatorname{Re}\sigma(A)\leqq-2\varkappa<0$. *Then there exist numbers ε_1 and ϱ_1 such that the following assertations are true:*

(i) *For each* $\varepsilon \in (0, \varepsilon_1]$, $\varrho \in (0, \varrho_1)$ *and* t_0 *there exists for* (3.1) *an integral manifold represented by an equation of form* $x_1 = \varepsilon H(t, u, z, \varepsilon)$ *where* H *is an operator-function defined and continuous on* $R \times X \times B_\varrho \times I_{\varepsilon_1}$ *and, moreover,* H *satisfies* (3.8)—(3.10).

(ii) *Every solution* $x = x(t)$, $y = y(t)$ *of* (1.1) *with* $x(t_0) = x_0, y(t_0) = y_0$, $\|y_0 - \varepsilon h(t_0, x_0, \varepsilon)\| \leqq \varrho$ *can be represented of form* (3.11), *where* $u = u(t)$, $u(t_0) = u_0$ *is a solution of* (1.2), u_0 *is a solution of* (4.2); $v = v(t)$ *is a solution of* (1.3) *with* $u = u(t)$, $v(t_0) = v_0 = y_0 - \varepsilon h(t_0, x_0, \varepsilon)$.

(iii) *If* $f(t, 0, 0, \varepsilon) = 0$, $g(t, 0, 0, \varepsilon) = 0$ *and the zero solution of* (1.2) *is stable (asymptotically stable, unstable), then the zero solution of* (1.1) *is stable (asymptotically stable, unstable).*

Note, that in the proof of this theorem we did not use the boundedness of A. So, Theorem 4.1 can be extended onto the system (1.1) with an unbounded operator A, if A is the generator of a strongly continuous linear semigroup $S(t)$ such that $\|S(t)\| \leqq Ke^{-\alpha t}$, $t \geqq 0$.

It should be observed that similar problems for systems with unbounded operators were studied in [2, 6].

The next result shows that, in principle, the operator-function H can be approximated to any degree of accuracy with respect to ε. Let $D(\varepsilon H) = \varepsilon \dfrac{\partial H}{\partial t} + \varepsilon \dfrac{\partial H}{\partial u} F(t, u, \varepsilon) + \dfrac{\partial H}{\partial v}(Av + \varepsilon Z(t, u, \varepsilon H, v, \varepsilon)) - f_1(t, u, \varepsilon H, v, \varepsilon)$. *If* $D(\varepsilon \bar{H}) = O(\varepsilon^{k+1})$ *then* $\|H - \bar{H}\| = O(\varepsilon^k)$.

The idea of the proof of this statement is very simple. Let us introduce a new variable $x_2 = x_1 - \varepsilon \bar{H}(t, u, z, \varepsilon)$; then for u, x_2, z we obtain the following system

$$\dot{u} = f(t, u, \varepsilon),$$

$$\dot{x}_2 = f_2(t, u, x_2, z, \varepsilon),$$

$$\varepsilon \dot{z} = Az + \varepsilon Z(t, u, x_2 + \varepsilon \bar{H}, z, \varepsilon),$$

where $f_1 = f_1(t, u, x_2 + \varepsilon \bar{H}, z, \varepsilon) - f_1(t, u, \varepsilon \bar{H}, z, \varepsilon) - \varepsilon \dfrac{\partial \bar{H}}{\partial z}[Z(t, u, x_2 + \varepsilon \bar{H}, z, \varepsilon) - Z(t, u, \varepsilon \bar{H}, z, \varepsilon)]$, $\varepsilon \bar{H} = \varepsilon \bar{H}(t, u, z, \varepsilon)$. This last system has an integral manifold $x_2 = \varepsilon H_{k+1}(t, u, z, \varepsilon)$ such that $H_{k+1} = O(\varepsilon^k)$. It means that the system (3.1) has the integral manifold $x_1 = \varepsilon H(t, u, z, \varepsilon) = \varepsilon \bar{H}(t, u, z, \varepsilon) + O(\varepsilon^{k+1})$.

In many problems, H can be found as asymptotic expansion

$$\varepsilon H = \varepsilon H_1(t, u, v) + \ldots + \varepsilon^k H_k(t, u, v) + O(\varepsilon^{k+1})$$

from the equation $D(\varepsilon H) = 0$. Note, that u_0 can be found as asymptotic expansion

$$u_0 = u_0(\varepsilon) = u_0^0 + \varepsilon u_0^1 + \ldots + \varepsilon^k u_0^k + O(\varepsilon^{k+1})$$

from (4.2). It is easy to see that $u_0^0 = x_0$, $u_0^1 = -H_1(t_0, x_0, y_0)$.

5. Linear systems. Consider the following system

$$\dot{x}_1 = A_{11}x_1 + A_{12}x_2 + f_1, \tag{5.1}$$
$$\varepsilon\dot{x}_2 = A_{21}x_1 + A_{22}x_2 + f_2,$$

where x_i, $f_i = f_i(t, \varepsilon)$ vary in the Banach space X_i, and $A_{ij} = A_{ij}(t, \varepsilon)$ are operator-functions $A_{ij}\colon X_j \to X_i$ $(i, j = 1, 2)$. Assume A_{ij} and f_i to have high order continuous and bounded derivatives with respect to t and ε, for $t \in R$, $\varepsilon \in [0, \varepsilon_0]$. Therefore, they can be represented as asymptotic expansions

$$A_{ij} = A_{ij}^{(0)}(t) + \varepsilon A_{ij}^{(1)}(t) + \ldots + \varepsilon^k A_{ij}^{(k)}(t) + O(\varepsilon^{k+1}),$$
$$f_i = f_i^{(0)}(t) + \varepsilon f_i^{(1)}(t) + \ldots + \varepsilon^k f_i^{(k)}(t) + O(\varepsilon^{k+1})$$

with smooth and bounded coefficients.

Let us suppose that the family $A_{22}^{(0)}(t)$, $t \in R$, is compact, the spectrum $\sigma(A_{22}^{(0)})$ of $A_{22}^{(0)}(t)$ satisfies the inequality

$$\operatorname{Re}\sigma(A_{22}^{(0)}) \leqq -2\alpha < 0, \quad t \in R \tag{5.2}$$

and there exists bounded operator $[A_{22}^{(0)}]^{-1}$. Under such assumptions there exists a transformation

$$x_1 = u + \varepsilon H(t, \varepsilon)v,$$
$$x_2 = v + L(t, \varepsilon)x_1 + l(t, \varepsilon) = (I + \varepsilon LH)v + Lu + l(t, \varepsilon),$$

analogous to (3.11) for the linear case. The new variables u, v satisfy the equations

$$\dot{u} = (A_{11} + A_{12}L)u + f_1 + A_{12}l, \tag{5.3}$$

$$\varepsilon\dot{v} = (A_{22} - \varepsilon LA_{12})v. \tag{5.4}$$

The operator-functions L, H and the function l can be found from the equations

$$\varepsilon\dot{L} + \varepsilon L(A_{11} + A_{12}L) = A_{21} + A_{22}L, \tag{5.5}$$

$$\varepsilon\dot{H} + H(A_{22} - \varepsilon LA_{12}) = \varepsilon(A_{11} + A_{12}L)H + A_{12}, \tag{5.6}$$

$$\varepsilon\dot{l} + \varepsilon Lf_1 = (A_{22} - \varepsilon LA_{12})l + f_2 \tag{5.7}$$

as asymptotic expansions $L = L^{(0)}(t) + \varepsilon L^{(1)}(t) + \ldots + \varepsilon^k L^{(k)}(t) + O(\varepsilon^{k+1})$,

$$H = H^{(0)}(t) + \varepsilon H^{(1)}(t) + \ldots + \varepsilon^{k-1}H^{(k-1)}(t) + O(\varepsilon^k),$$
$$l = l^{(0)}(t) + \varepsilon l^{(1)}(t) + \ldots + \varepsilon^k l^{(k)}(t) + O(\varepsilon^{k+1}).$$

It is a straightforward computation to obtain expressions for $L^{(i)}$, $H^{(i)}$, $l^{(i)}$ from (5.5)—(5.7).

Note that L is a bounded solution of the Riccati equation (5.5) on R and, therefore, satisfies the integral equation

$$L(t, \varepsilon) = \frac{1}{\varepsilon}\int_{-\infty}^{t} U(t, s, \varepsilon)\big[A_{21}(s, \varepsilon) - \varepsilon L(s, \varepsilon)\big(A_{11}(s, \varepsilon) + A_{12}(s, \varepsilon)L(s, \varepsilon)\big)\big]\,ds,$$

where U is the evolutional operator of the equation $\varepsilon\dot{x}_2 = A_{22}x_2$. Using (5.2) and the compactness of $A_{22}^{(0)}(t)$ we obtain

$$\|U(t, s, \varepsilon)\| \leqq Ke^{-(\alpha/\varepsilon)(t-s)}, \quad -\infty < s \leqq t < \infty. \tag{5.8}$$

For H and l we have the exact expressions

$$H = -\frac{1}{\varepsilon}\int_t^\infty V(t, s, \varepsilon)A_{12}(s, \varepsilon)W(s, t, \varepsilon)\,ds,$$

$$l = \frac{1}{\varepsilon}\int_{-\infty}^t W(t, s, \varepsilon)[f_2(s, \varepsilon) - \varepsilon L(s, \varepsilon)f_1(s, \varepsilon)]\,ds,$$

where V is the evolutional operator of the equation $\dot{x}_1 = (A_{11} + A_{12}L)x_1$ and W is the one of the equation $\varepsilon\dot{x}_2 = (A_{22} - \varepsilon LA_{12})x_2$. The improper integrals here converge by virtue of (5.8). As earlier, the stability of (5.3) is equivalent to the stability of (5.1).

In conclusion it should be noted that the stability and decomposition problems for finite dimensional systems were considered in [7].

References

[1] Ю. А. Митропольский, О. Б. Лыкова, *Интегральные многообразия в нелинейной механике*, Наука (Москва, 1973).

[2] D. Henry, *Geometric theory of semilinear parabolic equations, Lecture Notes in Math.*, 840, Springer-Verlag (Berlin—New York, 1981).

[3] В. В. Стрыгин, В. А. Соболев, Влияние геометрических и кинетических параметров и диссипации энергии на устойчивость ориентации спутников с двойным вращением, *Космические Исследования*, 3 (1976), 366—371.

[4] В. А. Соболев, К теории интегральных многообразий одного класса систем сингулярно возмущенных уравнений, Приближенные методы исследования дифференциальных уравнений и их приложения, *Куйбышевский Университет* (1980), 124—147.

[5] A. Kelley, The stable, center-stable, center, center-unstable and unstable manifolds, *J. Differential Equations*, 3 (1967), 546—570.

[6] О. Б. Лыкова, Принцип сведения в банаховом пространстве, *Укр. Матем. Ж.*. 4 (1971), 464—471.

[7] V. A. Sobolev, Integral manifolds and decomposition of singularly perturbed systems, *Systems Control Lett.*, 5 (1984), 169—179.

DEPARTMENT OF DIFFERENTIAL EQUATIONS
KUIBYSHEV STATE UNIVERSITY
KUIBYSHEV, U.S.S.R.

Acta Sci. Math., 51 (1987), 501

Erratum to "A generalization of a theorem of Dieudonné for k-triangular set functions"[1])

E. PAP

The following unfortunate mistakes appeared in this note.

1. Lemma 1 is true in the given form only for $k \geqq 1$ (what is sufficient for the proof of Theorem 2). In the case $0<k<1$ only

$$\mu\Big(\bigcup_{j=1}^{\infty} O_j\Big) \leqq \mu(O_1)+k \sum_{j=2}^{\infty} \mu(O_j)$$

holds instead of

$$\mu\Big(\bigcup_{j=1}^{\infty} O_j\Big) \leqq k \sum_{j=1}^{\infty} \mu(O_j);$$

consequently the proof also needs some appropriate modifications.

2. The last two sentences of the proof of Theorem 2 on the page 165 should be replaced by "Applying the preceding proof for $k>1$ we can verify Theorem 2 for this case, too."

3. On the page 165, 10th row from below, "$|\nu(A)-\nu(A')|<\varepsilon$" should be replaced by "$|\mu(A)-\mu(A')|<\varepsilon$".

4. On the page 162, in Lemma 2, "$\|x_{ii}\| \geqq \delta > 0$" should be replaced by "$\|x_{ii}\| > 0$".

INSTITUTE OF MATHEMATICS
ILIJE DJURIČIĆA 4
21 000 NOVI SAD, YUGOSLAVIA

[1]) *Acta Sci. Math.*, **50** (1986), 159—167.

Acta Sci. Math., 51 (1987) 503—529

Bibliographie

Donald J. Albers—G. L. Alexanderson—Constance Reid, International Mathematical Congresses. An illustrated History 1893—1986, Revised Edition, 64 pages, Springer-Verlag, New York—Berlin—Heidelberg, 1987.

This is a nice picture book covering the "World Congress" in Chicago, 1893, and all the International Congresses of Mathematicians beginning with the first in Zürich, 1897, and ending with the latest, the twentieth one in Berkeley, 1986. Each congress receives two pages (plus, in connection with his famous address in Paris, 1900, Hilbert himself an extra two) with three to five photos or drawings of illustrious mathematicians, alone or together, who played outstanding roles at the given congress or of characteristic buildings. It is the pictures that make the book nice. Not much can be said about the text. Using (sometimes fragmentary and irrelevant) citations, it tries to give a "feeling" of the given congress. The aspect of "the first American" of the three American authors pops up in an inordinate frequency. There are nine pages with the photographs of all Fields Medalists and a list of all the plenary lectures from 1893 to 1986.

The original edition of the book has been distributed during the Berkeley Congress. As it is made clear in Czesław Olech's Opinion [The Mathematical Intelligencer 9 (1987), 36—37], the present revised edition has become necessary mainly because of protests, including his own in the same Opinion, against "an unfair description of the previous Congress in Warsaw" in the original edition.

Sándor Csörgő (Szeged)

Hans Wilhelm Alt, Lineare Funktionalanalysis. Eine Anwendungsorientierte Einführung (Hochschultext), IX+292 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

The material of this book is based on the lectures in linear functional analysis held some years ago at Bonn University for students in the fifth semester. The book consists of an introductory part, ten chapters and three supplements. The text gives an introduction to the study of Banach and Hilbert spaces, linear functionals and the most important classes of linear operators. The supplements deal with measures, integrals and Sobolev spaces. At the end of the book the spectral theory of compact normal operators can be found. All of the chapters end with exercises and their solutions.

The book is highly recommended to students who are interested in functional analysis and its applications in physics.

L. Gehér (Szeged)

Analytic Theory of Continued Fractions, Proceedings Pitlochry and Aviemore, Scotland, 1985. Edited by W. Y. Thron (Lecture Notes in Mathematics, 1199), III+299 pages, Springer-Verlag, Berlin—Heidelberg—New York—London—Paris—Tokyo, 1986.

In the last decades special attention is paid to the new results in the theory of continued fractions. The success of the workshop held in Loen, Norway in 1981 speaks for itself. Therefore a second workshop was arranged in Pitlochry and Aviemore, Scotland in 1985. This proceedings volume is thus the successor of Lecture Notes in Mathematics, Vol. 932.

This volume contains a survey article entitled "Schur fractions, Perron—Carathéodory fractions and Szegő polynomials" by W. B. Jones, O. Njåstad and W. J. Thron and thirteen original research papers. The introduction of the survey article presents historical comments with a limited list of references. Two main topics are treated in the research papers. The first one is the convergence theory of continued fractions, the second one is the investigation of various types of continued fractions useful in solving Stieltjes, Hamburger and trigonometric moment problems. In general the articles give applications from different branches of mathematics. Perhaps the volume would have been more interesting if some of the papers had contained open questions or conjectures in explicit form.

L. Pintér (Szeged)

D. F. Andrews—A. M. Herzberg, Data: A Collection of Problems from Many Fields for the Student and Research Worker (Springer Series in Statistics), XX+442 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1985.

The ultimate aim of Statistics is to provide methods and tools for the analysis of real data. In this very useful book the authors collected a great number of concrete real data sets. There are seventy-one concrete problems presented in the book with data sets given in 100 tables and 11 figures. For each data set the source or sources of the data are given with a description by the authors or by a contributor who supplied the data. No direct reference to any particular type of analysis is given: the student or researcher may try his/her arsenal of tools for the analysis. Some of the data sets are well known, such as the last century data on the number of deaths by horsekicks in the Prussian Army (the name of L. von Bortkiewicz, to whom the horsekicks belong, is written all four times erroneously as Bortkewitsch: the authors did not always go back to the original source), the Fisher *Iris* data (with which the collection starts), the Canadian lynx trappings data, the coal-mining disasters data, the Federalist Papers data, or the Stanford heart transplant data. The majority of the data sets, however, is relatively new and unknown to a wider statistical public and is very interesting. The authors have invested a great care into the organizational work and the uniformization of the presentation. The result is a splendid volume of great interest, completely unique in its kind and a great service to the international statistical community.

Sándor Csörgő (Szeged)

Astrophysics of Brown Dwarfs. Proceedings, Fairfax, 1985, Edited by M. C. Kafatos, R. S. Harrington and S. P. Maran, 276 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1986.

This book includes the scientific papers presented at a Workshop, held at George Mason University, Fairfax, Virginia, in 1985.

The term "brown dwarf" is a boundary class of stars that are partially supported by nuclear burning and partially by thermal cooling. These objects, "super-Jupiters", bridge the range of masses between planets and normal stars. The mass of the brown dwarfs is less than 0.08 solar mass, and their surface temperatures are expected to be 1000—2000 K.

The book consist of two parts. In the first one the experimental works are presented. The observation of brown dwarfs is very difficult, owing to the faintness of these objects, and only one definite object has been found so far. The articles about the systematic search for a very nearby solar companion ("Nemezis" or "Shiva") are especially interesting. In the second, theoretical part aspects of planetary interior physics are extended to higher densities and pressures. The brown dwarfs are of great importance in the stellar evolution theory.

Presently, in the topic of the brown dwarfs there are more theories than objects. However, with the help of space telescopes and infrared techniques the detection of numerous stars of this kind are likely to be discovered soon.

K. Szatmáry (Szeged)

Werner Ballmann—Michael Gromov—Viktor Schroeder, Manifolds of Nonpositive Curvature (Progress in Mathematics, Vol. 61), 263 pages, Birkhäuser, Boston—Basel—Stuttgart, 1985.

This book is based on four lectures given by Mikhael Gromov in February 1981 at the College de France in Paris. The presentation is due to Viktor Schroeder who made a coherent text by writing down all the proofs in complete detail. He also added some background material to Lecture I and exposed the basic facts on symmetric spaces needed for Lecture IV. The articles included in this book summarize the recent progress of the theory of manifolds of nonpositive curvature. The lectures are: I. Simply connected manifolds of nonpositive curvature, II. Groups of isometries, III. Finiteness theorems, IV. Strong rigidity of locally symmetric spaces. The further papers are: Manifolds of higher rank (by W. Ballmann); Finiteness results for nonanalytic manifolds, Tits metric and the action of isometries at infinity, Tits metric and asymptotic rigidity, Symmetric spaces of non-compact type (by V. Schroeder).

Péter T. Nagy (Szeged)

J. L. Berggren, Episodes in the Mathematics of Medieval Islam, 97 figures and 20 plates, IX + 197 pages, Springer-Verlag, New York—Berlin—Heidelberg—London—Paris—Tokyo, 1986.

Many people know today that some mathematical terminologies have their origin in the medieval Islamic (or Arabic) civilization, such as algebra and algorithm. It is also well known that several ancient Greek mathematical and philosophical works became known for the Renaissance Europe via Arabic translations, but we know very little about the original Islamic mathematics: "no textbook on the history of mathematics in English deals with the Islamic contribution in more than a general way" as the author writes. This is unfortunate because they made important contributions to the development of decimal arithmetic, plane and spherical trigonometry, algebra (e.g. solving cubic equations) and interpolation and approximation of roots of equations.

The aim of the present book is to make an attempt to fill this gap. In spite of that it is not and cannot be a "General History of Mathematics in Medieval Islam", we are sure that this volume is a very important contribution to the subject.

In an introductory chapter the reader gets acquainted with the Islam's reception of foreign science, the four most famous Muslim scientists: Al-Khwārizmī, Al-Bīrūnī, 'Umar al-Khayyāmī and Al-Khāshī, and the most important sources. The other chapters deal with arithmetic, geometrical constructions, algebra, trigonometry and spherics. Each chapter is followed by a set of exercises and a bibliography.

We recommend this book primarily for students and teachers of mathematics, but everybody interested in the history of mathematics can read this well-illustrated book with joy.

Lajos Klukovits (Szeged)

Arthur L. Besse, Einstein Manifolds (Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, Bd. 10), XII+510 pages, Springer-Verlag, Berlin—Heidelberg—New York—London—Paris—Tokyo, 1987.

This book is intended to be a complete reference book of the differential theory of Einstein—Riemannian manifolds. In the author's opinion the Einstein metrics are the best candidates for nicest geometric structures on a given manifold which are very natural generalizations of Euclidean and classical non-Euclidean spaces. These manifolds have close relations with the geometries of constant curvature, but they have non-necessarily transitive isometry groups and thus their geometric properties can reflect characteristic non-homogeneous features. The indefinite semi-Riemannian analogies of these spaces are of basic importance in the modern physical space-time theory.

The book provides a self-contained treatment of many important topics of Riemannian geometry presented in a textbook for the first time, such as Riemannian submersions, Riemannian functionals and their critical points, the theory of Riemannian manifolds with distinguished holonomy group and Quaternion-Kähler manifolds. The central chapters of the book are devoted to the study of questions related to the Calabi conjecture made in 1954, whose solution given by S. T. Yau and T. Aubin in 1976 yields a large class of non-homogeneous compact Einstein manifolds. Corresponding to this conjecture the main problems treated in this book are related to the existence and uniqueness questions and principally to finding interesting examples of Einstein metrics. The book contains the formulation of the main open problems of this theory.

This excellent book is warmly recommended to everyone interested in Riemannian geometry and its applications in mathematical physics.

Péter T. Nagy (Szeged)

J. Bliedtner—W. Hansen, Potential Theory, An Analytic and Probabilistic Approach to Balayage (Universitext), XIII+434 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986.

Recently much attention has been paid to stochastic processes in modern analysis. The classical example is potential theory. Suppose we want to solve Dirichlet's problem in a domain U with smooth boundary ∂U, and $f\colon \partial U \to R$ is the continuous function we want to extend to U to a harmonic function. If $\{X_t^{(x)}\}$ is a two-dimensional Brownian motion starting at $x \in U$ and $\eta = \eta(\omega)$ is the first time when $\{X_t^{(x)}(\omega)\}$ hits ∂U, then $u(x) = E(f(X_\eta^{(x)}))$, the expectation of the random variable $f(X_{\eta(\omega)}^{(x)}(\omega))$, solves the problem. In fact, it follows from the Markov property of $\{X_t\}$ that u possesses the mean value property in U and it is clear that if $x \in U$ is close to $y \in \partial U$ then $\{X_t^{(x)}\}$ will hit a fixed neighbourhood (on ∂U) of y with high probability, so u has f as its boundary function.

The same idea works in many other classical problems. The book under review is devoted to the study of general balayage theory which is at the heart of these applications. The central objects are the so called balayage spaces which are certain closed function cones with the property that if u, v', v'' are in the space and $u \leq v' + v''$, then u has a representation $u = u' + u''$, $u' \leq v'$, $u'' \leq v''$. These spaces occur in different chapters in different equivalent forms as families of harmonic kernels, sub-Markov semigroups and as Hunt processes (regularized Markov processes).

The authors were very careful to clarify the abstract notions and results through concrete examples such as classical potential theory, Riesz potentials, discrete potential theory, translation on R and heat conduction in R^n. These relax the abstract setting; still one may encounter the usual drawbacks of too much generality when trying to use the book as a "Universitext". I feel that it is more appropriate to recommend this work to those who have past experience with both classical potential theory and stochastic processes. Then the new examples and different approaches of the book can be refreshing.

For further orientation here is a characteristic list of section headings: Classical Potential Theory, Function Cones, Choquet Boundary, Laplace Transforms, Supermedian Functions, Semigroups and Resolvents, Hyperharmonic Functions, Harmonic Kernels, Minimum Principle and Sheaf Properties, Markov Processes, Stopping Times, Balayage of Functions and Measures, Dirichlet Problem, Partial Differential Equations, Bauer Spaces, Semi-Elliptic PDE, Elliptic-Parabolic Differential Operators.

Vilmos Totik (Szeged)

Umberto Bottazzini, The Higher Calculus: A History of Real and Complex Analysis from Euler to Weierstrass, 332 pages, Springer-Verlag, New York—Berlin—Heidelberg—London—Paris—Tokyo, 1986.

It was just 300 years ago that Newton published his monumental work *Principia mathematica philosophiae naturalis,* in which he founded differential and integral calculus and revolutionized the science of functions, mathematical analysis. Thereafter an enormous development had started in this area of mathematics, whose summit was reached in the 19th century, often mentioned in the history of mathematics as the century of analysis. This excellent book gives a detailed account of the history of this splendid period.

Can a book on the history of mathematics be interesting for wide circles of readers? Having read Bottazzini's book it is easy to answer: yes, it can. The author does not restrict himself to dull reviewing the results but acquaints the reader with the outstanding mathematicians of the area as people with emotions. For example, we can read the letter of a 24-year-old mathematician to his old teacher written after his arrival at Paris in 1826: "Up to now I have only made the acquaintance of Legendre, Cauchy, and Hachatte, plus a few secondary but very able mathematicians, ... Legendre is an extremely amiable man, but unfortunately "as old as stones". Cauchy is crazy and there is nothing to be done with him, even though at the moment he is the mathematician who knows how mathematics must be done. His works are excellent, but he writes in a very confusing way. At first I understood virtually nothing of what he wrote, but now it goes better... Poisson, Fourier, Ampère, etc. etc. occupy themselves with nothing other than magnetism and other physical matters... Everyone works by himself without interesting himself in others. Everyone wants to teach and no one wants to learn. The most absolute egoism reigns everywhere." These are rather hard words, but the young man was named Abel. Nevertheless, the book is written mainly about mathematics itself and gives the milestones of the history of such big problems as the solution of the equations of vibrating string and of heat diffusion, expansion of functions into trigonometric series, the foundation of the theory of complex functions, etc. Special attention is paid to the development of the concept of a function. Who could think without reading the history of mathematics that the contemporary definition of a function, which is tought in every elementary school today and seems born with us like our instincts, is a result of a long process of thinking and debates, and that Euler still defined a function as follows: "A function of a variable quantity is an analytic expression composed in any way from this variable quantity and from numbers or constant quantities."

This well-written book will be a very valuable and enjoyable reading not only for students and experts of the history of mathematics, but for every student learning calculus and for every researcher in mathematics, since Poincaré is absolutely right when saying: "The true method of foreseeing the future of mathematics is to study its history and its actual state."

L. Hatvani (Szeged)

César Camacho—Alcides Lins Neto, Geometric Theory of Foliations, 205 pages, Birkhäuser, Boston—Basel—Stuttgart, 1985.

The theory of foliations is a part of differential topology investigating the decompositions of manifolds into a union of connected, disjoint submanifolds of the same dimension. The origin of this subject is the geometric theory of differential equations that has begun with the works of Painlevé, Poincaré and Bendixson in the last century. The authors say in the introduction: »The development of the theory of foliations was however provoked by the following question about the topology of manifolds proposed by H. Hopf in the 1930's: "Does there exist on the Euclidean sphere S^3 a completely integrable vector field, that is, a field X such that $X \cdot \operatorname{curl} X = 0$?" By Frobenius' theorem this question is equivalent to the following: "Does there exist on the sphere S^3 a two-dimensional foliation?"«

The present book which is a translation of the original Portuguese edition published in Brasil in 1980 has the purpose to give an introduction to the subject. The first four chapters treat the basic notions and properties of foliations (Differentiable Manifolds, Foliations, The Topology of Leaves, Holonomy and the Stability Theorems). Chapter V discusses the relations between foliations and fiber bundles. Chapter VI is devoted to the proof of Haefliger's theorem about analytical foliations of codimension one. Chapter VII contains the proof of Novikov's theorem on the existence of a compact leaf of a C^2 codimension one foliation on a compact three-dimensional manifold with finite fundamental group. Chapter VIII deals with foliations induced by group actions. There is an appendix containing the proof of Frobenius' theorem.

The book is highly recommended to anyone interested in differential topology and familiar with the material of standard courses on analysis, topology and geometry.

Péter T. Nagy (Szeged)

Leonard S. Charlap, Bieberbach Groups and Flat Manifolds (Universitext), X+242 pages, Springer-Verlag, New York—Berlin—Heidelberg—London—Paris—Tokyo, 1986.

The theory of the Euclidean space form problem, treated in this book, is originated from Hilbert's famous 18th problem about the classification of discrete Euclidean rigid motion groups with fundamental domains, or what is the same, of crystallographic groups. The early solution of this problem, given by L. Bieberbach in about 1910, can be summarized in modern language in the following way. The fundamental group of a compact flat Riemannian manifold is a Bieberbach group, i.e. a torsionfree group having a maximal abelian subgroup of finite index which is free abelian. The manifolds with isomorphic fundamental groups are affinely equivalent, the number of their equivalence classes is finite.

The author of this book has developed the classification theory of Euclidean space forms based on the description of the linear holonomy group of flat Riemannian connections in the early 1960's. The purpose of the present treatment is to give a selfcontained introduction and at the same time a reference book on this topic. Chapter I contains the presentation of Bieberbach's classical theory. Chapter II gives an elementary introduction to Riemannian geometry including the notion and fundamental properties of linear holonomy groups. There is given a formulation of Bieberbach's results in the language of differential geometry. Chapter III deals with the algebraic classification of Bieberbach groups. It is finished with the proof of the Anslander—Kuranishi theorem saying that any finite group is the holonomy group of a compact flat manifold. Chapter IV is devoted to the author's principal results about the space forms whose holonomy group has prime order. Chapter V discusses the properties of automorphism groups of flat manifolds.

The general results are illustrated with many examples. Open problems, conjectures, counterexamples and results related to the theory of nonflat manifolds are formulated throughout. This very nice book is really interdisciplinary, it uses tools of differential topology and geometry, algebraic number theory, cohomology of groups and integral representations.

Péter T. Nagy (Szeged)

Coherence, Cooperation and Fluctuations, Edited by F. Haake, L. M. Narducci and D. F. Walls (Cambridge Studies in Modern Optics, 5), VIII+456 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1986.

In 1963 Roy Glauber laid down the fundamentals of quantum optics by introducing the quantum concept of coherence and the coherent states of the radiation field. This fact, which is already a part of the history of physics, justifies the decision to devote one of the volumes of the Cambridge Studies in Modern Optics to the works honouring the 60-th birthday of R. Glauber.

In spite of the series title, besides optics, there are papers on statistical physics and nuclear physics too, as Glauber himself contributed also to the development of the latter fields with essential and fundamental results. Among the authors of the invited papers we find L. Kadanoff, J. Langer, H. Feshbach, F. T. Arecchi, N. Bloembergen, S. Haroche, L. Mandel, R. Pike, M. O. Scully. The majority of the 33 papers deal with quantum optics, and reading them one may really learn what is in focus at present time in the field of optical coherence, cooperation and fluctuations. The two other topics are treated less comprehensively in this volume. To have at hand the roots of the ideas presented in the book, the editors included the reprints of the 4 classic papers of R. Glauber: the two about quantum coherence, the time dependent statistics of the Ising model and the one about the optical model of nuclear reactions.

The book is recommended mainly for research workers in the areas of nonrelativistic field theory and quantum optics.

M. G. Benedict (Szeged)

M. Crampin—F. A. E. Pirani, Applicable Differential Geometry (London Mathematical Society Lecture Notes Series, 59) 385 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1986.

Traditional courses in differential geometry contain first the elementary theory of curves and surfaces in a Euclidean space and thereafter the notion of a differentiable manifold and the theory of differential geometric structures on it. Such an approach has the disadvantage that the notion of fibre bundles and the general theory of connections and Lie group actions can be treated only in lecture courses for final year graduate or postgraduate student audiences. But these techniques are needed in the modern applications of differential geometry in the foundation of mechanics, gauge field theories and gravitation theory.

The present book gives an introduction to these methods of differential geometry on the level of beginning graduate students. "The essential ideas are first introduced in the context of affine space; this is enough for special relativity and vectorial mechanics. Then manifolds are introduced and the essential ideas are suitably adapted; this makes it possible to go on to general relativity and canonical mechanics. The book ends with some chapters on bundles and connections which may be useful in the study of gauge fields and such matters." The treatment is illustrated with many examples motivated by the applications in mathematical physics. Each chapter is concluded with a brief summary of its contents.

The reviewer thinks that this excellent introduction will be especially useful if it is supplemented with parallel courses on analytical mechanics and relativity theory.

Péter T. Nagy (Szeged)

Luc Devroye, Non-Uniform Random Variate Generation, XVI+843 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

The importance of this comprehensive work can hardly be overemphasized. A large amount of today's research in statistics, operations research and computer science depends upon large scale Monte Carlo computer simulation. Also, this is almost the sole means of investigation in certain applied fields in engineering, experimental and even theoretical physics and chemistry, the life sciences and technology, but such "esoteric pure mathematics" as number theory is not devoid of Monte Carlo experimentation either. Yet all these depend upon sequences of numbers or vectors generated on the computer which are to be viewed as independent realizations of a random variable or vector with a prescribed distribution. Then you apply one of the greatest things of Nature (or, put it with less euphemism, a trivial fact of probability theory), the law of large numbers, and, modulo the problem at hand, you are done.

Now, borrowing some expressions from the characteristically lively language of Devroye, the "story has two halves". Any machine that can be called a computer nowadays sports with a random number generator that is claimed to be capable to produce sequences of independent random variables uniformly distributed on (0, 1). This will of course never be the case, and the theoretical and practical aspects of this problem belong to the circle of the deepest common puzzles of probability theory and algorithm theory. However, that machines do indeed have such generators is becoming more and more reasonable an assumption together with another, theoretically impossible assumption that the computer can store and manipulate real numbers.

Based on these two assumptions, the book is about "the second half of the story": how to generate random numbers with a prescribed non-uniform distribution most efficiently? The efficiency of a procedure is measured by the complexity of the algorithm which produces one such number. This notion is achieved by the author's third assumption that the fundamental operations in the computer (addition, multiplication, division, compare, truncate, move, generate a uniform random variate, exp, log, square root, arctan, sin and cos) all take one unit of time, and the complexity is simply the required time. The algorithms themselves are then investigated by providing lower and upper bounds for the their expected complexity or for the tails of their distribution.

Following an introduction into a few basic probabilistic facts, Chapters 2 and 3 present the general principles of random number generation such as the inversion, the rejection, the composition, the acceptance-complement, alias and table look-up methods and their various combinations, while Chapters 4, 5, 8 and 14 describe a bewilderingly vast amount of specialized algorithms. The procedures are then applied in Chapters 7, 9 and 10 for the generation of random numbers from the most important continuous and discrete parametric families of distributions or from large families of distributions given by a qualitative property of the density such as log-concavity, monotonicity, or unimodality. The whole Chapter 11 is on random vector generation and Chapter 6 is devoted to the generation of the Poisson and related random processes. So these two chapters create special dependence structures already, and Chapters 12 and 13 go on further in this direction, to the generation of various sampling without replacement plans and to the generation of random permutations, binary and free trees, partitions and graphs. Finally, Chapter 15 presents the Knuth-Yao theory of discrete distribution generating trees in a random bit model in which, instead of *Uniform* (0, 1) *random numbers*; *Binomial* (1, 1/2) random numbers are available.

The hundreds of generation algorithms in the book are all written as PASCAL programs and are intelligible without knowing anything special about this language. In fact, the book is completely independent of today's computer and programming technology and I am sure that the author's hopes that "the text will be as interesting in 1995 as in 1985" are entirelly well-developed.

The author outlines a course in computer science and another one in statistics that can be based on the book, moreover he proposes a "fun reading course on the development and use of inequalities". My own random fun reading course turned out to be most gratifying and enjoyable. Wherever opens, it is difficult to put down the book which is bound to become to be the basic reference in non-uniform random number generation. It contains a good number of new results and an enormous amount of knowledge from probability and statistics, computer science, operations research and complexity and algorithm theory, blended and arranged by masterly scholarship. Congratulations Luc!

Sándor Csörgő (Szeged)

Differential Equations in Banach Spaces, Proceedings of a Conference held in Bologna, July 2—5, 1985. Edited by A. Favini and E. Obrecht (Lecture Notes in Mathematics, 1223), VIII+299 pages, Springer-Verlag, Berlin—Heidelberg—New York—London—Paris—Tokyo, 1986.

When modelling the evolution in time of a physical system we have to decide how to describe the position of the system at an instant of time. For example, for a finite system of particles in classical mechanics a position is a point of R^n, while in the model of the vibrating string or the heat conduction problem we use to this end functions in $C^2([0, l]; R)$. Respectively, the model equation will be a system of ordinary differential equations and a partial differential equation. However, partial differential equations can be also considered as ordinary ones of the from $\dot{u}=Au$ which are written in the Banach space $C^2([0, l]; R)$ as a state space, and A is a differential operator in this space. This unification was inspired by the fact that the basic concepts and methods of the theory of ordinary differential equations (eigenvalue, Jordan form, exponents of a matrix, spectral theory, calculus of functions) have been developed for operators in Banach spaces by linear (and recently nonlinear) functional analysis. As it is also shown by these proceedings, the approach to differential equations as abstract equations in Banach spaces is a fruitful and very rapidly developing field. Among the topics discussed at the Conference are: regular and singular evolution equations, both linear and nonlinear, of parabolic and hyperbolic type, integro-differential equations, semi-group theory, control theory, wave equations, transmutation methods and fuchsian differential equations.

L. Hatvani (Szeged)

B. A. Dubrovin—A. T. Fomenko—S. P. Novikov, Modern Geometry — Methods and Applications: Part II. The Geometry and Topology of Manifolds, 430+XV pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1985.

This is the English translation, by Robert G. Burns, of the original Russian edition, Nauka, Moscow, 1979. The present book is the second volume of a whole series in which the authors' main aim is the modernization of the teaching of differential geometry at universities. From this point of view this work can be considered as one of the best texts which acquaints the readers with a large part of modern differential geometry in a very clear and didactical style.

This second volume is devoted mainly to differential topology. After the elementary study of real and complex manifolds, Lie groups, homogeneous spaces, the exposition turns to the Sard theorem and related fields such as Morse theory, embeddings and immersions, the degree of mappings and the intersection index of submanifolds. Furthermore two chapters deal with the fundamental groups and homotopy groups of manifolds. After these the theory of fibre bundles, connections, foliations and dynamical systems is developed. The last chapter deals with general relativity and also with Yang—Mills theory whose comprehensible survey has been absent from the literature.

This well-written excellent monograph can be highly recommended to students, mathematicians and users interested in modern differential geometry.

Z. I. Szabó (Budapest)

Sir Arthur Eddington, Space, Time and Gravitation. An outline of the general relativity theory (Cambridge Science Classics Series), XII+218 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1987.

This classic book on the general theory of relativity was published first in the exciting days of 1920, soon after the first objective tests of the new theory assumed historico-scientific values. The reader can understand how Sir Arthur Eddington, the creative participant of the development of this theory in mathematics, physics and philosophy, saw the problems of space, time and gravitation. This new reprint, which is the twelfth in a sequence, includes a foreword by Sir Hermann Bondi, describing the place of this book in its historical and scientific context. He says: "How does his writing strike us now, some sixty years after it first appeared in print? The beautiful English is as good as ever, the subject matter, the theories of relativity and gravitation, have not suffered relegation to the backburner, but are as integral a part of physics as in his day. Thus his book is still very good and very relevant." Everyone interested in the development of new ideas and viewpoints in the sciences will enjoy this book.

Péter T. Nagy (Szeged)

K. J. Falconer, The Geometry of Fractal Sets (Cambridge Tracts in Mathematics, 85), XIV+162 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1986.

From the introduction of the author: "Recently there has been a meteoric increase in the importance of fractal sets in the sciences. Mandelbrot pioneered their use to model a wide variety of scientific phenomena from the molecular to the astronomical .... Sets of fractional dimension also occur in diverse branches of pure mathematics."

This widespreading applicability aroused both scientists' and mathematicians' interest in fractals. The aim of this book is to give a rigorous mathematical treatment of the geometrical properties of sets of both integral and fractional Hausdorff dimension, and the author unites into a theory the complete collection of these results, which have previously been available only in technical papers.

The first chapter contains a very good general measure theoretic introduction, the definition of Hausdorff measure and dimension and basic covering results. There is an emphasis on the Vitali covering theorem which will be often used. In some "simple" cases the Hausdorff dimension and measure are calculated. The next three chapters discuss the density properties and existence of tangents. The notion of local densities are similar to the Lebesgue case but there is no analogue of the Lebesgue density theorem. It is proved that only the integral dimensional sets can be regular and in the integral case the regular "curve-like" sets and irregular "dust-like" sets are characterised. In the fifth chapter a very useful tool, comparable net measures, is presented and applied to construct a subset with finite s-measure of a set with infinite s-measure, and to calculate the Hausdorff measure of Cartesian products of sets. In the sixth chapter two fruitful theories from analysis, potential theory and Fourier transforms are applied to investigate the projection properties of s-sets. The next chapter discusses the interesting problem of Kakeya of finding a set with zero measure containing a line segment in every direction. It is demonstrated that the previously descrised theory is related by duality to Kakeya sets. The final chapter contains miscellaneous examples of fractal sets. Methods for constructing curves of fractional dimension and generating self-similar sets are presented and applications to number theory, convexity, dynamical systems and Brownian motion are shown.

Each chapter contains a problem set which complete the topics and may help the reader in understanding the basic methods. The book is recommended to pure mathematicians, but it may be useful to anybody interested in the application of fractals.

J. Kincses (Szeged)

D. J. H. Garling, A Course in Galois Theory, VIII+167 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1986.

This book deals with Galois theory at a level or somewhat higher than it is customarily presented for undergraduates of mathematics. In fact, the book grew out of a course of lectures the author gave for several years at Cambridge University. Pages 1 trough 36 give a concise account on the necessary prerequisites like groups, vector spaces, rings, unique factorization domains and irreducible polynomials. The rest of the text is devoted to the theory of fields and Galois theory. Besides the classical topics including the problems of solvability the general quintic and geometric constructibility, some extra material, rarely discussed in teaching activity, is also added. For example, Lüroth's theorem, the normal basis theorem and a procedure for determining the Galois group of a polynomial is included. The reader is challenged by more than 200 exercises.

This book is warmly recommended mainly for instructors and students and also for everyone interested in its topic.

Gábor Czédli (Szeged)

Geometrical and Statistical Aspects of Probability in Banach Spaces. Proceedings, Strasbourg 1985. Edited by X. Fernique, B. Heinkel, M. B. Marcus and P. A. Meyer (Lecture Notes in Mathematics, 1193), II+128 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986.

The volume starts with a short description of the significant work of the young Strasbourg probabilist Antoine Ehrhard, who died less than two weeks before the meeting, by C. Borell.

A short note by S. Guerre deals with almost exchangeable sequences, another one by M. B. Schwarz with mean square convergence of weak martingales. B. Heinkel is on the strong law of large numbers in smooth Banach spaces, while M. Ledoux and M. B. Marcus are on the almost sure uniform convergence of Gaussian and Rademacher infinite Fourier quadratic forms. The papers by M. Ledoux and J. E. Yukich present results for the central limit theorem in a Banach space in two different directions. The paper of P. Doukhan and J. R. Leon deals with the central limit theorem for empirical processes indexed by functions based on stationary, strongly mixing random elements and for the local time of Markov processes with an application to testing uniformity on a compact Riemannian manifold, while the comprehensive 37-page article by P. Massart is on the rate of convergence in the central limit theorem for general empirical processes indexed by functions satisfying certain entropy conditions.

Sándor Csörgő (Szeged)

Mikhael Gromov, Partial Differential Relations (Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, Band 9), IX+363 pages, Springer-Verlag, Berlin—Heidelberg—New York—London—Paris—Tokyo, 1986.

The purpose of this book is to give a systematic and selfcontained treatment of analytical, topological and differential geometric methods of the theory of undetermined partial differential equations or differential relations and of its applications to imbedding and immersion problems of Riemannian and symplectic manifolds. This theory has been developed in the last 20 years mainly by the author's initiatives and activity.

Part 1 contains a survey of the basic problems and results giving the most important motivations for the theory. Part 2 is devoted to the study of a construction method which is a homotopic deformation of a jet-section solution into a differentiable map satisfying the differential relation. Part 3 deals with the investigation of C^∞ isometric immersions of Riemannian, Pseudo-Riemannian and symplectic manifolds.

The author writes in the Forward: "Our exposition is elementary and the proofs of the basic results are selfcontained. However, there is a number of examples and exercises (of variable difficulty), where the treatment of a particular equation requires a certain knowledge of pertinent facts in the surrounding field." But in the reviewer's opinion the reader is presupposed to be familiar with the techniques of singularity theory, differential operators, differential geometry and topology on higher than an elementary level.

The book includes many new results and yields a good overwiew of this developing field of mathematics.

Péter T. Nagy (Szeged)

John Guckenheimer—Philip Holmes, Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields (Second Printing, Revised and Corrected; Applied Mathematical Sciences, 42), XVI+459 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

At the early stage of the history of mechanics the oscillations were studied as "small oscillations". It means that one considers the linearized equations of motion around the equilibrium or the periodic orbit investigated. But these linear equations can describe the behaviour of the motions only locally. For example, within this theory it is impossible to handle the interaction between two or more isolated equilibria or cycles, which are very common in differential equation models. The global description of trajectories demands the study of the original nonlinear models.

However, the theory of nonlinear differential equations, as contrasted with that of linear equations, is far from being complete. Here the qualitative approach is the most important and fruitful, which has been developed by marrying analysis and geometry.

Over the past few years there has been increasing a widespread interest in the engineering and applied science communities in such phenomena as bifurcations, strange attractors and chaos. The rigorous study of these phenomena needs a wide and deep mathematical background and this is provided by the modern theory of dynamical systems. This book gives an excellent introduction to this fairly sophisticated theory for those who do not have the necessary prerequisites to go directly at the research literature.

Chapter 1 provides a review of basic results in the theory of dynamical systems and differential equations. Chapter 2 presents four examples from nonlinear oscillations: the famous oscillators of van der Pol and Duffing, the Lorenz equations and a bouncing ball problem. By the aid of these examples the reader can get acquainted with the chaotic behaviour of solutions and the concept of the strange attractor: an attracting motion which is neither periodic nor even quasiperiodic. Chapter 3 contains a discussion of the methods of local bifurcation theory, including center manifolds and normal forms. Chapter 4 is devoted to the method of averaging, perturbation theory and the Kolmogorov—Arnold—Moser theory. In Chapter 5 the famous horseshoe map of Smale is discussed in a nice and intelligible way. Chapter 6 is concerned with global homoclinic and heteroclinic bifurcations. In the final chapter the degenerated local bifurcations are treated.

It is easy to understand that the second edition of this valuable text-book had become necessary. It has to be on the bookshelf of every mathematician and of every user of mathematics interested in the modern theory of differential equations, dynamical systems and their applications.

L. Hatvani (Szeged)

James M. Henle, An Outline of Set Theory (Problem Books in Mathematics), VIII+145 pages, Springer-Verlag, New York—Berlin—Heidelberg—London—Paris—Tokyo, 1986.

Set theory is full with charming exercises, brilliant constructions and problems which sometimes challenge even the experts. Henle's book misess them all and this is all the more unfortunate that there is no good published collection of (solved) problems in set theory (one should not count those feeble attempts that every now and then appear on the scene with completely trivial exercises).

More proper justice should however be given to this worthy book because its aim is different. I feel neither the "Problem Books" series nor the title are appropriate for this work, for this is not a problem book in the ordinary sense, nor it is about Set Theory. More appropriate title would be something like "Construction and properties of numbers", all sorts of numbers such as naturals, integers, rationals, reals, ordinals, cardinals, infinitesimals. The *spirit* is set theoretical and ultimately this is why I would rank very high Henle's book.

It introduces an outstanding pedagogical system: so called projects are assigned to students. These contain proofs, discoveries of theorems and concepts etc.; under the guidance of the teacher the students work alone, and through these projects they explore the field step by step like "real" researchers. They "experience the same dilemmas and uncertainties that faced the pioneers". Accordingly, the book consists of three parts, hints and solutions occupy the second and third ones.

In my opinion every good exercise book (cf. Pólya—Szegő's, Halmos's, Lovász's) should be based on similar principles; here however the method is applied to the very foundation and exposition of the subject. I saw the same method efficiently working at Ohio State University where selected high school students participated in university summer shools. Although I doubt that the ordinary math major would succesfully complete the projects in the book, the system is certainly applicable to the better ones.

It is unfortunate that the author mostly restricted himself to the dullest part of set theory: construction of numbers and operations between them (I must add, however, that due to J. von Neumann and A. Robinson, the construction of ordinals and infinitesimals is definitely an exception). This may be so because set theory is not too well adequate for the above method (after all infinite sets are not objects that you can experience with); number theory, geometry, elementary algebra etc. certainly are more suitable.

The main advantage of the book is that without disturbing formalism the author gets the students think and work in a way as a logician should do, and the presentation is extremely "clean" and accurate. For this reason I warmly recommend that every math major read the book even if they have already completed a course in set theory. The material can also be valuable for lecturers on set theory and logic.

There is one more reason for this strong recommendation, and this is the Goodstein—Kirby—Paris theorem discussed in the last chapter (it is about an extraordinary number theoretical iterative process that seemingly produces larger and larger numbers but somehow it always reaches 0; and this can be proved *only* using infinite numbers). In the style of the quotations in the book: All's Well that Ends Well.

Vilmos Totik (Szeged)

Homogenization and Effective Moduli of Materials and Media. Edited by J. L. Ericksen, D. Kinderlehrer, R. Kohn and J.-L. Lions (The IMA Volumes in Mathematics and its Applications, Volume 1), X+263 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

At the Institute for Mathematics and its Applications the year 1984—1985 was dedicated to the study of partial differential equations and continuum physics. This volume, the first one in a series, contains research papers presented at a workshop on homogenization of differential equations and the determination of effective moduli of materials and media. This up-to-date theme is interesting for mathematicians, physicists and for engineers equally well. The papers are well-organized. In general they contain the origin and the history of the investigated problem and after the discussion open questions are presented. The style is well-characterized by the following sentence taken from Luc Tartar's paper: "This mathematical model of some physical questions involving different scales will of course be questioned by some; it is natural that it be so but I hope that criticism will be made in a constructive way and so improve my understanding of continuum mechanics and physics (and maybe of mathematics)."

The titles of the papers are: Generalized Plate Models and Optimal Design. — The Effective Dielectric Coefficient of a Composite Medium: Rigorous Bounds From Analytic Properties. — Variational Bounds on Darcy's Constant. — Micromodeling of Void Growth and Collapse. — On Bounding the Effective Conductivity of Anisotropic Composites. Thin Plates with Rapidly Varying Thickness and their Relation to Structural Optimization. — Modelling the Properties of Composites by Laminates. — Wares in Bubbly Liquids. — Some Examples of Crinkles. — Mikrostructures and Physical Properties of Composites. — Remarks on Homogenization. — Variational Estimates for the Overall Response of an Inhomogeneous Nonlinear Dielectric.

L. Pintér (Szeged)

Mark Kac—Gian-Carlo Rota—Jacob T. Schwartz, Discrete Thoughts: Essays on Mathematics, Science, and Philosophy. Edited by Harry Newman (Scientists of Our Time), XII+264 pages, Birkhäuser, Boston—Basel—Stuttgart, 1986.

the reason people so often lie
is that they lack imagination:
they don't realize that the truth, too,
is a matter of invention.

This is how Rota starts the volume in his preface, translating nicely a three-line poem of Machado which summarizes the "prophetical warning" of the philosopher Ortega.

This fine composition of twenty-six essays should be read by every mathematician, statistician and computer scientist. It would be a little bit better still if every physicist, economist and historian of science could also read it, and the world, scientific or otherwise, would surely improve a trifle if in fact the whole intelligentsia read it. It is not just that three "grifted expositors of mathematics" came together as the jacket says, but these three illustrious thinkers really want to tell the truth, if not the whole truth, but nothing but the truth. And the world usually betters by telling the truth.

There are as many readings of such a text as readers. In the reviewer's reading the frame of this composition is constituted by the seven brilliant writings of the late Professor Kac (he died in the fall of 1984). These are: Mathematics: Tensions (essay No. 2), Statistics (4), Statistics and its history (5), Mathematics: Trends (8), Academic responsibility (13), Will computers replace humans? (18), Doing Away with Science (25). Heavier building blocks are brought by the six essays of Schwartz: The pernicious influence of mathematics on science (3), Computer science (7), The future of computer science (9), Economics, mathematical and empirical (10), Artificial intelligence (16), Computer-aided instruction (19) and by Rota's essay Combinatorics (6). The cohesion of and the paint on the structure is provided by Rota's shorter bookreviews and sketches: Complicating mathematics (11), Mathematics and its history (12), Husserl and the reform of logic (14), Husserl (15), Computing and its history (17), Misreading the history of mathematics (20), The wonderful world of Uncle Stan (21), Ulam (22), Kant (23) and Heidegger (24) and his Chapter 1 (Discrete thoughts) and Chapter 26 (More discrete thoughts).

The depth of thought, charm, experience and characteristic wit of Kac, the vehement cool logic of Schwartz and the broad knowledge and aphoristic penetration of Rota harmonize beautifully. The selection and ordering of the essays (presumably the work of the editor) to achieve the non-formalizable rythm of thought in the book, a composition instead of a collection, will not be possible for any artificial intelligence. The whole thing is more continuous than discrete.

No, Uncle Mark, computers will never replace a man like you were.

Sándor Csörgő (Szeged)

Serge Lang, Linear Algebra. Third edition (Undergraduate Texts in Mathematics) IX+285 pages, Springer-Verlag, New York—Berlin—Heidelberg—London—Paris—Tokyo, 1987.

The text is divided into twelve chapters. The first four chapters introduce vector spaces over subfields of complex numbers, the space of matrices, matrices of linear equations, the notion of linear operators and show the connection between matrices and linear operators. In Chapter 5 scalar products and orthogonality are defined; and applications to linear equations, bilinear and quadratic forms are given. Chapter 6 is devoted to give the notion and elementary properties of determinants. Chapter 7 studies the important special cases of linear operators, symmetric, Hermitian and unitary operators. Chapter 8 defines eigenvalues and eigenvectors, the characteristic polynomial, and gives the method of computing eigenvalues by finding the maximum and the minimum of quadratic forms on the unit sphere. In Chapter 9 polynomials of matrices and linear operators are defined. The main purpose of Chapter 10 is to prove the existence of triangulation of linear operators and especially of the diagonalization of unitary operators. Chapter 11 deals with the factorization of polynomials and as an application of this concludes the Jordan normal form of linear operators. Chapter 12 is devoted to the study of convex sets and proves the finite dimensional case of the Krein—Milman theorem. At the end of the book an Appendix can be found dealing with complex numbers.

The book can be used as a handbook for learning linear algebra.

L. Gehér (Szeged)

László Lovász—Michael D. Plummer, Matching Theory (North-Holland Mathematics Studies), XXXIII+544 pages, Akadémiai Kiadó, Budapest and North-Holland, Amsterdam, 1986.

Matching theory consists of only a part of graph theory, but one of its deepest and hardest parts. The history of matching is related to the four color problem, and since then it has been a focus of interest. Most of the general questions and methods of combinatorics are considered in matching theory, and many have a nice solution or application. Because of its complexity, matching theory really has the right to bear the title of theory.

The book starts with the most classical results. The first two chapters contain the very basic and very important results on bipartite graph matching and flow theory. The next three chapters deal with the structure of general graphs related to matching. The main result in this territory is the Edmonds—Gallai structure theorem which shows that from the point of view of matchings, every graph is built of several different kind of "bricks". These bricks are well known, thanks to the two authors' previous works. The next four chapters illuminate matching problems from different perspectives. The first one discusses the graph-theoretical consequences. The next chapter shows the very important connection with linear programming. The book discusses the description of a matching polytope, its facets and the dimension of the perfect matching polytope. This section includes the effect of the ellipsoid method on combinatorial optimization. There is one chapter on the related enumeration problems. Besides the basic results on permanents, pfaffians, and matching polynomials, there are some interesting applications of these results. One other chapter covers the algorithm-theoretic aspects of matchings, containing not only the important algorithms, but also their implementation. The final three chapters consider the generalizations of the question of matchings in graph theory and matroid theory. These discuss the problem of f-factors, vertex packing, hypergraph matching and matching of 2-polymatroids.

The book contains all of the important results which are in or related to matching theory. The many applications of this subject in other parts of combinatorics, and the wide variety of methods used ensure that the reader will not only learn about matching theory, but about most of the important parts of combinatorics. There are discussions of matroid theory, polyhedral combinatorics, enumerations, algorithm theory, and data structures. The corresponding chapters are not only good introductions to the fields but contain some of the most important results of current research. Sometimes the flow of results is interrupted by "boxes". These boxes contain remarks which go beyond the scope of the book or sketch a main underlying idea. These parts are very useful to understand how to fit the actual results or methods into the main stream of research. The cited references are a big help to the reader whose appetite has been whetted. These short guides are also very helpful for the reader who is untrained in combinatorics. If the reader takes the effort and studies mathematics by solving problems (this is harder but more rewarding), then there are a lot of exercises inserted in the text. Solving these exercises adds a lot of fun to the reading and gives good practice for the methods. If the reader requires more of a challenge, there are many open problems in the book. Both authors live and breathe matching theory. As a consequence, the chapters frequently suggest the most important directions of research, and the reader easily can find problems to think about.

Matching theory is only a small part of combinatorics. This might lead an outsider to think that this book is too specialized. (A very narrow, specialized subject gets no interest outside a small group.) This is not the case with this book. The relations between combinatorics and the classical fields, and the applications of combinatorics are an undiscovered part of the science. This book gives many examples of applications of combinatorics to other areas of mathematics and the physical sciences. Here are just some of them for appetizers: engineering, chemistry, physics, measure theory and topology.

Finally, the physical appearance of the book is pleasing, as it was typeset with the $\TeX$ system. The book is very important to any specialist in combinatorics. It is highly recommended to anybody who is interested in this new part of mathematics or who is working in a field which applies combinatorial methods.

Péter Hajnal (Szeged and Chicago)

New Developments in the Theory and Applications of Solitons, Proceedings of a Royal Society Meeting, London, 1984 November. Edited by Sir Michael Atiyah, J. D. Gibbon and G. Wilson, (Reprint from the Philosophical Transactions of the Royal Society, Ser. A. Vol. 315, p. 333—469), The Royal Society, London, 1985.

The exponential growth of the number of papers about solitons has become somewhat less steep in the last few years, nevertheless they are still the subject of intensive study. Over and over again new delicate details of soliton theory are discovered by pure mathematicians, and the sudden appearance of solitons is not a rare event in any field of physics.

This situation is well documented in the introductory lecture of these proceedings: "A survey of the origins and physical importance of soliton equations" given by J. D. Gibbon. Here past and present of solitons are outlined, and this is the lecture that can be recommended both to the beginner and to the specialist, in order to see how wide this field really is. More detailed investigations of some of the branches of mathematics and physics, where solitons play important role can be found in the other 8 papers of this volume. Half of them have purely mathematical character, and show the connection of soliton theory with such classical problems as the integrability of ordinary differential equations, as well as with modern fields like algebraic geometry and Kac—Moody algebras. The rest of the articles communicate on experiments and applications of soliton theory in laser physics, biomolecules, magnetic monopoles, fluid dynamics etc.

M. G. Benedict (Szeged)

New Directions in the Philosophy of Mathematics: An Anthology, Edited by Thomas Tymoczko, XVII+323 pages, Birkhäuser, Boston—Basel—Stuttgart, 1985.

The philosophy of mathematics has played an important role in philosophy going back to the ancient Greeks. This discipline has been radically changed about the turn of the century. The new dominant question (or the new paradigm, according to T. Kuhn's terminology) was: what is the foundation of mathematics?

Now, in the latest decades, as R. Hersh wrote: "We are still in the aftermath of the great foundationalist controverses of the early twentieth century. Formalism, intuitionism and logicism, each left its trace in the form of certain mathematical research program that ultimately made its own contribution to the corpus of mathematics itself."

In Part I, entitled Challenging Foundations, we can read five essays on the major perspectives on the philosophy of mathematics written by R. Hersh, I. Lakatos, H. Putnam, R. Thom and N. D. Goodman. They strongly criticize the foundationalist approach to the philosophy of mathematics. This part is followed by an interlude containing two writtings of G. Pólya, who was the forerunner of quasi-empiricism in mathematics. The essays in the second part demonstrate quasi-empiricism which is an increasingly popular approach to the recent philosophy of mathematics.

Part II deals with the reexamination of mathematical practice. It contains three sets of essays. The first set explores some general issues in mathematical practice, starting with the concept of informal proof. The authors are: Hao Wang, I. Lakatos, Ph. J. Davis and R. Hersh. The second set of essays focuses on the growth of mathematical knowledge, the development or change in the essential aspect of informal proof. The authors are: R. L. Wilder, Judith V. Grabiner and Ph. Kitcher. The final set continues the theme of informal proof and discusses the change due to the use of computers in mathematical research. The authors are: T. Tymoczko, R. A. de Millo, R. J. Lipton, A. J. Perlis and G. Chaitin.

All essays of the second part argue the philosophical relevance of mathematical practice. According to the editor's view: "The crucial step in approaching them is our willingness to conceive of mathematics as a rational human activity, that is, as a practice."

Each part and almost all essays have an introduction written by the editor which helps the reader in better understanding and offers a short summary.

As a recommendation we cite the closing paragraph of the editor's Introduction: "Although this anthology does not completely represent the philosophy of mathematics, it does, I belive, gather together some of the more exciting essays published recently in the field. In this instance, the whole really is greater than the sum of all its parts; each essay reinforces the others. One purpose in bringing these essays together is to demonstrate their collective force. The collection will have succeeded if it stimulates the reader — mathematician or philosopher, professional, apprentice or amateur — to rethink his or her conception of mathematics."

Lajos Klukovits (Szeged)

N. K. Nikol'skii, Treatise on the Shift Operator: Spectral Function Theory. With an Appendix by S. V. Hruščev and V. V. Peller. Translated from the Russian by Jeak Petree (Grundlehren der mathematischen Wissenschaften 273), XI+491 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986.

The title of the Introductory Lecture (chapters are called lectures) is: "What this book is about." A short, and thus by no means exhaustive, answer can be: about non-classical spectral theory in Hilbert space. The discussion is essentially based on the functional model for contractions due to Sz.-Nagy and Foiaş. This approach makes possible to use more function-theoretic tools, namely many properties of functions in Hardy classes, as in classical spectral theory. The central role of the shift operator in this model makes at once understandable why this work can be considered (again, not in its totality) as a "treatise on the shift operator".

Besides the introductory one the book contains eleven lectures. In the first parts of these lectures the shift operator in question appears as multiplication by the independent variable in the Hardy space H^2 of scalar valued functions on the disc. These parts can be considered as an introduction in an elementary fashion to the second ones, entitled "Supplements and Bibliographical Notes". These second parts contain more advanced studies extending the first parts in various directions and are written more condensedly and to a certain extent sketchily. Each lecture is ended by "Concluding Remarks" where a review of the literature and hints for unsolved problems complete the discussion.

It is hopeless to even try to sketch the rich contents of this book, the "Bibliography" lists about five hundred items! It may be informative to mention that the Carleson corona theorem plays a central role in the discussion. Another interesting method is the introduction of special Hankel and Toeplitz operators when studying the model operators.

The present book is not simply a translation of the original Russian one but it is an improved and considerably enlarged edition. Some parts have been revised and moreover, while the Russian original has contained only a single Appendix on the spectral multiplicity of operators of class C_0 the present edition contains four more ones. Appendix 2 presents the proof of all assertions on Hardy classes which are used in the text. Appendix 3 contains the modern proof of the Carleson corona theorem and its operator theoretic generalisation. Appendix 4 is devoted to Toeplitz and Hankel operators connected with the general orientation of the book. Appendix 5 entitled "Hankel operators of Schatten—von Neumann class and their application to stationary processes and best approximations" has been written by S. V. Hruščev and V. V. Peller. "List of Symbols", "Author Index" and "Subject Index" complete the book.

The reader needs to be familiar only with standard material in mathematical analysis taught usually in undergraduate courses. Because of the many interesting methods and the large material covered in this book, it can be warmly recommended to everybody who is interested in its topic. The special two-level structure of the discussion certainly helps the reader to orient himself. It is worth to glance trough this edition even for those who know the Russian original well, because of the improvements and Appendices mentioned above.

E. Durszt (Szeged)

Optimization and Related Fields, Proceedings of the "G. Stampacchia International School of Mathematics" held at Erice, Sicily, September 17—30, 1984, Edited by R. Conti, E. De Giorgi and F. Giannessi (Lecture Notes in Mathematics, 1190), VIII+419 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986.

To find extreme values of functions is perhaps the most important problem of mathematics derived from practice. It is the simplest case of this problem when the maximum or minimum of a smooth function of several variables in a domain is to be found. But loking for the extreme values of smooth functions on a closed set with a piece-wise smooth boundary, which is common e.g. in econometrics already requires a lot of special methods that constitue mathematical (nonlinear) programming. Similarly, in the calculus of variations some new problems have appeared recently in which the control parameters vary on closed sets with boundaries. These problems gave rise the "new calculus of variation", control theory or the theory of optimal processes. As it has turned out, functional analysis is suitable for investigating the deep common roots of these optimization problems seemingly independent at the first glance.

These lecture notes contain the invited talks of the meeting above, whose aim was to give an opportunity for promoting the exchange of ideas and for stimulating the interaction among various branches of optimization. The reader can find articles among others on gradient methods, homogenization problems in mechanics, Lagrange multipliers, equilibria in the theory of games, recent progress in the calculus of variations and optimal control problems and stability analysis in optimization.

L. Hatvani (Szeged)

Oscillation Theory, Computation, and Methods of Compensated Compactness. Edited by Constantine Dafermos, J. L. Ericksen, David Kinderlehrer, and Marshall Slemrod (The IMA Volumes in Mathematics and Its Applications, 2), IX+395 pages, Springer-Verlag, New York—Berlin—Heidelberg—London—Paris—Tokyo, 1986.

This volume is the proceedings of the Workshop held under the same title in the Institute for Mathematics and its Applications (University of Minnesota). The Workshop was an integral part of the 1984—85 IMA program on Continuum Physics and Partial Differential Equations. The subject-matter of the conference was the treatment of nonlinear hyperbolic systems of conservation laws, which is the most important problem of continuum mechanics. Both the analytical and numerical sides were emphasized, and special attention was paid to the new ideas of compensated compactness and oscillation theory. The proceedings contain articles among others on the nonlinear Schrodinger equation, total variation dimishing schemes, the weak convergence of dispersive difference schemes, the Korteweg de Vries equation, nonlinear geometric optics, commutation relations, and the interrelationship among mechanics, numerical analysis, compensated compactness and oscillation theory.

L. Hatvani (Szeged)

Pappus of Alexandria, Book 7 of the Collection, Edited with translation and commentary by Alexander Jones, in two Parts, with 308 Figures (Sources in the History of Mathematics and Physical Sciences, 8), X+748 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

Pappus of Alexandria flourished about 320 A.D. and had the opportunity to read all the books of the preceding ages in the liberary. To help the forthcoming generations in studying the works of the famous Greek mathematicians and astronomers, he has written detailed commentaries. If he thought a proof of a theorem too difficult, then he inserted a lemma to make it easier, and if an author considered only one of the possible cases, then Pappus supplied with similar proofs the remaining cases. We know the content of several Greek works from his commentaries only.

One of his most famous works is the Collection, which contains eight books and preserved in a tenth-century manuscript, Vaticanus gr. 218. This is defective at the beginning and end. We have lost (in Greek) Book 1, the first part of Book 2, and the end of Book 8. The Collection has often been regarded as a kind of encyclopedia of Greek mathematics, a compendium in which Pappus attempted to encompass all the most valuable accomplishments of the past.

Book 7 of the Collection is a companion to several geometrical treatises, which were supposed to equip the geometer with a "special resource" enabling him to solve geometrical problems. More precisely, they were to help him in a particular kind of mathematical argument called "analysis", which is a kind of reversal of the usual "synthetic" method of proof and construction.

In Part I we can read an Introduction: Pappus and the Collection, containing historical and textological remarks, an Introduction to Book 7, and the Greek text of Book 7 with a fresh English translation due to the editor.

Part II contains three essays on lost works that Pappus discusses: The Minor Works of Apollonius, Euclid's Porisms and The Loci of Aristaeus, Euclid and Eratosthenes. This part contains a general and a Greek index and the figures to the text.

We warmly recommend this valuable work to everybody who is interested in ancient mathematics.

Lajos Klukovits (Szeged)

H.-O.Peitgen—P. H. Richter, **The Beauty of Fractals**, XII+199 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986.

The theory of fractals is a rapidly developing part of mathematics nowadays. Mandelbrot's work indicated the turning point and his famous book aroused both scientists' and nonscientists' interest in fractals. The theory, besides the theoretical interest in itself has practical importance and the computer-generated colour pictures of fractals have aesthetic value. The authors aim was to unite these three aspects of fractals.

The book starts with the essay "Frontiers of Chaos" which, without any mathematical rigour, explains the background to the non-specialist. This is followed by eight special sections, each of which corresponds to a part of the essay and completes the topic considered there.

In the first special section the authors analyse the Verhulst dynamics which is a population growth model with one controlling parameter: $x_{n+1}=(1+r)x_n-rx_n^2$. Depending on the choice of the parameter r, the system may be convergent, periodic or, surprisingly, "chaotic". This "deterministic chaos" has become an important idea and directed the attention to aspects of complex analytical dynamical systems. Fatou and Julia extensively studied these processes during the first World War. In the second special section the definition and basic properties of Julia and Fatou sets are collected without proofs but with complete references for the interested readers. The works of Julia and Fatou "remained largely unknown, even to mathematicians, because without computer graphics it was almost inpossible to communicate the subtle ideas". They characterized the Julia set, which is the set of initial values for which the process behaves chaotic, in two ways. These results make the computergraphical generation of Julia sets possible and their properties become easy by looking at these pictures. This enables us "thinking in pictures" and the experimental computer results can help in arriving at new discoveries and conjectures. The philosophical contents of this kind of unity of science and art is discussed in detail in the essay. The third special section contains Sullivan's famous classification theorem of the components of the Fatou set. The authors give several examples from physics, biology and other fields to show that the quadratic dynamical systems have special importance. (From the dynamical point of view these are equivalent to the processes generated by the polynomials $p_c(z)=z^2+c$). From the general theory of Fatou and Julia it follows that the Julia set in this case is either connected or a Cantor set. Mandelbrot defined and investigated the set of c values for which the corresponding Julia set is connected. This strange set is named after him today. The fourth special section is devoted to the Mandelbrot set and an up-to-date list of known results and related problems are presented. When c is wandering in a component of the interior of the Mandelbrot set then the corresponding Julia set does not change topologycally but at branch points qualitative changes occur and crossing the boundary yields the most dramatic one, the Julia set becomes a Cantor set. This "transition from order into chaos" phenomenon is one of the central questions treated in the essay. In the fifth section the relationship between two-dimensional electrostatistics and quadratic processes are discussed. Potential theory is applied to obtain additional information about the structure of fractals. The maps contained in the book were coloured by calculating equipotential lines. Calculation of field lines is usually hard but in the case of the Mandelbrot set an efficient method, the Hubbard trees, is presented. In the next three special sections the Newton method for the complex and for the real case and a discrete Volterra—Lotka system are analyzed from the dynamical point of view. Surprisingly, the pictures of fractals in the complex and real cases are different. In the authors opinion the former are in baroque style and the latter are more modern shapes, and they state that something must be hidden behind this fact.

The second essay "Magnetism and complex boundaries" is an intuitive outline of a possible explanation of phase transition on the ground of the fractals. The next two special sections contain the physical and mathematical details.

The book contains papers of four invited contributors, the most distinguished experts of the field. B. Mandelbrot reports on the way that has led him to the discovery of the Mandelbrot set. A. Douady presents an outline of the known results and unsolved questious. The physicist G. Eilenberger describes the symbolic meaning of what the authors' pictures may have within the changing comprehension of nature. H. W. Franke, one of the pioneers of computer graphics, reports on his own experiences and draws a number of inferences from them.

There is a "Do it yourself" section at the end of the book. This contains some hints for interested readers who want to try to generate pictures on their own computer. The presentation of the book is nice, it contains 88 really beautiful pictures. The book is recommended to anybody, from pure mathematicians to the layman, who is interested in fractals.

J. Kincses (Szeged)

R. Michael Range, Holomorphic Functions and Integral Representations in Several Complex Variables (Graduate Text in Mathematics), XIX+386 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

"The subject of this book is Complex Analysis in Several Variables. This text begins at an elementary level with standard local results, followed by a thorough discussion of the various fundamental concepts of" complex convexity "related to the remarkable extension properties of holomorphic functions in more than one variable. It then continues with a comprehensive introduction to integral representations, and concludes with complete proofs of substantial global results on domains of holomorphy and on strictly pseudoconvex domains in C^n, including, for example, C. Fefferman's famous Mapping Theorem."

The book, written in a lucid style and offering the reader a wealth of material, is excellent for courses and seminars or for independent study. Much of this material was not readily accessible and the inclusion of such topics greatly enhances the value of the book. The most important prerequisities are: calculus in several real variables, complex analysis in one variable, Lebesgue measure and the elementary theory of Hilbert and Banach spaces and some basis facts of point set topology and algebra.

A good book has some characteristic features which run through it. In this work integral representations are the principal tools in developing the global theory. This presentation has several advantages. For example, as the author writes, it helps to bridge the gap between complex analysis in one and in several variables, it directly leads to deep global results and concrete integral representations lend themselves to estimations. The work presents the main developments of the last twenty years concerning integral representations. One of the other characteristic features of the book is the constant presence of historical comments. In the light of these comments the new notions and results become more natural and understandable. This is the most attractive peculiarity of the book for the reviewer. (One of the particularly valuable gems can be found at the end of Ch. IV. on the history of integral representations.) A further remarkable feature of this book is that it contains a relatively large number of exercises ranging from the routine to the very advanced ones. This is particularly important since the subject abounds in abstract theorems and has only a few worked examples.

L. Pintér (Szeged)

Patrick J. Ryan, Euclidean and non-Euclidean Geometry. An Analytic Approach, XVII+215 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1986.

Teachers of geometry can find nowadays some good text-books on Euclidean and non-Euclidean plane geometry which can serve as an introduction to combinatorial, algebraic or topological theories of transformation groups, to direct methods of non-Euclidean spaces and of the differential geometry of homogeneous manifolds. Since computational-analytical aspects of geometric theories have increasing importance in the applications in mathematical physics and computer graphics, there is a demand on an up-to-date analytical introduction to plane geometry. The present book gives a very well-written and useful treatment of this topic. It contains the fundamentals of Euclidean, spherical, elliptic and hyperbolic plane geometry using the methods of isometric, affine and projective transformation groups. At the same time it provides an arsenal of computational techniques and a certain attitude toward geometrical investigations. It aims to give an appropriate background for teachers of high school geometry and to prepare students for further study and research.

The book is self-contained for upper-level undergraduate mathematics students, the necessary knowledge is summarized in appendices. Only a familiarity with linear algebra and elementary transcendental functions is expected from the reader. The material is illustrated with many exercises, requiering specific numerical computations or supplying proofs that have been omitted. Some of them extend the results proved in the text.

The first main part is, of course, Euclidean plane geometry (Historical introduction, Plane Euclidean geometry, Affine transformations in the Euclidean plane, Finite groups of isometries of E^2). The second part contains: Geometry on the sphere, The projective plane P^2, Distance geometry on P^2. The last chapter is: The hyperbolic plane.

The book gives a very good basis for high school geometry teaching and a good introduction for graduate work in differential geometry or computer graphics.

Péter T. Nagy (Szeged)

Lewis H. Ryder, Quantum Field Theory, X+443 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1985.

For a long period, quantum field theory had meant only the quantum theory of electromagnetic fields. Other forces of nature as the weak and strong nuclear interactions resisted the formalism that proved to be so successful in the description of electromagnetism. The principle of local gauge invariance has overcome the difficulties, and the prominent achievements of gauge field theories of the seventies have reached the textbook level by now.

To find a good balance however, in a single book, between the several parts of this huge subject is not an easy task. This is the more so if among the author's aims is that the presentation should be intelligible by a graduate student. The Quantum Field Theory by L. Ryder has solved this problem succesfully. To read this volume it is enough to be familiar with quantum mechanics and special relativity. The text leads us with great pedagogical skill, step by step from elementary field theory to the renormalization of gauge fields. The emphasis of the presentation is on the path integral method. Besides introducing the fundamentals of quantum field theory, the author acquaints the reader with some modern mathematical tools as well. One may regret that certain more recent concepts (e.g. supersymmetry) are not found in the book, but the author probably wanted to include only those results that have more or less experimental basis.

The book is very well suited for teaching and studying this subject, and brings even the beginner close to present day field theory.

M. G. Benedict (Szeged)

M. Shirvani—B. A. F. Wehrfritz, Skew Linear Groups (London Mathematical Society Lecture Note Series, 118), 253 pages, Cambridge University Press, Cambridge—New York—Melbourne, 1986.

Skew linear groups arise naturally as a generalization of linear groups, by omitting the requirement of the commutativity of the corresponding field. One of the main problems in passing from linear groups to skew linear groups is that, at least presently, division rings are rather difficult to handle. The investigation of skew linear groups is a fairly young branch of algebra, in comparison with the theory of linear groups, however, in recent years it has expanded very rapidly. This book is the first monograph providing a systematic treatment of a number of results that were available, till now, in research papers only.

In Chapter 1 the basic concepts such as irreducibility, absolute irreducibility and unipotence are reviewed in the context of skew linear groups, and some groups with faithful skew linear representations are constructed. Chapter 2 discusses finite (and locally finite) skew linear groups, including the description of finite subgroups of division rings of characteristic zero, and the theorem that finite skew linear groups over division rings of characteristic zero have large metaabelian normal subgroups. Chapter 3 is devoted to skew linear groups over locally finite-dimensional division algebras, with the emphasis laid on nilpotence and solubility. In Chapter 4 the authors consider skew linear groups over division rings generated by a central subfield and a polycyclic-by-finite subgroup. Chapter 5 contains a detailed study of normal subgroups of absolutely irreducible skew linear groups. In Chapter 6 the book concludes with an application showing how the theory of skew linear groups may shed light on some known results on group rings.

To help the reader, the authors give a list of prerequisites for each chapter, a detailed notation index, and author and subject indices. This monograph is warmly recommended as a textbook for those wishing to get acquainted with the subject, and as a reference book as well.

Ágnes Szendrei (Szeged)

Michael Shub, Global Stability of Dynamical Systems, XII+150 pages, Springer-Verlag, New York—Berlin—Heidelberg—London—Paris—Tokyo, 1986.

The most characteristic property of an equilibrium position in a mechanical system is its stability or instability. The equilibrium position is stable if during its motion the system remains arbitrarily near the equilibrium state provided that it was near enough at the initial moment. Obviously, only the stable equilibria can be realized in practice, so the research for conditions of stability have started at the early stages of mechanics and mathematics. Later on the investigations have been extended to general dynamical systems and have created the Lyapunov Stability Theory. In this theory it is always assumed that the system itself is under ideal circumstances, i.e. it cannot be disturbed by outside effects. However, each system is under the action of certain small undefinable perturbations. Therefore, it is clear that one can expect only those properties of the model to be realized in reality which are not too sensitive to small changes in the model. In 1937 Andronov and Pontryagin introduced the concept of robustness or roughness of a system (nowadays it is called structural stability), which means that the topological structure of the trajectories does not change under small perturbations of the system. If we observe the trajectories only in a neighbourhood of a point then we talk about local stability. If the trajectories are observed on the whole manifold then the stability is global.

The central objective of the modern theory of Dynamical Systems is the description of the orbit structures of vector fields on a differentiable manifold. There exist, however, fields with extremely complicated orbit structures, thus one has to restrict the study to a subset of the space of vector fields. It is desirable that this subset should be open and dense, or as large as possible, and it should consist of structurally stable vector fields with simple enough orbit structure so that one could classify them. Due to the celebrated Hartman—Grobman Theorem, this program has been completely solved if the stability and the equivalence are ment locally.

To complete the above program in the global sense is much more difficult. As it was proved by Smale, on manifolds of dimensions higher than two the structurally stable fields are not dense, and the structure of the trajectories and their limit sets even for the stable fields can be extremely complicated. Their description is still an active area of research.

Shub's book gives an excellent account on the results of this area, most of which were available only in articles so far. The reader can get acquainted the central concepts, theorems and examples of the global theory of dynamical systems such as filtration, hyperbolic invariant sets, change recurrence, stable and center manifold theorems, Smale's Axiom A, symbolic dynamics, Markov partitions, Ω-stability theorems, Smale's horseshoe and the solenoid. It is highly recommended to anyone interested in dynamical systems and stability theory.

L. Hatvani (Szeged)

Gábor J. Székely, Paradoxes in Probability Theory and Mathematical Statistics (Mathematics and its Applications), XII+250 pages, Akadémiai Kiadó, Budapest and D. Reidel Publishing Company, Dordrecht, 1986.

This book is very unusual and, as far as I know, is unique in its kind. It endeavours "to show how the rapidly progressing and widely used branch of knowledge of the mathematics of randomness has developed from paradoxes". While this is a bit too much to be hoped for as it flaunts, and would be the greatest paradox of all had the author succeeded in doing so, his paradoxical vision is certainly a valid one and interesting. The result is a most enjoyable reading which is worth much more then two dozens of half-thought dishonest "introduction to probability and statistics" books published in so great a number nowadays.

Chapter 1 contains the discussion of 12 paradoxes or families of paradoxes from classical probability theory, while Chapters 2, 3 and 4 expose and treat 12, 6 and 12 paradoxes or families of paradoxes in mathematical statistics, the theory of stochastic processes and the foundations of probability theory, respectively. The discussion of each paradox is devided into five parts: the history, the formulation and the explanation of the paradox, remarks and references. Furthermore, the four chapters end, respectively, with 15, 16, 8 and 8 of what the author calls quickies which either did not fit into the main line of thought of the book to be discussed in such detail as the numbered paradoxes, or are adjecant curiosities, strange facts, gems.

The history and remark sections and the quicky passages contain a lot of interesting historical and cultural information, narrated in the easy, chatting style of the author, together with stories, anecdotes and gossip. Instead of fooling around for pages on end to say the same thing politely, for example, he is not afraid of very simply stating that "R. A. Fisher hated K. Pearson". Or, while discussing the paradox of the almost sure eventual extinction of a critical Galton—Watson process, he proposes an interesting system for the inheritance of family names to avoid the replacement of "nice old family names" by "more common dull ones like Smith, etc."

The book makes an easy and recreational reading but can be used more seriously as a supplementary reading to almost any course in probability and statistics. In fact, my math major students like the original Hungarian edition, of which the present English one is a revised and updated version.

Sándor Csörgő (Szeged)

Audrey Terras, Harmonic Analysis on Symmetric Spaces and Applications I., 341 + XII pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1985.

Harmonic analysis is one of the most useful areas of mathematics which made a deep influence on several other fields of mathematics and physics. The book demonstrates exactly this usefulness by presenting many applications in number theory, statistics, medicine, geophysics and quantum physics. This is the first volume of a series dealing with the harmonic analysis of the three classical geometries (euclidean, spherical and hyperbolic).

Besides the standard development of euclidean Fourier analysis we learn in the first chapter how to use this theory to the solution of the heat equation, to the examination of crystals, as well as zeta functions of algebraic number fields. In Chapter 2 spherical Fourier analysis is applied for the study of the hydrogen atom, for the sun's magnetic field and also for group representations and Radon transforms. The last chapter is devoted mainly to the fundamental domains of discrete subgroups of hyperbolic isometries, the Reolche—Selberg spectral resolution and the Selberg trace formula.

We recommend this excellent text-book to every mathematician, engineer, scientist and applied mathematician who is interested in harmonic analysis and in its applications.

Z. I. Szabó (Budapest)

The Craft of Probabilistic Modelling: A Collection of Personal Accounts, Edited by J. Gani (Applied Probability. A Series of the Applied Probability Trust), XIV+313 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

This is the first volume of the new series in the braces above with series editors J. Gani and C. C. Heyde. The beginning is indeed very nice. The volume contains nineteen essays from leading probabilists who, among other things, have distinguished themselves in applied probability model building. Each of the essays are preceded by a short biography. Some of the essays concentrate on the models themselves that the authors have built, others are entirelly autobiographical, while the rest is a combination of the two. Some of the writings are very dry, some are exceptionally lively. I don't single out any of the essays for special mention here because more then half of them are very close to my heart for one reason or other. In the grouping of the editor, the contributors are the following. Early craftsmen: D. G. Kendall, H. Solomon, E. J. Hannan, G. S. Watson; The craft organized: N. T. J. Bailey, J. W. Cohen, R. Syski, N. U. Prabhu, L. Takács, M. Kimura, P. Whittle, R. L. Disney; The craft in development: M. F. Neuts, D. Vere-Jones, K. R. Parthasarathy, M. Iosifescu, W. J. Ewens, R. L. Tweedie.

The book is a very enjoyable reading.

Sándor Csörgő (Szeged)

John B. Thomas, Introduction to Probability (Springer Texts in Electrical Engineering), X+247 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

This is a textbook designed for an introductory one-term course for undergraduate or beginning graduate students majoring in engineering, the social sciences or business administration. The only prerequisite is a solid standard calculus course. Contrary to the practice followed by dozens of texts with the same aim, the present one introduces the basic notions and formulates the corresponding theorems with very great care and rigour. Of course, not all the proofs can be given by formal arguments in this framework. These are someties substituted by very nice heuristic explanations. There is a great number of well-chosen, illustrative examples and the eight chapters (Introduction and preliminary concepts, Random variables, Distribution and density functions, Expectations and characteristic functions, The binomial, Poisson, and normal distributions, The multivariate normal distribution, The transformation of random variables, Sequences of random variables) each end with a good set of homework problems. The trend is towards engineering applications. Six short appendices on integration and matrix theory help the student. Instructors of courses of the type noted above will like the book. A clean and honest work.

Sándor Csörgő (Szeged)

Livres reçus par la rédaction

R. H. Abraham—C. D. Shaw, Dynamics — The geometry of behavior, Part 3: Global behavior (The Visual Mathematics Library: Vismath Volume 3), XV+123 pages, Aerial Press, Inc., Santa Cruz. — SFr. 82.00.

D. J. Alberts—G. L. Alexanderson—C. Reid, International Mathematical Congresses. An illustrated history 1893—1986, V+64 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1987. — DM 74,—.

H. W. Alt, Lineare Funktionanalyse. Eine anwendungsorientierte Eifühmng (Hochschultext), IX+292 Seiten, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985. — DM 34,—.

Analytic Theory of Continued Fraction II. Proceedings of a Seminar-Workshop held in Pitlochry and Aviemore, Scotland, June 13—29, 1985. Edited by W. J. Thron (Lecture Notes in Mathematics, 1199), VI+299 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 50,—.

D. F. Andrews—A. M. Herzberg, DATA: A collection of problems from many fields for the student and research worker (Springer Series in Statistics), XX+442 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985. — DM 138,—.

A. N. Andrianov, Quadratic forms and Hecke operators (Grundlehren der mathematischen Wissenschaften, Bd. 286), XII+374 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1987. — DM 184,—.

Astrophysics of Brown Dwarfs. Proceedings of a Workshop held at George Mason University, Fairfax, Virginia, October 14—15, 1985. Edited by M. C. Kafatos, R. S. Harrington, S. P. Maran, IX+276 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1986. — £ 25.00.

M. Berger, Geometry I—II. (Universitext). Translated from the French by M. Cole and S. Levy, XIII+428, X+406 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1987. — DM 74,—+74,—.

J. L. Berggren, Episodes in the mathematics of medieval Islam, XIV+197 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1987. — DM 58,—.

A. L. Besse, Einstein manifolds (Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge. Vol. 10), XII+510 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1987. — DM 198,—.

T. Beth—D. Jungnickel—H. Lenz, Design theory, 688 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1986. — £ 50.00.

J. Bliedtner—W. Hansen, Potential theory. An analytic and probabilistic approach to balayage (Hochschultext), XIII+435 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 84,—.

B. Bollobás, Combinatorics. Set systems, hypergraphs, families of vectors, and combinatorial probability, XII+177 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1986.

János Bolyai, Appendix. The theory of space, 239 pages, Akadémiai Kiadó, Budapest, 1987.

I. Borg—J. Lingoes, Multidimensional similarity structure analysis, XIV+390 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1987. — DM 78,—.

U. Bottazini, The higher calculus: A history of real and complex analysis from Euler to Weierstrass, VII+332 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 96,—.

D. Braess, Nonlinear approximation theory (Springer Series in Computational Mathematics, Vol. 7), XIV+290 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 148,—.

D. Bridges—F. Richman, Varieties of constructive mathematics (London Mathematical Society Lecture Note Series, 97), X+149 pages, Cambridge University Press, Cambridge—Lonon—New York—New Rochelle—Melbourne—Sydney, 1987. — £ 10.95.

L. S. Charlap, Bieberbach groups and flat manifolds (Universitext), XIII+242 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 78,—.

The Craft of Probabilistic Modelling. A Collection of Personal Accounts. Edited by J. Gani. With contributions by numerous experts (Applied Probability), XIV+313 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 108,—.

M. Crampin—F. A. E. Pirani, Applicable differential geometry (London Mathematical Society Lecture Note Series, 59), IV+394 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1986. — £ 17.50.

Detection of Changes in Random Processes. Edited by L. Telksnys. Translation Series in Mathematics and Engineering. XIII+226 pages, Optimization Software, Inc., New York—Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 188,—.

L. Devroye, Non-uniform random variate generation, XVI+843 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 164,—.

Differential Equations in Banach Spaces. Proceedings of a Conference held in Bologna, July 2—5, 1985. Edited by A. Favini and E. Obrecht (Lecture Notes in Mathematics, Vol. 1223), VIII+299 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 50,—.

W. Dittrich—M. Reuter, Selected topics in gauge theories (Lecture Notes in Physics, Vol. 244), V+315 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 41,—.

B. Eckmann, Selecta. Edited by M. A. Knus, G. Mislin, U. Stammbach, XII+835 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1987. — DM 178,—.

Sir Arthur Eddington, Space, time and gravitation. An outline of the general relativity theory (Cambridge Science Classic), XII+218 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1987. — £ 8.95.

A. T. Fomenko—D. B. Fuchs—V. L. Gutenmacher, Homotopic topology, 310 pages, Akadémiai Kiadó, Budapest, 1986.

G. K. Francis, A topological picturebook, XV+194 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1987. — DM 78,—.

F. R. Gantmacher, Matrizentheorie. Übersetzt nach der 2., erw. Auflage der russischen Ausgabe von H. Boseck, D. Soyka, K. Stengert, 654 Seiten, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 138,—.

D. J. H. Garling, A course in Galois theory, VIII+167 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1987. — £ 8.95.

M. B. Green—J. H. Schwarz—E. Witter, Superstring theory, Vol. 1. Introduction, Vol. 2. Loop amplitudes, anomalies & phenomanology (Cambridge Monographs on Mathematical Physics), X+469, XII+596 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Sydney, 1987. — £ 32.50+37.50.

M. Gromov, Partial differential relations (Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, Vol. 9), IX+363 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 148,— .

J. Guckenheimeer—P. Holmes, Nonlinear oscillations, dynamical systems, and bifurcations of vector fields (Applied Mathematical Sciences, Vol. 42), 2nd printing, XVI+459 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 104,—.

E. Hairer—S. P. Norsett—G. Wanner, Solving ordinary differential equations I. Nonstiff problems (Springer Series in Computational Mathematics, Vol. 8), XIV+480 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1987. — DM 124,—.

J. M. Henle, An outline of set theory (Problem Books in Mathematics), VIII+145 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 58,—.

Homogenization and Effective Moduli of Materials and Media. Edited by J. L. Ericksen, D. Kinderlehrer, R. Kohn, J.-L. Lions (The IMA Volumes in Mathematics and Applications, Vol. 1), X+268 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 58,—.

J. L. Koszul, Fibre boundles and differential geometry (Tata Institute of Fundamental Research: Lectures on Mathematics and Physics, Vol. 20), IV+127 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 20,—.

S. Lang, Linear algebra (Undergraduate Texts in Mathematics), 3rd edition, IX+285 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1987. — DM 78,—.

J. P. LaSalle, The stability and control of discrete processes (Applied Mathematical Sciences, Vol. 62), VII+150 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 49,80.

T. Matolcsi, A concept of mathematical physics. Models in mechanics, 335 pages, Akadémiai Kiadó, Budapest, 1986.

K. Murota, Systems analysis by graphs and matroids, IX+281 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1987. — DM 79,—.

N. K. Nikol'skiĭ, Treatise on the shift operator. Spectral function theory (Grundlehren der mathematischen Wissenschaften, Bd. 273), With an appendix by Hruščev and V. V. Peller, XI+491 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 184,—.

Oscillation Theory, Computation, and Methods of Compensated Compactness. Edited by C. Dafermos, J. L. Ericksen, D. Kinderlehrer, M. Slemrod (The IMA Volumes in Mathematics and its Applications, Vol. 2), IX+395 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 82,—.

Pappus of Alexandria, Book 7 of the collection. Part 1: Introduction, text, and translation. Part 2: Commentary, index, and figures. Edited with translation and commentary by A. Jones (Sources in the History of Mathematics and Physical Sciences, Vol. 8), XI+375, VII+371 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 258,—.

S. Pokorski, Gauge field theories (Cambridge Monographs on Mathematical Physics), XIII+394 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Sydney, 1987. — $ 89.50.

R. M. Range, Holomorphic functions and integral representations in complex variables (Graduate Texts in Mathematics, Vol. 108), XIX+386 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 128,—.

P. J. Ryan, Euclidean and non-Euclidean geometry, XVII+215 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Sydney, 1986. — $ 14.95.

L. Sachs, A guide to statistical methods and to the pertinent literature. Literatur zur angewandten Statistik, IX+212 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 45,—.

M. Shirvani—B. A. F. Wehrfritz, Skew linear groups (London Mathematical Society Lecture Note Series, 118), VII+253 pages, Cambridge University Press, Cambridge—London—New York —New Rochelle—Sydney, 1987. — £ 15.00.

M. Shub, Global stability of dynamical systems. Translated from the French by J. Christy, XII+150 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1987. — DM 74,—.

R. I. Soare, Recursively enumerable sets and degrees. A study of computable functions and computably generated sets (Perspectives in Mathematical Logic), XVIII+437 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1987. — DM 68,—.

G. J. Székely, Paradoxes in probability and mathematical statistics, XII+250 pages, Akadémiai Kiadó, Budapest, 1986.

J. B. Thomas, Introduction to probability (Springer Texts in Electrical Engineering), X+247 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986. — DM 60,—.

J. Wloka, Partial differential equations, XI+518 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Sydney, 1987. — $ 29.95.

H. P. Yap, Some topics in graph theory (London Mathematical Society Lecture Note Series, 108), VI+230 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Sydney, 1986.

TOMUS LI—1987—51. KÖTET

Bibliographie

V. I. ARNOLD, Catastrophe Theory. — Banach Spaces. — R. S. BLYTH—E. F. ROBINSON, Algebra Through Practice. — B. BOOSS—D. D. BLEECKER, Topology and Analysis. — R. CARMONA—H. KESTEN—J. B. WALSH, École d'Été de Probabilités de Saint-Flour XIV — 1984. — G. D. CROWN—MAUREEN H. FENRICK—R. J. VALENZA, Abstract Algebra. — L. DEVROYE, Lecture Notes on Bucket Algorithms. — H. J. EICHLER—P. GÜNTER—D. W. POHL, Laser Induced Dynamic Gratings. — Goodness-of-Fit Techniques. — W. HACKBUSCH, Multi-Grid Methods and Applications. — L. HÖRMANDER, The Analysis of Linear Partial Differential Operators Vol. III. Pseudo-Differential Operators, Vol. IV. Fourier Integral

INDEX

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 0324-6523 Acta Univ. Szeged
ISSN 0001 6969 Acta Sci. Math.

INDEX: 26024

87-2871—Szegedi Nyomda — Felelős vezető: Surányi Tibor igazgató

Felelős szerkesztő és kiadó: Leindler László
A kézirat a nyomdába érkezett: 1987. június 3.
Megjelenés: 1988. február

Példányszám: 1000. Terjedelem: 21,35 (A/5) ív
Készült monószedéssel, íves magasnyomással,
az MSZ 5601-24 és az MSZ 5602-55 szabvány szerint