ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 52

FASC. 1—2

SZEGED, 1988

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

52. KÖTET

FASC. 1—2

SZEGED, 1988

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 52

SZEGED, 1988

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

52. KÖTET

SZEGED, 1988

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

Acta Sci. Math., 52 (1988), 3—19

Group theoretic results in Clifford semigroups

J. D. P. MELDRUM

Clifford semigroups or strong semilattices of groups are a class of inverse semigroups which are obviously very closely related to groups. This paper attempts to exploit this close relationship. Petrich's characterization of congruences on inverse semigroups is analyzed in this special case to obtain a description of homomorphisms and their images in terms of the groups involved. Next, the idea of classes and closure operations due to P. HALL, which has proved very useful in group theory, is extended. Some results are obtained, but there are many interesting open problems left. This is applied to nilpotency of groups and a number of interesting results are extended, in particular Fitting's Theorem, the Hirsch-Plotkin Theorem and the characterization of nilpotent groups in terms of subnormal subgroups. Finally some remarks on solubility are made. The techniques demonstrated here should lead to a very large number of results being transferred.

This paper describes a technique for applying group theoretic ideas and results to Clifford semigroups mainly by giving some examples of it in action.

I would like to thank Drs. KOWOL and MITSCH for a preprint of their paper [4] and for stimulating conversation and, later, correspondence.

I would also like to thank Dr. O'CARROLL for much help.

We refer to Howie's book [3] for background on the subject. In this paper we are exclusively concerned with Clifford semigroups and we give a definition now to establish notation.

Definition. A semigroup S is a *Clifford semigroup* or *strong semilattice of groups* if S is the disjoint union of a set of groups $\{S_\alpha\colon \alpha\in E\}$, where E is a meet semilattice and for all α, β in E such that $\alpha\geqq\beta$, there exists a homomorphism $\varphi_{\alpha,\beta}\colon S_\alpha\to S_\beta$ satisfying

$$\varphi_{\alpha,\beta}\varphi_{\beta,\gamma}=\varphi_{\alpha,\gamma}\quad\text{for all}\quad \alpha\geqq\beta\geqq\gamma\quad\text{in}\quad E.$$

Received December 13, 1984, and in revised form July 9, 1987.

The homomorphisms $\{\varphi_{\alpha,\beta};\ \alpha\geqq\beta \text{ in } E\}$ are called the *linking homomorphisms*. For all α in E, $\varphi_{\alpha,\alpha}$ is the identity map on S_α. For s_1, s_2 in S, the product is defined by

$$s_1 s_2 = (s_1\varphi_{\alpha,\alpha\beta})(s_2\varphi_{\beta,\alpha\beta})$$

where $s_1\in S_\alpha$, $s_2\in S_\beta$, $\alpha\beta$ is the join in E and the product on the right is the product in the group $S_{\alpha\beta}$.

We denote the identity of S_α by e_α. Then $\{e_\alpha\colon \alpha\in E\}$ is a semilattice of idempotents isomorphic to E, and we will often denote it by $E(S)$ or even simply E. This will not cause any confusion. Note that e_α is central in S for all $\alpha\in E$.

It can be shown from Howie [3], and it is in any case well known, that Clifford semigroups form a variety of algebras, a subvariety of the variety of inverse semigroups. Petrich [7] has defined a concept of congruence pairs for inverse semigroups and related them to congruences. This enables a link to be made between congruences and a substructure which strongly resembles normal subgroups. This correspondence is analysed closely in the context of Clifford semigroups in section 1. In section 2, some applications are made of the concept of closure operations. In section 3, we deal with extensions of the idea of nilpotency from groups to Clifford semigroups, and finally we deal with solubility in the final section.

1. Congruences on Clifford semigroups

This material is a slight extension of the results of Petrich [7] as applied to Clifford semigroups. From now on, unless explicitly stated otherwise, all semigroups are assumed to be Clifford semigroups. Let S be a semigroup, with constituent groups $\{S_\alpha\colon \alpha\in E\}$, linking homomorphisms $\{\varphi_{\alpha,\beta}\colon \alpha\geqq\beta,\ \alpha,\beta\in E\}$ and semilattice of idempotents $\{e_\alpha\colon \alpha\in E\}$.

Definition 1.1. An inverse subsemigroup T of S is called *normal* if $a^{-1}Ta\subseteqq T$ for all $a\in S$ and *full* if $E\subseteqq T$.

This definition departs from standard practice, as usually normal subsemigroups are necessarily full. We do not require this.

Definition 1.2. A pair (ϱ, N) is called a *congruence pair* if N is a normal full subsemigroup and $ae\in N$, $e\varrho a^{-1}a$ implies $a\in N$, where $a\in S$, $e\in E$.

If we define

$$a\varkappa(\varrho, N)b \quad \text{if and only if} \quad a^{-1}a\varrho b^{-1}b,\ ab^{-1}\in N$$

then Petrich [7] shows that $\varkappa(\varrho, N)$ is a congruence on S and every congruence σ on S

is of this form, where

$$\varrho = \operatorname{tr} \sigma \quad \text{(the restriction of } \sigma \text{ to } E \times E),$$
$$N = \ker \sigma := \{s\sigma e\colon\ e \in E\}.$$

Our version is simpler than his because we take advantage of the fact that S is a Clifford semigroup. We now present some fairly straightforward results concerning the concepts that we have just defined. But first a useful notational device. If T is an inverse subsemigroup of S we write T_α for $T \cap S_\alpha$. Then $T = \bigcup_{\alpha \in E} T_\alpha$. In general some of the T_α may be empty. But T is full if and only if $T_\alpha \neq \emptyset$ for all $\alpha \in E$.

Lemma 1.3.

(i) *If N is a normal inverse subsemigroup of S then N_α is a normal subgroup of S_α for all $\alpha \in E$ such that $N_\alpha \neq \emptyset$.*

(ii) *Let N be an inverse subsemigroup of S. Then $NE \subseteqq N$ if and only if $N_\alpha \varphi_{\alpha,\beta} \subseteqq N_\beta$ for all $\alpha, \beta \in E$, $\alpha \geqq \beta$.*

(iii) *Let N be an inverse subsemigroup of S such that $NE \subseteqq N$. Then N is normal in S if and only if N_α is a normal subgroup of S_α for all $\alpha \in E$ such that $N_\alpha \neq \emptyset$.*

(iv) *Let N be a full inverse subsemigroup of S. Then N is normal in S if and only if N_α is normal in S_α for all $\alpha \in E$.*

(v) *The condition in Definition* 1.2 *is equivalent to: for all $\alpha, \beta \in E$ such that $e_\alpha \varrho e_\beta$ we have $N_{\alpha\beta} \varphi^{-1}_{\alpha,\alpha\beta} \subseteqq N_\alpha$.*

(vi) *If $NE \subseteqq N$ and $N_\alpha \neq \emptyset$, then* $\ker \varphi_{\alpha,\beta} \subseteqq N_\alpha$ *for all $\beta \leqq \alpha$, $\alpha, \beta \in E$.*

(vii) *Let ϱ be a congruence on E, N a normal full subsemigroup of S. Then (ϱ, N) is a congruence pair if and only if for all $\alpha, \beta \in E$ such that $e_\alpha \varrho e_\beta$ then $N_{\alpha\beta} \varphi^{-1}_{\alpha,\alpha\beta} \subseteqq N_\alpha$.*

These results can all be checked very easily and so no details of proof will be given. We now look at the minimum group congruence σ on S. Then σ is a congruence on S such that S/σ is a group and all group images of S can be factored through S/σ. See Howie [3] p. 139.

Lemma 1.4. *Let S be a Clifford semigroup with semilattice of idempotents E. Let $X \subseteqq E$ be a chain with the property that for all $\alpha \in E$ there exists $\beta \in X$ such that $\beta \leqq \alpha$. Then S/σ is the direct limit of the chain of groups*

$$\{S_\alpha, \varphi_{\alpha,\beta}\colon\ \alpha, \beta \in X\}.$$

Note that such a chain always exists. If E has a minimal element δ, then we can take $X = \{\delta\}$ and then $S/\sigma \cong S_\delta$. A special case will be used later.

Corollary 1.5. *Using the notation of Lemma* 1.4, *assume that $\varphi_{\alpha,\beta}$ is a monomorphism for all $\alpha, \beta \in X$. Then without loss of generality we may assume $S_\alpha \subseteqq S_\beta$ for all $\alpha \geqq \beta$ and $S/\sigma = \bigcup_{\alpha \in X} S_\alpha$.*

These results do not need proving as they seem well-known, and can in any case be checked quickly. To finish this section we consider homomorphic images of Clifford semigroups. We use ε to denote the identity congruence, i.e., $a\varepsilon b$ if and only if $a=b$. It is obvious from the definition that (ε, N) is a congruence-pair for all full normal subsemigroups N of S.

Lemma 1.6. *Let ϱ be a congruence on E. Then the least full normal subsemigroup $N(\varrho)$ such that $(\varrho, N(\varrho))$ is a congruence pair is defined by*

$$N(\varrho)_\alpha = \prod_{\alpha\varrho\beta} \ker \varphi_{\alpha,\alpha\beta}.$$

In particular if $\varphi_{\alpha,\alpha\beta}$ are monomorphisms for all $\alpha, \beta\in E$ such that $\alpha\varrho\beta$, then (ϱ, E) is a congruence pair.

Again this result is easy to prove, especially if we use Lemma 1.3.

Lemma 1.7. *Let N be a full normal inverse subsemigroup of S. Let $\varkappa=\varkappa(\varepsilon, N)$, and let $T=S/\varkappa$. Then $T_\alpha=S_\alpha/N_\alpha$ and $\theta_{\alpha,\beta}\colon T_\alpha\to T_\beta$ where $\alpha\geqq\beta$ is defined by $t\theta_{\alpha,\beta}=N_\beta s\varphi_{\alpha,\beta}$, where $t=N_\alpha s$,* i.e., *$\theta_{\alpha,\beta}$ is induced naturally by $\varphi_{\alpha,\beta}$.*

This follows easily from the definitions. We finally consider a general congruence pair.

Lemma 1.8. *Let (ϱ, N) be a congruence pair on S. Let $\varkappa=\varkappa(\varrho, N)$, $T=S/\varkappa$. Let $\lambda=\varkappa(\varrho, E)$ defined on T. Then*

$$T/\lambda \cong S/\varkappa(\varrho, N).$$

If $\{A_\gamma\colon \gamma\in C\}$ are the congruence classes of ϱ on E, then T/λ is obtained from T by replacing $\bigcup_{\alpha\in A_\gamma} T_\alpha$ by its maximal group homomorphic image T_γ, and for $\gamma, \delta\in C$, $\gamma\geqq\delta$, $\psi_{\gamma,\delta}$ is defined as the natural extension of the $\theta_{\alpha,\beta}$ for $\alpha\in A_\gamma$, $\beta\in A_\delta$.

Proof. We first note that, using the notation of Lemma 1.7, the homomorphism $\theta_{\alpha,\beta}\colon T_\alpha\to T_\beta$ is a monomorphism. Hence (ϱ, E) is a congruence pair on T by Lemma 1.6. Hence T_γ can be written as a union of a tower of groups as described in Lemma 1.4. This makes the definition of $\psi_{\gamma,\delta}$ easy to verify. All the rest is very easy to check.

2. Closure operations on classes

We use the ideas of classes of groups and closure operations as developed by P. Hall and apply them to Clifford semigroups. A good presentation of these can be found in Robinson [8] chapter 1, section 1. They have also been used in many other settings by many other people. In particular Cohn [1] uses them in the context of universal algebras.

The only condition imposed on a class $\mathfrak{X}$ of groups is that $\{e\}\in\mathfrak{X}$ and if $G\in\mathfrak{X}$ and $H\cong G$ then $H\in\mathfrak{X}$. A closure operation on classes of groups is a map A from classes of groups to classes of groups $A\colon \mathfrak{X}\to A\mathfrak{X}$ satisfying $A\mathfrak{X}\supseteq\mathfrak{X}$, $\mathfrak{X}\subseteq\mathfrak{Y}$ implies $A\mathfrak{X}\subseteq A\mathfrak{Y}$ and $AA\mathfrak{X}=A\mathfrak{X}$. A class $\mathfrak{X}$ is A-closed if $A\mathfrak{X}=\mathfrak{X}$. Any intersection of A-closed classes is A-closed. Hence to define A we only need to specify the A-closed classes. For then $A\mathfrak{X}=\bigcap\{\mathfrak{Y}\colon \mathfrak{Y}\supseteq\mathfrak{X},\ A\mathfrak{Y}=\mathfrak{Y}\}$. The concept of classes and closure operations can be transferred to any other algebraic structure, and, in particular, to Clifford semigroups.

Definition 2.1. For a class $\mathfrak{X}$ of groups, we define $\mathfrak{X}_S$ to be the class of Clifford semigroups given by

$$S\in\mathfrak{X}_S \quad \text{if and only if} \quad S_\alpha\in\mathfrak{X} \quad \text{for all} \quad \alpha\in E(S).$$

This gives the natural extension of the definition of a class of groups to a class of Clifford semigroups. We will see later that this extension of the definition is not always the most useful one. There is immediately a family of questions which can be posed.

Problem 2.2. Given a class $\mathfrak{X}$ of groups and a closure operation A on classes, determine whether $A\mathfrak{X}_S=(A\mathfrak{X})_S$. Alternatively if $A\mathfrak{X}=\mathfrak{X}$, is $A\mathfrak{X}_S=\mathfrak{X}_S$?

We will deal with a few cases of this problem, but there is a great deal more that can be done in this area. We first define the closure operations which we will be using, to cover both groups and Clifford semigroups.

The class $\mathfrak{X}$ is S closed if every substructure of an $\mathfrak{X}$ structure is itself an $\mathfrak{X}$-structure.

The class $\mathfrak{X}$ is Q closed (sometimes the symbol H is used) if every epimorphic image of an $\mathfrak{X}$ structure is itself an $\mathfrak{X}$ structure.

The class $\mathfrak{X}$ is R closed if given a structure Y such that a family of homomorphisms $\{\theta_i\colon i\in I\}$ exists with $Y\theta_i\in\mathfrak{X}$ for all $i\in I$ and $\bigcap_{i\in I}\ker\theta_i$ is trivial, then $Y\in\mathfrak{X}$. We say $\mathfrak{X}$ is residually closed.

The class $\mathfrak{X}$ is L closed if given a structure Y such that every finite subset $\{y_1, \ldots, y_n\}$ of Y is contained in an $\mathfrak{X}$ substructure of Y, then $Y\in\mathfrak{X}$.

The class $\mathfrak{X}$ is $N(N_0)$ closed if every structure Y which can be expressed as a product of a (finite) number of normal substructures is again in $\mathfrak{X}$.

Lemma 2.3. *Let* $S\mathfrak{X}=\mathfrak{X}$. *Then* $S\mathfrak{X}_S=\mathfrak{X}_S$.

Proof. Let $T\in S\mathfrak{X}_S$. Then there exists $U\in\mathfrak{X}_S$ such that T is a Clifford subsemigroup of U. Hence for all $\alpha\in E(U)$, T_α is a subgroup of U_α or is empty. But $U_\alpha\in\mathfrak{X}=S\mathfrak{X}$. Hence $T_\alpha\in\mathfrak{X}$ or is empty. Thus $T\in\mathfrak{X}_S$. Thus $\mathfrak{X}_S=S\mathfrak{X}_S$.

Example 2.4. Let $\mathfrak{X}$ be the class of finite p-groups for some prime p. Let the semilattice E be the set of negative integers with the natural order inducing the semilattice structure. So $(-n)\cdot(-m)=\min\{-n, -m\}$. Let S_{-n} be the cyclic group of order p^n, $\varphi_{-n,-m}$ for $n\leqq m$ be the natural embedding. Then S, the Clifford semigroup so defined has as maximal group homomorphic image the group C_{p^∞}, the Prüfer group of type p^∞, which is certainly not a finite p-group. So in this case $\mathfrak{X}=Q\mathfrak{X}$ but $Q\mathfrak{X}_S\neq\mathfrak{X}_S$.

The problem with Q closure occurs because group homomorphic images of Clifford semigroups include direct limits. This leads to the following result.

Lemma 2.5. *Let $\mathfrak{X}$ be a class of groups closed under the operation of taking direct limits. Then $\mathfrak{X}_S$ is Q closed.*

Proof. Let $S\in Q\mathfrak{X}_S$. Then S is the homomorphic image of a semigroup $T\in\mathfrak{X}_S$. From Lemma 1.8, we see that the component groups of S are obtained from those of T by taking homomorphic images and direct limits. Since the component groups of T lie in $\mathfrak{X}$ and $\mathfrak{X}$ is closed under direct limits, and hence Q closed, it follows that $S\in\mathfrak{X}_S$. This finishes the proof.

We next look at L closure. First we prove a result used later.

Lemma 2.6. *Let $\{s_1, \ldots, s_n\}$ be a finite subset of a Clifford semigroup S. Then the inverse subsemigroup of S generated by $\{s_1, \ldots, s_n\}$ is contained in the union of a finite number of finitely generated groups forming a semigroup.*

Proof. Let $E=E(S)$, and let X be the finite subset of E defined by $\alpha\in X$ if and only if $s_i\in S_\alpha$ for some i, $1\leqq i\leqq n$. Then X generates a finite subsemilattice Y of E. For all $\beta\in Y$, we define

$$Z_\beta=\{s_i\varphi_{\alpha,\beta}\colon\ 1\leqq i\leqq n,\ \alpha\geqq\beta,\ \alpha\in Y,\ s_i\in S_\alpha\}.$$

Then Z_β is a finite subset of S_β and so generates a finitely generated subgroup G_β of S_β. It is routine to check that the inverse subsemigroup of S generated by $\{s_1, \ldots, s_n\}$ is contained in $\bigcup_{\beta\in Y} G_\beta$, and this is a semigroup, which is all we wished to show.

Lemma 2.7. *Let $\mathfrak{X}=L\mathfrak{X}$. Then $L\mathfrak{X}_S=\mathfrak{X}_S$.*

Proof. Let $S\in L\mathfrak{X}_S$. We need to show that $S_\alpha\in\mathfrak{X}$ for all $\alpha\in E=E(S)$. Let $\{s_1, \ldots, s_n\}$ be a finite subset of S_α. Then $\{s_1, \ldots, s_n\}\subseteqq T\in\mathfrak{X}_S$, T an inverse subsemigroup of S. In particular $T_\alpha\supseteqq\{s_1, \ldots, s_n\}$ and lies in $\mathfrak{X}$. Thus $S_\alpha\in L\mathfrak{X}=\mathfrak{X}$. Hence the result is true.

Lemma 2.8. *Let $\mathfrak{X}=Q\mathfrak{X}=S\mathfrak{X}=L\mathfrak{X}$. Then $\mathfrak{X}_S=Q\mathfrak{X}_S=S\mathfrak{X}_S=L\mathfrak{X}_S$.*

Proof. Following Lemma 2.3 and Lemma 2.7, we only need to show that

$\mathfrak{X}_S = Q\mathfrak{X}_S$. Let $S \in Q\mathfrak{X}_S$, S a homomorphic image of $T \in \mathfrak{X}_S$. From Lemma 1.8, each S_α is obtained from $\{T_\beta\colon \beta \in E(T)\}$ by taking homomorphic images and unions of towers. Let $\{G_\gamma\colon \gamma \in X\}$ be a tower of groups in $\mathfrak{X}$, $G = \bigcup_{\gamma \in X} G_\gamma$. Then any finite subset of G is contained in G_γ for some γ, and $G_\gamma \in \mathfrak{X}$. Hence $G \in L\mathfrak{X} = \mathfrak{X}$. Thus each $S_\alpha \in \mathfrak{X}$ and $S \in \mathfrak{X}_S$.

For any class $\mathfrak{X}$ we denote by $V\mathfrak{X}$ the least variety containing $\mathfrak{X}$. It is a standard result from universal algebra that $V\mathfrak{X} = \mathfrak{X}$ if and only if $\mathfrak{X} = S\mathfrak{X} = Q\mathfrak{X} = R\mathfrak{X}$. (Cohn [1] IV. 3). We now state

Lemma 2.9. *Let $\mathfrak{X}$ be a class of groups. Then $\mathfrak{X}_S$ is a variety if and only if $\mathfrak{X}$ is a variety.*

This is an easy consequence of known results (Petrich [6]) or can be proved directly without much trouble.

Corollary 2.10. *$\mathfrak{X}$ is Q, R, S closed if and only if $\mathfrak{X}_S$ is Q, R, S closed.*

3. Nilpotency and its generalizations

Let $\mathfrak{N}$ be the class of nilpotent groups, and let $\mathfrak{N}_c$ be the class of nilpotent groups of nilpotency class at most c. Then $\mathfrak{N}_c$ is a variety and $\mathfrak{N} = \bigcup_{c \geq 1} \mathfrak{N}_c$. The most obvious generalization of $\mathfrak{N}$ to Clifford semigroups is $\mathfrak{N}_S$, but this leads to problems as we now see.

Example 3.1. Let G_n be a nilpotent group of nilpotency class exactly n, in particular let G_n be the group of $(n+1)\times(n+1)$ unitriangular matrices over some field F. Then we can embed G_n in G_{n+1} by mapping $(a_{ij}) \in G_n \to (b_{ij}) \in G_{n+1}$, where for $j > i$, $a_{ij} = b_{ij+1}$, $b_{ii+1} = 0$. Let S be the Clifford semigroup whose semilattice of idempotents is isomorphic to the negative integers with the natural order. Compare Example 2.4. For each $-n \in E$, let $S_{-n} = G_n$ and $\varphi_{-n,\,-m}$ be the embedding obtained from the embeddings outlined above. Then $S \in \mathfrak{N}_S$, but S has as a homomorphic image $G = \bigcup_{n \geq 1} G_n$, the maximal group homomorphic image of S. And G is not nilpotent, since it contains subgroups of arbitrarily high nilpotency class.

Because of this example, we make the following definition.

Definition 3.2. The class of *nilpotent Clifford semigroups* is defined to be

$$\hat{\mathfrak{N}} = \bigcup_{c \geq 1} (\mathfrak{N}_c)_S.$$

Hence $S\in\hat{\mathfrak{N}}$ if and only if $S_\alpha\in\mathfrak{N}_c$ for all $\alpha\in E=E(S)$, and some $c=c(S)$. This coincides with LALLEMENT's definition [5]. As KOWOL and MITSCH dealt with finite semigroups, either definition would have served. In the infinite case this definition leads to a more satisfactory theory. Denote $(\mathfrak{N}_c)_S$ by $\hat{\mathfrak{N}}_c$.

Lemma 3.3. *$\hat{\mathfrak{N}}$ is S and Q closed.*

Proof. Let $T\in S\hat{\mathfrak{N}}$. Then there exists $U\in\hat{\mathfrak{N}}$ and T is a subsemigroup of U. So $U\in\hat{\mathfrak{N}}_c$ and $S\mathfrak{N}_c=\mathfrak{N}_c$. By Lemma 2.3 $S\hat{\mathfrak{N}}_c=\hat{\mathfrak{N}}_c$, hence $T\in\hat{\mathfrak{N}}_c\subseteqq\hat{\mathfrak{N}}$. The case of Q closure follows the same pattern, using Corollary 2.10 since $\mathfrak{N}_c$ is a variety.

We now introduce upper and lower central series for Clifford semigroups which extend the corresponding ideas for groups, as was done in Kowol and Mitsch [4].

Definition 3.4. Let S be a Clifford semigroup, N_i full normal subsemigroups of S for $0\leqq i\leqq r$.

(i) $Z(S)$, the *centre* of S is defined by $Z(S)=\{x\in S\colon xs=sx$ for all $s\in S\}$.

(ii) Let H, K be inverse subsemigroups of S. Define $[H, K]$ to be the inverse subsemigroup of S generated by

$$\{[h, k] = h^{-1}k^{-1}hk\colon h\in H, k\in K\}.$$

(iii) A sequence

$$E(S) = N_0 \subseteqq N_1 \subseteqq \dots \subseteqq N_r = S$$

is called a central series of S if

$$N_i \subseteqq Z(S/\varkappa(\varepsilon, N_{i-1}))\theta_{i-1}^{-1}$$

for $1\leqq i\leqq r$, where θ_{i-1} is the natural homomorphism associated with $\varkappa(\varepsilon, N_{i-1})$.

(iv) The upper central series of S is defined inductively by

$$Z_0(S) = E(S),$$
$$Z_{i+1}(S)\theta_i = Z\big(S/\varkappa(\varepsilon, Z_i(S))\big),$$

for $i\geqq 0$, where θ_i is the natural homomorphism associated with $\varkappa\big(\varepsilon, Z_i(S)\big)$ and $Z_{i+1}(S)$ is maximal such.

(v) The lower central series of S is defined inductively by

$$\gamma_1(S) = S,$$
$$\gamma_{i+1}(S) = [S, \gamma_i(S)],$$

for $i\geqq 1$.

We now list some easy consequences of this composite definition.

Lemma 3.5. *Let S be a Clifford semigroup.*

(i) *$Z_i(S)$ is a normal full subsemigroup of S for all $i\geqq 0$.*

(ii) *$\gamma_i(S)$ is a normal full subsemigroup of S for all $i\geqq 1$.*

(iii) *$S/\varkappa(\varepsilon, N)$ is commutative if and only if $N \supseteqq \gamma_2(S)$, where N is a normal full subsemigroup of S.*

(iv) *$[s_1, s_2] \in E(S)$ if and only if $s_1 s_2 = s_2 s_1$.*

Proof. This is all easy to prove or can be deduced easily from Section 3 of Kowol and Mitsch [4].

Lemma 3.6. *Let S be a Clifford semigroup. Then*

$$\gamma_i(S) = \bigcup_{\alpha \in E} \gamma_i(S_\alpha).$$

Proof. Obviously $\gamma_i(S_\alpha) \subseteqq \gamma_i(S)$ for all $\alpha \in E$. Conversely we prove by induction on i that $\gamma_i(S) \subseteqq \bigcup_{\alpha \in E} \gamma_i(S_\alpha)$. This is true trivially for $i=1$. So assume that this is true for i. Let $s \in S$, $t \in \gamma_i(S)$. Then $[s, t] = s^{-1}t^{-1}st = (s\varphi_{\alpha,\alpha\beta})^{-1}(t\varphi_{\beta,\alpha\beta})^{-1} \cdot \cdot (s\varphi_{\alpha,\alpha\beta})(t\varphi_{\beta,\alpha\beta})$, where $s \in S_\alpha$, $t \in S_\beta$. So $[s, t] \in [S_{\alpha\beta}, \gamma_i(S_{\alpha\beta})]$ using the induction hypothesis. This suffices to prove the result since now the generators of $\gamma_{i+1}(S)$ lie in $\bigcup_{\alpha \in E} \gamma_{i+1}(S_\alpha)$ and this is easily checked to be a normal full subsemigroup.

Lemma 3.7. *The upper and lower central series of a Clifford semigroup are central series.*

Proof. This is immediate from Definition 3.4 and Lemma 3.5.

Theorem 3.8. *Let S be a Clifford semigroup with a central series*

$$E(S) = N_0 \subseteqq N_1 \subseteqq \dots \subseteqq N_r = S. \tag{3.9.}$$

Then $Z_r(S) = S$, $\gamma_{r+1}(S) = S$ and for all i, $0 \leqq i \leqq r$, $N_i \subseteqq Z_i(S)$ and $N_{r-i} \supseteqq \gamma_{1+i}(S)$.

Proof. We only need to prove the two inequalities which we do by induction on i. Both are true trivially for $i=0$. Assume that both are true for i. Let $x \in N_{i+1}$. Then $N_{i+1}\theta_i \subseteqq Z(S/\varkappa(\varepsilon, N_i))$, where θ_i is the natural homomorphism associated with $\varkappa(\varepsilon, N_i)$. Let $s \in S$. Then $(xs)\theta_i = x\theta_i s\theta_i = s\theta_i x\theta_i$ since (3.9) is central, and so $xs\varkappa(\varepsilon, N_i)sx$. Since $Z_i(S) \supseteqq N_i$, it follows that $xs\varkappa(\varepsilon, Z_i(S))sx$ for all $s \in S$. Thus $x\varphi_i \in Z(S/\varkappa(\varepsilon, Z_i(S))) = Z_{i+1}(S)\varphi_i$, where φ_i is the natural homomorphism associated with $\varkappa(\varepsilon, Z_i(S))$. Hence $x \in Z_{i+1}(S)$. Thus $N_{i+1} \subseteqq Z_{i+1}(S)$.

Let $x \in N_{r-i}$, $s \in S$ now. Then $xs\varkappa(\varepsilon, N_{r-i-1})sx$ as before. So $x^{-1}s^{-1}xs \in N_{r-i-1}$. Thus $[N_{r-i}, S] \subseteqq N_{r-i-1}$. Hence $\gamma_{2+i}(S) = [\gamma_{1+i}(S), S] \subseteqq [N_{r-i}, S] \subseteqq N_{r-i-1}$ using the induction hypothesis. This finishes the induction step for both inequalities and hence the proof of the theorem.

Corollary 3.9. *A Clifford semigroup S is in $\hat{\mathfrak{N}}$ if and only if there exist c and d such that $Z_c(S) = S$, $\gamma_{d+1}(S) = E(S)$ and the least such c and d satisfy $c = d$.*

This is the *nilpotency class* of S and is the least c such that $S\in\hat{\mathfrak{N}}_c$.

This result follows directly from Lemmas 3.6, 3.7 and Theorem 3.8. Notice the close connection with the work of Kowol and Mitsch [4], Section 4. We now prove a selection of theorems about nilpotency and its generalizations in Clifford semigroups by transferring the results from group theory. As source book for the group theoretic results any standard text book will serve. We mention particularly Hall [2], an excellent account of the particular areas under consideration here, but not widely available, and also Robinson [8] and Scott [10].

Theorem 3.10. *Let S be a nilpotent Clifford semigroup. Then elements of coprime order commute.*

Proof. The order of $s\in S$ is its order in S_α, where $s\in S_\alpha$, i.e., the least integer $n>0$ such that $s^n\in E(S)$. Let $s_1, s_2\in S$. If $s_1\in S_\alpha$, $s_2\in S_\beta$, then $s_1s_2= =s_1\varphi_{\alpha,\alpha\beta}s_2\varphi_{\beta,\alpha\beta}=s_2\varphi_{\beta,\alpha\beta}s_1\varphi_{\alpha,\alpha\beta}$, using the group theoretic result in $S_{\alpha\beta}$. Since the order of $s\varphi_{\gamma,\delta}$ divides the order of s, the result follows.

Theorem 3.11. *In a torsion nilpotent Clifford semigroup, the elements of order a power of p, a prime, form an inverse subsemigroup.*

Theorem 3.12. *In a nilpotent Clifford semigroup, the elements of finite order form an inverse subsemigroup, the torsion subsemigroup.*

These both follow immediately from Theorem 3.10, and the corresponding results from group theory. Most of the results from Section 4 of Kowol and Mitsch [4] can be obtained by transferring from group theory, and we will not repeat them here. The exception to this is Theorem 4.3 on the representation of an element of a nilpotent Clifford semigroup as a product of elements of prime power order.

Theorem 3.13. *Let S be a torsion nilpotent Clifford semigroup, and let $\{P_i: i\in I\}$ be the Sylow subsemigroups of S,* i.e., $P_i=\{s\in S$: *order of s is a power of $p_i\}$, where $\{p_i: i\in I\}$ are a set of distinct primes. If $s\in S$, then $s=a_1\dots a_n$ is a uniquely defined representation of s, where $a_i\in S_\alpha\cap P_i$, α is defined by $s\in S_\alpha$, $1\leqq i\leqq n$, a finite subset of I.*

This follows directly from the group theoretic result. This seems to be the only uniqueness result of this kind, applicable in general. But under very special circumstances, there is a maximal version of the theorem.

Theorem 3.14. *Let S be a torsion nilpotent Clifford semigroup such that $E=E(S)$ is a lattice with the maximal condition, and such that all linking homomorphisms are monomorphisms. Let $\{P_i: i\in I\}$ be the Sylow subsemigroups of S, where $\{p_i: i\in I\}$ are a set of distinct primes. If $s\in S$, then $s=b_1\dots b_n$ is a uniquely defined* *representation of s, where* $b_i \in P_i \cap S_{\beta(i)}$, *and* $\beta(i)$ *is defined by* $\beta(i)$ *is maximal in* E *such that* $b_i \varphi_{\beta(i),\alpha} = a_i$, *using the notation of Theorem* 3.13.

Proof. Since E is a lattice with the maximum condition, $\beta(i)$ is unique. Since $\varphi_{\beta(i),\alpha}$ is a monomorphism b_i is uniquely defined, since a_i is unique given Theorem 3.13.

From the proof of Theorem 3.14, it is obvious how examples could be constructed to show that $\beta(i)$ has to be uniquely defined, and that $\varphi_{\beta(i),\alpha}$ has to be a monomorphism, to obtain a unique "maximal" representation.

The next results we will prove are the Clifford semigroup theoretic versions of famous group theoretic results on nilpotency. The first is Fitting's Theorem, the one about normal nilpotent subgroups.

Lemma 3.15. *Let S be a Clifford semigroup. Let N be a normal inverse subsemigroup of S, T an inverse subsemigroup of S. Then* $NT = TN$ *is an inverse subsemigroup of S. Also* $(NT)_\alpha = N_\alpha T_\alpha$, *if* $TE = T$, *for all* $\alpha \in E$. *If T is normal, then so is NT.*

Proof. Let $n_1 t_1, n_2 t_2 \in NT$, where $n_i \in N$, $t_i \in T$, $i = 1, 2$. Then $n_1 t_1 n_2 t_2 = n_1 t_1 t_1^{-1} t_1 n_2 t_2 = n_1 t_1 n_2 t_1^{-1} t_1 t_2 = n_1 n_3 t_1 t_2 \in NT$. So NT is a subsemigroup. Let $tn \in TN$. Then $tn = tt^{-1}tn = tnt^{-1}t = n't \in NT$. Hence $TN \subseteq NT$. Similarly $NT \subseteq TN$. Thus $NT = TN$ is an inverse subsemigroup as $(nt)^{-1} = t^{-1}n^{-1} \in TN = NT$. We now show that $(NT)_\alpha = N_\alpha T_\alpha$. Certainly $N_\alpha T_\alpha \subseteq (NT)_\alpha$. Let $nt \in (NT)_\alpha$. Then there exist $\beta \geqq \alpha, \gamma \geqq \alpha$ such that $\beta\gamma = \alpha$, $n \in N_\beta$, $t \in T_\gamma$ and $nt = n\varphi_{\beta,\alpha} t\varphi_{\gamma,\alpha}$. But $n\varphi_{\beta,\alpha} \in N_\alpha$, $t\varphi_{\gamma,\alpha} \in T_\alpha$. Hence $(NT)_\alpha \subseteq N_\alpha T_\alpha$. Thus $N_\alpha T_\alpha = (NT)_\alpha$. Finally let T be also normal and let $nt \in NT$, $s \in S$. Then $s^{-1}nts = s^{-1}ntss^{-1}s = s^{-1}nss^{-1}ts \in NT$, for all $s \in S$. Hence the whole lemma is proved.

This result extends directly Lemma 2.4 of Kowol and Mitsch [4]. We now come to Fitting's Theorem.

Theorem 3.16. *Let S be a Clifford semigroup. The product of two normal nilpotent subsemigroups of S is normal and nilpotent.*

Proof. Let N, M be normal and nilpotent subsemigroups of S. Then NM is a normal subsemigroup by Lemma 3.15. Also $(NM)_\alpha = N_\alpha M_\alpha$ for all $\alpha \in E$. Suppose $N \in \hat{\mathfrak{N}}_c$, $M \in \hat{\mathfrak{N}}_d$, then $N_\alpha \in \mathfrak{N}_c$, $M_\alpha \in \mathfrak{N}_d$ and by standard group theory, $N_\alpha M_\alpha \in \mathfrak{N}_{c+d}$. Hence $(NM)_\alpha \in \mathfrak{N}_{c+d}$ for all $\alpha \in E$, and $NM \in \hat{\mathfrak{N}}_{c+d}$.

Corollary 3.17. *Let* $N \in \hat{\mathfrak{N}}_c$, $M \in \hat{\mathfrak{N}}_d$ *be normal nilpotent subsemigroups of S. Then* $NM \in \hat{\mathfrak{N}}_{c+d}$.

Corollary 3.18. *Let S be a Clifford semigroup which satisfies the maximal condition on normal subsemigroups. Then S contains a unique maximal normal nilpotent* *subsemigroup containing all normal nilpotent subsemigroups, called the Fitting subsemigroup.*

The next result which we extend is the Hirsch—Plotkin Theorem.

Theorem 3.19. *Let S be a Clifford semigroup. Then the product of two normal locally nilpotent subsemigroups is a normal locally nilpotent subsemigroup. There is a unique maximal normal locally nilpotent subsemigroup, containing all normal locally nilpotent subsemigroups, the Hirsch—Plotkin radical of S.*

Proof. Because of Lemma 3.15 we only need to show, for the first part, that if $N, M \in L\hat{\mathfrak{N}}$ are normal, then $NM \in L\hat{\mathfrak{N}}$. Since $N \in L\hat{\mathfrak{N}}$, it follows that N_α, M_α are locally nilpotent groups which are normal in S_α. Let $\{n_1 m_1, \ldots, n_r m_r \colon n_i \in N, m_i \in M\}$ be a finite subset of NM. Let $Z = \{n_1, \ldots, n_r, m_1, \ldots, m_r\}$. By Lemma 2.6, T, the inverse subsemigroup generated by Z, is generated by a finite set of elements of the form $n_i \varphi_{\beta,\alpha}, m_j \varphi_{\gamma,\delta}$. T_α is generated as a group by a finite set of the form $\{n_i \varphi_{\beta,\alpha}, m_j \varphi_{\gamma,\alpha}\}$, which is a finite subset of $N_\alpha M_\alpha$, the product of two locally nilpotent normal subgroups of S_α. Hence $N_\alpha M_\alpha$ is locally nilpotent by the Hirsch—Plotkin Theorem and thus T_α is nilpotent. T is the union of a finite number of groups of the form T_α. Hence we can find c such that $T_\alpha \in \mathfrak{N}_c$ for all T_α, and so $T \in \hat{\mathfrak{N}}_c$. Since $\{n_1 m_1, \ldots \ldots, n_r m_r\} \subseteqq T$, we have shown that $NM \in L\hat{\mathfrak{N}}$.

The last part follows as in the group case. The product of any finite set of normal locally nilpotent subsemigroups is locally nilpotent by the first part. Consider the product H of all the normal locally nilpotent subsemigroups of S. It is normal and any finite subset of H is contained in the product of a finite number of normal locally nilpotent subsemigroups which is locally nilpotent, hence is contained in a nilpotent subsemigroup. Thus H is locally nilpotent. This finishes the proof.

The next result which we extend is a well-known one concerning minimal normal subgroups of locally nilpotent groups.

Theorem 3.20. *Let S be a locally nilpotent Clifford semigroup, N a minimal normal subsemigroup of S. Then there exists a unique $\alpha \in E$ such that $N_\alpha \supset \{e_\alpha\}$ and $N_\alpha \subseteqq Z(S_\alpha)$, and for all $\beta \leqq \alpha$, we have* $\ker \varphi_{\alpha,\beta} \supseteqq N_\alpha$.

Proof. By Lemma 1.3, it is easy to see that if there are two elements α, β of E such that $N_\alpha \supset \{e_\alpha\}$, $N_\beta \supset \{e_\beta\}$, then N is not minimal. If S is locally nilpotent, then so is S_α. So N_α is a normal subgroup of S_α such that for all $\beta \in E$, $\beta \leqq \alpha$, $N_\alpha \subseteqq \ker \varphi_{\alpha,\beta}$. It follows that N_α can be replaced by any normal subgroup of S_α contained in it, and we would still have a minimal normal subsemigroup. Then minimality of N forces N_α to be a minimal normal subgroup of S_α, hence by group theory $N_\alpha \subseteqq Z(S_\alpha)$.

The last results about nilpotency which we will present concern normalizers.

Definition 3.21. Let T be an inverse subsemigroup of a Clifford semigroup S. The *normalizer* $N_S(T)$ of T in S is the unique largest inverse subsemigroup of S in which T is normal.

A priori $N_S(T)$ may not always exist. We will show that it does.

Lemma 3.22. *Let T be an inverse subsemigroup of a Clifford semigroup S. Then $N_S(T)$ always exists and is defined by*

$$N_S(T) = \{x \in S\colon\ x^{-1}Tx \subseteqq T\}.$$

Proof. If U defined to be $\{x\colon\ x^{-1}Tx \subseteqq T\}$ is an inverse subsemigroup, then it must be $N_S(T)$. Now U is obviously closed under products. Let $x \in U$. Then $xTx^{-1} \supseteqq xx^{-1}Txx^{-1} = Txx^{-1}$. But $x^{-1}Tx \subseteqq T$. So if $t \in T_\beta$ and $x \in S_\alpha$ then $x^{-1}tx = x^{-1}\varphi_{\alpha,\alpha\beta}t\varphi_{\beta,\alpha\beta}x\varphi_{\alpha,\alpha\beta} \in T$. Hence $T_{\alpha\beta} \neq \varphi$. So $txx^{-1} = t\varphi_{\beta,\alpha\beta}e_{\alpha\beta} = te_{\alpha\beta} \in T$, since $e_{\alpha\beta} \in T_{\alpha\beta} \subseteqq T$.

Definition 3.23. An inverse subsemigroup T of a Clifford semigroup S is called *subnormal* if there exists a sequence of inverse subsemigroups

$$T = T_0 \subseteqq T_1 \subseteqq \ldots \subseteqq T_n = S$$

such that T_i is normal in T_{i+1} for $0 \leqq i \leqq n-1$. The least length n of such a series is called the *index of subnormality.*

Theorem 3.24. *Let S be a nilpotent Clifford subsemigroup, T an inverse subsemigroup such that $TE \subseteqq T$. Then T is subnormal of index at most c where c is the nilpotency class of S.*

Proof. We show that if $\{Z_i\colon\ 0 \leqq i \leqq c\}$ is the upper central series of S, then TZ^i is normal in TZ_{i+1}, replacing TZ_c by S. Note that $Z_0 = E$, so $T = TE = TZ_0$. By Lemma 3.15 TZ_i is an inverse subsemigroup of S. Let $x \in Z_{i+1}, y \in TZ_i$. Then $x^{-1}yx = x^{-1}yy^{-1}yx = yy^{-1}x^{-1}yx = y[y, x] \in TZ_i$ since $y \in TZ_i$ and $[y, x] \in Z_i$ since $x \in Z_{i+1}$. Thus $Z_{i+1} \subseteqq N_S(TZ_i)$. This is enough to prove the result. If $i = c-1$, then $Z_c = S \subseteqq N_S(TZ_{c-1})$, $E \subseteqq N_S(TZ_{c-1})$ so $S = SE \subseteqq N_S(TZ_{c-1})$.

We could have used the group theoretic results and transferred them. But the details of the links to the group theory would be longer than the direct proof, which parallels very closely the group theory proof.

Theorem 3.25. *Let S be a Clifford semigroup with the property that all its full inverse subsemigroups are subnormal of index at most c. Then $S \in \hat{\mathfrak{N}}_d$ where d is a function of c.*

Proof. Let $\alpha \in E$ and consider U a subgroup of S_α. Let T be a full inverse subsemigroup of S such that $T_\alpha = U$, e.g. $T_\beta = \{e_\beta\}$ if $\beta \nleqq \alpha$, $T_\beta = U\varphi_{\alpha,\beta}$ if $\beta \leqq \alpha$.

Then $T=T_0 \subseteqq T_1 \subseteqq \dots \subseteqq T_c=S$ is a sequence such that T_i is normal in T_{i+1} for $0 \leqq i \leqq c-1$. In particular $T_{i,\alpha}$ is a normal subgroup of $T_{i+1,\alpha}$. Hence $T_{0,\alpha}=U$ is subnormal of index at most c in S_α. This is true for all subgroups of S_α. By Roseblade [9], S_α is nilpotent of class at most $f(c)=d$ say. Hence $S \in \hat{\mathfrak{N}}_d$.

Corollary 3.26. *Let S be a finite Clifford semigroup such that all its full inverse subsemigroups are subnormal. Then S is nilpotent.*

The result that gives as a sufficient condition for a finite group to be nilpotent that all its maximal subgroups are normal does not carry over in the most obvious way.

Example 3.27. Let E consist of three elements α, β and $\alpha\beta=\gamma$. With $S_\alpha \cong C_2 \cong S_\beta$ a cyclic group of order 2, S_γ the symmetric group on three symbols. Then $\varphi_{\alpha,\gamma}\colon S_\alpha \to \{e_\gamma, (12)\}$, $\varphi_{\beta,\gamma}\colon S_\beta \to \{e_\gamma, (13)\}$ defines $S=S_\alpha \cup S_\beta \cup S_\gamma$ as a Clifford semigroup. It is easy to check that the only maximal inverse subsemigroups are $E \cup S_\gamma \cup S_\alpha$ and $E \cup S_\gamma \cup S_\beta$, both normal. But S is not nilpotent.

We leave the reader to find some possible generalizations of this result.

4. Solubility

Let $\mathfrak{S}$ be the class of soluble groups, and $\mathfrak{S}_d$ the class of soluble groups of solubility class at most d. Then $\mathfrak{S}_d$ is a variety and $\mathfrak{S}=\bigcup_{d \geqq 1} \mathfrak{S}_d$. Example 3.1 shows that $\mathfrak{S}_S$ again leads to problems. The semigroup S of Example 3.1 is in $\mathfrak{S}_S$, but its maximal group homomorphic image G is not soluble, although it is a homomorphic image of S.

Definition 4.1. The class of *soluble* Clifford semigroups is defined to be

$$\hat{\mathfrak{S}} = \bigcup_{d \geqq 1} (\mathfrak{S}_d)_S.$$

Hence $S \in \hat{\mathfrak{S}}$ if and only if $S_\alpha \in \mathfrak{S}_d$ for all $\alpha \in E$ and some $d=d(S)$. Denote $(\mathfrak{S}_d)_S$ by $\hat{\mathfrak{S}}_d$. Lemma 3.3 extends very easily.

Lemma 4.2. *$\hat{\mathfrak{S}}$ is S and Q closed.*

Definition 4.3. Let S be a Clifford semigroup. The *derived series* of S is defined to be

$$\delta_0(S) = S, \quad \delta_{i+1}(S) = [\delta_i(S), \delta_i(S)].$$

A sequence

$$E(S) = N_r \subseteqq N_{r-1} \subseteqq \dots \subseteqq N_0 = S$$

is called an *abelian series* of S if N_i is normal in N_{i-1} and $N_{i-1}/\varkappa(\varepsilon, N_i)$ is commutative for $r \geqq i \geqq 1$.

Lemma 4.4. *Let S be a Clifford semigroup. Then $\delta_i(S)$ is a full normal subsemigroup of S for all $i \geqq 1$.*

Lemma 4.5. *Let S be a Clifford semigroup. Then*

$$\delta_i(S) = \bigcup_{\alpha \in E} \delta_i(S_\alpha).$$

Lemma 4.6. *The derived series of S is an abelian series.*

These results all follow in much the same way as the corresponding results at the beginning of Section 3.

Theorem 4.7. *Let S be a Clifford semigroup with an abelian series*

$$E(S) = N_r \subseteqq N_{r-1} \subseteqq \dots \subseteqq N_0 = S.$$

Then $N_i \supseteqq \delta_i(S)$ for all $i \geqq 0$ and $\delta_r(S) = E(S)$.

Proof. We prove the result by induction. Obviously $S = N_0 \supseteqq \delta_0(S) = S$. Assume that $N_i \supseteqq \delta_i(S)$. Then $N_i/\varkappa(\varepsilon, N_{i+1})$ is commutative and so $[s_1, s_2] \in N_{i+1}$ for all $s_1, s_2 \in N_i$. Hence by Lemma 3.5 (iv) $[s_1, s_2] \in N_{i+1}$ for all $s_1, s_2 \in \delta_i(S) \subseteqq N_i$. Then $\delta_{i+1}(S) \subseteqq N_{i+1}$. This gives the result by induction.

Corollary 4.8. *A Clifford semigroup S is in $\hat{\mathfrak{S}}$ if and only if there exists d such that $\delta_d(S) = E(S)$.*

The least such d satisfying this is called the *solubility class* of S. It is the least d such that $S \in \hat{\mathfrak{S}}_d$.

Lemma 4.9. *Let S be a Clifford semigroup. Let N be a normal full subsemigroup. Then $S/\varkappa(\varepsilon, N) \in \hat{\mathfrak{S}}_d$ if and only if $\delta_d(S) \subseteqq N$.*

Proof. It is immediate that if θ is a homomorphism, then $[s_1, s_2]\theta = [s_1\theta, s_2\theta]$. Hence $\delta_i(S/\varkappa(\varepsilon, N)) = \delta_i(S)\varkappa(\varepsilon, N)/\varkappa(\varepsilon, N)$ by a simple induction argument. Then $S/\varkappa(\varepsilon, N) \in \hat{\mathfrak{S}}_d$ by Corollary 4.8 if and only if $\delta_d(S/\varkappa(\varepsilon, N)) = E(S/\varkappa(\varepsilon, N))$, i.e. $\delta_d(S)\varkappa(\varepsilon, N) = E(S/\varkappa(\varepsilon, N))$. This is just $\delta_d(S) \subseteqq N$.

Theorem 4.10. *Let S be a Clifford semigroup. Let N be a normal full subsemigroup such that $N \in \hat{\mathfrak{S}}_d$ and $S/\varkappa(\varepsilon, N) \in \hat{\mathfrak{S}}_e$. Then $S \in \hat{\mathfrak{S}}_{d+e}$.*

Proof. By Lemma 4.9, $S/\varkappa(\varepsilon, N)$ is in $\hat{\mathfrak{S}}_e$ implies $\delta_e(S) \subseteqq N$. By a simple induction argument $\delta_i(N) \supseteqq \delta_{e+i}(S)$. But $N \in \hat{\mathfrak{S}}_d$ implies $\delta_d(N) = E(N)$ as N is full. So $\delta_{e+d}(S) = E(S)$ and $S \in \hat{\mathfrak{S}}_{d+e}$.

Theorem 4.11. *Let $N \in \hat{\mathfrak{S}}_c$, $M \in \hat{\mathfrak{S}}_d$ be normal soluble subsemigroups of S, a Clifford semigroup. Then $NM \in \hat{\mathfrak{S}}_{c+d}$.*

Proof. The proof follows closely that of Theorem 3.16. It would be instructive to develop a proof involving a more general version of Theorem 4.10 and paralleling the group theoretic proof.

Theorem 4.12. *Let S be a Clifford semigroup which satisfies the maximal condition on normal subsemigroups. Then S contains a unique maximal normal soluble subsemigroup containing all normal soluble subsemigroups.*

We present the locally soluble version of Theorem 3.17.

Theorem 4.13. *Let S be a locally soluble Clifford semigroup, N a minimal normal subsemigroup of S. Then there exists a unique* $\alpha \in E$ *such that* $N_\alpha \supset \{e_\alpha\}$, *N is commutative and for all* $\beta \leqq \alpha$, *we have* $\ker \varphi_{\alpha,\beta} \supseteqq N_\alpha$.

Proof. A minimal normal subgroup of a locally soluble group is abelian by a standard result from group theory. The same technique as in the proof of Theorem 3.20 now proves the result.

We will leave the extension of results from group theory here. There is obviously an almost inexhaustible supply of results which could be transferred, and there are also some traps for the unwary. Before finishing a few comments might be in order. Finite soluble group theory has a beautiful set of results in the formation theory of GASCHÜTZ. The right extension of this to finite Clifford semigroups should be an interesting exercise with pleasing results. The other point concerns nilpotent versus soluble groups. The laws of $\mathfrak{N}_c$ can be defined without reference to inverses. Using this LALLEMENT [5] showed that regular nilpotent semigroups were Clifford semigroups in $\hat{\mathfrak{N}}_c$. This might be expected because idempotents should be central in a nilpotent semigroup. The same could be done for solubility. There the natural expectation is for idempotents to commute. So it should be a theory naturally based in general inverse semigroups. This is what we hope to attempt soon.

References

[1] P. M. COHN, *Universal Algebra*, Harper and Row (New York 1965).

[2] P. HALL, *The Edmonton Notes on Nilpotent Groups*, London Queen Mary College Mathematics Notes, 1969.

[3] J. M. HOWIE, *An introduction to semigroup theory*, Academic Press (London—New York, 1976).

[4] K. KOWOL and H. MITSCH, Nilpotent inverse semigroups with central idempotents, *Trans. Amer. Math. Soc.*, **271** (1982), 437—444.

[5] G. LALLEMENT, On nilpotency and residual finiteness in semigroups, *Pacific J. Maths.*, **42** (1972), 693—700.

[6] M. PETRICH, *Inverse semigroups*, John Wiley and Sons (New York, 1984).

[7] M. PETRICH, Congruences on inverse semigroups, *J. Algebra*, **55** (1978), 231—256.

[8] D. J. S. Robinson, *Finiteness Conditions and Generalized Soluble Groups*, Vols. 1 and 2, Springer-Verlag (Berlin—Heidelberg—New York, 1972).

[9] J. E. Roseblade, On groups in which every subgroup is subnormal, *J. Algebra*, **2** (1965), 402—412.

[10] W. R. Scott, *Group Theory*, Prentice Hall (Englewood Cliffs, 1964).

DEPARTMENT OF MATHEMATICS
KING'S BUILDINGS
MAYFIELD ROAD
EDINBURGH EH9 3JZ, SCOTLAND

Acta Sci. Math., 52 (1988), 21—28

Semigroups with a universally minimal left ideal

ŠTEFAN SCHWARZ

A left ideal L of a semigroup S is called universally minimal if it is contained in every left ideal of S. In such a semigroup L is at the same time the kernel of S (i.e. the minimal two-sided ideal of S) and L itself is a left simple semigroup. We shall deal with the case that L is a left group.

For simplicity we introduce the following notation. A semigroup containing a universally minimal left ideal which is a left group will be called a *ULG-semigroup*. If L is a group, such semigroups are called *homogroups*. Let S be a semigroup and A an ideal of S. An endomorphism h of S onto A is called an *A-endomorphism* if h leaves the elements of A fixed.

In a forthcomming paper [5] I have been led in a quite natural way to the following class of semigroups: S is a ULG-semigroup with kernel L and S has an L-endomorphism. The main goal of this note is to show that such semigroups have a rather simple structure. Though there are several papers dealing with analogous (and even more general) questions (see, e.g. [1], [2], [3], [4]), I can find nowhere the results given below (at least not in an explicit formulation).

Throughout the paper we use the following notations. S is a ULG-semigroup, L is the kernel of S and $E=\{e_\nu|\nu\in M\}$ is the set of all idempotents of L (i.e. primitive idempotents of S). It is well-known that L can be written in the form $L=\bigcup_{\nu\in M} G_\nu$. Hereby each G_ν is a group (with identity element e_ν) and at the same time a minimal right ideal of S. We have $e_\alpha G_\nu=G_\alpha$, $G_\alpha G_\nu=G_\alpha$ (for any ν, $\alpha\in M$). Moreover each e_α ($\alpha\in M$) is a right identity of L.

In the sequel $|A|$ denotes the cardinality of A.

1. In order to make this note independent of [5] we give in Lemma 1 a modified version of a few results proved in [5].

Received May 18, 1984 and in revised form January 10, 1985.

Lemma 1. *Let S be a ULG-semigroup with kernel L and E the set of all idempotents of L. Then the following holds:*

a) *Any L-endomorphism of S can be written in the form $x \mapsto xe_\alpha$ $(x \in S,\ e_\alpha \in E)$.*

b) *If for some $e_\alpha \in E$ the mapping $x \mapsto xe_\alpha$ is an L-endomorphism of S, then $x \mapsto xe_\nu$ is an L-endomorphism of S for any $e_\nu \in E$.*

c) *The mapping $x \mapsto xe_\alpha$ is an L-endomorphism of S iff for any $x \in S$ we have $|xE|=1$.*

Proof. a) Let h be an L-endomorphism of S and $x \in S$. Since $xe_\alpha \in L$, we have $h(xe_\alpha)=h(x)\cdot h(e_\alpha)=xe_\alpha$, i.e. $h(x)e_\alpha=xe_\alpha$. Since $h(x)\in L$ and e_α is a right identity of L, we have $h(x)=xe_\alpha$.

b) By assumption we have $xe_\alpha ye_\alpha=xye_\alpha$ for any $x, y \in S$. Putting $y=e_\nu$ we have in particular $xe_\alpha e_\nu e_\alpha = xe_\nu e_\alpha$. Since $e_\alpha e_\nu e_\alpha = e_\alpha$ and $e_\nu e_\alpha = e_\nu$, we have $xe_\alpha = xe_\nu$ for any $x \in S$. Hence $xe_\nu ye_\nu = xe_\alpha ye_\alpha = xye_\alpha = xye_\nu$, i.e. $x \mapsto xe_\nu$ is an L-endomorphism of S.

c) If $x \mapsto xe_\alpha$ is an L-endomorphism, we have [by b)] $xe_\alpha = xe_\nu$ for any $\nu \in M$, hence $xe_\alpha = xE$ so that $|xE|=1$. Suppose conversely that $|xE|=1$ for any $x \in S$ and consider the product $xe_\alpha ye_\alpha$ $(x, y \in S,\ e_\alpha \in E)$. The element ye_α is contained in L, hence there is a group $G_\gamma \subset L$ such that $ye_\alpha \in G_\gamma$. Therefore (if e_γ is the identity element of G_γ) $e_\gamma ye_\alpha = ye_\alpha$. By assumption $xe_\alpha = xe_\gamma$, hence $xe_\alpha ye_\alpha = xe_\gamma ye_\alpha = xye_\alpha$. The mapping $x \mapsto xe_\alpha$ is an L-endomorphism. This proves the statement c).

Remark. To understand well the statement a) consider the ULG-semigroup S given by the multiplication table

	a	b	c
a	a	b	a
b	b	a	b
c	a	b	a.

Here $L=\{a, b\}$, $E=\{a\}$, hence S is a homogroup. S has an L-endomorphism φ_1: $x \mapsto xa$. Also φ_2: $x \mapsto xc$ is an endomorphism though here $c \notin E$. But φ_2 is the same endomorphism as φ_1. By c) whenever S has an L-endomorphism we can rewrite it in the form $x \mapsto xE$.

Needless to remark that the mapping $x \mapsto xe_\alpha$ need not be an endomorphism of S. But if it is an endomorphism, it is automatically an L-endomorphism. Hence the result of Lemma 1 can be reformulated as follows

Theorem 1. *Let S be a ULG-semigroup with kernel L. Then S has an L-endomorphism iff for any $x \in S$ we have $|xE|=1$.*

The condition $|xE|=1$ is a very simple one. If S is given by a multiplication table it can be immediately verified. But this condition does not reflect any structural property of S. The structure of such semigroups is given by Theorem 2. (A part of this theorem can be deduced from a result in [1].)

Theorem 2. *Let S be a ULG-semigroup with kernel L. Then S has an L-endomorphism iff S can be written as a union of disjoint right ideals of S each of which is a homogroup. The kernels of these homogroups are then isomorphic to one another.*

Proof. a) Suppose that S has an L-endomorphism. We use the notations introduced above. By Lemma 1 this endomorphism can be written in the form $x \mapsto xE$ $(x \in S)$. For any $\alpha \in M$ denote $R_\alpha = \{x \mid x \in S, xE \in G_\alpha\}$. Clearly $S = \bigcup_{\nu \in M} R_\nu$ and $R_\alpha \cap R_\beta = \emptyset$ if $\alpha \neq \beta$. Further $G_\alpha \subset R_\alpha$ (since $G_\alpha E = G_\alpha$).

We show that $R_\alpha R_\beta \subset R_\alpha$. Let $x \in R_\alpha$, $y \in R_\beta$, i.e., $xE \in G_\alpha$, $yE \in G_\beta$. Then $e_\beta yE = yE$ and $xyE = xe_\beta yE = xE \cdot yE \subset G_\alpha G_\beta = G_\alpha$. Hence $xy \in R_\alpha$, i.e. $R_\alpha R_\beta \subset R_\alpha$. In particular each R_α is a right ideal of S, since $R_\alpha S = R_\alpha \cdot [\bigcup_{\nu \in M} R_\nu] \subset R_\alpha$.

Finally we show that each R_α is a homogroup with kernel G_α. We have $G_\alpha \subset L \cap R_\alpha$, and since $G_\beta \cap R_\alpha = \emptyset$ for $\beta \neq \alpha$, this implies $G_\alpha = L \cap R_\alpha$. The intersection $L \cap R_\alpha$ is a two-sided ideal of R_α. Since it is a group, it is moreover the minimal two-sided ideal of R_α. Hence G_α is the kernel of R_α. This proves the first part of Theorem 2. Moreover it follows from the proof that the kernels of all R_α are isomorphic groups.

b) Suppose conversely that S is a ULG-semigroup with kernel L and S can be written as a union of disjoint right ideals of S in the form $S = \bigcup_{\mu \in N} R'_\mu$. Here we suppose that each R'_μ is a homogroup, hence the kernel of R'_μ is a group K_μ.

Write again $L = \bigcup_{\nu \in M} G_\nu$. Since $R'_\mu L \subset R'_\mu \cap L$, this latter intersection is not empty and it is a right ideal of S contained in L. Hence $L \cap R'_\mu$ is a union of some groups from the family $\{G_\nu\}_{\nu \in M}$. If a group $G_\varkappa$, $\varkappa \in M$, is contained in R'_μ, it is a minimal right ideal of R'_μ. Since a homogroup contains a unique minimal right ideal, we conclude $G_\varkappa = K_\mu$. Hence $L \cap R'_\mu$ contains exactly one group from the family $\{G_\nu\}_{\nu \in M}$ and we have $K_\mu = L \cap R'_\mu$. Otherwise expressed: To any R'_μ there exists an $\alpha \in M$ such that $L \cap R'_\mu = K_\mu = G_\alpha$.

Conversely: Any $e_\beta \in E$ is contained in some R'_μ, hence G_β is contained in R'_μ. Since G_β is a right ideal of S, it is also a right ideal of R'_μ and (since G_β is a group) it is a minimal right ideal of R'_μ. Since R'_μ is a homogroup, G_β is the kernel of R'_μ.

We conclude $|M| = |N|$ and we may write $S = \bigcup_{\nu \in M} R'_\nu$. Also the kernels of all R'_ν are isomorphic groups.

If $x \in S$, then there is a unique R'_ν such that $x \in R'_\nu$. We denote this homogroup R'_ν by $R^{(x)}$. The kernel of $R^{(x)}$ will be denoted by $G^{(x)}$ and the identity element of $G^{(x)}$ by $e^{(x)}$. Note that $R^{(x)} e^{(x)} = e^{(x)} R^{(x)} = G^{(x)}$.

To prove that S has an L-endomorphism it is sufficient, by Theorem 1, to show that $x \cdot e_\alpha = x \cdot e^{(x)}$ for any $x \in S$, $e_\alpha \in E$. Now $x \cdot e_\alpha \in R^{(x)} \cdot L \subset R^{(x)} \cap L = G^{(x)}$. Taking into account that e_α is a right unit in L and $e^{(x)}$ is the unit element of the group $G^{(x)}$ (the kernel of $R^{(x)}$), we have

$$xe_\alpha = e^{(x)} x \cdot e_\alpha = e^{(x)} \cdot x = e^{(x)} \cdot xe^{(x)} = xe^{(x)}. \tag{1}$$

This proves our statement.

Example 1. Suppose that S is a ULG-semigroup with kernel L, S has an L-endomorphism and S is defined by its multiplication table. To find the right ideals R_α mentioned in Theorem 1 we may proceed as follows. We collect all "rows" of the multiplication table containing a fixed chosen $e_\alpha \in E$ (i.e. all sets $\{u, uS\}$ containing e_α). Then $R_\alpha = \bigcup_u \{u, uS\}$. Clearly R_α is a right ideal of S, it contains e_α, and it follows from the proof that it cannot contain any other idempotent of L.

Consider, e.g., the semigroup S given by the following multiplication table:

	a	b	c	d	f
a	a	a	c	c	a
b	b	b	d	d	b
c	c	c	a	a	c
d	d	d	b	b	d
f	a	a	c	c	a

Here $L = E = \{a, b\}$. S has an L-endomorphism since $|x \cdot \{a, b\}| = 1$ for any $x \in S$. The idempotent a is contained in $\{a, aS\}, \{c, cS\}, \{f, fS\}$. Hence $R^{(a)} = \{a, c, f\}$. Analogously $R^{(b)} = \{b, bS\} \cup \{d, dS\} = \{b, d\}$. Finally $S = R^{(a)} \cup R^{(b)}$.

We shall return to this procedure in Section 3.

2. In Theorem 2 the right ideals R_ν have the property that their kernels are isomorphic groups. The question arises whether there are some other limitations concerning the ideals R_ν. The answer is no. To any family of homogroups $\{Q_\nu\}$ with isomorphic kernels we can construct at least one ULG-semigroup which has an L-endomorphism. We give a special construction and we do not attempt to find all such semigroups.

More precisely we have:

Theorem 3. *Let L_0 be a left group. Write $L_0 = G_0 \times E_0$, where G_0 is a group and E_0 a left zero semigroup. Let $\{Q_\nu \mid \nu \in M\}$ be a family of disjoint homogroups whereby each Q_ν has a kernel isomorphic to G_0 and $|E_0| = |M|$. Then there exists a ULG-semigroup S having the following properties:*

1) $S = \bigcup_{\nu \in M} Q_\nu$.

2) *Each Q_ν is a right ideal of S.*

3) *The kernel L of S is isomorphic to L_0 and S has an L-endomorphism.*

Proof. Denote the kernel of Q_ν by H_ν and denote the identity element of H_ν by e_ν. Suppose that $1\in M$. For every $\nu\in M$ let φ_ν be a fixed chosen isomorphism of H_1 onto H_ν. Define the mapping $\varphi_{\alpha\beta}\colon H_\alpha\to H_\beta$ by $\varphi_{\alpha\beta}=\varphi_\alpha^{-1}\varphi_\beta$. Then $\varphi_{\alpha\beta}$ is an isomorphism and $\varphi_{\alpha\alpha}$ is the identity mapping of H_α onto H_α. For any $a\in H_\alpha$ we have

$$(a\varphi_{\alpha\beta})\varphi_{\beta\gamma} = (a\varphi_\alpha^{-1}\varphi_\beta)\varphi_\beta^{-1}\varphi_\gamma = a\varphi_\alpha^{-1}\varphi_\gamma = a\varphi_{\alpha\gamma}.$$

In this way we get a set of mappings $\{\varphi_{\mu\nu}\}$ where $\varphi_{\alpha\beta}\varphi_{\beta\gamma}=\varphi_{\alpha\gamma}$ for any $\alpha, \beta, \gamma\in M$.

Note finally: Since $\varphi_{\alpha\beta}$ is an isomorphism, we have $(e_\alpha)\varphi_{\alpha\beta}=e_\beta$.

1) We now use the set of these mappings to define on $S=\bigcup_{\nu\in M} Q_\nu$ a multiplication (denoted by $*$). For $\alpha\neq\beta$ and $x\in Q_\alpha$, $y\in Q_\beta$, we define

$$x*y = (e_\alpha x)\cdot(e_\beta y)\varphi_{\beta\alpha},$$

while inside of each Q_α the multiplication remains unaltered.

The definition implies $x*y\in H_\alpha\cdot H_\alpha=H_\alpha$, hence for $\alpha\neq\beta$, $Q_\alpha*Q_\beta\subset H_\alpha$. Since $H_\alpha\subset Q_\alpha$, $(H_\beta)\varphi_{\beta\alpha}=H_\alpha$, we have $H_\alpha*H_\beta=H_\alpha$ and therefore for $\alpha\neq\beta$,

$$Q_\alpha*Q_\beta = Q_\alpha*H_\beta = H_\alpha*H_\beta = H_\alpha*Q_\beta = H_\alpha. \tag{2}$$

In order to show that S is a semigroup we have to check associativity.

a) Suppose first $\alpha\neq\beta$, $\beta\neq\gamma$ and $x\in Q_\alpha$, $y\in Q_\beta$, $z\in Q_\gamma$.

In the following we use: $x*y\in H_\alpha$ implies $e_\alpha(x*y)=x*y$ and $u*v\in H_\beta$ implies $e_\beta(u*v)=u*v$. We have:

$$x*(y*z) = x*[e_\beta y\cdot(e_\gamma z)\varphi_{\gamma\beta}] = e_\alpha x\cdot[e_\beta y\cdot(e_\gamma z)\varphi_{\gamma\beta}]\varphi_{\beta\alpha} =$$

$$= e_\alpha x\cdot(e_\beta y)\varphi_{\beta\alpha}\cdot(e_\gamma z)\varphi_{\gamma\alpha} = (x*y)\cdot(e_\gamma z)\varphi_{\gamma\alpha} = e_\alpha(x*y)\cdot(e_\gamma z)\varphi_{\gamma\alpha} = (x*y)*z.$$

b) Suppose next. $\alpha\neq\beta, \beta=\gamma$, and $x\in Q_\alpha$, $y\in Q_\beta$, $z\in Q_\beta$.

In the following we use $e_\beta y\in H_\beta$, hence $e_\beta y=e_\beta ye_\beta$. We have:

$$x*(y*z) = x*(yz) = e_\alpha x\cdot(e_\beta yz)\varphi_{\beta\alpha} = (e_\alpha x)(e_\beta ye_\beta z)\varphi_{\beta\alpha} =$$

$$= (e_\alpha x)\cdot(e_\beta y)\varphi_{\beta\alpha}\cdot(e_\beta z)\varphi_{\beta\alpha} = (x*y)\cdot(e_\beta z)\varphi_{\beta\alpha} = e_\alpha(x*y)\cdot(e_\beta z)\varphi_{\beta\alpha} = (x*y)*z.$$

c) Suppose finally $\alpha=\beta$, $\beta\neq\gamma$, and $x\in Q_\alpha$, $y\in Q_\alpha$, $z\in Q_\gamma$.

$$x*(y*z) = x*\left[(e_\alpha y)\cdot(e_\gamma z)\varphi_{\gamma\alpha}\right] = x\cdot(e_\alpha y)\cdot(e_\gamma z)\varphi_{\gamma\alpha}.$$

Now since $e_\alpha x\in H_\alpha$ we have $e_\alpha x=e_\alpha xe_\alpha$ and $e_\alpha xy=e_\alpha xe_\alpha y$. Also since $xe_\alpha y\in H_\alpha$ we have $e_\alpha xe_\alpha y=xe_\alpha y$. Hence $e_\alpha xy=xe_\alpha y$. We may write therefore:

$$x*(y*z) = e_\alpha xy\cdot(e_\gamma z)\varphi_{\gamma\alpha} = (xy)*z = (x*y)*z.$$

This proves that S is a semigroup.

2) The relation (2) implies $Q_\alpha * Q_\beta = H_\alpha \subset Q_\alpha$ for $\alpha \neq \beta$ and $Q_\alpha^2 \subset Q_\alpha$ (for any $\alpha \in M$). Next

$$Q_\alpha * S = Q_\alpha * [\bigcup_{\nu \in M} Q_\nu] \subset Q_\alpha,$$

so that each Q_α is a right ideal of S. Denote $L = \bigcup_{\mu \in M} H_\mu$, then by (2)

$$S * L = [\bigcup_{\nu \in M} Q_\nu] * [\bigcup_{\mu \in M} H_\mu] = \bigcup_{\nu \in M} H_\nu = L,$$

$$L * S = [\bigcup_{\mu \in M} H_\mu] * [\bigcup_{\nu \in M} Q_\nu] = \bigcup_{\mu \in M} H_\mu = L.$$

Hence L is a two-sided ideal of S.

To prove that L is a left group it is sufficient to show that for any $y \in S$ we have $L * y = L$. Now $y \in S$ implies $y \in Q_\beta$ for some $\beta \in M$. Denote $(e_\beta y)\varphi_{\beta\nu} = y_\nu \in H_\nu$. We have

$$L * y = [\bigcup_{\nu \in M} H_\nu] * y = \bigcup_{\nu \in M} [H_\nu * y] = \bigcup_{\nu \in M} [H_\nu \cdot (e_\beta y)\varphi_{\beta\nu}] =$$

$$= \bigcup_{\nu \in M} [H_\nu \cdot y_\nu] = \bigcup_{\nu \in M} H_\nu = L.$$

This proves that S is a ULG-semigroup with kernel L and clearly L is isomorphic to L_0.

3) It remains to show that S has an L-endomorphism. Denote by E the set of all idempotents contained in L. It is sufficient to show that for $x \in S$ we have $|x * E| = 1$. If $x \in S$ we have $x \in Q_\alpha$ for some $\alpha \in M$. Let $e_\gamma \in E$. Then

$$x * e_\gamma = (e_\alpha x) \cdot (e_\gamma)\varphi_{\gamma\alpha} = e_\alpha x e_\alpha.$$

The right-hand side is independent of e_γ, hence $|x * E| = 1$. This proves Theorem 3.

3. The procedure described in Example 1 can be carried out in any ULG-semigroup (even if S has not an L-endomorphism). To any minimal right ideal G_ν of a ULG-semigroup S there is a largest right ideal R_ν^* of S (containing G_ν) such that R_ν^* is a homogroup. This right ideal consists of all "rows" $\{u, uS\}$ containing e_ν but no other idempotent of E. If $e_\alpha \neq e_\beta$, then $R_\alpha^* \cap R_\beta^* = \emptyset$. The union $S^* = \bigcup_{\nu \in M} R_\nu^*$ is a right ideal of S. If S does not have an L-endomorphism, then S^* is a proper subset of S.

Lemma 2. *The set S^* consists exactly of those elements $x \in S$ for which $|xE| = 1$.*

Proof. a) Let $x \in S^*$, hence $x \in R_\alpha^*$ with suitably chosen $\alpha \in M$. We have $xE \subset R_\alpha^* L \subset R_\alpha^* \cap L = G_\alpha$. Note that in the homogroup R_α^* we have $xe_\alpha = e_\alpha x$ (for any $x \in R_\alpha^*$).

Let now e_γ be any element of E. Then $xe_\gamma \in G_\alpha$ implies $(xe_\gamma)e_\alpha = e_\alpha(xe_\gamma)$. This implies $xe_\gamma = e_\alpha(xe_\gamma) = (e_\alpha x)e_\gamma = (xe_\alpha)e_\gamma = x(e_\alpha e_\gamma) = xe_\alpha$. Hence $xe_\gamma = xe_\alpha$; therefore $xE = xe_\alpha$, i.e., $|xE| = 1$ for any $x \in S^*$.

b) Suppose conversely that $x \in S - S^*$. We have to show that $|xE| \geqq 2$. The right ideal $\{x, xS\}$ contains at least two idempotents of E, say e_α, e_β $(e_\alpha \neq e_\beta)$. (Note that any right ideal of a ULG-semigroup contains at least one minimal right ideal hence some of the groups $\{G_\gamma\}$.) Write $\{x, xS\} = \{e_\alpha, e_\beta, S_1\}$, where S_1 is a subset of S. (We do not exclude that S_1 contains some further elements of E.) Multiplying by E we have

$$\{xE, xSE\} = \{e_\alpha E, e_\beta E, S_1 E\}.$$

Since $SE = L, e_\alpha E = e_\alpha$, $e_\beta E = e_\beta$, we have

$$\{xE, xL\} = \{e_\alpha, e_\beta, L_1\},$$

where L_1 is a subset of L. Finally since $xE \subset xL$ we get

$$xL = \{e_\alpha, e_\beta, L_1\}.$$

Hence there are two elements $g \in L, g_1 \in L$, such that

$$xg = e_\alpha, \tag{3}$$

$$xg_1 = e_\beta. \tag{4}$$

Since $L = \bigcup_{\nu \in M} G_\nu$, there are two indices $\gamma, \delta \in M$ such that $g \in G_\gamma$, $g_1 \in G_\delta$. Denote by g^{-1} the element of G_γ for which $gg^{-1} = e_\gamma$ and by g_1^{-1} the element of G_δ for which $g_1 g_1^{-1} = e_\delta$. Then (3) and (4) imply

$$xgg^{-1} = e_\alpha g^{-1}, \quad xg_1 g_1^{-1} = e_\beta g_1^{-1},$$

hence

$$xe_\gamma = e_\alpha g^{-1} \in e_\alpha L = G_\alpha, \quad xe_\delta = e_\beta g_1^{-1} \in e_\beta L = G_\beta.$$

Since $G_\alpha \cap G_\beta = \emptyset$, the elements xe_γ, xe_δ are different elements (contained in L). Hence xE contains at least two different elements (namely xe_γ, xe_δ) so that $|xE| \geqq 2$. This proves Lemma 2.

The semigroup S^* (being a union of right ideals of S) is a right ideal of S. But we easily show that S^* is also a left ideal of S (hence a two-sided ideal of S). Suppose that $x \in S^*$, i.e. $|xE| = 1$. Then for any $s \in S$ $(sx)E = s(xE)$ and since xE is a unique element (contained in L), we conclude $|(sx)E| = 1$, i.e. $sx \in S^*$, hence $SS^* \subset S^*$.

We have proved:

Theorem 4. *Let S be a ULG-semigroup with kernel L. Denote by E the set of all idempotents of L. Then there exists a unique largest subsemigroup S^* of S containing*

L such that S^* has an L-endomorphism. The semigroup S^* is a two-sided ideal of S and it can be characterised by the following two equivalent conditions:

a) *S^* is the set of all $x \in S$ such that $|xE|=1$.*

b) *S^* is the union of (disjoint) largest right ideals of S each of which is a homogroup.*

Remark. The emphasis in the second characterization is on the fact that the right ideals in $S^* = \bigcup_{\alpha \in M} R^*_\alpha$ are right ideals of S (and not merely of S^*).

Example 2. Consider the ULG-semigroup S given by the multiplication table

	a	b	c	d
a	a	a	a	a
b	b	b	b	b
c	a	b	c	d
d	a	a	a	d.

Here $L=E=\{a, b\}$. The semigroup S has no L-endomorphism. The largest right ideal R^*_a containing the idempotent a which is a homogroup is $R^*_a=\{a, d\}$. Next R^*_b is $\{b\}$ itself. We have $S^*=\{a, d\} \cup \{b\}$. The element c cannot be contained in a right ideal which is a homogroup, since $\{c, cS\}$ contains both idempotents a and b.

It is worth noting that $R^*_a=\{a, d\}$ is a homogroup, but not the largest homogroup containing a. The largest homogroup containing a is the subsemigroup $\{a, d, c\}$. (Of course this semigroup is not a right ideal of S.)

References

[1] G. Lallement, Homomorphismes d'un demi-groupe sur un demi-groupe complètement 0-simple, *Seminaire Dubreil—Pisot*, 17e année 1963—64, No 14.

[2] G. Lallement, M. Petrich, A generalization of the Rees theorem in semigroups, *Acta Sci. Math.*, 30 (1969), 113—132.

[3] M. Petrich, *Lectures in semigroups*, Akademie Verlag (Berlin, 1977).

[4] B. M. Schein, Bands of monoids, *Acta Sci. Math.*, 36 (1974), 145—154.

[5] Š. Schwarz, Right composition of semigroups, *Math. Slovaca*, 36 (1986), 3—14.

PORUBSKÉHO 8
81106 BRATISLAVA, CZECHOSLOVAKIA

Acta Sci. Math., 52 (1988), 29—33

Principal tolerance trivial commutative semigroups

BEDŘICH PONDĚLÍČEK

Following I. CHAJDA [1] an algebra A is said to be (principal) tolerance trivial if every (principal) tolerance on A is a congruence. In [2] B. ZELINKA has shown that a commutative semigroup S is tolerance trivial if and only if either S is a group or card $S=2$.

In this paper we shall describe all commutative semigroups which are principal tolerance trivial. Non-defined terminology and notation may be found in [3] and [4].

Recall that a tolerance T on a commutative semigroup S is a reflexive and symmetric subsemigroup of the direct product $S\times S$. For $a, b\in S$ we denote by $T(a, b)$ the least tolerance on S containing (a, b), i.e. $T(a, b)$ is the principal tolerance on S generated by (a, b). We shall use the following notation: $(a, b)^m z=(a^m z, b^m z)$ for all $a, b, z\in S$ and for every positive integer m. The set of all idempotents of a commutative semigroup S is denoted by $E(S)$ and is partially ordered by: $e\leqq f$ if and only if $ef=e$. We write $e<f$ for $e\leqq f$ and $e\neq f$. We denote by G_e the maximal subgroup of S containing an idempotent e. The notation S^1 stands for S if S has an identity, otherwise it stands for S with an identity adjoined.

The following lemma is clear:

Lemma 1. *Let S be a commutative semigroup and $a, b\in S$, $a\neq b$. For $x, y\in S$, $x\neq y$, we have $(x, y)\in T(a, b)$ if and only if there exist $z\in S^1$ and a positive integer m such that either $(x, y)=(a, b)^m z$ or $(x, y)=(b, a)^m z$.*

Note 1. Let S be a zero semigroup, i.e. card $S^2=1$. Using Lemma 1 it is easy to show that S is principal tolerance trivial.

Note 2. Now, we give another example of a principal tolerance trivial commutative semigroup. Let G be a commutative periodic group and let A be a non-empty set. Suppose that $G\cap A=\emptyset$ and put $S=G\cup A$. Let a multiplication on S be defined as follows:

Received October 8, 1984.

a) If $e, f \in A$, then $ef=e$ for $e=f$ and $ef=h$ for $e \neq f$, where h denotes the identity of G.

b) If $e \in A$ and $g \in G$, then $eg=g=ge$.

c) If $g_1, g_2 \in G$, then the product $g_1 g_2$ is the same as in G.

It is easy to show that S is a commutative semigroup which is a semilattice of groups. Clearly $E(S)=A \cup \{h\}$, $G_e=\{e\}$ for all $e \in A$ and $G_h=G$.

Now, we shall prove that S is a principal tolerance trivial semigroup. Let $a, b \in S$, $a \neq b$. It suffices to show that the relation $T(a, b)$ is transitive.

Case 1. Suppose that $a, b \in A$. It follows from Lemma 1 that $T(a, b)= =R \cup R^{-1} \cup \mathrm{id}_S$, where $R=\{(a, b), (b, h), (h, a)\}$. Clearly $T(a, b)$ is transitive.

Case 2. Suppose that $a \in A$ and $b \in G$. Evidently $T(a, b)=T(b, a)$. Let (x, y), $(y, z) \in T(a, b)$ and $x \neq y$, $y \neq z$. It follows from Lemma 1 that $(x, y)=(a, b)^m u$ or $(x, y)=(b, a)^m u$ for some $u \in S^1$ and some positive integer m. Analogously we have $(y, z)=(a, b)^n v$ or $(y, z)=(b, a)^n v$ for some $v \in S^1$ and some positive integer n.

Subcase 2a. Assume that $x=au$, $y=b^m u=av$ and $z=b^n v$. Then $z=b^n v= =b^n av=b^{m+n} u$ and so, by Lemma 1, we have $(x, z)=(a, b)^{m+n} u \in T(a, b)$.

Subcase 2b. Assume that $x=au$, $y=b^m u=b^n v$ and $z=av$. If $u=v$, then $x=z$ and so $(x, z) \in T(a, b)$. We can suppose that $u \neq v$. If $u, v \in G$, then $b^m u=b^n v$ and so $uv^{-1}=b^{n-m}=b^r$ for some positive integer r, because the group G is periodic. By Lemma 1, we have $(x, z)=(u, v)=(b, a)^r v \in T(a, b)$. If $u \in G$ and $v \in S^1 \setminus G$, then $b^m u=b^n$ and so $u=b^{n-m}=b^r$ for some positive integer r. Hence we have $(x, z)= =(u, a)=(b, a)^r$ for $v \in \{1, a\}$ and $(x, z)=(u, h)=(b, a)^r v$ for $v \notin \{1, a\}$. This gives in both cases $(x, z) \in T(a, b)$. Analogously we can prove that $u \in S^1 \setminus G$ and $v \in G$ imply $(x, z) \in T(a, b)$. Let $u, v \in S^1 \setminus G$. Then it is easy to show that $(x, z) \in \{a, h\} \times \times \{a, h\}$. Since G is periodic, there exists a positive integer k such that $b^k=h$ and so $(a, h)=(a, b)^k$. Therefore we have $(x, z) \in T(a, b)$.

Subcase 2c. Assume that $x=b^m u$, $y=au=av$ and $z=b^n v$. Since b is a periodic element of G, there exists a positive integer r such that $b^{n-m}=b^r$. Thus we have $(x, z)=(b^m u, b^n v)=(a, b)^r b^m au \in T(a, b)$.

Subcase 2d. Assume that $x=b^m u$, $y=au=b^n v$ and $z=av$. Using the same method as in Subcase 2a we obtain that $(x, z) \in T(a, b)$.

Case 3. Suppose that $a, b \in G$. Let (x, y), $(y, z) \in T(a, b)$ and $x \neq y$, $y \neq z$. It follows from Lemma 1 that $x, y, z \in G$ and so $(x, z)=(x, y)(y^{-1}, y^{-1})(y, z) \in \in T(a, b)$.

Theorem. *A commutative semigroup S is principal tolerance trivial if and only if S satisfies one of the following conditions:*

(i) *S is group;*

(ii) *S is a zero semigroup;*

(iii) *S is of type defined in Note 2.*

Proof. Let S be a commutative semigroup. If S satisfies one of the conditions (i), (ii) or (iii), then S is principal tolerance trivial (see Notes 1 and 2).

Now, we shall prove the following lemmas, in which we shall suppose that the commutative semigroup S is principal tolerance trivial, card $S^2 \geqq 2$ and S is not a group.

Lemma 2. *If* $a \in S \setminus a^2S$, *then* a^2 *is a zero in* S.

Proof. Let $a \in S \setminus a^2S$. Then $a \neq a^2$ and, by Lemma 1, we obtain (a, a^2), $(a^2, a^3) \in T(a, a^2)$. Since $T(a, a^2)$ is transitive, we have $(a, a^3) \in T(a, a^2)$. According to Lemma 1, there exists a $u \in S^1$ such that $(a, a^3) = (a, a^2)u$ and so $a^3 = a^2u = a^2$. Put $h = a^2$. Clearly $h^2 = h = ah$. Now, we shall show that $hx = h$ for all $x \in S$. Assume that $hb \neq h$ for some $b \in S$. If $hb = a$, then $a \in a^2S$, which is a contradiction. We have $hb \neq a$. It is clear that $(hb, h) = (hb, a)a$. According to Lemma 1, we have (a, hb), $(hb, h) \in T(a, hb)$ and so $(a, h) \in T(a, hb) = T(a, a^2b)$. It follows from Lemma 1 that $(a, h) = (a, hb)u$ for some $u \in S^1$. Hence we have $h = ah = ahbu = ahb = hb$, a contradiction. Therefore h is a zero in S.

Lemma 3. *Let* S *have a zero* 0 *and let* $a, b \in S$. *If* $a^2 = 0 = b^2$ *and* $a \neq 0 \neq b$, *then* $ab = 0$.

Proof. Assume that $ab \neq 0$. If $a = ab$, then $a = ab^2 = 0$, a contradiction. We have $a \neq ab$. By Lemma 1, we obtain (a, ab), $(ab, 0) \in T(a, ab)$, because $(ab, 0) = (a, ab)b$. Hence we have $(a, 0) \in T(a, ab)$. If $a = abu$ for some $u \in S^1$, then $ab = 0$, a contradiction. Lemma 1 implies that $(a, 0) = (a, ab)u$ for some $u \in S^1$. Then $ab = aub = 0$, a contradiction.

Lemma 4. *Let* S *have a zero* 0 *and let* a, $e \in S$. *If* $a^2 = 0$, $e^2 = e$ *and* $a \neq 0 \neq e$, *then* $ae = 0$.

Proof. Assume that $ae \neq 0$. We have $(e, 0), (0, ae) \in T(e, 0)$ and so $(e, ae) \in T(e, 0)$. If $e = ae$, then $e = a^2e = 0$, a contradiction. Hence we have $e \neq ae$ and so, by Lemma 1, $e = 0$ or $ae = 0$, which is a contradiction.

Lemma 5. *S is regular.*

Proof. Suppose that S is not regular. From Lemma 2 it follows that S has a zero 0. Since card $S^2 \geqq 2$ by hypothesis, therefore there exist $a, b \in S$ such that $ab \neq 0$. According to Lemmas 2 and 3, a or b is a regular element of S. This implies that there exists an idempotent $e \neq 0$ in S. Evidently, S has an element $c \neq 0$, which is not regular. It follows from Lemma 2 that $c^2 = 0$ and Lemma 4 implies that $ce = 0$. Clearly $c \neq e$, and according to Lemma 1, we have (c, e), $(e, 0) \in T(e, c)$, because $(e, 0) = (e, c)e$. Thus $(c, 0) \in T(e, c)$. If $c = eu$ for some $u \in S^1$, then $0 = ce = c$, a contradiction. Hence, by Lemma 1, we obtain $(c, 0) = (c, e)u$ for some $u \in S^1$. Then $c=cu=cu^2$ and so $u^2 \neq 0$. Lemma 2 implies that u is regular, which means that $u=u^2v$ for some $v \in S$. Hence we obtain $uv \neq 0$ and $(uv)^2=uv$. According to Lemma 4, we have $cuv=0$ and so $c=cu=(cuv)u=0$, a contradiction.

Lemma 6. *If* $e \leqq f < g$ *for* $e, f, g \in E(S)$, *then* $e=f$.

Proof. Assume that $e<f$. Then $e<g$ and $(f, e)=(g, e)f$. It follows from Lemma 1 that $(f, e), (e, g) \in T(e, g)$ and so $(f, g) \in T(e, g)$. By Lemma 1, we have either $f=ez$ or $g=ez$ for some $z \in S^1$. If $f=ez$, then $e=ef=f$, a contradiction. If $g=ez$, then analogously $e=eg=g$, a contradiction.

Lemma 7. *$E(S)$ is of the type defined in Note 2.*

Proof. It follows from Lemma 5 that $E(S) \neq \emptyset$. If $\operatorname{card} E(S)=1$, then S is a group, which is a contradiction. Hence we have $\operatorname{card} E(S) \geqq 2$. Our statement follows from Lemma 6.

Lemma 8. *S is periodic.*

Proof. It follows from Lemma 5 that S is a semilattice of maximal subgroups G_e $(e \in E(S))$. Suppose that there exists a $c \in S$ which is not periodic. Then $c \in G_e$ for some $e \in E(S)$. Clearly $c \neq e$. It follows from Lemma 7 that there exists an $f \in E(S)$ such that either $f<e$ or $e<f$.

Case 1. $f<e$. According to Lemma 1, we have (c, f), $(f, c^2) \in T(f, c)$ and so $(c, c^2) \in T(f, c)$. It follows from Lemma 1 that either $c=fu$ or $c^2=fu$ for some $u \in S^1$. Then either $e=fuc^{-1}$ or $e=fu(c^{-1})^2$ (c^{-1} denotes the inverse element of c in G_e). This gives in both cases $e=ef=f$, a contradiction.

Case 2. $e<f$. Then we have $(c, e)=(c, f)e$ and so, by Lemma 1, we obtain $(f, c), (c, e) \in T(f, c)$. By hypothesis we have $(f, e) \in T(f, c)$. Lemma 1 implies that either $(f, e)=(f, c)^m u$ or $(f, e)=(c, f)^m u$ for some $u \in S^1$ and some positive integer m. If $f=fu$ and $e=c^m u$, then $e=ef=c^m uf=c^m f=(c^m e)f=c^m$ and so c is periodic, a contradiction. If $f=c^m u$, then $e=ef=ec^m u=c^m u=f$, a contradiction.

Lemma 9. *If* $h<e$, $e, h \in E(S)$, *then* $\operatorname{card} G_e=1$.

Proof. Assume that there exists a $c \in G_e$ such that $c \neq e$. It follows from Lemma 8 that $c^k=e$ for some positive integer k. By Lemma 1, we have $(c, h), (h, e) \in T(h, c)$ and so $(c, e) \in T(h, c)$. It follows from Lemma 1 that either $c=hu$ or $e=hu$ for some $u \in S^1$. If $c=hu$, then $e=c^k=hu^k$ and so $h=he=e$, a contradiction. If $e=hu$, then analogously we have $h=e$, a contradiction.

The proof of Theorem follows from Lemmas 5, 6, 7, 8 and 9.

Corollary 1. *A semilattice is principal tolerance trivial if and only if its length is not greater than two.*

It is known (see [5] and [6]) that the set $\mathscr{L}(S)$ of all tolerances on a semigroup S forms a complete algebraic lattice with respect to set inclusion.

Corollary 2. *Let S be a tolerance trivial commutative semigroup. Then the lattice $\mathscr{L}(S)$ is modular.*

Proof. If S is a commutative group, then $\mathscr{L}(S)$ is the lattice of all congruences on S and so $\mathscr{L}(S)$ is modular. If S is a zero semigroup, then $\mathscr{L}(S)$ is the lattice of all reflexive and symmetric relations on S and so $\mathscr{L}(S)$ is distributive. If S is of the type defined in Note 2, then it follows from Theorem 1 of [7] that $\mathscr{L}(S)$ is modular.

References

[1] I. Chajda, Tolerance trivial algebras and varieties, *Acta Sci. Math.*, **46** (1983), 35—40.
[2] B. Zelinka, Tolerance in algebraic structures. II, *Czechoslovak Math. J.*, **25** (1975), 175—178.
[3] A. H. Clifford and G. B. Preston, *The algebraic theory of semigroups*, Vol. I, Amer. Math. Soc. (Providence, R. I., 1961).
[4] M. Petrich, *Introduction to semigroups*, Merrill Publishing Company (Columbus, 1973).
[5] I. Chajda, Lattices of compatible relations, *Arch. Math. (Brno)*, **13** (1977), 89—96.
[6] I. Chajda and B. Zelinka, Lattices of tolerances, *Časopis Pěst. Mat.*, **102** (1977), 10—24.
[7] B. Pondělíček, Modularity and distributivity of tolerance lattices of commutative separative semigroups, *Czechoslovak Math. J.*, **35** (1985), 333—337.

FEL ČVUT
SUCHBÁTAROVA 2
166 27 PRAHA, CZECHOSLOVAKIA

Acta Sci. Math., 52 (1988), 35—45

On non-modular n-distributive lattices
I. Lattices of convex sets

A. P. HUHN

1. Introduction. A lattice is called n-distributive if it satisfies the identity

$$x\wedge\bigvee_{i=0}^{n}y_i=\bigvee_{j=0}^{n}\Big[x\wedge\bigvee_{\substack{i=0\\ i\neq j}}^{n}y_i\Big]. \tag{1}$$

A lattice satisfying the dual of (1) is called dually n-distributive. The class of n-distributive (respectively, dually n-distributive) lattices is denoted by Δ_n (respectively, ∇_n). n-distributive lattices were introduced to describe dimension like properties of modular lattices. Here we present some examples of non-modular n-distributive lattices. E^{n-1} denotes the $(n-1)$-dimensional Euclidean space and $\mathfrak{L}(E^{n-1})$ denotes its lattice of convex sets. Our first result describes how $\mathfrak{L}(E^{n-1})$ is situated in the classes Δ_m and ∇_m.

Theorem 1.1. $\mathfrak{L}(E^{n-1})\in(\Delta_n\setminus\Delta_{n-1})\cap(\nabla_n\setminus\nabla_{n-1})$.

The proof of n-distributivity in Section 2 is based on Carathéodory's theorem, while the dual n-distributivity is derived from Helly's theorem.

In Section 3 we strengthen part of this result. Let F denote the class of finite lattices.

Theorem 1.2. $\mathfrak{L}(E^{n-1})\in \mathrm{HSP}(\Delta_n\cap F)$.

In other words, $\mathfrak{L}(E^{n-1})$ is in the lattice variety (equational class) generated by the finite n-distributive lattices. The intuitive reason for Theorem 1.2 is that, if we restrict the operation of convex closure to a finite subset H of E^{n-1}, then this closure system has an n-distributive lattice of closed sets by Carathéodory's theorem, and this lattice resembles $\mathfrak{L}(E^{n-1})$ as H becomes large. We note that $\mathfrak{L}(E^{n-1})$ is also in the

Received June 5, 1985.

class $\mathrm{HSP}(\nabla_n \cap F)$. The proof of this theorem involves more geometry and will be published separately together with other Helly-type results.

Notice that the above sketch of the proof of Theorem 1.2 gives rise to a high variety of n-distributive lattices: associated with any finite subset of E^{n-1} there is an n-distributive lattice. The example given by the following theorem is of different character. Let $\bar{\mathfrak{L}}(E^{n-1})$ denote the lattice of closed convex sets of E^{n-1}. In Section 4 we prove:

Theorem 1.3. $\bar{\mathfrak{L}}(E^{n-1}) \in (\Delta_n \setminus \Delta_{n-1}) \cap (\nabla_n \setminus \nabla_{n-1})$.

Carathéodory's theorem provides also a new aspect to the study of modular n-distributive lattices. In Section 5 we characterize complete, complemented, modular, completely n-distributive lattices among all projective geometries as those satisfying a Carathéodory type condition. (Completely n-distributive lattices are defined in Section 5 in analogy with completely distributive lattices.) An unexpected consequence of our characterization is that this class of lattices (as well as the corresponding class of projective geometries) is self-dual.

Finally, in Section 6 we prove the following fact on modular n-distributive lattices:

Theorem 1.4. *Every modular n-distributive lattice is a member of* $\mathrm{HSP}(\Delta_n \cap F)$.

It is now natural to ask whether there are any further examples of non-modular n-distributive lattices in other branches of mathematics. It is not hard to show that the partition lattice of an $(n+1)$-element set is in $(\Delta_n \setminus \Delta_{n-1}) \cap (\nabla_n \setminus \nabla_{n-1})$. This example will be developed further in Part II of this paper, where graphs with an n-distributive (respectively, dually n-distributive) contraction lattice are characterized. Partition lattices occur as special cases, as they are the contraction lattices of complete graphs.

In an independent paper [3] Horst Gerstmann also considers nonmodular n-distributive lattices, defines complete and infinite n-distributive laws and characterizes the different sorts of n-distributivity of the closed sets of a closure space in terms of properties of the closure operator. Gerstmann's generalized distributive laws cover, beside the n-distributive laws, the concepts of (von Neumann) $\wedge$-continuity and of Scott-continuity.

2. The lattice of convex sets. We first quote the two classical theorems that are in the centre of this paper.

Helly's theorem. *Let $\mathscr{C}$ be a finite family of convex subsets of E^{n-1}. If any n elements of $\mathscr{C}$ have a non-empty intersection, then the intersection of the whole family $\mathscr{C}$ is not empty.*

Carathéodory's theorem. *Let H be a subset of E^{n-1} and let p be a point in E^{n-1}. If p is in the convex closure of H, then it is in the convex closure of an n element subset of H.*

We first prove that $\mathfrak{L}(E^{n-1})$ is n-distributive. Let $X, Y_0, Y_1, \ldots, Y_n \in \mathfrak{L}(E^{n-1})$. Let p be a point of E^{n-1} and assume that

$$p \in X \wedge \bigvee_{i=0}^{n} Y_i$$

(where the $\wedge$ and $\vee$ are the operations of $\mathfrak{L}(E^{n-1})$). Then, by Carathéodory's theorem there are n elements of the set union $\bigcup_{i=0}^{n} Y_i$, say $p_0, p_1, \ldots, p_{n-1}$, such that p is an element of their convex closure. If $p_j \in Y_{i_j}$, $j=0, 1, \ldots, n-1$, then p is also in $\bigvee_{j=0}^{n-1} Y_{i_j}$. Of course, $p \in X$, hence

$$p \in \bigvee_{j=0}^{n} [X \wedge \bigvee_{\substack{i=0 \\ i \neq j}}^{n} Y_i],$$

that is,

$$X \wedge \bigvee_{i=0}^{n} Y_i \subseteq \bigvee_{j=0}^{n} [X \wedge \bigvee_{\substack{i=0 \\ i \neq j}}^{n} Y_i].$$

The reverse inclusion is obvious.

Now we prove that the dual n-distributive law holds in $\mathfrak{L}(E^{n-1})$. Let $X, Y_0, Y_1, \ldots, Y_n \in \mathfrak{L}(E^{n-1})$. Let

$$p \in \bigwedge_{j=0}^{n} [X \vee \bigwedge_{\substack{i=0 \\ i \neq j}}^{n} Y_i].$$

Then there exist points $x_0, x_1, \ldots, x_n$ and $y_0, y_1, \ldots, y_n$ such that

$$x_j \in X, \quad y_j \in \bigwedge_{\substack{i=0 \\ i \neq j}}^{n} Y_i, \quad j = 0, 1, \ldots, n$$

and p is a convex linear combination of each pair x_j, y_j. Now a trivial induction over k yields that, whenever y is a convex linear combination of $y_0, y_1, \ldots, y_k$ $(k \leqq n)$ then there is a convex linear combination x of $x_0, x_1, \ldots, x_k$ such that p is a convex linear combination of x and y.

We are ready to apply Helly's theorem. Let Y_i' be the convex closure of $\{y_0, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n\}$. Then

$$y_j \in \bigwedge_{\substack{i=0 \\ i \neq j}}^{n} Y_i', \quad j = 0, 1, \ldots, n.$$

By Helly's theorem, the intersection of the Y_i' is not empty. Let

$$y \in \bigwedge_{i=0}^{n} Y_i'.$$

y is a convex linear combination of, say, $y_0, y_1, \ldots, y_{n-1}$. Applying our last observation, there is an x in the convex closure of $x_0, x_1, \ldots, x_{n-1}$ (hence also in X) such that p is in the convex closure of x and y:

$$p \in X \vee \bigwedge_{i=0}^{n} Y_i' \subseteq X \vee \bigwedge_{i=0}^{n} Y_i,$$

as claimed.

Finally, $\mathfrak{L}(E^{n-1})$ is not $(n-1)$-distributive, as the following counterexample shows: Let S be a simplex, let $x \in S$ such that x is not contained in any $(n-2)$-dimensional face of S, and let $y_0, y_1, \ldots, y_{n-1}$ be the extremal points of S. Then

$$\{x\} \wedge \bigvee_{i=0}^{n-1} \{y_i\} = \{x\} \neq \emptyset = \bigvee_{j=0}^{n-1} \Big[\{x\} \wedge \bigvee_{\substack{i=0 \\ i \neq j}}^{n-1} \{y_i\}\Big].$$

$\mathfrak{L}(E^{n-1})$ is not dually $(n-1)$-distributive either: Let X be a closed halfspace disjoint from S (S is also closed) and let $Y_0, Y_1, \ldots, Y_{n-1}$ be the $(n-2)$-dimensional faces of S. Then

$$X \vee \bigwedge_{i=0}^{n-1} Y_i = X \vee \emptyset = X,$$

which is a proper part of

$$\bigwedge_{j=0}^{n-1} \Big[X \vee \bigwedge_{\substack{i=0 \\ i \neq j}}^{n-1} Y_i\Big] = \bigwedge_{j=0}^{n-1} [X \vee \{y_j\}].$$

3. On the variety generated by all finite n-distributive lattices. In this section we prove Theorem 1.2 via the following three lemmas.

Lemma 3.1. *$\mathfrak{L}(E^{n-1}) \in \mathrm{HSP}(\mathfrak{L}_{\mathrm{fin}}(E^{n-1}))$. where $\mathfrak{L}_{\mathrm{fin}}(E^{n-1})$ denotes the set of all those convex sets of E^{n-1} that are the convex closures of a finite set of points.*

Proof. Every element of $\mathfrak{L}(E^{n-1})$ is a join of atoms and every atom of $\mathfrak{L}(E^{n-1})$ is compact by Carathéodory's theorem. Thus $\mathfrak{L}(E^{n-1})$ is algebraic. Furthermore, its compact elements are exactly the elements of $\mathfrak{L}_{\mathrm{fin}}(E^{n-1})$. Hence $\mathfrak{L}(E^{n-1})$ is isomorphic to the ideal lattice of $\mathfrak{L}_{\mathrm{fin}}(E^{n-1})$, whence it is in the variety generated by $\mathfrak{L}_{\mathrm{fin}}(E^{n-1})$.

In the above proof we implicitely made use of the fact that $\mathfrak{L}_{\mathrm{fin}}(E^{n-1})$ is a sublattice of $\mathfrak{L}(E^{n-1})$, that is, the intersection of two convex polytopes is a convex polytope, otherwise we could not have spoken of the *lattice* $\mathfrak{L}_{\mathrm{fin}}(E^{n-1})$.

Now let H be any finite subset of E^{n-1}, and let $\mathfrak{L}(H)$ denote the set of all those subsets X of H which are of the form $X=C\cap H$ with $C\subseteqq E^{n-1}$ convex. Clearly

$$\mathfrak{L}(H) = \{X(\subseteqq H) \mid X = (\operatorname{conv} X)\cap H\},$$

where "conv" denotes the operator associating with any set its convex hull. Now it is clear that $\mathfrak{L}(H)$ is a lattice relative to the inclusion and its operations $\vee^H$ and $\wedge^H$ are as follows.

$$X\vee^H Y = (\operatorname{conv} X \vee \operatorname{conv} Y)\cap H,$$

$$X\wedge^H Y = (\operatorname{conv} X \wedge \operatorname{conv} Y)\cap H = X\cap Y,$$

where $\vee$ and $\wedge$ are the operations in $\mathfrak{L}(E^{n-1})$.

Lemma 3.2. *$\mathfrak{L}(H)$ is n-distributive.*

Proof. Assume that $X, Y_0, Y_1, \ldots, Y_n \in \mathfrak{L}(H)$, $p\in H$, and

$$p\in X\wedge^H \bigvee_i{}^H Y_i.$$

As in the proof of Theorem 1.1, Carathéodory's theorem and the descriptions of $\vee^H$ and $\wedge^H$ before the Lemma yield that there is a $j\in\{0, 1, \ldots, n\}$ such that

$$p\in \bigvee_{\substack{i\\ i\neq j}}{}^H Y_i,$$

that is,

$$p\in \bigvee_j{}^H [X\wedge^H \bigvee_{\substack{i\\ i\neq j}}{}^H X_i],$$

proving the lemma.

The following lemma finishes the proof of Theorem 1.2.

Lemma 3.3. $\mathfrak{L}_{\mathrm{fin}}(E^{n-1})\in \mathrm{HSP}(\mathfrak{L}(H) | H\subseteqq E^{n-1}, |H|<\aleph_0)$.

Proof. Let $\mathscr{H}=\{H \mid H\subseteqq E^{n-1}, |H|<\aleph_0\}$. Let

$$L = \prod_{H\in\mathscr{H}} \mathfrak{L}(H),$$

and let M consist of all $a\in L$ for which there is a $P\in\mathfrak{L}_{\mathrm{fin}}(E^{n-1})$ with the property that for some $H_0\in\mathscr{H}$ and for all $H\in\mathscr{H}$ containing H_0, we have $a(H)=H\cap P$. If $a\in M$ and P has the above property, then P is called a support of a. The support of a is uniquely determined. Indeed, if $P\neq P'\in\mathfrak{L}_{\mathrm{fin}}(E^{n-1})$, $H_0, H_0'\in\mathscr{H}$, $a(H)=P\cap H$ for all $H_0\subseteqq H\in\mathscr{H}$ and $a(H)=P'\cap H$ for all $H_0'\subseteqq H\in\mathscr{H}$ then extend $H_0\cup H_0'$ to an $H\in\mathscr{H}$ that contains an element from the symmetric difference $P\triangle P'$. For this H we have $a(H)=P\cap H\neq P'\cap H=a(H)$, a contradiction.

We first prove that M is a sublattice of L. Let $a, b \in M$, let P_a and P_b be the supports of a and b, respectively, and choose H_a and H_b such that

$$a(H) = H \cap P_a \quad \text{if} \quad H_a \subseteqq H \in \mathscr{H}$$

and

$$b(H) = H \cap P_b \quad \text{if} \quad H_b \subseteqq H \in \mathscr{H}.$$

Let $H_0 \in \mathscr{H}$ contain the sets H_a and H_b and the sets of extremal points of P_a and of P_b. Then we have

$$\operatorname{conv}(H \cap P_a) = P_a, \quad \operatorname{conv}(H \cap P_b) = P_b$$

whenever $H_0 \subseteqq H \in \mathscr{H}$. Compute the values of $a \vee b$ and $a \wedge b$ at H (H as above).

$$(a \vee b)(H) = a(H) \vee^H b(H) = (H \cap P_a) \vee^H (H \cap P_b) =$$
$$= (\operatorname{conv}(H \cap P_a) \vee \operatorname{conv}(H \cap P_b)) \cap H = (P_a \vee P_b) \cap H.$$

Clearly $P_a \vee P_b \in \mathfrak{L}_{\mathrm{fin}}(E^{n-1})$, whence $a \vee b \in M$,

$$(a \wedge b)(H) = a(H) \wedge^H b(H) = (H \cap P_a) \cap (H \cap P_b) = H \cap (P_a \wedge P_b).$$

Applying that $P_a \wedge P_b \in \mathfrak{L}_{\mathrm{fin}}(E^{n-1})$, we obtain that $a \wedge b \in M$.

We have also obtained that the map $M \to \mathfrak{L}_{\mathrm{fin}}(E^{n-1})$, $a \mapsto P_a$ is a lattice homomorphism. For any $P \in \mathfrak{L}_{\mathrm{fin}}(E^{n-1})$, P is the support of the choice function a defined by $a(H) = P \cap H$. Hence $\mathfrak{L}_{\mathrm{fin}}(E^{n-1})$ is a homomorphic image of M, which completes the proof.

4. The lattice of closed convex sets. In this section we prove Theorem 1.3. The operations of $\bar{\mathfrak{L}}(E^{n-1})$ will be denoted as sum and product. Obviously, $XY = X \wedge Y$ and $X + Y$ is the topological closure of $X \vee Y$ if $X, Y \in \bar{\mathfrak{L}}(E^{n-1})$. Choose a point

$$p \in X \sum_{i=0}^{n} Y_i,$$

where $X, Y_0, Y_1, \ldots, Y_n \in \bar{\mathfrak{L}}(E^{n-1})$. Then $p \in X$ and $p = \lim_{m \to \infty} p_m$ for some $\{p_m\}_{m \in \mathbf{N}} \subseteqq \bigvee_{i=0}^{n} Y_i$. By Carathéodory's theorem, for every $m \in N$ there is a $j(m) \in \{0, 1, \ldots, n\}$ such that $p_m \in \bigvee_{\substack{i=0 \\ i \neq j(m)}}^{n} Y_i$. For at least one $k \in \{0, 1, \ldots, n\}$, $k = j(m)$ for infinitely many $m \in N$. Therefore, the subsequence $\{p_m\}_{j(m)=k}$ of $\{p_m\}_{m \in \mathbf{N}}$ is infinite and converges to p. Besides $p_m \in \bigvee_{\substack{i=0 \\ i \neq k}}^{n} Y_i$. Hence

$$p \in X \sum_{\substack{i=0 \\ i \neq k}}^{n} Y_i.$$

Thus

$$X \sum_{i=0}^{n} Y_i \subseteq \sum_{k=0}^{n} [X \sum_{\substack{i=0\\ i\neq k}}^{n} Y_i].$$

To prove the dual n-distributivity, we need a lemma.

Lemma 4.1. *Let $p, q, r \in E^{n-1}$. Then, for any $u \in \operatorname{conv}\{p, r\}$, $v \in \operatorname{conv}\{q, s\}$, and $x \in \operatorname{conv}\{p, q\}$, there exist $y \in \operatorname{conv}\{r, s\}$ and $z \in \operatorname{conv}\{u, v\}$ such that $z \in \operatorname{conv}\{x, y\}$.*

Proof. We may assume that $u \notin \{p, r\}$ and $v \notin \{q, s\}$ as otherwise the statement is trivial. The conditions of the lemma show that there exist real numbers $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2$ such that

$$q = \alpha_1 s + \alpha_2 v, \quad \alpha_1 + \alpha_2 = 1, \quad \alpha_1 \leqq 0,$$

$$p = \beta_1 r + \beta_2 u, \quad \beta_1 + \beta_2 = 1, \quad \beta_1 \leqq 0,$$

$$x = \gamma_1 q + \gamma_2 p, \quad \gamma_1 + \gamma_2 = 1, \quad \gamma_1, \gamma_2 \geqq 0.$$

Hence

$$x = \gamma_1 \alpha_1 s + \gamma_1 \alpha_2 v + \gamma_2 \beta_1 r + \gamma_2 \beta_2 u =$$

$$= (\gamma_1 \alpha_1 + \gamma_2 \beta_1)\left(\frac{\gamma_1 \alpha_1}{\gamma_1 \alpha_1 + \gamma_2 \beta_1} s + \frac{\gamma_2 \beta_1}{\gamma_1 \alpha_1 + \gamma_2 \beta_1} r\right) +$$

$$+ (\gamma_1 \alpha_2 + \gamma_2 \beta_2)\left(\frac{\gamma_1 \alpha_2}{\gamma_1 \alpha_2 + \gamma_2 \beta_2} v + \frac{\gamma_2 \beta_2}{\gamma_1 \alpha_2 + \gamma_2 \beta_2} u\right) = \delta_1 y + \delta_2 z,$$

where

$$\delta_1 = \gamma_1 \alpha_1 + \gamma_2 \beta_1, \quad \delta_2 = \gamma_1 \alpha_2 + \gamma_2 \beta_2,$$

$$y = \frac{\gamma_1 \alpha_1}{\gamma_1 \alpha_1 + \gamma_2 \beta_1} s + \frac{\gamma_2 \beta_1}{\gamma_1 \alpha_1 + \gamma_2 \beta_1} r,$$

$$z = \frac{\gamma_1 \alpha_2}{\gamma_1 \alpha_2 + \gamma_2 \beta_2} v + \frac{\gamma_2 \beta_2}{\gamma_1 \alpha_2 + \gamma_2 \beta_2} u.$$

This representation shows that $y \in \operatorname{conv}\{s, r\}$, $z \in \operatorname{conv}\{u, v\}$ (the coefficients are non-negative and sum up to 1). Finally, $\delta_1 + \delta_2 = 1$, $\delta_1 \leqq 0$ yield that $z \in \operatorname{conv}\{x, y\}$.

The following extension of this lemma is now proved by an easy induction over k.

Corollary. *Let $p_0, p_1, \dots, p_k, q_0, q_1, \dots, q_k, r_0, r_1, \dots, r_k \in E^{n-1}$. Assume $r_i \in \operatorname{conv}\{p_i, q_i\}$, $i = 0, 1, \dots, k$. Let $p \in \operatorname{conv}\{p_0, p_1, \dots, p_k\}$. Then there exist $q \in \operatorname{conv}\{q_0, q_1, \dots, q_k\}$ and $r \in \operatorname{conv}\{r_0, r_1, \dots, r_k\}$ such that $r \in \operatorname{conv}\{p, q\}$.*

Now we pass on to prove the dual n-distributivity of $\mathfrak{L}(E^{n-1})$. Let

$$p\in \prod_{j=0}^{n}[X+\prod_{\substack{i=0\\ i\neq j}}^{n} Y_i],$$

where $X, Y_0, Y_1, \ldots, Y_n\in\mathfrak{L}(E^{n-1})$. Then there exist sequences $\{p_{jm}\}_{m\in N}$, $j=0, 1, \ldots$ $\ldots, n$, each converging to p, such that

$$p_{jm}\in X\vee \prod_{\substack{i=0\\ i\neq j}}^{n} Y_i, \quad m\in N, \quad j=0, 1, \ldots, n.$$

Now choose, for all $m\in N$ and $j=0, 1, \ldots, n$,

$$x_{jm}\in X, \quad y_{jm}\in \prod_{\substack{i=0\\ i\neq j}}^{n} Y_i$$

such that p_{jm} is a convex linear combination of x_{jm} and y_{jm}. By Helly's theorem there exists an

$$y_m\in \prod_{i=0}^{n} Y_i$$

for all $m\in N$, and y_m can be chosen to be an element of conv $\{y_{0m}, y_{1m}, \ldots, y_{nm}\}$. Thus, by the Corollary, there exist points $x_m\in$conv $\{x_{0m}, x_{1m}, \ldots, x_{nm}\}$ and $p_m\in$ $\in$conv $\{p_{0m}, p_{1m}, \ldots, p_{nm}\}$ with $p_m\in$conv $\{x_m, y_m\}$ for all $m\in N$. Obviously, $p_m\to p$ as $m\to\infty$, thus p is in the topological closure of $\{p_m\}_{m\in N}$ and each p_m is a member of $X\vee \prod_{i=0}^{n} Y_i$. Hence

$$p\in X+\prod_{i=0}^{n} Y_i.$$

The counterexamples at the end of Section 2 also show that $\mathfrak{L}(E^{n-1})\notin\Delta_{n-1}$, ∇_{n-1}.

5. Complemented modular lattices revisited. n-distributivity of complemented modular lattices was studied in [4]. Here we add a result describing those projective geometries in which "Carathéodory's theorem holds". As it is well-known by Frink [2] there is a one-to-one correspondence between projective geometries and their subspace lattices, which are exactly the complete, complemented, modular, atomic lattices such that every atom is compact. It will be convenient to call *these lattices* projective geometries. We say that a projective geometry M satisfies the property (C_n) iff, for any atoms $p, p_1, \ldots, p_m$, $m\geqq n+1$ of M with $p\leqq \bigvee_{i=1}^{m} p_i$, there exist $i_1, i_2, \ldots, i_n\in\{1, 2, \ldots, m\}$ such that $p\leqq \bigvee_{j=1}^{n} p_{i_j}$.

A lattice is called infinitely n-distributive iff it satisfies the identity

$$x \wedge \bigvee_{i \in I} Y_i = \bigvee_{\substack{K \subseteq I \\ |K| = n}} [x \wedge \bigvee_{i \in K} Y_i]$$

for arbitrary index set I. It is called completely n-distributive iff the identity

$$\bigwedge_{i \in I} \bigvee_{j \in J_i} x_{ij} = \bigvee_{\varphi} \bigwedge_{i \in I} \bigvee_{j \in \varphi(i)} x_{ij}$$

holds in it for arbitrary I and J_i, $i \in I$ and $|J_i| \geqq n$, where the $\bigvee_\varphi$ at the right hand side is to be formed for all choice functions $\varphi\colon I \to \bigcup_{i \in I} P_n(J_i)$ (with $\varphi(i) \in P_n(J_i)$), where $P_n(J_i)$ denotes the set of n element subsets of J_i, $i \in I$. Now we are ready to state the main result of this section.

Theorem 5.1. *Let L be a complete complemented modular lattice. Then the following conditions are equivalent:*

(i) *L is a projective geometry satisfying (C_n);*
(ii) *L is atomic and infinitely n-distributive;*
(iii) *L is completely n-distributive,*
(iv) *L is isomorphic to a direct product of irreducible projective geometries of length $\leqq n$.*

Corollary. *The dual of a projective geometry satisfying (C_n) also satisfies (C_n). The dual of a completely n-distributive complemented modular lattice is also completely n-distributive.*

Proof. (i)$\Rightarrow$(iv). If (i) holds, then, by FRINK [2], Theorem 7, Corollary, L is a direct product of irreducible projective geometries L_γ, $\gamma \in \Gamma$. We show that L_γ must be of length $\leqq n$ for all $\gamma \in L$. Indeed, in the contrary case L_γ contains an independent set of $n+1$ atoms: $p_0, p_1, \dots, p_n$. By irreducibility, $p_0 \vee p_1 \geqq p_{01}$ for some atom $p_{01} \neq p_0, p_1$. We have also $p_0 \vee p_1 \vee p_2 \geqq p_{01} \vee p_2 \geqq p_{012}$ for some atom $p_{012} \neq p_{01}, p_2$. Clearly, $p_{012} \nleqq p_0 \vee p_1$ (otherwise $p_0 \vee p_1 \geqq p_{012} \vee p_{01} \geqq p_2$, a contradiction). Similarly, for $\{i, j\} = \{0, 1\}$, $p_{012} \nleqq p_i \vee p_2$ as otherwise $p_i \vee p_2 = p_i \vee p_{012} \vee p_2 = p_i \vee p_{01} \vee p_2 = p_j \vee p_{01} \vee p_2 \geqq p_j$. By induction, we find an atom $p_{01\dots n} \leqq p_0 \vee p_1 \vee \dots \vee p_n$ such that $p_{01\dots n} \nleqq p_0 \vee \dots \vee p_{i-1} \vee p_{i+1} \vee \dots \vee p_n$, $i = 0, 1, \dots, n$. This contradicts (C_n).

(iv)$\Rightarrow$(iii). Irreducible projective geometries of length $\leqq n$ are completely n-distributive (in fact, any meet of joins equals one of the meets of n element subjoins), hence so are their direct products.

(iii)$\Rightarrow$(ii). It is easily seen that complete n-distributivity implies infinite n-distributivity. So we only have to show that L is atomic. It suffices to show that every element of L is a join of elements of height $\leqq n$. Let $x \in L$ be of height greater than n. Consider all independent sets $\{x_{\gamma 0}, x_{\gamma 1}, \dots, x_{\gamma n}\}$, $\gamma \in \Gamma$ such that $\bigvee_{i=0}^{n} x_{\gamma i} = x$. As usual, H_n^Γ denotes the set of all mappings of the set Γ to $H_n=\{0, 1, \ldots, n\}$. By the complete n-distributive law,

$$x = \bigwedge_{\gamma\in\Gamma}\bigvee_{i=0}^{n} x_{\gamma i} = \bigvee_{m_1\in H_n^\Gamma}\cdots\bigvee_{m_n\in H_n^\Gamma}\bigwedge_{\gamma\in\Gamma}(x_{\gamma m_1(\gamma)}\vee\ldots\vee x_{\gamma m_n(\gamma)}).$$

We show that the elements

$$z_{m_1\ldots m_n} = \bigwedge_{\gamma\in\Gamma}\bigvee_{i=1}^{n} x_{\gamma m_i(\gamma)}$$

are of height $\leqq n$. Indeed, in the contrary case, some of the intervals $[0, z_{m_1\ldots m_n}]$ contains a chain of $n+1$ elements. Thus there is an independent set $\{x_1, x_2, \ldots, x_n\}$ such that $x_0' := \bigvee_{i=1}^{n} x_i < z_{m_1\ldots m_n}$ and $\bigwedge_{i=1}^{n} x_i = 0$. Let x_0 be a complement of x_0' in $[0, x]$. Then $\bigvee_{i=0}^{n} x_i = x$. Therefore, some of the joins $\bigvee_{i=0, i\neq j}^{n} x_i$ occurs in the $\wedge$-representation of $z_{m_1\ldots m_n}$. For $j=0$, this yields $x_0'\geqq z_{m_1\ldots m_n}$, a contradiction. If $j\neq 0$, then

$$x_0' = x_0'\wedge z_{m_1\ldots m_n} \leqq x_0'\wedge\bigvee_{\substack{i=0\\ i\neq j}}^{n} x_i = \bigvee_{\substack{i=0\\ i\neq 0, j}}^{n} x_i < \bigvee_{i=1}^{n} x_i = x_0'.$$

This contradiction yields (ii).

The implication (ii)$\Rightarrow$(i) being very easy, the proof is complete.

6. Modular lattices. In this section we prove Theorem 1.4. By a result of Faigle [1], every modular lattice M can be embedded into a modular lattice M' such that every element of M' is a join of compact completely join-irreducible elements. If we prove that M' is in $\mathrm{HSP}(\Delta_n\cap F)$, then the theorem follows. Let $\mathscr{P}$ be the set of all completely join-irreducible elements of M (these elements are all compact) and let $\mathscr{H}$ be the set of all finite subsets of $\mathscr{P}$. For any $H\in\mathscr{H}$, let M_H denote the set of all finite joins (in M') of elements of H. M_H is clearly a lattice relative to the ordering of M'. Let $\wedge^H$ and $\vee^H$ denote the operations in M_H (note that $\vee^H$ is the same as $\vee$). For any element $x\in M'$, and, for any $H\in\mathscr{H}$, let $x_H=\sup\{y \mid y\leqq x, y\in M_H\}$. Then

$$x\wedge y = \bigvee_{H\in\mathscr{H}}(x_H\wedge^H y_H)$$

and

$$x\vee y = \bigvee_{H\in\mathscr{H}}(x_H\vee^H y_H).$$

Indeed, observe that $x=\bigvee_H x_X$ and $H\subseteqq G\in\mathscr{H}$ implies $x_H\leqq x_G$. If $p\leqq x\wedge y$ for some $p\in\mathscr{P}$ then $x_H=y_H=p$ holds for $H=\{p\}$, whence $p\leqq p\wedge p=x_H\wedge^H y_H$. This proves the first equality. Now let $p\leqq x\vee y$. Then $p\leqq\bigvee_{H,K}(x_H\vee y_K)=\bigvee_H(x_H\vee y_H)=\bigvee_H(x_H\vee^H y_H)$, proving the second equality.

Assume that $p=q$ is an m-ary lattice identity holding in all finite n-distributive lattices. Then $p=q$ holds in all the lattices M_H. Let $x_1, x_2, \dots, x_m \in M'$, and let p^H and q^H be the realizations of p and q in M. Then

$$p(x_1, x_2, \dots, x_m) = \bigvee_{H \in \mathscr{H}} p^H((x_1)_H, (x_2)_H, \dots, (x_m)_H) =$$

$$= \bigvee_{H \in \mathscr{H}} q^H((x_1)_H, (x_2)_H, \dots, (x_m)_H) = q(x_1, x_2, \dots, x_m).$$

References

[1] U. FAIGLE, Frink's theorem for modular lattices, Preprint.

[2] O. FRINK, Complemented modular lattices and projective spaces of infinite dimension, *Trans. Amer. Math. Soc.*, **60** (1946), 452—467.

[3] H. GERSTMANN, n-Distributivgesetze, *Acta Sci. Math.*, **46** (1983), 99—113.

[4] A. HUHN, Two notes on n-distributive lattices, *Colloq. Math. Soc. János Bolyai*, **14** *Lattice Theory*, North-Holland (Amsterdam, 1976), 137—147.

[5] V. L. KLEE [editor], *Convexity, Proc. Sympos. Pure Math.*, Vol. **7**, Amer. Math. Soc. (Providence, R. I., 1963).

Acta Sci. Math., 52 (1988), 47—52

Free product of ortholattices

SYLVIA PULMANNOVÁ

The purpose of this paper is to prove a structure theorem for the free product of ortholattices. The method of BRUNS [1] for constructing a free ortholattice is combined with GRÄTZER's method for constructing the free product of lattices [2].

An ortholattice is a lattice L with a smallest element 0 and a largest element 1 and with an orthocomplementation $':L\to L$ such that

(i) $a''=a,\ \ a\in L,$

(ii) $a\leqq b$ implies $b'\leqq a',\ \ a,b\in L,$

(iii) $a\vee a'=1,\ \ a\wedge a'=0,\ \ a\in L.$

The free product of ortholattices is defined as follows.

Definition 1. Let $(L_i, 0_i, 1_i, '^i)$, $i\in I$, be a set of ortholattices. An ortholattice $(L, 0, 1, ')$ is a free product of the ortholattices L_i, $i\in I$, if

(i) for any $i\in I$, there is an injective homomorphism u_i: $L_i\to L$ which preserves the lattice operations and orthocomplementation so that each L_i can be considered as a subalgebra of L, and for $i,j\in I$, $i\neq j$, $L_i-\{0_i, 1_i\}$ and $L_j-\{0_j, 1_j\}$ are disjoint;

(ii) L is generated by $\bigcup\{u_i(L_i)\colon i\in I\}$;

(iii) for any ortholattice A and for a family of homomorphisms φ_i: $L_i\to A$, $i\in I$, there exists a homomorphism φ: $L\to A$ such that $\varphi\circ u_i$ agrees with φ_i for all $i\in I$.

Definition 2. Let X be an arbitrary set. The set $P(X)$ of polynomials over X is the smallest set satisfying (i) and (ii), where

(i) $X\subset P(X)$,

(ii) if $p,q\in P(X)$, then $p\vee q$ and $p\wedge q\in P(X)$.

For a lattice A we define $A^b=A\cup\{0^b, 1^b\}$, where $0^b, 1^b\notin A$, and we order A^b by the rules: $0^b<x<1^b$ for any $x\in A$, $x\leqq y$ in A^b if $x,y\in A$ and $x\leqq y$ in A. Thus

Received February 13, 1985.

$A^b \neq A$ and we have $a \wedge b = 0^b$ only if $a = 0^b$ or $b = 0^b$, and $a \vee b = 1^b$ only if $a = 1^b$ or $b = 1^b$.

Let $\{L_i \colon i \in I\}$ be a set of ortholattices. Put $Q = \cup \{L_i \colon i \in I\}$. We suppose that L_i and L_j are disjoint provided $i \neq j$, $i, j \in I$.

Definition 3. Let $P(Q)$ be the set of polynomials over Q. The upper i-cover of $p \in P(Q)$, $p^{(i)}$, is an element of $(L_i)^b$, defined as follows:

(i) for $a \in Q$ (i.e. $a \in L_i$ for exactly one $i \in I$), $a^{(j)} = a$ if $j = i$, $a^{(j)} = 1^b$ if $j \neq i$.

(ii) $(p \wedge q)^{(i)} = p^{(i)} \wedge q^{(i)}$ and $(p \vee q)^{(i)} = p^{(i} \vee q^{(i)}$, where $\wedge$ and $\vee$ is taken in $(L_i)^b$.

The definition of lower i-cover, $p_{(i)}$, is analogous, with 0^b replacing 1^b in (i).

It is clear that $p^{(i)} \neq 0_b$ and $p_{(i)} \neq 1^b$. An upper or lower i-cover is proper if it is not 1^b or 0^b.

Corollary 4. [2] *For any* $p \in P(Q)$ *and* $i \in I$ *we have that* $p_{(i)} \leqq p^{(i)}$, *and if* $p_{(i)}$ *and* $p^{(j)}$ *are proper and* $p_{(i)} \leqq p^{(j)}$, *then* $i = j$.

Definition 5. For $p, q \in P(Q)$, we put $p \subseteqq q$ if one of the following cases (i)—(vi) below occurs:

(i) $p = q$,

(ii) for some $i \in I$, $p^{(i)} \leqq q_{(i)}$,

(iii) $p = p_0 \wedge p_1$ where $p_0 \subseteqq q$ or $p_1 \subseteqq q$,

(iv) $p = p_0 \vee p_1$ where $p_0 \subseteqq q$ and $p_1 \subseteqq q$,

(v) $q = q_0 \wedge q_1$ where $p \subseteqq q_0$ and $p \subseteqq q_1$,

(vi) $q = q_0 \vee q_1$ where $p \subseteqq q_0$ or $p \subseteqq q_1$.

The rank $r(p)$ of a $p \in P(Q)$ is defined as follows: for $p \in Q$, $r(p) = 1$ and $r(p) = r(p_1) + r(p_2)$ if $p = p_1 \wedge p_2$ or $p = p_1 \vee p_2$.

Lemma 6. [2] *Let* $p, q, r \in P(Q)$ *and* $i \in I$. *Then*

(i) $p \subseteqq q$ *implies* $p_{(i)} \leqq q_{(i)}$ *and* $p^{(i)} \leqq q^{(i)}$.

(ii) $p \subseteqq q$ *and* $q \subseteqq r$ *imply* $p \subseteqq r$.

Since by 5 (i), $p \subseteqq p$ for any $p \in P(Q)$, the relation $\subseteqq$ is a quasiordering, and so we can define $p \equiv q$ iff $p \subseteqq q$ and $q \subseteqq p$, $p, q \in P(Q)$. We put

$$R(p) = \{q \colon q \in P(Q) \text{ and } p \equiv q\}, \quad R(Q) = \{R(p) \colon p \in P(Q)\},$$

$$R(p) \leqq R(q) \quad \text{if} \quad p \subseteqq q.$$

Lemma 7. [2] *$R(Q)$ is a lattice, and we have*

$$R(p) \wedge R(q) = R(p \wedge q), \quad R(p) \vee R(q) = R(p \vee q).$$

Furthermore, if $a, b, c, d \in L_i$, $i \in I$, *and if* $a \wedge b = c$, $a \vee b = d$ *in* L_i, *then* $R(a) \wedge R(b) = R(c)$ *and* $R(a) \vee R(b) = R(d)$.

As a consequence of Lemma 7 we get that $p \mapsto R(p)$, $p \in L_i$, is an embedding of L_i into $R(Q)$. Therefore, identifying $p \in L_i$ with $R(p)$ we get each L_i as a sublattice of $R(Q)$, and hence $Q \subset R(Q)$. It is also obvoius that the partial ordering induced by $R(Q)$ on Q agrees with the original partial ordering.

Let us add the set $\{0, 1\}$ to $P(Q)$ and let us define $0 \subseteqq p \subseteqq 1$ for any $p \in P(Q)$, $p \vee 0 = p$, $p \wedge 0 = 0$, $p \vee 1 = 1$, $p \wedge 1 = p$. Let us further define the map $'$ on $P(Q) \cup \{0, 1\}$ as follows: if $x \in L_i$ for some $i \in I$, put $x' = x'^i$; $1' = 0$, $0' = 1$, and recursively, $(a \wedge b)' = a' \vee b'$, $(a \vee b)' = a' \wedge b'$.

We note that the elements 0, 1 are different from the auxiliary elements 0^b and 1^b used in the definition of the lower and upper covers. In the following lemma we put $(0^b)' = 1^b$, $(1^b)' = 0^b$.

Lemma 8. *For any* $p \in P(Q)$, $(p')^{(i)} = (p_{(i)})'$ *and* $(p')_{(i)} = p^{(i)\prime}$.

Proof. We shall proceed by induction on $r(p)$. If $r(p) = 1$, then $p \in L_i$ for some $i \in I$, and $p^{(i)} = p_{(i)} = p$, $p^{(j)} = 1^b$, $p_{(j)} = 0^b$ for $j \neq i$. Therefore, $p^{(j)\prime} = (1^b)' = 0^b = p_{(j)}$ for $j \neq i$, and as $p' = p'^i$ is the orthocomplement of p in L_i, we have $(p')^{(j)} = 1^b$, $(p')_{(j)} = 0^b$ for $j \neq i$. From this we obtain that $(p^{(j)})' = 0^b = (p')_{(j)}$, $(p_{(j)})' = 1^b = (p')^{(j)}$ for $j \neq i$. Further, $(p')^{(i)} = p' = (p_{(i)})'$, $(p')_{(i)} = p' = (p^{(i)})'$. Now let $p = q \vee r$, then $p' = q' \wedge r'$, and $(p')^{(i)} = (q')^{(i)} \wedge (r')^{(i)} = (q_{(i)})' \wedge (r_{(i)})'$ by the induction hypothesis, so that $(p')^{(i)} = (q_{(i)} \vee r_{(i)})' = (p_{(i)})'$, and dually for $p = q \wedge r$. The proof of $(p')_{(i)} = (p^{(i)})'$ is similar.

Lemma 9. $a'' = a$ *for any* $a \in P(Q) \cup \{0, 1\}$, *and* $a \subseteqq b$ *implies* $b' \subseteqq a'$ *for any* $a, b \in P(Q) \cup \{0, 1\}$.

Proof. By the definition, $0'' = 1' = 0$, $1'' = 0' = 1$. If $a \in L_i$ for some $i \in I$, then obviously $a'' = a$. Let $a = b \wedge c$. Then $a' = b' \vee c'$, and $a'' = b'' \wedge c''$. By induction we obtain that $a'' = a$. For $a = b \vee c$ the situation is dual.

Now we shall prove the second statement. If $a = 0$ or $b = 1$, it is obvious. We shall suppose that $a, b \notin \{0, 1\}$ and proceed by induction on $r(a) + r(b)$. If $r(a) + r(b) = 2$, then $a \subseteqq b$ holds by 5 (i) or 5 (ii), so that $a, b \in L_i$ for some $i \in I$, and $a' = a'^i$, $b' = b'^i$, which implies that $b' \subseteqq a'$. Now let $r(a) + r(b) = r$, and let the statement hold for all $r(a) + r(b) < r$. If $a \subseteqq b$ holds by 5 (i), then $a' = b'$. If $a \subseteqq b$ holds by 5 (ii), then $a^{(i)} \leqq b_{(i)}$ for some $i \in I$. By Lemma 8, $(a^{(i)})' = (a')_{(i)}$ and $(b_{(i)})' = (b')^{(i)}$. Therefore $a^{(i)} \leqq b_{(i)}$ implies $(b')^{(i)} \leqq (a')_{(i)}$, which in turn implies that $b' \subseteqq a'$ by 5 (ii). If $a \subseteqq b$ by 5 (iii) with $a = a_0 \wedge a_1$, then $a_0 \subseteqq b$ or $a_1 \subseteqq b$, which implies by the induction hypothesis that $b' \subseteqq a_0'$ or $b' \subseteqq a_1'$. As $a' = a_0' \vee a_1'$, we get that $b' \subseteqq a'$ by 5 (vi). If $a \subseteqq b$ by 5 (iv) and $a = a_0 \vee a_1$, where $a_0 \subseteqq b$ and $a_1 \subseteqq b$, then $b'\subseteqq a_0'$ and $b'\subseteqq a_1'$ and this implies that $b'\subseteqq a_0'\wedge a_1'=a'$ by 5 (v). If $a\subseteqq b$ by 5 (v), then $b=b_0\wedge b_1$ and $a\subseteqq b_0$ and $a\subseteqq b_1$. This implies that $b_0'\subseteqq a'$ and $b_1'\subseteqq a'$, which implies that $b'=b_0'\vee b_1'\subseteqq a'$ by 5 (iv). If $a\subseteqq b$ by 5 (vi), where $b=b_0\vee b_1$ with $a\subseteqq b_0$ or $a\subseteqq b_1$, then $b_0'\subseteqq a'$ or $b_1'\subseteqq a'$, and therefore $b_0'\wedge b_1'\subseteqq a'$ by 5 (iii).

Following Bruns [1], we shall define the subset S of reduced elements in $P(Q)\cup\{0, 1\}$.

Definition 10. Define a subset S of $P(Q)\cup\{0, 1\}$ recursively as follows: a is in S if

(i) $a\in\{0, 1\}$ or $a\in\bigcup\{L_i-\{0_i, 1_i\}: i\in I\}$,

(ii) $a=b\vee c$ with $b, c\in S$ and $b'\nsubseteq a$, $c'\nsubseteq a$,

(iii) $a=b\wedge c$ with $b, c\in S$ and $a\nsubseteq b'$, $a\nsubseteq c'$.

Lemma 11. *The set S is closed under $'$.*

Proof. If $a\in\{0, 1\}$, then obviously $a'\in\{0, 1\}$. If $a\in L_i-\{0_i, 1_i\}$ for some $i\in I$, then $a'\in L_i-\{0_i, 1_i\}$ so that $a'\in S$. If $a=b\vee c$, $b, c\in S$ and $b'\nsubseteq a$, $c'\nsubseteq a$, then $a'=b'\wedge c'$ and $a'\nsubseteq b$, $a'\nsubseteq c$. By induction, $b', c'\in S$, and $a'\in S$ by 10 (iii). If $a=b\wedge c$ with $b, c\in S$ and $a\nsubseteq b'$, $a\nsubseteq c'$, then by induction, $b', c'\in S$, and $b\nsubseteq a'$, $c\nsubseteq a'$ implies that $a'\in S$ by 10 (ii).

Lemma 12. *If $a\in S-\{0, 1\}$ then $a^{(i)}\neq 0_i$ and $a_{(i)}\neq 1_i$ for all $i\in I$.*

Proof. We shall proceed by induction. If $a\in L_i-\{0_i, 1_i\}$ then $a^{(i)}=a_{(i)}=$ $=a\notin\{0_i, 1_i\}$, and $a_{(j)}=0^b$, $a^{(j)}=1^b$ for $j\neq i$. Now let $a=b\vee c$. Let us suppose that $a^{(i)}=0_i$ for some $i\in I$. Then $a^{(i)}=b^{(i)}\vee c^{(i)}$ implies that $b^{(i)}$ and $c^{(i)}$ are proper, and $b^{(i)}=c^{(i)}=0_i$, which contradicts the induction hypothesis. Now let $a=b\wedge c$, $b, c\in S$, $a\nsubseteq b'$, $a\nsubseteq c'$. If $a^{(i)}=0_i$, then $a^{(i)}=b^{(i)}\wedge c^{(i)}$ implies that $b^{(i)}$ or $c^{(i)}$ are proper, and $a^{(i)}=0_i\subseteqq(b^{(i)})'=(b')_{(i)}$ implies by 5 (ii) that $a\subseteqq b'$, a contradiction. Now let us suppose that $a_{(i)}=1_i$ for $a\in S$, $i\in I$. By Lemma 11, $a'\in S$, and by Lemma 8, $(a_{(i)})'=(a')^{(i)}=0_i$, which contradicts the above part of the proof.

Lemma 13. *For any $a\in S-\{1\}$, $a'\nsubseteq a$. If $a\in P(Q)$ and $b\in S-\{1\}$, then $a\nsubseteq b$ or $a'\nsubseteq b$.*

Proof. If $a\in S-\{1\}$ by 10 (i), then $a=0$ or $a\in L_i-\{0_i, 1_i\}$ for some $i\in I$. In both cases $a'\nsubseteq a$ holds. Now let us suppose that $a\in S-\{1\}$ and $a'\subseteqq a$ holds by 5 (ii). Then $(a')^{(i)}\leqq a_{(i)}$ for some $i\in I$. This implies that $(a')^{(i)}=(a_{(i)})'\leqq a_{(i)}$, but this is impossible by Lemma 12. Now let $a\in S$ by 10 (ii) with $a=b\vee c$. If $b'\wedge c'\subseteqq$ $\subseteqq b\vee c$ holds by 5 (iii), then $b'\subseteqq b\vee c$ or $c'\subseteqq b\vee c$, which contradicts 10 (ii). If $b'\wedge c'\subseteqq b\vee c$ by 5 (vi), then $b'\wedge c'\subseteqq b$ or $b'\wedge c'\subseteqq c$. From this it follows that $b'\subseteqq a$ or $c'\leqq a$, contradicting 10 (ii). If $a=b\wedge c$, then if $b'\vee c'\subseteqq b\wedge c$ by 5 (iv), then

$b' \subseteqq b \wedge c$ and $c' \subseteqq b \wedge c$. But this implies that $b' \subseteqq b$ and $c' \subseteqq c$ by 5 (v), contradicting the induction hypothesis. If $b' \vee c' \subseteqq b \wedge c$ by 5 (v), then $b' \vee c' \subseteqq b$ and $b' \vee c' \subseteqq c$, and this implies by 5 (iv) that $b' \subseteqq b$ and $c' \subseteqq c$, contradicting the induction hypothesis. Thus the first part of Lemma 13 is proved.

Finally, if $a \subseteqq b$ and $a' \subseteqq b$ with $a \in P(Q)$ and $b \in S - \{1\}$, then $a' \subseteqq b$ implies $b' \subseteqq a$, and this together with $a \subseteqq b$ gives $b' \subseteqq b$, which contradicts the first part of the proof.

Obviously, the relation $\subseteqq$ defined on $S \times S$ by Definition 5 together with the rule $0 \subseteqq x \subseteqq 1$ for all $x \in S$, is a quasiordering on S. Let Θ be the relation defined on $S \times S$ by $a \Theta b$ iff $a \subseteqq b$ and $b \subseteqq a$. We prove now that S/Θ is an ortholattice with $0/\Theta$ as the smallest and $1/\Theta$ as the largest element, $a/\Theta \vee b/\Theta = (a \vee b)/\Theta$ if $a \vee b \in S$ and $a/\Theta \vee b/\Theta = 1/\Theta$ if $a \vee b \notin S$, and, finally, that $a/\Theta \mapsto a'/\Theta$ is an orthocomplementation.

Let us define $a/\Theta \leqq b/\Theta$ iff $a \subseteqq b$, $a, b \in S$. Obviously, $\leqq$ is a partial ordering on S/Θ, and $0/\Theta$ and $1/\Theta$ are the smallest and largest element of S/Θ, respectively. If, for $a, b \in S$, the element $a \vee b \in S$, then $a/\Theta \vee b/\Theta = (a \vee b)/\Theta$ by Lemma 7. If, for $a, b \in S$, the element $a \vee b \notin S$, then $a' \subseteqq a \vee b$ or $b' \subseteqq a \vee b$ holds, and for every c in S such that $a, b \subseteqq c$ we get by 5 (iv) that $a, a' \subseteqq c$ or $b, b' \subseteqq c$. This implies by Lemma 13 that $c = 1$. Thus $1/\Theta$ is the supremum of a/Θ and b/Θ. For meets the situation is dual. Therefore, S/Θ is a lattice. For every $a \in S - \{0, 1\}$ the elements $a \vee a'$ and $a \wedge a'$ are not in S, and this implies that a'/Θ is the complement of a/Θ in S.

Theorem 14. *Let $\{L_i\colon i \in I\}$ be a set of ortholattices and let $Q = \bigcup \{L_i\colon i \in I\}$. Denote by $P(Q)$ the set of all polynomials over Q and by S the subset of $P(Q) \cup \{0, 1\}$ given by Definition 10. Finally, let Θ be the congruence relation defined by $a \Theta b$ iff $a \subseteqq b$ and $b \subseteqq a$. Then S/Θ is a free product of L_i, $i \in I$.*

Proof. Put $L = S/\Theta$. We have to prove that

(i) each L_i, $i \in I$, is a subalgebra of L and for $i, j \in I$, $i \neq j$, $L_i - \{0_i, 1_i\}$ and $L_j - \{0_j, 1_j\}$ are disjoint,

(ii) L is generated by $\bigcup \{L_i\colon i \in I\}$,

(iii) for any ortholattice A and for a family of homomorphisms $\varphi_i\colon L_i \to A$, $i \in I$, there exists a homomorphism $\varphi\colon L \to A$ such that φ agrees on L_i with φ_i for all $i \in I$.

(i) We have already proved that L is an ortholattice. Define $\psi_i\colon L_i \to L$ by $\psi_i(x) = x/\Theta \equiv R(x)$ if $x \in L_i - \{0_i, 1_i\}$, and $\psi_i(1_i) = 1/\Theta$, $\psi_i(0_i) = 0/\Theta$. Clearly, we have $\psi_i(x') \equiv \psi_i(x)'$, and $\psi_i(x \vee y) = \psi_i(x) \vee \psi_i(y)$ for $x, y \in L_i$. If $x \in L_i$, $x \neq 0_i$, then $x/\Theta \neq 0/\Theta$, which implies that ψ_i is an embedding.

(ii) is clear.

(iii) We define inductively a map $v\colon P(Q) \to A$ as follows: for $p \in Q$ we set

$\nu(p)=\varphi_i(p)$ if $p\in L_i$, $i\in I$. If $p=p_0\wedge p_1$ or $p=p_0\vee p_1$, $\nu(p_0)$ and $\nu(p_1)$ have already been defined, we set $\nu(p)=\nu(p_0)\wedge\nu(p_1)$ or $\nu(p)=\nu(p_0)\vee\nu(p_1)$, respectively.

We need the following lemma.

Lemma 15. *For $p\in P(Q)$ and $i\in I$, the following hold.*

(i) *If $p_{(i)}$ is proper, then $\nu(p_{(i)})\leqq\nu(p)$.*

(ii) *If $p^{(i)}$ is proper, then $\nu(p)\leqq\nu(p^{(i)})$.*

(iii) *$p\subseteqq q$ implies that $\nu(p)\leqq\nu(q)$.*

(iv) *$\nu(p')=\nu(p)'$ in A.*

Proof. (i)—(iii) The proof is the same as the proof of Lemma 9 in [2].

(iv) If $p\in Q$, then $p\in L_i$ for exactly one $i\in I$, and $\nu(p)=\varphi_i(p)$, so that $\nu(p')=\varphi_i(p')=\varphi_i(p)'=\nu(p)'$. If $p=p_0\wedge p_1$, then $\nu(p')=\nu(p_0'\vee p_1')=\nu(p_0')\vee\nu(p_1')=\nu(p_0)'\vee\nu(p_1)'$ by the induction hypothesis, which implies that $\nu(p')=(\nu(p_0)\wedge\nu(p_1))'=\nu(p)'$. The situation for $p=p_0\vee p_1$ is dual.

Now we can complete the proof of Theorem 14. Take a $p\in S$ and define $\varphi(p/\Theta)=\nu(p)$ if $p\in S-\{0,1\}$, and $\varphi(1/\Theta)=1$, $\varphi(0/\Theta)=0$ in A. φ is well-defined since if $p,q\in S-\{0,1\}$ and $p/\Theta=q/\Theta$, then $p\subseteqq q$ and $q\subseteqq p$, which implies by Lemma 15 that $\nu(p)=\nu(q)$. Further, $\varphi(p/\Theta\wedge q/\Theta)=\varphi((p\wedge q)/\Theta)=\nu(p\wedge q)=\nu(p)\wedge\nu(q)=\varphi(p/\Theta)\wedge\varphi(q/\Theta)$ if $p\wedge q\in S$, $p,q\in S-\{0,1\}$. Clearly, $\varphi(p/\Theta\wedge 0/\Theta)=\varphi(0/\Theta)=0=\varphi(p/\Theta)\wedge\varphi(0/\Theta)$, and $\varphi(p/\Theta\wedge 1/\Theta)=\varphi(p/\Theta)=\varphi(p/\Theta)\wedge\varphi(1/\Theta)$. If $p,q\in S$, and $p\wedge q\notin S$, then $p\wedge q\subseteqq p'$ or $p\wedge q\subseteqq q'$, so that $\nu(p\wedge q)\leqq\nu(p)'$ or $\nu(q)'$, which implies that $\nu(p\wedge q)=\nu(p)\wedge\nu(q)=0$. Hence, $\varphi(p/\Theta)\wedge\varphi(q/\Theta)=\nu(p)\wedge\nu(q)=0=\varphi(0/\Theta)=\varphi(p/\Theta\wedge q/\Theta)$. Further, $\varphi(p'/\Theta)=\nu(p')=\nu(p)'=\varphi(p/\Theta)'$ if $p\in S-\{0,1\}$, and $\varphi(1/\Theta)=\varphi(0/\Theta)'$. We see that $\varphi\colon S/\Theta\to A$ is a homomorphism. Finally, for $p\in L_i$, $p\neq 0_i, 1_i$, we have $\varphi(p/\Theta)=\varphi(\psi_i(p))=\nu(p)=\varphi_i(p)$, $\varphi(\psi_i(0_i))=\varphi(0/\Theta)=0=\varphi_i(0_i)$, $\varphi(\psi_i(1_i))=\varphi(1/\Theta)=1=\varphi_i(1_i)$, so that $\varphi\circ\psi_i=\varphi_i$. This completes the proof.

References

[1] G. Bruns, Free ortholattices, *Canad. J. Math.*, **28** (1976), 977—985.

[2] G. Grätzer, Free products and reduced free products of lattices, in: *Proc. Univ. Houston Lattice Theory Conf.* (1973), Dept. Math., Univ. of Houston (Houston, 1973); pp. 539—563.

MATEMATICKÝ ÚSTAV SAV
OBRANCOV MIERU 49
814 73 BRATISLAVA, CZECHOSLOVAKIA

Acta Sci. Math., 52 (1988), 53—67

Abstract spectral theory. II: Minimal characters and minimal spectrums of multiplicative lattices

N. K. THAKARE, C. S. MANJAREKAR and S. MAEDA

1. Introduction

A multiplicative lattice is a complete lattice in which there is defined a commutative, associative multiplication which distributes over arbitrary joins (i.e., $a(\bigvee_\alpha b_\alpha)=$ $=\bigvee_\alpha ab_\alpha$), $ab\leqq a\wedge b$ and the greatest element 1 acts as a multiplicative identity. Throughout this paper, let L denote a multiplicative lattice. In L an element p different from 1 is called prime if $ab\leqq p$ implies $a\leqq p$ or $b\leqq p$. A minimal element in the set of prime elements of L will be called a minimal prime element of L. A character of L is a homomorphism of L onto a two element chain $\mathbf{C}_2$. It was shown in [9] that an element a of L is prime if and only if there is a homomorphism φ of L onto $\mathbf{C}_2$ with $a=\bigvee\{x\colon \varphi(x)=0\}$. This means that a prime element of L can now be equivalently associated with a character of L, and so a prime element itself will be called a character of L. We denote by $\sigma(L)$ and $\pi(L)$ the sets of characters and minimal characters of L respectively.

This work is a continuation of the work initiated by THAKARE and MANJAREKAR [9]. Here we are concerned mainly with minimal characters of L and with the topology on the set $\pi(L)$ which is the restriction of the hull kernel topology introduced on the set $\sigma(L)$ (see [9]).

The studies of minimal prime ideals for commutative rings, commutative semigroups, distributive lattices, lattice ordered groups, f-rings and recently 0-distributive semilattices (THAKARE and PAWAR [11], [7]) have been carried out extensively. An attempt to unify these scattered studies was nicely made by KEIMEL [4]. Our study in this paper is close in spirit to the study [4], though however we carry out investigations to include many more novel notions the motivation for which stems from the desire to abstract available notions in commutative rings on the lines of DILWORTH [2].

Received December 27, 1984.

The notion a^* of an element a of L is defined as the join of annihilators of powers of a, and this concept plays an important role in the investigations of minimal characters in Sections 2 and 5. The concept of minimal characters belonging to an element, appeared in MURATA [5] and ANDERSON [1], is discussed in Section 3. We abstract the notion of an ideal B of a commutative ring R that is related to an ideal A of R, and this concept is used in the arguments on primary decompositions of elements of L in Section 4.

In the previous paper [9], we assumed that L always satisfies the following condition which is equivalent to the ascending chain condition:

(K) *Every element of L is compact.*

In this paper, we assume that condition (K) or some weaker ones according to the need.

We remark that for any $p\in\sigma(L)$, the existence of a maximal character q with $p\leqq q$ can be proved under the assumption that L satisfies (K) (see [9]) but the existence of a minimal character r with $r\leqq p$ can be proved without this assumption (because, if Q is a chain of characters then $p=\wedge Q$ is also a character).

2. Characters and minimal characters

A subset S of L is called *multiplicatively closed* if $a, b\in S$ implies $ab\in S$, and S is called *submultiplicatively closed* if for $a, b\in S$ there exists $c\in S$ with $c\leqq ab$. Without assuming the condition (K), the Separation Lemma can be stated as follows (cf. [9], Lemma 2.2):

Separation Lemma. *Let S be a submultiplicatively closed subset of L, and assume that every element of S is compact. If $S\cap[0, a]=\emptyset$ for some $a\in L$, then there exists a character p of L which is a maximal element of the set $\{x\in L\colon a\leqq x$ and $S\cap[0, x]=\emptyset\}$.*

In fact, this set has a maximal element p by Zorn's lemma since every element of S is compact, and we can prove that p is a character since S is submultiplicatively closed.

An element a of L is called *M-compact* if a^n are compact for infinitely many integer n. Every nilpotent element is M-compact. An idempotent is M-compact if and if it is compact.

Proposition 2.1. *If a is an M-compact element of L and if $a^n\nleqq b$ for every integer n, then there exists $p\in\sigma(L)$ such that $b\leqq p$ and $a\nleqq p$. Especially, if a is M-compact and is not nilpotent then there exists $p\in\sigma(L)$ such that $a\nleqq p$.*

Proof. The set $S=\{a^n\colon a^n \text{ is compact}\}$ is submultiplicatively closed and $S\cap[0,b]=\emptyset$. Hence, by the Separation Lemma there is $p\in\sigma(L)$ such that $b\leqq p$ and $S\cap[0,p]=\emptyset$. Then, $a\nleqq p$.

Corollary 2.2. *If the greatest element* 1 *of* L *is compact, then for any* $b\in L$ *with* $b<1$ *there exists* $p\in\sigma(L)$ *such that* $b\leqq p$.

Proof. Put $a=1$ in Proposition 2.1.

We need to introduce the following notation which is important in the arguments on minimal characters. For $a\in L$,

$$a^* = \bigvee\{x\in L\colon a^n x = 0 \text{ for some integer } n\}.$$

Evidently, $0^*=1$, $1^*=0$, and $a\leqq b$ implies $b^*\leqq a^*$.

Lemma 2.3. (i) *If* a^* *is compact, then* $a^n a^*=0$ *for some* n, *and* $a\wedge a^*$ *is nilpotent.*

(ii) *In the case that* 1 *is compact,* $a\in L$ *is nilpotent if and only if* $a^*=1$.

Proof. (i) The set $S=\{x\in L\colon a^n x=0 \text{ for some } n\}$ is an ideal, since $a^{m+n}(x\vee y)\leqq a^m x\vee a^n y$. Hence, if a^* is compact then $a^*\in S$. Thus, $a^n a^*=0$ for some n, and $(a\wedge a^*)^{n+1}=0$.

(ii) If $a^*=1$ then a is nilpotent by (i). The converse is evident.

Lemma 2.4. *Let* $a\in L$ *and* $p\in\sigma(L)$. $a\nleqq p$ *implies* $a^*\leqq p$. *(Hence,* $a\wedge a^*\leqq p$ *always.)*

Proof. Assume $a\nleqq p$. If $a^n x=0$ then we have $a^n x\leqq p$ and $a^n\nleqq p$. Hence, $x\leqq p$. Therefore, $a^*\leqq p$.

Using the condition (K), we now get a fundamental result with some interesting corollaries.

Theorem 2.5. *Assume that* L *satisfies* (K). *For* $a\in L$ *and* $p\in\sigma(L)$ *the following statements are equivalent:*

(1) $a^*\leqq p$;

(2) *there is some* $q\in\pi(L)$ *with* $q\leqq p$ *and* $a\nleqq q$.

Proof. (1)⇒(2): Let $S=\{a^n x\colon x\nleqq p,\ n=1,2,\dots\}$. Then S is multiplicatively closed. We have $0\notin S$, because if $a^n x=0$ then $x\leqq a^*\leqq p$ by (1). By the Separation Lemma there exists $r\in\sigma(L)$ such that $S\cap[0,r]=\emptyset$. Take $q\in\pi(L)$ such that $q\leqq r$. We have $r\leqq p$, since otherwise $ar\in S\cap[0,r]$, a contradiction. Also, $a\nleqq r$, since $a\in S$. Hence, $q\leqq p$ and $a\nleqq q$.

(2)⇒(1): If $q\leqq p$ and $a\nleqq q$, then $a^*\leqq q\leqq p$ by Lemma 2.4.

Corollary 2.6. *Assume that L satisfies* (K), *and let* $p \in \sigma(L)$. *If* $p^* \leqq p$ *then* p *is not minimal.*

Proof. If $p^* \leqq p$, there is $q \in \pi(L)$ with $q \leqq p$ and $p \nleqq q$ by Theorem 2.5. Thus, $q < p$, and p is not minimal.

As stated in the previous paper [9], the *hull kernel topology* on $\sigma(L)$ is given as follows. For $a \in L$ we put

$$V(a) = \{p \in \sigma(L): a \leqq p\}.$$

Since $V(0) = \sigma(L)$, $V(1) = \emptyset$, $V(a) \cup V(b) = V(ab)$ $(= V(a \wedge b))$ and $\cap_\alpha V(a_\alpha) = = V(\vee_\alpha a_\alpha)$, we obtain a topology on $\sigma(L)$ such that $\{V(a): a \in L\}$ is the family of all closed sets. It is easy to verify that the closure $\bar{R}$ of a subset R of $\sigma(L)$ coincides with $V(\wedge R)$.

Corollary 2.7. *Assume that L satisfies* (K), *and let* $a \in L$. $V(a^*)$ *is equal to the closure of the open set* $\sigma(L) - V(a)$.

Proof. By Lemma 2.4, we have $\sigma(L) - V(a) \subset V(a^*)$. Hence, it suffices to show that if $\sigma(L) - V(a) \subset V(x)$ then $V(a^*) \subset V(x)$. Let $p \in V(a^*)$. By Theorem 2.5 there is $q \in \pi(L)$ with $q \leqq p$ and $a \nleqq q$. Then, $q \in \sigma(L) - V(a) \subset V(x)$, and hence $x \leqq q \leqq p$. Hence $p \in V(x)$, and we obtain $V(a^*) \subset V(x)$.

The concept of regular characters was introduced by [3], [8] and [9], while its dual concept, coregular characters, appeared in [8] for bounded distributive lattices. A character $r \in \sigma(L)$ is called *coregular* if for $p, q \in \sigma(L)$, $r \leqq p$ and $q \leqq p$ together imply $r \leqq q$. The companion of Theorem 2.7 of [9] would now be proved.

Theorem 2.8. *Assume that L satisfies* (K). *For* $r \in \sigma(L)$ *the following five statements are equivalent:*

(1) r *is coregular;*

(2) *the set* $V(r)$ *is open;*

(3) $V(r) \cap V(r^*) = \emptyset$;

(4) $r \vee r^* = 1$;

(5) *there is* $x \in L$ *such that* $x \vee r = 1$ *and* $r^n x = 0$ *for some integer n.*

(We remark that (5)⇒(4)⇒(3)⇒(2)⇒(1) can be proved without the assumption (K).)

Proof. (5)⇒(4) is evident. (4)⇒(3): If $V(r) \cap V(r^*)$ had an element p then $r \vee r^* \leqq p < 1$, contradicting (4). (3)⇒(2): We have $V(r) \cup V(r^*) = \sigma(L)$ by Lemma 2.4. Hence, by (3) we have $V(r) = \sigma(L) - V(r^*)$, and then $V(r)$ is an open set. (2)⇒ (1): Let $r \leqq p$ and $q \leqq p$, and put $G = \sigma(L) - V(r)$. Since G is closed by (2), we have

$p\notin G=\bar{G}=V(\wedge G)$, and hence $\wedge G\not\leqq p$. As $q\leqq p$, we have $\wedge G\not\leqq q$, whence $q\notin G$. Hence, $r\leqq q$.

Next, we assume that L satifies (K). (4) implies (5), since $r^n r^*=0$ for some n by Lemma 2.3 (i). (1)⇒(4): If $r\vee r^*<1$, then there is $p\in\sigma(L)$ with $r\vee r^*\leqq p$ by Corollary 2.2. By Theorem 2.5 there is $q\in\pi(L)$ with $q\leqq p$ and $r\not\leqq q$, contradicting (1).

Recall the concept of multiplicative normal (i.e. M-normal) lattice introduced in [9]. A multiplicative lattice L is called *M-normal* if each character of L contains a unique minimal character of L. We shall have several characterizations of M-normal multiplicative lattices in the following two theorems.

Theorem 2.9. *The following two statements are equivalent:*

(1) *L is M-normal;*

(2) *every minimal character of L is coregular.*

If 1 *is compact,* (1) *is also equivalent to the following statement:*

(3) *$q_1\vee q_2=1$ for any distinct minimal characters q_1, q_2 of L.*

Proof. (1)⇒(2): Let $r\in\pi(L)$, and we take $p, q\in\sigma(L)$ with $r\leqq p$ and $q\leqq p$. There is $q'\in\pi(L)$ with $q'\leqq q$. Then, $r, q'\leqq p$, and hence $r=q'\leqq q$ by (1). Hence, r is coregular. (2)⇒(1): Let $p\in\sigma(L)$, $r_i\in\pi(L)$ $(i=1, 2)$ and $r_i\leqq p$. Since r_1 is coregular by (2), we have $r_1\leqq r_2$. Similarly we have $r_2\leqq r_1$, and hence $r_1=r_2$.

(1)⇒(3): Assume that 1 is compact. If $q_1\vee q_2<1$, then there is $p\in\sigma(L)$ with $q_1\vee q_2\leqq p$ by Corollary 2.2, and hence $q_1=q_2$ by (1). (3)⇒(1) is evident.

Recall that a topological space is called *extremally disconnected* if the closure of each open set is open.

Lemma 2.10. *A topological space X is extremally disconnected if and only if for open subsets G_1, G_2 of X, $G_1\cap G_2=\emptyset$ implies $\bar{G}_1\cap\bar{G}_2=\emptyset$.*

Proof. Assume that X is extremally disconnected. If $G_1\cap G_2=\emptyset$, then $\bar{G}_1\subset X-G_2$, since $X-G_2$ is closed. Hence, $G_2\subset X-\bar{G}_1$, Since $\bar{G}_1$ is open, we have $\bar{G}_2\subset X-\bar{G}_1$, and then $\bar{G}_1\cap\bar{G}_2=\emptyset$.

Next we shall prove the converse. For an open set G, we put $U=X-\bar{G}$. Then, U is open and $U\cap G=\emptyset$, and hence $\bar{U}\cap\bar{G}=\emptyset$. Hence, $\bar{U}\subset X-\bar{G}=U$, which implies that U is closed. Hence, $\bar{G}$ is open.

Theorem 2.11. *Assume that L satisfies* (K). *The following five statements are equivalent:*

(1) *L is M-normal;*

(2) *if G_1 and G_2 are open sets of $\sigma(L)$ with $G_1\cap G_2=\emptyset$ then $\bar{G}_1\cap\bar{G}_2=\emptyset$;*

(3) *$\sigma(L)$ is extremally disconnected;*

(4) $V(a^*)$ *is open for every* $a\in L$;

(5) *if* $V(a)\cup V(b)=\sigma(L)$ *then* $a^*\vee b^*=1$.

Proof. The equivalences (2)⇔(3) and (3)⇔(4) immediately follow from Lemma 2.10 and Corollary 2.7, respectively.

(1)⇒(2): Let G_1 and G_2 be open sets with $G_1\cap G_2=\emptyset$. We can put $G_i=\sigma(L)-V(a_i)$ for some $a_i\in L$ $(i=1,2)$. By Corollary 2.7, we have $\bar{G}_i=V(a_i^*)$. If $\bar{G}_1\cap\bar{G}_2$ had an element p, then $a_i^*\leqq p$ and by Theorem 2.5 there would exist $q_1,q_2\in\pi(L)$ with $q_i\leqq p$ and $a_i\not\leqq q_i$. By (1), we have $q_1=q_2$, which implies $q_1\in G_1\cap G_2$, a contradiction.

(2)⇒(5): Let $V(a)\cup V(b)=\sigma(L)$. Putting $G_1=\sigma(L)-V(a)$ and $G_2=\sigma(L)-V(b)$, we have $G_1\cap G_2=\emptyset$. By (2) we have $V(a^*)\cap V(b^*)=\bar{G}_1\cap\bar{G}_2=\emptyset$. Hence, $a^*\vee b^*=1$ by Corollary 2.2.

(5)⇒(1): Let $q_1,q_2\in\pi(L)$ with $q_1\neq q_2$, and we shall show $V(q_1^*)\cup V(q_2^*)=\sigma(L)$. For any $p\in\sigma(L)$, there is $q\in\pi(L)$ with $q\leqq p$. If $q\neq q_1$, then since $q_1\not\leqq q$ we have $q_1^*\leqq q\leqq p$ by Lemma 2.4. If $q=q_1$, then $q\neq q_2$ and hence we have $q_2^*\leqq p$. Thus, we get $V(q_1^*)\cup V(q_2^*)=\sigma(L)$, and then $q_1^{**}\vee q_2^{**}=1$ by (5). Since $q_i^*\not\leqq q_i$ by Corollary 2.6, we get $q_i^{**}\leqq q_i$ by Lemma 2.4. Hence, $q_1\vee q_2=1$, and there is no character which contains both q_1 and q_2.

3. Minimal characters belonging to an element

We consider a relation between characters and multiplicatively closed subsets. For $a\in L$, we put

$$C(a) = \{x\in L:\ x\not\leqq a\}.$$

(This notion was introduced in Nemitz [6].) The set of all multiplicatively closed subsets of L is denoted by $\mathcal{M}(L)$.

Lemma 3.1. $C(p)\in\mathcal{M}(L)$ *if and only if p is a character of L. The mapping* $p\mapsto C(p)$ *of* $\sigma(L)$ *into* $\mathcal{M}(L)$ *is one-to-one, and* $p\leqq q\Leftrightarrow C(p)\supset C(q)$.

Proof. Evident.

Lemma 3.2. *Let* $a\in L$, *and take* $M\in\mathcal{M}(L)$ *with* $M\cap[0,a]=\emptyset$.

(i) $\mathcal{U}=\{N\in\mathcal{M}(L):\ N\supset M$ *and* $N\cap[0,a]=\emptyset\}$ *has a maximal element.*

(ii) $N^*\in\mathcal{U}$ *is maximal in* $\mathcal{U}$ *if and only if for any* $x\in L$ *with* $x\notin N^*$ *there exists* $y\in N^*$ *such that* $x^ny\leqq a$ *for some integer* n.

Proof. (i) For any chain $\mathcal{V}\subset\mathcal{U}$, the union $\bigcup\{N:\ N\in\mathcal{V}\}$ belongs to $\mathcal{U}$. Hence, $\mathcal{U}$ has a maximal element by Zorn's lemma.

(ii) Let N^* be maximal and let $x \notin N^*$. The set $N_1 = \{x^n, y, x^n y\colon y \in N^*, n = 1, 2, \ldots\}$ is multiplicatively closed since $N^* \in \mathcal{M}(L)$, and $N_1 \supset N^*$. Moreover, $N_1 \neq N^*$, for $x \in N_1$ and $x \notin N^*$. Hence, by the maximality of N^* we have $N_1 \cap \cap [0, a] = \emptyset$. Then, there exists $y \in N^*$ such that $x^n y \leqq a$ for some n.

Next, take $N \in \mathcal{M}(L)$ with $N \supsetneqq N^*$, and take $x \in N - N^*$. If N^* satisfies the given condition, there exists $y \in N^*$ such that $x^n y \leqq a$ for some n. Then, $x^n y \in N \cap \cap [0, a]$. Hence, N^* is maximal in $\mathcal{U}$.

Recall the concept of minimal characters belonging to an element, which was initiated by Murata [5]. For $a \in L$ with $a < 1$, a minimal element of $V(a) = \{p \in \sigma(L)\colon a \leqq p\}$ is called a *minimal character belonging to a*. The set of all minimal characters belonging to a is denoted by $V_{\min}(a)$. For any chain Q in $V(a)$, we have $\wedge Q \in V(a)$. Hence, for any $p \in V(a)$ there is $q \in V_{\min}(a)$ with $q \leqq p$ by Zorn's lemma. We remark that $V_{\min}(0) = \pi(L)$.

Theorem 3.3. *Let $a \in L$ with $a < 1$ and let $p \in \sigma(L)$. If L satisfies* (K) *then the following statements are equivalent:*

(1) $p \in V_{\min}(a)$;

(2) *$C(p)$ is maximal in the set* $\{N \in \mathcal{M}(L)\colon N \cap [0, a] = \emptyset\}$;

(3) *$a \leqq p$ and there exists $x \in L$ such that $x \not\leqq p$ and $p^n x \leqq a$ for some integer n.*

Moreover, without assuming (K), *the statements* (2) *and* (3) *are equivalent, and* (2) *implies* (1).

Proof. (2)$\Leftrightarrow$(3): Putting $M = \{1\}$ in Lemma 3.2, (2) is equivalent to the following statement: "$a \leqq p$ and for any $x \leqq p$ there is $y \not\leqq p$ such that $x^n y \leqq a$ for some n". Evidently, this is equivalent to (3).

(2)$\Rightarrow$(1): If $a \leqq q \leqq p$ with $q \in \sigma(L)$, then $C(q) \in \mathcal{M}(L)$, $C(q) \cap [0, a] = \emptyset$ and $C(q) \supset C(p)$. Hence, $C(q) = C(p)$ by (2), and then $q = p$.

We assume (K) and prove (1)$\Rightarrow$(2). Put $\mathcal{U} = \{N \in \mathcal{M}(L)\colon N \cap [0, a] = \emptyset\}$. $C(p) \in \mathcal{U}$ by $a \leqq p$. If $C(p) \subset N \in \mathcal{U}$, then $N \cap [0, a] = \emptyset$, and by the Separation Lemma there is $q \in \sigma(L)$ with $a \leqq q$ and $N \cap [0, q] = \emptyset$. Then, $C(q) \supset N \supset C(p)$, and hence $p \geqq q$. Hence, $p = q$ by (1), and then $C(p) = N$. Thus, $C(p)$ is maximal in $\mathcal{U}$.

Theorem 3.4. *Let $a \in L$ with $a < 1$. If every finite product of elements of $V_{\min}(a)$ is compact (especially, if L satisfies* (K)*), then $V_{\min}(a)$ is a finite set.*

Proof. Assume that $V_{\min}(a)$ is an infinite set. The set M of all finite products of elements of $V_{\min}(a)$ is multiplicatively closed. If $b \in M$, then $b = p_1 \ldots p_n$ with $p_i \in V_{\min}(a)$, and by the assumption there is $q \in V_{\min}(a)$ which is different from all p_i. We have $b \not\leqq q$ since $p_i \not\leqq q$ for all i, and then $b \not\leqq a$. Thus, we have $M \cap [0, a] = \emptyset$. By the Separation Lemma there is $r \in \sigma(L)$ with $a \leqq r$ and $M \cap [0, r] = \emptyset$. But, we can take $r_0 \in V_{\min}(a)$ with $r_0 \leqq r$, and then $r_0 \in M \cap [0, r]$, a contradiction.

The concept of radicals is a classical notion of commutative ring theory and its abstract formulation has been attempted long back and is scattered in several papers in various forms (see for example MURATA [5] and ANDERSON [1]). Let us recall this concept in abstract form. The *radical* of an element $a \in L$, denoted by $\sqrt{a}$, is defined by

$$\sqrt{a} = \vee\{x \in L\colon x^n \leqq a \text{ for some integer } n\}.$$

Evidently, $a \leqq \sqrt{a}$ for any $a \in L$, and $p = \sqrt{p}$ if $p \in \sigma(L)$. Hence, we have $V(\sqrt{a}) = V(a)$.

Lemma 3.5. (i) *If $\sqrt{a}$ is compact then $\sqrt{a}^n \leqq a$ for some integer n.*

(ii) *If $\sqrt{a}$ and $\sqrt{b}$ are compact then $\sqrt{ab} = \sqrt{a \wedge b} = \sqrt{a} \wedge \sqrt{b}$.*

(iii) *If 1 is compact, then $a < 1$ implies $\sqrt{a} < 1$.*

Proof. (i) The set $S = \{x \in L\colon x^n \leqq a$ for some $n\}$ is an ideal, for $(x \vee y)^{m+n} \leqq \leqq x^m \vee y^n$. Hence, if $\sqrt{a}$ is compact then $\sqrt{a} \in S$.

(ii) Evidently, $\sqrt{ab} \leqq \sqrt{a \wedge b} \leqq \sqrt{a} \wedge \sqrt{b}$. By (i), $\sqrt{a}^m \leqq a$, $\sqrt{b}^n \leqq b$ for some m, n. Then, $(\sqrt{a} \wedge \sqrt{b})^{m+n} = (\sqrt{a} \wedge \sqrt{b})^m (\sqrt{a} \wedge \sqrt{b})^n \leqq \sqrt{a}^m \sqrt{b}^n \leqq ab$. Hence, $\sqrt{a} \wedge \sqrt{b} \leqq \sqrt{ab}$.

(iii) By Corollary 2.2, there is $p \in \sigma(L)$ with $a \leqq p$. Then, $\sqrt{a} \leqq \sqrt{p} = p < 1$.

Theorem 3.6. *Assume that L is generated by M-compact elements, that is, every element of L is a join of M-compact elements. For $a \in L$ with $a < 1$,*

$$\sqrt{a} = \wedge\{p\colon p \in V_{\min}(a)\} = \wedge\{p\colon p \in V(a)\}.$$

Proof. Evidently, $\wedge V_{\min}(a) = \wedge V(a)$, and $\sqrt{a} \leqq \wedge V(\sqrt{a}) = \wedge V(a)$. If $\sqrt{a} < < \wedge V(a)$, there would exist an M-compact element x such that $x \leqq \wedge V(a)$ and $x \nleqq \sqrt{a}$. Then, $x^n \nleqq a$ for every n, and by Proposition 2.1 there is $p \in \sigma(L)$ with $a \leqq p$ and $x \nleqq p$. This contradicts $x \leqq \wedge V(a)$.

Corollary 3.7. *Assume that L is generated by M-compact elements, and let $a \in L$ with $a < 1$. $V_{\min}(a)$ contains only one element if and only if $\sqrt{a}$ is a character.*

Proof. The "only if" part follows from the theorem, and the converse is evident.

We remark that the r-lattice introduced in [1] satisfies the assumption of this theorem, because any compact element of an r-lattice is M-compact by Theorem 2.1 of [1].

4. Related elements and associated characters of primary elements

We now take up a notion of one more related concept which is found in ring theory. The notion so far has not been pulled down to lattice theory nor has been abstracted in the sense of DILWORTH [2].

Let $a\in L$ with $a<1$. An element $b\in L$ is said to be *related to* a if there exists $x\in L$ such that $x\nleqq a$ and $bx\leqq a$. If b is related to a and $b'\leqq b$ then evidently b' is related to a. Hence, the set of all elements of L which are related to a is multiplicatively closed. Next, let $p\in\sigma(L)$. Evidently, b is related to p if and only if $b\leqq p$. Hence, the set of all elements of L which are unrelated to p coincides with $C(p)$ and hence it is multiplicatively closed.

Lemma 4.1. *Let $a\in L$ with $a<1$, and let $b\in L$.*

(i) *If $a=a_1\wedge\ldots\wedge a_n$ $(a_i<1)$ and if b is related to a, then b is related to a_i for some i.*

(ii) *If there exists $x\in L$ such that $x\nleqq a$ and $b^n x\leqq a$ for some integer n, then b is related to a.*

(iii) *Assume that $\sqrt{a}$ is compact. If b is related to $\sqrt{a}$ then b is related to a. Especially, $\sqrt{a}$ is related to a.*

Proof. (i) is evident.

(ii) If $x\nleqq a$ and $b^n x\leqq a$, then taking the smallest integer i such that $b^i x\leqq a$, we have $b^{i-1}x\nleqq a$ and $b(b^{i-1}x)\leqq a$ $(b^0=1)$. Hence, b is related to a.

(iii) By Lemma 3.5 (i), $\sqrt{a}^n\leqq a$ for some n. If $x\nleqq\sqrt{a}$ and $bx\leqq\sqrt{a}$, then $x^n\nleqq a$ and $b^n x^n\leqq\sqrt{a}^n\leqq a$. Hence, b is related to a by (ii).

Theorem 4.2. *Assume that L satisfies* (K), *and let $a\in L$ with $a<1$. Every minimal character p belonging to a is related to a.*

Proof. By Theorem 3.3, there is $x\in L$ such that $x\nleqq p$ and $p^n x\leqq a$ for some n. Then, we have $x\nleqq a$, for $a\leqq p$. Hence, p is related to a by Lemma 4.1. (ii).

Following DILWORTH [2], an element $q\in L$ with $q<1$ is called *primary* if $xy\leqq q$ implies $x\leqq q$ or $y^n\leqq q$ for some integer n.

Lemma 4.3. *If $q\in L$ is primary and if $\sqrt{q}$ is compact, then $\sqrt{q}\in\sigma(L)$ and $V_{\min}(q)=\{\sqrt{q}\}$. Moreover, $b\in L$ is related to q if and only if $b\leqq\sqrt{q}$.*

Proof. This can be proved by using the fact: $\sqrt{q}^n\leqq q$ for some n, and the details are omitted.

Hereafter in this section, we assume that

($*$) *For every primary element q of L the element $\sqrt{q}$ is compact.*

By this assumption, we have $\sqrt{q}^n \leqq q$ for some integer n, and $\sqrt{q}$ is the least element of $V(q)$. We call $\sqrt{q}$ the *character associated with* q.

As stated in [2], we have the following lemma (the proof is omitted).

Lemma 4.4. *If q_1, q_2 are primary elements associated with the same character p, then $q_1 \wedge q_2$ is also a primary element with the same associated character p.*

Following [2], an element $a \in L$ is said to have an *irredundant* (or normal) *primary decomposition,* if $a = q_1 \wedge \ldots \wedge q_m$ for some primary elements $q_1, \ldots, q_m$ and if this expression cannot be reduced further. Then, by Lemma 4.4, $q_1, \ldots, q_m$ are associated with distinct characters.

Remark 4.5. If $a \in L$ has an irredundant primary decomposition $a = q_1 \wedge \ldots \ldots \wedge q_m$ $(m \geqq 2)$, then a is not primary. This fact can be proved by the same way as [5], Lemma 7, since $\sqrt{a} = \sqrt{q_1} \wedge \ldots \wedge \sqrt{q_m}$ by Lemma 3.5 (ii).

Lemma 4.6. *Let $a \in L$ have an irredundant primary decomposition $a = q_1 \wedge \ldots \ldots \wedge q_m$ and put $p_i = \sqrt{q_i}$ $(p_i \in \sigma(L))$.*

(i) *For $p \in \sigma(L)$, $a \leqq p$ if and only if $p_i \leqq p$ for some i.*

(ii) *An element $c \in L$ is related to a if and only if $c \leqq p_i$ for some i.*

Proof. (i) Let $a \leqq p$. We have $p_i^{n_i} \leqq q_i$ for some integer n_i. Put $b = \prod_{i=1}^{m} p_i^{n_i}$. Since $b \leqq q_i$ for every i, we have $b \leqq a \leqq p$. Then, $p_i \leqq p$ for some i, since p is a character. The converse is evident.

(ii) If c is related to a, then $c \leqq \sqrt{q_i} = p_i$ for some i by Lemma 4.1 (i) and Lemma 4.3. Conversely, let $c \leqq p_i$ for some i. Putting $b = \bigwedge_{j \neq i} q_j$, we have $b > a$ since the decompositon is irredundant. Since $p_i^n \leqq q_i$ for some n, we have $c^n b \leqq q_i b \leqq \bigwedge_{j=1}^{m} q_j = a$. Hence, c is related to a by Lemma 4.1 (ii).

Theorem 4.7. *Let $a \in L$ have an irredundant primary decomposition $a = q_1 \wedge \ldots \ldots \wedge q_m$ and put $p_i = \sqrt{q_i}$. The set of all minimal elements of $\{p_1, \ldots, p_m\}$ coincides with $V_{\min}(a)$. The set of all maximal elements of $\{p_1, \ldots, p_m\}$ coincides with the set of all maximal elements of the set $\{x \in L\colon x$ is related to $a\}$.*

Proof. These statements immediately follow from Lemma 4.6.

Corollary 4.8. *If $a \in L$ has an irredundant primary decomposition, then every maximal element among all the elements related to a is a character containing a.*

For $a \in L$ and $p \in \sigma(L)$, we put

$$a(p) = \vee \{x \in L\colon xy \leqq a \text{ for some } y \nleqq p\}.$$

We now set ourselves to describe the elements $a(p)$.

Lemma 4.9. *If* $a \leqq p$ *then* $a \leqq a(p) \leqq p$. *If* $a \not\leqq p$ *then* $a(p)=1$.

Proof. Let $a \leqq p$. If $xy \leqq a$ and $y \not\leqq p$, then we have $x \leqq p$, since $xy \leqq p$. Hence, $a(p) \leqq p$. Moreover, $a \leqq a(p)$, since $a1 \leqq a$ and $1 \not\leqq p$. Next, $a \not\leqq p$ implies $a(p)=1$, since $1a \leqq a$.

Lemma 4.10. *Let* $a \in L$ *have an irredundant primary decomposition* $a=q_1 \wedge \ldots \wedge q_m$ *and put* $p_i=\sqrt{q_i}$. *For* $p \in \sigma(L)$, *if we put* $I(p)=\{i: p_i \leqq p\}$, *then* $a(p)=\wedge\{q_i: i \in I(p)\}$. ($a(p)=1$ *if* $I(p)=\emptyset$.)

Proof. Let $i \in I(p)$. If $xy \leqq a$ and $y \not\leqq p$, then since $p_i \leqq p$, we have $y \not\leqq p_i=\sqrt{q_i}$, and hence $y^n \not\leqq q_i$ for every n. Since $xy \leqq q_i$, we have $x \leqq q_i$. Thus, $a(p) \leqq q_i$. Put $b=\wedge\{q_i: i \in I(p)\}$. As above we get $a(p) \leqq b$. Next, since $p_j^{n_j} \leqq q_j$ for some n_j, we put $c=\prod\{p_j^{n_j}: j \notin I(p)\}$. Then, $c \not\leqq p$, since $p_j \not\leqq p$ for every $j \in I(p)$. We have $c \leqq \wedge\{q_j: j \notin I(p)\}$, and hence $bc \leqq a$. Therefore, $b \leqq a(p)$. (If $I(p)=\emptyset$ then we may put $b=1$.)

Theorem 4.11. *Let* $a \in L$ *have an irredundant primary decomposition* $a=q_1 \wedge \ldots \wedge q_m$ *and put* $p_i=\sqrt{q_i}$. *For* $p \in \sigma(L)$, $p=p_i$ *for some i if and only if* $a(p)<1$ *and p is maximal among all the elements related to* $a(p)$.

Proof. Let $p=p_k$ and put $I=\{i: p_i \leqq p_k\}$ ($I \neq \emptyset$, since $k \in I$). By Lemma 4.10, $a(p_k)$ has an irredundant primary decomposition $a(p_k)=\wedge\{q_i: i \in I\}$. Since p_k is maximal in $\{p_i: i \in I\}$, p_k is maximal in $\{x \in L: x \text{ is related to } a(p)\}$ by Theorem 4.7.

Conversely, if $a(p)<1$, then $I(p)=\{i: p_i \leqq p\}$ is non-empty and $a(p)$ has an irredundant primary decomposition $a(p)=\wedge\{q_i: i \in I(p)\}$. If p is maximal among the elements related to $a(p)$, then p coincides with a maximal element of $\{p_i: i \in I(p)\}$.

Corollary 4.12. *Any two irredundant primary decompositions of an element* $a \in L$ *have the same number of components and the same set of associated characters.*

5. Minimal spectrum

First we shall introduce a new concept. A character $p \in \sigma(L)$ is called *purely minimal* if $C(p)$ is maximal in the set $\{M \in \mathcal{M}(L): 0 \notin M\}$. It follows from Lemma 3.1 that any purely minimal character is minimal. The set of all purely minimal characters is denoted by $\pi_0(L)$. This is a subset of $\pi(L)$.

Theorem 5.1. (i) *For* $p \in \sigma(L)$ *the following four statements are equivalent:* (1) *p is purely minimal;*

(2) *there exists* $x \in L$ *such that* $x \not\leqq p$ *and* $p^n x = 0$ *for some integer n:*

(3) $p^* \not\leqq p$;

(4) *for any* $x \in L$, *p contains precisely one of x and* x^*;

(ii) *if L satisfies* (K), *then any minimal character is purely minimal, that is,* $\pi_0(L) = \pi(L)$.

Proof. The equivalence of (1) and (2) follows from Theorem 3.3 by putting $a=0$. The statement (ii) also follows from Theorem 3.3. The equivalence (2)$\Leftrightarrow$(3) and the implication (4)$\Rightarrow$(3) are evident. (3)$\Rightarrow$(4): If $x \not\leqq p$, then $x^* \leqq p$ by Lemma 2.4. If $x \leqq p$, then $p^* \leqq x^*$, and hence $x^* \not\leqq p$ by (3).

Corollary 5.2. *If* $p \in \sigma(L)$ *is purely minimal then* $p^{**} \leqq p$, *and* $x \leqq p$ *implies* $x^{**} \leqq p$.

Proof. Since $p^* \not\leqq p$ by Theorem 5.1, we have $p^{**} \leqq p$ by Lemma 2.4. If $x \leqq p$, then we have $x^* \geqq p^*$, and hence $x^{**} \leqq p^{**} \leqq p$.

The hull kernel topology on $\pi(L)$ is the induced topology of the hull kernel topology on $\sigma(L)$. $\pi(L)$ with this topology will be called the *minimal spectrum* of L. For any $a \in L$, the set $h(a) = \{p \in \pi(L): a \leqq p\}$ is called the *hull* of a. For any subset R of $\pi(L)$, the element $K(R) = \wedge \{p: p \in R\}$ is called the *kernel* of R. Then, a subset R of $\pi(L)$ is closed if and only if $R = h(a)$ for some $a \in L$. Evidently, $a \leqq K(h(a))$ for every $a \in L$, and for every $R \subset \pi(L)$, $h(K(R))$ is equal to the closure of R.

Now we get an important topological property of purely minimal characters.

Theorem 5.3. *If* $p \in \pi(L)$ *is purely minimal then p is an isolated point of* $\pi(L)$.

Proof. Put $G = \pi(L) - h(p^*)$. G is an open set, and $p \in G$ since $p^* \not\leqq p$. If $q \in \pi(L)$ and $q \neq p$, then $p \not\leqq q$ and hence $p^* \leqq q$ by Lemma 2.4. Hence we have $G = \{p\}$, and p is an isolated point.

Corollary 5.4. *The induced topology on* $\pi_0(L)$ *from* $\pi(L)$ *is discrete. If L satisfies* (K) *then the minimal spectrum* $\pi(L)$ *is discrete.*

Proof. These statements follow from Theorem 5.3 and Theorem 5.1 (ii) immediately.

Remark 5.5. If every finite product of elements of $\pi(L)$ is compact (especially, f L satisfies (K)), then $\pi(L)$ is a finite set. This follows from Theorem 3.4 by putting $a=0$.

Finally, we shall obtain several important results about hulls and nilpotent elements, assuming the condition (K).

Lemma 5.6. *Assume that L satisfies* (K). *For* $a \in L$ *and* $p \in \pi(L)$, $a \leqq p$ *if aud only if* $a^* \not\leqq p$. *Hence,* $h(a) = \pi(L) - h(a^*) = h(a^{**})$.

Proof. This follows from the property (4) in Theorem 5.1.

Theorem 5.7. *Assume that L satisfies* (K), *and let R be a subset of* $\pi(L)$. *If we put* $a=\vee\{p^*: p\in R\}$, *then* $R=h(a^*)=h(K(R))$.

Proof. If $p\in R$, then we have $p^*\leqq a$ and $p^*\nleqq p$, and then $a\nleqq p$. Conversely, if $a\nleqq p\in\pi(L)$, there exists $q\in R$ such that $q^*\nleqq p$. Then, $q\leqq p$ by Lemma 5.6, and hence $p=q\in R$. Therefore, $R=\pi(L)-h(a)=h(a^*)$. Next, we have $a^*\leqq K(h(a^*))=$ $=K(R)$, and hence $h(K(R))\subset h(a^*)=R\subset h(K(R))$.

Lemma 5.8. *Assume that L satisfies* (*K*).

(i) $\sqrt{0}$ *is the greatest nilpotent element and is equal to* $\wedge\{p: p\in\pi(L)\}$.

(ii) $x\in L$ *is nilpotent if and only if* $h(x)=\pi(L)$.

(iii) x^* *is nilpotent if and only if* $h(x)=\emptyset$.

(iv) $x\wedge x^*$ *is nilpotent for every* $x\in L$.

Proof. (i) follows from Lemma 3.5 (i) and Theorem 3.6. Evidently, (ii) follows from (i). (iii) follows from (ii), since $h(x^*)=\pi(L)-h(x)$. (iv) follows from Lemma 2.3 (i).

Theorem 5.9. *Assume that L satisfies* (*K*). *The following eight statements are equivalent:*

(1) *no nonzero element of L is nilpotent;*

(2) $\wedge\{p: p\in\pi(L)\}=0$;

(3) $x^*=\wedge\{p\in\pi(L): x\nleqq p\}$ *for every* $x\in L$;

(4) $x^*=K(h(x^*))$ *for every* $x\in L$;

(5) $x^{**}=K(h(x))$ *for every* $x\in L$;

(6) $x\leqq x^{**}$ *for every* $x\in L$;

(7) $x\wedge x^*=0'$ *for every* $x\in L$;

(8) $x^*=1$ *implies* $x=0$.

Proof. The equivalence of (1) and (2) follows from Lemma 5.8 (i). The equivalence of (3) and (4) follows from Lemma 5.6. (2)⇒(4): Putting $y=K(h(x^*))$, we have $x^*\leqq y$. If $x\leqq p\in\pi(L)$, then $y\leqq p$, since $x^*\leqq p$. Hence, $xy\leqq p$ for every $p\in\pi(L)$, and hence $xy=0$ by (2). Thus, $y\leqq x^*$. (4)⇒(5) is evident, since $h(x^{**})=h(x)$. (5)⇒(6) is evident. (6)⇒(8) is evident, since $1^*=0$. (8)⇒(1) follows from Lemma 2.3 (ii). (1)⇒(7) follows from Lemma 5.8 (iv). (7)⇒(8) is evident.

Theorem 5.10. *Assume that L satisfies* (*K*) *and that no nonzero element of L is nilpotent.*

(i) *L is pseudo-complemented and x^* is a pseudo-complement of x for any $x \in L$.*

(ii) *For $x, y \in L$, $h(x) \subset h(y)$ if and only if $x^* \leqq y^*$. Hence, $h(x)=h(y)$ if and only if $x^*=y^*$.*

(iii) *$x^{***}=x^*$ for every $x \in L$.*

(iv) *For $a \in L$, the following four statements are equivalent:*

(1) *$a=a^{**}$ (following* [10], *a may be called normal);*

(2) *$a=b^*$ for some $b \in L$;*

(3) *$a=K(h(a))$;*

(4) *a is the kernel of some subset of $\pi(L)$.*

Proof. (i) If $y \wedge x=0$, then $xy=0$ and hence $y \leqq x^*$. Then, by (7) of Theorem 5.9, x^* is the greatest element of the set $\{y \in L: y \wedge x=0\}$.

(ii) If $h(x) \subset h(y)$, then $h(x^*)=\pi(L)-h(x) \supset \pi(L)-h(y)=h(y^*)$, and hence $x^*=K(h(x^*)) \leqq K(h(y^*))=y^*$ by (4) of Theorem 5.9. Conversely, if $x^* \leqq y^*$, then $h(x^*) \supset h(y^*)$ and then $h(x) \subset h(y)$.

(iii) By (6) of Theorem 5.9, we have $x^* \leqq (x^*)^{**}$, and moreover $x^{**} \geqq x$ implies $(x^{**})^* \leqq x^*$.

(iv) (1)$\Rightarrow$(2) and (3)$\Rightarrow$(4) are trivial. (1) and (3) are equivalent by (5) of Theorem 5.9. (2)$\Rightarrow$(1): If $a=b^*$ then $a^{**}=b^{***}=b^*=a$ by (iii). (4)$\Rightarrow$(3): If $a=K(R)$ for some $R \subset \pi(L)$, then we have $h(a)=R$ by Theorem 5.7.

References

[1] D. D. Anderson, Abstract commutative ideal theory without chain condition, *Algebra Universalis*, **6** (1976), 131—145.

[2] R. P. Dilworth, Abstract commutative ideal theory, *Pacific J. Math.*, **12** (1962), 481—498.

[3] K. H. Hofmann and K. Keimel, A general character theory for partially ordered sets and lattices, *Mem. Amer. Math. Soc.*, No. 122, 1972.

[4] K. Keimel, A unified theory of minimal prime ideals, *Acta Math. Acad. Sci. Hungar.*, **23** (1972), 51—69.

[5] K. Murata, Primary decomposition of elements in compactly generated integral multiplicative lattices, *Osaka J. Math.*, **7** (1970), 97—115.

[6] W. C. Nemitz, Irreducibility in implicative semilattices, *Algebra Universalis*, **9** (1979), 244—249.

[7] Y. S. Pawar and N. K. Thakare, The space of minimal prime ideals in 0-distributive lattices, *Period. Math. Hungar.*, **13** (1982), 309—319.

[8] H. Simmons, Reticulated rings, *J. Algebra*, **66** (1980), 169—192.

[9] N. K. Thakare and C. S. Manjarekar, Abstract spectral theory: Multiplicative lattices in which every character is contained in a unique maximal character, in: *Algebra and its applications*, Marcel Dekker Inc. (New York 1984), pp. 265—276.

[10] N. K. THAKARE and S. K. NIMBHORKAR, Spaces of minimal prime ideals of a ring without nilpotent elements, *J. Pure Appl. Algebra,* **27** (1983), 75—85.

[11] N. K. THAKARE and Y. S. PAWAR, Minimal prime ideals in 0-distributive semilattices, *Period. Math. Hungar.,* **13** (1982), 237—246.

(N. K. T.)
JAI HIND COLLEGE OF ARTS,
SCIENCE AND COMMERCE
DHULE 424 002, INDIA

(C. S. M.)
DEPARTMENT OF MATHEMATICS
VIVEKANAND COLLEGE
KOLHAPUR 416 002, INDIA

(S. M.)
DEPARTMENT OF MATHEMATICS
EHIME UNIVERSITY
BUNKYO-CHO, MATSUYAMA, 790
JAPAN

Current address (N. K. T.):
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF POONA
GANESHKING, PUNE-471 007
INDIA

Acta Sci. Math., 52 (1988), 69—79

Varieties and quasivarieties, generated by two-element preprimal algebras, and their equivalences

KLAUS DENECKE*

Dedicated to Professor H.-J. Hoehnke on his 63rd birthday

1. Introduction

The subsequent considerations on universal algebras are stimulated by the following situation in the variety of Boolean algebras: It is generated by the two-element Boolean algebra **2** which has the property that every function defined on the two-element set $\{0, 1\}$ is a term function of **2**. This property corresponds to the functional completeness of classical propositional calculus since the class of Boolean algebras constitutes a semantical basis for classical logics. As a generalization one defines a finite nontrivial algebra $\mathbf{A}=\langle A; F\rangle$ to be primal if every function on A is a term function of $\mathbf{A}$. Then many properties of Boolean algebras carry over immediately to varieties generated by a primal algebra. This is already implied by the categorical equivalence between any variety which is generated by a primal algebra and the variety of Boolean algebras.

This equivalence is generalized now in two directions: firstly to preprimal algebras and secondly to quasivarieties. The term functions of a preprimal algebra $\mathbf{A}=\langle A; F\rangle$ constitute a dual atom in the lattice of closed classes of functions defined on A. All two-element preprimal algebras were determined by E. L. Post [11]. Identifying algebras with the same term functions we obtain exactly the following two-element preprimal algebras (up to isomorphisms):

$$\mathbf{C}_3 = \langle\{0, 1\};\ \wedge,\ +,0\rangle,\quad \mathbf{A}_1 = \langle\{0, 1\};\ \wedge, \vee, 0, 1\rangle,$$
$$\mathbf{D}_3 = \langle\{0, 1\};\ d, x+y+z, N\rangle,\quad \mathbf{L}_1 = \langle\{0, 1\};\ +, N, 0, 1\rangle.$$

Received January 7, 1985, and in revised form July 3, 1985.

*) During the preparation of this paper, the author worked at the Institut für Züchtungsforschung der Akademie der Landwirtschaftswissenschaften der DDR, Quedlinburg.

Here $\wedge, \vee, +, N$ are the Boolean operations conjunction, disjunction, addition mod 2, and negation. Further d is the ternary operation with $d(x, y, z)=(x\wedge y)\vee \vee(x\wedge z)\vee(y\wedge z)$. Our main result is the following: A quasivariety is equivalent to the quasivariety generated by one of the two-element preprimal algebras if and only if it is generated by a preprimal algebra of a special form. The result can be applied in non-classical logics and in electrical circuit theory. Consider a variety $V_{2'}$ generated by a two-element algebra and assume $V_{2'}=\mathrm{ISP}(\mathbf{2}')$ (I—isomorphisms, S—subalgebras, P—direct products), i.e., assume the quasivariety $QV_{2'}=\mathrm{ISP}(\mathbf{2}')$ generated by $\mathbf{2}'$ agrees with the variety generated by $\mathbf{2}'$. In [2] the algebras $B\in\mathrm{ISP}(\mathbf{2}')$ are called pure dyadic algebras. Boolean algebras and Boolean rings, distributive lattices, implication algebras, median algebras, and Boolean groups are well-known examples of pure dyadic algebras. Let $\mathbf{B}(X)\in V_{2'}$ be the free algebra freely generated by $X=\{x_1, \ldots, x_n\}$, and let $\mathbf{p}, \mathbf{q}$ be two terms of $\mathbf{B}(X)$. The fact that every algebra of $V_{2'}$ is isomorphic to a subdirect power of $\mathbf{2}'$ implies that $\mathbf{p}, \mathbf{q}\in\mathbf{B}(X)$ are identical if for all homomorphisms h: $\mathbf{B}(X)\to\mathbf{2}'$ one has $h(\mathbf{p})=h(\mathbf{q})$. In the case of Boolean algebras this property is meaningful in the complexity theory of Boolean functions and the truth table method of classical logics ([8]). Let $\mathscr{K}$ be a variety which, as a category, is equivalent to $V_{2'}$. Then there is a map t from the n-ary terms of $V_{2'}$ to the n-ary terms of $\mathscr{K}$ such that

(i) $t(\mathbf{x}_i)=\mathbf{x}_i$,

(ii) if α and β are self-maps of $\{1, \ldots, n\}$ and $V_{2'}$ satisfies $\mathbf{p}(x_{\alpha 1}, \ldots, x_{\alpha n})= =\mathbf{p}(x_{\beta 1}, \ldots, x_{\beta n})$, then $\mathscr{K}$ satisfies $(t\mathbf{p})(x_{\alpha 1}, \ldots, x_{\alpha n})=(t\mathbf{q})(x_{\beta 1}, \ldots, x_{\beta n})$.

It follows that $\mathscr{K}$ satisfies $(t\mathbf{p})(x_{\alpha 1}, \ldots, x_{\alpha n})=(t\mathbf{q})(x_{\beta 1}, \ldots, x_{\beta n})$ if $h(\mathbf{p})=h(\mathbf{q})$ holds for all homomorphisms h: $\mathbf{B}(X)\to\mathbf{2}'$.

2. Preliminaries

Let A be a nonempty finite set. The collection of n-ary operations on A will be denoted by $O_A^{(n)}$ $(n\geqq 1)$. We set $O_A=\bigcup_{n\geqq 1} O_A^{(n)}$. Let ϱ be an h-ary relation on A $(h\geqq 1)$, i.e. $\varrho\subseteqq A^h$. Let Pol ϱ denote the set of all operations from O_A preserving ϱ, i.e. all operations $f\in O_A$ such that ϱ is a subalgebra of $\langle A; f\rangle^h$. A ternary operation $d\in O_A^{(3)}$ is called a majority function if for all $x, y\in A$ we have

$$d(x, x, y) = d(x, y, x) = d(y, x, x) = x.$$

We adopt the terminology of [7] except that polynomials will be called term functions. $T(\mathbf{A})$ denotes the set of term functions of an algebra $\mathbf{A}=\langle A;\ F\rangle$. $\mathbf{A}$ is said to be primal if $T(\mathbf{A})=O_A$. $\mathbf{A}$ is order complete if there is a lattice order $\leqq$ on A such that Pol $\leqq\, =T(\mathbf{A})$. $\mathbf{A}$ is said to be preprimal if $T(\mathbf{A})\neq O_A$ and the algebra

$\langle A;\ F\cup\{f\}\rangle$ is primal for every operation $f\in O_A\setminus T(\mathbf{A})$. By a compatible relation of an algebra $\mathbf{A}=\langle A;\ F\rangle$ we mean a relation ϱ on A such that $F\subseteq \operatorname{Pol}\varrho$. The compatible binary reflexive and symmetric relations on $\mathbf{A}$ are called tolerance relations of $\mathbf{A}$. We say a relation ϱ generates an algebra $\mathbf{A}$ if $T(\mathbf{A})=\operatorname{Pol}\varrho$, and we write $\mathbf{A}\varrho$ for any such algebra.

For $2\leqq h<\infty$ let $\sigma_h=\{(a_1, \ldots, a_h)\in A^h:\ a_i\neq a_j,\ 1\leqq i<j\leqq h\}$. Furthermore, we set $\iota_h=A^h\setminus\sigma_h$. An h-ary relation ϱ on A $(h\geqq 3)$ is totally reflexive if $\varrho\supseteq\iota_h$. A binary relation on A is called trivial if $\varrho=\iota_2$ or $\varrho=A^2$.

We say that an algebra is tolerance-free if it has no nontrivial tolerance relation. An algebra $\mathbf{A}=\langle A;\ F\rangle$ is said to be semiprimal if every operation on A admitting all subalgebras of $\mathbf{A}$ is a term function of $\mathbf{A}$ and demiprimal if $\mathbf{A}$ has no proper subalgebra and every operation on A admitting all automorphisms of $\mathbf{A}$ is a term function of $\mathbf{A}$. We need the following result from [1].

Theorem 2.1. *Let $\mathbf{A}=\langle A;\ F\rangle$ be a finite algebra with a majority term function. Then an operation on A is a term function of $\mathbf{A}$ iff it preserves all compatible binary relations of $\mathbf{A}$.*

From Theorem 2.1 we obtain immediately the following

Corollary 2.2. *Let $\mathbf{A}=\langle A;\ F\rangle$ be a finite algebra with a majority term function. Then $\mathbf{A}$ is primal iff it has no nontrivial compatible binary relation. Moreover, $\mathbf{A}$ is preprimal iff it has a nontrivial compatible binary relation and for any two nontrivial compatible relations ϱ_1 and ϱ_2 of $\mathbf{A}$ we have* $\operatorname{Pol}\varrho_1=\operatorname{Pol}\varrho_2$.

We need the following list of preprimal algebras ([12], [5]):

$\mathbf{A}_{\leqq}$, where $\leqq$ is a lattice order on A, hence $\mathbf{A}_{\leqq}$ is order complete,

$\mathbf{A}_{\{b\}}$, where $\{b\}$ is a one-element subalgebra of $\mathbf{A}_{\{b\}}$, hence $\mathbf{A}_{\{b\}}$ is semiprimal,

$\mathbf{A}_{s_2}$, where s_2 is a permutation on A without invariant elements and with cycles of the same length 2, hence $\mathbf{A}_{s_2}$ is demiprimal, $|A|=2m$, $m\in N$,

$\mathbf{A}_{\alpha_m}$, where $\alpha_m=\{(x, y, z, e):\ e=x+y+z\}$, $x+y+z$ is the operation of a Boolean 3-group $\mathbf{G}_3^m=\langle A;\ x+y+z\rangle$ with $|A|=2^m$, $m\in N$, $m\geqq 1$.

Clearly, $\mathbf{A}_1$, $\mathbf{C}_3$, $\mathbf{D}_3$ and $\mathbf{L}_1$ are preprimal algebras of these forms with $|A|=2$.

Let $\mathcal{L}$ and $\mathcal{K}$ be quasivarieties which are equivalent as categories, i.e., there are functors $G\colon \mathcal{K}\to\mathcal{L}$ and $H\colon \mathcal{L}\to\mathcal{K}$, and for each $\mathbf{A}\in\mathcal{K}$ and $\mathbf{B}\in\mathcal{L}$ there are isomorphisms $\alpha_{\mathbf{A}}\colon \mathbf{A}\to HG(\mathbf{A})$ and $\beta_{\mathbf{B}}\colon \mathbf{B}\to GH(\mathbf{B})$ such that for each $g\colon \mathbf{A}\to\mathbf{A}'$ in $\mathcal{K}$ and each $h\colon \mathbf{B}\to\mathbf{B}'$ in $\mathcal{L}$ the following diagrams commute:

$$\begin{array}{ccc} \mathbf{A} & \xrightarrow{\ g\ } & \mathbf{A}' \\ {\scriptstyle\alpha_{\mathbf{A}}}\downarrow & & \downarrow{\scriptstyle\alpha_{\mathbf{A}'}} \\ HG(\mathbf{A}) & \xrightarrow{HG(g)} & HG(\mathbf{A}') \end{array} \qquad \begin{array}{ccc} \mathbf{B} & \xrightarrow{\ h\ } & \mathbf{B}' \\ {\scriptstyle\beta_{\mathbf{B}}}\downarrow & & \downarrow{\scriptstyle\beta_{\mathbf{B}'}} \\ GH(\mathbf{B}) & \xrightarrow{GH(h)} & GH(\mathbf{B}') \end{array}$$

The question arises, which properties of a quasivariety carry over to equivalent quasivarieties? Necessary conditions are given by

Theorem 2.2. [3] *Let $\mathcal{L}$ and $\mathcal{K}$ be quasivarieties which are equivalent as categories via the functors $G: \mathcal{K} \to \mathcal{L}$ and $H: \mathcal{L} \to \mathcal{K}$.*

(1) *If $\mathbf{A} \in \mathcal{L}$ is a finite algebra, then $H(\mathbf{A})$ is a finite algebra.*

(2) *For all $\mathbf{A} \in \mathcal{L}$ the subalgebra lattices of $\mathbf{A}$ and $H(\mathbf{A})$ are isomorphic. Therefore the subalgebra lattices of $\mathbf{A}^2$ and $H(\mathbf{A}^2)$ are isomorphic and since $H(\mathbf{A}^2)$ is isomorphic to $H(\mathbf{A})^2$, the subalgebra lattices of $\mathbf{A}^2$ and $H(\mathbf{A})^2$ are isomorphic.*

(3) *H maps subdirectly irreducible algebras to subdirectly irreducible algebras, simple algebras to simple algebras, and tolerance-free algebras to tolerance-free algebras.*

(4) *If $\mathcal{L}$ is the variety generated by some algebra $\mathbf{A}$, then $\mathcal{K}$ is the variety generated by $H(\mathbf{A})$.*

(5) *If $\mathcal{L}$ and $\mathcal{K}$ are varieties and if in $\mathcal{L}$ there exists a majority term then in $\mathcal{K}$ there also exists a majority term; i.e. if $\mathcal{L}$ is the variety generated by $\mathbf{A}$ and $\mathbf{A}$ has a majority function among its term functions then $H(\mathbf{A})$ also has a majority function among its term functions.*

3. Tolerance-free algebras having majority term functions

The two-element preprimal algebras $\mathbf{C}_3, \mathbf{A}_1$ and $\mathbf{D}_3$ have majority functions among their algebraic functions ([4]) and admit no nontrivial tolerance relation. By [4] the quasivarieties generated by $\mathbf{C}_3$, $\mathbf{A}_1$ and $\mathbf{D}_3$ agree with the varieties generated by these algebras. Therefore, by Theorem 2.2 (3), (4), (5), varieties equivalent as categories to $V_{\mathbf{C}_3}, V_{\mathbf{A}_1}, V_{\mathbf{D}_3}$ are generated by tolerance-free algebras $H(\mathbf{C}_3)$, $H(\mathbf{A}_1)$, and $H(\mathbf{D}_3)$ having majority functions among their term functions. In order to characterize varieties equivalent to $V_{\mathbf{C}_3}, V_{\mathbf{A}_1}, V_{\mathbf{D}_3}$ we give some properties for tolerance-free algebras having majority term functions.

For a binary relation on A define two n-ary relations ϱ_n and ϱ_n' $(2 \leqq n \leqq |A|)$ as follows:

$$\varrho_n = \{(a_1, \ldots, a_n) \in A^n \colon (a_i, u) \in \varrho,\ i = 1, \ldots, n, \text{ for some } u \in A\},$$
$$\varrho_n' = \{(a_1, \ldots, a_n) \in A^n \colon (o, a_i) \in \varrho,\ i = 1, \ldots, n, \text{ for some } o \in A\}.$$

Lemma 3.1. *Let ϱ be a binary relation on A preserved by a majority function $d \in O_A^{(3)}$. If $\varrho \circ \varrho^{-1} = A^2$ $(\varrho^{-1} \circ \varrho = A^2)$, then $\varrho_n = A^n$ $(\varrho_n' = A^n)$ for every $n = 2, \ldots, |A|$.*

Proof. We prove the lemma by induction on n. Clearly, $\varrho_2 = \varrho \circ \varrho^{-1} = A^2$. Suppose that $\varrho_{n-1} = A^{n-1}$, $2 \leqq n \leqq |A|$. From the definition of ϱ_n it follows that $\varrho_n \supseteq \iota_n$, i.e. ϱ_n is totally reflexive. Now, if $(a_1, \ldots, a_n) \in A^n$ then $(a_2, a_2, a_3, a_4, \ldots, a_n) \in \varrho_n$, $(a_1, a_1, a_3, a_4, \ldots, a_n) \in \varrho_n$ and $(a_1, a_2, a_2, a_4, \ldots, a_n) \in \varrho_n$. Therefore

$(a_1, \ldots, a_n)=(d(a_2, a_1, a_1), d(a_2, a_1, a_2), d(a_3, a_3, a_2), d(a_4, a_4, a_4), \ldots, d(a_n, a_n, a_n))\in$ $\in \varrho_n$. Hence $\varrho_n=A^n$. (Similarly, we can prove that $\varrho^{-1}\circ\varrho=A^2$ implies $\varrho_n'=A^n$, $n=2, \ldots, |A|$.)

Lemma 3.2. *Let* $\mathbf{A}=\langle A;\ F\rangle$ *be a tolerance-free algebra admitting a majority term function, and let* ϱ *be a binary nontrivial reflexive compatible relation of* $\mathbf{A}$. *Then* ϱ *is a lattice order.*

Proof. $\varrho\cap\varrho^{-1}$ $(\subseteq\varrho)$ is a tolerance relation of $\mathbf{A}$ distinct from A^2. Therefore $\varrho\cap\varrho^{-1}=\iota_2$, i.e. ϱ is antisymmetric. $\varrho\circ\varrho^{-1}$ and $\varrho^{-1}\circ\varrho$ are tolerance relations distinct from ι_2. Therefore, $\varrho\circ\varrho^{-1}=\varrho^{-1}\circ\varrho=A^2$, which by Lemma 3.1 implies that $\varrho_{|A|}=\varrho'_{|A|}=A^{|A|}$. Hence there are elements $0, 1\in A$ such that $(a, 1)\in\varrho$ and $(0, a)\in\varrho$ for every $a\in A$. Let d be a majority term function of $\mathbf{A}$. It is known [6] that $d(0, a, b)=a\wedge b$ and $d(1, a, b)=a\vee b$ are the infimum and supremum of a and b with respect to ϱ. Finally we show that ϱ is transitive. Let $(a, b)\in\varrho$ and $(b, c)\in\varrho$. Then $d(0, a, b)=a\wedge b=a$ and $d(1, b, c)=b\vee c=c$. Therefore $(a, c)=(d(0, a, b), d(1, b, c))\in\varrho$, which completes the proof.

Lemma 3.3. *Let* $\mathbf{A}=\langle A;\ F\rangle$ *be a tolerance-free algebra with a majority term function admitting no proper subalgebra. Let* ϱ *be a binary nontrivial symmetric compatible relation of* $\mathbf{A}$ *with* $\varrho\cap\iota_2=\emptyset$. *Then* $\varrho=\{(a, s(a)):\ a\in A\}$ *where* s *is an automorphism of* $\mathbf{A}$ *without fixed points and with cycles of equal length* 2.

Proof. Since $\varrho\circ\varrho^{-1}$ and $\varrho^{-1}\circ\varrho$ are tolerance relations of $\mathbf{A}$ it follows that $\varrho\circ\varrho^{-1}, \varrho^{-1}\circ\varrho\in\{\iota_2, A^2\}$. If $\varrho\circ\varrho^{-1}=A^2$, then by Lemma 3.1 $\varrho_{|A|}=A^{|A|}$. Thus there is a $u\in A$ such that $(a, u)\in\varrho$ for every $a\in A$, implying that $(u, u)\in\varrho$, a contradiction. Similarly we can prove that $\varrho^{-1}\circ\varrho\neq A^2$. Hence $\varrho\circ\varrho^{-1}=\varrho^{-1}\circ\varrho=\iota_2$, which implies that $\varrho=\{(a, s(a)):\ a\in A\}$ for a permutation s on A. Clearly, s has no fixed point $(\varrho\cap\iota_2=\emptyset)$. From $\varrho=\varrho^{-1}$ one gets $\varrho^2=\iota_2$. Therefore each cycle of s has length 2.

The proof of the next lemma is given in [6].

Lemma 3.4. *Let* $\mathbf{A}=\langle A;\ F\rangle$ *be a tolerance-free algebra having a majority term function. Then* $\mathbf{A}$ *has at most two compatible lattice orders* ϱ *and* ϱ^{-1}.

Lemma 3.5. *Let* $\mathbf{A}=\langle A;\ F\rangle$ *be an algebra with a majority term function and exactly one proper subalgebra which moreover has exactly one element. Let* $\{b\}$ *be the one-element subalgebra of* $\mathbf{A}$. *Suppose* $\mathbf{A}$ *has exactly three nontrivial binary compatible relations. Then* $\mathbf{A}$ *is a semiprimal algebra of the form* $\mathbf{A}_{\{b\}}$ *and thus preprimal.*

Proof. $\{b\}\times\{b\}$, $A\times\{b\}$, and $\{b\}\times A$ are all nontrivial compatible binary relations of $\mathbf{A}$. Therefore, by Theorem 2.1, $T(\mathbf{A})=\mathrm{Pol}\,(\{b\}\times\{b\})\cap\mathrm{Pol}\,(A\times\{b\})\cap$

$\cap \mathrm{Pol}(\{b\}\times A)=\mathrm{Pol}(\{b\})$, i.e. $\mathbf{A}$ is a semiprimal algebra of the form $\mathbf{A}_{\{b\}}$ and thus preprimal.

We are ready to formulate and prove our first theorem.

Theorem 3.6. *Let* $\mathbf{P}$ *be one of the two-element algebras* $\mathbf{A}_1, \mathbf{C}_3, \mathbf{D}_3$, *and let* $V_{\mathbf{P}}$ *be the variety generated by* $\mathbf{P}$. *Let* $\mathcal{K}$ *be a variety equivalent as a category to* $V_{\mathbf{P}}$. *Then* $\mathcal{K}$ *is generated by one of the preprimal algebras* $\mathbf{A}_{\leqq}$, $\mathbf{A}_{\{b\}}$ *or* A_{s_2}.

Proof. Let $\mathcal{K}$ be a quasivariety which is equivalent as a category to the quasivariety $QV_{\mathbf{P}}$ via some functors $G\colon \mathcal{K}\to QV_{\mathbf{P}}$ and $H\colon QV_{\mathbf{P}}\to\mathcal{K}$. Since $\mathbf{P}$ has a term function which is a majority function, by a result of Jónsson [10], we have $QV_{\mathbf{P}}=V_{\mathbf{P}}$. By Theorem 2.2, $\mathcal{K}$ is the variety generated by the finite algebra $H(\mathbf{P})$ and $H(\mathbf{P})$ is tolerance-free, having a term function which is a majority function. $H(\mathbf{A}_1)$ and $H(\mathbf{D}_3)$ have no proper subalgebras and $H(\mathbf{C}_3)$ has exactly one (one-element) subalgebra. By Theorem 2.2 (2), the subalgebra lattices of $\mathbf{P}^2$ and $H(\mathbf{P})^2$ are isomorphic. Therefore $H(\mathbf{D}_3)$ has exactly one nontrivial compatible binary relation ϱ and $\varrho\cap\iota_2=\emptyset$ holds. By Lemma 3.3, Theorem 2.1, and Corollary 2.2 $H(\mathbf{D}_3)$ is a demiprimal preprimal algebra of the form $\mathbf{A}_{s_2}$. Further, $H(\mathbf{A}_1)$ has exactly two binary nontrivial compatible relations which are reflexive. By Lemma 3.2, Lemma 3.4, Theorem 2.1, and Corollary 2.2 $H(\mathbf{A}_1)$ is an order-complete preprimal algebra $\mathbf{A}_{\leqq}$. $H(\mathbf{C}_3)$ has exactly three nontrivial binary compatible relations. By Lemma 3.5, $H(\mathbf{C}_3)$ is a semiprimal preprimal algebra of the form $\mathbf{A}_{\{b\}}$.

4. Dualities and full dualities of quasivarieties

The next statements concern the category equivalence of a quasivariety generated by any preprimal algebra of the form $\mathbf{A}_{\leqq}$, $\mathbf{A}_{\{b\}}$, $\mathbf{A}_{s_2}$, $\mathbf{A}_{\alpha_m}$ to the quasivariety generated by a two-element preprimal algebra $\mathbf{A}_1, \mathbf{C}_3, \mathbf{D}_3, \mathbf{L}_1$. These considerations rest upon concepts and results of Davey—Werner [3] on dualities and equivalences of quasivarieties.

Let $\mathbf{C}=\langle C;\ F\rangle$ be a finite algebra and let $\mathcal{L}=\mathrm{ISP}(\mathbf{C})$ be the quasivariety generated by $\mathbf{C}$. Let $\mathsf{C}=\langle C;\ \tau, R\rangle$ be a topological relational structure where R is a set of compatible relations of $\mathbf{C}$, and τ is the discrete topology on C. Let $\mathcal{X}$ be the class of all topological relational structures of the same type as C. For $\mathsf{X}, \mathsf{Y}\in\mathcal{X}$ a morphism $X\to Y$ is a map between the carrier sets of X, Y, which preserves the defining relations of X, Y. Let $\mathcal{X}(X, Y)$ denote the set of all continuous morphisms $X\to Y$. A mapping $\Phi\in\mathcal{X}(X, Y)$ is an embedding if it is one-to-one, closed, and for each relation $r\in R$ and $x_1, \ldots, x_n\in X$ we have

$$(\Phi(x_1), \ldots, \Phi(x_n))\in r \Rightarrow (x_1, \ldots, x_n)\in r.$$

An onto-embedding is an isomorphism in $\mathscr{X}$. Let $\mathbf{X}\in\mathscr{X}$ and $Y\subseteq X$. $\mathbf{Y}$ is a closed substructure if the inclusion map $Y\to X$ is an embedding. A power of $\mathbf{C}$ is always endowed with the product topology and the pointwise relations, i.e. the sets

$$\langle i;\, p\rangle := \{x\in C^I : x(i) = p\} \quad \text{with } i\in I \text{ and } p\in C$$

form a subbasis for the topology on C^I. For $x_1, \ldots, x_n\in C^I$ one has

$$(x_1, \ldots, x_n)\in r \Leftrightarrow (\forall i\in I)(x_1(i), \ldots, x_n(i))\in r.$$

The subclass of $\mathscr{X}$ consisting of all members isomorphic to a closed substructure of a power of $\mathbf{C}$ is denoted by $\mathscr{R}$. Symbolically, we write $\mathscr{R}=\mathrm{ISP}(\mathbf{C})$.

The following lemma shows the interconnection between the categories $\mathscr{L}$ and $\mathscr{R}$.

Lemma 4.1. *There exists a pair of adjoint contravariant functors* $D\colon \mathscr{L}\to\mathscr{R}$, $E\colon \mathscr{R}\to\mathscr{L}$.

A pair (D, E) as in Lemma 4.1 is called a protoduality. The protoduality is called a duality if for each algebra $\mathbf{A}$ in $\mathscr{L}$ the embedding $e_A\colon \mathbf{A}\to ED(\mathbf{A})$ is an isomorphism.

Let $\mathscr{R}_0\subseteq\mathscr{R}$ be the subcategory consisting of all structures isomorphic to some closed substructure of a power of $\mathbf{C}$. Then the duality (D, E) is called a full duality between $\mathscr{L}$ and $\mathscr{R}_0$ if for all $\mathbf{X}\in\mathscr{R}_0$ the embedding $\varepsilon_X\colon \mathbf{X}\to DE(\mathbf{X})$ is an isomorphism. $\mathbf{C}$ is said to be injective in $\mathscr{R}_0$ (with respect to some class $\mathscr{I}$ of embeddings) if for each embedding $\sigma\colon \mathbf{X}\to\mathbf{Y}$ in $\mathscr{R}_0$ $(\sigma\in\mathscr{I})$, every continuous morphism $\varphi\colon \mathbf{X}\to\mathbf{C}$ extends to a continuous morphism $\psi\colon \mathbf{Y}\to\mathbf{C}$ with $\psi\circ\sigma=\varphi$.

The next statements rest upon the following two conditions (IB) and (EF).

(IB) For every substructure $\mathbf{X}$ of a finite power $\mathbf{C}^n$ of $\mathbf{C}$, each morphism $\varphi\colon \mathbf{X}\to\mathbf{C}$ extends to a term function $\bar{\varphi}\colon C^n\to C$ of $\mathbf{C}$.

(EF) If $\mathbf{X}$ is a proper substructure of some finite $\mathbf{Y}\in\mathscr{R}_0$ then there exist two different morphisms $\varphi, \psi\colon \mathbf{Y}\to\mathbf{C}$ such that $\varphi/X=\psi/X$.

Lemma 4.2. *Let* $\mathscr{L}=\mathrm{ISP}(\mathbf{C})$ *for a finite algebra* $\mathbf{C}=\langle C;\, F\rangle$. *Let* $\mathbf{C}=\langle C;\, \tau, R\rangle$ *be a (finite) relational structure where R is a finite set of compatible relations on* $\mathbf{C}$ *and* $\mathscr{R}=\mathrm{ISP}(\mathbf{C})$. *Suppose the conditions* (IB) *and* (EF) *hold. Then the protoduality* (D, E) *is a full duality between* $\mathscr{L}$ *and* $\mathscr{R}_0$, *and* $\mathbf{C}$ *is injective in* $\mathscr{R}_0$.

Now we assume that $\mathbf{C}$ admits a majority term function.

Lemma 4.3. *Let* $\mathbf{C}=\langle C;\, F\rangle$ *be a finite algebra with a majority term function. Let R be the set of all binary compatible relations on* $\mathbf{C}$. *Then the protoduality* (D, E) *is a duality between* $\mathscr{L}$ *and* $\mathscr{R}_0$, *and* $\mathbf{C}$ *is injective in* $\mathscr{R}_0$. *If* (EF) *holds,* (D, E) *is a full duality between* $\mathscr{L}$ *and* $\mathscr{R}_0$.

We are ready to apply the preceding duality theory to obtain dualities or even full dualities for varieties (quasivarieties) generated by two-element preprimal algebras.

Theorem 4.4. *Let* $\mathbf{2}_P=\langle\{0,1\};\ F\rangle$ *be a two-element preprimal algebra* $(\mathbf{2}_P\in\{\mathbf{A}_1, \mathbf{C}_3, \mathbf{D}_3, \mathbf{L}_1\})$. *Let* $\boldsymbol{2}_P=\langle\{0,1\};\ \varrho\rangle$ *be a finite relational structure with* $F=\mathrm{Pol}\,\varrho$ *and* $\mathcal{R}=\mathrm{ISP}(\boldsymbol{2}_P)$. *Then the protoduality is a full duality between* $\mathcal{L}$ *and* $\mathcal{R}_0$ *and* $\boldsymbol{2}_P$ *is injective in* $\mathcal{R}_0$.

Proof. By Corollary 2.2 for any two nontrivial compatible relations ϱ_1, ϱ_2 of a preprimal algebra $\mathbf{A}=\langle A;\ F\rangle$ we have $F=\mathrm{Pol}\,\varrho_1=\mathrm{Pol}\,\varrho_2$. Therefore we can set $\boldsymbol{2}_P=\langle\{0,1\};\ \varrho\rangle$ with $F=\mathrm{Pol}\,\varrho$. The algebras $\mathbf{A}_1$, $\mathbf{C}_3$, and $\mathbf{D}_3$ have majority term functions. In view of Lemma 4.3 it is sufficient to prove that condition (EF) is satisfied. We define $\boldsymbol{A}_1=\langle\{0,1\};\ \leqq\rangle$, $\boldsymbol{C}_3=\langle\{0,1\};\ 0\rangle$, $\boldsymbol{D}_3=\langle\{0,1\};\ N\rangle$. In the first case, if $\boldsymbol{X}\subset\boldsymbol{Y}\in\mathcal{R}_0$, Y finite, and $a\in Y\setminus X$, then both $(a]=\{y\in Y:\ y\leqq a\}$ and $(a)=\{y\in Y:\ y<a\}$ are ideals such that $X\cap(a]=X\cap(a)$. Thus $\varphi,\psi\colon Y\to\{0,1\}$, $\varphi(x)=0\Leftrightarrow x\leqq a$, $\psi(x)=0\Leftrightarrow x<a$ are two order-preserving maps which agree on X. In the second case, let $\boldsymbol{X}\subset\boldsymbol{Y}$ be a substructure of a finite $\boldsymbol{Y}\in\mathcal{R}_0$, i.e. $0\in\boldsymbol{X}$ and let $\varphi,\psi\colon Y\to C_3$ with $\varphi(x)=0$ and

$$\psi(x)=\begin{cases}0 & \text{if } x\in X\\ 1 & \text{if } x\notin X.\end{cases}$$

Then φ and ψ are morphisms, $\varphi\neq\psi$ but $\varphi/X=\psi/X$.

Now we consider $\mathbf{D}_3$. Let $\boldsymbol{X}\subset\boldsymbol{Y}\in\mathcal{R}_0$, Y finite, i.e. $N\boldsymbol{X}\subseteq\boldsymbol{X}$ where N is a permutation on Y with cycles of the same length 2 and without fixed points. Then we consider two proper subsets $X_1, X_2\subset X$ with $X_1=\{x\in X:\ Nx\in X_2\}$, $X_2=\{x\in X:\ Nx\in X_1\}$, $0\in X_1$, $1\in X_2$, $N0=1$. From $Nx\neq x$, $x\in Y$ it follows $X_1\cap X_2=\emptyset$. Further, we have $X_1\cup X_2=X$, X_1 and X_2 can be extended to Y_1 and Y_2, respectively, such that $Y_1=\{x\in Y:\ Nx\in Y_2\}$, $Y_2=\{x\in Y:\ Nx\in Y_1\}$, $Y_1\cap Y_2=\emptyset$, $Y_1\cup Y_2=Y$. We choose

$$\varphi(x)=\begin{cases}0 & \text{if } x\in X_1\\ 1 & \text{if } x\in X_2\\ 0 & \text{if } x\in Y_1\setminus X_1\\ 1 & \text{if } x\in Y_2\setminus X_2\end{cases},\quad \psi(x)=\begin{cases}0 & \text{if } x\in X_1\\ 1 & \text{if } x\in X_2\\ 1 & \text{if } x\in Y_1\setminus X_1\\ 0 & \text{if } x\in Y_2\setminus X_2\end{cases}.$$

φ and ψ are two distinct morphisms which agree on X.

Finally, we consider $\mathbf{L}_1=\langle\{0,1\},+,\mathrm{N},0,1\rangle$. Let $\mathcal{L}=\mathrm{ISP}(\mathbf{L}_1)$ be the quasivariety generated by $\mathbf{L}_1$ $(\mathcal{L}\neq V_{\mathbf{L}_1})$. The term functions of $\mathbf{L}_1$ are exactly all Boolean functions which preserve $\alpha=\{(x,y,z,e):\ e=x+y+z\}$. Here $x+y+z$ is the ternary operation of the Boolean 3-group $\mathbf{G}_3=\langle\{0,1\};\ x+y+z\rangle$. For $\mathbf{L}_1=\mathbf{G}_3$ condition (IB) is satisfied. $\mathrm{ISP}(\mathbf{G}_3)$ is the variety of Boolean 3-groups. $\boldsymbol{X}$ being a proper subalgebra of a finite Boolean 3-group $\mathbf{Y}\in\mathcal{R}_0$, we choose a maximal subgroup $\mathbf{Z}$ of $\mathbf{Y}$ containing $\mathbf{X}$. $\mathbf{Y}\setminus\mathbf{Z}$ is simple and thus isomorphic to $\mathbf{L}_1$. Hence we have two homomorphisms $\mathbf{Y}\to\mathbf{L}_1$ with kernels $\mathbf{Z}$ and $\mathbf{Y}$, respectively, which therefore agree on X. Thus condition (EF) is satisfied.

5. Application of the Equivalent Quasivarieties Theorem

In this section we prove that the quasivarieties generated by the preprimal algebras $\mathbf{A}_{\leqq}$, $\mathbf{A}_{\{b\}}$, $\mathbf{A}_{s_2}$, $\mathbf{A}_{\alpha_m}$, respectively, are equivalent as categories to the varieties (quasivarieties) generated by the two-element preprimal algebras $\mathbf{A}_1$, $\mathbf{C}_3$, $\mathbf{D}_3$, $\mathbf{L}_1$. We need the following Equivalent Quasivarieties Theorem [3].

Theorem 5.1. *Assume that the protoduality (D, E) is a full duality between $\mathcal{L}$ and $\mathcal{R}_0$ and assume further that $\mathbf{C}$ is injective in $\mathcal{R}_0$. Then a quasivariety $\mathcal{K}$ is equivalent as a category to the quasivariety $\mathcal{L}$ if and only if the following conditions are satisfied:*

(i) *there is a finite algebra* $\mathbf{Q}$ *in* $\mathcal{K}$ *and a family R of compatible relations on* $\mathbf{Q}$ *such that* $\mathbf{Q}=\langle Q;\ R\rangle$ *is an object of* $\mathcal{R}_0$,

(ii) (a) $\mathcal{K}=\mathrm{ISP}(\mathbf{Q})$,

(b) $\mathbf{C}$ *is isomorphic to a subalgebra of a power of* $\mathbf{Q}$,

(iii) $\mathbf{Q}$ *is injective in* $\mathcal{R}_0$ *(or equivalently,* $\mathbf{Q}$ *is a retract of a finite power of* $\mathbf{C}$*)*,

(iv) *for each positive integer n every morphism* $\mathbf{Q}^n\to\mathbf{Q}$ *is a term function on* $\mathbf{Q}$.
If $\mathcal{K}$ is equivalent as a category to $\mathcal{L}$, then $\mathbf{Q}$ *above can be chosen to be* $H(\mathbf{C})$.

Let $\mathbf{2}_P=\langle\{0, 1\};\ F\rangle$ be a two-element preprimal algebra and let $\mathbf{2}_P=\langle\{0, 1\};\ \varrho\rangle$ be a relational structure with $F=\mathrm{Pol}\ \varrho$. We set $\mathcal{L}=\mathrm{ISP}(\mathbf{2}_P)$ and $\mathcal{R}=\mathrm{ISP}(\mathbf{2}_P)$. By Theorem 4.4 (D, E) is a full duality between $\mathcal{L}$ and $\mathcal{R}_0$ and $\mathbf{2}_P$ is injective in $\mathcal{R}_0$. In order to apply Theorem 5.1 for the proof that the quasivariety generated by one of the preprimal algebras $\mathbf{A}_{\leqq}$, $\mathbf{A}_{\{b\}}$, $\mathbf{A}_{s_2}$, $\mathbf{A}_{\alpha_m}$ is equivalent as a category to the quasivariety $\mathcal{L}$ one has to show that conditions (i)—(iv) are satisfied.

Lemma 5.2. *The variety generated by a preprimal algebra* $\mathbf{A}_{\leqq}$ *is category equivalent to* $V_{\mathbf{A}_1}$.

Proof. By Theorem 3.6 $\mathcal{K}=\mathrm{ISP}(\mathbf{A}_{\leqq})$ is the variety generated by $\mathbf{A}_{\leqq}$. It is clear that $\mathbf{C}=\mathbf{A}_1=\langle\{0, 1\};\ \leqq\rangle$, $\mathbf{Q}=\mathbf{A}_{\leqq}$, $\mathbf{Q}=\mathbf{A}_{\leqq}=\langle A;\ \leqq\rangle$ fulfil the conditions (i), (ii) (a), and (iv). $\mathbf{A}_1$ is isomorphic to the substructure of $\mathbf{A}_{\leqq}$ consisting of the least and the greatest element with respect to $\leqq$, i.e. (ii) (b) holds. Then the lattice $P(A)$ of all subsets of A is isomorphic to a finite power of $\mathbf{A}_1$, and the maps σ and τ

given by

$$\sigma\colon \mathbf{A}_{\leqq} \to P(A),\quad \sigma(a) = \{x\in A\colon (x, a)\in \leqq \text{ for all } a\in A\},$$

$$\tau\colon P(A) \to \mathbf{A}_{\leqq},\quad \tau(B) = \sup B \text{ for all } B \subseteq A,$$

are order preserving and such that $\sigma\circ\tau = 1_{\mathbf{A}_{\leqq}}$. Hence (iii) holds.

Lemma 5.3. *The variety generated by a preprimal algebra $\mathbf{A}_{\{b\}}$ is category equivalent to $V_{\mathbf{C}_3}$.*

Proof. By Theorem 3.6 we have $\mathcal{K} = \mathrm{ISP}(\mathbf{A}_{\{b\}}) = V_{\mathbf{A}_{\{b\}}}$. For $\mathbf{C} = \mathbf{C}_3 = \langle\{0, 1\};\ 0\rangle$, $\mathbf{Q} = \mathbf{A}_{\{b\}}$, $\mathbf{Q} = \mathbf{A}_{\{b\}} = \langle A;\ b\rangle$, conditions (i), (ii) (a), and (iv) hold. $\mathbf{C}_3$ is isomorphic to a substructure of $\mathbf{A}_{\{b\}}$ consisting of b and any other element of A. Hence (ii) (b) holds. We choose a positive integer n such that $|A| \leqq 2^n$. Then there exist a monomorphism $\sigma\colon \mathbf{A}_{\{b\}} \to \langle\{0, 1\}^n;\ 0\rangle$ and an epimorphism $\tau\colon \langle\{0, 1\}^n;\ 0\rangle \to \mathbf{A}_{\{b\}}$ such that $\sigma\circ\tau = 1_{\mathbf{A}_{\{b\}}}$. Hence (iii) holds.

Lemma 5.4. *The variety generated by a preprimal algebra $\mathbf{A}_{s_2}$ is category equivalent to $V_{\mathbf{D}_3}$.*

Proof. By Theorem 3.6, we have $\mathcal{K} = \mathrm{ISP}(\mathbf{A}) = V_{\mathbf{A}_{s_2}}$. For $\mathbf{C} = \mathbf{D}_3 = \langle\{0, 1\};\ N\rangle$, $\mathbf{Q} = \mathbf{A}_{s_2}$, $\mathbf{Q} = \mathbf{A}_{s_2} = \langle A;\ N\rangle$, conditions (i), (ii) (a), and (iv) hold. $\mathbf{C}_3$ is isomorphic to a substructure of $\mathbf{A}_{s_2}$ consisting of any two elements a, b, $a \neq b$, of A with $Na = b$, $Nb = a$ ($|A| = 2k$). Hence (ii) (b) holds. We choose n such that $|A| \leqq 2^n$. Without restriction of generality we choose $\mathbf{A}_{s_2} = \langle\{0, 1, \ldots, 2k-1\};\ N\rangle$ with $N = (01)(23)\ldots(2k-1\ 2k)$, and $\mathbf{2}^n = \langle\{a_0, a_1, \ldots, a_{2^n-1}\},\ N\rangle$. Then we can define a monomorphism $\sigma\colon \mathbf{A}_{s_2} \to \mathbf{2}^n$ by $\sigma(i) = a_i$, $i = 0, \ldots, 2k-1$, and an epimorphism $\tau\colon \mathbf{2}^n \to \mathbf{A}_{s_2}$ by $\tau(a_i) = i$ for $i = 0, \ldots, 2k-1$ and $\tau(a_{2k+i}) = i$ for $i = 0, \ldots$ $\ldots, 2^n - 2k$ such that $\sigma\circ\tau = 1_{\mathbf{A}_{s_2}}$. Hence (iii) holds.

Lemma 5.5. *A quasivariety $\mathcal{K}$ is category equivalent to the quasivariety generated by $\mathbf{L}_1$ if and only if it is generated by a preprimal algebra $\mathbf{A}_{\alpha_m}$.*

Proof. Let $\mathcal{L} = \mathrm{ISP}(\mathbf{L}_1)$ be the quasivariety generated by $\mathbf{L}_1$. By Theorem 4.4, for $\mathbf{C} = \mathbf{L}_1 = \mathbf{G}_3 = \langle\{0, 1\};\ x+y+z\rangle$, $\mathcal{R} = \mathrm{ISP}(\mathbf{L}_1)$ the protoduality (D, E) is a full duality between $\mathcal{L}$ and $\mathcal{R}_0$, and $\mathbf{L}_1$ is injective in $\mathcal{R}_0$.

Let $\mathcal{K}$ be equivalent to $\mathcal{L} = \mathrm{ISP}(\mathbf{L}_1)$. Then by Theorem 5.1 (i), there exist a finite algebra $\mathbf{Q}$ in $\mathcal{K}$ and a family R of compatible relations of $\mathbf{Q}$ such that $\mathbf{Q} = \langle Q;\ R\rangle$ is an object of $\mathcal{R}_0$, i.e. $\mathbf{Q}$ is a Boolean 3-group and therefore $\mathbf{Q}$ is a finite power of the two-element Boolean 3-group. By (iv), $\mathbf{Q}$ is a preprimal algebra of the form $\mathbf{A}_{\alpha_m}$ with $\alpha_m = \{(x, y, z, e)\colon e = x+y+z\}$ and $x+y+z$ the operation of a Boolean 3-group $\mathbf{G}_3^m = \langle A;\ x+y+z\rangle$, $|A| = 2^m$, $m > 1$. Conversely, let $\mathrm{ISP}(\mathbf{A}_{\alpha_m})$ be the quasivariety generated by $\mathbf{A}_{\alpha_m}$. Taking $\mathbf{Q} = \mathbf{A}_{\alpha_m}$, $\mathbf{Q} = \mathbf{G}_3^m$, (i), (ii) (a), (b), and (iv) are satisfied. Since G_3 is injective in $\mathscr{R}_0$, $\mathbf{Q}=\mathbf{G}_3^m$ also is injective in $\mathscr{R}_0$. Hence (iii) holds and $\mathrm{ISP}(\mathbf{A}_{\alpha_m})$ is equivalent to $\mathrm{ISP}(\mathbf{L}_1)$.

Finally, by Lemmas 5.2—5.5 and Theorem 3.6 we obtain

Theorem 5.6. *A quasivariety is category equivalent to the quasivariety generated by a two-element preprimal algebra iff it is generated by a preprimal algebra of one of the forms* $\mathbf{A}_{\leq}$, $\mathbf{A}_{\{b\}}$, $\mathbf{A}_{s_2}$ $(|A|=2k)$, $\mathbf{A}_{\alpha_m}$ $(|A|=2^m)$.

References

[1] K. A. Baker and A. F. Pixley, Polynomial interpolation and the Chinese Remainder Theorem for algebraic systems, *Math. Z.*, **143** (1975), 165—174.

[2] G. Birkhoff, Some applications of universal algebra, in: *Universal Algebra* (Proc. Conf. Esztergom, 1977), Colloq. Math. Soc. János Bolyai, vol. 26, North-Holland (Amsterdam, 1982); pp. 107—128.

[3] B. A. Davey and H. Werner, Dualities and equivalences for varieties of algebras, in: *Contributions to Lattice Theory* (Proc. Conf. Szeged, 1980), Colloq. Math. Soc. János Bolyai, vol. 33, North-Holland (Amsterdam, 1983); pp. 101—275.

[4] K. Denecke, *Algebraische Fragen nichtklassischer Aussagenkalküle*, Dissertation, Päd. Hochschule Potsdam (1978).

[5] K. Denecke, *Preprimal Algebras*, Akademie-Verlag (Berlin, 1982).

[6] K. Denecke and L. Szabó, Characterization of algebras having majority term functions and admitting no proper subalgebras and no nontrivial tolerance relations, in: *Lectures in Universal Algebra* (Proc. Conf. Szeged, 1983), Colloq. Math. Soc. J. Bolyai, vol. 43, North-Holland (Amsterdam, 1986), pp. 497—506.

[7] G. Grätzer, *Universal Algebra*, 2nd ed., Birkhäuser-Verlag (Basel, 1979).

[8] H.-J. Hoehnke, Struktursätze der Algebra und Kompliziertheit logischer Schemata. I, *Math. Nachr.*, **61** (1974), 15—53.

[9] T. K. Hu, Stone duality for primal algebra theory, *Math. Z.*, **110** (1969), 180—198.

[10] B. Jónsson, Algebras whose congruence lattices are distributive, *Math. Scand.*, **21** (1967), 110—121.

[11] E. L. Post, *The two-valued iterative systems of mathematical logic*, Ann, Math. Studies, vol. 5, Amer. Math. Soc. (Princeton, 1941).

[12] I. Rosenberg, Über die funktionale Vollständigkeit in den mehrwertigen Logiken, *Rozpr. ČSAV, Řada Math. Přír. Věd*, **80** (1970), 3—93.

[13] W. Taylor, Characterizing Mal'cev conditions, *Algebra Universalis*, **3** (1977), 351—391.

SEKTION MATHEMATIK/PHYSIK
PÄDAGOGISCHE HOCHSCHULE "KARL LIEBKNECHT"
AM NEUEN PALAIS
1571 POTSDAM, GDR

Acta Sci. Math., 52 (1988), 81—84

Examples of local uniformity of congruences

IVAN CHAJDA

Following [6], a congruence Θ on an algebra A is *uniform* if every two congruence classes of Θ have the same cardinality. An algebra A is *uniform* if each $\Theta \in \mathrm{Con}\, A$ has this property. A class of algebras is *uniform* if every algebra of this class has this property.

It is well known that groups and Boolean algebras are uniform. Moreover, every variety generated by quasi-primal algebras (i.e. a discriminator variety, see [7]) is uniform, see [6] or Theorem 2.2 in [7]. Some classes of uniform algebras are depicted also in [3]. Although such "nice" varieties are uniform, W. TAYLOR [6] proved that the class of uniform varieties is not definable by a Mal'cev condition. He introduces the following concept: an algebra A is *weakly uniform* if for every cardinal $\mathfrak{m}$ there exists a cardinal $\mathfrak{n}$ such that whenever B_1 and B_2 are congruence classes of some $\Theta \in \mathrm{Con}\, A$, if $\mathrm{card}\, B_1 \leqq \mathfrak{m}$ then $\mathrm{card}\, B_2 \leqq \mathfrak{n}$. It was proven in [6] that the class of varieties of weakly uniform algebras is definable by a Mal'cev condition.

For algebras with a nullary operation, we can give a local version of uniformity:

Definition. An algebra A with a nullary operation c is *c-locally uniform* if for each element $a \in A$ and each $\Theta \in \mathrm{Con}\, A$, $\mathrm{card}\, [a]_\Theta \leqq \mathrm{card}\, [c]_\Theta$. A class K of algebras of the same type with a nullary operation c is *c-locally uniform* if each $A \in \mathscr{K}$ has this property.

It is clear that every uniform algebra with a nullary operation c is c-locally uniform and every c-locally uniform algebra is weakly uniform with $\mathfrak{n} = \mathrm{card}\, [c]_\Theta$.

Recall that an algebra A is *regular* if every two congruences on A coincide whenever they have a congruence class in common. An algebra A with a nullary operation c is *weakly regular (with respect to c)* if every two congruences $\Theta, \Phi \in \mathrm{Con}\, A$ coincide whenever $[c]_\Theta = [c]_\Phi$. A class $\mathscr{K}$ of algebras is *(weakly) regular* if each $A \in \mathscr{K}$ has this property.

Received January 23, 1985.

Proposition (Lemma 2.6 in [5]). *Every uniform variety is regular.*

We can prove a similar result for c-locally uniform algebras.

Theorem 1. Let $\mathcal{K}$ be a class of algebras of the same type with a nullary operation c closed under homomorphic immages. If $\mathcal{K}$ is c-locally uniform, then $\mathcal{K}$ is weakly regular with respect to c.

Proof. Let c be a nullary operation of a c-locally uniform class $\mathcal{K}$. Let $\mathcal{K}$ be closed under homomorphic immages. Suppose $A \in \mathcal{K}$, $\Theta_1, \Theta_2 \in \mathrm{Con}\, A$ and

$$(*) \qquad [c]_{\Theta_1} = [c]_{\Theta_2}.$$

In this case we have clearly $[c]_{\Theta_1 \wedge \Theta_2} = [c]_{\Theta_1 \vee \Theta_2} = [c]_{\Theta_1} = [c]_{\Theta_2}$; without loss of generality, we can assume $\Theta_1 \leqq \Theta_2$. Denote by ω the identity relation on A/Θ_1. By $(*)$, the congruences $\omega = \Theta_1/\Theta_1$ and Θ_2/Θ_1 of $A/\Theta_1 \in \mathcal{K}$ have the same congruence class containing the nullary operation $[c]_{\Theta_1}$ of A/Θ_1. Thus

$$\mathrm{card}\, [c]_{\Theta_2/\Theta_1} = \mathrm{card}\, [c]_\omega = 1.$$

Since A/Θ_1 is c-locally uniform, we have

$$1 \leqq \mathrm{card}\, [a]_{\Theta_2/\Theta_1} \leqq \mathrm{card}\, [c]_{\Theta_2/\Theta_1} = 1$$

for each $a \in A$, thus $\Theta_2/\Theta_1 = \omega$, i.e. $\Theta_2 = \Theta_1$.

The aim of this paper is to show that there exist important classes of finite algebras which are c-locally uniform but not uniform. By Theorem 1, they must be weakly regular. By [4], such algebras can be found among Heyting algebras, implication algebras and other types of lattice ordered algebras with pseudocomplementation.

An algebra $\langle L; \vee, \wedge, \cdot, 1\rangle$ with three binary and one nullary operations is an *rp-algebra* if $\langle L; \vee, \wedge, 1\rangle$ is a lattice with greatest element 1 and $\cdot$ satisfies the following identities:

$$(**) \qquad x \cdot x = 1, \quad (x \cdot y) \wedge y = y, \quad (x \cdot y) \wedge x = x \wedge y.$$

Theorem 2. *The class of all finite rp-algebras is* 1*-locally uniform but not uniform.*

Proof. Let $\mathcal{K}$ be a class of all finite rp-algebras. Clearly $\mathcal{K}$ is not uniform, because, e.g. the three-element chain $C = \{0, a, 1\}$, $0 < a < 1$, with a binary operation $\cdot$ defined by

$a \cdot 0 = 0$, $1 \cdot 0 = 0$ and $x \cdot y = 1$ for all other combinations of variables

is an rp-algebra but the partition $\{0\}$, $\{a, 1\}$ forms a congruence on C which is not uniform.

We prove that $\mathscr{K}$ is 1-locally uniform. Let $A\in\mathscr{K}$, $z\in A$, $\Theta\in\mathrm{Con}\, A$. Since A is a finite lattice, the congruence class $[z]_\Theta$ contains a greatest element a. Put $\varphi(x)=a\cdot x$. We prove that φ is an injection of $[z]_\Theta$ into $[1]_\Theta$. If $x\in[z]_\Theta$, then $\langle x, a\rangle\in\Theta$. Since φ is an algebraic function over A, it follows that $\langle\varphi(x), \varphi(a)\rangle=\langle\varphi(x), a\cdot a\rangle=\langle\varphi(x), 1\rangle\in\Theta$, i.e. $\varphi(x)\in[1]_\Theta$. Thus $\varphi\colon [z]_\Theta\to[1]_\Theta$. Suppose $\varphi(x)=\varphi(y)$ for $x, y\in[z]_\Theta$. Then $a\cdot x=a\cdot y$, whence $a\wedge(a\cdot x)=a\wedge(a\cdot y)$. By $(**)$, this yields $a\wedge x=a\wedge y$. Since $x\leqq a, y\leqq a$, we obtain $x=y$. Thus φ is an injection, and therefore $\mathrm{card}\,[z]_\Theta\leqq\mathrm{card}\,[1]_\Theta$.

Let L be a lattice and $a, b\in L$. An element $x\in L$ is called a *relative pseudocomplement of a with respect to b if x* is the greatest element satisfying $a\wedge x=a\wedge b$; denote it by $a*b$. A lattice L is *relatively pseudocomplemented* if $a*b$ exists for each $a, b\in L$. Then clearly L has a greatest element 1, and $a*a=1$ for each $a\in L$. Clearly the operation $*$ satisfies the identities $(**)$, i.e. we obtain the following

Corollary 1. *Every finite relatively pseudocomplemented lattice is* 1-*locally uniform.*

Note that a finite lattice is relatively pseudocomplemented if and only if it is distributive. Corollary 1 implies immediately (for the definition, see e.g. [7])

Corollary 2. *Every finite Heyting algebra is* 1-*locally uniform.*

Remark. By [4], a Heyting algebra is regular if and only if it is a Boolean algebra. Every three-element chain $0<a<1$ with a pseudocomplementation is a Heyting algebra which is not uniform.

Following [1], an algebra $\langle A; \cdot\rangle$ with one binary operation is an *implication algebra* if it satisfies

$$(x\cdot y)\cdot x = x, \quad (x\cdot y)\cdot y = (y\cdot x)\cdot x, \quad x\cdot(y\cdot z) = y\cdot(x\cdot z).$$

As it was proven in [1], every implication algebra A has a nullary operation 1 such that $a\cdot a=1$ for each $a\in A$.

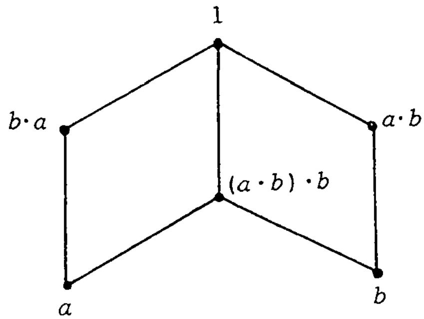
Lemma 1. *Every implication algebra is a* $\vee$-*semilattice with greatest element* 1 *with respect to the operation* $a\vee b=(a\cdot b)\cdot b$.

For the proof, see Theorem 3 and Theorem 4 in [1].

Lemma 2. (Theorem 5 in [1]). *Let A be an implication algebra and* $a, b\in A$. *If p is any lower bound for a and b (with respect to the semilattice ordering), then the infimum* $a\wedge b$ *of a and b exists, and* $a\wedge b=[a\cdot(b\cdot p)]\cdot p$.

Theorem 3. *The class of all finite implication algebras is* 1-*locally uniform but not uniform.*

Proof. Let A be free implication algebra with two free generators a, b. By the Corollary of Theorem 2 in [1], A has the following diagram (as a $\vee$-semilattice):



Clearly the equivalence Θ given by the partition

$$\{b\cdot a, 1\},\quad \{a, (a\cdot b)\cdot b\},\quad \{a\cdot b\},\quad \{b\}$$

is a congruence on A which is not uniform.

Let A be a finite implication algebra, $\Theta\in\text{Con } A$ and $z\in A$. By Lemma 1, there exists a greatest element a in $[z]_\Theta$. Put $\varphi(x)=a\cdot x$. Clearly $\varphi(a)=a\cdot a=1$. If $x\in[z]_\Theta$, then $\langle x, a\rangle\in\Theta$ which implies $\langle\varphi(x), \varphi(a)\rangle=\langle\varphi(x), 1\rangle\in\Theta$, i.e. $\varphi(x)\in[1]_\Theta$. Thus φ is a mapping of $[z]_\Theta$ into $[1]_\Theta$.

We prove that φ is an injection on $[z]_\Theta$. Suppose $x, y\in[z]_\Theta$ and $\varphi(x)=\varphi(y)$. Then $a\cdot x=a\cdot y$. Since $x\leqq a$, $x\leqq a\cdot x$ and $y\leqq a$, $y\leqq a\cdot y$, therefore x is a lower bound of a and $a\cdot x$, y, is a lower bound of a and $a\cdot y$. By Lemma 2, $a\wedge a\cdot x$ and $a\wedge a\cdot y$ exist, and $a\cdot x=a\cdot y$ implies that $a\cdot x\wedge a=a\cdot y\wedge a$. By Lemma 2, $a\cdot x\wedge a= =[(a\cdot x)(a\cdot x)]\cdot x=1\cdot x=x$, and analogously $a\cdot y\wedge a=y$. Hence $x=y$.

References

[1] J. C. Abbott, *Semi-Boolean algebra, Mat. Vestnik*, 4 (19) (1967), 177—198.

[2] S. Burris, H. P. Sankappanavar, *A Course in Universal Algebra*, Springer-Verlag (Berlin—Heidelberg—New York, 1981).

[3] I. Chajda, Notes on the Big Dipper identity, *Ann. Univ. Sci. Budapest, R. Eötvös Nom.*, **29** (1986), 97—101.

[4] J. C. Varlet, Regularity in p-algebras and p-semilattices, in: *Banach Center Publ.*, Vol. 9, Polish Sci. Publ. (Warszawa, 1982); pp. 369—382.

[5] R. W. Quackenbush, Algebras with minimal spectrum, *Algebra Univesalis*, **10** (1980), 117—129.

[6] W. Taylor, Uniformity of congruences, *Algebra Universalis*, **4** (1974), 342—360.

[7] H. Werner, *Discriminator algebras*, Studien zur Algebra und ihre Anwendungen, Band 6, Akademie-Verlag (Berlin, 1979).

TŘÍDA LIDOVÝCH MILICÍ 22
750 00 PŘEROV, CZECHOSLOVAKIA

Acta Sci. Math., 52 (1988), 85—92

On the category of S-posets

SYED M. FAKHRUDDIN

0. Introduction. Generalizing usual posets as well as semilattices both of which have been treated from categorical viewpoint in [5] and [8], we study in this article the category of posets acted on by a pomonoid S and the action satisfying the usual properties. Our main results are:

(i) adjunctions from our category to the category of usual posets,

(ii) a structure theorem for projective S-posets, and finally

(iii) if S is a pogroup, our category admits injective hulls.

1. The category of S-posets — SPOS. Let S be a pomonoid which is not necessarily commutative and let E be a poset. We call E a left *S-poset* (the adjective "left" would be omitted in the sequel) if S acts on E in such a way that (i) the action is monotonic in each of the variables, (ii) for $s, t \in S$ and $x \in E$ we have $s(tx)=(st)x$ and (iii) $ex=x$ where e is the identity of S and sx stands for the result of the action of s on x. Let us call such an order on E an *S-order*. A *morphism* from an S-poset E to another S-poset F is a monotonic map which preserves S-action. The class of S-posets and morphisms evidently forms a category, which we denote by SPOS.

2. Congruences in SPOS. An equivalence relation Θ on an S-poset E is called a *congruence* if Θ is compatible with the S-action on E and the quotient set E/Θ can be endowed with an S-order so that the canonical surjection is a morphism in SPOS. Let now Θ be an equivalence on E compatible with S-action and $\mathbf{A}=\{A_1, A_2, \ldots, A_n\}$ be a finite sequence of distinct equivalence classes of Θ. $\mathbf{A}$ is called a *Θ-chain* if each class in $\mathbf{A}$ contains an element which is smaller than some element of the following class. Then Θ is a *congruence iff no element belonging to a member in a Θ-chain is smaller than an element of a previous member in that chain* (cf. [4], p. 177 or [1], p. 42).

If Θ is a congruence on E then the induced S-order on E/Θ is given by $[a]\leqq[b]$ iff there is a Θ-chain from $[a]$ to $[b]$. Moreover, every equivalence relation R on E

Received October 10, 1984, and in revised form April 15, 1987.

compatible with S-action generates a congruence Θ_R: $a\,\Theta_R b$ iff there is an R-chain from $[a]$ to $[b]$ and another from $[b]$ to $[a]$ (cf. [1], p. 46 or [4], p. 182). Finally if Θ_1 and Θ_2 are congruences on E such that $\Theta_1 \leqq \Theta_2$ then the unique map $E/\Theta_1 \to E/\Theta_2$ is a morphism.

3. Standard constructions in SPOS. Let E be an S-poset. The usual notions of *S-subposet, S-subposet generated by a subset X of E and convex S-subposet,* etc., can be defined in the obvious way. The convex S-subposet generated by X will be denoted by $\langle X \rangle$.

The usual definitions of monics and epimorphisms carry over to SPOS. However, an epimorphism need not be surjective. Let $f\colon E \to F$ be a surjective morphism. Then one can check that $\ker f = \Theta$ is a congruence over E and E/Θ, equipped with the smallest order making the natural surjection $E \to E/\Theta$ an order-preserving map, is isomorphic to F. Let $\{E_i\}_{i\in I}$ be a family of S-posets. Then the categorical *product* is the usual cartesian product with product order and the *coproduct* is the disjoint union.

The *equalizer* of $f, g\colon E \to F$ is $j\colon G \to E$ where j is the natural injection and G is given by $G = \{x \mid x \in E\colon f(x) = g(x)\}$. The *coequalizer* is $F \to F/\Theta$ where Θ is the congruence generated by the binary relation R over F, where aRb iff there exists an $x \in E$ such that the sets $\{f(x), g(x)\}$ and $\{a, b\}$ are the same in F.

By [7], Theorem 1 and its dual on page 109 we have

3.1. Theorem. SPOS *has arbitrary limits and colimits.*

There is another construction which is peculiar to SPOS. Given a family of S-posets $\{E_i\}_{i\in P}$ indexed by a poset P, the *ordinal sum* $\coprod^{\circ}_{i\in P} E_i$ of the family is the disjoint union and obvious S-action; the order relation is now given for $x, y \in \coprod^{\circ}_{i\in P} E_i$ by $x < y$ if $x \in E_i$ and $y \in E_j$ with $i < j$ or else $x \leqq y$ in $E_i = E_j$. The ordinal sum has the universal mapping property (UMP): given a family $f_i\colon E_i \to F$ of morphisms such that for $x \in E_i$, $y \in E_j$ with $f_i(x) \leqq f_j(y)$ in F, there exists a unique morphism $f\colon \coprod^{\circ}_{i\in P} E_i \to F$ with $f \cdot j_i = f_i$ where j_i is the canonical injection of E_i into the ordinal sum.

4. Free S-**posets.** Let P be a poset. Then a *free S-poset over P* is a pair (E, φ) where E is an S-poset and $\varphi\colon P \to E$ is a monotonic map such that for every monotonic map $\psi\colon P \to F$ into an S-poset F, there is a unique morphism $f\colon E \to F$ such that $\psi = f \cdot \varphi$.

4.1. Theorem. *Given a poset P there exists a free S-poset E over P and E is unique up to isomorphism.*

Proof. Let $E=\coprod^{\circ}_{i\in P} S_i$ where $S_i=S$ for each i and $\varphi\colon P\to E$ given by $\varphi(i)=e$, the identity of S in S_i. Then the UMP of the ordinal sum implies the UMP for the pair (E, φ) as given above and the uniqueness is clear.

We shall denote the free S-poset over P by $\mathbf{F}(P)$. It is isomorphic to $P\times S$ where $P\times S$ is the product poset with lexicographic order and S-action only on the second component.

The subset $B=\{b_i=(i, e)\}$ of $P\times S$ has the property: every element of $P\times S$ is a unique multiple of one (and only one) member of B and if $b_i<b_j$ then for any $s, t\in S$, $sb_i<tb_j$. If we call such a family an *ordered base*, then clearly an S-poset E is free over a poset P iff E has an ordered base $\{x_i\}_{i\in P}$ indexed by P. In this case $Sx_i=\langle x_i\rangle$. The poset P is called the *order type* of the free S-poset E. Then two free S-posets E and F are isomorphic in SPOS exactly if their order types are isomorphic in POS — the category of posets.

Not all S-posets are free even if S is a pogroup. For example, let E be the set Z of all integers and S be the full permutation group of E. Then S acquires a poorder from the natural order of Z and the resulting S-poset E is not free.

Let E be an S-poset. Consider the free S-poset over the poset E, $\mathbf{F}(E)=E\times S$ with the map $\varphi\colon E\to E\times S$ defined by $\varphi(x)=(x, e)$, then there is a unique morphism $\Pi\colon E\times S\to E$ defined by $\Pi((x, s))=sx$ such that $\Pi\cdot\varphi=I_E$ and we have $\mathbf{F}(E)/\ker\Pi\cong E$. Hence

4.2. Proposition. *Every S-poset is the quotient of a free S-poset.*

Remark. For a systematic study of standard constructions in ordered algebras we refer the reader to [2] and [3].

5. Some functors. An ordinary poset can be considered as an S-poset with trivial S-action. Let POS denote the category of posets and $\mathbf{U}$ be the inclusion functor. In this section, we shall find a left adjoint $\mathbf{H}$ to $\mathbf{U}$ and study the properties of the resulting adjunction.

First observe that a morphism from an S-poset to a poset is just a monotonic map which is constant on each orbit Sx for $x\in E$.

5.1. Proposition. *Let E be an S-poset. Then there is a poset $\mathbf{H}(E)$ and a morphism $h_E\colon E\to\mathbf{H}(E)$ such that for any morphism $f\colon E\to X$ into a poset X, there exists a unique monotonic map $\hat{f}\colon \mathbf{H}(E)\to X$ with $\hat{f}\cdot h_E=f$.*

Proof. Let Θ be the congruence on E generated by the binary relation aRb iff there exists $s\in S$ such that $sa=b$. More specifically define $x\Theta y$ for $x, y\in E$ if there exist elements $x=a_0, a_1, \ldots, a_n=y$ such that $Sa_i\cap Sa_{i+1}\neq\emptyset$. Let $\mathbf{H}(E)=E/\Theta$ and h_E be the natural morphism: $E\to E/\Theta$. Suppose $f\colon E\to X$ is a morphism into a poset X. Then clearly $\ker f=\Theta$ and so there exists a monotonic map $\hat{f}$: $\mathbf{H}(E)\to X$ such that $\hat{f}\cdot h_E=f$.

Now if f: $E\to F$ is a morphism in SPOS, the above construction implies that there exists a unique montonic map f_H: $\mathbf{H}(E)\to\mathbf{H}(F)$ such that $f_H\cdot h_E=h_F\cdot f$ and the correspondence $(\mathbf{H}(\ ),\ (\ )_H)$ defines a functor from SPOS to POS. Let $\eta_E=U\cdot h_E$: $E\to\mathbf{UH}(E)$ be the natural homomorphism. Then we have

5.2. Theorem. **H** *is left adjoint to* **U**.

Proof. The correspondence η: $I_{\mathrm{SPOS}}\to\mathbf{UH}$ is clearly a natural transformation such that η_E: $E\to\mathbf{UH}(E)$ is universal $\to$ from E to U for every E in SPOS. Then the assignment $\varphi f=\mathbf{U}f\cdot\eta_E$: $E\to\mathbf{U}(X)$ for f: $\mathbf{H}(E)\to X$ establishes a bijective correspondence between the respective hom-sets. Now the theorem follows (by [7], Theorem 2, condition (i), p. 81).

Remark. The unit of this adjunction is η and the counit ε: $\mathbf{HU}\to I_{\mathrm{POS}}$ is the natural order isomorphism.

Now let us discuss the associated monad of this adjunction ([7], p. 134). This is given by $\langle\mathbf{UH};\ \eta\colon I_{\mathrm{SPOS}}\to\mathbf{UH};\ \mu\colon \mathbf{UHUH}\to\mathbf{UH}\rangle$ where μ assigns to every object E in SPOS the map $\mathbf{U}\cdot\varepsilon_{\mathbf{H}(E)}$: $\mathbf{UHUH}(E)\to\mathbf{UH}(E)$ given by the rule $[[x]]$ mapped to $[x]$ for each $x\in E$.

If $\langle T,\eta,\mu\rangle$ is a monad in a category X, then an Eilenberg—Moore algebra (in short: EM-algebra) is a pair (x, h) where x is an object (the underlying object of the algebra) and h is an arrow h: $Tx\to x$ of X (called the structure map of the algebra) with the following properties:

(i) hTh: $T^2x\to x$ is the same as $h\cdot\mu_x$ (associative law),

(ii) $h\cdot\eta_x$: $x\to T_x\to x$ is the identity on x ([7], p. 136).

Hence applying this general definition to our situation, we find that an Eilenberg—Moore algebra for the monad above is a pair (E, g) where E is an S-poset and g is a left inverse for h_E such that the associative law above holds.

A morphism f: $(E, g)\to(E', g')$ of Eilenberg—Moore algebras is a morphism in SPOS such that $g'\cdot f_k=f\cdot g$.

Now consider the category of EM-algebras $(\mathrm{SPOS})^T$. This gives rise to an adjunction $\langle\mathbf{H}^T,\mathbf{U}^T,\eta^T,\varepsilon^T\rangle$: $(\mathrm{SPOS})\to(\mathrm{SPOS})^T$ in which $\mathbf{H}^T$ and $\mathbf{U}^T$ are given by the respective assignments

$$\mathbf{U}^T\ \begin{array}{ccc}(E,g) & \mapsto & E\\ \downarrow f & & \downarrow f\\ (E',g') & \mapsto & E'\end{array}\quad\text{and}\quad \mathbf{H}^T\ \begin{array}{ccc}E & \mapsto & (\mathbf{UH}(E),\mu_E)\\ \downarrow f & & \downarrow \mathbf{UH}(f)\\ E' & \mapsto & (\mathbf{UH}(E'),\mu)\end{array}$$

and $\eta^T=\eta$ and $\varepsilon^T(E, g)=g$ for each algebra in $(\mathrm{SPOS})^T$ (cf. [7], Theorem 1, p. 136). The monad defined by this new adjunction on SPOS is the same as the original monad.

Also, this new adjunction is related to the original adjunction by the comparison functor. This functor **K** is from POS to $(\mathrm{SPOS})^T$ with $\mathbf{U}^T\mathbf{K}=\mathbf{U}$ and $\mathbf{KH}=\mathbf{H}^T$. This is defined by $\mathbf{K}(P)=\langle \mathbf{U}(P), \mathbf{U}\varepsilon_P\rangle$ for any P in POS and $\mathbf{K}(f)=\mathbf{U}(f)$: $\langle \mathbf{U}(P), \mathbf{U}\varepsilon_P\rangle \to \langle \mathbf{U}(P), \mathbf{U}\varepsilon_{P'}\rangle$ for any morphism f in POS ([7], Theorem 1, pp. 138, 139). When this functor **K** is an isomorphism, the functor **U** is called *monadic*. In the present case **U** is indeed monadic, and we shall indicate the proof.

A functor $G: A \to X$ *creates coequalizers* for a parallel pair $f, g: a \to b$ in A when to each coequalizer $u: Gb \to z$ of Gf, Gg in X there is a unique object c and a unique arrow $e: b \to c$ with $Gc=z$ and $Ge=u$ and when, moreover, this unique arrow is a coequalizer of f and g. Also a fork $a \overset{f}{\underset{g}{\rightrightarrows}} b \to c$ in a category is called an *absolute coequalizer* if it remains a coequalizer under the action of any functor. Hence in particular it is a coequalizer. By Beck's theorem ([7], Theorem 1, p. 147), the functor $\mathbf{U}$: POS$\to$SPOS is monadic iff **U** creates coequalizers for those parallel pairs f, g in POS for which $\mathbf{U}f$ and $\mathbf{U}g$ has an absolute coequalizer in SPOS. Now this is easily verified, since a coequalizer is surjective both in POS and SPOS. Hence we have

5.3. Theorem. *The inclusion functor* **U** *is monadic.*

On the other hand, let **F** be the free S-poset construction. Then it is easily seen that **F** defines a functor from POS to SPOS and let **V** be the forgetful functor from SPOS to POS. The map $\Phi_P: P \to \mathbf{F}(P)$ associated with **F** is a natural transformation from I_{POS} to **F**. Let $\delta_P = \mathbf{V}\Phi_P: P \to \mathbf{VF}(P)$. Then δ_P is a universal arrow from P to **V**. Hence we conclude

5.4. Theorem. **F** *is left adjoint to* **V**.

The unit of this adjunction is δ and the counit is the canonical epimorphism (Prop. 4.2) $\Pi: \mathbf{FV} \to I_{\mathrm{SPOS}}$. The associated monad is given by $\langle \mathbf{VF}, S: I_{\mathrm{POS}} \to \mathbf{VF}, \sigma: \mathbf{VFVF} \to \mathbf{VF}\rangle$ where σ assigns to every object P in POS the map $\mathbf{V}\Pi_{F(P)}$ from $\mathbf{VFVF}(P) \to \mathbf{VF}(P)$ given by the rule $((p, e), e)$ of $\mathbf{VFVF}(P)$ is mapped into (p, e) of $\mathbf{VF}(P)$ for P in POS.

Now an Eilenberg—Moore algebra for the monad above is a pair (P, h) where P is a poset and $h: \mathbf{VF}(P) \to P$ is a left inverse for $\delta: P \to \mathbf{VF}(P)$. Using the method of ([7], Theorem 1, p. 152) we can show

5.5. Theorem. *The forgetful functor* **V** *is monadic.*

Summarising we have

5.6. Theorem. *In* SPOS *the functor* **UV** *which trivialises S-action has a left adjoint* **FH**.

6. Projective *S*-posets. An S-poset E is called *projective* if every epimorphism to E is a retraction.

By 4.2 an S-poset E is projective iff it is a retract of a free S-poset. The following theorem gives a characterisation of a projective S-poset similar to the one valid for projective modules.

6.1. Theorem. *Let E be an S-poset. Then E is projective iff there exist maps $h\colon E\to E$, $g\colon E\to S$ with the following properties:*

(i) *h is a monotonic map which is constant on $\langle x\rangle$ for $x\in E$.*

(ii) *g preserves S-action, and if $x, y\in E$, $x\leqq y$, $h(x)=h(y)$ then $g(x)\leqq g(y)$.*

(iii) *$g(x)h(x)=x$ for every $x\in E$.*

Proof. Suppose E is projective. Then $\Pi\colon E\times S\to E$ given by $\Pi((x, s))=sx$ is a retraction, so there exists an S-morphism $f\colon E\to E\times S$ where $f(x)=(h(x), g(x))$ such that $\Pi((h(x), g(x)))=g(x)h(x)=x$ for every $x\in E$. Thus it remains only to check conditions (i) and (ii) above.

Since f is monotonic h is also monotonic. Also, if $y=tx$ for some $t\in S$ then $f(tx)=(h(tx), g(tx))=tf(x)=t(h(x), g(x))=(h(x), tg(x))$. Thus $h(tx)=h(x)$ and $g(tx)=tg(x)$. If, however, $ax\leqq y\leqq bx$ then $f(ax)\leqq f(y)\leqq f(bx)$. Therefore $(h(x), ag(x))\leqq(h(y), g(y))\leqq(h(x), bg(x))$. Thus $h(x)=h(y)$ and $ag(x)\leqq g(y)\leqq bg(x)$.

Conversely, given h and g a priori satisfying the above conditions, define $f\colon E\to E\times S$ by $f(x)=(h(x), g(x))$. Then by (iii) $(\Pi\cdot f)(x)=x$ for every $x\in E$. Also $f(tx)=(h(tx), g(tx))=(h(x), tg(x))=t(h(x), g(x))=tf(x)$ and if $x<y$ then $h(x)<h(y)$ or else $g(x)\leqq g(y)$. Then $(h(x), g(x))\leqq(h(y), g(y))$ and thus f is an S-morphism and E is a retract of a free S-poset, so it is projective.

Since the map h factors through $h_E\colon E\to \mathbf{H}(E)$, we have a different, but equivalent formulation of the theorem above.

6.2. Theorem. *Let E be an S-poset. Then E is projective iff there exist maps $h'\colon \mathbf{H}(E)\to E$ and $g\colon E\to S$ with the following properties:*

(i) *h' is a montonic map.*

(ii) *g preserves S-action and if $x\leqq y$, $h'([x])=h'([y])$, then $g(x)\leqq g(y)$.*

(iii) *$g(x)h'([x])=x$ for every $x\in E$, where $[x]$ is the class of x in $H(E)$ for $x\in E$.*

Example. If $E=X\times S$, the free S-poset over X, then $h\colon X\times S\to X\times S$ is given by $h((x, s))=(x, e)$ and $g\colon E\to S$ by $g((x, s))=s$.

Call an ideal J in S projective if J is a projective S-poset. Then we have

6.3. Theorem. *An S-poset E is projective iff E is isomorphic to an ordinal sum of the form $\coprod\limits_{i\in I}^{\circ} J_i z_i$ where z_i is a suitable element of E and J_i is a projective ideal of S with the property*

(i) *there exists an $s_i\in J_i$ such that $s_i z_i=z_i$, and*

(ii) *$a\leqq b$ in J_i exactly if $az_i\leqq bz_i$.*

Proof. Projective property is stable under isomorphism and ordinal sum, hence sufficiency is clear.

Conversely let E be projective and $h: E \to E$, $g: E \to S$ be the functions given in Theorem 6.1. Then the equivalence classes E_i of h are convex S-subposets. Let $h_i = h|_{E_i}$ and $g_i = g|_{E_i}$. By the previous result E_i is projective with the aid of the maps g_i and h_i; also h_i is constant on E_i and $(g_i h_i)(x) = x$, so g_i is an isomorphism. Thus $g_i(E_i) = J_i$ is a projective ideal of S and if $h_i(E_i) = z_i \in E_i$ then $J_i \cong J_i z_i = E_i$. Consider $E' = \coprod_{i \in I}^{\circ} E_i \cong \coprod_{i \in I}^{\circ} J_i z_i$. Then E' is projective and as sets $E = E'$. However, the identity map $x = g_i(x) z_i$ is a bimorphism, so in particular an epimorphism from E to E' and since E' is projective this is an isomorphism.

6.4. Corollary. *Over a pogroup G, all projective G-posets are free.*

7. Complete S-posets — completion — injectivity. An S-poset E is *complete* if E is a complete lattice and given a family of elements $\{x_i\}$ in E and $s \in S$ we have $s(\vee x_i) = \vee s x_i$ where $\vee$ denotes the supremum. A morphism between complete posets is *complete* if it preserves supremum of arbitrary family of elements.

A *completion* of an S-poset E is a pair (E^*, φ) where E^* is a complete S-poset and $\varphi: E \to E^*$ is a monomorphism with the property that $\varphi(x) < \varphi(y)$ exactly if $x < y$ in E and for any other pair (F, ψ) with the above data there exists a unique complete morphism $f: E^* \to F$ such that $f \cdot \varphi = \psi$.

7.1. Theorem. *Every S-poset E admits a completion, which is unique up to isomorphism.*

Proof. The proof is essentially the same as that of Theorem 2 in [6].

Now call a monomorphism f of an S-poset *strict* if $f(x) < f(y)$ exactly if $x < y$. An S-poset E is *injective* if given a strict monomorphism $g: A \to B$ and a morphism $f: A \to E$, there exists an extension of f to B, $h: B \to E$ such that $h \cdot g = f$.

7.2. Proposition. *An injective S-poset is complete.*

Proof. If E is an injective S-poset and (E^*, φ) its completion then by the definition applied to the identity morphism on E, φ is a coretraction. Hence E is already complete.

For a converse, we have

7.3. Proposition. *Let G be a pogroup. Then a complete G-poset is injective.*

Proof. Let E be a complete G-poset and $g: A \to B$ be a strict monomorphism of S-posets and $f: A \to E$ be a morphism. For $b \in B$ we define $h(b) = \bigvee_{g(a) \leq b} f(a)$. Now $h(b)$ exists in E and clearly h is monotonic; moreover, since g is strict, we have

$h\cdot g=f$. For $s\in G$ we have

$$h(sb) = \bigvee_{g(x)\leqq sb} f(x) \geqq \bigvee_{g(a)\leqq b} f(sa) = s\Big(\bigvee_{g(a)\leqq b} f(a)\Big) = sh(b).$$

Further $sh(b)=sh(s^{-1}(sb))\geqq ss^{-1}h(sb)$ which gives $sh(b)\geqq h(sb)$. Thus $h(sb)=$ $=sh(b)$ and h is an S-morphism.

Noting that a minimal injective extension is a hull, we have

7.4. Corollary. *If G is a pogroup, then the category of G-posets admits injective hulls.*

References

[1] T. S. Blyth and M. F. Janowitz, *Residuation theory*, Pergamon Press (Oxford, 1972).
[2] G. Czédli and A. Lenkehegyi, On a class of ordered algebras and quasi order distributivity, *Acta Sci. Math.*, **46** (1983), 41—54.
[3] G. Czédli and A. Lenkehegyi, On congruence n-distributivity of ordered algebras, *Acta. Math. Hung.*, **41** (1—2) (1983), 17—26.
[4] M. L. Dubriel-Jacotin, L. Lesieur and R. Croisot, Leçons sur la théorie des treillis des structures algebriques ordonnées et des treillis géométriques, Gauthier—Villars (Paris, 1953).
[5] A. Horn and N. Kimira, Category of semilattices, *Algebra Universalis*, **1** (1971), 26—38.
[6] V. S. Krishnan, Les algebras partiellement ordonnees et leurs extensions, *Bull. Soc. Math. France*, **78** (1950), 235—263.
[7] S. Mac Lane, Category for the Working Mathematician, Springer-Velag (Berlin—Heidelberg—New York, 1971).
[8] M. Sekanina, Category of posets, *Arch. Math.* (Brno), **4** (1968), 25—59.

DEPARTMENT OF MATHEMATICS
WEST VIRGINIA UNIVERSITY
MORGANTOWN, WV 26506, U.S.A.

Acta Sci. Math., 52 (1988), 93—100

On the minimal ring containing the boundary of a convex body

IMRE BÁRÁNY

1. Let $K \subset \mathbf{R}^2$ be a convex compact set with boundary C. For each point $x \in K$ there exist a minimal circular disc $B(R(x), x)$ containing K and a maximal circular disc $B(r(x), x)$ contained in K, where $B(r, x)$ denotes the disc with radius r and center x.

The function $R(x) - r(x)$ attaines its minimal value in a unique point $x_0 \in K$. This was shown by BONNESEN [1], Bonnesen and FENCHEL [2]. So the circular ring around x_0 with radii $R(x_0)$ and $r(x_0)$, respectively, is the *minimal ring* containing the boundary C of K.

This result was used by Bonnesen and Fenchel [2] to sharpen the isoperimetric inequality in $\mathbf{R}^2$. Later I. VINCZE [7] showed that

$$\frac{\min\{R(x):\ x \in K\}}{R(x_0)} \geqq \frac{\sqrt{3}}{2} \tag{1}$$

and

$$\frac{\max\{r(x):\ x \in K\}}{r(x_0)} < 2 \tag{2}$$

and these inequalities are sharp.

Answering a question due to I. Vincze we generalize the inequalities (1) and (2) to arbitrary dimension. To do so we need a theorem characterizing the minimal ring in $\mathbf{R}^d$. For $d=2$ and $d=3$ such a theorem was found by Bonnesen [1] and by KRITIKOS [4]. The main tool in the proof of our results is the use of convex analysis (see: Йоффе — Тихомиров [3] and ROCKEFELLAR [5]).

2. Again, let $K \subset \mathbf{R}^d$ be a convex compact set with boundary C. $B(r, x)$ stands for the ball with radius r and center x. For $x \in K$ we define

$$R(x) = \min\{R:\ B(R, x) \supseteqq K\},$$
$$r(x) = \max\{r:\ B(r, x) \subseteqq K\}.$$

Received January 18, 1985.

It is easy to see that the maximum and minimum above exist, so the definition is correct. Moreover, this means that for each $x \in K$ there exist points p and q such that $p, q \in C$ and $\|x-p\|=R(x)$ and $\|x-q\|=r(x)$. In this case we say that p supports $R(x)$ and q supports $r(x)$.

Theorem 1. *There exists a point $x_0 \in K$ in which the function $R(x)-r(x)$ attaines its minimal value. This point x_0 is unique.*

The set $\{x \in \mathbf{R}^d : r(x_0) \leqq \|x-x_0\| \leqq R(x_0)\}$ is called the *minimal ring* containing C. The characterization theorem for the minimal ring is this:

Theorem 2. *The point $x_0 \in K$ is the center of the minimal ring if and only if there are points $p_1, \ldots, p_k \in C$ supporting $R(x_0)$ and $q_1, \ldots, q_l \in C$ supporting $r(x_0)$ $(k, l \geqq 1)$ such that*

$$\operatorname{conv}\left\{\frac{p_i - x_0}{R(x_0)} : i = 1, \ldots, k\right\} \cap \operatorname{conv}\left\{\frac{q_i - x_0}{r(x_0)} : j = 1, \ldots, l\right\} \neq \emptyset,$$

where conv *denotes the convex hull.*

There is a certain converse to this theorem. We describe it when $x_0 = 0$.

Theorem 3. *Let $p_1, \ldots, p_k$, $q_1, \ldots, q_l$ be vectors in $\mathbf{R}^d$ such that*
(i) $\|p_1\| = \ldots = \|p_k\| = R \geqq r$,
(ii) $\|q_1\| = \ldots = \|q_l\| = r > 0$,
(iii) $\{p_i/R : i=1, \ldots, k\} \cap \operatorname{conv} \{q_j/r : j=1, \ldots, l\} \neq \emptyset$,
(iv) *each p_i is contained in the halfspaces*

$$\{x \in \mathbf{R}^d : \langle q_j, q_j - x \rangle \geqq 0\} \quad (j = 1, \ldots, l).$$

In this case there exists a convex compact set $K \subset \mathbf{R}^d$ for which $R(x)-r(x)$ attaines its minimal value at $x_0=0$, $R(0)=R$, $r(0)=r$ and $R(0)$ is supported by $p_1, \ldots, p_k \in C$ and $r(0)$ is supported by $q_1, \ldots, q_l \in C$.

Now we give the generalization of the inequalities (1) and (2).

Theorem 4. *For $d \geqq 3$, $\max r(x)/r(x_0)$ is not bounded from above. On the other hand, for $d \geqq 3$,*

$$\min R(x)/R(x_0) \geqq \frac{1}{2}\left(\cos^2 \alpha_0 + \cos \alpha_0 - 1 + \frac{1}{\cos \alpha_0}\right) \approx 0.8054,$$

where $\alpha_0 \in (0, \pi/2)$ is the root of the equation $\sin^2 \alpha - 2\cos^3 \alpha = 0$. This inequality is sharp.

3. This section contains the proofs. We start with some simple facts and observations.

Claim 1.

$$R(x) = \max_{p\in K} \|x-p\| = \max_{p\in C} \|x-p\|,$$

$$r(x) = \inf_{p\notin K} \|x-p\| = \min_{p\in C} \|x-p\|,$$

and the points in which the maximum (minimum) is attained support $R(x)$ ($r(x)$, respectively).

Claim 2.

(a)
$$R\left(\frac{x_1+x_2}{2}\right) \leqq \frac{1}{2}(R(x_1)+R(x_2))$$

and if equality holds here, then there is a unique $p\in C$ supporting $R((x_1+x_2)/2)$ and this point lies on the straight line through x_1 and x_2, and p supports $R(x_1)$ and $R(x_2)$ as well.

(b)
$$r\left(\frac{x_1+x_2}{2}\right) \geqq \frac{1}{2}(r(x_1)+r(x_2))$$

Proof. (a) Let $p\in C$ be a point of support for $R((x_1+x_2)/2)$. Then $p\in B(R(x_1), x_1)\cap B(R(x_2), x_2)$ and the triangle-inequality proves the claim.

(b) Obviously conv $(B(r(x_1), x_1)\cup B(r(x_2), x_2))\subseteqq K$ and an easy calculation shows that

$$B\left(\frac{r(x_1)+r(x_2)}{2}, \frac{x_1+x_2}{2}\right) \subseteqq \operatorname{conv}(B(r(x_1), x_1)\cup B(r(x_2), x_2)).$$

Proof of Theorem 1. By Claim 2, $R(x)$ is a convex, $r(x)$ is a concave function. So $R(x)-r(x)$ is convex and attaines its infimum. What we have to show is the uniqueness of the minimum. This will be done by showing that $x_1, x_2\in K$, $x_1\neq x_2$ and $R(x_1)-r(x_1)=R(x_2)-r(x_2)=h$ implies that $R((x_1+x_2)/2)-r((x_1+x_2)/2)<h$.

Convexity implies that $R((x_1+x_2)/2)-r((x_1+x_2)/2)\leqq h$, so assume, by way of contradiction, that $R((x_1+x_2)/2)-r((x_1+x_2)/2)=h$. Then by Claim 2, we have $R((x_1+x_2)/2)=1/2\,(R(x_1)+R(x_2))$ and a unique point $p\in C$ supporting $R(x_1)$, $R(x_2)$ and $R((x_1+x_2)/2)$ and p lies on the straight line through x_1 and x_2. Without loss of generality we suppose that x_2 lies between x_1 and p on this line. By our assumption $R(x_1)-r(x_1)=R(x_2)-r(x_2)$, so $B(r(x_2), x_2)\subseteqq B(r(x_1), x_1)$, and then there is a unique point $q\in C$ supporting $r(x_2)$ and this point lies on the line segment joining x_2 and p. But K contains the set conv $(B(r(x_1), x_1)\cup\{p\})$ and this set contains q in its interior. This contradicts the assumption

$$R\left(\frac{x_1+x_2}{2}\right)-r\left(\frac{x_1+x_2}{2}\right) = h. \quad \square$$

For fixed $p\in C$ define $Z(p)$ as the set of unit outer normals to K at p, i.e.,

$$Z(p) = \{z\in R^d:\ \|z\| = 1, \langle z, p\rangle = \max_{x\in K}\langle z, x\rangle\}.$$

Define now

$$\Gamma = \{(p, z)\in R^d\times R^d:\ z\in Z(p)\}.$$

It is clear that Γ is compact.

Claim 3. (a) $R(x) = \max\{\langle z, p-x\rangle:\ (p, z)\in\Gamma\}$,

(b) $r(x)=\min\{\langle z, p-x\rangle:\ (p, z)\in\Gamma\}$.

Proof. (a) Clearly for each $(p, z)\in\Gamma$

$$\langle z, p-x\rangle \leqq \|z\|\cdot\|p-x\| = \|p-x\| \leqq R(x).$$

If p_0 supports $R(x)$, then $(p_0, ((p_0-x)/\|p_0-x\|)\in\Gamma$ and

$$\left\langle \frac{p_0-x}{\|p_0-x\|}, p_0-x\right\rangle = R(x).$$

(b) Trivially $\langle z, p-x\rangle\geqq r(x)$ for each $(p, z)\in\Gamma$. On the other hand it is easy to check that if p_0 supports $r(x)$, then $Z(p_0)=\{p_0-x/\|p_0-x\|\}$ and

$$\left\langle \frac{p_0-x}{\|p_0-x\|}, p_0-x\right\rangle = r(x). \quad \square$$

Using Claim 3 the function $r\colon K\to R^1$ can be extended over the whole space $\mathbf{R}^d$. It is again easy to see that the extended $r(x)$ is concave, and so the function $R(x)-r(x)$ $(x\in\mathbf{R}^d)$ attaines its minimal value at $x_0\in K$ only.

To prove Theorem 2 we need some definitions and theorem from convex analysis.

Definition. Let $f\colon \mathbf{R}^d\to\mathbf{R}$ be a convex function. The set

$$\partial f(x) = \{x^*\in\mathbf{R}^d:\ \langle x^*, z-x\rangle \leqq f(z)-f(x)\ \text{(for every } z\in\mathbf{R}^d)\}$$

is the *subgradient of f at x.*

It is well-known that the subgradient of a finite convex function is nonempty, convex and compact.

Theorem A (Fenchel, Rockafellar—Moreau, see [5]). *Let $f\colon \mathbf{R}^d\to\mathbf{R}$ be convex, $g\colon \mathbf{R}^d\to\mathbf{R}$ concave functions, finite over the whole space. Then $f(x)-g(x)$ attains its minimum at x_0 if and only if*

$$0\in\partial f(x_0)+\partial(-g)(x_0).$$

Here the last addition is meant in the Minkowski sense; $(-g)$ is a convex function so $\partial(-g)(x_0)$ is its subgradient at x_0.

Theorem B (Йоффе — Тихомиров [3]). *Assume Γ is compact and the map $\gamma\mapsto(x_\gamma^*, a_\gamma)\in\mathbf{R}^d\times\mathbf{R}$ is continuous. Let $f(x)=\sup\{\langle x_\gamma^*, x\rangle+a_\gamma\colon \gamma\in\Gamma\}$. Then $f\colon \mathbf{R}^d\to\mathbf{R}$ is a finite convex function and $\partial f(x_0)=\operatorname{conv}\{x_\gamma^*\colon \gamma\in\Gamma$ and $\langle x_\gamma^*, x_0\rangle+a_\gamma=f(x_0)\}$.*

Now we are ready to prove Theorem 2.

Proof of Theorem 2. First by Theorem B

$$\partial R(x_0) = \operatorname{conv}\{-z\colon (p, z)\in\Gamma, \langle z, p-x_0\rangle = R(x_0)\},$$

$$\partial(-r)(x_0) = \operatorname{conv}\{z\colon (p, z)\in\Gamma, \langle z, p-x_0\rangle = r(x_0)\}.$$

By Theorem A, $R(x)-r(x)$ is minimal at x_0 if and only if for some $x^*\in\mathbf{R}^d$, $x^*\in\partial R(x_0)$ and $-x^*\in\partial(-r)(x_0)$. But $x^*\in\partial R(x_0)$ is the same as $x^*=-\sum_{i=1}^k \alpha_i z_i$ for some $\alpha_i\geqq 0$, $\sum_1^k \alpha_i=1$ and z_i with $(p_i, z_i)\in\Gamma$, $\langle z_i, p_i-x_0\rangle=R(x_0)$.

This is true if and only if $z_i=p_i-x_0/\|p_i-x_0\|$, i.e., if p_i supports $R(x_0)$. Similarly $-x^*\in\partial(-r)(x_0)$ is equivalent to $-x^*=\sum_{j=1}^l \beta_j w_j$ for some $\beta_j\geqq 0$, $\sum_1^l \beta_j=1$ and w_j with $(q_j, w_j)\in\Gamma$, $\langle w_j, q_j-x_0\rangle=r(x_0)$. In this case, again $w_j=(q_j-x_0)/\|q_j-x_0\|$ and q_j supports $r(x_0)$. These conditions imply that $R(x)-r(x)$ is minimal at x_0 if and only if there exist points $p_1, \dots, p_k\in C$ supporting $R(x_0)$ and $q_1, \dots, q_l\in C$ supporting $r(x_0)$ such that

$$\operatorname{conv}\left\{\frac{p_i-x_0}{R(x_0)}\colon i = 1, \dots, k\right\}\cap\operatorname{conv}\left\{\frac{q_j-x_0}{r(x_0)}\colon j = 1, \dots, l\right\} \neq \emptyset.$$

So we are finished with the proof. We mention that $k=1$ (or $l=1$) implies that K is a ball. Further, it can be shown that if $\operatorname{conv} P\cap\operatorname{conv} Q\neq\emptyset$ for some $P, Q\in\mathbf{R}^d$, then there are subsets $P'\subseteqq P$ and $Q'\subseteqq Q$ such that $\operatorname{conv} P'\cap\operatorname{conv} Q'\neq\emptyset$ and $|P'|+|Q|'\leqq d+2$. This means that we can suppose $k+l\leqq d+2$ in Theorem 2.

I mention here that the "only if" part of Theorem 2 can be proved in a simpler way: Set $P=\{(p_i-x_0)/R(x_0)\colon i=1, \dots, k\}$ and $Q=\{(q_j-x_0)/r(x_0)\colon j=1, \dots, l\}$. If $\operatorname{conv} P\cap\operatorname{conv} Q=\emptyset$, then there is a hyperplane separating P and Q with normal $a\in\mathbf{R}^d$, say. One can easily see that $R(x_0)>R(x_0+a)$ and $r(x_0)<r(x_0+a)$ which shows that $R(x)-r(x)$ cannot attain its minimal value at x_0.

Proof of Theorem 3. Set

$$K_{\min} = \operatorname{conv}(B(r, 0)\cup\{p_1, \dots, p_k\}).$$

$$K_{\max} = B(R, 0)\cap\bigcap_{j=1}^l \{x\colon \langle q_j, q_j-x\rangle \geqq 0\}.$$

It is easy to see that both $K_{\min}$ and $K_{\max}$ satisfy the conditions of Theorem 2 with $x_0=0$ and $p_1, \ldots, p_k, q_1, \ldots, q_l$. Moreover, any convex compact set K with $K_{\min} \subseteq \subseteq K \subseteq K_{\max}$ will do the same.

Proof of Theorem 4. *First part.* We construct a convex compact set $K \subset \mathbf{R}^d$ for each $d \geqq 3$ such that $\max r(x)/r(x_0)$ is "large".

Let $\bar{p}_1, \bar{p}_2, q_1, q_2$ be the vertices of a square such that $\|\bar{p}_1\| = \|\bar{p}_2\| = \|q_1\| = \|q_2\| = 1$ and the length of the diagonals $\bar{p}_1\bar{p}_2$ and q_1q_2 is $2-\varepsilon$ (where $\varepsilon > 0$ is small). The hyperplanes $\langle q_1, q_1 - x\rangle = 0$ and $\langle q_2, q_2 - x\rangle = 0$ meet in an affine flat A. The halflines starting from the origin in directions $\bar{p}_1$ and $\bar{p}_2$ meet A in the points $p_1 = R\bar{p}_1$ and $p_2 = R\bar{p}_2$. Consider the set $K_{\max}$ from Theorem 3 with p_1, p_2 and q_1, q_2. A simple calculation shows that

$$R(0) = \left(\varepsilon - \frac{\varepsilon^2}{4}\right)^{-1}, \quad r(0) = 1, \quad \text{and} \quad \max r(x) = \left(\varepsilon - \frac{\varepsilon^2}{4}\right)^{-1/2}.$$

So we have

$$\frac{\max r(x)}{r(x_0)} = \left(\varepsilon - \frac{\varepsilon^2}{4}\right)^{-1/2}$$

which indeed tends to infinity as $\varepsilon \to 0$.

Second part. Let $K \subset \mathbf{R}^d$ $(d \geqq 3)$ be convex compact body and suppose that $R(x) - r(x)$ attaines its minimal value at $x_0 = 0$ and $r(x_0) = 1$, $R(x_0) = R$. By Theorem 2 there exist points $p_1, \ldots, p_l$ supporting $R(x_0)$ and $q_1, \ldots, q_l$ supporting $r(x_0)$ with

$$\operatorname{conv}\{p_i/R\colon i = 1, \ldots, k\} \cap \operatorname{conv}\{q_j\colon j = 1, \ldots, l\} \neq \emptyset,$$

and we may assume $k, l \geqq 2$, $k + l \leqq d + 2$. Then $\operatorname{conv}\{p_1, \ldots, p_k\}$ is a simplex whose nearest point to the origin is p_0 say. Clearly $\|p_1 - p_0\| = \ldots = \|p_k - p_0\|$ and the angle between the vectors p_i and p_0 is the same for each i. Denote this angle by α.

Now the halfspaces $\langle q_j, q_j - x\rangle \geqq 0$ $(j = 1, \ldots, l)$ have to contain the simplex $\operatorname{conv}\{p_1, \ldots, p_k\}$ and so the point p_0 as well. On the other hand, for some $j = 1, \ldots, l$ the angle between the vectors q_j and p_0 is not larger than α for otherwise

$$\operatorname{conv}\{p_i/R\colon i = 1, \ldots, k\} \cap \operatorname{conv}\{q_j\colon j = 1, \ldots, l\} = \emptyset.$$

This implies that

$$0 \leqq \langle q_j, q_j - p_0\rangle = 1 - \langle q_j, p_0\rangle =$$
$$= 1 - \|q_j\| \cdot \|p_0\| \cos(\sphericalangle q_j 0 p_0) \leqq 1 - R\cos^2\alpha.$$

Consider now $\min_x R(x) = \varrho$ and set $R(\bar{x}) = \varrho$, $\bar{x} \in K$. Then $B(\varrho, \bar{x})$ contains the points $p_1, \ldots, p_k$ and the ball $B(1, 0)$, so it contains the point $\bar{p}_0 = -p_0/\|p_0\|$ as well. We are going to give an estimation from below for the radius of the smallest ball containing the points $\bar{p}_0, p_1, \ldots, p_k$. It is clear that the smallest ball containing

$p_1, \ldots, p_k$ is $B(R\sin\alpha, p_0)$ and so $R\sin\alpha \leqq \varrho$. However if $\|\bar{p}_0 - p_0\| = R\cos\alpha + 1 > R\sin\alpha$, then $B(R\sin\alpha, p_0)$ does not contain $\bar{p}_0$. In this case, using some elementary geometry, we get the estimation

$$\varrho \geqq \frac{1+2R\cos\alpha + R^2}{2(1+R\cos\alpha)}.$$

Define now

$$f(R, \alpha) = \begin{cases} \sin\alpha & \text{if} \quad R\sin\alpha \geqq R\cos\alpha + 1, \\ \dfrac{1+2R\cos\alpha+R^2}{2R(1+R\cos\alpha)} & \text{otherwise} \end{cases}$$

where $R \geqq 1$, $0 \leqq \alpha \leqq \pi/2$ and $R\cos^2\alpha \leqq 1$.

What we have to do is to find the minimum of f in the domain determined by these inequalities. This is a routine calculation. The main steps are:

1) for R fixed $f(R, \alpha)$ is monotone non-decreasing, so the minimum is attained on the curve $R\cos^2\alpha = 1$;

2) on this curve the minimum of f is equal to

$$\frac{1}{2}(\cos^2\alpha_0 + \cos\alpha_0 - 1 + \cos^{-1}\alpha_0)$$

where α_0 is the solution of the equation $\sin^2\alpha - 2\cos^3\alpha = 0$ with $0 \leqq \alpha_0 \leqq \pi/2$.

This proves that

$$\frac{\min R(x)}{R(x_0)} \geqq \frac{1}{2}\left(\cos^2\alpha_0 + \cos\alpha_0 - 1 + \frac{1}{\cos\alpha_0}\right). \tag{4}$$

Finally we give an example showing that equality can occur here for $d = 3, 4, \ldots$. Again, let $\bar{p}_1, \bar{p}_2, q_1, q_2$ be the vertices of a square such that the diagonals $\bar{p}_1, \bar{p}_2$ and q_1, q_2 meet in a point q and the angle between q and $\bar{p}_1, \bar{p}_2, q_1, q_2$ equals α_0. Now set $p = \cos^{-2}\alpha_0\, \bar{p}_1$ and $p_2 = \cos^{-2}\alpha_0\, \bar{p}_2$ and apply Theorem 3 with the vectors p_1, p_2, q_1, q_2 to get the convex compact set $K_{\min}$. An easy calculation shows that for $K_{\min}$ (4) holds with equality.

Acknowledgement. The author is indebted to professor I. VINCZE for raising the problem and for fruitful discussions.

References

[1] T. Bonnesen, Über das isoperimetrische Defizit ebener Figuren, *Math. Ann.*, **91** (1924), 252—268.
[2] T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper*, de Gruyter (Berlin, 1956).
[3] А. Д. Йоффе, В. М. Тихомиров, *Теория экстремальных задач*, Наука (Москва, 1974).
[4] N. Kritikos, Über konvexe Flachen und einschließende Kugeln, *Math. Ann.* **96** (1927), 583—586.
[5] T. R. Rockafellar, *Convex Analysis*, Princeton (1970).
[6] J. Szőkefalvi-Nagy, Konvexe Kurven und einschließende Kreisringe, *Acta. Sci. Math. (Szeged)*, **10** (1941—1943), 174—184.
[7] St. Vincze, Über den Minimalkreisring einer Eilinie, *Acta Sci. Math. (Szeged)*, **11** (1947), 133—138.
[8] I. Vincze, Über Kreisringe, die eine eilinien einschließen, *Studia Sci. Math. Hungar*, **9** (1974), 155—159.

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

Acta Sci. Math., 52 (1988), 101—111

Об одном интерполяционном процессе Эрмита—Фейера при ультрасферических узлах

Д. Л. БЕРМАН

1. Пусть C множество всех функций, непрерывных в $[-1, 1]$. Для матрицы чисел

$$\text{(m)} \qquad \{x_k^{(n)}\}, \quad k=\overline{1,n}, \quad n=1,2,\ldots, \quad -1<x_n^{(n)}<x_{n-1}^{(n)}<\ldots<x_1^{(n)}<1,$$

строим полином $H_n(f, x)$ степени $2n-1$, однозначно определяющийся из условий $H_n(f, x_k^{(n)})=f(x_k^{(n)})$, $H_n'(f, x_k^{(n)})=0$, $k=1, 2, \ldots, n$. Классическая теорема Л. Фейера [1] утверждает, что, если n-я строчка матрицы (m) состоит из чисел

$$x_k^{(n)}=\cos\frac{(2k-1)\pi}{2n}, \quad k=1,2,\ldots,n, \quad n=1,2,\ldots, \tag{1}$$

то для любой $f\in C$ выполняется равномерно в $[-1, 1]$ соотношение

$$H_n(f, x)\to f(x), \quad n\to\infty. \tag{2}$$

Хорошо известно, что процесс $\{H_n(f, x)\}$ называется интерполяционным процессом Эрмита—Фейера.

Пусть полином $H_n(f, x)$ построен для n-й строчки произвольной матрицы узлов вида (m). Наряду с полиномом $H_n(f, x)$ рассмотрим полином $F_n(f, x)$ степени $2n+3$, который однозначно определяется из условий

$$F_n(f, x_k^{(n)})=f(x_k^{(n)}); \quad F_n(f, \pm1)=f(\pm1); \quad F_n'(f, x_k^{(n)})=F_n'(f, \pm1)=0,$$
$$k=1,2,\ldots,n.$$

В [2]—[3] автор изучал процесс $\{F_n(f, x)\}$ для случая узлов

$$x_0^{(n+2)}=1;\ x_k^{(n+2)}=\cos\big((2k-1)\pi/2n\big), \quad k=\overline{1,n}, \quad x_{n+1}^{(n+2)}=-1, \quad n=1,2,\ldots.$$

Оказалось, что этот процесс, построенный для $f(x)=|x|$, расходится в точке $x=0$. В [4] было доказано, что он расходится всюду в $(-1, 1)$. Такое же утверж-

Поступило 5-ого декабря 1984.

дение имеет место для $f(x)=x^2$ и для $f(x)=x$ при $x\neq 0$ [5]—[6]. В [7]—[8] изучался процесс Эрмита—Фейера при узлах

$$(\mathrm{m}_1)\qquad x_0^{(n+1)}=1,\quad x_k^{(n+1)}=\cos((2k-1)\pi/2n),\quad k=\overline{1,n},\quad n=1,2,\ldots,$$

$$(\mathrm{m}_2)\qquad x_{n+1}^{(n+1)}=-1,\quad x_k^{(n+1)}=\cos((2k-1)\pi/2n,\quad k=\overline{1,n},\quad n=1,2,\ldots,$$

Было доказано, что процесс Эрмита—Фейера, построенный для $f(x)=|x|$ при узлах (m_1) расходится в каждой точке из $[-1,1)$. Если же этот процесс построить при узлах (m_2), то он расходится всюду в $(-1,1]$.

На первый взлгяд может показаться, что эти отрицательные результаты связаны с отсутствием производной у функции $f(x)=|x|$ в точке $x=0$. Но это не так, ибо в [9] установлено, что процесс Эрмита—Фейера при узлах (m_1) для $f(x)=x$ расходится всюду в $[-1,1)$. С другой стороны, простой проверкой можно убедиться, что процесс Эрмита—Фейера при узлах (m_1) для $f(x)=(x-1)^2$ сходится равномерно в $[-1,1]$. Поэтому возникает вопрос о нахождении необходимых и достаточных условий для функции для равномерной сходимости процесса Эрмита—Фейера при матрице узлов (m_1). Аналогичный вопрос возникает для матрицы узлов (m_2). Этим вопросам, в основном, и посвящена эта заметка. Аналогичная задача возникает также для процесса $\{F_n(f,x)\}$. Она изучалась в [10]. Рассмотрение будем вести для некоторого класса матриц узлов, включающего матрицы узлов из корней ультрасферических полиномов $\{J_n^{(\alpha)}(x)\}$, где $-1/2\leqq\alpha<0$. Недавно R. Bojanic [13] изучал эту задачу для узлов (1), что соответствует тому, что $\alpha=-1/2$. Следует подчеркнуть, что наше рассмотрение совершенно элементарное и не пользуется асимптотическими формулами для полиномов Якоби.

2. Хорошо известно, что при любой матрице узлов (m) полином $H_n(f,x)$ может быть представлен в виде

$$(3)\qquad H_n(f,x)=\sum_{k=1}^{n}f(x_k^{(n)})h_k^{(n)}(x),\quad h_k^{(n)}(x)=h_k(x)=[l_k^{(n)}(x)]^2V_k^{(n)}(x),$$

$$l_k^{(n)}(x)=l_k(x)=\frac{\omega_n(x)}{(x-x_k^{(n)})\omega_n'(x_k^{(n)})},\quad \omega_n(x)=\omega(x)=\prod_{k=1}^{n}(x-x_k^{(n)}),$$

$$V_k^{(n)}(x)=V_k(x)=1-\omega''(x_k^{(n)})(x-x_k^{(n)})(\omega'(x_k^{(n)}))^{-1}.$$

Из однозначности полинома $H_n(f,x)$ следует, что

$$\sum_{k=1}^{n}h_k^{(n)}(x)=1,\quad n=1,2,\ldots$$

Будем говорить, что матрица узлов (m) обладает свойством (F), если выполняются условия:

1) $h_k^{(n)}(1) \geqq 0, \quad k=\overline{1, n}, \quad n=1,2, \ldots;$

2) $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} x_k^{(n)} h_k^{(n)}(1) = 1;$

3) существует конечный предел $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} [l_k^{(n)}(1)]^2.$

Обозначим через Δ подмножество из C, состоящее из всех функций $f(x)$, имеющих левую производную $f'(1)$. Введем функционал

$$\alpha_n(f) = H_n'(f, 1)/2 + (\omega_n'(1)/\omega_n(1))[f(1) - H_n(f, 1)], \tag{4}$$

где $\omega_n(x) = \prod\limits_{i=1}^{n}(x - x_i^{(n)})$ и числа $\{x_i^{(n)}\}_{i=1}^{n}$ составляют n-ую строчку матрицы (m). Справедлива следующая лемма.

Лемма. *Пусть функционал $\alpha_n(f)$ построен при матрице узлов* (m), *обладающей свойством* (F). *Тогда для любой $f \in \Delta$ существует конечный $\lim\limits_{n\to\infty} \alpha_n(f)$ и выполняется равенство*

$$\lim_{n\to\infty} \alpha_n(f) = ((1+d)/2) f'(1), \tag{5}$$

где $d = \lim\limits_{n\to\infty} \sum\limits_{i=1}^{n} [l_i^{(n)}(1)]^2.$

Доказательство. Мы часто опускаем верхний индекс n ради простоты письма. Очевидно, что $d \geqq 0$ — конечное число. Из (3) получим, что

$$H_n'(f, 1) = \sum_{k=1}^{n} f(x_k^{(n)})[l_k^2(1) V_k'(1) + 2 l_k(1) l_k'(1) V_k(1)]. \tag{6}$$

Очевидно, что

$$l_k^2(1) V_k'(1) = -\frac{\omega_n^2(1)\omega_n''(x_k)}{(\omega_n'(x_k))^3 (1-x_k)^2}.$$

После простых вычислений имеем

$$l_k'(1) = -\frac{\omega_n(1)}{\omega_n'(x_k)(1-x_k)^2}\left(1 - \frac{\omega_n'(1)}{\omega_n(1)}(1-x_k)\right).$$

Поэтому

$$l_k^2(1) V_k'(1) + 2 l_k(1) l_k'(1) V_k(1) = \left(\frac{\omega''(x_k)}{\omega'(x_k)} - \frac{2}{1-x_k}\right) l_k^2(1) + \frac{2\omega'(1)}{\omega(1)} h_k(1).$$

Отсюда и из (6) получим, что

$$(7)\qquad H'_n(f,1) = -\sum_{k=1}^{n}\frac{f(x_k)}{1-x_k}h_k(1) - \sum_{k=1}^{n}\frac{f(x_k)}{1-x_k}l_k^2(1) + \frac{2\omega'(1)}{\omega(1)}H_n(f,1).$$

Из (4) и (7) вытекает, что

$$(8)\qquad \alpha_n(f) = -\frac{1}{2}\sum_{k=1}^{n}\frac{f(x_k)}{1-x_k}h_k(1) - \frac{1}{2}\sum_{k=1}^{n}\frac{f(x_k)}{1-x_k}l_k^2(1) + \frac{\omega'(1)}{\omega(1)}f(1).$$

Положим в (8) $f(x)\equiv 1$ и учтем, что из (4) следует, что в этом случае $\alpha_n(f)=0$. Стало быть, из (8) выводим, что

$$(9)\qquad \frac{\omega'(1)}{\omega(1)} = \frac{1}{2}\sum_{k=1}^{n}\frac{1}{1-x_k}h_k(1) + \frac{1}{2}\sum_{k=1}^{n}\frac{1}{1-x_k}l_k^2(1).$$

Подставляя (9) в (8), получим, что

$$(10)\qquad \alpha_n(f) = \frac{1}{2}\sum_{k=1}^{n}\frac{f(1)-f(x_k)}{1-x_k}h_k(1) + \frac{1}{2}\sum_{k=1}^{n}\frac{f(1)-f(x_k)}{1-x_k}l_k^2(1) = S_1^{(n)} + S_2^{(n)}.$$

По условию существует $f'(1)$. Поэтому по $\varepsilon>0$ можно найти такое $\delta>0$, что

$$(11)\qquad \left|\frac{f(1)-f(x_k)}{1-x_k} - f'(1)\right| < \varepsilon,$$

если $1-x_k<\delta$. Так как $\sum_{k=1}^{n}h_k(1)=1$, то

$$(12)\qquad S_1 - \frac{f'(1)}{2} = \frac{1}{2}\sum_{k=1}^{n}\left(\frac{f(1)-f(x_k)}{1-x_k} - f'(1)\right)h_k(1).$$

Из (11) и (12) вытекает, что

$$(13)\qquad \left|S_1 - \frac{f'(1)}{2}\right| \leqq \frac{\varepsilon}{2}\sum_{1-x_k<\delta}|h_k(1)| + \frac{1}{2}\sum_{1-x_k\geqq\delta}\left|\frac{f(1)-f(x_k)}{1-x_k} - f'(1)\right| |h_k(1)|.$$

Согласно условию 1) теоремы 1 $h_k(1)\geqq 0$, $k=\overline{1,n}$. Поэтому из (13) получаем, что

$$(14)\qquad |S_1 - f'(1)/2| \leqq \varepsilon/2 + (1/2)\big(2\|f\|/\delta + |f'(1)|\big)\sum_{1-x_k\geqq\delta}h_k(1),$$

где $\|f\| = \max_{-1\leqq x\leqq 1}|f(x)|$. Заметим, что

$$(15)\qquad \sum_{1-x_k\geqq\delta}h_k(1) \leqq (1/\delta)\sum_{k=1}^{n}(1-x_k)h_k(1),$$

и что для матрицы обладающей свойством (F)

(16) $$\lim_{n\to\infty}\sum_{k=1}^{n}(1-x_k^{(n)})h_k^{(n)}(1)=0.$$

Из (14) и (15—(16) выводим, что

(17) $$\lim_{n\to\infty}S_1^{(n)}=f'(1)/2.$$

Рассмотрим теперь $S_2^{(n)}$. Очевидно, что

(18) $$S_2^{(n)}-\frac{f'(1)\,d_n}{2}=\frac{1}{2}\sum_{k=1}^{n}{}'\left(\frac{f(1)-f(x_k)}{1-x_k}-f'(1)\right)l_k^2(1),$$

где $d_k=\sum_{k=1}^{n}l_k^2(1)$. Ясно, что

(19) $$\left|S_2^{(n)}-\frac{f'(1)\,d_n}{2}\right|\leqq\frac{\varepsilon}{2}\sum_{k=1}^{n}l_k^2(1)+\frac{1}{2}\sum_{1-x_k\geqq\delta}\left|\frac{f(1)-f(x_k)}{1-x_k}-f'(1)\right|l_k^2(1).$$

Поскольку матрица узлов обладает свойством (F), то существует такая константа $C_1>0$, что $\sum_{k=1}^{n}[l_k^{(n)}(1)]^2\leqq C_1$, $n=1,2,\ldots$. Поэтому из (19) выводим

(20) $$|S_2^{(n)}-f'(1)\,d_n/2|\leqq\varepsilon C_1/2+(1/2)(2\|f\|/\delta+|f'(1)|)\sum_{1-x_k\geqq\delta}l_k^2(1).$$

Очевидно, что

(21) $$\sum_{1-x_k\geqq\delta}l_k^2(1)\leqq(1/\delta)\sum_{k=1}^{n}(1-x_k)l_k^2(1).$$

Из тождества

$$x=\sum_{k=1}^{n}x_kh_k(x)+\sum_{k=1}^{n}(x-x_k)\,l_k^2(x)$$

следует, что

(22) $$\sum_{k=1}^{n}(1-x_k)l_k^2(1)=1-\sum_{k=1}^{n}x_k\,h_k(1).$$

Из условия 2) матрицы узлов, обладающей свойством (F) и из (22) получим, что

(23) $$\lim_{n\to\infty}\sum_{k=1}^{n}(1-x_k^{(n)})[l_k^{(n)}(1)]^2=0.$$

Поэтому из (20), (21), (23) выводим, что

(24) $$\lim_{n\to\infty}S_2^{(n)}=f'(1)\,d/2,$$

ибо $\lim_{n\to\infty}d_n=d$. Из (10), (17), (24) вытекает (5).

3. Интерполяционный полином $Q_n(f, x)$ Эрмита—Фейера степени $2n+1$ для точек $x_n^{(n)}<x_{n-1}^{(n)}<\ldots<x_1^{(n)}<1$ определяется однозначно из условий

$$Q_n(f, x_k^{(n)}) = f(x_k^{(n)}), \quad Q'_n(f, x_k^{(n)}) = 0, \quad k = \overline{1, n}, \quad Q_n(f, 1) = f(1), \quad Q'_n(f, 1) = 0.$$

Положим

$$r_n(f, x) = r_n = Q_n(f, x) - H_n(f, x). \tag{25}$$

Из определения полиномов $Q_n(f, x)$ и $H_n(f, x)$ имеем, что

$$r_n = \omega^2(x)(Ax+B), \tag{26}$$

где A и B определяются из системы уравнений

$$\omega^2(1)(A+B) = f(1) - H_n(f, 1),$$
$$2\omega(1)\omega'(1)(A+B) + A\omega^2(1) = -H'_n(f, 1).$$

Отсюда и из (26), после простых вычислений, получим, что

$$r_n = 2(\omega_n^2(x)/\omega_n^2(1))(1-x)\alpha_n(f) + (\omega_n^2(x)/\omega_n^2(1))(f(1) - H_n(f, 1)), \tag{27}$$

где $\alpha_n(f)$ определяется согласно (4). Теперь можно доказать следующую теорему*.

Теорема 1. *Пусть интерполяционный процесс* $\{Q_n(f, x)\}$ *построен для* $f \in \Delta$ *при матрице узлов* (m), *обладающей свойствами:*

1) $h_k^{(n)}(x) \geqq 0, \ |x| \leqq 1,$

2) $\lim\limits_{n\to\infty} \sum\limits_{k=2}^{n} (x_k^{(n)})^i h_k^{(n)}(x) = x^i, \ i=1, 2,$ *равномерно в* $[-1, 1]$.

3) *Существует конечный предел* $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} [l_k^{(n)}(1)]^2$.

4) *Выполняется неравенство* $|\omega_n(x)| \leqq C_2|\omega_n(1)|$, *где* $|x| \leqq 1$, C_2 — *константа.*

5) $|\omega_n(-1)| = |\omega_n(1)|$, *где* $\omega_n(x) = \prod\limits_{k=1}^{n} (x - x_k^{(n)})$ *и* $\{x_k^{(n)}\}_{k=1}^n$ *составляют n-ую строчку матрицы* (m).

Тогда для равномерной сходимости процесса $\{Q_n(f, x)\}$ *к* $f(x)$ *в* $[-1, 1]$ *необходимо и достаточно, чтобы выполнялось условие* $f'(1)=0$.

Доказательство. Докажем сперва достаточность. Из (3) и из условия 1) теоеремы 1 следует, что оператор $H_n(f, x)$ положительный. Поэтому из условия 2) теоремы 1 и из равенства $\sum\limits_{k=1}^{n} h_k(x) = 1$, в силу теоремы П. П. Коровкина

*) Отметим, что все результаты этой статьи без труда переносятся на случай, когда процесс Эрмита—Фейера строится для матрицы узлов $-1<x_n^{(n)}<x_{n-1}^{(n)}<\ldots<x_1^{(n)}$, $n=1, 2, \ldots$. Условие $f'(1)=0$ заменяется условием $f'(-1)=0$.

[12], заключаем, что для любой $f\in C$ выполняется равномерно в $[-1, 1]$ соотношение $H_n(f, x)\to f(x)$, $n\to\infty$.

Стало быть, нужно доказать, что при $f'(1)=0$ выполняется в $[-1, 1]$ равномерно соотношение

$$r_n(f, x) \to 0, \quad n \to \infty. \tag{28}$$

Из условий теоремы 1 непосредственно следует, что

$$|r_n(f, x)| \leqq 4C_2|\alpha_n(f)|+C_2|f(1)-H_n(f, 1)|. \tag{29}$$

Ясно, что матрица узлов обладает свойством (F). Поэтому применима лемма. Отсюда, поскольку $f'(1)=0$, то получаем, что $\lim\limits_{n\to\infty} \alpha_n(f)=0$. Кроме того, из (2) следует, что $f(1)-H_n(f, 1)\to 0$, $n\to\infty$. Поэтому из (29) вытекает (28).

Необходимость. Положим в (27) $x=-1$ и учтем, что по условию $|\omega_n(-1)|=|\omega_n(1)|$. Поэтому из (27) получим, что

$$Q_n(f, -1)-H_n(f, -1) = 4\alpha_n(f)+\big(f(1)-H_n(f, 1)\big). \tag{30}$$

По условию $\lim\limits_{n\to\infty} Q_n(f, -1)=f(-1)$. Кроме того, согласно 2) имеем $\lim\limits_{n\to\infty} H_n(f, \pm 1)=f(\pm 1)$. Стало быть, из (30) заключаем, что $\lim\limits_{n\to\infty} \alpha_n(f)=0$. Отсюда в силу леммы выводим, что $\big((1+d)/2\big)f'(1)=0$. Так как $d\geqq 0$, то отсюда получаем, что $f'(1)=0$.

4. Пусть n-я строчка матрицы (m) состоит из корней полинома $\omega_n(x)= =\omega(x)=\prod\limits_{i=1}^{n}(x-x_i^{(n)})$. Согласно Л. Фейеру [11] матрица (m) является ϱ-нормальной, если существует такое число $\varrho>0$, что всюду в $[-1, 1]$ выполняется неравентсво

$$V_k(x) = 1-(x-x_k^{(n)})\omega_n''(x_k^{(n)})\big(\omega_n'(x_k^{(n)})\big)^{-1} > \varrho > 0, \quad k=\overline{1, n}, \quad n=1, 2, \ldots,$$

где $\{x_k^{(n)}\}_{k=1}^n$ — корни $\omega_n(x)$. Л. Фейер [11] доказал, что, если матрцца (m) составлена из корней полиномов Якоби $J_n^{(\alpha_n, \beta_n)}(x)$, где $-1<\alpha_n$, $\beta_n<-\gamma<0$, $n=1, 2, \ldots$, а γ — сколь угодно малое фиксированное число, то она ϱ-нормальная. Г. Грюнвальд [14] доказал, что при ϱ-нормальной матрице узлов (m) для любой $f\in C$ выполняется в $[-1, 1]$ равномерно соотношение (2). Поэтому из теоремы 1 вытекает

Теорема 2. *Пусть матрица узлов* (m) *ϱ-нормальная и пусть выполняются условия* 3), 4), 5) *из теоремы* 1. *Тогда для равномерной сходимости процесса* $\{Q_n(f, x)\}$ *к* $f(x)$ *в* $[-1, 1]$ *необходимо и достаточно, чтобы выполнялось условие* $f'(1)=0$. Дадим приложение теоремы 2 к случаю, когда матрица узлов (m) составлена из корней ультрасферических полиномов $J_n^{(\alpha)}(x)$, $-1/2\leqq\alpha<0$. Для этого нужна следующая теорема Л. Фейера [15]

Теорема (Л. Фейера). *Если n-я строчка матрицы* (m) *составлена из корней полинома Якоби* $J_n^{(\alpha,\beta)}(x)$, $n=1, 2, \ldots$ *и* α, β *удовлетворяют неравенствам* $-1<\alpha, \beta<0$, *то справедливы равентсва*

(31) $$\lim_{n\to\infty}\sum_{k=1}^{n}[l_k^{(n)}(1)]^2 = -1/\alpha;$$

(32) $$\lim_{n\to\infty}\sum_{k=1}^{n}[l_k^{(n)}(-1)]^2 = -1/\beta;$$

(33) $$\lim_{n\to\infty}\sum_{k=1}^{n}[l_k^{(n)}(x)]^2 = 1, \quad |x|<1.$$

У Фейера [15] равенства (31) и (32) имеют следующий вид:

(34) $$\lim_{n\to\infty}\sum_{k=1}^{n}[l_k^{(n)}(1)]^2 = 1/(1-2\beta);$$

(35) $$\lim_{n\to\infty}\sum_{k=1}^{n}[l_k^{(n)}(-1)]^2 = 1/(1-2\alpha),$$

ибо он пользуется обозначениями Стилтьеса [16], где α и β соответствуют в наших обозначениях $(\beta+1)/2$ и $(\alpha+1)/2$. Равенства (34) и (35) приводятся также в работе Г. Грюнвальда [14]. По поводу доказательства этой теоремы Фейер [15] пишет „Auf dem beweis werde ich hier nicht eingehen". У Грюнвальда [14] также нет доказательства. Мне неизвестно, где изложено доказательство упомянутой теоремы Фейера. Поэтому я здесь вкратце издожу её доказательство. Идея этого доказательства, для случая полиномов Лежандра, принадлежит Л. Фейеру [15].

Рассмотрим сперва случай, когда $x=1$. Введем функцию

(36) $$f(x) = (1+x)/\varphi(x),$$

где $\varphi(x)=(1+\beta)(1-x)-\alpha(1+x)$. Очевидно, что $\varphi(-1)=2(1+\beta)>0$, ибо $\beta>-1$. $\varphi(1)=-2\alpha<0$, ибо по условию $\alpha<0$. Поскольку $\varphi(x)$ — линейная функция от x, то отсюда заключаем, что $\varphi(x)>0$, в $[-1, 1]$. Значит функция (36) непрерывна в $[-1, 1]$. Поэтому согласно упомянутой теореме Г. Грюнвальда [14] выполняется равенство

(37) $$\lim_{n\to\infty}\sum_{k=1}^{n} f(x_k^{(n)})V_k^{(n)}(1)[l_k^{(n)}(1)]^2 = f(1).$$

Но функция (36) выбрана так, что $f(x_k^{(n)})V_k^{(n)}(1)=1$, $k=\overline{1, n}$, $n=1, 2, \ldots$. Стало быть, (37) принимает вид:

$$\lim_{n\to\infty}\sum_{k=1}^{n}[l_k^{(n)}(1)]^2 = -1/\alpha.$$

Аналогичным образом доказывается равенство (32). Докажем теперь равенство (33). Итак, пусть a — фиксированное число из $(-1, 1)$. Введем функцию

$$f_1(x) = (1-x^2)/\varphi_1(x), \tag{38}$$

где $\varphi_1(x)=\varphi_1(x, \alpha, \beta)=(\alpha+\beta+1)x^2+[(\alpha-\beta)-(\alpha+\beta+2)a]x+1-a(\alpha+\beta)$. Поскольку $\varphi_1(x, \alpha, \beta)$ — линейная функция от α и β, то нетрудно проверить, что при $x\in[-1, 1]$, $-1+\gamma_1\leqq\alpha\leqq 0$, $-1+\gamma_1\leqq\beta\leqq 0$, $\varphi(x, \alpha, \beta)>0$, при этом $\gamma_1>0$ — сколь угодно малое число. Согласно теореме Г. Грюнвальда [14] имеем, что

$$\lim_{n\to\infty}\sum_{k=1}^{n} f_1(x_k^{(n)})V_k^{(n)}(a)[l_k^{(n)}(a)]^2 = f_1(a), \quad |a| < 1, \tag{39}$$

ибо $f_1(x)$ непрерывна в $[-1, 1]$. Функция $f_1(x)$ выбрана таким образом, что $f_1(x_k^{(n)})V_k^{(n)}(a)=1$, $k=\overline{1, n}$. Поэтому (39) принимает вид:

$$\lim_{n\to\infty}\sum_{k=1}^{n} [l_k^{(n)}(a)]^2 = f_1(a), \quad |a| < 1. \tag{40}$$

Из (38) видно, что $f_1(a)=1$. Поэтому (40) совпадает с (33).

Пользуясь теоремой Л. Фейера [15] можно доказать такую теорему

Теорема 3. *Пусть n-я строчка матрицы* (m) *состоит из корней ультрасферических полиномов $J_n^{(\alpha)}(x)$, где*

$$-1/2 \leqq \alpha < 0, \tag{41}$$

и пусть $f\in\Delta$. Тогда для равномерной сходимости процесса $\{Q_n(f, x)\}$ к $f(x)$ в $[-1, 1]$ необходимо и достаточно, чтобы выполнялось условие $f'(1)=0$.

Доказательство. Эта теорема непосредственно следует из теоремы 2, ибо все условия теоремы 2 выполнены. Действительно, при выполнении (41) матрица узлов (m) ϱ-нормальная. Как известно, при $\alpha\geqq -1/2$, $|J_n^{(\alpha)}(x)|\leqq|J_n^{(\alpha)}(1)|$, $|x|\leqq 1$. Стало быть, выполняется условие 4) из теоремы 1. Для ультрасферических полиномов $|J_n^{(\alpha)}(-x)|=|J_n^{(\alpha)}(x)|$. Поэтому выполняется условие 5) из теоремы 1. В силу (41) и (31) выполняется условие 3) из теоремы 1. Итак, теорема 3 доказана.

В связи с этой теоремой возникает вопрос о нахождении аналога этой теоремы, когда неравенства (41) заменяются условием $\alpha\in(-1, \infty)\setminus[-1/2, 0)$. Вероятно, для решения этого вопроса будут полезные исследования Szabados [17] и P. Vértesi [18]—[19].

Замечание. Как видно из доказательства леммы условие, что существует конечный $\lim_{n\to\infty} d_n=d<\infty$ можно заменить условием, что существует

$\overline{\lim\limits_{n\to\infty}}\, d_n=d<\infty$ и тогда равенство (5) заменится равенством

$$\lim_{n\to\infty}\left(\alpha_n(f)-((1+d_n)/2)f'(1)\right)=0,$$

которое достаточно для доказательства теореамы 1. Поэтому в теореме 1 можно равенство $\lim\limits_{n\to\infty} d_n=d<\infty$ заменить равенством $\overline{\lim\limits_{n\to\infty}}\, d_n=d<\infty$. Хорошо известно [11], что для ϱ-нормальной матрицы узлов $d_n\leqq 1/\varrho$. Поэтому в этом случае $\overline{\lim\limits_{n\to\infty}}\, d_n<\infty$. Стало быть, из теоремы 2 можно исключить условие 3). Л. Фейер [11] доказал что, если матрица узлов (m) составлена из корней полиномов Якоби $J_n^{(\alpha,\beta)}(x)$, где $-1<\alpha,\beta<-\gamma$, γ — сколь угодно малое фиксированное число, то

$$\sum_{k=1}^{n}[l_k^{(n)}(x)]^2\leqq\max(-1/\alpha,-1/\beta),\quad |x|\leqq 1,$$

Поэтому при доказательстве теоремы 3 можно обойтись без теоремы Л. Фейера (см. стр. 7). Выражаю благодарность референту за полезные замечания.

Литература

[1] L. Fejér, Über Interpolation, *Gött. Nachr.*, **1916**, 66—91.
[2] Д. Л. Берман, К теории интерполяции, *ДАН СССР*, **163** (1965), 551—554.
[3] Д. Л. Берман, К теории интерполяции функции действительного переменного, *Известия вузов, Матем.*, **1967**, №1, 15—20.
[4] Д. Л. Берман, Всюду расходящийся расширенный интерполяционный процесс Эрмита—Фейера, *Известия вузов, Матем.*, **1975**, №9, 84—87.
[5] Д. Л. Берман, Исследование интерполяционного процесса Эрмита—Фейера, *ДАН СССР*, **187**, (1969), 241—244.
[6] Д. Л. Берман, Об одном всюду расходящемся интерполяционном процессе Эрмита—Фейера, *Известия вузов, Матем.*, **1970**, №1, 3—8.
[7] Д. Л. Берман, К теории интерполяции функции действительного переменного, *Известия вузов, Матем.* **1969**, №8, 10—16.
[8] Д. Л. Берман, Всюду расходящиеся интерполяционные процессы Эрмита—Фейера, *Известия вузов, Матем.*, **1978**, №7, 3—4.
[9] Д. Л. Берман, Исследоваиие сходимости всевозможных вариантов расширенного интерполяционного процесса Эрмита—Фейера, *Известия вузов, Матем.*, **1975**, №8, 97—101.
[10] Д. Л. Берман, О расширенном интерполяционном процессе Эрмита—Фейера, *Известия вузов, Матем.*, **1981**, №8, 5—13.
[11] L. Fejér, Lagrangesche Interpolation und die zugehörigen konjugierten Punkte, *Math. Ann.*, **106** (1932), 1—55.
[12] П. П. Коровкин, *Линейные операторы и теория приближений*, Физматгиз (Москва, 1959).
[13] R. Bojanic, Necessary and sufficient conditions for the convergence of the extended Hermite—Fejér interpolation process, *Acta Math. Hung.*, **36** (1980), 271—279.
[14] G. Grünwald, On the theory of interpolation, *Acta Math.*, **75** (1943), 219—245.

[15] L. Fejér, Bestimmung derjenigen Abszissen eines Intervalles, für welche die Quadratsumme der Grundfunktionen ein möglichst kleines Maximum besitzt, *Ann. Scuola Norm. Sup. Pisa* (2), **1** (1932), 263—276.

[16] Г. Сеге, *Ортогональные многочлены* (Москва, 1962).

[17] J. Szabados, On Hermite—Fejér interpolation for the Jacobi abscissas, *Acta Math. Acad. Sci. Hungar*, **23** (1972), 449—464.

[18] P. Vértesi, Hermite—Fejér type interpolation. III, *Acta Math. Acad. Sci. Hungar.*, **34** (1979), 67—84.

[19] P. Vértesi, Hermite—Fejér type interpolation. IV. (Convergence criteria for Jacobi abscissas), *Acta Math. Acad. Sci. Hungar.*, **39** (1982), 83—93.

БАССЕЙНАЯ УЛ. 68, КВ. 90.
192238 ЛЕНИНГРАД-238, СССР

Acta Sci. Math., 52 (1988), 113—116

A note on optimal interpolation with rational functions

THEODORE KILGORE

Introduction. This note provides a further application of results derived in [7], which dealt with polynomial interpolation.

Let **Y** be the space of rational functions whose numerators are of degree n or less, with denominator

$$Q(t) = (t-t_{n+1})\dots(t-t_{n+m}).$$

If nodes of interpolation $t_0, \dots, t_n$ are chosen on an interval, $[a, b]$ such that

$$a = t_0 < t_1 < \dots < t_n = b,$$

and such that

$$t_{n+k} \notin [t_0, t_n] \quad \text{for} \quad k \in \{1, \dots, n\},$$

it is possible to construct *fundamental functions* $y_0, \dots, y_n$ such that $y_i(t_j) = \delta_{ij}$ (Kronecker delta) for $i \in \{0, \dots, n\}$ and for $j \in \{0, \dots, n\}$, by means of the formula

$$y_i(t) = \frac{Q(t_i)}{Q(t)} \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{(t-t_j)}{(t_i-t_j)}.$$

One defines an interpolating projection $L\colon C[a, b] \to \mathbf{Y}$ by

$$Lf = \sum_{i=0}^{n} f(t_i) y_i \quad \text{for} \quad f \in C[a, b].$$

Clearly, L is bounded, and

$$\|L\| = \Big\| \sum_{i=0}^{n} |y_i| \Big\|.$$

Our purpose here is to minimize $\|L\|$.

Notation. We define, for $i \in \{1, \dots, n\}$, X_i to be the function (in **Y**) which agrees with $\sum_{j=0}^{n} |y_j|$ on the interval $[t_{i-1}, t_i]$, $\lambda_i = X_i(T_i)$, and T_i as the point in (t_{i-1}, t_i)

Received March 11, 1985.

at which λ_i is attained. We note that

$$X_i'(T_i) = 0, \quad \text{for} \quad i \in \{1, \ldots, n\}.$$

Results.

Theorem. *If interpolation is done on an interval* $[a, b]$ *with rational functions having denominator*

$$Q(t) = (t - t_{n+1}) \ldots (t - t_{n+m})$$

and nodes of interpolation $t_0, \ldots, t_n$, *such that*

$$a = t_0 < \ldots < t_n = b < t_{n+1} < \ldots < t_{n+m},$$

then

(i) *interpolation of minimal norm is characterized by the Bernstein condition* [1] *that* $\lambda_1 = \ldots = \lambda_n$, *which is produced by a unique choice of nodes;*

(ii) *the quantities* $\lambda_1, \ldots, \lambda_n$ *obey the Erdős condition* [3] *that if one of them is greater than the common value given in* (i), *another is less;*

(iii) *the norm of interpolation is governed by the ratio* $(b-a)/(t_{n+1}-b)$. *Specifically, the norm increases without bound as* $b \mapsto t_{n+1}$ *and decreases as* $b \to a$, *with lower limit equal to the norm of optimal Lagrange interpolation with polynomials of degree n or less.*

Corollary 1. *The above theorem also holds when the space of interpolation consists of all multiples of the function*

$$(t - t_{n+1})^{k_1} \ldots (t - t_{n+m})^{k_m}$$

by a polynomial of degree n or less, with $k_j < 0$ *for* $j \in \{1, \ldots, n\}$.

Corollary 2. *Some or all of the points* $t_{n+1}, \ldots, t_{n+m}$ *can be to the left of* t_0 *as well as to the right of* t_n, *and the above results are still valid.*

Proof of Theorem. One notes that the functions

$$\partial \lambda_i / \partial t_j = -y_j(T_i) X_i'(t_j), \quad i \in \{1, \ldots, n\}, \quad j \in \{0, \ldots, n\}$$

exist and are continuous in $t_0, \ldots, t_n$. The points $T_0, \ldots, T_n$, of course, depend in an analytic fashion upon the nodes.

All of our results will follow from properties of various submatrices of

$$A = (\partial \lambda_i / \partial t_j)_{i,j=1}^n,$$

which represents the derivative of the function

$$(t_1, \ldots, t_n) \mapsto (\lambda_1, \ldots, \lambda_n).$$

We define A_i for $i\in\{1, \ldots, n\}$ to be the matrix obtained by deleting the i^{th} column and n^{th} row of A. To prove (i) and (ii) of the theorem, it suffices to show

(1) $\det A_i \neq 0$ for $i\in\{1, \ldots, n\}$ for arbitrary $t_0, \ldots, t_{n+m}$,

and

(2) $\det A_i$ alternates in sign on $\{1, \ldots, n\}$.

To prove (iii), it is enough to prove

(3) $\det A \neq 0$.

To show (1) and (2), we first perform some row and column cancellations. For $j\in\{1, \ldots, n\}$, the j^{th} row of A is given by

$$-y_j(T_1)X_1'(t_j)\ldots -y_j(T_n)X_n'(t_j).$$

It is possible therefore to multiply the j^{th} row by the "denominator" of y_j, namely by

$$\frac{1}{Q(t_j)}\sum_{\substack{l=0\\ l\neq j}}^{n+m}(t_j-t_l).$$

When this procedure has been completed, the i^{th} column, for $i\in\{1, \ldots, n+m\}$ is of the form

$$\frac{1}{Q(T_i)}\prod_{\substack{j=0\\ j\neq 1}}^{n}(T_i-t_j)X_i'(t_1)$$

$$\ldots$$

$$\frac{1}{Q(T_i)}\prod_{\substack{j=0\\ j\neq n}}^{n}(T_i-t_j)X_i'(t_n),$$

and the non-zero quantity $\prod_{j=0}^{n}(T_i-t_j)$ may be divided from the i^{th} column. Following this operation by multiplication of the i^{th} column by $Q(T_i)$, the matrix is left in the form

$$B=\begin{bmatrix}\dfrac{X_1'(t_1)}{t_1-T_1} & \cdots & \dfrac{X_n'(t_1)}{t_1-T_n}\\ \cdots & \cdots & \cdots\\ \dfrac{X_1'(t_n)}{t_n-T_1} & \cdots & \dfrac{X_n'(t_n)}{t_n-T_n}\end{bmatrix}.$$

Now, it is possible to multiply the j^{th} row by $(Q(t_j))^2$, and the expression

$$q_i(t)=\frac{X_i'(t)}{t-T_i}(Q(t))^2,\quad i\in\{1, \ldots, n+m\}$$

is a polynomial of degree $n+m-2$ or less which is evaluated at the successive points $t_1, \ldots, t_n$ down the i^{th} column of the matrix.

Clearly, the roots of the polynomials $(t-T_i)q_i(t)$ will strictly interlace on the interval $[T_1, \infty]$, and it is possible to choose points $T_{n+1}, \ldots, T_{n+m}$ with

$$T_n < t_n < T_{n+1} < \ldots < T_{n+m}$$

such that the following conditions are satisfied by $q_1, \ldots, q_n$.

(i) Each polynomial q_i has exactly one root in each of the subintervals (T_j, T_{j+1}), $j \in \{1, \ldots, n+m-1\}$ of the interval $[T_1, T_{n+m}]$, except that q_i has no root in (T_i, T_{i+1}) for $i \in \{1, \ldots, n+m-1\}$, nor in (T_{i-1}, T_i) for $i \in \{2, \ldots, n+m\}$.

(ii) $q_i(T_j) \neq 0$ for $i, j \in \{1, \ldots, n+m\}$.

Proposition. *Let polynomials* $q_1, \ldots, q_{n+m}$ *and points* $T_1, \ldots, T_{n+m}$ *satisfy* (i) *and* (ii), *and let points* $t_1, \ldots, t_{n-1}$ *be situated so that*

$$T_1 < t_1 < T_2 < \ldots < T_{n-1} < t_{n-1} < T_n.$$

Then, for $k \in \{1, \ldots, n\}$,

$$\det\left(q_i(t_j)\right)_{\substack{1 \le i \le n,\ 1 \le j \le n-1 \\ i \neq k}} \neq 0.$$

A proof of this Proposition appears in [7].

At this point, (1) and (2) follow. To prove (3), we need only to note that, in the present context, $n-1$ may be replaced by n in the above Proposition, with $k=n+1$ and the proposition still holds, permitting one to analyse what occurs as $t_n \to t_0$ or $t_n \to t_{n+1}$, subject to the condition $\lambda_1 = \ldots = \lambda_n$.

This completes the proof of the Theorem. Corollary 1 is now established by a re-examination of the steps of matrix cancellation, leading to a similar system of polynomials $q_1, \ldots, q_n$. Details of a similar argument appear in [7]. Corollary 3 can clearly be obtained by a slight modification of the above Proposition.

References

[1] S. Bernstein, Sur la limitation des valeurs d'un polynome $P_n(x)$ de degre n sur tout un segment par ses valeurs en $n+1$ points du segment, *Isv. Akad. Nauk. SSSR*, **7** (1931), 1025—1050.

[2] C. De Boor and A. Pinkus, Proof of the conjectures of Bernstein and Erdős concerning polynomial interpolation, *J. Approx. Theory*, **24** (1978), 289—303.

[3] P. Erdős, Some remarks on polynomials, *Bull. Amer. Math. Soc.*, **53** (1947), 1169—1176.

[4] T. A. Kilgore, Optimization of the Lagrange interpolation operator, *Bull. Amer. Math. Soc.*, **83** (1977), 1069—1071.

[5] T. A. Kilgore, A characterization of the Lagrange interpolating projection with minimal Tchebycheff norm, *J. Approx. Theory*, **24** (1978), 273—288.

[6] T. A. Kilgore, Optimal interpolation with incomplete polynomials, *J. Approx. Theory*, **41** (1984), 279—290.

[7] T. A. Kilgore, Optimal interpolation with polynomials having fixed roots, *J. Approx. Theory*, submitted for publication.

THE OHIO STATE UNIVERSITY
MARION, OHIO, USA

Acta Sci. Math., 52 (1988), 117—140

Operators of Toeplitz and Hankel type

VLASTIMIL PTÁK and PAVLA VRBOVÁ

In the present note the authors investigate abstract analoga of classical Toeplitz and Hankel operators and extend to these more general classes some of the results known from the classical theory. The investigation is based on the use of isometric dilations of contractions and on the properties of their Wold decompositions. In particular the unitary part of the isometric dilation plays a decisive role. To explain the genesis and motivation of our investigation let us recall some of the classical facts which are essential for our considerations.

We denote by $T(\varphi)$ the Toeplitz operator on H^2 defined for $f\in H^2$ by the formula $T(\varphi)f=P_+\varphi f$ where P_+ stands for the projection operator of L^2 onto H^2 and φ is an L^∞ function, the symbol of $T(\varphi)$. The projection onto the orthogonal complement $H^2_-=L^2\ominus H^2$ will be denoted by P_-. Since $P_+\bar{z}P_-=0$ we have $P_+\bar{z}P_+\varphi(z)zf(z)=$ $=P_+\varphi(z)f(z)$ for every $f\in H^2$. If S stands for the shift operator (multiplication by z) on H^2 this relation may be restated in the form

$$S^*T(\varphi)S = T(\varphi)$$

and it turns out that the relation $S^*AS=A$ is characteristic for Toeplitz operators on H^2.

There is another important class of operators which may be characterized by a similar relation. Hyponormal operators are defined by the inequality $TT^*\leqq T^*T$ and may accordingly be characterized by the existence of a contraction C such that $T^*=CT$. Hence

$$CTC^* = T^*C^* = (CT)^* = T$$

so that T satisfies a relation of the same type.

In a paper on hyponormal operators [4] C. FOIAȘ and B. SZ.-NAGY used dilation theory to show that for each hyponormal operator T acting on a Hilbert space $\mathfrak{H}$ there exists a normal operator N on a suitable Hilbert space $\mathfrak{G}$, a unitary operator

Received November 20, 1984, and in revised form April 23, 1986.

U on $\mathfrak{G}$ and a contraction X: $\mathfrak{H}\to\mathfrak{G}$ such that $T=X^*NX$ while $N^*=UN$ and $\|X^*Ug\|\leqq\|X^*g\|$ for all $g\in\mathfrak{G}$.

The relation $T=X^*NX$ is clearly an analogon of the formula $T=P_+\varphi|H^2$: X replaces the injection operator $H^2\to L^2$, N replaces the symbol φ and X^* plays the role of a projection of $\mathfrak{G}$ onto $\mathfrak{H}$. Starting from this observation these two authors developed a theory of Toeplitz type operators [3] where the L^∞ function φ is replaced by an abstract symbol.

It is interesting to note that the relation $N^*=UN$ implies $N=UNU^*$. Indeed, $N=(UN)^*=N^*U^*=(UN)U^*$; it is one of the purposes of the present note to explain the importance of this relation.

In the classical case there is a parallel theory for Hankel operators. Starting from a $\varphi\in L^\infty$ we define $H(\varphi)$: $H^2\to H^2_-$ by the formula $H(\varphi)f=P_-\varphi f$. Since $P_-zP_+=0$ we have

$$H(\varphi)zf = P_-z\varphi(z)f(z) = P_-zP_-\varphi(z)f(z) = P_-zH(\varphi)f$$

so that

$$H(\varphi)S = ZH(\varphi)$$

if Z denotes the operator $g\to P_-zg$ on H^2_-. Again, this relation turns out to be characteristic for Hankel operators from H^2 into H^2_-.

In the present paper we intend to show that the class of Hankel operator also has an abstract analogon and propose to out line a theory of symbols for operators of Toeplitz and Hankel types.

To obtain the symbol for an operator A on H^2 satisfying the relation $S^*AS=A$ we first use this relation to extend A to the whole of L^2; it turns out that this extension commutes with the shift so that it coincides with the operator of multiplication by an L^∞ function φ. The operator A appears then as a compression to H^2 of this multiplication operator $M(\varphi)$.

In the sequel we shall view the symbol of A as this multiplication operator rather than the function generating it — this is possible in view of the isometric isomorphism between L^∞ taken as an algebra and the corresponding algebra of multiplication operators.

To obtain a symbol for an operator X: $H^2\to H^2_-$ satisfying $XS=ZX$ we use first the theorem on intertwining dilations to obtain an operator from H^2 into L^2 intertwining S and $M(z)$; extending its domain of definition to the whole of L^2 we obtain an operator which commutes with $M(z)$ and which yields the original operator as a compression, this time from H^2 into H^2_-.

Observe that $Z=P_-M(z)|H^2_-$ and that $M(\bar{z})=M(z)^*$ is the minimal isometric dilation of S^*. A similar situation obtains in the general case.

In a manner of speaking the construction of symbols for generalized Toeplitz and Hankel operators proceeds — in its early stages — along similar lines as in the classical theory; at a certain point, however, difficulties present themselves which have no counterpart in the classical case. In particular, a relation of the type $XS=ZX$ alone is not sufficient to characterize a class with satisfactory properties. We intend to show that, in the general case of abstract Toeplitz and Hankel operators, it is also possible to construct a symbol which is characterized by a certain commutativity relation and as a compression of which the given operator may be reconstructed.

The investigations of B. Sz.-Nagy and C. Foiaş indicate the important role played by the space $\mathfrak{R}$, the unitary part of the Wold decomposition of the isometric dilation of the contraction T by means of which the abstract Toeplitz operator is defined. The results of the present note seem to confirm the hypothesis that this space forms the natural domain of definition for operators which should play the role of an abstract symbol both for Toeplitz and Hankel operators. The main difficulty seems to lie in the fact that the Wold decomposition is trivial in the classical case, the isometric dilation of S^* being unitary, so that little help can be expected from immediate analogies.

It turns out that the methods presented below work even in the more general case when a Toeplitz operator X is defined by the relation $X=T_1XT_2^*$ where T_1 and T_2 are two arbitrary contractions acting on the spaces $\mathfrak{H}_1$ and $\mathfrak{H}_2$ which may be different from each other in general. In this manner we hope to eliminate results whose validity is essentially based on the equality $T_1=T_2$; at the same time, this generality does not seem to be excessive. We still obtain analoga of the Kronecker theorem as well as of the identity

$$T(\varphi\psi)-T(\varphi)T(\psi) \doteq H(\varphi^*)^*H(\psi).$$

In a paper on operator equations [1] R. G. DOUGLAS considered operators satisfying $X=T_1XT_2^*$. His investigations proceed along different lines; nevertheless, his ideas provided inspiration for some of our methods.

The paper is divided into five sections. In the first section we list some technical facts from dilation theory which will be needed in the main text.

Section two contains a short exposé of the theory of Toeplitz operators. In spite of the fact that the emphasis of this note is on Hankel operators it is, in our opinion, useful to include this short section. Our approach differs in details from that of Sz.-Nagy and Foiaş, the differences being motivated by the necessity to prepare the ground for the theory of Hankel operators. Since we intend to represent Toeplitz and Hankel operators as compressions of their symbols (like in the classical case) we use the term symbol in a slightly different way — nevertheless there is a one-to-one correspondence between symbols in our sense and those used by Sz.-Nagy and Foiaş. This makes it possible to present a unified theory for both types of operators.

Section 3 contains the definition and basic properties of Hankel opreators including a generalization of the Nehari theorem. The last two chapters are devoted to an analogy of analytic symbols and to a generalization of the Kronecker theorem.

1. Preliminaries

We start by recalling some of the properties of the minimal isometric dilation of a contraction. Let $\mathfrak{H}$ be a Hilbert space, $T\in\mathscr{B}(\mathfrak{H})$ be a contraction. We denote by U the minimal isometric dilation of T on the space $\mathfrak{K}^+$, i.e. an isometry U defined on a space $\mathfrak{K}^+\supset\mathfrak{H}$ satisfying

$$T^n = P(\mathfrak{H})U^n|\mathfrak{H} \quad \text{for} \quad n = 0, 1, \ldots$$

and

$$\mathfrak{K}^+ = \bigvee_{k\geqq 0} U^k\mathfrak{H}.$$

We shall denote by $P(\mathfrak{L})$ the orthogonal projection of $\mathfrak{K}^+$ onto a subspace $\mathfrak{L}\subset\mathfrak{K}^+$.

Any two minimal isometric dilations of a given contraction are unitarily equivalent.

We shall frequently use the following facts:

(1) $TP(\mathfrak{H})=P(\mathfrak{H})U$;

(2) $U^*\mathfrak{H}\subset\mathfrak{H}$ and $U^*|\mathfrak{H}=T^*$;

(3) the subspace $\mathfrak{H}^\perp=\mathfrak{K}^+\ominus\mathfrak{H}$ can be decomposed as follows

$$\mathfrak{H}^\perp = \mathfrak{L}\oplus U\mathfrak{L}\oplus U^2\mathfrak{L}\oplus\ldots,$$

where $\mathfrak{L}=((U-T)\mathfrak{H})^-$;

(4) $U\mathfrak{H}^\perp\subset\mathfrak{H}^\perp$ and $U|\mathfrak{H}^\perp$ is a unilateral shift of multiplicity $\dim\mathfrak{L}$;

(5) the sequence $\{P(\mathfrak{H}^\perp)U^{*n}\}$ tends to zero in the strong operator topology;

(6) T^* is an isometry if and only if the minimal isometric dilation of T is a unitary operator;

(7) let W be a unitary operator on a Hilbert space $\mathfrak{G}$, let $\mathfrak{M}\subset\mathfrak{G}$ be a subspace invariant with respect to W; then the restriction of W^* to the W^* invariant subspace of $\mathfrak{G}$ generated by $\mathfrak{M}$ is the minimal isometric dilation of the operator $(W|\mathfrak{M})^*$.

The reader is referred to [2] for proofs of (1)—(4).

For lack of space the proofs of the remaining results in this section have to be left to the reader.

If S is an arbitrary isometry on a Hilbert space $\mathfrak{K}$ then the Wold decomposition applies. In other words, the space $\mathfrak{K}$ can be decomposed into a direct sum of two subspaces reducing with respect to S,

$$\mathfrak{K} = \Big(\bigcap_{n\geqq 0} S^n\mathfrak{K}\Big)\oplus\big((\mathfrak{K}\ominus S\mathfrak{K})\oplus(S\mathfrak{K}\ominus S^2\mathfrak{K})\oplus\ldots\big)$$

so that the restriction of S to the first subspace is a unitary operator and the restriction to the second one is a unilateral shift.

Now, let $\mathfrak{R}$ be the reducing subspace in the Wold decomposition of the minimal isometric dilation U on $\mathfrak{K}^+$ on which U is unitary, i.e. $\mathfrak{R}=\bigcap_{n\geq 0} U^n\mathfrak{K}^+$. Then we have (see [2]):

(8) $UP(\mathfrak{R})=P(\mathfrak{R})U,\ U^*P(\mathfrak{R})=P(\mathfrak{R})U^*$;

(9) the sequence of projections $\{U^nU^{*n}\}_0^\infty$ is decreasing,
$P(\mathfrak{R})\leqq U^nU^{*n}$ for $n=0,1,\ldots$
and
$P(\mathfrak{R})k=\lim_{n\to\infty} U^nU^{*n}k$ for all $k\in\mathfrak{K}^+$;

(10) $P(\mathfrak{R})h=\lim_{n\to\infty} U^nT^{*n}h$ for all $h\in\mathfrak{H}$.

There are two subspaces of the space $\mathfrak{R}$ which play an important role in our investigations, namely, $(P(\mathfrak{R})\mathfrak{H})^-$ and $\mathfrak{H}\cap\mathfrak{R}$. Denote by R the restriction of U onto the subspace $\mathfrak{R}$.

1.1. Lemma (see also [3]). *The operator U^* maps $P(\mathfrak{R})\mathfrak{H}$ into itself and $U^*|(P(\mathfrak{R})\mathfrak{H})^-$ is an isometry. The sequence of linear manifolds $\{R^nP(\mathfrak{R})\mathfrak{H}\}_{-\infty}^\infty$ is increasing and*

$$\mathfrak{R}=\Big(\bigcup_{n\geq 0} R^nP(\mathfrak{R})\mathfrak{H}\Big)^-.$$

If T is a contraction on a Hilbert space $\mathfrak{H}$ then $\mathfrak{H}$ can be uniquely decomposed into an orthogonal sum of two subspaces reducing T, $\mathfrak{H}=\mathfrak{H}_u\oplus\mathfrak{H}_s$ such that $T|\mathfrak{H}_u$ is unitary and $T|\mathfrak{H}_s$ is completely non-unitary. We have

$$\mathfrak{H}_u=\{h\in\mathfrak{H}\colon \|T^nh\|=\|T^{*n}h\|=\|h\| \text{ for all } n\geqq 0\}.$$

See [2].

1.2. Lemma. *We have*

$$\mathfrak{H}\cap\mathfrak{R}=\mathfrak{H}\cap P(\mathfrak{R})\mathfrak{H}=\{h\in\mathfrak{H}\colon \|T^{*n}h\|=\|h\| \textit{ for all } n\geqq 0\}=$$
$$=\{h\in\mathfrak{H}\colon T^nT^{*n}h=h \textit{ for all } n\geqq 0\}.$$

The subspace $\mathfrak{H}\cap\mathfrak{R}$ is invariant with respect to U^* and $U^*|\mathfrak{H}\cap\mathfrak{R}$ is an isometry whose Wold decomposition has the form

$$\mathfrak{H}\cap\mathfrak{R}=\mathfrak{H}_u\oplus(\mathfrak{N}\oplus U^*\mathfrak{N}\oplus U^{*2}\mathfrak{N}\oplus\ldots)$$

where $\mathfrak{N}=(\mathfrak{H}\cap\mathfrak{R})\ominus U^*(\mathfrak{H}\cap\mathfrak{R})$.

We close this section with two results of a different character which we shall use later. The first is a technical proposition based on the following observation. We have, for each complex number α,

$$U(1-\alpha T)-(1-\alpha U)T=U-T.$$

If $|\alpha|<1$ this relation implies

$$(1-\alpha U)^{-1}(U-T)(1-\alpha T)^{-1} = (1-\alpha U)^{-1}U|\mathfrak{H}-T(1-\alpha T)^{-1}.$$

1.3. Proposition. *Let T be a contraction on a Hilbert space with the minimal isometric dilation U. Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be complex numbers of modulus less than 1. Then*

$$U^n(1-\alpha_1 U)^{-1}\ldots(1-\alpha_n U)^{-1}|\mathfrak{H}-T^n(1-\alpha_1 T)^{-1}\ldots(1-\alpha_n T)^{-1} =$$

$$= \sum_{k=1}^{n} U^{k-1}(1-\alpha_1 U)^{-1}\ldots(1-\alpha_k U)^{-1}(U-T)T^{n-k}(1-\alpha_k T)^{-1}\ldots(1-\alpha_n T)^{-1}.$$

1.4. Proposition. *Let $\mathfrak{H}_1, \mathfrak{H}_2, \mathfrak{K}_1, \mathfrak{K}_2$ be Hilbert spaces, $X\in\mathscr{B}(\mathfrak{H}_1, \mathfrak{H}_2)$, $A_1\in\mathscr{B}(\mathfrak{H}_1, \mathfrak{K}_1)$, $A_2\in\mathscr{B}(\mathfrak{H}_2, \mathfrak{K}_2)$. If $|(Xh_1, h_2)|\leqq\|A_1h_1\|\cdot\|A_2h_2\|$ for all $h_1\in\mathfrak{H}_1$, $h_2\in\mathfrak{H}_2$ then there exists a contraction operator $C\colon(\mathrm{Ran}\, A_1)^-\to(\mathrm{Ran}\, A_2)^-$ for which $X=A_2^*CA_1$.*

2. Toeplitz operators and their symbols

Consider two contractions $T_1\in\mathscr{B}(\mathfrak{H}_1)$, $T_2\in\mathscr{B}(\mathfrak{H}_2)$; denote by U_1 and U_2 their minimal isometric dilations acting on the spaces $\mathfrak{K}_1^+$, $\mathfrak{K}_2^+$ respectively. We denote by $\mathfrak{R}_1$ and $\mathfrak{R}_2$ the subspaces of $\mathfrak{K}_1^+$ and $\mathfrak{K}_2^+$ which reduce U_1 and U_2 to their unitary parts R_1 and R_2. We denote by $P(\mathfrak{Z})$ the orthogonal projection of $\mathfrak{K}_j^+$ onto a subspace $\mathfrak{Z}\subset\mathfrak{K}_j^+$.

2.1. Proposition. *Consider the set $\mathscr{S}(T_1, T_2)$ of all operators $Z\in\mathscr{B}(\mathfrak{R}_2, \mathfrak{R}_1)$ satisfying the condition*

$$ZR_2 = R_1Z,$$

and the set $\mathscr{S}'(T_1, T_2)$ of all operators $Y\in\mathscr{B}(\mathfrak{K}_2^+, \mathfrak{K}_1^+)$ satisfying

$$Y = U_1YU_2^*.$$

If $Y\in\mathscr{B}(\mathfrak{K}_2^+, \mathfrak{K}_1^+)$ then the following four conditions are equivalent:

1° $Y\in\mathscr{S}'(T_1, T_2)$;

2° $YU_2 = U_1Y$ *and* $Y=YP(\mathfrak{R}_2)$;

3° $YU_2^*=U_1^*Y$ *and* $Y=P(\mathfrak{R}_1)Y$;

4° $Y=\lim U_1^nP(\mathfrak{H}_1)YP(\mathfrak{H}_2)U_2^{*n}$ *in the strong operator topology.*

Furthermore

5° *if* $Z\in\mathscr{S}(T_1, T_2)$ *then* $ZP(\mathfrak{R}_2)\in\mathscr{S}'(T_1, T_2)$;

6° *if* $Y\in\mathscr{S}'(T_1, T_2)$ *then* $Y|\mathfrak{R}_2\in\mathscr{S}(T_1, T_2)$.

Proof. If $Y\in\mathscr{B}(\mathfrak{K}_2^+, \mathfrak{K}_1^+)$ satisfies $Y=U_1YU_2^*$ then $YU_2=U_1Y$. Also, for each $x\in\mathfrak{K}_2^+$,

$$YP(\mathfrak{R}_2)x = \lim YU_2^nU_2^{*n}x =$$

$$= \lim U_1^nYU_2^{*n}x = \lim Yx = Yx.$$

On the other hand, $YU_2=U_1Y$ and $YP(\mathfrak{R}_2)=Y$ implies

$$U_1YU_2^* = YU_2U_2^* = YP(\mathfrak{R}_2)U_2U_2^* = Y$$

since $U_2U_2^*\geqq P(\mathfrak{R}_2)$. This proves the equivalence of 1° and 2°.

$Y\in\mathscr{S}'(T_1, T_2)$ if and only if $Y^*\in\mathscr{S}'(T_2, T_1)$. The last inclusion is equivalent to $Y^*U_1=U_2Y^*$ and $Y^*=Y^*P(\mathfrak{R}_1)$. Taking adjoints we obtain the equivalence of 1° and 3°.

The implication 4°$\Rightarrow$1° is obvious. On the other hand, if $Y=U_1YU_2^*$ then, for $n\geqq 0$,

$$Y = U_1^nYU_2^{*n} =$$

$$= U_1^nP(\mathfrak{H}_1)YP(\mathfrak{H}_2)U_2^{*n}+U_1^n(1-P(\mathfrak{H}_1))YU_2^{*n}+$$

$$+U_1^nP(\mathfrak{H}_1)Y(1-P(\mathfrak{H}_2))U_2^{*n} =$$

$$= U_1^nP(\mathfrak{H}_1)YP(\mathfrak{H}_2)U_2^{*n}+U_1^n(1-P(\mathfrak{H}_1))U_1^{*n}Y+$$

$$+U_1^nP(\mathfrak{H}_1)Y(1-P(\mathfrak{H}_2))U_2^{*n}.$$

Both $(1-P(\mathfrak{H}_1))U_1^{*n}$ and $(1-P(\mathfrak{H}_2))U_2^{*n}$ tend to zero in the strong operator topology.

Now suppose $Z\in\mathscr{B}(\mathfrak{R}_2, \mathfrak{R}_1)$ satisfies $ZR_2=R_1Z$. Then $Y=ZP(\mathfrak{R}_2)$ satisfies $YU_2=ZP(\mathfrak{R}_2)U_2=ZU_2P(\mathfrak{R}_2)=U_1ZP(\mathfrak{R}_2)=U_1Y$ and $YP(\mathfrak{R}_2)=Y$. It follows from 2° that $Y\in\mathscr{S}'(T_1, T_2)$.

If $Y\in\mathscr{S}'(T_1, T_2)$ we have, for each n and each $x\in\mathfrak{K}_2^+$, $Yx=U_1^nYU_2^{*n}x\in U_1^n\mathfrak{K}_1^+$ so that the range of Y is contained in $\mathfrak{R}_1$. Since $Y=U_1YU_2^*$ we have $YU_2=U_1Y$ and, in view of the inclusion $Yx\in\mathfrak{R}_1$ for each x, this implies

$$(Y|\mathfrak{R}_2)R_2 = R_1(Y|\mathfrak{R}_2)$$

as asserted. The proof is complete.

2.2. Remark. The correspondence between elements of sets $\mathscr{S}$ and $\mathscr{S}'$ described in 5° and 6° is contractive in both directions and so it is an isometric linear mapping.

2.3. Definition. An element of the set $\mathscr{S}'(T_1, T_2)$ will be called a symbol with respect to T_1, T_2.

2.4. Proposition. *Let* $Y=\mathscr{S}'(T_1, T_2)$ *be a symbol. Denote by*

$$A = P(\mathfrak{H}_1)Y|\mathfrak{H}_2, \quad B = P(\mathfrak{H}_1^\perp)Y|\mathfrak{H}_2.$$

Then

$$A = T_1 A T_2^*, \tag{12}$$

$$(U_1 | \mathfrak{H}_1^\perp)^* B = B T_2^*. \tag{13}$$

Moreover, there exists a positive K such that A satisfies the estimate

$$\|A h_2\| \leqq K \cdot \|P(\mathfrak{R}_2) h_2\| \tag{14}$$

for all $h_2 \in \mathfrak{H}_2$ and similarly, B satisfies

$$(B h_2, h_1^\perp) \leqq K \cdot \|P(\mathfrak{R}_2) h_2\| \cdot \|P(\mathfrak{R}_1) h_1^\perp\| \tag{15}$$

for $h_2 \in \mathfrak{H}_2$, $h_1^\perp \in \mathfrak{H}_1^\perp$.

Proof. Using the relation $T_1 P(\mathfrak{H}_1) = P(\mathfrak{H}_1) U_1$ we have, for $h_2 \in \mathfrak{H}_2$,

$$T_1 P(\mathfrak{H}_1) Y T_2^* h_2 = P(\mathfrak{H}_1) U_1 Y U_2^* h_2 = P(\mathfrak{H}_1) Y h_2$$

which proves (12). Similarly, using the inclusion $U_1^* \mathfrak{H}_1 \subset \mathfrak{H}_1$,

$$\begin{aligned} B T_2^* h_2 &= P(\mathfrak{H}_1^\perp) Y U_2^* h_2 = P(\mathfrak{H}_1^\perp) U_1^* Y h_2 = \\ &= P(\mathfrak{H}_1^\perp) U_1^* P(\mathfrak{H}_1^\perp) Y h_2 + P(\mathfrak{H}_1^\perp) U_1^* P(\mathfrak{H}_1) Y h_2 = \\ &= P(\mathfrak{H}_1^\perp) U_1^* P(\mathfrak{H}_1^\perp) Y h_2 = P(\mathfrak{H}_1^\perp) U_1^* B h_2 = \\ &= (U_1 | \mathfrak{H}_1^\perp)^* B h_2. \end{aligned}$$

The estimates (14) and (15) with $K = \|Y\|$ are immediate consequences of the relation

$$Y = Y P(\mathfrak{R}_2) = P(\mathfrak{R}_1) Y P(\mathfrak{R}_2).$$

It is interesting to observe that the estimate (14) is a consequence of (12). On the other hand, we shall see that condition (13) alone does not imply (15).

2.5. Remark. *If $A = T_1 A T_2^*$ then*

$$\|A h_2\| \leqq \|A\| \|P(\mathfrak{R}_2) h_2\|$$

for each $h_2 \in \mathfrak{H}_2$.

Proof. For each $h_2 \in \mathfrak{H}_2$ and each natural number n,

$$A h_2 = T_1^n A T_2^{*n} h_2$$

so that $\|A h_2\| \leqq \|A\| \|T_2^{*n} h_2\| = \|A\| \cdot \|U_2^n T_{2n}^* h_2\|$. Since $P(\mathfrak{R}_2) h_2 = \lim U_2^n T_2^{*n} h_2$ the assertion follows.

2.6. Example. Let $\mathfrak{H}_1, \mathfrak{H}_2$ be Hilbert spaces, $T_1 \in \mathscr{B}(\mathfrak{H}_1)$ be such that T_1^* is a nonunitary isometry. Then the minimal isometric dilation U_1 of T_1 is a unitary operator acting on the space $\mathfrak{K}_1^+$, $\mathfrak{R}_1 = \mathfrak{K}_1^+$ and the operator $V_1^* = (U_1 | \mathfrak{H}_1^\perp)^*$ has a nontrivial kernel.

The operator $T_2=0$ on $\mathfrak{H}_2$ is a contraction whose isometric dilation U_2 is a unilateral shift on the space $\mathfrak{K}_2^+$. The subspace $\mathfrak{R}_2$ reduces to $\{0\}$ and the operator $V_2^*=(U_2|\mathfrak{H}_2^\perp)^*$ again has a nontrivial kernel.

Take an arbitrary nonzero operator $X\in\mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1^\perp)$ such that $\operatorname{Ran} X\subset\operatorname{Ker} V_1^*$. Then obviously $V_1^*X=0=XT_2^*$ and $\|X^*h_1^\perp\|\leqq\|X^*\|\|h_1^\perp\|=\|X^*\|\|P(\mathfrak{R}_1)h_1^\perp\|$ for all $h_1^\perp\in\mathfrak{H}_1^\perp$. Since $\mathfrak{R}_2=\{0\}$ the operator X does not satisfy $\|Xh_2\|\leqq k\cdot\|P(\mathfrak{R}_2)h_2\|$ for all $h_2\in\mathfrak{H}_2$ with any positive k.

Similarly, let Y be any nonzero operator from $\mathscr{B}(\mathfrak{H}_1, \mathfrak{H}_2^\perp)$ which is zero on $(T_1^*\mathfrak{H}_1)^-$ and with values in $\operatorname{Ker} V_2^*$. Then $V_2^*Y=0=YT_1^*$, $\|Yh_1\|\leqq\|Y\|\|h_1\|=\|Y\|\cdot\|P(\mathfrak{R}_1)h_1\|$ for all $h_1\in\mathfrak{H}_1$, but Y^* does not satisfy $\|Y^*h_2^\perp\|\leqq k\|P(\mathfrak{R}_2)h_2^\perp\|$ on $\mathfrak{H}_2^\perp$ with any positive k.

2.7. Definition. Denote by

$$\mathscr{T}(T_1, T_2) = \{A\in\mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1)\colon A=T_1AT_2^*\}.$$

Operators from the set $\mathscr{T}(T_1, T_2)$ will be called *Toeplitz operators with respect to* T_1, T_2.

Further, operators $B\in\mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1^\perp)$ satisfying (13), i.e. $(U_1|\mathfrak{H}_1^\perp)^*B=BT_2^*$ and

$$\|(Bh_2, h_1^\perp)\| \leqq \gamma\|P(\mathfrak{R}_2)h_2\|\,\|P(\mathfrak{R}_1)h_1^\perp\|$$

for all $h_2\in\mathfrak{H}_2$, $h_1^\perp\in\mathfrak{H}_1^\perp$ and a suitable constant γ will be called *Hankel operators with respect to* T_1, T_2. Similarly, the family of all Hankel operators will be denoted by $\mathscr{H}(T_1, T_2)$.

2.8. Lemma. *Suppose* $X\in\mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1)$ *satisfies the relation*

$$X = T_1XT_2^*.$$

Then there exists exactly one operator $\tilde{X}\colon \mathfrak{K}_2^+\to\mathfrak{H}_1$ *with the following three properties:*

1° $\tilde{X}=T_1\tilde{X}U_2^*$,

2° $X=\tilde{X}|\mathfrak{H}_2$,

3° $\|\tilde{X}\|=\|X\|$.

Conversely, if $\tilde{X}\in\mathscr{B}(\mathfrak{K}_2^+, \mathfrak{H}_1)$ *satisfies* 1° *and if* X *is defined by* 2° *then* $X=T_1XT_2^*$ *and* $\|X\|=\|\tilde{X}\|$.

Proof. Suppose first that $\tilde{X}$ satisfies 1° and X is defined by 2°. Then, for $h_2\in\mathfrak{H}_2$,

$$Xh_2 = \tilde{X}h_2 = T_1\tilde{X}U_2^*h_2 = T_1\tilde{X}T_2^*h_2 = T_1XT_2^*h_2.$$

It follows that $X=T_1XT_2^*$.

Further, given $n\geqq 0$, $h_2\in\mathfrak{H}_2$, we have

$$\tilde{X}U_2^nh_2 = T_1^n\tilde{X}U_2^{*n}U_2^nh_2 = T_1^n\tilde{X}h_2 = T_1^nXh_2.$$

This together with $\mathfrak{K}_2^+ = \operatorname{span}_{n\geq 0} U_2^n h_2$ proves that there is at most one $\tilde{X}$ satisfying 1° and 2° for a given X.

Now, suppose $X \in \mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1)$ satisfies $X = T_1 X T_2^*$. To prove the existence of $\tilde{X}$ it would be sufficient to define it for finite sums of elements $U_2^k h$ $(k \geq 0, h \in \mathfrak{H}_2)$ which form a dense subset in $\mathfrak{K}_2^+$. Let $m \geq k$, $h \in \mathfrak{H}_2$, then

$$T_1^k X h = T_1^k (T_1^{m-k} X T_2^{*m-k}) h = T_1^m X U_2^{*m-k} h = T_1^m X U_2^{*m-k} (U_2^{*k} U_2^k) h =$$

$$= T_1^m X U_2^{*m} U_2^k h$$

and consequently,

$$\sum_0^m T_1^k X h_k = T_1^M X U_2^{*M} \Big(\sum_0^m U_2^k h_2\Big)$$

for each $M \geq m$ and $h_k \in \mathfrak{H}_2$. In particular,

$$\Big\|\sum_0^m T_1^k X h_k\Big\| \leq \|X\| \, \Big\|\sum_0^m U_2^k h_k\Big\|.$$

It follows that the operator $\tilde{X}$ defined on $\lim_{k\geq 0} U_2^k \mathfrak{H}_2$ by $\tilde{X} \sum_0^m U_2^k h_k = \sum_0^m T_1^k X h_k$ is well defined, $\tilde{X} h = X h$ for $h \in \mathfrak{H}_2$ and $\|\tilde{X}\| \leq \|X\|$ so that $\|\tilde{X}\| = \|X\|$. Moreover,

$$T_1 \tilde{X} U_2^* \Big(\sum_0^m U_2^k h_k\Big) = T_1 \tilde{X} \sum_1^m U_2^{k-1} h_k + T_1 \tilde{X} U_2^* h_0 =$$

$$= T_1 \sum_1^m T_1^{k-1} X h_k + T_1 X T_2^* h_0 = \sum_1^m T_1^k X h_k + T_1 X T_2^* h_0 =$$

$$= \sum_1^m T_1^k X h_k + X h_0 = \tilde{X} \sum_0^m U_2^k h_k.$$

The proof is complete.

2.9. Remark. The preceding lemma can be reformulated in a dual version. Namely, if $X \in \mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1)$ satisfies $X = T_1 X T_2^*$ then $X^* \in \mathscr{B}(\mathfrak{H}_1, \mathfrak{H}_2)$ satisfies $X^* = T_2 X^* T_1^*$. It follows that there exists exactly one operator $W^* \in \mathscr{B}(\mathfrak{K}_1^+, \mathfrak{H}_2)$ such that

$$W^* = T_2 W^* U_1^*, \quad X^* = W^* | \mathfrak{H}_1, \quad \|X^*\| = \|W^*\|,$$

or equivalently,

$$W \in \mathscr{B}(\mathfrak{H}_2, \mathfrak{K}_1^+), \quad W = U_1 W T_2^*,$$

$$X = P(\mathfrak{H}_1) W, \quad \|X\| = \|W\|.$$

2.10. Theorem. *Suppose* $X \in \mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1)$ *satisfies*

$$X = T_1 X T_2^*.$$

Then there exists exactly one operator $Y\in\mathscr{B}(\mathfrak{K}_2^+, \mathfrak{K}_1^+)$ *with the following properties*

1° Y *is a symbol with respect to* T_1, T_2,
2° $X=P(\mathfrak{H}_1)Y|\mathfrak{H}_2$,
3° $\|X\|=\|Y\|$.

The operator Y will be called the symbol of X.

Proof. According to 2.8 there exists an $\tilde{X}\in\mathscr{B}(\mathfrak{K}_2^+, \mathfrak{H}_1)$ such that $\tilde{X}=T_1\tilde{X}U_2^*$, $X=\tilde{X}|\mathfrak{H}_2$ and $\|\tilde{X}\|=\|X\|$. Again, according to 2.9 there exists a $Y\in\mathscr{B}(\mathfrak{K}_2^+, \mathfrak{K}_1^+)$ such that

$$Y = U_1YU_2^*, \quad P(\mathfrak{H}_1)Y|\mathfrak{H}_2 = \tilde{X}|\mathfrak{H}_2 = X$$

and

$$\|Y\| = \|\tilde{X}\| = \|X\|.$$

The rest of the proof is straightforward.

Proposition 2.1 and Theorem 2.10 show that there is a one-to-one correspondence between $\mathscr{S}(T_1, T_2)$, $\mathscr{S}'(T_1, T_2)$ and $\mathscr{T}(T_1, T_2)$. Summing up, we have the following

2.11. Theorem. *Let* β: $\mathscr{B}(\mathfrak{K}_2^+, \mathfrak{K}_1^+)\to\mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1)$ *be defined by*

$$\beta Y = P(\mathfrak{H}_1)Y|\mathfrak{H}_2 \quad \textit{for} \quad Y\in\mathscr{B}(\mathfrak{K}_2^+, \mathfrak{K}_1^+).$$

Then β *maps* $\mathscr{S}'(T_1, T_2)$ *isometrically onto* $\mathscr{T}(T_1, T_2)$.

The inverse mapping α *of the restriction of* β *to* $\mathscr{S}'(T_1, T_2)$ *assigns to a Toeplitz operator* $X\in\mathscr{T}(T_1, T_2)$ *its symbol and*

$$\alpha X = \lim U_1^n XP(\mathfrak{H}_2)U_2^{*n}$$

in the strong operator topology.

Proof. Suppose X belongs to $\mathscr{T}(T_1, T_2)$ and that X is generated by a symbol $Y\in\mathscr{S}'(T_1, T_2)$ so that $X=P(\mathfrak{H}_1)Y|\mathfrak{H}_2$. Since $Y=\lim U_1^nP(\mathfrak{H}_1)YP(\mathfrak{H}_2)U_2^{*n}$ and $P(\mathfrak{H}_1)YP(\mathfrak{H}_2)=XP(\mathfrak{H}_2)$ we have $Y=\lim U_1^nXP(\mathfrak{H}_2)U_2^{*n}$.

3. Hankel operators

In this section we intend to develop an analogous theory for generalized Hankel operators. To obtain a symbol for operators of this type we shall apply Lemma 2.8 again, this time to a certain operator of Toeplitz type which we shall construct using the theorem on intertwining dilations; as a consequence of the nonuniqueness of the intertwining dilation a situation analogous to the classical case presents itself: a Hankel operator has more than one symbol in general.

The theory is based on the following lemma, a particular case of which is already contained in [5].

3.1. Lemma. *Let* $\mathfrak{M}_1, \mathfrak{M}_2$ *be two Hilbert spaces,* G_1, G_2 *isometries on* $\mathfrak{M}_1, \mathfrak{M}_2$ *respectively. Denote by* $W_i \in \mathscr{B}(\mathfrak{N}_i)$ *the minimal isometric dilation of* G_i^* *so that the* W_i *are unitary* $(i=1, 2)$.

Suppose $C \in \mathscr{B}(\mathfrak{M}_2, \mathfrak{M}_1)$ *satisfies the relation* $G_1^* C = C G_2$. *Then there exists an operator* $D: \mathfrak{N}_2 \to \mathfrak{N}_1$ *such that*

$$D = W_1^* D W_2^*, \quad \|D\| = \|C\|$$

and

$$C = P_{\mathfrak{M}_1}^{\mathfrak{N}_1} D | \mathfrak{M}_2.$$

Proof. The operator G_2 is its own minimal isometric dilation. By the theorem on intertwining dilations [2] there exists a $\tilde{D}: \mathfrak{M}_2 \to \mathfrak{N}_1$ such that $W_1 \tilde{D} = \tilde{D} G_2$, $P_{\mathfrak{M}_1}^{\mathfrak{N}_1} \tilde{D} = C$ and $\|\tilde{D}\| = \|C\|$. Since W_1 is unitary we may write $\tilde{D} = W_1^* \tilde{D} G_2 = W_1^* \tilde{D} (G_2^*)^*$.

Now apply Lemma 2.8 in the situation $T_1 = W_1^*$, $T_2 = G_2^*$, $\mathfrak{H}_1 = \mathfrak{N}_1$, $\mathfrak{H}_2 = \mathfrak{M}_2$. It follows that there exists a $D: \mathfrak{N}_2 \to \mathfrak{N}_1$ such that $D = W_1^* D W_2^*$, $\tilde{D} = D | \mathfrak{M}_2$ and $\|\tilde{D}\| = \|D\|$. Hence $C = P_{\mathfrak{M}_1}^{\mathfrak{N}_1} \tilde{D} = P_{\mathfrak{M}_1}^{\mathfrak{N}_1} D | \mathfrak{M}_2$ and $\|D\| = \|C\|$.

A linear transformation A from $\mathfrak{H}_2$ into $\mathfrak{H}_1^\perp$ is said to be $\mathfrak{R}$-bounded if there exists a constant α such that

$$|(Ah, k)| \leqq \alpha \|P(\mathfrak{R}_2) h\| \, \|P(\mathfrak{R}_1) k\|$$

for all $h \in \mathfrak{H}_2$ and all $k \in \mathfrak{H}_1^\perp$. The minimum of all α for which the above inequality holds will be called the $\mathfrak{R}$-norm of A and will be denoted by $\|A\|_{\mathfrak{R}}$. Clearly every $\mathfrak{R}$-bounded operator A is norm bounded and its norm does not exceed the $\mathfrak{R}$-norm.

3.2. Theorem. *Suppose* $X \in \mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1^\perp)$ *satisfies*

$$V_1^* X = X T_2^*,$$

where V_1 *is the restriction of* U_1 *to* $\mathfrak{H}_1^\perp$ *and the domination condition*

$$|(X h_2, h_1^\perp)| \leqq \|X\|_{\mathfrak{R}} \|P(\mathfrak{R}_2) h_2\| \, \|P(\mathfrak{R}_1) h_1^\perp\|$$

holds for all $h_2 \in \mathfrak{H}_2$ *and* $h_1^\perp \in \mathfrak{H}_1^\perp$.

Then there exists an operator $Y \in \mathscr{B}(\mathfrak{K}_2^+, \mathfrak{K}_1^+)$ *with the following properties*

$$Y = U_1 Y U_2^*, \quad \|Y\| = \|X\|_{\mathfrak{R}}$$

and

$$X = P(\mathfrak{H}_1^\perp) Y | \mathfrak{H}_2.$$

Proof. Introduce the abbreviations $A_1 = P(\mathfrak{R}_1) | \mathfrak{H}_1^\perp$, $A_2 = P(\mathfrak{R}_2) | \mathfrak{H}_2$, $\mathfrak{M}_1 = (P(\mathfrak{R}_1) \mathfrak{H}_1^\perp)^-$, $\mathfrak{M}_2 = (P(\mathfrak{R}_2) \mathfrak{H}_2)^-$. According to Proposition 1.4 there exists an operator $C \in \mathscr{B}(\mathfrak{M}_2, \mathfrak{M}_1)$ such that $\|C\| = \|X\|_{\mathfrak{R}}$ and $X = A_1^* C A_2$. Thus $V_1^* A_1^* C A_2 =$

$=A_1^* CA_2 T_2^*$. Consider first the product $V_1^* A_1^*$. We have $A_1 V_1 = P(\mathfrak{R}_1) U_1 | \mathfrak{H}_1^\perp = (U_1|\mathfrak{M}_1) P(\mathfrak{R}_1)|\mathfrak{H}_1^\perp = (U_1|\mathfrak{M}|) A_1$ so that $V_1^* A_1^* = A_1^* (U_1|\mathfrak{M}_1)^*$. Furthermore, for $h_2 \in \mathfrak{H}_2$,

$$A_2 T_2^* h_2 = P(\mathfrak{R}_2) U_2^* h_2 = U_2^* P(\mathfrak{R}_2) h_2 = U_2^* A_2 h_2.$$

Thus

$$A_1^* (U_1 | \mathfrak{M}_1)^* C A_2 h_2 = V_1^* A_1^* C A_2 h_2 = A_1^* C A_2 T_2^* h_2 = A_1^* C U_2^* A_2 h_2.$$

Since A_1^* is injective on $\mathfrak{M}_1 = (\text{Ran } A_1)^-$ we have

$$(U_1 | \mathfrak{M}_1)^* C = C(U_2^* | \mathfrak{M}_2).$$

The minimal isometric dilation W_1 of the coisometry $(U_1|\mathfrak{M}_1)^*$ is unitary. Since $\mathfrak{M}_2 \subset \mathfrak{R}_2$ the operator $(U_2^*|\mathfrak{M}_2)$ is an isometry.

Now apply Lemma 3.1 with $G_1 = U_1|\mathfrak{M}_1$, $G_2 = U_2^*|\mathfrak{M}_2$. Since $G_2^* = P(\mathfrak{M}_2) U_2 | \mathfrak{M}_2$, its minimal isometric dilation is U_2 on the smallest U_2 invariant subspace of $\mathfrak{K}_2^+$ containing $\mathfrak{M}_2$: this is $\mathfrak{R}_2$. Thus $W_2 = U_2|\mathfrak{R}_2$, $\mathfrak{N}_2 = \mathfrak{R}_2$.

Since $G_1^* = P(\mathfrak{M}_1) U_1^* | \mathfrak{M}_1$ we have $W_1 = U_1^*$ on the smallest U_1^* invariant subspace of $\mathfrak{K}_1^+$ containing $P(\mathfrak{R}_1)\mathfrak{H}_1^\perp$: thus $\mathfrak{N}_1 \subset \mathfrak{R}_1$ but the inclusion may be a strict one. By Lemma 3.1 there exists a D: $\mathfrak{R}_2 \to \mathfrak{R}_1$ such that $D = U_1 D U_2^* | \mathfrak{R}_2$, $\|D\| = \|C\|$ and

$$C = P_{\mathfrak{M}_1}^{\mathfrak{R}_1} D | \mathfrak{M}_2.$$

Finally, set $Y = DP(\mathfrak{R}_2)$. Then

$$Y = DP(\mathfrak{R}_2) = U_1 D U_2^* P(\mathfrak{R}_2) = U_1 Y U_2^*,$$

$$\|Y\| = \|D\| = \|C\| = \|X\|_{\mathfrak{R}}$$

and

$$X = A_1^* C A_2 = A_1^* C P(\mathfrak{R}_2) | \mathfrak{H}_2 = A_1^* P_{\mathfrak{M}_1}^{\mathfrak{R}_1} D P(\mathfrak{R}_2) | \mathfrak{H}_2 = A_1^* P_{\mathfrak{M}_1}^{\mathfrak{R}_1} Y | \mathfrak{H}_2.$$

To complete the proof it suffices to show that $A_1^* P_{\mathfrak{M}_1}^{\mathfrak{R}_1} = P(\mathfrak{H}_1^\perp)|\mathfrak{R}_1$. Indeed, for $r_1 \in \mathfrak{R}_1$, $h_1^\perp \in \mathfrak{H}_1^\perp$,

$$(A_1^* P_{\mathfrak{M}_1}^{\mathfrak{R}_1} r_1, h_1^\perp) = (P_{\mathfrak{M}_1}^{\mathfrak{R}_1} r_1, P(\mathfrak{R}_1) h_1^\perp) = (r_1, P(\mathfrak{R}_1) h_1^\perp) =$$
$$= (r_1, h_1^\perp) = (P(\mathfrak{H}_1^\perp) r_1, h_1^\perp).$$

Notation: Suppose $Y \in \mathscr{S}'(T_1, T_2)$. We shall denote by $T(Y)$ and $H(Y)$ the corresponding Toeplitz and Hankel operators, i.e.

$$T(Y) = P(\mathfrak{H}_1) Y | \mathfrak{H}_2$$

and

$$H(Y) = P(\mathfrak{H}_1^\perp) Y | \mathfrak{H}_2.$$

The well-known identity for products of Toeplitz operators extends to the abstract case without any change.

3.3. Proposition. *Let T_1, T_2, T_3 be contractions acting on spaces $\mathfrak{H}_1, \mathfrak{H}_2, \mathfrak{H}_3$ respectively. If $Y \in \mathcal{S}'(T_1, T_2)$, $Z \in \mathcal{S}'(T_3, T_1)$ then $ZY \in \mathcal{S}'(T_3, T_2)$ and*

$$T(ZY) - T(Z)T(Y) = H(Z^*)^* H(Y).$$

Proof. Consider $Y \in \mathcal{B}(K_2^+, K_1^+)$, $Z \in \mathcal{B}(K_1^+, K_3^+)$ satisfying $Y = U_1 Y U_2^*$, $Z = U_3 Z U_1^*$. Then

$$ZY = U_3 Z U_1^* U_1 Y U_2^* = U_3 Z Y U_2^*$$

and

$$P(\mathfrak{H}_3) ZY | \mathfrak{H}_2 = P(\mathfrak{H}_3) Z P(\mathfrak{H}_1) Y | \mathfrak{H}_2 + P(\mathfrak{H}_3) Z P(\mathfrak{H}_1^\perp) Y | \mathfrak{H}_2 =$$

$$= T(Z)T(Y) + H(Z^*)^* H(Y).$$

3.4. Definition. The operators $Y \in \mathcal{S}'(T_1, T_2)$ for which the corresponding Hankel operator $H(Y)$ is zero will be called *analytic symbols.* Thus Y is analytic if and only if Y maps $\mathfrak{H}_2$ into $\mathfrak{H}_1$. The set of all analytic symbols with respect to T_1, T_2 will be denoted by $\mathcal{A}(T_1, T_2)$.

Obviously

$$\mathcal{H}(T_1, T_2) = \mathcal{T}(T_1, T_2) / \mathcal{A}(T_1, T_2)$$

in the sense of isomorphism of linear spaces.

The classical theorem of Z. Nehari may be formulated as follows. We denote by $\{e_j\}$ the natural basis of L^2 and consider a linear operator A defined on the algebraic linear span of the $\{e_j\}$ with nonnegative indices taking its values in H^2_-. Furthermore, we assume the existence of a sequence of complex numbers $a_0, a_1, \ldots$ such that

$$(Ae_k, e_j) = a_{k+j}$$

for $k \geqq 0$ and $j < 0$. Then the Nehari theorem asserts that the operator A is the Hankel operator corresponding to some $\varphi \in L^\infty$ if and only if A is bounded.

We intend to show that the Nehari theorem has an analogon in the general situation described in the preceding sections. In the abstract theory, however, the boundedness condition has to be replaced by a stronger one — this boundedness condition reduces to ordinary boundedness in the classical case but is different from it in general. It is only in the present generality that the role played by the spaces $\mathfrak{R}$ as well as their meaning for the theory manifests itself; since $\mathfrak{R}^+ = \mathfrak{R}$ in the scalar case, it is not so easy to see the essential features of the classical results which make the theory work.

Using the notion of $\mathfrak{R}$-boundedness it is possible to formulate the following extension of the Nehari theorem.

3.5. Theorem. *Suppose $\mathfrak{M} \subset \mathfrak{H}_2$ is such that the linear span $\mathfrak{H}_0$ of all elements of the form $T_2^{*k} m$, $k \geqq 0$, $m \in \mathfrak{M}$ is dense in $\mathfrak{H}_2$. Let X: $\mathfrak{H}_0 \to \mathfrak{H}_1^\perp$ be a linear trans-*

formation which satisfies

$$V_1^* X h = X T_2^* h$$

for all $h \in \mathfrak{H}_0$.

Then the following assertions are equivalent:

1° X *is* $\mathfrak{R}$*-bounded;*

2° X *is a Hankel operator.*

Moreover, if X *satisfies* 1° *or* 2° *and* $X = H(Y)$ *with a* $Y \in \mathscr{S}'(T_1, T_2)$ *then*

$$\|H(Y)\|_{\mathfrak{R}} = \operatorname{dist}(Y, \mathscr{A}(T_1, T_2))$$

and the infimum is attained.

Proof. If 1° is satisfied then X can be regarded as an operator acting on the whole space $\mathfrak{H}_2$. Thus X is a Hankel operator and according to Theorem 3.2 there exists a symbol Y such that $X = H(Y)$ and $\|X\|_{\mathfrak{R}} = \|Y\|$. To complete the proof it is sufficient to observe that $H(Y+A) = H(Y)$ for all $A \in \mathscr{A}(T_1, T_2)$.

4. Symbols

One of the interesting questions to be asked in the context of the abstract theory is a more detailed description of the set of all symbols. We can only give partial results in this direction: we do give, however, a complete characterization of those pairs T_1, T_2, for which nonzero Toeplitz operators exist. This question is equivalent to that of the existence of non zero symbols and will be given in terms of the spaces $\mathfrak{R}_1$ and $\mathfrak{R}_2$, the unitary parts in the Wold decomposition of the minimal isometric dilations of T_1 and T_2. The answer is particularly interesting in the case $T_1 = T_2 = T$. The nonzero Toeplitz operators exist if and only if $\mathfrak{R} \neq \{0\}$. The situation is considerably more complicated in the case of analytic symbols. More delicate considerations are necessary this time; we show that it is possible to reformulate conditions for the existence of nontrivial analytic symbols in a form which may not be much easier to verify but which provides, in principle, a complete description of the set of all analytic symbols.

Consider now the particular case where $T_1 = T_2$; it is interesting to characterize those contractions T for which the corresponding set of Toeplitz operators consists of the zero operator only. In other words, to characterize those contractions $T \in \mathscr{B}(\mathfrak{H})$ for which $X \in \mathscr{B}(\mathfrak{H})$ and $X = TXT^*$ implies $X = 0$.

4.1. Proposition. *Let T be a contraction on a Hilbert space $\mathfrak{H}$. Then these are equivalent:*

1° *the only operator* X *satisfying* $X = TXT^*$ *is the zero operator;*

2° $\lim T^{*n} h = 0$ *for each* $h \in \mathfrak{H}$*;*

3° $P(\mathfrak{R})\mathfrak{H}=0$;
4° $P(\mathfrak{H})\mathfrak{R}=0$;
5° $P(\mathfrak{H})P(\mathfrak{R})P(\mathfrak{H})=0$;
6° $\mathfrak{R}=0$.

Proof. Assume 1°. According to 2.1 the projection $P(\mathfrak{R})$ is a symbol so that $X=P(\mathfrak{H})P(\mathfrak{R})|\mathfrak{H}$ is a Toeplitz operator. Since $X=0$ we have also $P(\mathfrak{H})P(\mathfrak{R})P(\mathfrak{H})=$ $=0$. Since $P(\mathfrak{H})P(\mathfrak{R})P(\mathfrak{H})=P(\mathfrak{H})P(\mathfrak{R})(P(\mathfrak{H})P(\mathfrak{R}))^*$ the condition 5° implies 4°. If 4° is satisfied we have $P(\mathfrak{R})P(\mathfrak{H})=0$ as well. Now assume 3°. According to Lemma 1.1 we have $\mathfrak{R}=(\bigcup_{n\geq 0} U^n P(\mathfrak{R})\mathfrak{H})^-$ so that $\mathfrak{R}=0$. The implication $6°\Rightarrow 2°$ follows from (10) and the implication $2°\Rightarrow 1°$ is obvious.

Let us remark that condition 5° appears implicitly in the paper of R. G. Douglas [1]. The ideas used in the proof of Theorem 3 in [1] may be used to describe existence conditions even in the case of operators Toeplitz with respect to possibly different T_1 and T_2. To this end it will be convenient to recall a definition.

Consider two unitary operators $U_1\in\mathscr{B}(\mathfrak{H}_1)$ and $U_2\in\mathscr{B}(\mathfrak{H}_2)$ with spectral measures E_1 and E_2 respectively. Following R. G. Douglas we shall say that the operators U_1 and U_2 are relatively singular if, for each $h_1\in\mathfrak{H}_1$ and $h_2\in\mathfrak{H}_2$, the measures $(E_1(\cdot)h_1, h_1)$ and $(E_2(\cdot)h_2, h_2)$ are mutually singular.

According to R. G. Douglas [1] the set of operators intertwining U_1 and U_2 is trivial if and only if U_1 and U_2 are relatively singular.

Using this notion it is possible to formulate conditions for the existence of Toeplitz operators.

4.2. Proposition. *The following assertions are equivalent:*
1° *the only operator $X\in\mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1)$ satisfying $X=T_1XT_2^*$ is the zero operator;*
2° *either one of the subspaces $\mathfrak{R}_1, \mathfrak{R}_2$ is trivial or the unitary operators R_1 and R_2 are relatively singular.*

Proof. In view of what has been said above it suffices to observe that, according to 2.11 and Remark 2.2 condition 1° is satisfied if and only if the only operator intertwining R_1 and R_2 is the zero operator.

In the classical theory analytic Toeplitz operators may be characterized by the relation $XS=SX$. The corresponding relation $T_1^*X=XT_2^*$ does not guarantee, in general, that X is (T_1, T_2) Toeplitz; we list below some supplementary condition which, together with the above relation, make X Toeplitz in which case the corresponding symbol is analytic.

4.3. Proposition. *Suppose $X\in\mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1)$ satisfies*

$$XT_2^* = T_1^*X. \tag{16}$$

Then the operator $P(\mathfrak{H}_1)P(\mathfrak{R}_1)X$ *belongs to* $\mathscr{T}(T_1, T_2)$ *and the following four conditions are equivalent:*

1° $X\in\mathscr{T}(T_1, T_2)$,

2° $X=P(\mathfrak{H}_1)P(\mathfrak{R}_1)X$,

3° $X=P(\mathfrak{H}_1\cap\mathfrak{R}_1)X$,

4° Ran $X\subset\mathfrak{H}_1\cap\mathfrak{R}_1$.

Moreover, if X satisfies (16) *and one of the conditions* 2°, 3°, 4° *then X is a Toeplitz operator whose symbol is analytic.*

On the other hand, if $Y\in\mathscr{S}'(T_1, T_2)$ *is analytic then the corresponding Toeplitz operator X satisfies* (16) *and the conditions* 2°, 3°, 4°.

Proof. Consider an $X\in\mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1)$ satisfying $XT_2^*=T_1^*X$. Then

$$T_1P(\mathfrak{H}_1)P(\mathfrak{R}_1)XT_2^* = P(\mathfrak{H}_1)U_1P(\mathfrak{R}_1)T_1^*X = P(\mathfrak{H}_1)U_1P(\mathfrak{R}_1)U_1^*X = \\ = P(\mathfrak{H}_1)P(\mathfrak{R}_1)X$$

so that the operator $P(\mathfrak{H}_1)P(\mathfrak{R}_1)X$ is Toeplitz.

Now, assume (16) and 1°. Then, for $h_2\in\mathfrak{H}_2$ and each natural number n,

$$Xh_2 = T_1^nXT_2^{*n}h_2 = T_1^nT_1^{*n}Xh_2 = P(\mathfrak{H}_1)U_1^nT_1^{*n}Xh_2 \xrightarrow[n\to\infty]{} P(\mathfrak{H}_1)P(\mathfrak{R}_1)Xh_2.$$

This proves the implication 1°⇒2°.

If 2° is satisfied then

$$X = (P(\mathfrak{H}_1)P(\mathfrak{R}_1))^nP(\mathfrak{H}_1)X \xrightarrow[n\to\infty]{} P(\mathfrak{H}_1\cap\mathfrak{R}_1)X$$

so that 3° is satisfied. The equivalence of 3° and 4° is obvious as well as the implications 4°⇒2°⇒1°.

Again, assume 4° and (16). Let Y be a symbol corresponding to X. Then, according to 2.1

$$Yh_2 = \lim U_1^nXT_2^{*n}h_2 = \lim U_1^nT_1^{*n}Xh_2 = P(\mathfrak{R}_1)Xh_2 = Xh_2\in\mathfrak{H}_1$$

for all $h_2\in\mathfrak{H}_2$, so that Y is an analytic symbol.

It remains to show that the Toeplitz operator X corresponding to an analytic symbol Y satisfies (16). Since $X=Y|\mathfrak{H}_2$ we have Ran $X=$Ran $Y|\mathfrak{H}_2\subset\mathfrak{H}_1\cap\mathfrak{R}_1$ and

$$XT_2^*h_2 = XU_2^*h_2 = YU_2^*h_2 = U_1^*Yh_2 = T_1^*Xh_2$$

for $h_2\in\mathfrak{H}_2$. The proof is complete.

The following example shows that the condition (16) alone does not imply 2°.

4.4. Example. Let us take $T_i=0$ on a Hilbert space $\mathfrak{H}_i$ $(i=1, 2)$. Then any $X\in\mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1)$ satisfies (16). Since both $\mathfrak{R}_1$ and $\mathfrak{R}_2$ are trivial, the only Toeplitz operator with respect to T_1, T_2 is the zero operator.

Now let us turn to existence conditions for analytic symbols. To this end we introduce some notation. The space $\mathfrak{H}_1 \cap \mathfrak{R}_1$ is invariant with respect to U_1^* and the restriction of U_1^* to it is an isometry. Let us denote by $\mathfrak{M}_1$ and $\mathfrak{N}_1$ the unitary part and the wandering subspace respectively in the Wold decomposition of $U_1^*|\mathfrak{H}_1 \cap \mathfrak{R}_1$. Similarly, U_2^* maps the subspace $P(\mathfrak{R}_2)\mathfrak{H}_2^-$ into itself and the restriction of U_2^* to it is an isometry; we denote by $\mathfrak{M}_2$ and $\mathfrak{N}_2$ the analogous subspaces for the Wold decomposition of $U_2^*|P(\mathfrak{R}_2)\mathfrak{H}_2^-$. Using this notation, we intend to prove the following

4.5. Theorem. *Nontrivial analytic symbols with respect to T_1 and T_2 exist if and only if the following three conditions are satisfied:*

1° *$\mathfrak{M}_1$ and $\mathfrak{M}_2$ are both nontrivial and the unitary operators $U_1^*|\mathfrak{M}_1$ and $U_2^*|\mathfrak{M}_2$ are not relatively singular;*

2° *$\mathfrak{N}_1$ and $\mathfrak{N}_2$ are both nontrivial;*

3° *$\mathfrak{M}_1$ and $\mathfrak{N}_2$ are both nontrivial and the spectral measure E of $U_1^*|\mathfrak{M}_1$ is not concentrated on a set of Lebesgue measure zero.*

Proof. In view of the one-to-one correspondence between the set of all symbols and the set of all Toeplitz operators, the set $\mathscr{A}(T_1, T_2)$ will be nontrivial if and only if the corresponding set $\mathscr{T}^a(T_1, T_2)$ of Toeplitz operators is nontrivial. According to 4.3 this set consists of all $X \in \mathscr{B}(\mathfrak{H}_2, \mathfrak{H}_1 \cap \mathfrak{R}_1)$ satisfying $XT_2^* = T_1^* X$. We shall establish a one-to-one linear correspondence between elements of the set $\mathscr{T}^a(T_1, T_2)$ and certain triangular matrices. To simplify the notation we shall write $\mathfrak{L}_1 = (\mathfrak{H}_1 \cap \mathfrak{R}_1) \ominus \mathfrak{M}_1$, $\mathfrak{L}_2 = (P(\mathfrak{R}_2)\mathfrak{N}_2)^- \ominus \mathfrak{M}_2$. To each $X \in \mathscr{T}^a(T_1, T_2)$ we assign a matrix

$$\vartheta X = \begin{bmatrix} Y_{11} & Y_{12} \\ 0 & Y_{22} \end{bmatrix}$$

defined by the following relations

$$Y_{11} = P(\mathfrak{M}_1)Y|\mathfrak{M}_2, \quad Y_{12} = P(\mathfrak{M}_1)Y|\mathfrak{L}_2, \quad Y_{22} = P(\mathfrak{L}_1)Y|\mathfrak{L}_2,$$

where $Y \in \mathscr{S}'(T_1, T_2)$ is the symbol corresponding to X.

Now denote by $\mathscr{M}$ the set of all matrices of the form

$$\begin{bmatrix} M_{11} & M_{12} \\ 0 & M_{22} \end{bmatrix}$$

with $M_{11} \in \mathscr{B}(\mathfrak{M}_2, \mathfrak{M}_1)$, $M_{12} \in \mathscr{B}(\mathfrak{L}_2, \mathfrak{M}_1)$, $M_{22} \in \mathscr{B}(\mathfrak{L}_2, \mathfrak{L}_1)$ such that the following relations are satisfied

$$(U_1^*|\mathfrak{M}_1)M_{11} = M_{11}(U_2^*|\mathfrak{M}_2), \tag{17}$$

$$(U_1^*|\mathfrak{M}_1)M_{12} = M_{12}(U_2^*|\mathfrak{L}_2), \tag{18}$$

$$(U_1^*|\mathfrak{L}_1)M_{22} = M_{22}(U_2^*|\mathfrak{L}_2). \tag{19}$$

We intend to show that ϑ is an injective mapping of the set $\mathcal{T}^a(T_1, T_2)$ onto $\mathcal{M}$.

Let us consider an $X\in\mathcal{T}^a(T_1, T_2)$ with the corresponding symbol Y. Since $Y=YP(\mathfrak{R}_2)$ we have

$$YP(\mathfrak{R}_2)\mathfrak{H}_2 = Y\mathfrak{H}_2 = X\mathfrak{H}_2 \subset \mathfrak{H}_1\cap\mathfrak{R}_1.$$

Furthermore, the relation $U_1^*Y=YU_2^*$ implies that $\vartheta X\in\mathcal{M}$.

Consider now the operators $Z=P(\mathfrak{L}_1)Y|\mathfrak{M}_2$ and $S_1=U_1^*|\mathfrak{L}_1$. Using $U_1^*Y= =YU_2^*$ again we have also $S_1Z=Z(U_2^*|\mathfrak{M}_2)$ so that $Z^*S_1^{*n}=(U_2^*|\mathfrak{M}_2)^{*n}Z^*$ for every natural number n. Given $m\in\mathfrak{M}_2$ we have

$$\|Z^*m\| = \|(U_2^*|\mathfrak{M}_2)^{*n}Z^*m\| =$$
$$= \|Z^*S_1^{*n}m\| \leqq \|Z^*\|\,\|S_1^{*n}m\| \to 0,$$

so that $Z^*=0$ and $Z=0$ as well. Thus, for each $h_2\in\mathfrak{H}_2$, X can be decomposed as follows

$$Xh_2 = YP(\mathfrak{R}_2)h_2 = Y_{11}P(\mathfrak{M}_2)P(\mathfrak{R}_2)h_2+$$
$$+Y_{12}P(\mathfrak{L}_2)P(\mathfrak{R}_2)h_2+Y_{22}P(\mathfrak{L}_2)P(\mathfrak{R}_2)h_2.$$

Hence $\vartheta X=0$ implies $X=0$ and ϑ is injective.

On the other hand, each $M\in\mathcal{M}$ defines an operator from $P(\mathfrak{R}_2)\mathfrak{H}_2^-$ into $\mathfrak{H}_1\cap\mathfrak{R}_1$. The relations (17), (18), (19) imply that

$$\begin{bmatrix} U_1^*|\mathfrak{M}_1 & 0 \\ 0 & U_1^*|\mathfrak{L}_1\end{bmatrix}\begin{bmatrix} M_{11} & M_{12} \\ 0 & M_{22}\end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} \\ 0 & M_{22}\end{bmatrix}\begin{bmatrix} U_2^*|\mathfrak{M}_2 & 0 \\ 0 & U_2^*|\mathfrak{L}_2\end{bmatrix},$$

so that $U_1^*M=MU_2^*|P(\mathfrak{R}_2)\mathfrak{H}_2^-$. If we set $Xh_2=MP(\mathfrak{R}_2)h_2$ for $h_2\in\mathfrak{H}_2$ then $X\in\mathcal{T}^a(T_1, T_2)$ and $\vartheta X=M$.

In view of the isomorphism between $\mathcal{T}^a(T_1, T_2)$ and $\mathcal{M}$ our problem is equivalent to that of describing conditions for $\mathcal{M}$ to be nontrivial. An element $M\in\mathcal{M}$ is nonzero if and only if at least one of its entries is nonzero.

If $M_{11}\neq 0$ then clearly both $\mathfrak{M}_1$ and $\mathfrak{M}_2$ must be nontrivial subspaces; at the same time M_{11} is a nonzero operator intertwining the unitary operators $U_1^*|\mathfrak{M}_1$ and $U_2^*|\mathfrak{M}_2$ and this yields condition 1°. On the other hand, if condition 1° is satisfied, there exists a nonzero operator $Z\in\mathcal{B}(\mathfrak{M}_2, \mathfrak{M}_1)$ for which

$$(U_1^*|\mathfrak{M}_1)Z = Z(U_2^*|\mathfrak{M}_2);$$

then

$$\begin{bmatrix} Z & 0 \\ 0 & 0\end{bmatrix}\in\mathcal{M}.$$

If $M_{22}\neq 0$ both its domain and range must be nontrivial, hence condition 2°. Conversely, if condition 2° holds, take a vector $g\in\mathfrak{R}_2$ and a vector $h\in\mathfrak{R}_1$. It is easy to see that $((U_2^pg,\ U_2^qg)=0$ for all integers $p\neq q$. The sequence U_1^kh, $k\in\mathbf{Z}$

possesses the same property so that it is possible to define an operator $Y \in \mathcal{B}(\mathfrak{K}_2^+, \mathfrak{K}_1^+)$ by the formula

$$Yx = \sum_{-\infty}^{\infty} (x, U_2^k g) U_1^k h;$$

clearly $Y = U_1 Y U_2^*$. If $k \geqq 1$ we have

$$(U_2^k g, \mathfrak{H}_2) = (g, U_2^{*k} \mathfrak{H}_2) = (g, P(\mathfrak{R}_2) U_2^{*k} \mathfrak{H}_2) =$$
$$= (g, U_2^{*k} P(\mathfrak{R}_2) \mathfrak{H}_2) = 0,$$

so that $Yx = \sum_{k \leqq 0} (x, U_2^k g) U_1^k h \subset \mathfrak{H}_1$ for $x \in \mathfrak{H}_2$. Thus $Y \in \mathcal{A}(T_1, T_2)$.

Consider the case $M_{12} \neq 0$; it follows that $\mathfrak{L}_2 \neq \{0\}$. The operator $S_2 = U_2^* | \mathfrak{L}_2$ is a unilateral shift so that the minimal isometric dilation of S_2^* is a unitary operator W with the following properties (see [2], Ch. 2, Theorem 6.4):

(i) the spectral measure $E_W(\cdot)$ of W is equivalent to Lebesgue measure,

(ii) for each nonzero $z \in \mathfrak{L}_2$, the measure $(E_W(\cdot)z, z)$ is equivalent to Lebesgue measure.

Since M_{12} satisfies (18) we have $M_{12}^* = S_2^* M_{12}^* (U_1^* | \mathfrak{M}_1)$ so that $M_{12}^* \in \mathcal{T}(S_2^*, (U_1^* | \mathfrak{M}_1)^*)$. The corresponding symbol G satisfies $G = WG(U_1^* | \mathfrak{M}_1)$ so that $G(U_1^* | \mathfrak{M}_1) = W^* G$ and this implies condition 3°.

On the other hand, if condition 3° holds there exists a nonzero vector $x \in \mathfrak{M}_1$ and a set M of positive Lebesgue measure for which $(E(M)x, x) > 0$. Furthermore, if z is an arbitary vector in $\mathfrak{N}_2$ the measure $(E_W(\cdot)z, z)$ is equivalent to Lebesgue measure. It follows that there exists a nonzero operator K defined on $\mathfrak{M}_1$ which intertwines W and $(U_1^* | \mathfrak{M}_1)$, $K(U_1^* | \mathfrak{M}_1) = WK$. Hence $K = WK(U_1^* | \mathfrak{M}_1)^*$ so that $K \in \mathcal{S}'(S_2^*, U_1^* | \mathfrak{M}_1)$ and the corresponding Toeplitz operator $T(K)$ satisfies (18). Accordingly,

$$\begin{bmatrix} 0 & T(K) \\ 0 & 0 \end{bmatrix} \in \mathcal{M}.$$

The proof is complete.

4.6. Corollary. *If T_1 is completely nonunitary then $\mathcal{A}(T_1, T_2)$ is nontrivial if and only if both $\mathfrak{N}_1$ and $\mathfrak{N}_2$ are nontrivial.*

Proof. It follows from Lemma 1.2 that $\mathfrak{M}_1 = \mathfrak{H}_u(T_1)$. If T_1 is completely nonunitary then $\mathfrak{M}_1 = \mathfrak{H}_u(T_1) = \{0\}$. It follows from the preceding theorem that $\mathcal{A}(T_1, T_2)$ is nontrivial if and only if 2° is satisfied.

Of course it is possible to reformulate the existence conditions for analytic symbols in a manner analogous to Proposition 4.2 The problem does not become any easier in this reformulation; nevertheless, it provides some more insight into the structure of these symbols.

4.7. Proposition. *Let* $T_1 \in \mathscr{B}(\mathfrak{H}_1)$, $T_2 \in \mathscr{B}(\mathfrak{H}_2)$ *be two contractions. Let us denote by* $\mathfrak{G}_1$ *the smallest* U_1 *reducing subspace containing* $\mathfrak{H}_1 \cap \mathfrak{R}_1$. *Then these are equivalent:*

1° $\mathscr{A}(T_1, T_2) = \{0\}$,

2° *the unitary operators* R_2 *and* $U_1|\mathfrak{G}_1$ *are relatively singular.*

Proof. If Y is an analytic symbol then $Y = U_1 Y U_2^*$ and Y maps $\mathfrak{R}_2 = \bigvee_{n \geq 0} P(\mathfrak{R}_2) U_2^n \mathfrak{H}_2$ into $\bigvee_{n \geq 0} U_1^n(\mathfrak{H}_1 \cap \mathfrak{R}_1)$ which is nothing more than the smallest reducing subspace $\mathfrak{G}_1$ for U_1 containing $\mathfrak{H}_1 \cap \mathfrak{R}_1$. Thus

$$Y|\mathfrak{R}_2 = (U_1|\mathfrak{G}_1)(Y|\mathfrak{R}_2)(U_2|\mathfrak{R}_2)^*.$$

5. Rational symbols and the theorem of Kronecker

It might seem that there is little hope that a reasonable extension to this generality of the algebraic notion of rational function would be possible. We intend to show in this section that such an extension does exist and that it may be used to obtain a generalization of the theorem of Kronecker.

We shall use an abbreviation: if p is a polynomial of degree n, we shall write p_1 for the polynomial defined by the relation $p_1(x) = x^n p(1/x)$.

5.1. Proposition. *Suppose* $Y \in \mathscr{A}(T_1, T_2)$ *and let* q *be a polynomial of degree* n *with roots of modulus less then* 1, $q(x) = (x - \alpha_1)\ldots(x - \alpha_n)$.

Then $q(R_1^*)^{-1} Y$ *is a symbol and the corresponding Hankel operator may be expressed as follows*

$$H\big(q(R_1^*)^{-1}Y\big) =$$
$$= \sum_{k=1}^{n} U_1^{k-1}(1 - \alpha_1 U_1)^{-1}\ldots(1 - \alpha_k U_1)^{-1}(U_1 - T_1)T_1^{n-k}(1 - \alpha_k T_1)^{-1}\ldots(1 - \alpha_n T_1)^{-1} Y|\mathfrak{H}_2$$

or in an equivalent form

$$\sum_{k=1}^{n} (U_1^* - \alpha_{k+1})\ldots(U_1^* - \alpha_n) q_1(U_1)^{-1} U_1^{n-1}(U_1 - T_1)\cdot$$
$$\cdot T_1^{n-1} q_1(T_1)^{-1} Y (T_2^* - \alpha_{k-1})\ldots(T_2^* - \alpha_1).$$

Proof. Since $Y = P(\mathfrak{R}_1)Y$ we have

$$U_1 q(R_1^*)^{-1} Y U_2^* = R_1 q(R_1^*)^{-1} Y U_2^* = q(R_1^*)^{-1} R_1 Y U_2^* = q(R_1^*)^{-1} Y$$

so that $q(R_1^*)^{-1} Y$ is a symbol.

Since R_1 is unitary we have

$$q(R_1^*)^{-1} = \prod_1^n (R_1^{-1}-\alpha_j)^{-1} = R_1^n(1-\alpha_1 R_1)^{-1}\dots(1-\alpha_n R_1)^{-1} = $$
$$= R_1^n q_1(R_1)^{-1}$$

and

$$P(\mathfrak{H}_1^\perp)q(R_1^*)^{-1}Y|\mathfrak{H}_2 = P(\mathfrak{H}_1^\perp)R_1^n q_1(R_1)^{-1}Y|\mathfrak{H}_2 = P(\mathfrak{H}_1^\perp)U_1^n q_1(U_1)^{-1}Y|\mathfrak{H}_2 =$$
$$= U_1^n q_1(U_1)^{-1}Y|\mathfrak{H}_2 - P(\mathfrak{H}_1)U_1^n q_1(U_1)^{-1}Y|\mathfrak{H}_2 = U_1^n q_1(U_1)^{-1}Y|\mathfrak{H}_2 -$$
$$- T_1^n q_1(T_1)^{-1}Y|\mathfrak{H}_2.$$

Now, *it* suffices to apply Proposition 1.3.

5.2. Theorem. *Let H be a Hankel operator, $H \in \mathscr{H}(T_1, T_2)$. Then the following assertions are equivalent:*

1° *the range of H is finite dimensional;*

2° *there exists a polynomial q with roots of modulus less than* 1 *and an analytic symbol $Y \in \mathscr{A}(T_1, T_2)$ such that*

2.1° $$H = H(q(R_1^*)^{-1}Y)$$

and one of the two following equivalent conditions is satisfied

2.2° $d_j = \dim (U_1 - T_1)T_1^j(1-\alpha_1 T_1)^{-1}\dots(1-\alpha_{j+1}T_1)^{-1}Y\mathfrak{H}_2 < \infty$ *for* $j = 0, \dots, n-1$

where $\alpha_1, \dots, \alpha_n$ are the roots of q, $\deg q = n$,

2.3° $\dim (U_1 - T_1)T_1^{n-1}(1-\alpha_1 T_1)^{-1}\dots(1-\alpha_n T_1)^{-1}Y\mathfrak{H}_2 < \infty.$

If these conditions are satisfied then

$$\dim \operatorname{Ran} H \leqq d_0 + d_1 + \dots + d_{n-1}.$$

Proof. The range of H is invariant with respect to V_1^*. If it is finite-dimensional there exists a polynomial q such that $q(V_1^*|\operatorname{Ran} H) = 0$ so that $q(V_1^*)H = 0$.

Since V_1 is a unilateral shift both V_1 and V_1^* have no eigenvalues on the unit circle. Hence we can assume that all the roots of q lie inside the unit disc. If Z is any symbol for H, i.e. $H = P(\mathfrak{H}_1^\perp)Z|\mathfrak{H}_2$, $Z \in \mathscr{S}'(T_1, T_2)$ we have

$$0 = q(V_1^*)H = Hq(T_2^*) = P(\mathfrak{H}_1^\perp)Zq(U_2^*)|\mathfrak{H}_2 = P(\mathfrak{H}_1^\perp)q(U_1^*)Z|\mathfrak{H}_2.$$

Hence $q(U^*)Z\mathfrak{H}_2 \subset \mathfrak{H}_1$. Since the range of Z is contained in $\mathfrak{R}_1$ it follows that $Y = q(U_1^*)Z$ is an analytic symbol and $Y = q(R_1^*)Z$ whence $Z = q(R_1^*)^{-1}Y$ which proves 2.1°.

The range of the operator $P(\mathfrak{L}_1)q_1(U_1)H$ is also finite dimensional and it follows from Proposition 5.1 that it is equal to the space $(U_1 - T_1)T_1^{n-1}q_1(T_1)^{-1}Y\mathfrak{H}_2$. Thus condition 2.3° is satisfied.

Let us show now, that, for any polynomial q with roots inside the unit disc and any analytic symbol Y, condition 2.3° implies 2.2°. Since Y is an analytic symbol we have $Y\mathfrak{H}_2 \subset \mathfrak{H}_1 \cap \mathfrak{N}_1$. On the other hand T_1^* is an isometry on $\mathfrak{H}_1 \cap \mathfrak{N}_1$ and $T_1 T_1^* h = h$ for all $h \in \mathfrak{H}_1 \cap \mathfrak{N}_1$. Using these facts we can write, for $|\alpha| < 1$,

$$Y\mathfrak{H}_2 = (1-\alpha T_1)^{-1}(1-\alpha T_1)Y\mathfrak{H}_2 =$$
$$= (1-\alpha T_1)^{-1}(T_1 T_1^* - \alpha T_1)Y\mathfrak{H}_2 = (1-\alpha T_1)^{-1} T_1 (T_1^* - \alpha)Y\mathfrak{H}_2 =$$
$$= T_1(1-\alpha T_1)^{-1} Y(T_2^* - \alpha)\mathfrak{H}_2 \subseteqq T_1(1-\alpha T_1)^{-1} Y\mathfrak{H}_2.$$

It is easy to deduce from the just established relation that 2.3° implies 2.2°.

Assume that 2.2° is satisfied for a polynomial q with roots inside the unit disc and some analytic symbol Y. Then, according to Proposition 5.1, the Hankel operator $H\big(q(R_1^*)^{-1}Y\big)$ is finite dimensional and $\dim \operatorname{Ran} H\big(q(R_1^*)^{-1}Y\big) \leqq d_0 + d_1 + \dots + d_{n-1}$.

The proof is complete.

5.3. Corollary. *Suppose* $\dim \mathfrak{L}_1 < \infty$. *Given a symbol of the form*

$$q(R_1^*)^{-1}Y,$$

where q is a polynomial of degree n (with roots inside the unit disc) and $Y \in \mathscr{A}(T_1, T_2)$, *condition* 2.2° *is automatically satisfied and*

$$\dim \operatorname{Ran} H\big(q(R_1^*)^{-1}Y\big) \leqq n \dim \mathfrak{L}_1.$$

The corollary applies in particular in the case where $\dim \mathfrak{L}_1 = 1$. Furthermore, for classical Hankel operators it is more natural to view the symbol as an equivalence class in L^∞/H^∞ rather than as an individual function; in conformity with this point of view it seems natural to define a rational symbol as a class which contains a rational function, or equivalently, a class which contains a quotient h/q, $h \in H^\infty$, q a polynomial. In view of this it is not unnatural to use the name rational symbol for operators of the form $q(R_1^*)^{-1}Y$, Y analytic.

Theorem 5.2 appears thus as an extension of the well-known theorem of Kronecker. It is natural to ask whether the assumption 2.2° in Theorem 5.2 is essential for the validity of the generalized Kronecker theorem. We limit ourselves to stating that there exist examples which show that ranges of Hankel operators with rational symbols may be both finite and infinite dimensional if $\dim \mathfrak{L}_1$ is infinite.

References

[1] R. G. DOUGLAS, On the operator equation $S^*XT=X$ and related topics, *Acta Sci. Math.*, **30** (1969), 19—32.

[2] B. SZ.-NAGY, C. FOIAŞ, *Harmonic analysis of operators on Hilbers space*, North Holland—Akadémiai Kiadó, (Amsterdam—Budapest, 1970).

[3] B. SZ.-NAGY, C. FOIAŞ, Toeplitz type operators and hyponormality, in *Dilation theory, Toeplitz operators and other topics, Operator Theory: Advances and Applications*, Vol. **11**, Birkhäuser Verlag (Basel, 1983), 371—378.

[4] B. SZ-NAGY, C. FOIAŞ, An application of dilation theory to hyponormal operators, *Acta Sci. Math.*, **37** (1975), 155—159.

[5] L. B. PAGE, Bounded and compact vectorial Hankel operators, *Trans. Amer. Math. Soc.*, **150** (1970), 529—539.

INSTITUTE OF MATHEMATICS
CZECHOSLOVAK ACADEMY OF SCIENCES
ŽITNÁ 25
115 67 PRAHA 1, CZECHOSLOVAKIA

Acta Sci. Math., 52 (1988), 141—144

Normalcy is a superfluous condition in the definition of G-finiteness*

JOSEPH M. SZŰCS

Dedicated to Professor Károly Tandori on his 60th birthday

Let M be a W^*-algebra and let G be a group of $*$-automrophisms of M. In [2] we have proved that if there exists a faithful G-invariant *normal* state φ on M, then for every $t \in M$, the w^*-closure of the convex hull of the orbit of t under G contains a unique G-invariant element t^G and the mapping $t \to t^G$ is normal. (In fact, we have proved this result under the more general assumption that the family of G-invariant normal states on M is faithful, i.e., M is G-finite [2]. If M is σ-finite, for example, if M is an operator algebra in a separable Hilbert space, then this assumption obviously implies the existence of a faithful G-invariant normal state on M.) In the present paper we shall prove that the assumption of normalcy of φ is superfluous in this theorem (cf. Theorem). Under additional hypotheses, we shall also prove that φ itself is a normal state (cf. Corollary 1). Furthermore, we shall prove some converse results (cf. Corollaries 2 and 3).

For the general theory of W^*-algebras, we refer the reader to [1] and [3].

At the end of the paper we shall make two comments on our paper [4].

Theorem. *Let M be a W^*-algebra and let G be a group of $*$-automorphisms of M. If there exists a faithful G-invariant state φ on M, then there exists a faithful G-invariant normal state ψ on M, i.e., M is G-finite.*

Proof**. Let $\varphi = \varphi_n + \varphi_S$ be the canonical decomposition of φ into normal part φ_n and singular part φ_s [3]. Consider an element $g \in G$. Then $\varphi_n(g\cdot)$ is normal due to

* This work was supported in part by organized research money granted by Texas A&M University at Galceston.

** The author's first proof of this theorem was much more complicated. This proof originated from a comment by R. R. Smith at a seminar at Texas A&M University, College Station.

Received January 8, 1985.

the continuity properties of g. On the other hand, $\varphi_s(g\cdot)$ is singular, since a positive linear form μ on M is singular if and only if every nonzero projection $p\in M$ majorizes a nonzero projection $q\in M$ such that $\mu(q)=0$ [3]. Since φ is g-invariant and the decomposition into normal and singular parts is unique, we obtain that φ_n is g-invariant (for all $g\in G$). Furthermore, φ_n is faithful. For let p be a nonzero projection in M. Since φ_s is singular, there exists a nonzero subprojection q of p in M, such that $\varphi_s(q)=0$. Then $\varphi_n(p)\geqq\varphi_n(q)=\varphi(q)-\varphi_s(q)=\varphi(q)>0$ because φ was assumed to be faithful. Summing up, we can choose $\psi=\varphi_n$.

Corollary 1. *Let M be a W^*-algebra and G a group of $*$-automorphisms of M. Suppose that for every $t\in M$, the norm-closed convex hull of the orbit Gt of t under G contains at least one G-invariant element. If φ is a G-invariant faithful state on M and the restriction of φ to the fixed-point algebra M^G is normal, then φ is normal.*

Proof. According to Theorem, M is G-finite [2]. Consequently, the G-invariant element, say t^G, in the norm-closed convex hull of Gt is unique [2]. Moreover, the mapping $t\to t^G$: $M\to M^G$ is normal [2]. Since φ is G-invariant and norm-continuous, $\varphi(t)=\varphi(t^G)$ $(t\in M)$. Therefore, the mapping $t\to\varphi(t)$: $M\to\mathbf{C}$ is the composite mapping of $t\to t^G$: $M\to M^G$ and $t\to\varphi(t)$: $M^G\to\mathbf{C}$. Since both of these mappings are normal, φ is normal.

Corollary 2. *Let M be a W^*algebra and G a group of $*$-automorphisms of M. Suppose that for every $t\in M$, the w^*-closed (norm-closed) convex hull of the orbit Gt of t under G contains exactly one G-invariant element, say t^G. If $t^G\neq 0$ for $t\geqq 0$, $t\neq 0$, then M is G-finite.*

Proof. The mapping $t\to t^G$: $M\to M^G$ is linear. In the case of the norm-closed convex hull, this can be proved as follows. The homogeneity of the mapping $t\to t^G$ is obvious. To prove its linearity, let t, $s\in M$ and let $\varepsilon>0$ be a given number. There exists a v_0 in the convex hull conv G of G, such that $\|v_0(t)-t^G\|<\varepsilon/2$. Similarly, there exists $v_1\in\operatorname{conv} G$, such that $\|v_1v_0(s)-s^G\|<\varepsilon/2$. Since every element of G has norm 1, we have $\|v_1v_0(t)-t^G\|<\varepsilon/2$. Consequently, $\|v_1v_0(t+s)-(t^G+s^G)\|\leqq\leqq\|v_1v_0(t)-t^G\|+\|v_1v_0(s)-s^G\|<\varepsilon$. Since $\varepsilon>0$ was arbitrary, this proves that $(t+s)^G=t^G+s^G$.

In the case of the w^*-closed convex hull, the linearity of $t\to t^G$ can be proved as follows. The homogeneity of $t\to t^G$ is obvious. Let us verify its additvity. Let s, $t\in M$. Then there exists a net v_n in conv G, such that $\lim_n v_n(s)=s^G$. Since the unit ball of M is w^*-compact, there exists a subnet v_k of v_n, such that $t_k=\lim_k v_k(t)$ exists. Then $\lim_k v_k(s+t)=\lim_k v_k(s)+\lim_k v_k(t)=\lim_n v_n(s)+t_k=s^G+t_k$ belongs to the w^*-closed convex hull of $G(s+t)$. By the definition of $(t_k)^G$, there is a net w_n in conv G, such that $\lim_n w_n(t_k)=(t_k)^G$. Then $\lim_n w_n(s^G+t_k)=\lim_n[w_n(s^G)+w_n(t_k)]=\lim_n[s^G+$

$+w_n(t_k)]=s^G+\lim_n w_n(t_k)=s^G+(t_k)^G$. Consequently, $(s^G+t_k)^G=s^G+(t_k)^G$. Since s^G+t_k belongs to the w^*-closed convex hull of $G(s+t)$, we have $(s+t)^G=(s^G+t_k)^G$. Therefore, $(s+t)^G=s^G+(t_k)^G$. Similarly, since t_k belongs to the w^*-closed convex hull of $G(t)$, we have $t^G=(t_k)^G$. Summing up, we have obtained that $(s+t)^G= =s^G+(t_k)^G=s^G+t^G$, which was to be proved.

So far we have proved that $t\to t^G$: $M\to M^G$ is linear. On the other hand, it is evident that $[g(t)]^G=t^G$ for every $g\in G$, $t\in M$ and $t^G=t$ for $t\in M^G$, the G-fixed-point algebra in M.

Now let φ_0 be a normal state on M^G. Let $\varphi(t)=\varphi_0(t^G)$ for $t\in M$. Then φ is a G-invariant state on M. Let p be the support of φ_0. Then $p\in M^G$ and $(ptp)^G=pt^Gp$. Consequently, φ is faithful on pMp, by the hypotheses of the corollary and by the faithfulness of φ_0 on pM^Gp. Since φ is invariant under the restriction of G to pMp, Theorem can be applied. We obtain that pMp is finite with respect to the restriction of G to pMp. This implies [2] that φ is a G-invariant normal state on M with support p. Since $\sup p=1$ if φ_0 runs over all normal states of M^G, we obtain that M is G-finite [2].

Corollary 3. *Let M be a W^*-algebra and G a group of $*$-automorphisms of M. If τ: $M\to M^G$ is a G-invariant faithful positive linear mapping which leaves M^G elementwise fixed, then M is G-finite.*

Proof. It is similar to the end of the proof of Corollary 2.

Remarks. 1. The proof of one half of Corollary 2 does not require Theorem:

Let M be a W^*-algebra and G a group of $*$-automorphisms of M. Suppose that for every $t\in M$, the norm-closed convex hull of the orbit Gt of t under G contains exactly one G-invariant element, say t^G. If $t^G\neq 0$ for $t\geqq 0$, $t\neq 0$, then M is G-finite (and $t\to t^G$ is a normal positive linear mapping of M onto M^G).

Proof. As in the proof of Corollary 2, we first prove that $t\to t^G$ is a linear mapping. This done, let φ_0 be a normal positive linear form on M^G and let p denote the support of φ_0. Then $(ptp)^G=pt^Gp$ and $t\to\varphi_0(t^G)$ is a faithful positive linear form φ on pMp, invariant under the restriction of G to pMp. Let e be a nonzero projection in pMp, such that $\varphi(e\cdot e)$ is normal [1]. Then $\varphi(\cdot e)\in M^G$. Let $v_n\in\operatorname{conv} G$ be such that $\|v_n(e)-e^G\|\to 0$ as $n\to\infty$. We have $\varphi(\cdot v_n(e))\in M^G$ by the G-invariance of φ and by the fact that $\varphi(\cdot e)M_*$. Then the norm limit of $\varphi(\cdot v_n(e))$ in M_* is $\varphi(\cdot e^G)$, since $\varphi\in M^*$. Therefore, $\varphi(\cdot e^G)\in M_*$. Consequently, $t\to\varphi(e^Gte^G)=\varphi_0((e^Gte^G)^G)=\varphi_0(e^Gt^Ge^G)$ is a normal positive linear form on M. Since $e^G\leqq p$, $e^G\in M^G$, we obtain that $t\to e^Gt^Ge^G$ is normal on M. If φ_0 runs over all normal forms on M^G, we obtain that every nonzero projection $p\in M^G$ majorizes a nonzero projection $e \in M$ (it is $e^G \in M^G$) such that $t \to te^G e$ is normal. This implies that $t \to t^G$ is normal on M and thus M is G-finite [2].

2. The assumption of Theorem that φ is faithful is essential. Indeed, let G be an abstract infinite Abelian group. Then G acts naturally on $M = l^\infty(G)$ as a group of $*$-automorphisms. A G-invariant state on M is noting else but an invariant mean on G. We know that there are ifinitely many invariant means on G, none of wich are normal (actually, they are singular).

Finally, the author would like to make two comments on his paper [4]. The first comment is that in Proposition 2 and in its corollary the assumption that M is σ-finite should be replaced by the assumption that the predual of M is separable.

The second comment is that all the results of the above mentioned paper remain valid if G is only assumed to be an amenable group (instead of an Abelian one). Indeed, if $U_n \subset G$ is a summing sequence [5], then it is easy to prove that under the hypotheses of Lemma 1, the sequence $\frac{1}{|U_n|} \sum_{g \in U_n} g(t)$ w^*-converges to t^G for every $t \in B^*$. The remaining results of the paper can be extended to amenable groups G without altering the proofs.

References

[1] J. Dixmier, *Les algèbres d'opérateurs dans l'espaces Hilbertiens (Algèbres de von Neumann)*, 2nd ed., Gauthier-Villars (Paris, 1969).

[2] I. Kovács and J. Szűcs, Ergodic theorems in von Neumann algebras, *Acta Sci. Math.*, **27** (1966) 233—246.

[3] Sh. Sakai, *C*-algebras*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 60, Springer-Verlag (Berlin—Heidelberg—New York, 1971).

[4] J. Szűcs, Some weak-star ergodic theorems, *Acta Sci. Math.*, **45** (1983), 389—394.

[5] F. P. Greenleaf, Ergodic theorems and the construction of summing sequences in amenable locally compact groups, *Comm. Pure Appl. Math.*, **26** (1973), 29—46.

TEXAS A&M UNIVERSITY AT GALVESTON
DEPARTMENT OF GENERAL ACADEMICS
MITCHELL CAMPUS
GALVESTON, TEXAS 77553, U.S.A.

Acta Sci. Math., 52 (1988), 145—155

Commutative GW^*-algebras

JÁNOS KRISTÓF

GW^*-algebras (i.e. generalized W^*-algebras) were introduced in [2]. In this paper the structure and the spectral properties of commutative GW^*-algebras will be examined in detail.

I. Preliminaries

Here we give a short summary of our former results concerning GW^*-algebras.

The vector space of the linear forms on the $*$-algebra A will be denoted by A^* and the weak $\sigma(A^*, A)$ topology relates to the canonical duality between A^* and A.

If A is a unital $*$-algebra (whose unit is denoted by $\mathbf{1}$ throughout this paper) and P is a set of positive linear forms on A then the set $\{f \in P | f(\mathbf{1}) \leqq 1\}$ will be denoted by the symbol $P(1)$. Further, assuming that $P(1)$ is non-void and bounded in the $\sigma(A^*, A)$ topology, $\|\cdot\|_P$ denotes the mapping from A into $\mathbf{R}_+$ defined by

$$\|x\|_P := \sup_{f \in P(1)} \sqrt{f(x^*x)}$$

for all $x \in A$. It is obvious that $\|\cdot\|_P$ is a seminorm on A; the dual seminorm is denoted by $\|\cdot\|'_P$.

If S is a subset of A^* then the linear subspace of A^* spanned by S and the convex hull of S is denoted by sp (S) and co (S), respectively, while the $\sigma(A^*, A)$-closed linear subspace of A^* spanned by S and the $\sigma(A^*, A)$-closed convex hull of S is denoted by $\widetilde{\text{sp}}(S)$ and $\widetilde{\text{co}}(S)$, respectively. If the elements of S are $\|\cdot\|_P$-continuous forms (where P is a set of positive linear forms on A such that $P(1)$ is non void and $\sigma(A^*, A)$-bounded) then the $\|\cdot\|'_P$-closed linear subspace of A^* spanned by S and the $\|\cdot\|'_P$-closed convex hull of S is denoted by $\overline{\text{sp}}(S)$ and $\overline{\text{co}}(S)$, respectively, provided there is no danger of confusion as for P.

Received January 9, 1985.

If f is a linear form on the $*$-algebra A then for every $x \in A$ we define the linear forms $x \cdot f$ and $f \cdot x$ on A as the mappings $y \mapsto f(xy)$ and $y \mapsto f(yx)$, respectively. If $f \in A^*$ and $x, y \in A$ then $x \cdot f \cdot y$ stands for $(x \cdot f) \cdot y$.

Definition. The pair (A, P) is called a *weak* GW^**-algebra* if A is a unital $*$-algebra and P is a separating set of positive linear forms on A satisfying:

(I) $P(1)$ is non-void and $\sigma(A^*, A)$-bounded;

(II_w) $\mathbf{R}_+ P \subset P$ and $x^* \cdot P \cdot x \subset \widetilde{\overline{\mathrm{co}}}(P)$ for all $x \in A$;

(III) $x \cdot P \subset \overline{\mathrm{sp}}(P)$ for all $x \in A$;

(IV) A is sequentially complete with respect to the uniform structure defined by the $\sigma(A, \mathrm{sp}(P))$ topology.

The pair (A, P) is called a GW^**-algebra* if it is a weak GW^*-algebra and instead of (II_w) satisfies the more restrictive condition:

(II) $\mathbf{R}_+ P \subset P$ and $x^* \cdot P \cdot x \subset \overline{\mathrm{co}}(P)$ for all $x \in A$.

Finally, the pair (A, P) is referred to as a *complete* GW^**-algebra* if it satisfies:

(IV_s) A is quasi complete with respect to the uniform structure defined by the $\sigma(A, \mathrm{sp}(P))$ topology.

The most important elementary facts concerning weak GW^*-algebras are the following. If (A, P) is a weak GW^*-algebra then:

— A is a C^*-algebra whose C^*-norm coincides with $\|\cdot\|_P$, that is why we refer to $\|\cdot\|_P$ as the C^*-norm of A (cf. [1] and [2]);

— the $\sigma(A, \mathrm{sp}(P))$ and $\sigma(A, \overline{\mathrm{sp}}(P))$ topologies coincide in every C^*-norm bounded subset of A (cf. [1] Lemma 1);

— the multiplication of A is C^*-norm boundedly left and right continuous in the $\sigma(A, \mathrm{sp}(P))$ topology (cf. [1] Lemma 2);

— the involution of A is proper and continuous in the $\sigma(A, \mathrm{sp}(P))$ topology;

— the set of projections (i.e. self-adjoint idempotent elements) of A, equipped with the natural ordering: $g \leqq h \Leftrightarrow g = hg$ and the orthocomplementation: $e^\perp := \mathbf{1} - e$, is a σ-complete orthomodular lattice admitting a separating set of σ-additive states (cf. [2] Theorem 1);

— the partial isometries are countably summable in A and, consequently, the equivalence of projections is countably additive in A (cf. [2] Proposition 2).

Here we deduce an important auxiliary result for general (not necessarily commutative) weak GW^*-algebras.

Proposition 1. *Let (A, P) be a weak GW^*-algebra. Then the order in A defined as $x \leqq y$ iff $f(y-x) \in \mathbf{R}_+$ $(f \in P)$ coincides with the algebraic order of the C^*-algebra A.*

Proof. Since the elements of P are positive linear forms on A, we have obviously $x \geqq 0$ with respect to the order defined by P, if $x \geqq 0$ in the C^*-algebra A.

Conversely, suppose that $x \geqq 0$ with respect to the order defined by P. Since the set of positive linear forms f on A satisfying $f(x) \in \mathbf{R}_+$ is $\sigma(A^*, A)$-closed, we have $f(x) \geqq 0$ for every $f \in \widetilde{\overline{co}}(P)$. Since $f(x) \in \mathbf{R}_+$ $(f \in P)$, we have $f(x^*) = \overline{f(x)} = f(x)$, hence $x = x^*$ since P separates the points of A. We know that A is a C^*-algebra thus we may write $x = x^+ - x^-$, where x^+ and x^- denotes the positive and negative part of the self-adjoint element x, respectively. Then the positive square root $\sqrt{x^-}$ of x^- exists in A and it is well known that the set $\{\sqrt{x^-}, x^+, x^-\}$ is commutative; moreover, $x^+x^- = x^-x^+ = 0$. Fixed a linear form f in P, we have $(\sqrt{x^-}) \cdot f \cdot (\sqrt{x^-}) \in \widetilde{\overline{co}}(P)$ thus

$$0 \leqq ((\sqrt{x^-}).f.(\sqrt{x^-}))(x) = f(\sqrt{x^-}(x^+ - x^-)\sqrt{x^-}) = = f(x^- x^+ - (x^-)^2) = -f((x^-)^2) \leqq 0,$$

i.e. $f((x^-)^* x^-) = 0$ $(f \in P)$. Since P separates the points of A and the involution of A is proper, it follows that $x^- = 0$ thus $x = x^+$ is a positive element in the C^*-algebra A.

II. A type of commutative GW*-algebras

If $\mathscr{B}$ is a σ-algebra of subsets of the set T then $\mathscr{F}_{\mathbf{C}}^b(T, \mathscr{B})$ will denote the set of bounded complex valued $\mathscr{B}-\mathscr{B}(\mathbf{C})$ measurable functions defined on T. The set $\mathscr{F}_{\mathbf{C}}^b(T, \mathscr{B})$ will always be thought of equipped with the pointwise defined algebraic structure and the sup-norm on T (denoted by $|||\cdot|||_T$), thus $\mathscr{F}_{\mathbf{C}}^b(T, \mathscr{B})$ will be regarded as a commutative unital C^*-algebra.

It is known that given a σ-algebra of subsets of the set T and a finitely additive mapping $\Theta: \mathscr{B} \to \mathbf{C}$, the following statements are equivalent:

— Θ is bounded, i.e. $\sup_{E \in \mathscr{B}} |\Theta(E)| < +\infty$;

— there is a unique continuous linear form $\hat{\Theta}$ on $\mathscr{F}_{\mathbf{C}}^b(T, \mathscr{B})$ (called the integral on $\mathscr{F}_{\mathbf{C}}^b(T, \mathscr{B})$ defined by Θ) such that $\hat{\Theta}(\chi_E) = \Theta(E)$ for all $E \in \mathscr{B}$.

Moreover, Θ is σ-additive if and only if the integral $\hat{\Theta}$ defined by Θ satisfies the condition:

(L) For every uniformly bounded sequence $(\varphi_n)_{n \in \mathbf{N}}$ of functions in $\mathscr{F}_{\mathbf{C}}^b(T, \mathscr{B})$, if $\varphi_n \to 0$ pointwise on T then $\hat{\Theta}(\varphi_n) \to 0$.

Lemma. *Let $\mathscr{B}$ be a σ-algebra of subsets of the set T and P the set of integrals on $\mathscr{F}_{\mathbf{C}}^b(T, \mathscr{B})$ defined by positive σ-additive set functions on $\mathscr{B}$. Then P is a separating set of positive linear forms on the unital $*$-algebra $A := \mathscr{F}_{\mathbf{C}}^b(T, \mathscr{B})$, P satisfies* (I) *and* sp (P) *is a $\|\cdot\|_P'$-closed set.*

Proof. Since $\{\hat{\delta}_t \mid t \in T\} \subset P$, the set P separates the points of A. On the other hand, $P(1) = \{\hat{\mu} \mid \mu: \mathscr{B} \to \mathbf{R}_+$ σ-additive and $\mu(T) = \hat{\mu}(1) \leqq 1\}$, thus for every $\varphi \in A$

and $\hat{\mu}\in P$ we have the inequality $|\hat{\mu}(\varphi)|\leqq\mu(T)|||\varphi|||_T$ showing that $P(1)$ is $\sigma(A^*, A)$-bounded and non void.

Now we prove that $\|\cdot\|_P=|||\cdot|||_T$. Indeed, if $\varphi\in A$ then

$$\|\varphi\|_P := \sup_{\hat{\mu}\in P(1)} \sqrt{\hat{\mu}(\varphi^*\varphi)} = \sup_{\hat{\mu}\in P(1)} \sqrt{\hat{\mu}(|\varphi|^2)} \leqq \sup_{\hat{\mu}\in P(1)} \sqrt{\mu(T)}\,|||\varphi|||_T \leqq |||\varphi|||_T,$$

i.e. $\|\cdot\|_P\leqq|||\cdot|||_T$. Conversely, if $\varphi\in A$ and $c<|||\varphi|||_T$ then there is a point t in T such that $c<|\varphi(t)|=\sqrt{\hat{\delta}_t(\varphi^*\varphi)}\leqq\|\varphi\|_P$, i.e. $|||\cdot|||_T\leqq\|\cdot\|_P$.

Let $\Theta\in\overline{\mathrm{sp}}\,(P)$ and choose a sequence $(\Theta_n)_{n\in\mathbf{N}}$ in $\mathrm{sp}\,(P)$ with the property $\|\Theta_n-\Theta\|'_P\to 0$. We have to show that $\Theta\in\mathrm{sp}\,(P)$. With regard to our former considerations, it sufficies to prove that for every uniformly bounded sequence $(\varphi_n)_{n\in\mathbf{N}}$ in A, if $\varphi_n\to 0$ pointwise on T then $\Theta(\varphi_n)\to 0$. If $n, m\in\mathbf{N}$ then

$$|\Theta(\varphi_m)| \leqq |\Theta(\varphi_m)-\Theta_n(\varphi_m)|+|\Theta_n(\varphi_m)| \leqq \|\Theta-\Theta_n\|'_P|||\varphi_m|||_T+|\Theta_n(\varphi_m)|.$$

If $\varepsilon>0$ is arbitrary then there is a number N_0 in $\mathbf{N}$ such that $\|\Theta-\Theta_{N_0}\|'_P\leqq\varepsilon/2(M+1)$ where $M:=\sup_{m\in\mathbf{N}}|||\varphi_m|||_T$. Since $\Theta_{N_0}\in\mathrm{sp}\,(P)$ we have $\Theta_{N_0}(\varphi_m)\to 0$ $(m\to+\infty)$ thus there is a number N in $\mathbf{N}$ with the property that $|\Theta_{N_0}(\varphi_m)|\leqq\varepsilon/2$ for $m\in\mathbf{N}$, $m\geqq N$. Then the above inequality implies that $|\Theta(\varphi_m)|\leqq\varepsilon$ for $m\in\mathbf{N}$, $m\geqq N$, i.e. $\Theta(\varphi_m)\to 0$.

Theorem 1. *Let $\mathscr{B}$ be a σ-algebra of subsets of the set T, $A:=\mathscr{F}^b_{\mathbf{C}}(T, \mathscr{B})$ and P the set of integrals on A defined by positive σ-additive set functions on $\mathscr{B}$. Then (A, P) is a commutative GW^*-algebra.*

Proof. With regard to our Lemma we have only to prove that the pair (A, P) satisfies (II), (III) and (IV). If $\varphi\in A$ and $\hat{\mu}\in P$ then $\varphi^*\cdot\hat{\mu}\cdot\varphi=|\varphi|^2\mu$ where $|\varphi|^2\mu$ is the positive σ-additive set function on $\mathscr{B}$ defined as: $E\mapsto\hat{\mu}(|\varphi|^2\chi_E)$, thus $\varphi^*\cdot\hat{\mu}\cdot\varphi\in P$ and, consequently, $\varphi\cdot\hat{\mu}\in P-P+iP-iP\subset\mathrm{sp}\,(P)$, i.e. (A, P) verifies (II) and (III).

In order to prove (IV), let $(\varphi_n)_{n\in\mathbf{N}}$ be a sequence in A such that $(\hat{\mu}(\varphi_n))_{n\in\mathbf{N}}$ is convergent for every $\hat{\mu}\in P$. Since $\hat{\delta}_t\in P$ $(t\in T)$, there is a unique function $\varphi\colon T\to\mathbf{C}$ with the property that $\varphi_n\to\varphi$ pointwise on T. From this we infer that φ is necessarily $\mathscr{B}-\mathscr{B}(\mathbf{C})$ measurable. We intend to show that $\varphi\in A$ and $\varphi_n\to\varphi$ in the $\sigma(A, \mathrm{sp}\,(P))$ topology. In order to prove this we first define for all $n\in\mathbf{N}$ the linear form $\tilde{\varphi}_n\colon \mathrm{sp}\,(P)\to\mathbf{C}$; $\Theta\mapsto\Theta(\varphi_n)$. On account of our Lemma, $\mathrm{sp}\,(P)$ will be considered a Banach space whose norm equals $\|\cdot\|'_P$. Then $\tilde{\varphi}_n$ is a continuous linear form on the Banach space $\mathrm{sp}\,(P)$ for every $n\in\mathbf{N}$ and, by our assumption, the sequence $(\tilde{\varphi}_n)_{n\in\mathbf{N}}$ is pointwise convergent in $\mathrm{sp}\,(P)$. Consequently, the theorem of Banach—Steinhaus implies that $\sup_{n\in\mathbf{N}}\|\tilde{\varphi}_n\|<+\infty$. If $n\in\mathbf{N}$ and $c<|||\varphi_n|||_T$ then there is a point t in T such that $c<|\varphi_n(t)|=|\tilde{\varphi}_n(\hat{\delta}_t)|\leqq\|\hat{\delta}_t\|'_P\|\tilde{\varphi}_n\|=\|\tilde{\varphi}_n\|$, since $\|\hat{\delta}_t\|'_P=1$ holds obviously, thus $|||\varphi_n|||_T \leqq \|\tilde{\varphi}_n\|$ showing that the sequence $(\varphi_n)_{n\in\mathbf{N}}$ is uniformly bounded in T. From this we obtain that the mapping φ is bounded, i.e. $\varphi\in A$.

Finally, if $\hat{\mu}\in P$ then the theorem of Lebesgue applied to the measure μ and the uniformly bounded, pointwise convergent sequence $(\varphi_n)_{n\in\mathbf{N}}$ result in $\hat{\mu}(\varphi_n)\to\hat{\mu}(\varphi)$, i.e. $\varphi_n\to\varphi$ in the $\sigma(A, \mathrm{sp}\,(P))$ topology.

This theorem provides a great deal of commutative GW^*-algebras that are not $*$-isomorphic to any W^*-algebra.

III. On the Gelfand representation of commutative GW*-algebras

If T is a compact Hausdorff space then $\mathscr{C}_{\mathbf{C}}(T)$ and $\mathscr{M}_{\mathbf{C}}(T)$ will denote the vector space of complex continuous functions defined on T and the vector space of complex Radon measures on T, respectively. Then $\mathscr{C}_+(T)$ and $\mathscr{M}_+(T)$ denote the convex cone of positive elements in $\mathscr{C}_{\mathbf{C}}(T)$ and $\mathscr{M}_{\mathbf{C}}(T)$, respectively. The complex vector space $\mathscr{C}_{\mathbf{C}}(T)$ will always be thought of equipped with the pointwise defined multiplication and conjugation, i.e. $\mathscr{C}_{\mathbf{C}}(T)$ will be considered a commutative unital $*$-algebra. It is well known that $\mathscr{C}_{\mathbf{C}}(T)$ is a C^*-algebra whose C^*-norm equals the sup-norm $|||\cdot|||_T$ on T.

Given a commutative unital C^*-algebra A, the celebrated representation theorem of Gelfand and Naimark assures that A and $\mathscr{C}_{\mathbf{C}}(X(A))$ are isometrically $*$-isomorphic C^*-algebras, where $X(A)$ denotes the compact Hausdorff space whose underlying set is the set of non zero multiplicative linear forms on A and whose topology is the well known Gelfand topology (cf. [3] ch. I, §6, Theorem 1). The Gelfand isomorphism between A and $\mathscr{C}_{\mathbf{C}}(X(A))$ is denoted usually by $\mathscr{G}_A$; we have $(\mathscr{G}_A(x))(\chi)=\chi(x)$ for all $x\in A$ and $\chi\in X(A)$.

In this section the structure of the compact Hausdorff space $X(A)$ will be examined in the case when (A, P) is a commutative GW^*-algebra.

Proposition 2. *Let T be a compact Hausdorff space, $P\subset\mathscr{M}_+(T)$ and suppose that $(\mathscr{C}_{\mathbf{C}}(T), P)$ is a weak GW^*-algebra. Then*

(i) $T=\big(\bigcup_{\mu\in P}\mathrm{Supp}\,\mu\big)^-$ *and* $\sup_{\mu\in P}\mu(G)>0$ *for every non-void open subset G of T.*

(ii) *The interior of a closed G_δ-set in T is closed.*

(iii) *If F is a closed G_δ-set in T and there is a measure μ in P such that $\mu(F)>0$ then the interior $\mathring{F}$ of F is non-void, i.e. F is not nowhere dense in T.*

Proof. (i) Let G be a non-void open subset of T. Then there is a function $\varphi\in\mathscr{C}_+(T)$ such that $0\leqq\varphi\leqq 1$, $\mathrm{Supp}\,\varphi\subset G$ and $\varphi\neq 0$. Since P is a separating set, there exists a measure μ in P with the property $\mu(\varphi)>0$. Then we have $\mu(G)\geqq$

$\geqq \mu(\varphi) > 0$. This proves the second part of (i) and the first part of our assertion is an easy consequence of the second part.

(ii) Let F be a closed G_δ-set in T. Then there is a sequence of functions $(\varphi_n)_{n\in\mathbf{N}}$ in $\mathscr{C}_+(T)$ such that $\varphi_n \geqq \varphi_{n+1}$ $(n\in\mathbf{N})$ and $\varphi_n \to \chi_F$ pointwise on T. If $\mu\in P$ then $(\mu(\varphi_n))_{n\in\mathbf{N}}$ is a decreasing sequence of positive real numbers thus the sequentially completeness of $\mathscr{C}_C(T)$ in the $\sigma(\mathscr{C}_C(T), \mathrm{sp}(P))$ topology now gives the existence of a function φ in $\mathscr{C}_C(T)$ such that $\mu(\varphi_n)\to\mu(\varphi)$ for all $\mu\in P$. Since $\mu\in P$ implies $\mu(\varphi)\geqq 0$ and $\mu(\varphi_n)\geqq\mu(\varphi)$ $(n\in\mathbf{N})$, by Proposition 1 we obtain that $\varphi_n\geqq\varphi\geqq 0$ $(n\in\mathbf{N})$. From this we conclude that $\varphi\leqq\chi_F$. If $\varphi'\in\mathscr{C}_+(T)$ and $\varphi'\leqq\chi_F$ then $\varphi'\leqq \varphi_n$ $(n\in\mathbf{N})$ thus $\mu(\varphi')\leqq\mu(\varphi_n)$ and $\mu(\varphi')\leqq\lim_n\mu(\varphi_n)=\mu(\varphi)$ for every $\mu\in P$, i.e. applying again Proposition 1, we find that $\varphi'\leqq\varphi$. This means that

$$\varphi = \sup\{\varphi' \colon \varphi'\in\mathscr{C}_+(T), \varphi' \leqq \chi_F\}. \tag{1}$$

If $n\in\mathbf{N}$ then $\inf(n\varphi, 1)\leqq\chi_F$ and $\inf(n\varphi, 1)\in\mathscr{C}_+(T)$ thus by (1) we obtain $\inf(n\varphi, 1)\leqq\varphi$. Then we have

$$\chi_{[\varphi>0]} = \sup_{n\in\mathbf{N}}\big(\inf(n\varphi, 1)\big) \leqq \varphi \leqq \chi_F$$

showing that $\varphi=1$ on the set $[\varphi>0]$ thus $\varphi=1$ on the set $\mathrm{Supp}\,\varphi=[\varphi>0]^-$ as well. Since $\varphi=0$ on $T\setminus\mathrm{Supp}\,\varphi$ we deduce that $\chi_{\mathrm{Supp}\,\varphi}=\varphi\in\mathscr{C}_+(T)$, i.e. $\mathrm{Supp}\,\varphi$ is an open-closed subset of T and $\mathrm{Supp}\,\varphi\subset F$ thus $\mathrm{Supp}\,\varphi\subset\mathring{F}$. We claim that $\mathring{F}$ equals $\mathrm{Supp}\,\varphi$. On the contrary, suppose that $\mathrm{Supp}\,\varphi\neq\mathring{F}$. Then $\mathring{F}\setminus\mathrm{Supp}\,\varphi$ is a non-void open subset of T thus there is a mapping $\varphi'\in\mathscr{C}_+(T)$ such that $0\leqq\varphi'\leqq 1$, $\mathrm{Supp}\,\varphi'\subset\mathring{F}\setminus\mathrm{Supp}\,\varphi$ and $\varphi'\neq 0$. Then $\varphi+\varphi'\in\mathscr{C}_+(T)$ and $\varphi+\varphi'\leqq\chi_F$ thus by (1) we have $\varphi+\varphi'\leqq\varphi$ in contradiction to $\varphi'\neq 0$. This proves that $\mathrm{Supp}\,\varphi=\mathring{F}$, i.e. the interior of the closed G_δ-set F is closed in T.

(iii) If F is a closed G_δ-set in T and $\mu\in P$ is a measure such that $\mu(F)>0$ then, applying the notations introduced in the proof of (ii), we obtain

$$\mu(\varphi) = \lim_n \mu(\varphi_n) = \mu(\chi_F) = \mu(F)$$

thus $\varphi\neq 0$, i.e. $\emptyset\neq\mathrm{Supp}\,\varphi=\mathring{F}$.

Corollary 1. *Let T be a compact Hausdorff space and let $P\subset\mathscr{M}_+(T)$ be a set such that $(\mathscr{C}_C(T), P)$ is a weak GW^*-algebra. Then the open-closed subsets of T form a basis for the topology of T and the closure of every open F_σ-set is open in T. Particularly,* $\mathrm{Supp}\,\varphi$ *is open-closed for all* $\varphi\in\mathscr{C}_C(T)$.

Proof. Let t be an arbitrary point of T and G an open neighbourhood of t. Then we can choose a function $\varphi\in\mathscr{C}_+(T)$ with the property that $0\leqq\varphi\leqq 1$, $\mathrm{Supp}\,\varphi\subset G$ and t is in the interior of $[\varphi=1]$. Since $[\varphi=1]$ is G_δ in T, by Proposition

2 we deduce that the interior of $[\varphi=1]$ is open-closed and contained in G. This means that at every point of T there is a basis consisting of open-closed sets, or equivalently, the topology of T has a basis formed by open-closed sets.

The second part of our assertion is a simple reformulation of (ii) in Proposition 2.

Theorem 2. *Let (A, P) be a commutative weak GW^*-algebra. Then A is a C^*-algebra whose underlying $*$-algebra is a Rickart $*$-algebra. Consequently, the set of projectors in A is total in the topology defined by the C^*-norm of A.*

Proof. Compare Corollary 1 with Theorems 1, ch. I, § 6. in [3] and 1.8 in [4].

IV. Spectral theorem for commutative GW*-algebras

If T is a compact Hausdorff space then $\mathscr{B}_0(T)$ denotes the σ-algebra in T generated by the closed G_δ subsets of T; $\mathscr{B}_0(T)$ is usually referred to as the Baire σ-algebra of T. On the other hand, a mapping $\varphi\colon T\to\mathbf{C}$ is called a Baire function if $\varphi^{-1}(E)\in\in\mathscr{B}_0(T)$ for every Borel set E in $\mathbf{C}$. It can be shown without difficulty that $\mathscr{B}_0(T)$ coincides with the least σ-algebra in T with respect to which every continuous complex valued function defined on T is measurable.

Let T be a compact Hausdorff space; for every countable ordinal number α we define by ω_1-induction the function space $\mathscr{C}_{\mathbf{C}}^{\alpha}(T)$ as follows:

— $\mathscr{C}_{\mathbf{C}}^{0}(T):=\mathscr{C}_{\mathbf{C}}(T)$,

— if $0<\alpha<\omega_1$ then $\varphi\in\mathscr{C}_{\mathbf{C}}^{\alpha}(T)$ if and only if φ is a function $T\to\mathbf{C}$ such that there is a sequence $(\varphi_n)_{n\in\mathbf{N}}$ in $\bigcup_{\beta<\alpha}\mathscr{C}_{\mathbf{C}}^{\beta}(T)$ which is uniformly bounded and pointwise converges to φ in T.

Then we define $\mathscr{C}_{\mathbf{C}}^{\infty}(T):=\bigcup_{\alpha<\omega_1}\mathscr{C}_{\mathbf{C}}^{\alpha}(T)$. It is easy to show that $\mathscr{C}_{\mathbf{C}}^{\infty}(T)==\mathscr{F}_{\mathbf{C}}^{b}(T,\mathscr{B}_0(T))$, i.e. $\mathscr{C}_{\mathbf{C}}^{\infty}(T)$ consists of the bounded complex valued Baire functions defined on T and a subset E of T belongs to $\mathscr{B}_0(T)$ if and only if $\chi_E\in\mathscr{C}_{\mathbf{C}}^{\infty}(T)$. In the sequel the sequence of function spaces $(\mathscr{C}_{\mathbf{C}}^{\alpha}(T))_{\alpha<\omega_1}$ will be referred to as the *standard graduation* of $\mathscr{C}_{\mathbf{C}}^{\infty}(T)$.

According to Theorem 1 and the fact that $\mathscr{C}_{\mathbf{C}}^{\infty}(T)=\mathscr{F}_{\mathbf{C}}^{b}(T,\mathscr{B}_0(T))$, the pair $(\mathscr{C}_{\mathbf{C}}^{\infty}(T), P)$ is a commutative GW^*-algebra, where P is the set of integrals on $\mathscr{C}_{\mathbf{C}}^{\infty}(T)$ defined by positive σ-additive set functions on $\mathscr{B}_0(T)$.

Lemma 2. *If T is a compact Hausdorff space, $P\subset\mathscr{M}_{\mathbf{C}}(T)$ and φ is a universally integrable complex valued function defined on T then the relation $\int_T\varphi\,\mathrm{d}\mu=0$ $(\mu\in P)$ implies that $\int_T\varphi\,\mathrm{d}\mu=0$ for all $\mu\in\overline{\mathrm{sp}}(P)$, where $\overline{\mathrm{sp}}(P)$ is the closure of $\mathrm{sp}(P)$ in $\mathscr{M}_{\mathbf{C}}(T)$ in the measure norm topology.*

Proof. Since the mapping $\mathcal{M}_{\mathbf{C}}(T)\to\mathbf{C}$, $\Theta\mapsto\int_T\varphi\,\mathrm{d}\Theta$ is a measure-norm continuous linear form on $\mathcal{M}_{\mathbf{C}}(T)$, the assertion is obviously true.

Lemma 3. *Let T be a compact Hausdorff space and let $P\subset\mathcal{M}_+(T)$ be a set such that $(\mathscr{C}_{\mathbf{C}}(T), P)$ is a GW^*-algebra. If $\varphi\in\mathscr{C}^\infty_{\mathbf{C}}(T)$, $\varphi^b\in\mathscr{C}_{\mathbf{C}}(T)$ and $\int_T\varphi\,\mathrm{d}\mu=\mu(\varphi^b)$ for all $\mu\in P$ then we have $|||\varphi^b|||_T\leqq|||\varphi|||_T$.*

Proof. Let t be a fixed point of T and $\mathscr{B}_t$ denote the basis at t of T consisting of open-closed subsets of T (see Proposition 2, Corollary 1). With regard to (i) in Proposition 2, to every $E\in\mathscr{B}_t$ there is a measure μ_E in P such that $\mu_E(E)>0$. Let μ_E be such a measure and put $\lambda_E:=\chi_E\mu_E/\mu_E(E)$ for every $E\in\mathscr{B}_t$. Then $\lambda_E\in\overline{\mathrm{sp}}(P)$ by (III), and it is easy to see that the continuity of φ^b in t implies that $\lim_{E,\mathscr{B}_t}\lambda_E(\varphi^b)=\varphi^b(t)$. Now Lemma 2 yields that $\int_T\varphi\,\mathrm{d}\lambda_E=\lambda_E(\varphi^b)$ for all $E\in\mathscr{B}_t$, since the measure-norm closure of sp (P) in $\mathcal{M}_{\mathbf{C}}(T)$ equals $\overline{\mathrm{sp}}$ (P) (viz. $|||\cdot|||_T=\|\cdot\|_P$). From this we infer that

$$|\varphi^b(t| = \lim_{E,\mathscr{B}_t}|\lambda_E(\varphi^b)| = \lim_{E,\mathscr{B}_t}\Big|\int_T\varphi\,\mathrm{d}\lambda_E\Big| \leqq |||\varphi|||_T,$$

i.e. $|||\varphi^b|||_T\leqq|||\varphi|||_T$.

Proposition 3. *Let T be a compact Hausdorff space, $P\subset\mathcal{M}_+(T)$ and suppose that $(\mathscr{C}_{\mathbf{C}}(T), P)$ is a GW^*-algebra. Then to every bounded complex valued Baire function φ defined on T there is a unique continuous function φ^b defined on T with the property that $\varphi=\varphi^b$ a.e. for all $\mu\in P$.*

Proof. Since P separates the points of $\mathscr{C}_{\mathbf{C}}(T)$, the uniqueness of φ^b is obvious. The existence of φ^b will be shown by the use of the standard graduation of $\mathscr{C}^\infty_{\mathbf{C}}(T)$. Assume that $\varphi\in\mathscr{C}^\infty_{\mathbf{C}}(T)$ and by ω_1-induction we show that for every $\alpha<\omega_1$, if $\varphi\in\mathscr{C}^\alpha_{\mathbf{C}}(T)$ then there is a function $\varphi^b\in\mathscr{C}_{\mathbf{C}}(T)$ such that $\varphi=\varphi^b$ a.e., for all $\mu\in P$.

The assertion holds for $\alpha=0$, evidently. Suppose that $0<\alpha<\omega_1$ and the assertion is true for every $\beta<\alpha$. Since $\varphi\in\mathscr{C}^\alpha_{\mathbf{C}}(T)$, there is a uniformly bounded sequence $(\varphi_n)_{n\in\mathbf{N}}$ in $\bigcup_{\beta<\alpha}\mathscr{C}^\beta_{\mathbf{C}}(T)$ such that $\varphi_n\to\varphi$ pointwise on T. With regard to our induction hypothesis, for every $n\in\mathbf{N}$ we can define a function φ_n^b in $\mathscr{C}_{\mathbf{C}}(T)$ such that $\varphi_n=\varphi_n^b$ a.e., for all $\mu\in P$. Now Lemma 3 gives that $|||\varphi_n^b|||_T\leqq|||\varphi_n|||_T$ $(n\in\mathbf{N})$ so the sequence $(\varphi_n^b)_{n\in\mathbf{N}}$ in $\mathscr{C}_{\mathbf{C}}(T)$ is also uniformly bounded.

If $\mu\in P$ then the theorem of Lebesgue applied to μ and the sequence $(\varphi_n)_{n\in\mathbf{N}}$ implies $\int_T\varphi_n\,\mathrm{d}\mu\to\int_T\varphi\,\mathrm{d}\mu$. On the other hand, $\int_T\varphi_n\,\mathrm{d}\mu=\mu(\varphi_n^b)$ $(n\in\mathbf{N})$ thus we obtain

$$\lim_n\mu(\varphi_n^b) = \int_T\varphi\,\mathrm{d}\mu \quad (\mu\in P). \tag{2}$$

The sequentially completeness of $\mathscr{C}_{\mathbf{C}}(T)$ in the $\sigma(\mathscr{C}_{\mathbf{C}}(T), \operatorname{sp}(P))$ topology now results in the existence of a function $\varphi^b \in \mathscr{C}_{\mathbf{C}}(T)$ such that

$$\lim_n \mu(\varphi_n^b) = \mu(\varphi^b) \quad (\mu \in P). \tag{3}$$

Comparing (2) and (3) we deduce that $\mu(\varphi^b) = \int_T \varphi \, d\mu$ for every $\mu \in P$. According to Lemma 3, $\Theta(\varphi^b) = \int_T \varphi \, d\Theta$ for all $\Theta \in \overline{\operatorname{sp}}(P)$. If $\mu \in P$ and $\psi \in \mathscr{C}_{\mathbf{C}}(T)$ then by (III) we have $\psi\mu \in \overline{\operatorname{sp}}(P)$ thus $(\varphi\mu)(\psi) = \int_T \varphi \, d(\psi\mu) = (\psi\mu)(\varphi^b) = (\varphi^b\mu)(\psi)$, i.e. $\varphi\mu = \varphi^b\mu$ for all $\mu \in P$. This shows that $\varphi = \varphi^b$ a.e., for every $\mu \in P$.

We call the attention to the fact that Proposition 3 holds only for commutative GW^*-algebras and not for commutative *weak* GW^*-algebras.

Theorem 3. *Let (A, P) be a commutative GW^*-algebra. Then there is a unique $*$-homomorphism $\Theta^P \colon \mathscr{C}_{\mathbf{C}}^{\infty}(X(A)) \to A$ preserving the unit elements satisfying*

$$f(\Theta^P(\varphi)) = \int_{X(A)} \varphi \, d(f \circ \mathscr{G}_A^{-1}) \tag{4}$$

for all $f \in P$ and $\varphi \in \mathscr{C}_{\mathbf{C}}^{\infty}(X(A))$.

Remark. Note that $f \circ \mathscr{G}_A^{-1} \in \mathscr{M}_+(X(A))$ for every positive linear form f on A.

Proof. The uniqueness of Θ^P follows from (4) and the fact that P separates the points of A. In order to prove the existence of Θ^P we first mention that the pair $(\mathscr{C}_{\mathbf{C}}(X(A)), P \circ \mathscr{G}_A^{-1})$ is a commutative GW^*-algebra. Then, by Proposition 3, we can define the mapping

$$\mathscr{C}_{\mathbf{C}}^{\infty}(X(A)) \to \mathscr{C}_{\mathbf{C}}(X(A)), \quad \varphi \mapsto \varphi^b$$

satisfying $\varphi = \varphi^b$ a.e., for every $\mu \in P \circ \mathscr{G}_A^{-1}$ and $\varphi \in \mathscr{C}_{\mathbf{C}}^{\infty}(X(A))$. It is routine to check that this mapping is a $*$-homomorphism between $\mathscr{C}_{\mathbf{C}}^{\infty}(X(A))$ and $\mathscr{C}_{\mathbf{C}}(X(A))$ preserving the unit elements. For every $\varphi \in \mathscr{C}_{\mathbf{C}}^{\infty}(X(A))$ we define $\Theta^P(\varphi) := \mathscr{G}_A^{-1}(\varphi^b)$. Then Θ^P is a unit preserving $*$-homomorphism between $\mathscr{C}_{\mathbf{C}}^{\infty}(X(A))$ and A, evidently. If $f \in P$ and $\varphi \in \mathscr{C}_{\mathbf{C}}^{\infty}(X(A))$ then $f \circ \mathscr{G}_A^{-1} \in P \circ \mathscr{G}_A^{-1}$ thus $\varphi = \varphi^b$ a.e., for $f \circ \mathscr{G}_A^{-1}$, showing that the equality holds for φ and f.

Of course, Theorem 3 can be appreciated as the global (or better to say, collective) spectral theorem for commutative GW^*-algebras. In order to formulate an individual version of the spectral theorem, we note that the spectrum of an element x in a unital algebra A is usually denoted by $\operatorname{Sp}_A(x)$, or, if no confusion arises as for the algebra, the letter A is omitted. It is well known that given a unital C^*-algebra A, to every normal element x of A there is a unique unit preserving $*$-homomorphism $\Theta_x \colon \mathscr{C}_{\mathbf{C}}(\operatorname{Sp}(x)) \to A$ such that $\Theta_x(\mathrm{id}_{\operatorname{Sp}(x)}) = x$ and Θ_x is an isometry whose range equals the C^*-subalgebra of A generated by the set $\{1, x, x^*\}$ (cf. [3] ch. I, § 6, Proposition 5).

Theorem 4. *Let (A, P) be a commutative GW^*-algebra and $x \in A$. Then there exists a unique unit preserving $*$-homomorphism Θ_x^P: $\mathcal{F}_C^b(\mathrm{Sp}(x), \mathcal{B}(\mathrm{Sp}(x))) \to A$ which is an extension of Θ_x and satisfies*

$$f(\Theta_x^P(\varphi)) = \int_{\mathrm{Sp}(x)} \varphi \, d(f \circ \Theta_x) \quad (f \in P), \tag{5}$$

for every bounded complex valued Borel function φ defined on $\mathrm{Sp}(x)$.

Remark. Note that $f \circ \Theta_x \in \mathcal{M}_+(\mathrm{Sp}(x))$ for every positive linear form on A.

Proof. The set P separates the points of A, thus the uniqueness of Θ_x^P follows from (5), evidently.

Since $\mathrm{Sp}(x)$ is a metrisable compact topological space, the σ-algebra $\mathcal{B}(\mathrm{Sp}(x))$ of Borel sets in $\mathrm{Sp}(x)$ coincides with the σ-algebra $\mathcal{B}_0(\mathrm{Sp}(x))$ of Baire sets in $\mathrm{Sp}(x)$. Consequently, we have $\mathcal{F}_C^b(\mathrm{Sp}(x), \mathcal{B}(\mathrm{Sp}(x))) = \mathcal{C}_C^\infty(\mathrm{Sp}(x))$. Since the mapping $\mathcal{G}_A(x)$ is a continuous function from $X(A)$ onto $\mathrm{Sp}(x)$, the operator

$$\mathcal{G}_A(x)^{\#} : \mathcal{C}_C^\infty(\mathrm{Sp}(x)) \to \mathcal{C}_C^\infty(X(A)), \quad \varphi \mapsto \varphi \circ \mathcal{G}_A(x)$$

is an injective unit preserving $*$-homomorphism between the C^*-algebras $\mathcal{C}_C^\infty(\mathrm{Sp}(x))$ and $\mathcal{C}_C^\infty(X(A))$. Then we put

$$\Theta_x^P := \Theta^P \circ \mathcal{G}_A(x)^{\#},$$

where Θ^P denotes the $*$-homomorphism between $\mathcal{C}_C^\infty(X(A))$ and A, introduced in Theorem 3. Thus Θ_x^P is a unit preserving $*$-homomorphism between $\mathcal{C}_C^\infty(\mathrm{Sp}(x))$ and A. It remained to prove the equality (5). Let there be given a linear form $f \in P$ and a function $\varphi \in \mathcal{C}_C^\infty(\mathrm{Sp}(x))$. Then, by the definition of Θ_x^P, we have

$$\begin{aligned} f(\Theta_x^P(\varphi)) &= f(\Theta^P(\mathcal{G}_A(x)^{\#}(\varphi))) = f(\Theta^P(\varphi \circ \mathcal{G}_A(x))) = \\ &= \int_{X(A)} \varphi \circ \mathcal{G}_A(x) \, \mathrm{d}(f \circ \mathcal{G}_A^{-1}) = \int_{X(A)} \varphi \, \mathrm{d}(\mathcal{G}_A(x)(f \circ \mathcal{G}_A^{-1})), \end{aligned} \tag{6}$$

where $\mathcal{G}_A(x)(f \circ \mathcal{G}_A^{-1})$ denotes the Radon measure on $\mathrm{Sp}(x)$, which is the image of the measure $f \circ \mathcal{G}_A^{-1} \in \mathcal{M}_+(X(A))$ established by the continuous function $\mathcal{G}_A(x)$. It is obvious that the mapping

$$\mathcal{C}_C(\mathrm{Sp}(x)) \mapsto A, \quad \psi \mapsto \mathcal{G}_A^{-1}(\psi \circ \mathcal{G}_A(x))$$

is a unit preserving $*$-homomorphism between $\mathcal{C}_C(\mathrm{Sp}(x))$ and A which assigns x to $\mathrm{id}_{\mathrm{Sp}(x)}$, so the uniqueness of Θ_x results in $\Theta_x(\psi) = \mathcal{G}_A^{-1}(\psi \circ \mathcal{G}_A(x))$ for all $\psi \in \mathcal{C}_C(\mathrm{Sp}(x))$. Thus we obtain $(f \circ \Theta_x)(\psi) = (f \circ \mathcal{G}_A^{-1})(\psi \circ \mathcal{G}_A(x))$ for every

$\psi\in\mathscr{C}_{\mathbf{C}}(\mathrm{Sp}\,(x))$ showing that $f\circ\Theta_x=\mathscr{G}_A(x)(f\circ\mathscr{G}_A^{-1})$. Comparing this equality with (6), we finally deduce that (5) holds for every $f\in P$ and $\varphi\in\mathscr{C}_{\mathbf{C}}^{\infty}(\mathrm{Sp}\,(x))$.

At last we mention that both the $*$-homomorphisms Θ^P and Θ_x^P introduced in Theorem 3 and Theorem 4, respectively, depend essentially on P.

References

[1] J. Kristóf, C^*-norms defined by positive linear forms, *Acta Sci. Math. (Szeged)*, **50** (1986), 427—432.

[2] J. Kristóf, On the projection lattice of GW^*-algebras, *Studia Sci. Math. Hungar.*, (to appear).

[3] N. Bourbaki, *Éléments de Mathématique, Théories Spectrales*, Hermann (Paris, 1967).

[4] S. Berberian, *Baer $*$-rings*, Springer-Verlag (Berlin—Heidelberg—New York, 1972).

DEPARTMENT OF APPLIED ANALYSIS
EÖTVÖS LORÁND UNIVERSITY
MÚZEUM KRT. 6—8
1088 BUDAPEST, HUNGARY

Acta Sci. Math., 52 (1988), 157—163

The invariance principle for functionals of sums of martingale differences

I. SZYSZKOWSKI

1. Introduction. Let $\{(X_{ni}, F_{ni}),\ 1\leqq i\leqq k_n\},\ n\geqq 1$, be a double array of square-integrable random variables whose rows are martingale difference sequences (MDS), i.e. for each $n\geqq 1$ the rv's $X_{ni},\ 1\leqq i\leqq k_n$, given on some probability space $(\Omega, \mathcal{A}, P)$ with sub-σ-fields $F_{n0}\subset F_{n1}\subset\ldots\subset F_{nk_n}$, are such that X_{ni} is F_{ni}-measurable and $E(X_{ni}|F_{n,i-1})=0$ a.s. for every $1\leqq i\leqq k_n$. Define

$$S_{nk} = \sum_{i=1}^{k} X_{ni}, \quad \sigma_{ni}^2 = E(X_{ni}^2|F_{n,i-1}),$$

$s_{nk}^2=ES_{nk}^2$ and $S_{nk}=s_{nk}^2=0$ if $k=0,\ n\geqq 1$. Let us observe that without loss of generality we may and do assume that for every $n\geqq 1$, $EX_{ni}^2\neq 0,\ 1\leqq i\leqq k_n,\ s_n^2=s_{nk_n}^2=1$, where $k_n\to\infty$ as $n\to\infty$.

Let $D[0, 1]$ be the space of functions defined on [0, 1] that are right-continuous and have left hand limits, endowed with the Skorohod J_1-topology (cf. [1, §14]). By W we will denote the Wiener measure on $D[0, 1]$ with the corresponding Wiener process $\{W(t)\colon 0\leqq t\leqq 1\}$.

Let F_M be the space of functions defined on $[0, 1]\times(-\infty, \infty)$ satisfying the following condition: there exists an absolute constant M such that if $f\in F_M$, then f and its derivatives satisfy inequalities of the form

$$(1)\qquad |Df(s, x)| \leqq M(1+|x|^{\alpha}),$$

where D denotes either the identity operator or a first derivative and α is some positive constant.

Define a random function $W_n(t),\ 0\leqq t\leqq 1$, by

$$(2)\qquad W_n(t) = S_{n, m_n(t)}, \quad n \geqq 1,$$

where $m_n(t)=\max\{i\leqq k_n\colon s_{ni}^2\leqq t\},\ t\in[0, 1]$.

Received November 27, 1984, and in revised form February 21, 1985.

We shall give sufficient conditions for the weak convergence of the process $\{Z_n(t)=\sum_{i=0}^{m_n(t)-1} f_n(s_{ni}^2, S_{ni})X_{n,i+1},\ 0\leqq t\leqq 1\}$, in Skorohod's space $D[0, 1]$, to the process $\left\{\int_0^t f(s, W(s))dW(s):\ 0\leqq t\leqq 1\right\}$ in $D[0, 1]$, which we denote by $\{Z(t),\ 0\leqq t\leqq 1\}$.

The results obtained are generalizations or extensions of those given in [1, Theorem 16.1], [3, p. 179], [2], [4] and [5].

2. Limit theorems. Suppose there exists a double array $\{C_{ni},\ 1\leqq i\leqq k_n,\ n\geqq 1\}$ of nonnegative numbers such that

(3) $$\sigma_{ni}^2 \leqq C_{ni}, \quad \text{a.s.} \quad 1 \leqq i \leqq k_n, \quad n \geqq 1.$$

and set

$$W_n^* = \sum_{i=0}^{m_n(t)} C_{ni}, \quad t\in[0, 1], \quad n \geqq 1 \quad (C_{n0} = 0).$$

The main result of this paper is given in the following

Theorem 1. *Let* $\{(X_{nk}, F_{nk}),\ 1\leqq k\leqq k_n\}$, $n\geqq 1$, *be a double array of random variables whose rows are martingale difference sequences such that* $s_n^2=1$, $n\geqq 1$. *Assume*

(4) *the finite dimensional distributions of* $\{W_n,\ n\geqq 1\}$ *converge weakly, as* $n\to\infty$, *to those of* $\{W(t),\ 0\leqq t\leqq 1\}$,

(5) *there exists an array of nonnegative numbers satisfying* (3) *such that for every* $t_1, t_2\in[0, 1]$, $t_2-t_1\geqq m(n)$, $n\geqq 1$,

$$W_n^*(t_2)-W_n^*(t_1) \leqq [F(t_2)-F(t_1)]^r,$$

where $m(n)=\min\{EX_{ni}^2:\ 1\leqq i\leqq k_n\}$, *F is a nondecreasing continuous function on* $[0, 1]$ *and* $r>1/2$ *is some positive constant.*

Then $Z_n\to Z$ *as* $n\to\infty$, *in* $D[0, 1]$, *provided that* $f, f_n\in F_M$, $n\geqq 1$, *and for every* $s\in[0, 1]$

(6) $$Df_n(s, x) \to Df(s, x), \quad as \quad n \to \infty$$

uniformly in x on every finite interval. Here, the stochastic integral in the definition of $Z(t)$ *is taken in the* L^2*-sense.*

From Theorem 1 we get the following

Theorem 2. *Assume* $\{(X_i, F_i),\ i\geqq 1\}$ *is a square-integrable martingale difference sequence such that* $EX_i^2=1$, $i\geqq 1$, *and*

(7) $$\sup_i E(X_i^2\,|\,F_{i-1}) \leqq M, \quad a.s.$$

for some positive constant M. If (4) *holds with* $W_n(t)=\sum_{i=0}^{[nt]} X_i/\sqrt{n}$ *then, in* $D[0, 1]$,

$$(8) \qquad (1/\sqrt{n}) \sum_{i=0}^{[nt]-1} f_n(i/n, S_i/\sqrt{n}) X_{i+1} \xrightarrow{\mathcal{D}} \int_0^t f(s, W(s))\, dW(s), \quad \textit{as} \quad n \to \infty,$$

provided (6) *holds as well.*

To prove Theorem 2 we note that, in this case, (5) is satisfied with $X_{nk}=(X_1+\dots+X_k)/\sqrt{n}$, $C_{nk}=M/n$, $1 \leqq k \leqq k_n=n$, $F(t)=2t$, $r=1$, $m_n(t)=[nt]$. Thus Theorem 2 follows from Theorem 1. It is easy to see that Theorem 16.1 in [1] is a consequence of Theorem 2 (it is enough to put $f_n=f\equiv 1$, $n \geqq 1$).

We note that a necessary and sufficient condition for (4) to hold is given in Theorem 7.7 [1, p. 49]. Furthermore, if $W_n=\{W_n(t)\colon 0 \leqq t \leqq 1\}$ converges weakly, in $D[0, 1]$, to a standard Wiener process $W=\{W(t), 0 \leqq t \leqq 1\}$, then (4) also holds. On the other hand, the assertion of Theorem 1 implies the weak convergence of W_n, as $n \to \infty$, to W. Thus the assumption (4) is necessary for (6) to hold. For example, it is well known that if $\{(X_{ni}, F_{ni}), 1 \leqq i \leqq k_n\}$, $n \geqq 1$, is a double array of square-integrable random variables whose rows are martingale difference sequences satisfying the Lindeberg condition and $\sum_{i=1}^{k_n} \sigma_{ni}^2 \xrightarrow{P} 1$, then (4) holds. Moreover, one can easily observe that every sequence $\{X_n, n \geqq 1\}$ of independent random variables, with $EX_n=0$, $EX_n^2=1$, $n \geqq 1$, satisfying the central limit theorem also satisfies the assumptions of Theorem 2. It should also be mentioned here that the assumptions (1) and (6) concerning the functions f_n, $n \geqq 1$, and f are very general. Some examples of such functions can be found in [3, Section 5].

To give a better illustration of the meaning of Theorem 1, let us note that from a very special case of it we immediately obtain the following assertions. If $\{(X_i, F_i), i \geqq 1\}$ is a sequence of random variables with $EX_i=1$, $i \geqq 1$, and satisfy (4) and (7), then in $D[0, 1]$,

$$\Big\{n^{-1} \sum_{1 \leqq i<j \leqq [nt]} X_i X_j,\ 0 \leqq t \leqq 1\Big\} \xrightarrow{\mathcal{D}} \Big\{\int_0^t W(s)\, dW(s),\ 0 \leqq t \leqq 1\Big\}$$

and

$$\Big\{n^{-3/2} \sum_{i=1}^{[nt]} (i-1) X_i,\ 0 \leqq t \leqq 1\Big\} \xrightarrow{\mathcal{D}} \Big\{\int_0^t s\, dW(s),\ 0 \leqq t \leqq 1\Big\}$$

as $n \to \infty$. The first assertion follows from Theorem 2 with $f_n(t, x)=f(t, x)=x$, and the second one with $f_n(t, x)=f(t, x)=t$. The distributions of the integrals

$$\int_0^t W(s)\, dW(s) \quad \text{and} \quad \int_0^t s\, dW(s)$$

are well known. For example,

$$\int_0^t W(s)\,dW(s) = (W^2(t)-t)/2 .$$

Remark. We note that condition (5) implies

(9) $$W_n^*(1) \leqq K,\ n \geqq 1, \quad \text{for some constant } K.$$

Moreover, by (5),

$$\max_{1\leqq i\leqq k_n} C_{ni} \leqq \sup\{[F(t_2)-F(t_1)]^r\colon\ t_2-t_1 = m(n)\},\ n \geqq 1,$$

and, by (3), $EX_{ni}^2 \leqq C_{ni}$, and $\lim_{n\to\infty} m(n) = 0$, so that

(10) $$\max_{1\leqq i\leqq k_n} EX_{ni}^2 \to 0, \quad \text{as} \quad n \to \infty,$$

because the function F is uniformly continuous.

3. Auxiliary lemmas. Let for every function $f\in F_M$

$$f^C(s, x) = f(s, x)I([-C, C])(x), \quad s\in[0, 1],$$

where C is a positive constant and $I(A)(\cdot)$ denotes the indicator function of the set A, and set

$$\|(x, y)\|_2 = (x^2+y^2)^{1/2}, \quad (x, y)\in R^2.$$

Lemma 1. *Let* $\{f_n,\ n\geqq 1\}$ *be a sequence of functions such that* $f_n\in F_M$, $n\geqq 1$, *and let* $0=p_0<p_1<\ldots<p_r=t$, $t=t_0<t_1<\ldots<t_b=s$, $0\leqq t<s\leqq 1$, *be partitions of the intervals* $[0, t]$ *and* $[t, s]$, *respectively. Assume that for each n the* MDS $\{(X_{ni}, F_{ni}),\ 1\leqq i\leqq k_n\}$ *satisfies the assumptions of Theorem* 1. *Then, for every* $\varepsilon>0$ *and each* $C>0$,

(11) $$\lim_{\gamma\to 0}\lim_{n\to\infty} P_1(\varepsilon, \gamma, n, C) = 0,$$

where

$$\gamma = \max_{1\leqq i\leqq r}(p_i-p_{i-1})+\max_{1\leqq i\leqq b}(t_i-t_{i-1})$$

and

$$P_1(\varepsilon, \gamma, n, C) = P\Big(\Big\|\Big(\sum_{i=0}^{m_n(t)} f_n^C(s_{ni}^2, S_{ni})X_{n,i+1}-\sum_{j=0}^{r-1} f_n^C(p_j, W_n(p_j))(W_n(p_{j+1})-W_n(p_j)),$$

$$\sum_{i=m_n(t)+1}^{m_n(s)} f_n^C(s_{ni}^2, S_{ni})X_{n,i+1}-\sum_{j=0}^{b-1} f_n^C(t_j, W_n(t_j))(W_n(t_{j+1})-W_n(t_j))\Big)\Big\|_2 > \varepsilon\Big) =$$

$$= P(\|(X(n, \gamma, 0, t), X(n, \gamma, t, s))\|_2 > \varepsilon) = P(\|(X_1, X_2)\|_2 > \varepsilon).$$

Proof. To prove Lemma 1 it is enough to show that

(12) $$\lim_{\gamma\to 0}\lim_{n\to\infty} EX^2(n, \gamma; t, s) = 0,$$

because, in the same way, we can prove that (12) holds with $X(n, \gamma, 0, t)$ and then $P_1(\varepsilon, \gamma, n, C) \leqq \varepsilon^{-2}(EX_1^2 + EX_2^2)$.

Let, for every i $(m_n(t_j) < i \leqq m_n(t_{j+1}))$, $W_{ij} = W_{ij}^{(n)} = f_n^C(s_{ni}^2, S_{ni}) - f_n^C(t_j, W_n(t_j))$. Then we have

$$P_1(\varepsilon, \gamma, n, C) = P\Big(\Big|\sum_{j=0}^{b-1} \sum_{i=m_n(t_j)+1}^{m_n(t_{j+1})} W_{ij} X_{n,i+1}\Big| > \varepsilon\Big).$$

On the other hand, for every $i < i'$ $(m_n(t_j) \leqq i < m_n(t_{j+1}),\ m_n(t_{j'}) \leqq i' < m_n(t_{j'+1}))$

$$EW_{ij} X_{n,i+1} W_{i'j'} X_{n,i'+1} = EW_{ij} X_{n,i+1} W_{i'j'} E(X_{n,i'+1} | F_{ni'}) = 0.$$

Thus

$$EX_2^2 = \sum_{j=0}^{b-1} \sum_{i=m_n(t_j)+1}^{m_n(t_{j+1})} EW_{i-1,j}^2 X_{ni}^2 = \sum_{j=0}^{b-1} \sum_{i=m_n(t_j)+1}^{m_n(t_{j+1})} EW_{i-1,j}^2 E(X_{ni}^2 | F_{n,i-1}) \leqq$$

$$\leqq \sup_{i,j} EW_{ij}^2 \Big(\sum_{i=m_n(t)}^{m_n(s)} C_{ni}\Big) \leqq W_n^*(1) \sup_{i,j} EW_{ij}^2.$$

Hence, by (9),

$$EX_2^2 \leqq K \sup_{i,j} EW_{ij}^2. \tag{13}$$

Let us observe that, by (1), for every $f \in F_M$ and (s, x), $(s_1, x_1) \in [0, 1] \times R$,

$$|f^C(s, x) - f^C(s_1, x_1)| \leqq K_C(|s - s_1| + |x - x_1|), \tag{14}$$

where K_C is an absolute positive constant which depends only on C. Thus, for every $m_n(t_j) < i \leqq m_n(t_{j+1})$, $0 \leqq j \leqq b-1$,

$$\begin{aligned} EW_{ij}^2 &\leqq 2K_C^2\{|s_{ni}^2 - t_j|^2 + E(S_{ni} - S_{nm_n(t_j)})^2\} \leqq \\ &\leqq 2K_C^2\{(t_{j+1} - t_j + \max_{1 \leqq i \leqq k_n} EX_{ni}^2)^2 + (t_{j+1} - t_j + \max_{1 \leqq i \leqq k_n} EX_{ni}^2)\}. \end{aligned} \tag{15}$$

Taking into account (10) and (15) we obtain (12).

Lemma 2. *Let f, f_n, $n \geqq 1$, be functions satisfying the assumptions of Theorem* 1. *If the assumptions of Lemma* 1 *are also satisfied, then for every* $C > 0$

$$\lim_{\gamma \to 0} \lim_{n \to \infty} P_2(\varepsilon, \gamma, n, C) = 0, \tag{16}$$

where

$$P_2(\varepsilon, \gamma, n, C) = P\Big(\Big\|\Big(\sum_{j=0}^{r-1} \{f_n^C(p_j, W_n(p_j)) - f^C(p_j, W_n(p_j))\}(W_n(p_{j+1}) - W_n(p_j)),$$

$$\sum_{j=0}^{b-1} \{f_n^C(t_j, W_n(t_j)) - f^C(t_j, W_n(t_j))\}(W_n(t_{j+1}) - W_n(t_j))\Big)\Big\|_2 > \varepsilon\Big) =$$

$$= P(\|(X_1', X_2')\|_2 > \varepsilon), \quad \textit{say}.$$

Proof. Again, it is enough to show that

$$\lim_{\gamma\to 0}\lim_{n\to\infty} E(X_2')^2 = 0.$$

Let, for every $0\leqq j\leqq b$, $V_{nj}(x)=f_n^C(t_j, x)-f^C(t_j, x)$. We have

$$E(X_1')^2 = \sum_{j=0}^{b-1} E\{(V_{nj}(W_n(t_j)))^2 \sum_{i=m_n(t_j)+1}^{m_n(t_{j+1})} E(X_{ni}^2 | F_{nm_n}(t_j))\}.$$

Let $R_n = \max_{0\leqq j\leqq b-1} \sup_x V_{nj}^2(x)$. Then, by (6), $R_n\to 0$ as $n\to\infty$. Thus

$$EX_1'^2 \leqq R_n \sum_{j=0}^{b-1} \sum_{i=m_n(t_j)+1}^{m_n(t_{j+1})} EX_{ni}^2 = R_n(EW_n^2(s)-EW_n^2(t)) \leqq R_n \to 0, \quad \text{as} \quad n\to\infty.$$

Lemma 3. *Let the assumptions of Lemma* 1 *and Theorem* 1 *be satisfied. Then for any given* $C>0$,

$$\Big(\sum_{j=0}^{r-1} f^C(p_j, W_n(p_j))(W_n(p_{j+1})-W_n(p_j)),\ \sum_{j=0}^{b-1} f^C(t_j, W_n(t_j))(W_n(t_{j+1})-W_n(t_j))\Big) \xrightarrow{\mathcal{D}}$$

(17)

$$\xrightarrow{\mathcal{D}} \Big(\sum_{j=0}^{r-1} f^C(p_j, W(p_j))(W(p_{j+1})-W(p_j)),\ \sum_{j=0}^{b-1} f^C(t_j, W(t_j))(W(t_{j+1})-W(t_j))\Big)$$
$$\text{as} \quad n\to\infty,$$

where $\{W(t)\colon 0\leqq t\leqq 1\}$ *is a standard Wiener process in* $D[0, 1]$.

The assertion of Lemma 3 follows from (4) and Theorem 5.1 [1].

Lemma 4. *If* $f\in F_M$, *then for every* $\varepsilon>0$ *and any given* $C>0$

$$P\Big(\Big\|\Big(\sum_{j=0}^{r-1} f^C(p_j, W(p_j))(W(p_{j+1})-W(p_j)) - \int_0^t f^C(x, W(x))\,dW(x),$$

(18)

$$\sum_{j=0}^{b-1} f^C(t_j, W(t_j))(W(t_{j+1})-W(t_j)) - \int_t^s f^C(x, W(x))\,dW(x)\Big\|_2 > \varepsilon\Big) \to 0$$

as

$$\gamma = \max_{0\leqq i\leqq r-1}(p_{i+1}-p_i) + \max_{0\leqq j\leqq t-1}(t_{j+1}-t_j) \to 0,$$

where $0=p_0<p_1<\ldots<p_r=t$, $t=t_0<t_1<\ldots<t_b=s$, $0\leqq t<s\leqq 1$ *are partitions of the intervals* $[0, t]$ *and* $[t, s]$, *respectively.*

The proof of Lemma 4 is essentialy the same that is given in [4].

4. Proof of Theorem 1. Let us observe that

(19) $$P(\max_{1\leqq i\leqq k_n} |S_{ni}| > C) \leqq C^{-2}.$$

Furthermore

(20) $$P(\sup_{0\leqq t\leqq 1}|W(t)|>C)\to 0 \quad \text{as} \quad C\to\infty.$$

Thus, taking into account (11) and (16)—(20) we get

(21) $$(Z_n(t), Z_n(s)-Z_n(t))\xrightarrow{\mathscr{D}}(Z(t), Z(s)-Z(t)) \quad \text{as} \quad n\to\infty,$$

for every $0\leqq t<s\leqq 1$. Clearly, using this method we may prove that the finite dimensional distributions of $\{Z_n, n\geqq 1\}$ converge weakly, as $n\to\infty$, to those of $\{Z(t): 0\leqq t\leqq 1\}$.

To complete the proof, we have to verify the tightness condition. We use Theorem 15.6 in [1]. From this theorem and (19) we infer that it suffices to show

(22) $$E(Z_n^C(t)-Z_n^C(t_1))^2(Z_n^C(t_2)-Z_n^C(t))^2\leqq [F(t_2)-F(t_1)]^{2r},$$

for any $t_1\leqq t\leqq t_2$, $n\geqq 1$, $C>0$, where

$$Z_n^C(t)=\sum_{i=0}^{m_n(t)-1} f_n^C(s_{ni}^2, S_{ni})X_{n,i+1}, \quad t\in[0, 1].$$

We first note that, by (3) and (1),

$$E(Z_n^C(t)-Z_n^C(t_1))^2(Z_n^C(t_2)-Z_n^C(t))^2\leqq$$
$$\leqq K(C)(W_n^*(t)-W_n^*(t_1))(W_n^*(t_2)-W_n^*(t))\leqq 4^{-1}K(C)(W_n^*(t_2)-W_n^*(t_1))^2,$$

where $K(C)$ is some positive constant which depends only on C. Hence, by assumption (5) condition (22) holds, because in the case $t_2-t_1<m(n)$, $Z_n^C(t)=Z_n^C(t_1)$ or $Z_n^C(t)=Z_n^C(t_2)$.

Acknowledgement. I would like to thank the referee for helpful comments.

References

[1] P. Billingsley, *Convergence of Probability Measures*, Wiley (New York, 1968).

[2] Z. Rychlik, I. Szyszkowski, Weak convergence of functionals of sums for martingale differences, *Bull. Acad. Polon. Sci., Ser. Sci. Math. Astronom. Phys.*, 31 (1983), 407—413.

[3] A. B. Skorohod, N. P. Slobodenyuk, *Limit theorems for random walks, Nauka* (Kiev, 1970). (Russian)

[4] K. Yoshihara, A weak convergence theorem for functionals of sums of martingale differences, *Yokohama Math. J.*, **26** (1978), 101—107.

[5] K. Yoshihara, A weak convergence theorem for functionals of sums of independent random variables, *Stochastic Process. Appl.*, **12** (1982), 293—299.

INSTITUTE OF MATHEMATICS
MARIA CURIE-SKŁODOWSKA UNIVERSITY
UL. NOWOTKI 10
20-031 LUBLIN, POLAND

Acta Sci. Math., 52 (1988), 165—170

On the group of analytic automorphisms of the unit ball of J^*-algebras

JOSÉ M. ISIDRO(*)

1. Introduction

Our purpose is to present an elementary method to integrate a certain Riccati differential equation that plays an important role in the study of the unit ball of J^*-algebras of operators as symmetric spaces. Our approach consists in the use of Potapov's generalized Möbius transformations together with some elementary facts in the theory of holomorphic functions between Banach spaces. These methods have proved to be successful in the study of J^*-algebras and in some other questions, too ([1], [2], [3]).

Let $\mathfrak{H}$ and $\mathfrak{K}$ be complex Hilbert spaces and denote by $\mathcal{U}$ a J^*-algebra of bounded linear operators $X\colon \mathfrak{H}\to\mathfrak{K}$. That is, by definition, $\mathcal{U}$ is a closed complex subspace of $\mathcal{L}(\mathfrak{H}, \mathfrak{K})$ such that $AB^*C+CB^*A\in\mathcal{U}$ whenever $A, B, C\in\mathcal{U}$. Let $B(\mathcal{U})$ be the open unit ball of $\mathcal{U}$ and assume that $A\in\mathcal{U}$ and $X\in B(\mathcal{U})$ are given. Then, we consider the Riccati initial value problem:

$$(*)\qquad \frac{d}{dt}y(t) = A-y(t)A^*y(t), \quad y(0) = X, \quad y(t)\in B(\mathcal{U}),$$

where A^* stands for the adjoint of A. We give an explicit formula for the maximal solution $y_A(t; X)$ of $(*)$ in terms of the initial value X and the parameter A. See also ([3], page 57) and ([4], page 509) where other (but non elementary) approaches to the problem can be found.

We recall the following principal property of J^*-algebras [1]:

Given $M\in B(\mathcal{U})$, the Möbius transformation

$$(1)\qquad T_M(X) = (1-MM^*)^{-1/2}(X+M)(1+M^*X)^{-1}(1-M^*M)^{1/2}, \quad X\in B(\mathcal{U})$$

Received November 2, 1984 and in revised form September 24, 1985.

(*) Supported by CAICYT, contract 1525/82.

is a holomorphic automorphism of $B(\mathcal{U})$. Moreover, we have

$$(2) \qquad T_M(0) = M, \quad T_{-M} = T_M^{-1}, \quad T_M(X)^* = T_{M^*}(X^*)$$

and

$$(3) \qquad dT_M(X)Y = (1-MM^*)^{1/2}(1+XM^*)^{-1}Y(1+M^*X)^{-1}(1-M^*M)^{1/2}$$

for $X \in B(\mathcal{U})$ and $Y \in \mathcal{U}$.

Here, positive and negative square roots are defined by the usual power series expansions and, at each ocurrence, 1 denotes the identity operator on the appropiate underlying Hilbert space.

Furthermore, we recall from [5] or [8] the following basic facts concerning $\operatorname{Aut} B(\mathcal{U})$, the group of holomorphic automorphisms of $B(\mathcal{U})$:

Let the vector field $f(X)\dfrac{\partial}{\partial X}$ be complete in $B(\mathcal{U})$ and denote by $y(t, X)$ the solution of

$$(4) \qquad \frac{d}{dt}y(t) = f[y(t)], \quad y(0) = X, \quad y(t) \in B(\mathcal{U})$$

Then, for each fixed $t \in \mathbf{R}$, the mapping $X \to y(t, X)$ is an element of $\operatorname{Aut} B(\mathcal{U})$. Moreover, the mapping $t \mapsto y(t, \cdot)$ is a continuous one-parameter group of automorphisms of $B(\mathcal{U})$ and we have $f(X) = \dfrac{d}{dt}\Big|_0 y(t, X)$ for $X \in B(\mathcal{U})$.

2. The main result: one-parameter groups

Let us fix arbitrarily any operator $A \in \mathcal{U}$. By the polar decomposition [6], there is a partial isometry $W \in \mathcal{L}(\mathfrak{H}, \mathfrak{K})$ such that $A = WP$ where $P =: (A^*A)^{1/2}$ and $E =: W^*W$ is a projector onto the closure of the range of P. Let $\operatorname{tgh}(t) = \sum_{n=0}^{\infty} a_{2n+1} t^{2n+1}$, $t \in \mathbf{R}$, be the power series expansion of the function hyperbolic tangent tgh and define

$$\operatorname{tgh}(tP) =: \sum_{n=0}^{\infty} a_{2n+1}(tP)^{2n+1}, \quad t \in \mathbf{R}.$$

Then $\operatorname{tgh}(tP) \in \mathcal{L}(\mathfrak{H})$ and $\|\operatorname{tgh}(tP)\| \leqq \operatorname{tgh}\|tP\| < 1$ for all $t \in \mathbf{R}$.

2.1. Proposition. *For $t \in \mathbf{R}$, the operator $F(t) =: W \operatorname{tgh}(tP)$ satisfies $F(t) \in B(\mathcal{U})$ and the mapping $t \mapsto F(t)$ is continuous.*

Proof. One has

$$F(t)=W\,\mathrm{tgh}\,(tP)=W\sum_{n=0}^{\infty}a_{2n+1}(tP)^{2n+1}=$$

$$=\sum_{n=0}^{\infty}a_{2n+1}t^{2n+1}WP(A^*A)^n=\sum_{n=0}^{\infty}a_{2n+1}t^{2n+1}A(A^*A)^n\in\mathcal{U}.$$

Moreover, $\|F(t)\|\leqq\|W\|\cdot\|\mathrm{tgh}\,(tP)\|<1$ so that $F(t)\in B(\mathcal{U})$. Obviously, $t\mapsto F(t)$ is continuous.

2.2. Proposition. *Let the operators M, $N\in B(\mathcal{U})$ be given with*

(5) $$MN^*=NM^*,\quad M^*N=N^*M.$$

Then we have $T_M\circ T_N=T_{T_M(N)}$.

Proof. By Cartan's uniqueness theorem, it suffices to show that the automorphisms $T_M\circ T_N$ and $T_{T_M(N)}$ have the same image and the same derivative at the origin 0.

From (2) we obtain $(T_M\circ T_N)0=T_M(N)=T_{T_M(N)}(0)$. On the other hand, from (3) we get

(6) $$dT_{T_M(N)}(0)X=\big(1-T_M(N)T_M(N)^*\big)^{1/2}X\big(1-T_M(N)^*T_M(N)\big)^{1/2}$$

where, by ([1], p. 22)

(7) $$1-T_M(N)^*T_M(N)=$$
$$=(1-M^*M)^{1/2}(1+N^*M)^{-1}(1-N^*N)(1+M^*N)^{-1}(1-M^*M)^{1/2}.$$

Using (2) together with (7) we obtain

(7′) $$1-T_M(N)T_M(N)^*=1-T_{M^*}(N^*)^*T_{M^*}(N^*)=$$
$$=(1-MM^*)^{1/2}(1+NM^*)^{-1}(1-NN^*)(1+NM^*)^{-1}(1-MM^*)^{1/2}.$$

Now, from the assumption (5) we see that the operators MM^*, NM^* and NN^* commute; thus the operators $(1-MM^*)^{1/2}$, $(1+NM^*)^{-1}$ and $(1-NN^*)$ also commute and (7′) yields

$$1-T_M(N)T_M(N)^*=(1-MM^*)(1+NM^*)^{-2}(1-NN^*),$$

whence

$$[1-T_M(N)T_M(N)^*]^{1/2}=(1-MM^*)^{1/2}(1+NM^*)^{-1}(1-NN^*)^{1/2}.$$

In a similar manner

$$[1-T_M(N)^*T_M(N)]^{1/2}=(1-N^*N)^{1/2}(1+MN^*)^{-1}(1-M^*M)^{1/2}.$$

Substitution in (6) gives

$$(8) \qquad dT_{T_M(N)}(0)X =$$
$$= (1-MM^*)^{1/2}(1+NM^*)^{-1}(1-NN^*)^{1/2}X(1-N^*N)^{1/2}(1+MN^*)^{-1}(1-M^*M)^{1/2}.$$

By the chain rule and (3) we have

$$dT_M\circ T_N(0)X = dT_M(N)\circ dT_N(0)X =$$
$$= (1-MM^*)^{1/2}(1+NM^*)^{-1}(1-NN^*)^{1/2}X(1-N^*N)^{1/2}(1+MN^*)^{-1}(1-M^*N)^{1/2}$$

which is the same as (8).

Let us fix any $A\in\mathcal{U}$ and consider the operator

$$F(t) =: W\operatorname{tgh}(tP)\in B(\mathcal{U}), \quad t\in\mathbf{R}.$$

2.3. Proposition. *The mapping* $\mathbf{R}\to\operatorname{Aut}B(\mathcal{U})$ *given by* $t\mapsto T_{F(t)}$ *is a continuous one-parameter group of Möbius transformations.*

Proof. Since the mappings $\mathbf{R}\to B(\mathcal{U})$ and $B(\mathcal{U})\to\operatorname{Aut}B(\mathcal{U})$ given respectively by $t\mapsto F(t)$ and $M\mapsto T_M$ are continuous, so is the composite.

Obviously, we have $T_{F(0)}=\mathrm{id}_{B(\mathcal{U})}$. Let us fix $s, t\in\mathbf{R}$ arbitrarily. As $E=W^*W$ is a projector onto the (closure of) the range of P, the operators $M=:F(s)$ and $N=:F(t)$ satisfy

$$MN^* = W\operatorname{tgh}(sP)\operatorname{tgh}(tP)W^* = W\operatorname{tgh}(tP)\operatorname{tgh}(sP)W^* = NM^*,$$
$$M^*N = \operatorname{tgh}(sP)W^*W\operatorname{tgh}(tP) = \operatorname{tgh}(sP)\operatorname{tgh}(tP) = \operatorname{tgh}(tP)\operatorname{tgh}(sP) =$$
$$= \operatorname{tgh}(tP)W^*W\operatorname{tgh}(sP) = N^*M,$$

and we can apply Proposition 2.2. Therefore

$$T_{F(s)}\circ T_{F(t)} = T_M\circ T_N = T_{T_M(N)}$$

and, in order to obtain the result, it suffices to show that

$$T_M(N) = W\operatorname{tgh}(s+t)P.$$

By the spectral calculus we have

$$(N+M)(1+M^*N)^{-1} = W(\operatorname{tgh}tP+\operatorname{tgh}sP)(1+\operatorname{tgh}tP\operatorname{tgh}sP)^{-1} =$$
$$= W\operatorname{tgh}(s+t)P.$$

Since the operator $\operatorname{tgh}(s+t)P$ obviously commutes with $(1-M^*M)^{1/2}=$

$=(1-\mathrm{tgh}^2\, tP)^{1/2}$, we have

$$T_M(N) = (1-MM^*)^{1/2}(N+M)(1+M^*N)^{-1}(1-M^*M)^{1/2} = \tag{9}$$
$$= (1-MM^*)^{1/2}[W\,\mathrm{tgh}\,(s+t)P](1-M^*M)^{1/2} =$$
$$= (1-MM^*)^{-1/2}W(1-M^*M)^{1/2}\,\mathrm{tgh}\,(s+t)P.$$

As W is a partial isometry, we have $WE=WW^*W=W$ and, as E is a projector onto the range of P, $EP=P=PE$. Therefore

$$E\,\mathrm{tgh}\,(sP) = \mathrm{tgh}\,(sP) = (\mathrm{tgh}\, sP)E.$$

Let us set $Q=:\mathrm{tgh}\,(sP)$. Then

$$1-MM^* = 1-WQ^2W^*;\quad 1-M^*M = 1-Q^2$$

and

$$(1-MM^*)^{1/2}W(1-M^*M)^{1/2} = (1-WQ^2W^*)^{-1/2}W(1-Q^2)^{1/2} =$$
$$= \left[\sum_{n=0}^{\infty}(-1)^n\binom{-1/2}{n}(WQ^2W^*)^n\right]W(1-Q^2)^{1/2} =$$
$$= \left[\sum_{n=0}^{\infty}(-1)^n\binom{-1/2}{n}WQ^{2n}E\right](1-Q^2)^{1/2} =$$
$$= \left[\sum_{n=0}^{\infty}(-1)^n\binom{-1/2}{n}WEQ^{2n}\right](1-Q^2)^{1/2} =$$
$$= WE(1-Q^2)^{-1/2}(1-Q^2)^{1/2} = W.$$

Substitution in (9) gives the result.

2.4. Theorem. *Let $\mathscr{U}$ be any J^*-algebra. Let $A\in\mathscr{U}$ be arbitrarily given and write $P=:(A^*A)^{1/2}$, $A=WP$ and $F(t)=:W\,\mathrm{tgh}\,(tP)$ for $t\in\mathbf{R}$. Then, the mapping $\mathbf{R}\to\mathrm{Aut}\,B(\mathscr{U})$ given by $t\mapsto T_{F(t)}$ is a continuous one-parameter group of automorphisms of $B(\mathscr{U})$ whose associated vector field is $f_A(X)\frac{\partial}{\partial X}=(A-XA^*X)\frac{\partial}{\partial X}$.*

Proof. By Proposition 2.3, the mapping $t\mapsto T_{F(t)}$ is a continuous one-parameter group of Möbius transformations. Therefore, the mapping

$$X\mapsto \frac{d}{dt}\Big|_0 T_{F(t)}(X),\quad X\in B(\mathscr{U})$$

is a holomorphic vector field that is complete in $B(\mathscr{U})$. Now an easy calculation gives

$$\frac{d}{dt}\Big|_0 T_{F(t)}(X) = A-XPW^*X = A-X(WP)^*X = A-XA^*X.$$

2.5. Corollary. *If* $\mathcal{U}$ *is any* J^**-algebra and* $A\in\mathcal{U}$, *then* $t\mapsto T_{F(t)}$ *is the maximal solution of the initial value problem*

$$\frac{d}{dt}y(t) = A - y(t)A^*y(t), \quad y(0) = X, \quad Y(t)\in B(\mathcal{U})$$

for $X\in B(\mathcal{U})$.

References

[1] L. A. Harris, Bounded symmetric homogeneous domains in infinite dimensional spaces, *Lecture Notes in Math.* **364**, Springer-Verlag (Berlin, 1973), 666—674.

[2] L. A. Harris, Banach algebras with involution and Möbius transformations, *J. Funct. Anal.*, **11** (1972), 1—16.

[3] W. Kaup, Algebraic characterization of symmetric complex Banach manifolds, *Math. Ann.*, **288** (1977), 39—64.

[4] W. Kaup, A Riemann mapping theorem for bounded symmetric domains in complex Banach spaces, *Math. Z.*, **183** (1983), 503—529.

[5] H. Upmeier, Über die Automorphismen-gruppen beschränkter Gebiete in Banach-räumen, *Dissertation* (Tübingen, 1975).

[6] M. A. Naimark, *Normed rings*, Noordhoff (Groningen, 1964).

[7] V. P. Potapov, The multiplicative structure of J-contractive matrix functions, *Amer. Math. Soc. Transl.*, **15** (1960), 131—243.

[8] J. P. Vigue, Le group des automorphismes analytiques d'un espace de Banach complexe, *Ann. Sci. École Norm. Sup.* (4), **9** (1976), 203—282.

FACULTAD DE MATEMÁTICAS
UNIVERSIDAD DE SANTIAGO
SANTIAGO DE COMPOSTELA, SPAIN

Acta Sci. Math., 52 (1988), 171—178

Sur un théorème de J. Bourgain

C. FINET

Introduction

J. Bourgain a démontré (voir [1] et [1']) que le dual de l'algèbre du disque a la propriété de Grothendieck. Nous étendons ce résultat aux algèbres séparables uniformes A sur un compact X telles que :

(1) Il existe une suite (φ_n) dans le spectre de A telle que pour tout n, φ_n ait une unique mesure représentante μ_n sur X (les mesures μ_n étant mutuellement singulières);

(2) Il n'y a pas de mesure non nulle dans $A^\perp$ qui soit orthogonale à toutes les mesures μ_n. (Où $A^\perp=\{\mu\in\mathscr{C}(X)' : \mu|_A=0\}$).

L'article comprend trois parties. La première consiste en des rappels notamment sur les algèbres uniformes. Nous nous intéressons à l'opérateur de conjugaison dans le cadre des algèbres dites « ω^*-de Dirichlet ». Dans la deuxième partie, nous établissons quelques propriétés de la projection de Riesz d'où nous tirons des conséquences analogues à celles obtenues par J. Bourgain dans [1]. La dernière partie de ce travail est consacrée à la démonstration du résultat principal et à quelques exemples et problèmes.

I. Rappels

Soient X, Y deux espaces de Banach. Notons $B(X, Y)$ l'espace des opérateurs linéaires bornés de X dans Y, $\Pi_p(X, Y)$ l'espace des opérateurs p-sommants de X dans Y et X' le dual de X. Rappelons la définition des algèbres « ω^*-de Dirichlet » (voir [2]).

Définition 1. Soit (X, μ) un espace de probabilité. On dit que A est une algèbre « ω^*-de Dirichlet » si A est une sousalgèbre de $L^\infty(\mu)$, contenant les constantes, telle que la mesure μ soit multiplicative sur l'algèbre A et $A+\bar{A}$ préfaiblement dense dans $L^\infty(\mu)$.

Reçu le 3 janvier, 1985.

Notons que si A est une algèbre uniforme sur X et μ l'unique mesure représentante de φ dans le spectre de A, alors A est une algèbre « ω^*-de Dirichlet ». Pour $p\in[1, +\infty[$, l'espace $H^p(\mu)$ est la fermeture de A dans $L^p(\mu)$, et $H^\infty(\mu)$ est la fermeture préfaible de A dans $L^\infty(\mu)$. Sur $L^p(\mu)$ $(1<p<\infty)$, on définit l'opérateur de conjugaison (noté « $\sim$ ») qui possède les principales propriétés de la transformée de Hilbert (« classique ») (voir [2]) : on commence par définir la conjuguée $\tilde{f}$ pour une fonction f dans $A+\bar{A}$: f s'écrit alors de façon unique : $f=f_1+C+\bar{f}_2$ avec $C\in\mathbf{C}$ et $f_1, f_2\in A_0$ où $A_0=\{f\in A, \int f\,d\mu=0\}$). On pose $\tilde{f}=i\bar{f}_2-if_1$.

Définition 2. Pour chaque $p\in]1, \infty[$, il existe dans $L^p(\mu)$ un unique opérateur continu, que l'on appellera « opérateur de conjugaison », (et que l'on notera « $\sim$ »), qui coincide avec l'application « $\sim$ » définie précédemment sur $A+\bar{A}$.

Rappelons les principales propriétés de *l'opérateur de conjugaison :*

1. Si $1<p<\infty$, pour tout $f\in L^p(\mu)$, $f+i\tilde{f}\in H^p(\mu)$.
2. If existe une constante M telle que si $p\in]1, \infty[$ et $f\in L^p(\mu)$, on ait :

$$\|\tilde{f}\|_p \leqq Mpp'\|f\|_p \quad (\text{où } p' = (p/p-1)).$$

3. Si $f\in \operatorname{Re} L^\infty(\mu)$, alors $\exp t(f+i\tilde{f})\in H^\infty(\mu)$, pour tout $t\in\mathbf{R}$.

De plus, il existe ([2]) un unique opérateur « $\sim$ » défini sur $L^1(\mu)$ qui coincide avec « $\sim$ » définie sur $L^p(\mu)$ $(p>1)$ et de type faible (1—1).

II. Lemmes preparatoires

Dans ce qui suit (X, μ) est un espace de probabilité et A une algèbre « ω^*-de Dirichlet ». Nous définissons la projection de Riesz (notée R). Soient $p\in]1, \infty[$, et $f\in L^p(\mu)$, on pose :

$$R(f) = 2^{-1}\left(f+i\tilde{f}+\int f\right).$$

Des propriétés de l'opérateur de conjugaison, on déduit le lemme suivant.

Lemme 3. (1) *Soient $1<p<\infty$ et $p'=p/(p-1)$, si $\|R\|_p$ désigne la norme de la projection de Riesz considérée comme un opérateur sur $L^p(\mu)$, alors il existe une constante C (indépendante de p) telle que*

$$\|R\|_p \leqq Cpp'.$$

(2) *La projection de Riesz est de type faible* (1—1).

Mentionnons quelques propriétés faciles qui nous seront utiles par la suite.

Lemme 4. (a) *Pour tout* $f \in L^p(\mu)$ $(1<p<\infty)$, $\tilde{\tilde{f}} = -f + \int f$.
(b) *La conjuguée d'une constante est la fonction nulle.*
(c) *Soit* $1<p<\infty$, *pour tout* $f \in H^p(\mu)$, $R(f)=f$.

On aura aussi besoin du

Lemme 5. *Pour tous* f *et* $g \in L^2(\mu)$, *on a*

$$\int fRg = \int gR_- f, \tag{1}$$

où R_- *désigne la projection de Riesz négative :*

$$R_-(f) = 2^{-1}\left(f - i\tilde{f} + \int f\right).$$

Démonstration. Prouver l'égalité (1) revient à démontrer :

$$\int f\tilde{g} + g\tilde{f} = 0.$$

Supposons d'abord f et g dans $A+\bar{A}$ et à valeurs réelles. On peut écrire :

$$f = f_1 + \int f + \bar{f}_1 = 2\operatorname{Re} f_1 + \int f, \quad f_1 \in A_0\,;$$
$$g = 2\operatorname{Re} g_1 + \int g, \quad g_1 \in A_0.$$

Dès lors, $\tilde{f} = 2\operatorname{Im} f_1$ et $\tilde{g} = 2\operatorname{Im} g_1$. Ainsi,

$$\int f\tilde{g} + g\tilde{f} = \int \left(2\operatorname{Re} f_1 + \int f\right) 2\operatorname{Im} g_1 + \left(2\operatorname{Re} g_1 + \int g\right) 2\operatorname{Im} f_1$$
$$= \int 4(\operatorname{Re} f_1 \operatorname{Im} g_1 + \operatorname{Re} g_1 \operatorname{Im} f_1) = 4\int \operatorname{Im}(f_1 g_1) = 0.$$

On obtient alors facilement le résultat souhaité pour $f, g \in A+\bar{A}$. Par densité de $A+\bar{A}$ dans $L^2(\mu)$, l'égalité (1) est vérifiée pour tous $f, g \in L^2(\mu)$.

De la propriété de type faible (1—1) de R, on déduit le

Lemme 6. *Si* $f \in L^1(\mu)$, $\omega \in L^\infty_+(\mu)$, $0<\alpha<1$, *on a :*

$$\int |R_-(f)|^\alpha \omega \leq C.(1-\alpha)^{-1} \|\omega\|_1^{1-\alpha} \|\omega\|_\infty^\alpha \|f\|_1^\alpha.$$

D'autre part, on a

Lemme 7. *Si* K *est un sous-ensemble mesurable de* X *et* $\varepsilon>0$, *alors il existe deux fonctions* φ *et* ψ *dans* $H^\infty(\mu)$ *telles que :*

(1) $|\varphi| + |\psi| \leq 1$;

(2) $|\varphi(z) - 1/5| \leq \varepsilon$ *pour* $z \in K$;

(3) $|\psi(z)| \leq \varepsilon$ *pour* $z \in K$;

(4) $\|\varphi\|_1 \leq C(\log \varepsilon^{-1})^2 \mu(K)$;

(5) $\|1-\psi\|_2 \leq C(\log \varepsilon^{-1}) \mu(K)^{1/2}$.

Démonstration. Soit $\tau=1-(1-\varepsilon)\chi_K$, $\log\tau=(\log\varepsilon)\chi_K\in\operatorname{Re}L^\infty(\mu)$. Donc $f=\exp(\log\tau+i\widetilde{\log\tau})\in H^\infty(\mu)$ (propriété 3 de l'opérateur de conjugaison). L'opérateur de conjugaison est borné en norme L^2, ainsi :

$$\|1-f\|_2\leqq\|1-\tau\|_2+\|\log\tau\|_2\leqq(1+\log\varepsilon^{-1})\mu(K)^{1/2}.$$

On choisit $\varphi=5^{-1}(1-f)^2\in H^\infty(\mu)$. Soient $G=1-|\varphi|$ et $g=G\exp(i\widetilde{\log G})\in H^\infty(\mu)$ puisque $\log G\in\operatorname{Re}L^\infty(\mu)$. On prend alors $\psi=f\cdot g\in H^\infty(\mu)$ et les cinq conditions sont remplies par les fonctions φ et ψ de $H^\infty(\mu)$.

Nous obtenons, dès lors, le même lemme de découpage que dans [1].

Lemme 8. *Il existe une constante C telle que pour toute fonction f dans* $L^1_+(\mu)$, $\int f=1$ *et* $0<\delta<1$, *il existe des scalaires positifs* (c_i) *et des suites* (θ_i), (τ_i) *de fonctions de* $H^\infty(\mu)$ *telles que :*

(1) $\|\theta_i\|_\infty\leqq C$;

(2) $\|\sum|\tau_i|\|_\infty\leqq C$;

(3) $|\tau_i|f\leqq c_i$ *p.s.* ;

(4) $\sum c_i\|\tau_i\|_1\leqq C\cdot\delta^{-C}$;

(5) $\int|1-\sum\theta_i\tau_i^2|f\leqq\delta.$

III. Le resultat principal

1. Nous étudions les opérateurs p-sommants définis sur des algèbres uniformes vérifiant les conditions 1, 2 citées dans l'introduction. Notons A une telle algèbre. Nous établissons le théorème de décomposition suivant :

Théorème 9. *Tout opérateur T p-sommant sur A* $(p\geqq1)$ *se décompose comme suit :* $T=T_1+T_2$ *où* T_1, T_2 *vérifient les propriétés suivantes :*

(1) $\pi_p(T_1)^p+\pi_p(T_2)^p\leqq\pi_p(T)^p$;

(2) *Il existe une suite d'opérateurs* $S_n:\mathscr{C}(X)\to A$ *telle que* (T_2S_n) *converge en norme* π_p *vers un opérateur* $\tilde T_2$ *vérifiant* $T_2=\tilde T_2\circ j$ *où j est l'injection canonique de A dans* $\mathscr{C}(X)$;

(3) *La première composante* T_1 *s'étend à* $H^\infty(m)$ *où m est une des mesures représentantes* μ_i.

Démonstration. 1) Comme T est p-sommant sur A $(p\geqq1)$, par le théorème de factorisation de Pietsch ([5]), il existe une mesure de probabilité μ sur X telle que pour tout $\varphi\in A$:

$$\|T(\varphi)\|\leqq\pi_p(T)\|\varphi\|_{L^p(\mu)}.$$

La mesure μ admet la décomposition de Lebesgue : $\mu=\mu_a+\mu_s$ avec $\mu_a\ll\alpha$ et $\mu_s\perp\alpha$ où

$\alpha=\sum_{n=1}^{\infty}2^{-n}\mu_n$. On a : $\mu_a=\sum_{n=1}^{\infty}h_n\mu_n$. Puisque μ_a n'est pas nul, on peut supposer $h_1\mu_1$ non nul. Soit $\sigma=\sum_{n=2}^{\infty}h_n\mu_n$. Il existe $L=\bigcup_{n\geqq 1}F_n$, F_n fermé tel que $\mu_1(L)=0=|\sigma+\mu_s|(X\setminus L)$. On applique alors le lemme de Forelli ([3]): il existe $(\psi_k)\subset A$ telle que $\|\psi_k\|_\infty\leqq 1$, (ψ_k) converge ponctuellement vers 0 dans L et (ψ_k) converge vers 1 (μ_1-presque partout). Dès lors, si on définit $T_1(\varphi)=\lim_{k\to\infty}T(\varphi\psi_k)$ et $T_2=T-T_1$, on obtient :

(2) $$\|T_1(\varphi)\|\leqq\pi_p(T)\|\varphi\|_{L^p(h_1\mu_1)},\quad \|T_2(\varphi)\|\leqq\pi_p(T)\|\varphi\|_{L^p(\sigma+\mu_s)}.$$

Et donc :

$$\pi_p(T_1)^p+\pi_p(T_2)^p\leqq\pi_p(T)^p.$$

2) On désire étendre l'opérateur T_2 à tout l'espace $\mathscr{C}(X)$. Soit alors $F=\bigcup_n K_n$, K_n compact, $K_n\subset K_{n+1}$ tel que $\mu_s(X\setminus F)=0=\alpha(F)$. Puisque A est séparable et $A^\perp\subset L^1(\alpha)$, on peut définir une extension linéaire préservant la norme : E_n : $\mathscr{C}(K_n)\to A$ ([4]). Soit l'opérateur restriction : R_n : $\mathscr{C}(X)\to\mathscr{C}(K_n)$. On pose $S_n=E_nR_n$. Et on montre alors facilement que la suite (T_2S_n) converge en norme π_p.

3) De l'inégalité (2), on déduit que T_1 s'étend en un opérateur sur $H^\infty(h_1\mu_1)$ et donc en un opérateur sur $H^\infty(\mu_1)$.

Le théorème est ainsi démontré.

2. Le résultat principal de notre travail est le

Théorème 10. *Si une algèbre séparable uniforme A sur un compact X vérifie les deux conditions suivantes :*

(1) Il existe une suite (φ_n) dans le spectre de A telle que pour tout n, φ_n ait une unique mesure représentante μ_n sur X (les mesures μ_n sont mutuellement singulières);

(2) Il n'y a pas de mesure non nulle dans $A^\perp$ qui soit orthogonale à toutes les mesures μ_n.

Alors le dual de A a la propriété de Grothendieck, c'est-à-dire $B(A, l^1)=\Pi_2(A, l^1)$.

Démonstration. La démonstration du théorème est basée sur l'inégalité d'interpolation suivante.

Proposition 11. *Soit T 2-sommant sur A. Soient $2<q<\infty$ et θ tels que*

$$\frac{1}{q'}=\theta+\frac{1-\theta}{2}\quad\left(q'=\frac{q}{q-1}\right).$$

Alors, pour tout $0\leqq\varphi<\theta$, on a l'inégalité

(A) $$i_q(T)\leqq C(\theta-\varphi)^{-1}\|T\|^\varphi\quad\pi_2(T)^{1-\varphi}$$

(où $i_q(T)$ désigne la norme q-intégrale de l'opérateur T).

Pour démontrer cette proposition, il suffit en fait ([1]) d'établir le lemme suivant.

Lemme 12. *Sous les hypothèses de la proposition et pour* $0<\delta<1$, *l'opérateur* T *a une décomposition* $T=I+S$ *où*

(1) I *est strictement q-intégral et*

$$\text{(B)} \qquad i_q(I) \leqq \|R\|_2(\theta-\varphi)^{-1}\delta^{-C(1-\varphi)/2}\|T\|^{\varphi}\pi_2(T)^{1-\varphi};$$

$$\text{(2)} \qquad \|S\| \leqq C\|T\| \quad et \quad \pi_2(S) \leqq C\delta^{1/2}\pi_2(T).$$

Démonstration du lemme. En vertu du théorème 9, on peut identifier T avec sa composante T_1. Donc T s'étend à $H^\infty(m)$ (où m est une des mesures représentantes μ_i), il existe $f\in L^1_+(m)$, $\int f=1$ telle que pour tout $\varphi\in H^\infty(m)$, on ait

$$\|T(\varphi)\| \leqq \pi_2(T) \quad \|\varphi\|_{L^2(f\,dm)}.$$

Soient (c_i), (θ_i), (τ_i) les suites obtenues par application du lemme 8 à la fonction f et $0<\delta<1$. Définissons

$$I(\varphi) = T\big(\varphi \sum_i \theta_i\tau_i^2\big), \quad S = T-I.$$

On obtient ([1]) : $\|S\|\leqq C\|T\|$ et $\pi_2(S)\leqq C\cdot\delta^{1/2}\cdot\pi_2(T)$. On désire étendre I à $\mathscr{C}(X)$. Soit pour $\varphi\in A+\bar{A}$:

$$I(\varphi) = T\big(\sum_i \tau_i R(\theta_i\tau_i\varphi)\big),$$

$$\|I(\varphi)\| = \sup_F \Big|\int \big(\sum \tau_i R(\theta_i\tau_i\varphi)\big)F\Big|$$

pour $F\in L^1(m)$ tel que

$$\Big|\int \varphi F\,dm\Big| \leqq \|T\varphi\| \text{ pour tout } \varphi\in H^\infty(m).$$

On procède alors comme dans [1], $A+\bar{A}$ jouera dans ce cadre, le rôle de l'ensemble des polynômes trigonométriques. L'opérateur I s'étend à $\mathscr{C}(X)$ (densité de $A+\bar{A}$ dans $\mathscr{C}(X)$) et vérifie l'inégalité (B).

Revenons maintenant à la démonstration du théorème. Du théorème fondamental de Grothendieck $B(\mathscr{C}, l^1)=\Pi_2(\mathscr{C}, l^1)$ et de l'inégalité d'interpolation (A), nous obtenons l'équivalence des normes opérateur et 2-sommante pour les opérateurs de rang fini de A dans l^1. Puisque l^1 a la propriété d'approximation bornée, on obtient : $B(A, l^1)=\Pi_2(A, l^1)$.

3. Conséquences, exemples et problémes. Dans ce qui suit, A désignera une algèbre séparable uniforme vérifiant les conditions 1, 2 du théorème 10. Notons que l'on a démontré en fait le résultat suivant.

Corollaire 13. *Si* Y *est un espace de Banach ayant la propriété d'approximation bornée et tel que* $B(\mathscr{C}(X), Y)=\Pi_2(\mathscr{C}(X), Y)$ *alors* $B(A, Y)=\Pi_2(A, Y)$.

Corollaire 14. *Si Y est un espace de Banach de cotype 2 ayant la propriété d'approximation bornée, alors* $B(A, Y) = \Pi_2(A, Y)$.

Corollaire 15. *Si* $T \in \Pi_2(A, Y)$, *T s'étend en un opérateur $\tilde{T}$ sur $\mathscr{C}(X)$ tel que* $\|\tilde{T}\| \leqq C\|T\| \log(\Pi_2(T)/\|T\|)$.

Corollaire 16. *Tout opérateur de rang n sur A s'étend en un opérateur $\tilde{T}$ sur $\mathscr{C}(X)$ tel que* $\|\tilde{T}\| \leqq C(\log n)\|T\|$.

Corollaire 17. *Si X est un sous-espace de dimension n de A, complémenté par une projection P, X est un P_λ-espace avec* $\lambda \leqq C(\log n)\|T\|$.

Nous mentionnons quelques exemples d'algèbres vérifiant les hypothèses du théorème 10. Soit K un compact du plan complexe. $P(K)$ est l'algèbre uniforme des fonctions à valeurs complexes qui sont limites uniformes sur K de polynômes en z. $A(K)$ est l'algèbre des fonctions continues sur K et analytiques sur l'intérieur de K. $R(K)$ est l'algèbre des fonctions limites uniformes sur K de fonctions rationnelles avec pôles sur $\mathbf{C}/K$. Les algèbres suivantes (considérées sur leur frontière de Shilov) vérifient les hypothèses du théorème ([6]) : $P(K)$ pour tout compact K, $A(K)$ quand le complémentaire de K est connexe et $R(K)$ quand $R(K)$ est une algèbre de Dirichlet.

Remarque. Soient $0<r<R$ et $K_1=\{z\in \mathbf{C}\colon r \leqq |z| \leqq R\}$. $R(K_1)$ en tant qu'espace de Banach est isomorphe à l'algèbre du disque et le dual de $R(K_1)$ a la propriété de Grothendieck. Par contre, l'algèbre $R(K_1)$ ne vérifie par les hypothèses du théorème 10.

Problèmes. Si A a la propriété de Grothendieck et est isomorphe à sa c_0-somme directe, A' est de cotype 2. Il serait donc intéressant de savoir si, sous les hypothèses du théorème, l'algèbre A est isomorphe à sa c_0-somme directe.

Si A est une algèbre uniforme et μ une unique mesure représentante, l'espace $L^1(\mu)|_{H^1(\mu)}$ est-il de cotype 2 ?

Les résultats de J. Bourgain et ce travail conduisent à la question : tout opérateur 0-sommant de A (algèbre uniforme vérifiant les conditions du théorème 10) est-il nucléaire ?

Bibliographie

[1] J. Bourgain, Opérateurs sommants sur l'algèbre du disque, *Séminaire de Géométrie des espaces de Banach, Paris VII* (1982).

[1'] J. Bourgain, New Banach space properties of the disc algebra and H^∞, *Acta Math.*, **152** (1984), 1—48.

[2] I. I. Hirschman, Jr. et R. Rochberg, Conjugate function theory in weak* Dirichlet algebras *J. Funct. Anal.*, **16** (1974), 359—371.

[3] G. Lumer, *Algèbre de fonctions et espaces de Hardy*, Springer-Verlag (Berlin—Heidelberg—New York, 1968).
[4] A. Pelczynski, *Banach spaces of analytic functions and absolutely summing operators*, CBMS Regional Conf. Ser. in Math.
[5] A. Pietsch, Absolute p-summierende Abbildungen in normierten räumen, *Studia Math.*, **28** (1967), 333—353.
[6] P. Wojtaszczyk, On weakly compact operators from some uniform algebras, *Studia Math.*, **64** (2) (1979), 105—116.

UNIVERSITÉ DE L'ETAT À MONS
FACULTÉ DES SCIENCES
AVENUE MAISTRIAU 15
7000 MONS BELGIQUE

Acta Sci. Math., 52 (1988), 179—190

On perturbations of boundary value problems for nonlinear elliptic equations on unbounded domains

L. SIMON

Introduction

In [1] it has been proved the existence of variational solutions of boundary value problems for the elliptic equation

$$\sum_{|\alpha|\leq m} (-1)^{|\alpha|} D^\alpha f_\alpha(x, u, \dots, D^\beta u, \dots) +$$
$$+ \sum_{|\alpha|\leq l} (-1)^{|\alpha|} D^\alpha g_\alpha(x, u, \dots, D^\beta u, \dots) = F, \quad x\in\Omega$$

where Ω is a possibly unbounded domain in $\mathbf{R}^n$; $|\beta|\leq m$; l is an integer with the property $l<m-(n/p)(1-p+\varrho)$; p and ϱ are real numbers such that $1<p<\infty$, $p-1<\varrho\leq p$. Functions f_α satisfy the same conditions as in [2] and g_α satisfy (essentially)

$$g_\alpha(x, \xi)\xi_\alpha \geq 0,$$
$$|g_\alpha(x, \xi)| \leq K(\xi')(C_1(x)+|\xi''|^\varrho)$$

where $\xi=(\xi', \xi'')$ and ξ' contains those coordinates ξ_β of ξ for which $|\beta|<m-(n/p)$, $C_1\in L^{p/\varrho}(\Omega)$.

In the present paper we give some stability results for solutions of the above problem. These results are connected with [3] and with several works referred in [3] where perturbation of other boundary value problems and variational inequalities has been considered.

1. Preliminaries

Let $\Omega\subset\mathbf{R}^n$ be a (possibly unbounded) domain, $p>1$, m a positive integer. Assume that Ω has the weak cone property (see [4]), and for all sufficiently large μ, there exists a bounded $\Omega_\mu\subset\Omega$ with the weak cone property such that $\Omega_\mu\supset\{x\in\Omega: |x|<\mu\}$. Denote by $W_p^m(\Omega)$ the usual Sobolev space of real valued functions u

Received February 21, 1985.

whose distributional derivatives of order $\leqq m$ belong to $L^p(\Omega)$. The norm on $W_p^m(\Omega)$ is defined by

$$\|u\| = \Big\{\sum_{|\alpha|\leqq m} \int_\Omega |D^\alpha u|^p\,dx\Big\}^{1/p}$$

where

$$\alpha = (\alpha_1, \ldots, \alpha_n), \quad |\alpha| = \sum_{j=1}^n \alpha_j, \quad D_j = \frac{\partial}{\partial x_j},$$

$$D^\alpha = D_1^{\alpha_1} \ldots D_n^{\alpha_n}.$$

Let N and M be the number of multiindices α satisfying $|\alpha|\leqq m$ and $|\alpha|\leqq m-1$, respectively. The vectors $\xi=(\xi_0, \ldots, \xi_\beta, \ldots)\in \mathbf{R}^N$ will be written in the form $\xi=(\eta, \zeta)$, where $\eta\in\mathbf{R}^M$ consists of those ξ_β for which $|\beta|\leqq m-1$. Assume that:

I. Functions $f_{\alpha,j}\colon \Omega\times\mathbf{R}^N\to\mathbf{R}$ ($|\alpha|\leqq m$; $j=0, 1, 2, \ldots$) satisfy the Carathéodory conditions, i.e. they are measurable with respect to x for each fixed $\xi\in\mathbf{R}^N$ and continuous with recpect to ξ for almost all $x\in\Omega$.

II. There exist a constant $c_1>0$ and a function $K_1\in L^q(\Omega)$ (where $1/p+1/q=1$) such that

$$|f_{\alpha,j}(x, \xi)| \leqq c_1|\xi|^{p-1}+K_1(x)$$

for all $|\alpha|\leqq m$, $j=0, 1, 2, \ldots$, a.e. $x\in\Omega$ and all $\xi\in\mathbf{R}^N$.

III. For all $(\eta, \zeta), (\eta, \zeta')\in\mathbf{R}^N$ with $\eta\in\mathbf{R}^M$, $\zeta\neq\zeta'$ and a.e. $x\in\Omega$ ($j=0, 1, 2, \ldots$)

$$\sum_{|\alpha|=m} [f_{\alpha,j}(x, \eta, \zeta)-f_{\alpha,j}(x, \eta, \zeta')](\xi_\alpha-\xi'_\alpha) > 0.$$

IV. There exist a constant $c_2>0$ and a function $K_2\in L^1(\Omega)$ such that for a.e. $x\in\Omega$ and all $\xi\in\mathbf{R}^N$

$$\sum_{|\alpha|\leqq m} f_{\alpha,j}(x, \xi)\xi_\alpha \geqq c_2|\xi|^p-K_2(x) \quad (j = 0, 1, 2, \ldots).$$

V. $\lim_{j\to\infty} \xi^{(j)} = \xi^{(0)}$ implies

$$\lim_{j\to\infty} f_{\alpha,j}(x, \xi^{(j)}) = f_{\alpha,0}(x, \xi^{(0)})$$

for a.e. $x\in\Omega$ and all $|\alpha|\leqq m$.

VI. Functions $p_{\alpha,j}, r_{\alpha,j}\colon \Omega\times\mathbf{R}^N\to\mathbf{R}$

$$(|\alpha| \leqq l, j = 0, 1, 2, \ldots)$$

satisfy the Carathéodory conditions and

$$g_{\alpha,j} = p_{\alpha,j}+r_{\alpha,j}.$$

VII. $p_{\alpha,j}(x, \xi)\xi_\alpha\geqq 0$ and $|r_{\alpha,j}(x, \xi)|\leqq h_\alpha(x)$ for all $|\alpha|\leqq l$, $\xi\in\mathbf{R}^N$ and a.e. $x\in\Omega$ where $h_\alpha\in L^{p/q}(\Omega)$, $j=0, 1, 2, \ldots$.

VIII. There exist a continuous function K_3 and $C_1 \in L^{p/\varrho}(\Omega)$ such that

$$|p_{\alpha,j}(x,\xi)| \leqq K_3(\xi')\big(C_1(x)+|\xi''|^\varrho\big) \quad j=0,1,2,\ldots$$

for all $|\alpha| \leqq l$, $\xi=(\xi',\xi'')\in \mathbf{R}^N$ and a.e. $x\in\Omega$ (ξ' contains those ξ_β for which $|\beta|< m-(n/p)$; $p-1<\varrho\leqq p$, $l<m-(n/p)(1-p+\varrho)$).

IX. $\lim_{j\to\infty} \xi^{(j)}=\xi^{(0)}$ implies

$$\lim_{j\to\infty} p_{\alpha,j}(x,\xi^{(j)})=p_{\alpha,0}(x,\xi^{(0)}), \quad \lim_{j\to\infty} r_{\alpha,j}(x,\xi^{(j)})=r_{\alpha,0}(x,\xi^{(0)})$$

for a.e. $x\in\Omega$ and all $|\alpha|\leqq l$.

X. V is a closed subspace of $W_p^m(\Omega)$ with the property: $v\in V$, $\varphi\in C_0^\infty(\mathbf{R}^n)$ imply that $\varphi v\in V$. (By $C_0^\infty(G)$ is denoted the set of infinitely differentiable functions with compact support contained in G.)

XI. $F_j\in V'$ ($j=0,1,2,\ldots$), i.e. F_j is a linear continuous functional on V and

$$\lim_{j\to\infty} \|F_j-F_0\|_{V'}=0.$$

Remarks. 1. Assume that I—IV, VI—VIII are fulfilled for $j=0$, i.e. $f_{\alpha,0}$, $g_{\alpha,0}$ satisfy conditions of the existence theorem in [1]. Further suppose that $f_{\alpha,j}$, $g_{\alpha,j}$ ($j=1,2,\ldots$) satisfy I, VI such that

$$\lim_{j\to\infty}\big[\sup_{\xi\in\mathbf{R}^N}|f_{\alpha,j}(x,\xi)-f_{\alpha,0}(x,\xi)|\big]=0 \quad \text{for a.e.} \quad x\in\Omega,$$

$$\sup_{\xi\in\mathbf{R}^N}|f_{\alpha,j}(x,\xi)-f_{\alpha,0}(x,\xi)| \leqq \varphi(x) \quad \text{for a.e.} \quad x\in\Omega$$

where $\varphi\in L^q(\Omega)$, $j=1,2,\ldots$;

$$\lim_{j\to\infty}\big[\sup_{\xi\in\mathbf{R}^N}|g_{\alpha,j}(x,\xi)-g_{\alpha,0}(x,\xi)|\big]=0 \quad \text{for a.e.} \quad x\in\Omega,$$

$$\sup_{\xi\in\mathbf{R}^N}|g_{\alpha,j}(x,\xi)-g_{\alpha,0}(x,\xi)| \leqq \psi(x) \quad \text{for a.e.} \quad x\in\Omega$$

where $\psi\in L^{p/\varrho}(\Omega)$, $j=1,2,\ldots$.

Then I, II, IV—VIII are satisfied for $f_{\alpha,j}, g_{\alpha,j}$ ($j=1,2,\ldots$) with $p_{\alpha,j}:=p_{\alpha,0}$, $r_{\alpha,j}:=(g_{\alpha,j}-g_{\alpha,0})+r_{\alpha,0}$.

2. If there is a constant $c>0$ such that for a.e. $x\in\Omega$, all (η,ζ), $(\eta,\zeta')\in\mathbf{R}^N$

$$\sum_{|\alpha|=m}[f_{\alpha,0}(x,\eta,\zeta)-f_{\alpha,0}(x,\eta,\zeta')](\xi_\alpha-\xi'_\alpha) \geqq c|\zeta-\zeta'|^p$$

and

$$|[f_{\alpha,j}(x,\eta,\zeta)-f_{\alpha,j}(x,\eta,\zeta')]-[f_{\alpha,0}(x,\eta,\zeta)-f_{\alpha,0}(x,\eta,\zeta')]| \leqq$$
$$\leqq d_j|\zeta-\zeta'|^{p-1} \quad (j=1,2,\ldots)$$

where $\lim_{j\to\infty} d_j=0$, then $f_{\alpha,j}$ satisfy III for sufficiently large j.

Lemma 1. *Assume that* $u_j \to u$ *weakly in* V *and for any bounded domain* $\omega \subset \Omega$

(1.1)
$$\lim_{j\to\infty} \int_\omega h_j \, dx = 0,$$

where

(1.2)
$$h_j(x) = \sum_{|\alpha|=m} [f_{\alpha,j}(x, u_j, \ldots, D^\gamma u_j, \ldots, D^\beta u_j, \ldots) -$$
$$- f_{\alpha,j}(x, u_j, \ldots, D^\gamma u_j, \ldots, D^\beta u, \ldots)](D^\alpha u_j - D^\alpha u),$$

$|\gamma|<m$, $|\beta|=m$. *Then there is a subsequence* (u_{j_k}) *of* (u_j) *such that* $D^\beta u_{j_k} \to D^\beta u$ *a.e. in* Ω *for all* β *with* $|\beta| \leqq m$ *and for any bounded* $\omega \subset \Omega$, $u_{j_k} \to u$ *with respect to the norm of* $W_p^m(\omega)$.

Proof. Since $u_j \to u$ weakly in V there is a subsequence (u_{j_k}) of (u_j) such that for $|\gamma|<m$

$$D^\gamma u_{j_k} \to D^\gamma u \quad \text{a.e. in } \Omega$$

and

(1.3)
$$\lim_{k\to\infty} \|D^\gamma u_{j_k} - D^\gamma u\|_{L^p(\omega)} = 0$$

for any bounded subdomain ω of Ω (see e.g. [5] and [4]). Further, by assumption III $h_j \geqq 0$ and so (1.1) and Fatou's lemma imply that $h_j \to 0$ a.e. in ω. Thus there exists $\omega_0 \subset \omega$ of measure 0 such that for $x \in \omega \setminus \omega_0$

(1.4)
$$|D^\beta u(x)| < \infty, \ |K_1(x)| < \infty, \ |K_2(x)| < \infty,$$

(1.5)
$$D^\gamma u_{j_k}(x) \to D^\gamma u(x) \quad (|\gamma| < m), \ h_{j_k}(x) \to 0, \quad k \to \infty.$$

Set

$$\xi^{(k)}(x) = (\ldots, D^\beta u_{j_k}(x), \ldots)$$

where $|\beta|=m$. By assumptions II, IV, V and (1.4), (1.5) we have

(1.6)
$$h_{j_k}(x) \geqq \sum_{|\alpha|=m} f_{\alpha,j_k}(x, u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) D^\alpha u_{j_k} -$$
$$- \sum_{|\alpha|=m} |f_{\alpha,j_k}(x, u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) D^\alpha u| -$$
$$- \sum_{|\alpha|=m} |f_{\alpha,j_k}(x, u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots, D^\beta u, \ldots)(D^\alpha u_{j_k} - D^\alpha u)| \geqq$$
$$\geqq c_2 |\xi^{(k)}(x)|^p - c_3(x)[1 + |\xi^{(k)}(x)|^{p-1} + |\xi^{(k)}(x)|]$$

if $x \in \omega \setminus \omega_0$ where $|\gamma|<m$, $|\beta|=m$. (For a fixed $x \in \omega \setminus \omega_0$, $D^\gamma u_{j_k}(x)$ and $f_{\alpha,j_k}(x, u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots, D^\beta u, \ldots)$ are convergent and thus they are bounded.) By (1.5) $(h_{j_k}(x))$ is bounded for a fixed $x \in \omega \setminus \omega_0$, thus (1.6) implies that $(\xi^{(k)}(x))$ is bounded, too. Consequently, for a fixed $x \in \omega \setminus \omega_0$, $(\xi^{(k)}(x))$ contains a subsequence which converges to a vector $\xi^*(x)$.

Now we show that

$$\xi^*(x) = \xi(x) = (\dots, D^\beta u(x), \dots). \tag{1.7}$$

Indeed, applying (1.2) to the subsequence of $(h_{j_k}(x))$ with $k \to \infty$, by (1.5) and assumption V we obtain

$$0 = \sum_{|\alpha|=m} [f_{\alpha,0}(x, u(x), \dots, D^\gamma u(x), \dots, \xi^*(x)) -$$
$$-f_{\alpha,0}(x, u(x), \dots, D^\gamma u(x), \dots, \xi(x))][\xi^*_\alpha(x) - \xi_\alpha(x)]$$

which implies (1.7) in virtue of assumption III.

So we have shown that all convergent subsequences of the bounded sequence $(\xi^{(k)}(x))$ tend to $\xi(x)$. Therefore, $\lim_{k\to\infty} \xi^{(k)}(x) = \xi(x)$ if $x \in \omega \setminus \omega_0$ and thus, by (1.5) $D^\beta u_{j_k} \to D^\beta u$ a.e. in ω for all β satisfying $|\beta| \leqq m$. Since ω is an arbitrary bounded subset of Ω we have

$$D^\beta u_{j_k} \to D^\beta u \quad \text{a.e. in} \quad \Omega \quad \text{if} \quad |\beta| \leqq m. \tag{1.8}$$

By using notations

$$F_k(x) = \sum_{|\alpha|=m} f_{\alpha,j_k}(x, u_{j_k}, \dots, D^\beta u_{j_k}, \dots) D^\alpha u_{j_k},$$
$$F_0(x) = \sum_{|\alpha|=m} f_{\alpha,0}(x, u, \dots, D^\beta u, \dots) D^\alpha u,$$

from (1.1) one obtains that

$$\int_\omega F_k\, dx - \sum_{|\alpha|=m} \int_\omega f_{\alpha,j_k}(x, u_{j_k}, \dots, D^\gamma u_{j_k}, \dots, D^\beta u_{j_k}, \dots) D^\alpha u\, dx -$$
$$- \sum_{|\alpha|=m} \int_\omega f_{\alpha,j_k}(x, u_{j_k}, \dots, D^\gamma u_{j_k}, \dots, D^\beta u, \dots) D^\alpha (u_{j_k} - u)\, dx \to 0,$$

i.e.

$$\int_\omega F_k\, dx - \int_\omega F_0\, dx - \tag{1.9}$$
$$- \sum_{|\alpha|=m} \int_\omega [f_{\alpha,j_k}(x, u_{j_k}, \dots, D^\gamma u_{j_k}, \dots, D^\beta u_{j_k}, \dots) -$$
$$-f_{\alpha,0}(x, u, \dots, D^\gamma u, \dots, D^\beta u, \dots)] D^\alpha u\, dx -$$
$$- \sum_{|\alpha|=m} \int_\omega f_{\alpha,j_k}(x, u_{j_k}, \dots, D^\gamma u_{j_k}, \dots, D^\beta u, \dots) D^\alpha (u_{j_k} - u)\, dx \to 0.$$

By assumptions II, V, (1.8), Hölder's inequality and Vitali's theorem the third term in (1.9) converges to 0. Furthermore, (1.8), assumptions II, V, (1.3) and Vitali's theorem imply that

$$f_{\alpha,j_k}(x, u_{j_k} \dots, D^\gamma u_{j_k}, \dots, D^\beta u, \dots) \to f_{\alpha,0}(x, u, \dots, D^\gamma u, \dots, D^\beta u, \dots)$$

in the norm of $L^q(\omega)$. Since $\lim_{k\to\infty} D^\alpha(u_{j_k}-u)\to 0$ weakly in $L^p(\Omega)$ one finds that the fourth term in (1.9) converges to 0, too.

Therefore, from (1.9) it follows that

$$\lim_{k\to\infty} \int_\omega F_k\,dx = \int_\omega F_0\,dx. \tag{1.10}$$

By assumption IV

$$F_k(x) \geqq c_2 \sum_{|\beta|=m} |D^\beta u_{j_k}(x)|^p - K_2(x).$$

Thus for functions $G_k=F_k+K_2$, $G_0=F_0+K_2$ we have

$$G_k(x) \geqq c_2 \sum_{|\beta|=m} |D^\beta u_{j_k}(x)|^p \geqq 0, \tag{1.11}$$

and by (1.10)

$$\lim_{k\to\infty} \int_\omega G_k\,dx = \int_\omega G_0\,dx. \tag{1.12}$$

(1.8) and assumption V imply that $G_k\to G_0$ a.e. in ω, thus from (1.11), (1.12) it follows that

$$G_k \to G_0 \quad \text{in} \quad L^1(\omega) \tag{1.13}$$

(see [6]). Consequently, (1.8), (1.11) and Vitali's theorem imply that, for $|\beta|=m$, $D^\beta u_{j_k}\to D^\beta u$ in $L^p(\omega)$, and the proof of Lemma 1 is complete.

Assume that instead of III condition

III'. $\sum_{|\alpha|\leqq m} [f_{\alpha,j}(x,\xi)-f_{\alpha,j}(x,\xi')](\xi_\alpha-\xi'_\alpha) > 0$

is fulfilled if $\xi\neq\xi'$.

An easy modification of the proof of Lemma 1 gives

Lemma 2. *Suppose that* $u_j\to u$ *weakly in* V *and*

$$\lim_{j\to\infty} \int_\Omega \tilde{h}_j\,dx = 0,$$

where

$$\tilde{h}_j(x) = \sum_{|\alpha|\leqq m} [f_{\alpha,j}(x,u_j,\dots,D^\beta u_j,\dots)-f_{\alpha,j}(x,u,\dots,D^\beta u,\dots)](D^\alpha u_j-Du).$$

Then there is a subsequence (u_{j_k}) *of* (u_j) *such that* $u_{j_k}\to u$ *with respect to the norm of* $W_p^m(\Omega)$.

2. Stability results

Theorem 1. *Assume that conditions I—XI are fulfilled and* $u_j \in V$ *is a solution of*

$$\sum_{|\alpha|\leqq m} \int_\Omega f_{\alpha,j}(x, u_j, \ldots, D^\beta u_j, \ldots) D^\alpha v\, dx + \tag{2.1}$$

$$+ \sum_{|\alpha|\leqq l} \int_\Omega g_{\alpha,j}(x, u_j, \ldots, D^\beta u_j, \ldots) D^\alpha v\, dx = \langle F_j, v\rangle$$

for all $v \in V$ $(j = 1, 2, \ldots)$.

Then there is a subsequence (u_{j_k}) *of* (u_j) *which converges weakly in* V *to a solution* $u \in V$ *of* (2.1) *for* $j=0$. *Moreover,* $D^\beta u_{j_k} \to D^\beta u$ *a.e. in* Ω *if* $|\beta| \leqq m$, *and for arbitrary bounded* $\omega \subset \Omega$, $u_{j_k} \to u$ *strongly in* $W_p^m(\omega)$.

If solution u *of* (2.1) *for* $j=0$ *is unique then* $u_j \to u$ *weakly in* V *and strongly in* $W_p^m(\omega)$ *for any bounded* $\omega \subset \Omega$.

Remark. According to [1], for any $F_j \in V'$ there exists at least one solution $u_j \in V$ of (2.1).

Proof of Theorem 1. Applying (2.1) to $v = u_j$, by assumptions IV, VI, VII we obtain that

$$c_2 \|u_j\|_V^p - \int_\Omega K_2(x)\, dx - \sum_{|\alpha|\leqq l} \|h_\alpha\|_{L^{p/\varrho}(\Omega)} \|D^\alpha u_j\|_{L^{q_1}(\Omega)} \leqq \|F_j\|_{V'} \|u_j\|_V \tag{2.2}$$

where q_1 is defined by $1/(p/\varrho) + 1/q_1 = 1$.
By an imbedding theorem (see e.g. [4]) for

$$|\alpha| \leqq l(< m - (n/p)(1-p+\varrho)), \quad v \in W_p^m(\Omega) \quad \text{we have}$$

$$\|D^\alpha v\|_{L^{q_1}(\Omega)} \leqq c \|v\|_{W_p^m(\Omega)} \tag{2.3}$$

(c is a constant) because $q_1 < np/(n-(m-l)p)$. Thus (2.2) and $p>1$ imply that (u_j) is bounded in V. Therefore, there exist a subsequence (u_{j_k}) of (u_j) and $u \in V$ such that

$$u_{j_k} \to u \quad \text{weakly in} \quad V, \tag{2.4}$$

$$D^\gamma u_{j_k} \to D^\gamma u \quad \text{a.e. in} \quad \Omega \quad \text{for} \quad |\gamma| \leqq m-1 \tag{2.5}$$

(see [5]).

Consider an arbitrary bounded domain $\omega \subset \Omega$ and a function $\Theta \in C_0^\infty(\mathbf{R}^n)$ such that $\Theta \geqq 0$ and $\Theta(x) = 1$ for $x \in \omega$. By the theorems on compact imbedding (see e.g. [4]) it may be supposed that

$$D^\gamma u_{j_k} \to D^\gamma u \text{ in } L^p(\Omega \cap \operatorname{supp} \Theta) \quad \text{for} \quad |\gamma| \leqq m-1 \tag{2.6}$$

and

(2.7) $$D^\gamma u_{j_k} \to D^\gamma u \quad \text{in} \quad L^{q_1}(\Omega \cap \operatorname{supp} \Theta) \quad \text{for} \quad |\gamma| \leqq l,$$

where q_1 is defined by $1/(p/\varrho)+1/q_1=1$ $\big(l<m-(n/p)(1-p+\varrho)\big)$. By a "diagonal process" the subsequence (u_{j_k}) can be chosen so that (2.6), (2.7) are true for any fixed $\Theta \in C_0^\infty(\mathbf{R}^n)$.

In virtue of assumption X $\Theta(u_{j_k}-u) \in V$ and thus from (2.1) one obtains

(2.8) $$\sum_{|\alpha| \leqq m} \int_\Omega f_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) D^\alpha[(\Theta(u_{j_k}-u)]\, dx +$$
$$+ \sum_{|\alpha| \leqq l} \int_\Omega g_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) D^\alpha[\Theta(u_{j_k}-u)]\, dx =$$
$$= \langle F_{j_k}, \Theta(u_{j_k}-u)\rangle.$$

Since $(u_{j_k}-u) \to 0$ weakly in V

(2.9) $$\Theta(u_{j_k}-u) \to 0 \quad \text{weakly in } V.$$

From (2.8) it follows that

(2.10) $$\sum_{|\alpha|=m} \int_\Omega [f_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) -$$
$$- f_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots, D^\beta u, \ldots]\, \Theta D^\alpha(u_{j_k}-u)\, dx =$$
$$= \sum_{|\alpha|=m} \int_\Omega f_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots, D^\beta u, \ldots)\, \Theta D^\alpha(u-u_{j_k})\, dx +$$
$$+ \sum_{|\alpha|=m} \int_\Omega f_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) \sum_{|\gamma| \leqq m-1} c_\gamma D^\gamma(u-u_{j_k}) D^{\alpha-\gamma}\Theta\, dx +$$
$$+ \sum_{|\alpha| \leqq m-1} \int_\Omega f_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) D^\alpha[\Theta(u-u_{j_k})]\, dx +$$
$$+ \sum_{|\alpha| \leqq l} \int_\Omega g_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) D^\alpha[\Theta(u-u_{j_k})]\, dx +$$
$$+ \langle F_{j_k}, \Theta(u_{j_k}-u)\rangle \quad (|\gamma| < m,\ |\beta| = m).$$

Now we show that all the terms on the right-hand side of (2.10) converge to 0 as $k \to \infty$. By (2.4), $D^\alpha(u_{j_k}-u) \to 0$ weakly in $L^p(\Omega)$. Furthermore, from (2.5) and assumption V we get

(2.11) $$\Theta f_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots, D^\beta u, \ldots) \to$$
$$\to \Theta f_{\alpha, 0}(x, u, \ldots, D^\gamma u, \ldots, D^\beta u, \ldots)$$

a.e. in Ω, and, consequently, by assumption II, (2.6) and Vitali's theorem (2.11) is valid in $L^q(\Omega)$ norm, too. Thus the first term in (2.10) converges to 0.

By assumptions I, II the functions

$$f_{\alpha, j_k}(x; u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots)$$

are bounded in $L^q(\Omega)$, hence (2.6) implies that the second and third terms in (2.10) converge to 0 as $k\to\infty$.

From assumptions VI—VIII it follows that

$$g_{\alpha,j_k}(x, u_{j_k}, \dots, D^\gamma u_{j_k}, \dots, D^\beta u_{j_k}, \dots)$$

is bounded in $L^{p/q}(\Omega\cap\operatorname{supp}\Theta)$, thus (2.7) implies that the fourth term in (2.10) converges to 0 as $k\to\infty$. Finally, for the last term we have

$$|\langle F_{j_k}, \Theta(u_{j_k}-u)\rangle| \leqq |\langle F_{j_k}-F_0, \Theta(u_{j_k}-u)\rangle| +$$
$$+|\langle F_0, \Theta(u_{j_k}-u)\rangle| \leqq \|F_{j_k}-F_0\|_{V'} \|\Theta(u_{j_k}-u)\|_V + |\langle F_0, \Theta(u_{j_k}-u)\rangle|,$$

thus assumption XI, (2.9) imply that also the last term in (2.10) converges to 0 as $k\to\infty$.

Thus we have shown that the term on the left-hand side of (2.10) converges to 0 as $k\to\infty$. By assumption III and $\Theta\geqq 0$ we find that (1.1) is valid for a subsequence of (h_j). Consequently, from Lemma 1 we obtain that (u_{j_k}) contains a subsequence $(u_{j_k'})$ such that

(2.12) $$D^\beta u_{j_k'} \to D^\beta u \quad \text{a.e. in } \Omega$$

if $|\beta|\leqq m$, and for any bounded $\omega\subset\Omega$

(2.13) $$(u_{j_k'}) \to u \quad \text{in} \quad W_p^m(\omega).$$

(2.12) and assumption V implies that

$$f_{\alpha,j_k'}(x, u_{j_k'}, \dots, D^\beta u_{j_k'}, \dots) \to f_{\alpha,0}(x, u, \dots, D^\beta u, \dots)$$

a.e. in Ω. Therefore, assumption II the boundedness of $\|u_{j_k'}\|_V$, Hölder's inequality and Vitali's theorem imply that for any $v\in V$

(2.14) $$\lim_{k\to\infty} \sum_{|\alpha|\leqq m} \int_\Omega f_{\alpha,j_k'}(x, u_{j_k'}, \dots, D^\beta u_{j_k'}, \dots) D^\alpha v\, dx =$$
$$= \sum_{|\alpha|\leqq m} \int_\Omega f_{\alpha,0}(x, u, \dots, D^\beta u, \dots) D^\alpha v\, dx.$$

By using assumption IX and (2.12) we find $g_{\alpha,j_k'}(x, u_{j_k'}, \dots, D^\beta u_{j_k'}, \dots)\to$ $\to g_{\alpha,0}(x, u, \dots, D^\beta u, \dots)$ a.e. in Ω and thus, by assumptions VI—VIII, (2.3), Hölder's inequality and Vitali's theorem we find that for any $v\in V$

$$\lim_{k\to\infty} \sum_{|\alpha|\leqq l} \int_\Omega g_{\alpha,j_k'}(x, u_{j_k'}, \dots, D^\beta u_{j_k'}, \dots) D^\alpha v\, dx =$$
$$= \sum_{|\alpha|\leqq l} \int_\Omega g_{\alpha,0}(x, u, \dots, D^\beta u, \dots) D^\alpha v\, dx.$$

Thus from (2.1), (2.14), assumption XI it follows that u is a solution of (2.1) for $j=0$ and, by (2.4), (2.12), (2.13), the proof of the first statement of Theorem 1 is complete.

If solution u of problem (2.1) for $j=0$ is unique but "$u_j \to u$ weakly in V" is not true then there are $G \in V'$, a positive number ε and a subsequence (u'_j) of (u_j) such that

$$|Gu'_j - Gu| > \varepsilon, \quad j = 1, 2, \dots . \tag{2.15}$$

Applying the first statement of Theorem 1 to (u'_j) instead of (u_j) we find that there is a subsequence (u''_j) of (u'_j) which converges weakly in V to a solution of (2.1) for $j=0$, i.e. $u''_j \to u$ weakly in V (because the solution of (2.1) for $j=0$ is unique). But this is impossible because of (2.15). It can be proved similarly that then $u_j \to u$ strongly in $W_p^m(\omega)$ for any bounded $\omega \subset \Omega$.

Theorem 2. *Assume that conditions I—II, III', IV—XI are fulfilled and $u_j \in V$ is a solution of (2.1). Then there is a subsequence (u_{j_k}) of (u_j) which converges strongly in V to a solution $u \in V$ of (2.1) for $j=0$. If the solution u of (2.1) for $j=0$ is unique then (u_j) also converges to u strongly in V.*

Proof. Assumption III' implies III thus all conditions of Theorem 1 are fulfilled. Consequently, by Theorem 1 there is a subsequence (u_{j_k}) of (u_j) such that

$$u_{j_k} \to u \quad \text{weakly in } V \tag{2.16}$$

and

$$D^\beta u_{j_k} \to D^\beta u \quad \text{a.e. in } \Omega \qquad \text{for} \quad |\beta| \le m, \tag{2.17}$$

where u is a solution of (2.1) for $j=0$.

Now we show that the sequence (u_{j_k}) satisfies the condition of Lemma 2. Since u_{j_k} is a solution of (2.1) with $j=j_k$, $v=u_{j_k}$ and u is a solution of (2.1) with $j=0$, $v=u$, we have

(2.18)

$$\sum_{|\alpha|\le m} \int_\Omega [f_{\alpha,j_k}(x, u_{j_k}, \dots, D^\beta u_{j_k}, \dots) - f_{\alpha,j_k}(x, u, \dots, D^\beta u, \dots)](D^\alpha u_{j_k} - D^\alpha u)\,dx =$$
$$= \sum_{|\alpha|\le m} \int_\Omega f_{\alpha,j_k}(x, u, \dots, D^\beta u, \dots)(D^\alpha u - D^\alpha u_{j_k})\,dx +$$
$$+ \sum_{|\alpha|\le m} \int_\Omega [f_{\alpha,0}(x, u, \dots, D^\beta u, \dots) - f_{\alpha,j_k}(x, u_{j_k}, \dots, D^\beta u_{j_k}, \dots)]\, D^\alpha u\,dx +$$
$$+ \sum_{|\alpha|\le l} \int_\Omega [g_{\alpha,0}(x, u, \dots, D^\beta u, \dots) D^\alpha u - g_{\alpha,j_k}(x, u_{j_k}, \dots, D^\beta u_{j_k}, \dots) D^\alpha u_{j_k}]\,dx +$$
$$+ \{\langle F_{j_k}, u_{j_k}\rangle - \langle F_0, u\rangle\}.$$

Applying Vitali's theorem, Hölder's inequality, assumptions I, II, V and (2.16), (2.17), we find that the first and second terms on the right-hand side of (2.18) converge to 0 as $k \to \infty$. By assumption XI and (2.16) we have

$$|\langle F_{j_k}, u_{j_k}\rangle - \langle F_0, u\rangle| \leqq |\langle F_{j_k} - F_0, u_{j_k}\rangle| + |\langle F_0, u_{j_k} - u\rangle| \leqq$$
$$\leqq \|F_{j_k} - F_0\|_{V'} \|u_{j_k}\|_V + |\langle F_0, u_{j_k} - u\rangle| \to 0 \quad \text{as} \quad k \to \infty.$$

Furthermore, (2.17) and assumption IX yield

$$p_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) D^\alpha u_{j_k} \to p_{\alpha,0}(x, u, \ldots, D^\beta u, \ldots) D^\alpha u$$

a.e. in Ω. In virtue of Fatou's lemma and assumption VII we get the inequality

$$\text{(2.19)} \qquad \int_\Omega p_{\alpha,0}(x, u, \ldots, D^\beta u, \ldots) D^\alpha u \, dx \leqq$$
$$\leqq \liminf_{k\to\infty} \int_\Omega p_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) D^\alpha u_{j_k} \, dx.$$

Assumptions VII, IX, (2.17), Hölder's inequality and Vitali's theorem imply that

$$\lim_{k\to\infty} \int_\Omega r_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) D^\alpha u_{j_k} \, dx =$$
$$= \int_\Omega r_{\alpha,0}(x, u, \ldots, D^\beta u, \ldots) D^\alpha u \, dx.$$

Hence and from (2.19) it follows that

$$\limsup_{k\to\infty} \sum_{|\alpha|\leqq l} \int_\Omega [g_{\alpha,0}(x, u, \ldots, D^\beta u, \ldots) D^\alpha u -$$
$$- g_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) D^\alpha u_{j_k}] \, dx =$$
$$= \limsup_{k\to\infty} \sum_{|\alpha|\leqq l} \int_\Omega [p_{\alpha,0}(x, u, \ldots, D^\beta u, \ldots) D^\alpha u -$$
$$- p_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) D^\alpha u_{j_k}] \, dx \leqq$$
$$\leqq \sum_{|\alpha|\leqq l} \int_\Omega p_{\alpha,0}(x, u, \ldots, D^\beta u, \ldots) D^\alpha u \, dx +$$
$$+ \sum_{|\alpha|\leqq l} \limsup_{k\to\infty} \int_\Omega [-p_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) D^\alpha u_{j_k}] \, dx \leqq 0.$$

In virtue of (2.18) we have shown that

$$\limsup_{k\to\infty} \sum_{|\alpha|\leqq m} \int_\Omega [f_{\alpha, j_k}(x, u_{j_k}, \ldots, D^\beta u_{j_k}, \ldots) -$$
$$- f_{\alpha, j_k}(x, u, \ldots, D^\beta u, \ldots)](D^\alpha u_{j_k} - D^\alpha u) \, dx \leqq 0.$$

Hence by assumption III′ it follows that (u_{j_k}) satisfies the conditions of Lemma 2 and there is a subsequence $(u_{j'_k})$ of (u_{j_k}) such that $u_{j'_k} \to u$ in $W_p^m(\Omega)$. This completes the proof of the first statement of Theorem 2. The case when the solution u of (2.1) for $j=0$ is unique can be treated in the same way as in the proof of Theorem 1.

References

[1] L. Simon, On boundary value problems for nonlinear elliptic equations on unbounded domains, *Publ. Math., Debrecen,* **34** (1987), 75—81.

[2] F. E. Browder, Pseudo-monotone operators and nonlinear elliptic boundary value problems on unbounded domains, *Proc. Nat. Acad. Sci. USA,* **74** (1977), 2659—2661.

[3] L. Boccardo and F. Donati, Existence and stability results for solutions of some strongly nonlinear constrained problems, *Nonlinear Anal.* **5** (1981), 975—988.

[4] О. В. Бесов и В. П. Ильин и С. М. Никольский, *Интегральные представления функций и теоремы вложения,* Наука, (Москва, 1975).

[5] D. E. Edmunds and J. R. L. Webb, Quasilinear elliptic problems in unbounded domains, *Proc. Roy. Soc. London Ser. A,* 337 (1973), 397—410.

[6] И. П. Натансон, *Теория функций вещественной переменной,* Наука (Москва, 1974).

DEPARTMENT OF APPLIED ANALYSIS
EÖTVÖS LORÁND UNIVERSITY
MÚZEUM KRT. 6—8
1088 BUDAPEST, HUNGARY

Acta Sci. Math., 52 (1988), 191—205

A characterization of weak convergence of weighted multivariate empirical processes

J. H. J. EINMAHL, F. H. RUYMGAART and J. A. WELLNER

1. Introduction

The characterization of weak convergence of the one-dimensional weighted empirical process indexed by points is obtained by CHIBISOV [5] and O'REILLY [11]. Later, SHORACK [16] and SHORACK and WELLNER [17] wanted to give a new, "elementary" proof of this so called Chibisov—O'Reilly theorem but their proofs were not correct wihout additional monotonicity conditions on the weight functions. This was pointed out in CSÖRGŐ, CSÖRGŐ, HORVÁTH and MASON [6] (pp. 25—27). SHORACK and WELLNER [17] also gave a characterization of weak convergence of the one-dimensional weighted empirical process indexed by rectangles. Their proof, however, is again only correct with an additional monotonicity condition on the weight function. Recently a new approximation of the empirical process is established in CSÖRGŐ CSÖRGŐ, HORVÁTH and MASON [7] which among others yields a proof of the Chibisov—O'Reilly theorem.

The aforementioned theorems can be generalized in two directions: (I) the case of dependent and/or non-identically distributed random variables and (II) the multivariate case. Case I has been studied by ALEXANDER [1], ALY, BEIRLANT and HORVÁTH [3] and BEIRLANT and HORVÁTH [4]. In our paper, which is a revision of the technical report EINMAHL, RUYMGAART and WELLNER [9], we study case II, i.e. we derive necessary and sufficient conditions on the weight functions for weak convergence of weighted multivariate empirical processes; these processes are indexed by quadrants (points) and rectangles respectively. Our main tools are exponential probability inequalities for the empirical process. The paper is a continuation of RUYMGAART and WELLNER [14], [15], where the basic tools are already presented but attention is focussed on strong convergence properties.

Received November 8, 1984.

During the preparation of the earlier version of this work we became aware of recent developments in this area, especially the work of ALEXANDER [1], already quoted before, on weighted empirical processes based on non-i.i.d. random elements and indexed by Vapnik—Chervonenkis classes of sets. Although his results are of impressive generality, also this author needs a rather unnatural monotonicity condition which we can avoid everywhere, i.e. though his theorems allow more general indexing classes, our theorems allow more general weight functions. Very recently, ALEXANDER [2] also obtained our (stronger) version of the multivariate characterization theorem for points.

In order to be more explicit we need to present the basic notation. Let $X_1^{(n)}, \dots, X_n^{(n)}$, $n \in \mathbf{N}$, be a triangular array of i.i.d. random vectors that are uniformly distributed on $[0, 1]^d$, $d \in \mathbf{N}$. Adopting the notation in OREY and PRUITT [12] we shall write $x = \langle x_1, \dots, x_d \rangle = \langle x_j \rangle = \langle x(j) \rangle \in \mathbf{R}^d$ if it is desirable to display the coordinates of x. If $x_j = \xi$ for all j we simply write $\langle \xi \rangle$. For $x, y \in \mathbf{R}^d$ we write $x \leqq y$ if $x_j \leqq y_j$ for all j and $x < y$ if $x \leqq y$ and $x \neq y$. It has some advantage to denote the half-open rectangles $(x_1, y_1] \times \dots \times (x_d, y_d]$ by $R(x, y)$ rather than $(x, y]$. The classes

(1.1) $$\mathscr{R}_0 = \{R(\langle 0 \rangle, y)\colon R(\langle 0 \rangle, y) \subset [0, 1]^d\}, \quad \mathscr{R} = \{R(x, y)\colon R(x, y) \subset [0, 1]^d\},$$

of all half-open quadrants respectively rectangles in the unit square will play an important role. We will write $|t| = t_1 \times \dots \times t_d$, $|dt|$ for Lebesgue measure on $[0, 1]^d$ and $|R|$ for the Lesbesgue measure of a rectangle R. Using this notation for the uniform underlying d.f. F we have

(1.2) $$F(t) = |t|, \; t \in [0, 1]^d.$$

Given any function $\Lambda\colon \mathbf{R}^d \to \mathbf{R}$ and an arbitrary rectangle $R = R(x, y)$ we write

(1.3) $$\Lambda\{R\} = \Lambda\{R(x, y)\} = \Delta_x^y \Lambda,$$

extending the difference operator Δ_x^y, usually applied only to distribution functions.

The weight functions will be always restricted to the class

(1.4) $$\mathscr{Q}^* = \{q\colon [0, 1] \to [0, \infty) \text{ with } q \text{ continuous and non-decreasing, } q > 0 \text{ on } (0, 1]\}.$$

The subclasses that will appear in our characterization are

(1.5) $$\mathscr{Q}_0 = \Big\{q \in \mathscr{Q}^*\colon \int_0^1 \sigma^{-1} \exp\big(-\lambda q^2(\sigma)/\sigma\big)\, d\sigma < \infty \text{ for all } \lambda > 0\Big\},$$

(1.6) $$\mathscr{Q}_k = \Big\{q \in \mathscr{Q}^*\colon q(\sigma)/\sqrt{\sigma(\log(1/\sigma))^k} \to \infty \text{ as } \sigma \downarrow 0\Big\}, \quad k \in \mathbf{N}.$$

Occosionally it will be convenient to use

(1.7) $$\mathscr{Q} = \{q \in \mathscr{Q}^*\colon (\cdot)^{-1/2} q(\cdot) \text{ non-increasing on } (0, 1]\}.$$

The (reduced multivariate) empirical process (indexed by points) is defined by

$$U_n(t) = n^{1/2}(\hat{F}_n(t)-|t|), \quad t\in[0, 1]^d, \tag{1.8}$$

where the empirical d.f. $\hat{F}_n$ is based on $X_1^{(n)}, \ldots, X_n^{(n)}$ and defined by $n\hat{F}_n(t)= = \# \{1\leq i\leq n\colon X_i^{(n)}\in R(\langle 0\rangle, t)\}$, $t\in[0, 1]^d$. It is well-known that $U_n\to_d U$, as $n\to\infty$, where U denotes the standard tied-down d-parameter Brownian motion. The so called Skorokhod construction ensures the existence of processes, equal in law to the U_n and U above and all defined on the same probability space, for which this convergence in distribution may be even replaced by almost sure convergence in the supremum norm. Without loss of generality we can and will assume that the present U_n and U are obtained from the Skorokhod construction so that we have

$$\sup_{t\in[0,1]^d} |U_n(t)-U(t)| \to_{\text{a.s.}} 0, \quad \text{as} \quad n\to\infty. \tag{1.9}$$

In view of (1.3) it will be clear that we even have

$$\sup_{R\in\mathscr{R}} \left|U_n\{R\}-U\{R\}\right| \to_{\text{a.s.}} 0, \quad \text{as} \quad n\to\infty. \tag{1.10}$$

It is the purpose of this paper to give necessary and sufficient conditions on the weight functions q and $\tilde{q}$ in order that

$$\sup_{R\in\mathscr{C}} \left|U_n\{R\}-U\{R\}\right|/q(|R|)\tilde{q}(1-|R|) \to_p 0, \quad \text{as} \quad n\to\infty, \tag{1.11}$$

where either $\mathscr{C}=\mathscr{R}_0$ (Section 2) or $\mathscr{C}\subset\mathscr{R}$ (Section 3).

Since for $R=R(\langle 0\rangle, t)\in\mathscr{R}_0$ we have $U_n\{R(\langle 0\rangle, t)\}=U_n(t)$ and $|R(\langle 0\rangle, t)|=|t|$, the random variable in (1.11) could as well be represented by means of the time points $t\in[0, 1]^d$ instead of the quadrants. More generally, a similar remark holds true for $R=R(s, t)\in\mathscr{R}$ provided we allow the time points to be of dimension $2d$. Let us write $\bar{s}=\langle\bar{s}_j\rangle=\langle 1-s_j\rangle$ and note that

$$\begin{aligned} F\{R(s, t)\} &= P(X_i^{(n)}\in R(s, t)) = \\ &= P(1-X_{i,1}^{(n)} \leq \bar{s}_1, \ldots, 1-X_{i,d}^{(n)} \leq \bar{s}_d, X_{i,1}^{(n)} \leq t_1, \ldots, X_{i,d}^{(n)} \leq t_d) \equiv \\ &\equiv \bar{F}(\bar{s}, t) = \begin{cases} |t+\bar{s}-1| = |t-s|, & \text{for} \quad s<t, \quad s, t\in[0, 1]^d, \\ 0, & if \quad s<t \quad \text{is not fulfilled;} \end{cases} \end{aligned} \tag{1.12}$$

cf. Kiefer and Wolfowitz [10]. Let $\bar{U}_n$ denote the reduced empirical process based on the vectors $(1-X_{i,1}^{(n)}, \ldots, 1-X_{i,d}^{(n)}, X_{i,1}^{(n)}, \ldots, X_{i,d}^{(n)})$ in $[0, 1]^{2d}$, for $i=1, \ldots, n$. Now it suffices for our purposes to consider

$$\frac{\bar{U}_n(\bar{s}, t)}{q(|t-s|)\tilde{q}(1-|t-s|)} \quad \text{instead of} \quad \frac{U_n\{R(s, t)\}}{q(|R(s, t)|)\tilde{q}(1-|R(s, t)|)}. \tag{1.13}$$

This will be called the point representation for rectangles.

To conclude this section we present, in the next paragraph, our basic inequality which can be found in RUYMGAART and WELLNER [14], [15]. The main results are presented in Section 2 and 3. They are derived under the assumption that the d.f. of the $X_i^{(n)}$ is uniform. We conjecture, however, that extension to the case that F has a density w.r.t. to Lebesgue measure that is bounded away from 0 and ∞ is possible. Let ψ: $[0, \infty)\to[0, \infty)$ be the decreasing function defined by

(1.14) $$\psi(\lambda) = 2\lambda^{-2}\int_0^\lambda \log(1+\sigma)\,d\sigma,\ \lambda > 0;\ \psi(0) = 1.$$

See SHORACK and WELLNER [17] for elementary properties of ψ.

Theorem 1.1 (basic inequality). *Let $R\in\mathscr{R}$ with $|R|\leqq 1/2$. Then we have*

(1.15) $$P(\sup_{S\subset R}|U_n\{S\}| \geqq \lambda) \leqq 2^{2d+4}\exp\left(\frac{-\lambda^2}{32|R|}\psi\left(\frac{\lambda}{4|R|n^{1/2}}\right)\right),\quad \lambda \geqq 0,$$

where $s\in\mathscr{R}$.

2. Weight functions for quadrants (points)

We first derive a useful inequality that should be compared with Inequality 1.1 in SHORACK and WELLNER [17]; see also RUYMGAART and WELLNER [14] (Corollary 2.3). For the proof a special countably infinite partition of $(0, 1]^d$ will be used that becomes arbitrarily fine near the lower boundary of this set. This kind of partition is motivated by O'REILLY [11]; see also SHORACK and WELLNER [17]. This partition is the collection of rectangles

(2.1) $$\mathscr{P} = \{R(\langle(1/2)^{k(j)}\rangle, \langle(1/2)^{k(j)-1}\rangle): \langle k(j)\rangle\in\mathbf{N}^d\}.$$

For any $R(a, b)\in\mathscr{P}$ we have the useful property

(2.2) $$\frac{|a|}{|b|} = \frac{(1/2)^{k(1)+\dots+k(d)}}{(1/2)^{(k(1)-1)+\dots+(k(d)-1)}} = (1/2)^d = \theta(d) = \theta\in(0, 1);$$

notice that θ is independent of the particular rectangle in the partition.

For any $0<\alpha\leqq\beta\leqq 1$ let us introduce the subclass

(2.3) $$\mathscr{P}_{\alpha,\beta} = \{R(a, b)\in\mathscr{P}: |b| \geqq \alpha, |a| < \beta\},$$

consisting of all rectangles having a non-empty intersection with the set $\{t\in[0, 1]^d: \alpha\leqq|t|\leqq\beta\}$. The inclusions

(2.4) $$\{\alpha \leqq |t| \leqq \beta\} \subset \bigcup_{R\in\mathscr{P}_{\alpha,\beta}} R \subset \{\theta\alpha \leqq |t| \leqq \beta/\theta\}$$

are immediate.

Inequality 2.1. *Let us choose any* $0<\alpha\leqq\beta\leqq\theta/2=1/2^{d+1}$. *For any* $q\in\mathcal{Q}$ *and* $\lambda\geqq 0$ *we have*

(2.5)
$$P(\sup_{\alpha\leqq|t|\leqq\beta}|U_n(t)|/q(|t|)\geqq\lambda)\leqq$$
$$\leqq 2^{3d+4}\int_{\theta\alpha}^{\beta/\theta}\frac{(\log 1/\sigma)^{d-1}}{\sigma}\exp\left(\frac{-\theta\lambda^2q^2(\sigma)}{32\sigma}\psi\left(\frac{\lambda q(\alpha)}{4\alpha n^{1/2}}\right)\right)d\sigma.$$

Proof. It follows from the monotonicity of q and from Theorem 1.1 that

(2.6)
$$P(\sup_{\alpha\leqq|t|\leqq\beta}|U_n(t)|/q(|t|)\geqq\lambda)\leqq P(\max_{R(a,b)\in\mathcal{P}_{\alpha,\beta}}\sup_{t\in R(a,b)}|U_n(t)|/q(|a|)\geqq\lambda)\leqq$$
$$\leqq\sum_{R(a,b)\in\mathcal{P}_{\alpha,\beta}}P(\sup_{t\in R(a,b)}|U_n(t)|\geqq\lambda q(|a|))\leqq$$
$$\leqq 2^{2d+4}\sum_{R(a,b)\in\mathcal{P}_{\alpha,\beta}}\exp\left(\frac{-\lambda^2q^2(|a|)}{32|b|}\psi\left(\frac{\lambda q(|a|)}{4|b|n^{1/2}}\right)\right).$$

In view of (2.2) and because $(\cdot)^{-1/2}q(\cdot)$ is non-increasing we may bound the first factor in the exponent in (2.6) below by

(2.7)
$$\lambda^2q^2(|a|)/32|b|\geqq\theta\lambda^2q^2(|t|)/32|t|,\quad\text{for}\quad t\in R(a,b).$$

Using the monotonicity of q and ψ and $q\in Q$, the second factor in the exponent in (2.6) may be bounded below by

(2.8)
$$\psi(\lambda q(|a|)/4|b|n^{1/2})\geqq\psi(\lambda q(\alpha)/4\alpha n^{1/2}),\quad\text{for}\quad R(a,b)\in\mathcal{P}_{\alpha,\beta}.$$

When we use

(2.9)
$$1=2^d/|b|\int_{R(a,b)}|dt|\leqq 2^d\int_{R(a,b)}1/|t|\,|dt|,\quad\text{for}\quad R(a,b)\in\mathcal{P},$$

at the transition from summation to integration we find, by combining (2.4), (2.6)—(2.8) that

(2.10)
$$P(\sup_{\alpha\leqq|t|\leqq\beta}|U_n(t)|/q(|t|)\geqq\lambda)\leqq$$
$$\leqq 2^{3d+4}\int_{\{\theta\alpha\leqq|t|\leqq\beta/\theta\}}\frac{1}{|t|}\exp\left(\frac{-\theta\lambda^2q^2(|t|)}{32|t|}\psi\left(\frac{\lambda q(\alpha)}{4\alpha n^{1/2}}\right)\right)|dt|.$$

To complete the proof we use the change of variables $\sigma=s_1=|t|$, $s_2=t_2$, ..., $s_d=t_d$ with Jacobian $(\prod_{j=2}^{d}s_j)^{-1}$ to compute the integral on the right hand side of (2.10). This yields as an upper bound for the right hand side of (2.10)

(2.11)
$$\int_{\theta\alpha}^{\beta/\theta}\left(\int_\sigma^1\cdots\int_\sigma^1\frac{1}{s_2\dots s_d}ds_2\dots ds_d\right)\frac{1}{\sigma}\exp\left(\frac{-\theta\lambda^2q^2(\sigma)}{32\sigma}\psi\left(\frac{\lambda q(\alpha)}{4\alpha n^{1/2}}\right)\right)d\sigma,$$

which is easily seen to be equal to the expression on the right in (2.5).

Theorem 2.1. *Let* $F(t)=|t|$, $t\in[0,1]^d$, $d\in\mathbf{N}$, *and* $q\in\mathcal{Q}^*$. *Then we have*

(2.12) $$\sup_{0\le|t|\le1}|U_n(t)-U(t)|/q(|t|)\to_p 0, \quad as \quad n\to\infty,$$

if and only if $q\in\mathcal{Q}_{d-1}$.

Proof. The theorem is well-known for $d=1$: see O'Reilly [11]. Hence we assume $d\ge2$. The notation

(2.13) $$g(\sigma)=q(\sigma)/\sqrt{\sigma(\log 1/\sigma)^{d-1}}, \quad \sigma>0,$$

will be used in both parts of the proof.

($\Leftarrow$) Suppose that $q\in\mathcal{Q}_{d-1}$. Following Shorack and Wellner [17] (p. 649) we can and will assume without loss of generality that

(2.14) $$g(\cdot)\le\sqrt{(\log 1/(\cdot))^{d-1}} \quad \text{and} \quad g\downarrow \quad \text{on} \quad (0,1] \quad (\text{hence } q\in\mathcal{Q}).$$

For any $0<\delta\le(1/2)^{d+1}$ we have

(2.15) $$\sup_{0\le|t|\le1}|U_n(t)-U(t)|/q(|t|)\le\sum_{k=1}^{5}Y_{nk},$$

where, with $\alpha_n=q^2(1/n)$, $\beta_n=(d-1)!\cdot(n(\log n)^{d-1})^{-1}$ and $\gamma\in(0,\infty)$, the r.v.'s Y_{nk} are given by

(2.16) $$Y_{n1}=\sup_{0\le|t|\le\beta_n/\gamma}|U_n(t)|/q(|t|),$$

(2.17) $$Y_{n2}=\sup_{\beta_n/\gamma\le|t|\le\alpha_n}|U_n(t)|/q(|t|),$$

(2.18) $$Y_{n3}=\sup_{\alpha_n\le|t|\le\delta}|U_n(t)|/q(|t|),$$

(2.19) $$Y_{n4}=\sup_{0\le|t|\le\delta}|U(t)|/q(|t|),$$

(2.20) $$Y_{n5}=\sup_{0\le|t|\le1}|U_n(t)-U(t)|/q(\delta).$$

It will be shown that for any $\varepsilon>0$ and each $k=1,\ldots,5$ there exist $\gamma=\gamma(\varepsilon)$, $\delta=\delta(\varepsilon)$ and $n(\varepsilon)\in\mathbf{N}$ such that

(2.21) $$P(Y_{nk}\ge\varepsilon)\le\varepsilon, \quad \text{for} \quad n\ge n(\varepsilon).$$

To show (2.21) for $k=1$ let $|X|_{1:n}=\min\{|X_1^{(n)}|,\ldots,|X_n^{(n)}|\}$. Note that $P(|X|_{1:n}\le\beta_n/\gamma)\to1-\exp(-1/\gamma)$, as $n\to\infty$, so that $P(|X|_{1:n}\le\beta_n/\gamma)\le\varepsilon$ for γ sufficiently large. Under the condition $\sup_{0\le|t|\le\beta_n/\gamma}\hat{F}_n(t)=0$, which is fulfilled with probability $\ge1-\varepsilon$ according to the remark just made, it is easy to see that

(2.22) $$Y_{n1}\le n^{1/2}\sup_{0\le|t|\le\beta_n/\gamma}|t|/q(|t|)\le n^{1/2}(\beta_n/\gamma)^{1/2}\{g(\beta_n/\gamma)(\log n)^{(d-1)/2}\}^{-1}<\varepsilon,$$

for n sufficiently large. Hence it follows that

$$P(Y_{n1} \geqq \varepsilon) \leqq P\Big(\sup_{0\leqq|t|\leqq\beta_n/\gamma} \hat{F}_n(t) > 0\Big) + \tag{2.23}$$
$$+P\Big(\sup_{0\leqq|t|\leqq\beta_n/\gamma} |U_n(t)|/q(|t|) \geqq \varepsilon \,\Big|\, \sup_{0\leqq|t|\leqq\beta_n/\gamma} \hat{F}_n(t) = 0\Big) \leqq \varepsilon,$$

for n sufficiently large.

For $k=2$ the left hand side of (2.21) is for any $\gamma_1\in(0, \infty)$ bounded above by

$$P\Big(\sup_{\beta_n/\gamma\leqq|t|\leqq\alpha_n} |U_n(t)|/|t|^{1/2} \geqq \varepsilon g(\alpha_n)(\log 1/\alpha_n)^{(d-1)/2}\Big) \leqq \tag{2.24}$$
$$\leqq P\Big(\sup_{\beta_n/\gamma\leqq|t|\leqq\alpha_n} |U_n(t)|/|t|^{1/2} \geqq \gamma_1(\log n)^{(d-1)/2}\Big),$$

for $n\geqq n_1=n_1(\gamma_1)$. Hence, applying Inequality 2.1 with $q(\cdot)=(\cdot)^{1/2}$, we see that there exist $c_1, \ldots, c_4\in(0, \infty)$ such that the last expression in (2.24) is in turn bounded above by

$$c_1(\log n)^d \exp\big(-c_2\gamma_1^2(\log n)^{d-1}\psi(c_3\gamma_1\gamma^{1/2}(\log n)^{d-1})\big) \leqq \tag{2.25}$$
$$\leqq c_1(\log n)^d \exp(-c_4\gamma_1\gamma^{-1/2}\log\log n) \leqq \varepsilon,$$

provided γ_1 and n are chosen sufficiently large.

Inequality 2.1 may be directly applied to Y_{n3} with $\alpha=\alpha_n$ and $\beta=\delta$. The integral in the resulting upper bound decreases to 0 as $\delta\downarrow 0$, since $q\in\mathcal{Q}_{d-1}$ implies that

$$\int_0^1 (1/\sigma^2)\exp\big(-\lambda q^2(\sigma)/\sigma\big)\,d\sigma < \infty, \quad \text{for all} \quad \lambda > 0; \tag{2.26}$$

see Shorack and Wellner [17], ((1.9), (1.15) and (1.26)).

According to Orey and Pruitt [12] (Theorem 2.2) the function λq is point upper class for U, for all $\lambda>0$. This yields

$$\sup_{0\leqq|t|\leqq\delta} |U(t)|/q(|t|) \to_{\text{a.s.}} 0, \quad \text{as} \quad \delta\downarrow 0, \tag{2.27}$$

which entails (2.21) for $k=4$. The validity of (2.21) for $k=5$ is immediate from (1.9).

($\Rightarrow$) Let β_n be as before. We obviously have

$$\sup_{0\leqq|t|\leqq 1} |U_n(t)-U(t)|/q(|t|) \geqq \sup_{0\leqq|t|\leqq\beta_n} |U_n(t)-U(t)|/q(|t|) = Y. \tag{2.28}$$

Using the remark below (2.21) we see that with probability larger than 1/2 we have

$$Y \geqq \Big\{n^{1/2}(n^{-1}-\beta_n) - \sup_{0\leqq|t|\leqq\beta_n} |U(t)|\Big\}/q(\beta_n) \leqq \tag{2.29}$$
$$\geqq \big(2n^{1/2}q(\beta_n)\big)^{-1} \geqq \big(3((d-1)!)^{1/2}g(\beta_n)\big)^{-1}$$

for all large n, where for the second inequality again Theorem 2.2 in Orey and Pruitt [12] is applied.

The assumption that $\sup_{0\leq|t|\leq 1} |U_n(t)-U(t)|/q(|t|)\to_p 0$, as $n\to\infty$, jointly with (2.28), (2.29) and the fact that q is nondecreasing, implies that $q\in\mathcal{Q}_{d-1}$.

Theorem 2.2. *Let* $F(t)=|t|$, $t\in[0,1]^d$, $d\in\mathbf{N}$ *and* $\tilde{q}\in\mathcal{Q}^*$. *Then we have*

$$\sup_{0\leq|t|\leq 1} |U_n(t)-U(t)|/\tilde{q}(1-|t|) \to_p 0, \quad as \quad n\to\infty, \tag{2.30}$$

if and only if $\tilde{q}\in\mathcal{Q}_0$.

Proof. Suppose $\tilde{q}\in\mathcal{Q}_0$. Starting with the equalities

$$U_n(t) = -U_n\{R(\langle 0\rangle, t)^c\} \quad \text{and} \quad U(t) = -U\{R(\langle 0\rangle, t)^c\} \tag{2.31}$$

we obtain using the union-intersection principle

$$|U_n(t)-U(t)| \leq \sum_{i\in\mathcal{I}} \left|U_n\{R_i(t)\}-U\{R_i(t)\}\right|, \tag{2.32}$$

where the $R_i(t)$'s are rectangles and $\mathcal{I}$ a finite index set. This yields

$$\sup_{0\leq|t|\leq 1} |U_n(t)-U(t)|/\tilde{q}(1-|t|) \leq \sum_{i\in\mathcal{I}} \sup_{0\leq|t|\leq 1} \left|U_n\{R_i(t)\}-U\{R_i(t)\}\right|/\tilde{q}(1-|t|). \tag{2.33}$$

It turns out to be convenient to split this sum into two parts. Define $\mathcal{I}_0$ as the set of all $i\in\mathcal{I}$ such that $R_i(t)$ is $(0,1]^{j-1}\times(t_j,1]\times(0,1]^{d-j}$ for some $1\leq j\leq d$. Write $\mathcal{I}_1=\mathcal{I}\setminus\mathcal{I}_0$. For $i\in\mathcal{I}_0$ we have

$$\sup_{0\leq|t|\leq 1} \left|U_n\{R_i(t)\}-U\{R_i(t)\}\right|/\tilde{q}(1-|t|) \leq \sup_{0\leq|t|\leq 1} \left|U_n\{R_i(t)\}-U\{R_i(t)\}\right|/\tilde{q}(|R_i(t)|). \tag{2.34}$$

Application of Theorem 2.1 with $d=1$ (the case $d=1$ is symmetrical) completes the proof for this part of the sum.

Now let $i\in\mathcal{I}_1$. Define dimension $(R_i(t))=\#\{j\colon R_i(t)$ depends on $t_j\}$. Suppose dimension $(R_i(t))=l$, $2\leq l\leq d$. By symmetry considerations, studying

$$\sup_{0\leq|t|\leq 1} \left|U_n\{R_i(t)\}-U\{R_i(t)\}\right|/\tilde{q}(1-|t|)$$

is equivalent with studying

$$\sup_{0\leq|t|\leq 1} |U_n(t')-U(t')|/\tilde{q}(1-|\langle 1\rangle-t|),$$

where t' is t restricted to $[0,1]^l$ in the way suggested above.

Define $\xi=\max_{1\leq j\leq d} t_j$. We have

$$\tilde{q}(1-|\langle 1\rangle-t|) \geq \tilde{q}(\xi), \tag{2.35}$$

and for small values of ξ

$$\tilde{q}(\xi) \geq \sqrt{\xi}, \tag{2.36}$$

because $\tilde{q}\in\mathcal{Q}_0$, using an argument similar to SHORACK and WELLNER [17] ((a) on p. 648). Define $q\in\mathcal{Q}_{l-1}$ in the following way:

$$q(\sigma) = \sup_{0\leq\tau\leq\sigma} \sqrt{\tau(\log 1/\tau)^l}.$$

Using $\xi^l\geq|t'|$, it is easy to see that

$$\sqrt{\xi} \geq \sqrt{|t'|(\log 1/|t'|)^l} = q(|t'|) \tag{2.37}$$

for small values of $|t'|$. The assertions (2.35)—(2.37) entail that

$$\sup_{t'\in[0,1]^l} |U_n(t')-U(t')|/q(|t'|) \to_p 0, \quad \text{as} \quad n\to\infty,$$

implies

$$\sup_{0\leq|t|\leq 1} |U_n(t')-U(t')|/\tilde{q}(1-|\langle 1\rangle - t|) \to_p 0, \quad \text{as} \quad n\to\infty.$$

Combining this with Theorem 2.1 completes the "if" part of the proof.

The "only if" part is clear from the "only if" part in the one-dimensional case by restricting the supremum e.g. to points of the form $t=\langle t_1, 1, ..., 1\rangle$.

Combining Theorems 2.1 and 2.2 yields

Corollary 2.1. *Let* $F(t)=|t|$, $t\in[0, 1]^d$, $d\in\mathbb{N}$ *and* $q, \tilde{q}\in\mathcal{Q}^*$. *Then we have*

$$\sup_{0\leq|t|\leq 1} |U_n(t)-U(t)|/q(|t|)\tilde{q}(1-|t|) \to_p 0, \quad as \quad n\to\infty, \tag{2.38}$$

if and only if both $q\in\mathcal{Q}_{d-1}$ *and* $\tilde{q}\in\mathcal{Q}_0$.

3. Weight functions for rectangles

Extending an example in SHORACK and WELLNER [17] to the multivariate case we have

$$\sup_{R\in\mathcal{R}} |U_n\{R\}|/q(|R|) = \infty, \quad \text{a.s.} \tag{3.1}$$

for any $q\in\mathcal{Q}^*$ with $q(0)=0$. For this reason $|R|$ should be bounded away from 0 when the growth of the empirical process for small rectangles $|R|$ is studied.

Our first goal is to obtain a suitable modification of Inequality 2.1. The special countably infinite partition of $(0, 1]^{2d}\setminus\{\bar{F}=0\}$ that will be used now becomes arbitrarily fine near the lower boundary of this set; for $d=1$ this boundary is the line segment joining (0, 1) and (1, 0). This partition cannot be written as a product of a partition of (0, 1] like (2.1), but it can be written as a product of a partition of a subset of $(0, 1]^2$, namely the set $A=\{(x, y)\in(0, 1]^2: x+y>1\}$. So we know the partition completely if we define it on A.

Let us first introduce a sequence $\bar{\mathscr{P}}_1''$, $\bar{\mathscr{P}}_2''$, ... of partitions of $(0, 1]^2$ consisting of a finite number of half-open squares. More specifically, let

(3.2) $$\bar{\mathscr{P}}_n'' = \{R(\langle(1/2)^{n+1}(k(j)-1)\rangle, \langle(1/2)^{n+1}k(j)\rangle), \langle k(j)\rangle \in \{1, \ldots, 2^{n+1}\}^2\}.$$

Let us next define recursively

(3.3)
$$\bar{\mathscr{P}}_1' = \{R\in\bar{\mathscr{P}}_1'': R\subset \{(x, y)\in(0, 1]^2: (1/2) < x+y-1 \leqq 1\}\},$$
$$\bar{\mathscr{P}}_n' = \{R\in\bar{\mathscr{P}}_n'': R\subset [\{(x, y)\in(0, 1]^2: (1/2)^n < x+y-1 \leqq 3\cdot(1/2)^n\}\setminus \bigcup_{R\in\bar{\mathscr{P}}_{n-1}'} R]\},$$
$$\text{for } n \geqq 2,$$

and finally the desired partition of A by

(3.4) $$\bar{\mathscr{P}}' = \bigcup_{n=1}^{\infty} \bar{\mathscr{P}}_n'.$$

We now obtain the partition of $(0, 1]^{2d}\setminus\{\bar{F}=0\}$ by taking the product of $\bar{\mathscr{P}}'$ taking the co-ordinates s_j and t_j together to form $(0, 1]^2$, $1\leqq j\leqq d$. Denote this partition as $\bar{\mathscr{P}}$.

For any $R(a, b)\in\bar{\mathscr{P}}$ we have the property

(3.5) $$\bar{F}(a)/\bar{F}(b) \geqq (1/2)^d = \bar{\theta}(d) = \bar{\theta}\in(0, 1).$$

Again for $0<\alpha\leqq\beta\leqq 1$ we introduce

(3.6) $$\bar{\mathscr{P}}_{\alpha,\beta} = \{R(a, b)\in\bar{\mathscr{P}}: \bar{F}(b) \geqq \alpha, \bar{F}(a) < \beta\}$$

and remark

(3.7) $$\{\alpha \leqq \bar{F}(\bar{s}, t) \leqq \beta\} \subset \bigcup_{R\in\bar{\mathscr{P}}_{\alpha,\beta}} R \subset \{\bar{\theta}\alpha \leqq \bar{F}(\bar{s}, t) \leqq \beta/\bar{\theta}\}.$$

Inequality 3.1. *Let us choose any* $0<\alpha\leqq\beta\leqq\bar{\theta}/2=(1/2)^{d+1}$. *For any* $q\in\mathscr{Q}$ *and* $\lambda\geqq 0$ *we have*

(3.8) $$P(\sup_{\alpha\leqq\bar{F}(\bar{s},t)\leqq\beta} |\bar{U}_n(\bar{s}, t)|/q(|t-s|) \geqq \lambda) \leqq$$
$$\leqq 2^{4d+4}\cdot 3^{2d} \int_{\bar{\theta}\alpha}^{\beta/\bar{\theta}} \frac{(\log 1/\sigma)^{d-1}}{\sigma^2} \exp\left(\frac{-\lambda^2\bar{\theta}}{32}\frac{q^2(\sigma)}{\sigma}\psi\left(\frac{\lambda q(\alpha)}{4\alpha n^{1/2}}\right)\right) d\sigma.$$

Proof. The same reasoning as in the proof of Inequality 2.1. yields

(3.9) $$P(\sup_{\alpha\leqq\bar{F}(\bar{s},t)\leqq\beta} |\bar{U}_n(\bar{s}, t)|/q(|t-s|) \geqq \lambda) \leqq$$
$$\leqq 2^{2d+4} \sum_{R(a,b)\in\bar{\mathscr{P}}_{\alpha,\beta}} \exp\left(\frac{-\lambda^2 q^2(\bar{F}(a))}{32\bar{F}(b)}\psi\left(\frac{\lambda q(\bar{F}(a))}{4\bar{F}(b)n^{1/2}}\right)\right).$$

In this case we have, moreover, that

$$\lambda^2 q^2(\bar F(a))/32\bar F(b) \geqq \bar\theta\lambda^2 q^2(\bar F(t))/32\bar F(t), \quad \text{for} \quad t\in R(a, b); \tag{3.10}$$

$$\psi(\lambda q(\bar F(a))/4\bar F(b)n^{1/2}) \geqq \psi(\lambda q(\alpha)/4n^{1/2}\alpha), \quad \text{for} \quad R(a, b)\in\bar{\mathcal{P}}_{\alpha,\beta}. \tag{3.11}$$

The way of construction of $\bar{\mathcal{P}}$ entails

$$1 \leqq 2^{2d}\cdot 3^{2d} \int_R (1/\bar F(t))^2\,|dt| \quad \text{for} \quad R\in\bar{\mathcal{P}}. \tag{3.12}$$

Combination of (3.7) and (3.9)—(3.12) yields

$$P(\sup_{\alpha\leqq F(\bar s, t)\leqq\beta} |\bar U_n(\bar s, t)|/q(|t-s|) \geqq \lambda) \leqq \tag{3.13}$$

$$\leqq 2^{4d+4}\cdot 3^{2d} \int_{\{\bar\theta\alpha\leqq|t-s|\leqq\beta/\bar\theta\}} (1/\bar F(\bar s, t))^2 \exp\left(\frac{-\bar\theta^2 q^2(\bar F(\bar s, t))}{32\bar F(\bar s, t)}\psi\left(\frac{\lambda q(\alpha)}{4n^{1/2}\alpha}\right)\right)|d(\bar s, t)|.$$

To complete the proof let us recall formula (1.12) for $\bar F(\bar s, t)$. The change of variables $u_j=t_j+\bar s_j-1$ and $v_j=t_j-\bar s_j$ for $1\leqq j\leqq d$, with Jacobian $(1/2)^d$, yields as upper bound for the integral in (3.13)

$$\int_{\{\bar\theta\alpha\leqq|u|\leqq\beta/\bar\theta\}} \frac{1}{|u|^2}\cdot\exp\left(\frac{-\bar\theta\lambda^2 q^2(|u|)}{32|u|}\psi\left(\frac{\lambda q(\alpha)}{4n^{1/2}\alpha}\right)\right)|du|. \tag{3.14}$$

Another change of variables, similar to the one above (2.11), completes the proof.

Theorem 3.1. *Let* $F(t)=|t|$, $t\in[0, 1]^d$, $d\in\mathbf{N}$, *and* $q\in\mathcal{Q}^*$. *For any fixed* $\gamma\in(0, \infty)$ *we have*

$$\sup_{\gamma\log n/n\leqq|R|\leqq 1} |U_n\{R\}-U\{R\}|/q(|R|) \to_p 0, \quad \text{as} \quad n\to\infty, \tag{3.15}$$

if and only if $q\in\mathcal{Q}_1$.

Proof. ($\Leftarrow$) Suppose that $q\in\mathcal{Q}_1$. Like in the proof of Theorem 2.1 the notation

$$g(\sigma) = q(\sigma)/\sqrt{\sigma\log 1/\sigma}, \quad \sigma > 0, \tag{3.16}$$

will be used. We can and will assume without loss of generality that (2.14) holds true (for q as in (3.16)) with $\sqrt{(\log 1/(\cdot))^{d-1}}$ replaced by $\sqrt{\log 1/(\cdot)}$. We have for any $0<\delta\leqq(1/2)^{d+1}$ that

$$\sup_{\gamma\log n/n\leqq|R|\leqq 1} |U_n\{R\}-U\{R\}|/q(|R|) \leqq \sum_{k=1}^{4} Z_{nk}, \tag{3.17}$$

where with $\alpha_n = q^2(n^{-1})$ and $\beta_n = \gamma \log n/n$ the r.v.'s Z_{nk} are given by

$$Z_{n1} = \sup_{\beta_n \leqq |R| \leqq \alpha_n} |U_n\{R\}|/q(|R|), \tag{3.18}$$

$$Z_{n2} = \sup_{\alpha_n \leqq |R| \leqq \delta} |U_n\{R\}|/q(|R|), \tag{3.19}$$

$$Z_{n3} = \sup_{0 \leqq |R| \leqq \delta} |U\{R\}|/q(|R|), \tag{3.20}$$

$$Z_{n4} = \sup_{0 \leqq |R| \leqq 1} |U_n\{R\} - U\{R\}|/q(\delta). \tag{3.21}$$

Again it will be shown that for any $\varepsilon > 0$ and each $k = 1, 2, 3, 4$ there exist $\delta = \delta(\varepsilon)$ and $n(\varepsilon) \in \mathbf{N}$ such that

$$P(Z_{nk} \geqq \varepsilon) \leqq \varepsilon \quad \text{for} \quad n \geqq n(\varepsilon). \tag{3.22}$$

For $k=1$ the left-hand side of (3.22) is bounded above by

$$\begin{aligned} &P\Big(\sup_{\beta_n \leqq |R| \leqq \alpha_n} |U_n\{R\}|/|R|^{1/2} \geqq \varepsilon g(\alpha_n)(\log 1/\alpha_n)^{1/2}\Big) \leqq \\ &\leqq P\Big(\sup_{\beta_n \leqq |R| \leqq \alpha_n} |U_n\{R\}|/|R|^{1/2} \geqq \gamma_1 (\log n)^{1/2}\Big) \end{aligned} \tag{3.23}$$

for $\gamma_1 \in (0, \infty)$ arbitrary and $n \geqq n_1 = n_1(\gamma_1)$. Using the point representation for rectangles we can apply Inequality 3.1. This yields the existence of $c_1, \ldots, c_4 \in (0, \infty)$ such that the last expression of (3.23) is bounded above by

(3.24)
$$c_1 \cdot n (\log n)^{d-2} \exp\left(-c_2 \gamma_1^2 \log n \psi(c_3 \gamma_1)\right) \leqq c_1 n (\log n)^{d-2} \exp\left(-c_4 \gamma_1 \log \gamma_1 \log n\right) \leqq \varepsilon,$$

provided γ_1 and n are chosen sufficiently large.

To handle Z_{n2} we can again use Inequality 3.1. The integral in the resulting upper bound decreases to 0 as $\delta \downarrow 0$ since $q \in \mathcal{Q}_1$ implies

$$\int_0^1 \frac{(\log 1/\sigma)^{d-1}}{\sigma^2} \exp\left(\frac{-\lambda q^2(\sigma)}{\sigma}\right) d\sigma < \infty, \quad \text{for all} \quad \lambda > 0, \quad d \in \mathbf{N}, \tag{3.25}$$

by a slight modification of the proof of Proposition 3.1 in Shorack and Wellner [17].

Using Theorem 2.1 in Orey and Pruitt [12] we can treat Z_{n3} in the same way as Y_{n4} in the preceding section. We also have similarity between Z_{n4} and Y_{n5} using (1.10) instead of (1.9).

($\Rightarrow$) For this half of the proof we refer to Csörgő, Csörgő, Horváth and Mason [7] (pp. 87—89) where the proof is given for the quantile process and the one-dimensional empirical process. Their proof immediately carries over to the multivariate empirical process; the generalizations of the results required in that paper can be found in Einmahl [8] (p. 2) and Pyke [13] (p. 340) respectively.

We note in passing that the analogue for rectangles of Proposition 2.1 in O'Reilly [11] can be obtained using some of the ideas in the proof of Theorem 3.1: *Let* $d\in\mathbf{N}$ *and* $q\in\mathcal{Q}^*$. *Then we have*

$$\lim_{\delta\downarrow 0}\sup_{|R|\leqq\delta}|U\{R\}|/q(|R|)=0 \quad a.s. \tag{3.26}$$

if and only if $q\in\mathcal{Q}_1$.

For any $\gamma\in(0,\infty)$, define $U_{n,\gamma}$, a process indexed by rectangles, by

$$U_{n,\gamma}\{R\}=U_n\{R\}\,1_{[\gamma\log n/n,1]}(|R|),\quad R\in\mathcal{R}. \tag{3.27}$$

Combining Theorem 3.1 and (3.26) yields

Corollary 3.1. *Let* $F(t)=|t|$, $t\in[0,1]^d$, $d\in\mathbf{N}$ *and* $q\in\mathcal{Q}^*$. *For any fixed* $\gamma\in(0,\infty)$ *we have*

$$\sup_{R\in\mathcal{R}}|U_{n,\gamma}\{R\}-U\{R\}|/q(|R|)\to_p 0,\quad as\quad n\to\infty, \tag{3.28}$$

if and only if $q\in\mathcal{Q}_1$.

Theorem 3.2. *Let* $F(t)=|t|$, $t\in[0,1]^d$, $d\in\mathbf{N}$ *and* $\tilde{q}\in\mathcal{Q}^*$. *Then we have*

$$\sup_{0\leqq|R|\leqq 1}|U_n\{R\}-U\{R\}|/\tilde{q}(1-|R|)\to_p 0,\quad as\quad n\to\infty, \tag{3.29}$$

if and only if $\tilde{q}\in\mathcal{Q}_0$.

Proof. ($\Leftarrow$) To avoid difficulties with notations and technicalities we restrict ourselves to the case $d=2$. Without any mathematical problems the proof can be extended to arbitrary d. (See also the proof of Theorem 2.2.)

Let us first remark that for $0<\delta<1$

$$\sup_{0\leqq|R|\leqq 1}|U_n\{R\}-U\{R\}|/\tilde{q}(1-|R|)\leqq \tag{3.30}$$

$$\leqq\sup_{0\leqq|R|\leqq 1}|U_n\{R\}-U\{R\}|/\tilde{q}(\delta)+\sup_{1-\delta\leqq|R|\leqq 1}|U_n\{R\}-U\{R\}|/\tilde{q}(1-|R|).$$

The first term of the last expression causes no problems, so we focus on the second term. Let us choose R with $|R|\geqq 1-\delta$ and angular points a_1, a_2, a_3, a_4 starting at the upper vertex and moving clockwise. Remark that $|a_1|\geqq 1-\delta$ and $|a_2|, |a_3|, |a_4|<\delta$. Using the inequality

$$|U_n\{R\}-U\{R\}|/\tilde{q}(1-|R|)\leqq\sum_{i=1}^{4}|U_n(a_i)-U(a_i)|/\tilde{q}(1-|R|) \tag{3.31}$$

we see that we only have to handle $\sup_{1-\delta\leqq|R|\leqq 1}|U_n(a_i)-U(a_i)|/\tilde{q}(1-|R|)$ for $i=1,2,3,4$. Using $\tilde{q}(1-|R|)\geqq\tilde{q}(1-|a_1|)$ we can apply Theorem 2.2 to handle the case $i=1$. With the same technique as used in the proof of this theorem we can also treat the cases $i=2, 3, 4$.

($\Rightarrow$) Theorem 2.2 together with the remark that (3.29) implies (2.30) yields this part of the proof.

Combining Theorem 3.1, Theorem 3.2 and Corollary 3.1 yields

Corollary 3.2. *Let* $F(t)=|t|$, $t\in[0, 1]^d$, $d\in\mathbf{N}$ *and* $q, \tilde{q}\in\mathcal{Q}^*$. *For any fixed* $\gamma\in(0, \infty)$ *the following three statements are equivalent:*

(3.32) $$\sup_{\gamma \log n/n \leqq |R| \leqq 1} \left|U_n\{R\} - U\{R\}\right|/q(|R|)\tilde{q}(1-|R|) \to_p 0, \quad as \quad n \to \infty,$$

(3.33) $$\sup_{R\in\mathcal{R}} \left|U_{n,\gamma}\{R\} - U\{R\}\right|/q(|R|)\tilde{q}(1-|R|) \to_p 0, \quad as \quad n \to \infty,$$

(3.34) $$q\in\mathcal{Q}_1 \quad and \quad \tilde{q}\in\mathcal{Q}_0.$$

Acknowledgement. We are grateful to Wim Vervaat for some helpful remarks concerning the tied-down Brownian sheet.

References

[1] K. S. Alexander, Some limit theorems for weighted and non-identically distributed empirical processes, Ph. D. dissertation, M. I. T., 1982.

[2] K. S. Alexander, The central limit theorem for empirical processes on Vapnik—Chervonenkis classes, Technical report, University of Washington, Seattle, 1984.

[3] E. A. A. Aly, J. Beirlant and L. Horváth, Strong and weak approximations of k-spacings processes, *Z. Wahrsch. Verw. Gebiete*, **66** (1984), 461—484.

[4] J. Beirlant and L. Horváth, Approximations of m-overlapping spacings processes, *Scand. J. Statist.*, to appear.

[5] D. M. Chibisov, Some theorems on the limiting behaviour of the empirical distribution function, *Sel. Transl. Math. Statist. Prob.*, **6** (1964), 147—156.

[6] M. Csörgő, S. Csörgő, L. Horváth and D. M. Mason, An asymptotic theory for empirical reliability and concentration processes, Technical Report no. 2, Lab. for research in statistics and probability, Ottawa, 1983.

[7] M. Csörgő, S. Csörgő, L. Horváth and D. M. Mason, Weighted empirical and quantile processes, in: Technical Report no. 24, Lab. for research in statistics and probability, Ottawa, 1984.

[8] J. H. J. Einmahl, A strong law for the oscillation modulus of the multivariate empirical process. II, Report 8422, Math. Inst., Kath. Univ., Nijmegen, 1984.

[9] J. H. J. Einmahl, F. H. Ruymgaart and J. A. Wellner, Criteria for weak convergence of weighted multivariate empirical processes, Report 8336, Math. Inst., Kath. Univ., Nijmegen, 1983.

[10] J. Kiefer and J. Wolfowitz, On the deviations of the empiric distribution function of vector chance variables, *Trans. Amer. Math. Soc.*, **87** (1958), 173—186.

[11] N. E. O'Reilly, On the weak convergence of empirical processes in sup-norm metrics, *Ann. Probability*, **2** (1974), 642—651.

[12] S. Orey and W. E. Pruitt, Sample functions of the N-parameter Wiener process, *Ann. Probability*, **1** (1973), 138—163.

[13] R. Pyke, Partial sums of matrix arrays and Brownian sheets, in: *Stochastic Analysis*, eds. E. F. Harding and D. G. Kendall, Wiley (New York, 1972); pp. 331—348.

[14] F. H. Ruymgaart and J. A. Wellner, Growth properties of multivariate empirical processes, Report 8202, Math. Inst., Kath. Univ., Nijmegen, 1982.

[15] F. H. Ruymgaart and J. A. Wellner, Some properties of weighted multivariate empirical processes, *Statist. Decisions*, **2** (1984), 199—223.

[16] G. R. Shorack, Weak convergence of empirical and quantile processes in sup-norm metrics via *KMT*-constructions, *Stochastic Process. Appl.*, 9 (1979), 95—98.

[17] G. R. Shorack and J. A. Wellner, Limit theorems and inequalities for the uniform empirical processes indexed by intervals, *Ann. Probability*, **10** (1982), 639—652.

(J. H. J. E.)
DEPARTMENT OF MED. INF. AND STATIST.
R.U. LIMBURG
6200 MD MAASTRICHT, THE NETHERLANDS

(F. H. R.)
DEPARTMENT OF MATHEMATICS
CATHOLIC UNIVERSITY
NIJMEGEN, THE NETHERLANDS

(J. A. W.)
DEPARTMENT OF STATISTICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195, U.S.A.

Acta Sci. Math., 52 (1988), 207—214

Large deviations of the empirical characteristic function

HEINZ-DIETER KELLER

1. Introduction. Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed (i.i.d.) random variables defined on a probability space $(\Omega, \mathfrak{A}, P)$ and taking values in $\mathbf{R}$ with common distribution function $F(x)$, $x \in \mathbf{R}$, and characteristic function

$$c(t) = \int_{\mathbf{R}} e^{itx}\, dF(x), \quad t \in \mathbf{R}.$$

The n^{th} empirical characteristic function (e.c.f.) of the sequence is

$$c_n(t) = (1/n) \sum_{j=1}^{n} e^{itX_j} = \int_{\mathbf{R}} e^{itx}\, dF_n(x), \quad t \in \mathbf{R},$$

where $F_n(x)$, $x \in \mathbf{R}$, denotes the empirical distribution function (e.d.f.) of $X_1, \ldots, X_n$. Csörgő [2], [3] and Marcus [8] gave necessary and sufficient conditions for the weak convergence of the empirical characteristic process $\sqrt{n}(c_n(t) - c(t))$ in the space of continuous complex-valued functions on a compact interval. Csörgő and Totik [4] solved the problem of consistency. The present investigation deals with the problem of large deviations of the e.c.f.. More precisely, let $S \subset \mathbf{R}$ and $T_n = \sup_{t \in S} |c_n(t) - c(t)|$, $n \in \mathrm{N}$. We shall derive asymptotic expressions for the limit

$$\lim_{n \to \infty} (1/n) \log P\{T_n \geqq \varepsilon\}, \quad \varepsilon > 0.$$

Theorems on probabilities of large deviations for related statistics are contained in the work of Groeneboom (see e.g. [6]) and many other authors, a powerful theory being available now. But such results only yield first order terms in an expansion of logarithms of large deviation probabilities, whereas our representation immediately gives higher order terms and can be used for the computation of the (relative) asymptotic Bahadur efficiency. Although some doubt exists as to the value of the concept of Bahadur efficiency, the present work was partly motivated by it (cf. [7]).

Received November 23, 1984.

2. Results. If $p\in[0,+\infty)$, $\varepsilon\in(0,1)$, let $J(0,\varepsilon)=+\infty$, $J(1-\varepsilon,\varepsilon)=-\log(1-\varepsilon)$,

$$J(p,\varepsilon)=$$

$$=\begin{cases}(p+\varepsilon)\log((p+\varepsilon)/p)+(1-p-\varepsilon)\log((1-p-\varepsilon)/(1-p)) & \text{if } 0<p<1-\varepsilon\\ +\infty & \text{if } 1-\varepsilon<p.\end{cases}$$

Lemma 1. *Suppose* $\varepsilon\in(0,1)$. *Then*

$$\varlimsup_{n\to\infty}(1/n)\log P\{\sup_{x\in\mathbf{R}}|F_n(x)-F(x)|\geqq\varepsilon\}\leqq-\min\{J(p,\varepsilon):\ 0<p\leqq1-\varepsilon\}.$$

Proof. Let $U_1, U_2, \ldots$ be a sequence of i.i.d. $U(0,1)$ random variables defined on a probability space $(\Omega^*, \mathfrak{A}^*, P^*)$. Denote the e.d.f. of the sample $U_1, \ldots, U_n$ by G_n. If $u\in[0,1]$ let $F^{-1}(u)=\inf\{x:\ F(x)\geqq u\}$. Then $F^{-1}(u)\leqq x$ if and only if $u\leqq F(x)$. X_1 and $F^{-1}(U_1)$ are identically distributed. Hence we get

$$P\{\sup_{x\in\mathbf{R}}|F_n(x)-F(x)|\geqq\varepsilon\}=P^*\{\sup_{x\in\mathbf{R}}|G_n(F(x))-F(x)|\geqq\varepsilon\}\leqq$$
$$\leqq P^*\{\sup_{0\leqq x\leqq1}|G_n(x)-x|\geqq\varepsilon\}.$$

This completes our proof since

$$\lim_{n\to\infty}(1/n)\log P^*\{\sup_{0\leqq x\leqq1}|G_n(x)-x|\geqq\varepsilon\}=-\min\{J(p,\varepsilon):\ 0<p\leqq1-\varepsilon\}$$

(cf. [6], Example 1.3.1, p. 21).

Before stating our main theorem let us introduce the random vector $Y_j(t)=(\cos(tX_j)-\operatorname{Re}c(t),\ \sin(tX_j)-\operatorname{Im}c(t))$ with its Laplace transform $M_t(\theta)=\int\exp(\langle\theta, Y_1(t)\rangle)dP$, $\theta\in\mathbf{R}^2$, for $t\in\mathbf{R}$, $j\in\mathbf{N}$. If $t\in\mathbf{R}$, $\varepsilon>0$, $\theta\in\mathbf{R}^2$, let $h_t(\varepsilon,\theta)=\inf\{\exp(-r\varepsilon)M_t(r\theta):\ r\geqq0\}$ and if $t\in\mathbf{R}$, $\varepsilon>0$, let $i_t(\varepsilon)=\log(\sup\{h_t(\varepsilon,\theta):\ \theta\in\mathbf{R}^2, \|\theta\|=1\})$. Let $C(S)$ be the space of continuous functions on S, AP the space of all almost periodic functions.

Now the following theorem holds:

Theorem 2. *Let the subset* S *be compact and let* $i(\varepsilon)=\sup\{i_t(\varepsilon):\ t\in S\}$ *for each* $\varepsilon>0$. *Then* $\lim_{n\to\infty}(1/n)\log P\{T_n\geqq\varepsilon\}=i(\varepsilon)$.

Proof. *If* $t\in S$, we get by Theorem 7 of Sethuraman [10] that

$$i_t(\varepsilon)=\lim_{n\to\infty}(1/n)\log P\{|c_n(t)-c(t)|\geqq\varepsilon\}\leqq\varliminf_{n\to\infty}(1/n)\log P\{T_n\geqq\varepsilon\}.$$

Hence $i(\varepsilon)\leqq\varliminf_{n\to\infty}(1/n)\log P\{T_n\geqq\varepsilon\}$.

Now let $k\in\mathbf{N}$ be arbitrary, and let us cover the set S by a finite number of open balls $B(k_j, 1/k)$ with center $k_j\in S$ and radius $1/k$, $1\leqq j\leqq k_*$. Writing

$$S_{n,k}=\sup\{|c_n(t)-c(t)-(c_n(t^*)-c(t^*))|:\ t, t^*\in S, |t-t^*|<1/k\},$$

we have

$$T_n \leqq \max_{1\leqq j\leqq k_*} |c_n(k_j)-c(k_j)|+S_{n,k} \quad \text{for each} \quad n\in\mathbf{N}.$$

Let $\delta\in(0,\varepsilon)$ be given. Then

$$\overline{\lim_{n\to\infty}}\,(1/n)\log P\{T_n\geqq\varepsilon\}\leqq$$

$$\leqq \overline{\lim_{n\to\infty}}\,(1/n)\log\big[P\big\{\max_{1\leqq j\leqq k_*}|c_n(k_j)-c(k_j)|+S_{n,k}\geqq\varepsilon,\ \delta>S_{n,k}\big\}+$$

$$+P\big\{\max_{1\leqq j\leqq k_*}|c_n(k_j)-c(k_j)|+S_{n,k}\geqq\varepsilon,\ S_{n,k}\geqq\delta\big\}\big]\leqq$$

$$\leqq \overline{\lim_{n\to\infty}}\,(1/n)\log\big[P\big\{\max_{1\leqq j\leqq k_*}|c_n(k_j)-c(k_j)|\geqq\varepsilon-\delta\big\}+P\{S_{n,k}\geqq\delta\}\big]\leqq$$

$$\leqq \overline{\lim_{n\to\infty}}\,(1/n)\log\big[2\cdot\max\big\{P\big\{\max_{1\leqq j\leqq k_*}|c_n(k_j)-c(k_j)|\geqq\varepsilon-\delta\big\},\ P\{S_{n,k}\geqq\delta\}\big\}\big]=$$

$$=\max\big\{\overline{\lim_{n\to\infty}}\,(1/n)\log P\big\{\max_{1\leqq j\leqq k_*}|c_n(k_j)-c(k_j)|\geqq\varepsilon-\delta\big\},\ \overline{\lim_{n\to\infty}}\,(1/n)\log P\{S_{n,k}\geqq\delta\}\big\}.$$

Looking for a bound for the first $\overline{\lim}$ in this expression, we get

$$\overline{\lim_{n\to\infty}}\,(1/n)\log P\big\{\max_{1\leqq j\leqq k_*}|c_n(k_j)-c(k_j)|\geqq\varepsilon-\delta\big\}\leqq$$

$$\leqq \overline{\lim_{n\to\infty}}\,(1/n)\log\sum_{j=1}^{k_*}P\{|c_n(k_j)-c(k_j)|\geqq\varepsilon-\delta\}\leqq\max_{1\leqq j\leqq k_*}i_{k_j}(\varepsilon-\delta)\leqq i(\varepsilon-\delta).$$

The second $\overline{\lim}$ requires some computations concerning $S_{n,k}$. Let $t, t^*\in S$, $|t-t^*|<1/k$ and $\lambda>0$ be given, where λ is a continuity point of F such that $\delta/16>1-F(\lambda)+F(-\lambda)$. Then we get

$$\big|c_n(t)-c(t)-\big(c_n(t^*)-c(t^*)\big)\big|\leqq$$

$$\leqq\Big|\int_{\{|x|>\lambda\}}(e^{itx}-e^{it^*x})\,d\big(F_n(x)-F(x)\big)\Big|+\Big|\int_{\{|x|\leqq\lambda\}}(e^{itx}-e^{it^*x})\,d\big(F_n(x)-F(x)\big)\Big|.$$

Now

$$\Big|\int_{\{|x|>\lambda\}}(e^{itx}-e^{it^*x})\,d\big(F_n(x)-F(x)\big)\Big|\leqq$$

$$\leqq\int_{\{|x|>\lambda\}}|e^{itx}-e^{it^*x}|\,dF_n(x)+\int_{\{|x|>\lambda\}}|e^{itx}-e^{it^*x}|\,dF(x)\leqq$$

$$\leqq 2|F_n(\lambda)-F(\lambda)|+2|F_n(-\lambda)-F(-\lambda)|+4\big(1-F(\lambda)+F(-\lambda)\big).$$

Let $K=\max\{|t|: t\in S\}$ and $\lambda^*=2(1+K\lambda)\lambda$. Using integration by parts,

$$\Big|\int_{-\lambda}^{+\lambda}(e^{itx}-e^{it^*x})\,d\big(F_n(x)-F(x)\big)\Big| =$$

$$= \Big|(e^{it\lambda}-e^{it^*\lambda})\big(F_n(\lambda-)-F(\lambda-)\big)-(e^{it(-\lambda)}-e^{it^*(-\lambda)})\big(F_n(-\lambda)-F(-\lambda)\big)-$$

$$-i\int_{-\lambda}^{+\lambda}(F_n(x)-F(x))(te^{itx}-t^*e^{it^*x})\,dx\Big| \leqq$$

$$\leqq 2|F_n(\lambda-)-F(\lambda)|+2|F_n(-\lambda)-F(-\lambda)|+|t-t^*|\lambda^*\cdot\sup_{x\in\mathbf{R}}|F_n(x)-F(x)| \leqq$$

$$\leqq 2|F_n(\lambda-)-F(\lambda)|+2|F_n(-\lambda)-F(-\lambda)|+(\lambda^*/k)\sup_{x\in\mathbf{R}}|F_n(x)-F(x)|.$$

Summing up

$$S_{n,k} \leqq 2|F_n(\lambda)-F(\lambda)|+2|F_n(\lambda-)-F(\lambda)|+4|F_n(-\lambda)-F(-\lambda)|+$$

$$+4\big(1-F(\lambda)+F(-\lambda)\big)+(\lambda^*/k)\sup_{x\in\mathbf{R}}|F_n(x)-F(x)|.$$

Hence $\overline{\lim}_{n\to\infty} 1/n\log P\{T_n\geqq\varepsilon\}$ is bounded by the maximum of $i(\varepsilon-\delta)$,

$$\overline{\lim_{n\to\infty}}\,(1/n)\log P\{|F_n(\lambda)-F(\lambda)|\geqq\delta/16\},\quad \overline{\lim_{n\to\infty}}\,(1/n)\log P\{|F_n(\lambda-)-F(\lambda)|\geqq\delta/8\},$$

$$\overline{\lim_{n\to\infty}}\,(1/n)\log P\{|F_n(-\lambda)-F(-\lambda)|\geqq\delta/16\}\quad\text{and}$$

$$\overline{\lim_{n\to\infty}}\,(1/n)\log P\Big\{\sup_{x\in\mathbf{R}}|F_n(x)-F(x)|\geqq(k/\lambda^*)\big(\delta/4-4(1-F(\lambda)+F(-\lambda))\big)\Big\}.$$

If we let first k and then λ tend to infinity, we get

$$\overline{\lim_{n\to\infty}}\,(1/n)\log P\{T_n\geqq\varepsilon\}\leqq i(\varepsilon-\delta).$$

This can be seen from the equality

$$\overline{\lim_{n\to\infty}}\,(1/n)\log P\{|F_n(x)-F(x)|\geqq\varepsilon\} =$$

$$= -\min\big\{J(F(x),\varepsilon),\,J\big(1-F(x),\varepsilon\big)\big\},\quad x\in\mathbf{R},\quad \varepsilon>0,$$

and Lemma 1 or directly.

Finally, $\delta\in(0,\varepsilon)$ was arbitrary. Hence i is continuous from the left by [10] Theorem 7 $\big(\mathscr{X}=C(S)\big)$ and [10] Lemma 3. This gives the result.

Example 3. If $L(X_1)=B(1,p)$, i.e. $P\{X_1=0\}=p$ and $P(X_1=1\}=1-p=:q$, some straightforward computations lead to the equality

$$i_t(\varepsilon)=\begin{cases}\max\{-J(p,\varepsilon/a_t),\,-J(q,\varepsilon/a_t)\} & \text{if}\quad \varepsilon<a_t\\ -\infty & \text{otherwise,}\end{cases}$$

where $a_t=(2(1-\cos t))^{1/2}$. Hence $i(\varepsilon)=\max\{-J(p,\varepsilon/a), \ -J(q,\varepsilon/a)\}$, if $T>0$, $S=[-T,+T]$ and $a=\max\{a_t\colon t\in S\}>\varepsilon$.

That T_n converges to zero almost surely even in the case $S=\mathbf{R}$ when F is purely discrete was pointed out by FEUERVERGER and MUREIKA [5]. We are now able to derive the corresponding large deviation generalization of Theorem 2.

Theorem 4. *Let F be purely discrete. If $S=\mathbf{R}$ and $i(\varepsilon)=\sup\{i_t(\varepsilon)\colon t\in S\}>-\infty$ for each $\varepsilon>0$, then* $\lim_{n\to\infty}(1/n)\log P\{T_n\geqq\varepsilon\}=i(\varepsilon)$.

Proof. With the same conclusion as in the proof of Theorem 2 we get $i(\varepsilon)\leqq\varliminf_{n\to\infty}(1/n)\log P\{T_n\geqq\varepsilon\}$. Now F is purely discrete. Hence there exist $N\in\mathbf{N}\cup\{+\infty\}$, $p_k\geqq 0$ and pairwise distinct $a_k\in\mathbf{R}$ with $P\{X_1=a_k\}=p_k$, $k\in\mathbf{N}$, $1\leqq k<N+1$, and $\sum_{k=1}^{N}p_k=1$. Let $\delta\in(0,\varepsilon)$, $m\in\mathbf{N}$, $m<N+1$, and $\gamma>0$ be given and let f denote the function $f(t)=\sum_{k=1}^{m}|1-e^{ita_k}|^2$, $t\in\mathbf{R}$. Since f is almost periodic, there exists an $L=L(\gamma^2)>0$ such that every interval of the real axis of length not smaller than L contains at least one ε-almost period, i.e. a number τ satisfying $|f(t+\tau)-f(t)|<\gamma^2$ for all $t\in\mathbf{R}$. Hence, if $t\in\mathbf{R}$ is fixed now, we can choose an ε-almost period from the open interval $(-t,-t+L)$. Then we get

$$|c_n(t)-c(t)|\leqq|c_n(t+\tau)-c(t+\tau)|+\big|c_n(t)-c(t)-\big(c_n(t+\tau)-c(t+\tau)\big)\big|\leqq$$
$$\leqq\sup_{t\in(0,L)}|c_n(t)-c(t)|+\big|c_n(t)-c(t)-\big(c_n(t+\tau)-c(t+\tau)\big)\big|.$$

It follows from Theorem 2 that

$$\varlimsup_{n\to\infty}(1/n)\log P\Big\{\sup_{t\in(0,L)}|c_n(t)-c(t)|\geqq\varepsilon-\delta\Big\}\leqq i(\varepsilon-\delta).$$

Next we study the second term which is a.s.

$$\Big|(1/n)\sum_{j=1}^{n}\sum_{k=1}^{N}\big((e^{ita_k}-e^{i(t+\tau)a_k})(I_{\{X_j=a_k\}}-p_k)\big)\Big|\leqq$$

$$\leqq(1/n)\sum_{j=1}^{n}\sum_{k=1}^{N}|e^{ita_k}-e^{i(t+\tau)a_k}|\cdot|I_{\{X_j=a_k\}}-p_k|\leqq$$

$$\leqq(1/n)\sum_{j=1}^{n}\Big[\sum_{k=1}^{m}|e^{ita_k}-e^{i(t+\tau)a_k}|\cdot|I_{\{X_j=a_k\}}-p_k|+2\sum_{k=m+1}^{N}|I_{\{X_j=a_k\}}-p_k|\Big]\leqq$$

$$\leqq(1/n)\sum_{j=1}^{n}\Big[\Big(f(\tau)\cdot\sum_{k=1}^{m}|I_{\{X_j=a_k\}}-p_k|^2\Big)^{1/2}+2\sum_{k=m+1}^{N}|I_{\{X_j=a_k\}}-p_k|\Big]\leqq$$

$$\leqq(1/n)\sum_{j=1}^{n}\Big[\gamma\Big(\sum_{k=1}^{m}|I_{\{X_j=a_k\}}-p_k|^2\Big)^{1/2}+2\sum_{k=m+1}^{N}|I_{\{X_j=a_k\}}-p_k|\Big]=(1/n)\sum_{j=1}^{n}Z_j,$$

where

$$Z_j = \gamma\Big(\sum_{k=1}^{m} |I_{\{X_j=a_k\}} - p_k|^2\Big)^{1/2} + 2 \sum_{k=m+1}^{N} |I_{\{X_j=a_k\}} - p_k|,$$

$1 \leqq j \leqq n$. It follows from Theorem 3.1 of BAHADUR [1] that

$$\overline{\lim_{n\to\infty}} (1/n) \log P\Big\{(1/n) \sum_{j=1}^{n} Z_j \geqq \delta\Big\} = \inf\Big\{\Big(\log \int e^{rZ_1}\, dP\Big) - r\delta \colon r \geqq 0\Big\}.$$

But now, since $t \in \mathbf{R}$ was arbitrary, the preceding inequalities lead to

$$\overline{\lim_{n\to\infty}} (1/n) \log P\{\sup_{t\in\mathbf{R}} |c_n(t) - c(t)| \geqq \varepsilon\} \leqq$$

$$\leqq \max\Big\{i(\varepsilon-\delta), \Big(\log \int e^{rZ_1}\, dP\Big) - r\delta\Big\} \quad \text{for all} \quad r \geqq 0.$$

If we let first γ converge to zero, then m tend to N, and finally r go to $+\infty$, we get

$$\overline{\lim_{n\to\infty}} (1/n) \log P\{T_n \geqq \varepsilon\} \leqq i(\varepsilon - \delta).$$

The closing step in the proof of Theorem 2 yields the desired result ($\mathscr{X} \subset AP$!).

Having Theorems 2 and 4, we are finally interested in an expansion of the limits of the logarithmic large deviation probabilities.

Lemma 5. *Let* $t \in \mathbf{R}$, $A_t = E\cos(tX_1)$, $B_t = E\sin(tX_1)$, $C_t = E\cos^2(tX_1)$, $D_t = E(\sin tX_1 \cdot \cos tX_1)$, $E_t = E\sin^2(tX_1)$ *and*

$$\sigma_t^2 = (1/2)[(C_t - A_t^2) + (E_t - B_t^2)] + [(1/4)((C_t - A_t^2) - (E_t - B_t^2))^2 + (D_t - A_t B_t)^2]^{1/2}.$$

Then $\sigma_t^2 = \sup\{\mathrm{Var}\langle\theta, Y_1(t)\rangle \colon \theta \in \mathbf{R}^2,\ \|\theta\| = 1\}$ *and*

$$i_t(\varepsilon) = -\varepsilon^2/2\sigma_t^2 + o(\varepsilon^2) \quad \textit{as} \quad \varepsilon \downarrow 0.$$

Proof. For $\theta = (\theta_1, \theta_2) \in \mathbf{R}^2$ we have

$$\mathrm{Var}\langle\theta, Y_1(t)\rangle = E\big((\theta_1(\cos(tX_1) - \mathrm{Re}\, c(t)) + \theta_2(\sin(tX_1) - \mathrm{Im}\, c(t)))^2\big).$$

Defining $a = E(\cos(tX_1) - \mathrm{Re}\, c(t))^2$, $b = E\big((\cos(tX_1) - \mathrm{Re}\, c(t))(\sin(tX_1) - \mathrm{Im}\, c(t))\big)$, $c = E(\sin(tX_1) - \mathrm{Im}\, c(t))^2$ and

$$\mathscr{B} = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

we get $\mathrm{Var}\langle\theta, Y_1(t)\rangle = \theta\mathscr{B}\theta^{\mathrm{T}}$. Hence $\sup\{\mathrm{Var}\langle\theta, Y_1(t)\rangle \colon \theta \in \mathbf{R}^2,\ \|\theta\| = 1\}$ is the greatest eigenvalue of $\mathscr{B}$, which is equal to

$$(1/2)(a+c) + [(1/4)(a-c)^2 + b^2]^{1/2}.$$

This proves the first part of our lemma, since $a = C_t - A_t^2$, $b = D_t - A_t B_t$ and $c = E_t - B_t^2$. The remaining expansion follows from Lemma 2.2 of JAMMALAMADAKA RAO [9].

Now Lemma 5 immediately yields expansions for the functions i appearing in our Theorems 2 and 4.

Theorem 6. *Let* $S\subset\mathbf{R}$ *be arbitrary,* $i(\varepsilon)=\sup\{i_t(\varepsilon):\ t\in S\}$ *and* $\sigma^2:=\sup\{\sigma_t^2:\ t\in S\}$, *where* $0<\sigma^2<+\infty$. *Then* $i(\varepsilon)=-\varepsilon^2/2\sigma^2+o(\varepsilon^2)$ *as* $\varepsilon\downarrow 0$.

Proof. Let $(\varepsilon_n)_{n\in\mathbf{N}}$ be a sequence of positive real numbers converging to zero. Then Lemma 5 yields

$$-1/2\sigma_t^2=\lim_{n\to\infty} i_t(\varepsilon_n)/\varepsilon_n^2\leqq \varliminf_{n\to\infty} i(\varepsilon_n)/\varepsilon_n^2 \quad \text{for all} \quad t\in S.$$

This implies $-1/2\sigma^2\leqq\varliminf\limits_{n\to\infty} i(\varepsilon_n)/\varepsilon_n^2$. Now

$$i_t(\varepsilon)=\log\Big(\sup\Big\{\inf_{r\geqq 0}\exp(-r\varepsilon)\int e^{r\langle\theta,Y_1(t)\rangle}\,dP:\ \theta\in\mathbf{R}^2,\ \|\theta\|=1\Big\}\Big)=$$

$$=\sup\Big\{\inf\Big\{-r\varepsilon+\log\int e^{r\langle\theta,Y_1(t)\rangle}\,dP:\ r\geqq 0\Big\}:\ \theta\in\mathbf{R}^2,\ \|\theta\|=1\Big\}\leqq$$

$$\leqq\sup\Big\{-(\varepsilon/\sigma^2)\varepsilon+\log\int\exp\big((\varepsilon/\sigma^2)\langle\theta,Y_1(t)\rangle\big)\,dP:\ \theta\in\mathbf{R}^2,\ \|\theta\|=1\Big\}$$

for all $\varepsilon>0$.

Let n be chosen large enough such that we have $\varepsilon_n<\sigma^2/4$. Then

$$\int\exp\big((\varepsilon_n/\sigma^2)\langle\theta,Y_1(t)\rangle\big)\,dP=$$

$$=1+(\varepsilon_n^2/2\sigma^4)\operatorname{Var}\langle\theta,Y_1(t)\rangle+\sum_{\nu=3}^{\infty}\frac{1}{\nu!}(\varepsilon_n/\sigma^2)^\nu\int\big(\langle\theta,Y_1(t)\rangle\big)^\nu\,dP\leqq$$

$$\leqq 1+\varepsilon_n^2/2\sigma^2+\sum_{\nu=3}^{\infty}\frac{1}{\nu!}(2\varepsilon_n/\sigma^2)^\nu\leqq 1+\varepsilon_n^2/2\sigma^2+8\varepsilon_n^3/\sigma^6,\quad \text{if}\quad \theta\in\mathbf{R}^2,\ \|\theta\|=1.$$

Since

$$i_t(\varepsilon_n)\leqq-\varepsilon_n^2/\sigma^2+\log(1+\varepsilon_n^2/2\sigma^2+8\varepsilon_n^3/\sigma^6)\leqq-\varepsilon_n^2/2\sigma^2+8\varepsilon_n^3/\sigma^6,$$

we have

$$\varlimsup_{n\to\infty} i_t(\varepsilon_n)/\varepsilon_n^2\leqq-1/2\sigma^2.$$

Combining this result with the first inequality we get the desired expansion.

Note added in proof. Theorem 2 can also be derived from [10] Theorem 2 by taking the set functions $f_{t,\theta}(x)=\theta_1\cos(tx)+\theta_2\sin(tx)$, $t\in S$, $\theta=(\theta_1,\theta_2)\in\mathbf{R}^2$, $\|\theta\|=1$, $x\in\mathbf{R}$.

References

[1] R. R. Bahadur, On the asymptotic efficiency of tests and estimates, *Sankhyā*, **22** (1960), 229—252.

[2] S. Csörgő, Limit behaviour of the empirical characteristic function, *Ann. Probability*, 9 (1981), 130—144.

[3] S. Csörgő, Multivariate empirical characteristic functions, *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **55** (1981), 203—229.

[4] S. Csörgő and V. Totik, On how long interval is the empirical characteristic function uniformly consistent?, *Acta Sci. Math.*, **45** (1983), 141—149.

[5] A. Feuerverger and R. A. Mureika, The empirical characteristic function and its applications, *Ann. Statist.*, 5 (1977), 88—97.

[6] P. Groeneboom, Large deviations and asymptotic efficiences, Mathematical Centre Tracts 118, Mathematical Centre, Amsterdam, 1980.

[7] H.-D. Keller, Einige Untersuchungen zur empirischen charakteristischen Funktion und deren Anwendungen, Dissertation, Dortmund, 1979.

[8] M. B. Marcus, Weak convergence of the empirical characteristic function, *Ann. Probability*, 9 (1981), 194—201.

[9] J. S. Rao, Bahadur efficiencies of some tests for uniformity on the circle, *Ann. Math. Statist.*, 43 (1972), 468—479.

[10] J. Sethuraman, On the probability of large deviations of families of sample means, *Ann. Math. Statist.*, **35** (1964), 1304—1316.

FACHBEREICH IV/ANGEWANDTE MATHEMATIK
TRIER UNIVERSITY
POSTFACH 3825
5500 TRIER, FEDERAL REPUBLIC OF GERMANY

Acta Sci. Math., 52 (1988), 215—230

Bibliographie

Ralph H. Abraham—Christopher D. Shaw, Dynamics — The Geometry of Behavior, Part 3: Global Behavior (The Visual Mathematics Library, 3), XI+123 pages, Aerial Press, Inc., Santa Cruz, California.

At the defence of a Ph. D. thesis on topological dynamics one of the referees criticized the author not presenting figures enough in his work. A sharp debate broke out about the question whether or not figures are necessary in articles or books on dynamics. Some people said "no" arguing that every drawing takes us in to some extent, it is in the way of the abstraction oversimplifying the circumstances. By the way, in his original work, *Mécanique analitique* Lagrange used no diagrams. Other people (including the reviewer) said that the geometrical ideas having been appeared in dynamics nowadays should be visualized in some way. Abraham's and Shaw's book shows that this purpose can be realized on a very high level. Their pictures do not restrict the abstraction, quite the contrary, they help the reader imagine and assimilate very abstract concepts and phenomena.

In talking among themselves mathematicians universally use the so called "dynamic picture technique": a picture is drawn slowly, line-by-line, along with a spoken narrative. The coordination between the phases of the picture and the narrative is very important in the process of comprehension. The book preserves the dynamics of the live presentation. If the final picture is sophisticated, the reader can find its intermediate phases with appropriate comments. A typical example is the section on the famous and mysterious Lorenz attractor, which is not so mysterious after having read and watched the section. Yes, the book has to be read and looked at alternately, and the interaction of reading and watching results a deep and quick understanding.

The book contains chapters and sections on attractors, separatrices, generic properties, structural stability, heteroclinic and homoclinic tangles, horseshoes and nontrivial recurrence. As an excellent supplement to the standard monographs in the field, it should be on the bookshelf of each student, user of mathematics or mathematician studying or teaching dynamics.

László Hatvani (Szeged)

A. N. Andrianov, Quadratic Forms and Hecke Operators (Grundlehren der mathematischen Wissenschaften, 286), XII+374 pages, Springer-Verlag, Berlin—Heidelberg—New York—London—Paris—Tokyo, 1987.

In the classical theory of quadratic forms remarkable multiplicative properties of the number of integral representations of integers by positive definite integral quadratic forms were discovered. To explain these properties, E. Hecke had introduced operators in 1973, which were named later after him. Hecke operators are classically linear operators acting on the space of modular forms of one variable. This concept may be generalized in a natural way to multivariate modular forms. Using this idea, many interesting multiplicative properties of the number of integral representations of quadratic forms of more than one variable by quadratic forms were discovered in the last 50 years.

The purpose of this book — as the author writes in the preface — is to present in the form of a self-contained text-book the contemporary state of the theory of Hecke operators on the spaces of holomorphic modular forms of integral weight (the Siegel modular forms) for congruence subgroups of integral symplectic groups.

The book is divided into five chapters. Three short appendices with the required knowledge about symmetric matrices, about quadratic spaces and about modules in quadratic fields make it complete.

The content of the book is briefly as follows. In Chapter 1 theta-series of positive definite quadratic forms are introduced and their automorphic properties are studied. Looking at all functions which satisfy similar transformations as the theta-series, the space of modular forms is defined in Chapter 2. This way makes it possible to study a lot of properties of theta-series using the nice analytic expansions of modular forms. Chapter 3 deals with Hecke rings. This concept is defined first abstractly, for pairs (Γ, S), where S is a multiplicative semigroup and Γ is a suitable subgroup of S. The special properties of the most interesting Hecke rings of the general linear groups, of the symplectic groups and of the triangular subgroup of the symplectic groups are studied in detail. Chapter 4 is devoted to the study of the multiplicative properties of the Fourier coefficients of modular forms. The most important tools to get such relations are Hecke operators, introduced also here. The last chapter deals with the action of Hecke operators on theta-series. Here, there are not proved final, general results on the multiplicative properties of the Fourier coefficients of theta-series but rather a possible way is shown to study this problem. So this book does not have a happy end, but I think, it will inspire further research on this topic.

This book is written in a clear, well readable style. I want to emphasize the few introductory sentences explaining the goal and methods before each section. I find the exercises another valuable component of the book.

This volume is designed for graduate students and researchers who wish to work in the arithmetic theory of automorphic forms.

Attila Pethő (Debrecen)

M. Berger, Geometry I—II, (Universitext), XIII+428 pages, X+406 pages, Springer-Verlag, Berlin—Heidelberg—New York—London—Paris—Tokyo, 1987.

There are a lot of books on geometry but only few of them include all part of geometry and also written clearly, using modern terminology but do not lose in the labirinth of formalism. Here is an excellent book which certainly satisfies these conditions. It is the translation of the French book "Géométrie" originally published in five volumes. The book contains the detailed discussion of classical geometries and beside this it is a unified reference source for all the subfields of geometry. The author's aim was threefold as he writes: "to emphasize the visual, or 'artistic' aspect of geometry, by using figures in abundance; to accompany each new notion with as interesting a result as possible, preferably one with a simple statement but a non-trivial proof; finally, to show that this simple-looking mathematics does not belong in a museum, that it is an everyday tool in advanced mathematical research, and that occasionally one encounters unsolved problems at even the most elementary level".

It is hopeless to give even a short summary of the material discussed, so let us mention only some of the most delicate parts which usually omitted from textbooks: the classification of crystallographic groups, the classification of regular polytopes in arbitrary dimension, Cauchy's theorem on the rigidity of convex polyhedra, the discussion of polygonal billiards, Poncelet's theorem on polygons inscribed in a conic, the Villarceau circles on the torus, Clifford parallelism, the isoperimetric inequality in arbitrary dimension, the simplicity of the orthogonal group, the theorem of Witt and Cartan-Dieudonné.

In each chapter there are a great number of exercises which are usually more difficult than those in comparable books. The solutions of the most difficult ones and other exercises can be found in the companion volume "Problems in Geometry".

This book can be used in different ways. Teachers and students can use it for introductory course and some parts of it for higher-level course. It also serves as a handbook for researchers in geometry.

J. Kincses (Szeged)

T. Beth—D. Jungnickel—H. Lenz, Design Theory, 688 pages, Cambridge University Press, London—New York—New Rochelle—Melbourne—Sydney, 1986.

The main concepts and ideas of modern Design Theory are presented in this book.

Chapter I is a general introduction to the different topics of Design Theory. This part of the book provides those algebraic, geometric and parametric properties of certain incidence structures which are important for an advanced study of them. The second chapter is concerned with the techniques of deriving necessary parametric conditions which have to be fulfilled by an incidence structure of a given type. (Some titles from this chapter: Fisher's inequality for pairwise balanced designs, symmetric designs, generalizations for Fisher's inequality.) Since it is sometimes heplful to use the group of automorphisms of a design, the Chapter III deals with the connections between groups and designs. Separated chapter is devoted to Witt designs, which have been constructed with special Steiner systems and the Mathieu groups. (These are the only known finite t-transitive permutation groups with $t>3$, except for the symmetric and alternating groups.) For those readers who are familiar with non-elementary groups theory Chapter 5 is a nice application with the higly transitive groups. Further two chapters present the difference sets and the regular symmetric designs. Chapter 8 deals with various direct constructions of designs. In Chapter 9 some important recursive reconstruction methods are developed which will be applied to mutually orthogonal Latin squares and pairwise balanced designs. The next part provides more advanced existence and non-existence results for transversal designs. Separated chapter is devoted for the proof of Wilson's main theorem concerning the existence of an $S_\lambda(2, K, v)$. In the last chapter after returning to the discussion of automorphism groups an extensive literature is presented on characterisation problems.

An extensive bibliography of about 500 titles — all quoted in the previous sections — has been included.

The reader is expected to be familiar only with basic algebra but otherwise the work is self-contained. It is suitable for advanced courses and a reference book for private study, too. The proofs of several fundamental theorems have been simplified and many advanced results are presented. Last we notice that the book achived its aim: "to provide some of the necessary mathematical background for anyone working in Communication Engineering, Optimization, Statistical Planning, Computer Science and Signal Processing".

G. Galambos (Szeged)

Béla Bollobás, Combinatorics (Set systems, hypergraphs, families of vectors, and combinatorical probability), XII+180 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sidney, 1986.

Béla Bollobás has published formerly "Graph Theory", an introductory text, and two research monographs, "Extremal Graph Theory" and "Random Graphs". "Combinatorics" is a book whose main theme is the study of subsets of finite sets.

This book is an expanded account of a first-year graduate course in combinatorics but it contains considerably more material than one could reasonably hope to cover in a one semester course, this gives the lecturer ample freedom to slant the lectures to his taste.

The contents of the book (the list of section headings) present the topics very well: 1. Notation, 2. Representing Sets, 3. Sperner Systems, 4. The Littlewood — Offord Problem, 5. Shadows, 6. Random Sets, 7. Intersecting Hypergraphs, 8. The Turán Problem, 9. Saturated Hypergraphs, 10. Well-Separated Systems, 11. Helly Families, 12. Hypergraphs with a given number of Disjoint Edges, 13. Intersecting Families, 14. Factorizing Complete Hypergraphs, 15. Weakly Saturated Hypergraphs, 16. Isoperimetric Problems, 17. The Trace of a Set System, 18. Partitioning Sets of Vectors, 19. The Four Functions Theorem, 20. Infinite Ramsey Theorem.

Generally an initial combinatorics textbook contains very little for these topics, but ones are as worthy of consideration as any, in view of their fundamental nature and elementary structure.

The sections are short summaries of the topics, with their main theorems and with elegant and beautiful proofs, those which may be called the gems of the theory.

The reader can consolidate his understanding of the material by tackling over one hundred exercises. If a researcher wants to know more about a special topic, he (or she) finds many articles on the basis of references.

Zoltán Blázsik (Szeged)

Detection of changes in random processes (Edited by L. Telksnys) Optimization Software, Inc. Publications Division, New York, 1986.

Changepoint problems have originally arisen in the context of quality control, where one typically investigates the output of a production line and would wish to signal deviation from an acceptable average output level while investigating the data. Such situations can usually be modelled by saying that we have a random process $\{X(t), 0 \leq t \leq T\}$ and we wish to detect whether the probabilistic behaviour of $\{X(t), 0 \leq t \leq \tau\}$ and $\{X(t), \tau \leq t \leq T\}$ is the same. Not surprisingly, changepoint problems have been studied by many researchers from theoretical as well as applied points of view.

The book under review is a new addition to the literature on this subject. It contains 25 papers on detection of changes in random processes. The papers give a nice summary on recent progress in the Soviet Union on these problems. The references reflect intensive activity in this field. The authors of the volume cite a lot of papers on changepoint problems published in the Soviet Union. However, they do not seem to be aware of results which have appeared in Western journals.

The translation of this collection makes results of researchers working in the Soviet Union readily available for a wider audience. This translation series of Optimization Software Inc. is a great service for the mathematical community.

Lajos Horváth (Ottawa, Canada)

Dietrich Braess, Nonlinear Approximation Theory (Springer Series in Computational Mathematics), XIV+290 pages, Springer-Verlag, Berlin—Heidelberg—New York—London—Paris—Tokyo, 1986.

The monograph is based on the lectures given by the author to fourth year students at German universities. The material of these lectures is widened by additional one so that the book is a useful text not only for students but for researchers interested in approximation theory, too.

The prerequisites consist essentially of a good basic knowledge of analysis and functional analysis.

The book has been organized so that the sections recommended mostly for researchers (so as rational approximations, exponential sums, spline functions with free nodes) are independent of each other.

Let us give a short detail of the chapters pointing out just the main topic of them.

Chapter I is a review of well-known results from the linear theory. Chapter II contains the functional analytic approach (properties of Chebyshev sets; Kolmogorov criterion for suns). Chapter III is devoted to the methods of local analysis (critical points; nonlinear approximation in Hilbert spaces; Gauss—Newton method). Chapter IV is consisting of the methods of global analysis (the uniqueness theorem for Haar manifolds; concepts of the classification of critical points). In Chapter V the rational approximation is included (existence of best approximation; Chebyshev approximation by rational functions; rational interpolation; Padé approximation and moment problems; degree of rational approximation; the computation of best rational approximation). Chapter VI is devoted to the approximation by exponential sums (existence of best approximation; interpolation). Chapter VII contains Chebyshev approximation by γ-polynomials (Descartes family; approximation by proper γ-polynomials and by extended γ-polynomials; local best approximation). An finally Chapter VIII is dealing with the approximation by spline functions with free nodes (spline functions; Chebyshev approximation by spline functions; monosplines of least L_1, L_p and L_∞ norms).

The book is pretty well organized, its style is clear. Hopefully it can certainly be a very useful text for both researchers and students.

József Németh (Szeged)

Walter Dittrich—Martin Reuter, Selected Topics in Gauge Theories (Lecture Notes in Physics, 244), 315 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986.

This volume contains a collection of lectures and seminar talks given by the authors at Tübingen University and elsewhere. The material is organized into 16 chapters which are devoted to various aspects of chiral anomalies, topological objects like instantons and skyrmions, effective actions, background field methods and other topics of current interest in gauge theories. The material is presented in an unorthodox way: standard explanations (which can be found in textbooks) are omitted to a large extent, whereas computational details are completely given. The only general prerequisite is some grounding in quantum field theory, however, to get better acquainted with the background of the topics presented here, the reader should first consult some of the references cited at the end of each chapter.

The book is particularly recommended to those who are looking for a good introduction to topological aspects and chiral anomalies in gauge theories. The manner of presentation makes it ideally suited to the needs of graduate students.

L. Gy. Fehér (Szeged)

Beno Eckmann Selecta, Edited by M. A. Knus, G. Mislin and U. Stammbach, XII+835 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1987.

The edition of 65 selected papers of Beno Eckmann is in honor of his work on the occasion of his seventieth birthday. The volume contains the representatives of his research papers. Some of his survey articles have also been included, which are exceptional in their art of presenting mathematical ideas to non-specialists. Professor Eckmann writes in his Biographical notes: "Under the wonderful guidance of Heinz Hopf I then got my doctoral thesis work. It was characteristic of Hopf's views on our science that this meant not only learning algebraic topology — then a very young field — but also getting acquainted with group theory, differential geometry, and algebra in the 'abstract' sense of the Emmy Noether school. The combination of these fields, considered at that time to be largely separated from each other, remained a constant challange during all my later work." Really, it is the characteristic feature of the fundamental results and all scientific activity of Beno Eckmann that the mentioned fields represent a unified, organically connected subject in mathematics. The most competent classification of the directions of his research can be formulated by the titles of his comments to the selected papers: Homotopy groups and fiber spaces; Continuous solutions of linear equations; Cohomology of groups; Homological algebra, transfer; Duality in homotopy theory; Duality groups, Poincaré duality.

Péter T. Nagy (Szeged)

A. T. Fomenko—D. B. Fuchs—V. L. Gutemacher, Homotopic Topology, 310 pages, Akadémiai Kiadó, Budapest, 1986.

This book is a translation of the Russian original which based on the lectures held at the Moscow University. The authors' main aim "was to dig a tunnel for the ignorant from the basic terms to the 'height of heights' — the Adams spectral sequence, and it was a lucky chance that this tunnel led through a few reefs of gold". This aim is completely fulfilled.

The first chapter contains the basic ideas of homotopy theory. First the general constructions are presented: natural group structures on the sets $\pi(X, Y)$, homotopy groups, covering spaces, fibratons and homotopy sequences and then the homotopy of CW-complexes are studied in details. The second chapter introduces the general homology theory. This is started with singular homology and cohomology of topological spaces, especially the computation of the homology groups of CW-complexes and then the connections between homology and homotopy groups are studied, namely Hurwitz's theorems are proved. The chapter ends with the obstruction theory. The third chapter deals with the construction of the spectral sequences of filtered spaces and with their applications to the calculation of homology groups. The subject of the fourth chapter is the discussion of cohomology operations. After the general constructions some particular but very important cases are presented namely the Steenrod squares and Steenrod algebras. Finally the fifth chapter is fully devoted to the Adams spectral sequence and to its applications.

The presentation of the material is clear, the proofs, even of the most abstract theorems, are as geometric as possible. The book is fully illustrated by A. Fomenko's pictures which are organic part of it. Each of them gives an intuitive insight into a complicated construction or shows the main point of a proof. The book contains also a great number of exercises which help to understand the main concepts and extend the theory.

The book is recommended to everybody interested in homotopy theory but it can be useful for researchers in topology and related fields.

J. Kincses (Szeged)

George K. Francis, A Topological Picturebook, XV+194 pages, Springer-Verlag, New York—Berlin—Heidelberg—London—Paris—Tokyo, 1987.

This book is about how to draw mathematical pictures. Many mathematicians and teachers would like to draw pictures, but they belive that they can not do it. No this book teaches everybody to draw, but gives some method how one can imagine and draw some figures in mathematics.

The author believes that: "There are some rules, based on differential geometry, which can be distilled into practical routines for 'calculating' how to draw a picture." He proves his idea using many examples from different objects of mathematics.

It is noteworthy that each chapter is a "picture story", i.e. tells a topological story matching the picture.

This very nice book with 87 illustrations is warmly recommended to all teachers of mathematics and mathematicians who would like to illustrate their lectures.

Árpád Kurusa (Szeged)

Felix R. Gantmacher, Matrizentheorie, 654 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986.

This book is the German translation of the Russian original edition appeared in 1966.

The text is divided into two parts, the fist of which (chapters 1—10) deals with general theory of matrices and the second one is devoted to special questions and applications. Chapters 1—8 give the theory of matrices in general finite dimensional vector spaces. Chapters 9 and 10 investigate special matrices, linear operators, quadratic and Hermitian forms in inner product spaces. Chapters 11—14 deal with complex symmetric, antisymmetric and orthogonal matrices, matrices with non-negative elements, regularity criteria and localization of characteristic roots. Chapter 15 presents applications of the theory of matrices for systems of linear differential equations. The last chapter is devoted to Routh—Hurwitz problem and joined questions.

The book is recommended not only to mathematicians but to every specialist interested in application of mathematics.

László Gehér (Szeged)

M. B. Green—J. H. Schwarz—E. Witten, Superstring Theory, Volume 1: Introduction, X+469 pages; **Volume 2: Loop Amplitudes, Anomalies and Phenomenology,** XII+596 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1987.

Recently there has been an enormous and even growing interest in superstring theory. No wonder, superstring theory is the most promising candidate to reconcile general relativity with quantum mechanics and to unify the fundamental interactions. There is a widely felt need for a systematic exposition of the subject. This two volume text written by outstanding experts on string theory is intended to meet this need.

Volume 1 is a self-contained introduction to string theory. It starts off with an introductory chapter in which the authors explain what string theory is, present its historical background and general philosophy concentrating on bosonic strings. The next two chapters develop the theory of a free bosonic string in detail. All the four approaches (covariant, light cone, path integral and BRST) of quantization are presented here. Chapters 4 and 5 are devoted to questions concerning worldsheet and space-time supersymmetry in string theory, i.e. the fermionic degrees of freedom are introduced. In Chapter 6 the authors describe how gauge symmetries can be introduced in string theory. This is essential to make the link with the real world. Finally, this volume contains a detailed discussion of the evaluation of scattering amplitudes in the tree approximation.

Volume 2 contains a number of topics from current research papers. Chapters 8 and 9 deal with one-loop amplitudes in bosonic string and in superstring theory respectively. A large amount of space is given to questions concerning anomalies in effective field theory. The authors investigate the emergence of effective field theory and possible mechanisms of compactification of extra dimensions. The necessary differential and algebraic geometric background material is presented in cosiderable detail in separate chapters. In the final, 16^{th} chapter the authors illustrate how the machinery of algebraic geometry can be used to understand the properties of four dimensional models obtained from

$D=10$ effective field theory via compactification. They discuss how topological formulae can fix the number of generations, the couplings and symmetries of elementary particle interactions.

The authors write in the preface: "We hope that these two volumes will be useful for a wide range of readers, ranging from those who are motivated mainly by curiosity to those who actually wish to do research on string theory." There is no doubt that this excellent book will become a standard reference on string theory. It is a need for everybody interested in this very exciting subject.

László Gy. Fehér (Szeged)

E. Hairer—S. P. Norsett—G. Wanner, Solving Ordinary Differential Equations I. Nonstiff Problems (Springer Series in Computational Mathematics, 8), XIII+480 pages, Springer-Verlag, Berlin—Heidelberg—New York—London—Paris—Tokyo, 1987.

Nowadays many mathematicians dealing with pure mathematics also have a personal computer of big capacity and efficiency on their desks. So dealing with differential equations one is strongly tempted to get or firm conjectures via computer experiments. (The most exciting problem of the last decade in the theory of dynamical systems, the chaotic behavior has been discovered by such an experiment.) This activity needs precise and fast numerical methods of solving differential equations, so there is a great interest in them among mathematicians and users of mathematics. The present monograph will satisfy these demands.

The first chapter gives a survey of the "Classical Mathematical Theory" of differential equations from Newton and Leibniz to limit cycles and strange attractors. Fortunately, it does not repeat the standard way of recalling the basic theorems, it is written markedly by numerical analysts. The reader can find existence theorems using iteration methods and Taylor series, and the very first proof of the convergence of Euler's method due to Cauchy, which has recently been discovered on fragmentary notes and was never published in Cauchy's lifetime.

The second chapter contains the one step methods, i.e. the Runge—Kutta and extrapolation methods. Besides the classical ones, the modern procedures with practical error estimation and stepsize control are presented such as Dormand and Prince formulae, the embedded Runge—Kutta methods, the newest Nyström type methods for the second order equations, etc. Special section is devoted to delay differential equations and their applications (infectious disease modelling, enzyme kinetics, population dynamics, etc.).

The third chapter is concerned with the multistep methods and general linear methods. The order, stability and convergence properties are studied. The various available codes are compared by using numerical examples.

The book is concluded by an appendix containing the FORTRAN codes of some very new effective procedures treated in the book. They can be obtained from the Authors also on IBM diskette on payment of 15 Swiss Franks.

László Hatvani (Szeged)

Arthur Jones—Alistair Gray—Robert Hutton, Manifolds and Mechanics (Australian Mathematical Society Lecture Series, 2), IV+166 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1987.

We learned from the classical texts of mechanics (see e.g. P. E. Appel's and E. T. Whittaker's books) that the motions of a holonomic system with n degrees of freedom could be described by the Lagrange equation of second kind, in which the Lagrangian function is defined and differentiable on an open set in the configuration space $\mathbf{R}^n$. However, it may often happen that a single equation defined on an open set describes the motions only locally. For example, in the case of the double plain pendulum the configuration space is a two dimensional torus, which cannot be mapped by any single one-to-one function onto an open set in $\mathbf{R}^2$. But we can find an "atlas" for the entire torus with "charts" giving coordinates only for some parts of the torus. In the other words, wanting to study the motions globally one needs the differentiable manifold technique. But the text-books based upon this approach (e.g. R. Abraham's and J. E. Marsden's or V. I. Arnold's books) demands essentially more than the standard undergraduate advanced calculus texts give. This gap has been bridged by the present excellent lecture notes.

The first part is an easy mathematical introduction, in which the reader can get acquainted with such concepts as differentiable manifold, tangent space, tangent boundle, double tangent, etc. In the second part the authors show how the theory can be used for the development of the theory of Lagrangian mechanics directly from Newton's law, and give some applications (the spherical pendulum, rigid bodies).

This well-written book is highly recommended to students, applied mathematicians and theoretical physicists as well as to mathematicians interested in applications of the modern mathematics.

László Hatvani (Szeged)

Hüseyin Kocak, Differential and Difference Equations through Computer Experiments (With Disketts Containing PHASER: An Animator/Simulator for Dynamical Systems for IBM Personal Computers), XV+224 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

Nowadays the "strange attractor" is a key word of both theoretical and applied dynamical systems. It is an attracting set of the phase space that is more complicated than an equilibrium point or a limit cycle studied by classics. And this kind of attractors has been discovered by using numerical integration of a "simple" polynomial differential equation. E. Lorenz, a mathematician-meteorologist was investigating the motion of a layer heated from below. Using a routine numerical algorithm he got a strange attractor and noticed that the solutions behaved at almost random. Despite the strong efforts of many mathematicians, most of the properties noticed have not been proved yet theoretically. So it can be understood that computer experiment is becoming a very important tool in the theory of dynamical systems, in other words, the computer is becoming the mathematician's laboratory. Kocak's book makes this tool available also for those scholars not having any programming knowledge.

The first part gives a synopsis of the facts from the theory of differential equations, difference equations and numerical methods that are prerequisite for the book. The second part is a handbook of PHASER. It should be noted here that the program is a masterpiece. Let us cite the author to describe how it works and what it does: "It is an extremely versatile and easy-to-use program, incorporating state-of-the-art software technology (menus, windows, etc.) in its user interface. The user first creats, with the help of a menu, a suitable window configuration for displaying a combination of views-phase portraits, text of equations, Poincaré sections, etc. Next, the user can specify, from another menu, various choices in preparation for numerical computations. He or she can choose, for instance, to study from a library of many dozen equations, and then compute solutions of these equations with different initial conditions or step sizes, while interactively changing parameters in the equations. From yet another menu, these solutions can be manipulated graphically. For example, the user can rotate the images, take sections, etc. During simulations, the solutions can be saved in various ways: as a hardcopy image of the screen, as a printed list, or in a form that can be reloaded into PHASER at a later time for demonstrations for further work."

The third part briefly describes the over sixty differential and difference equations stored in the permanent library of PHASER, among them the Lorenz equation, van der Pol's oscillator, Lotka—Volterra equation, Mathieu's equation, the restricted problem of three bodies on the plane. One can meet with different kinds of bifurcations, strange attractors, homoclinic orbits etc. Moreover, phaser provides a menu entry for adding new equations to the library without any programming knowledge so that each user can easily enlarge the library according to personal needs.

Summing up, this unusual book with the diskettes gives an invaluable help for using computers in teaching, research and application of differential equations.

László Hatvani—János Karsai (Szeged)

J. L. Koszul, Lectures on Fibre Bundles and Differential Geometry, (Tata Institute of Fundamental Research, Lectures on Mathematics and Physics, 20) IV + 127 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986.

The first edition in 1960 of these Lectures was one of the first explanations of the general connection theory making a significant influence on the further development of both differential geometry and the applications in mathematical physics. In the present time the wide-ranging interest of fibre bundle technique and of the notion of connections on principal and vector bundles has increased considerably and the present "classical" treatment of this modern theory can serve as a very good introduction to the differential geometric methods used in the mathematical manifold and Lie group theory and in their applications in Yang—Mills theory and in the related fields. The first two Chapters are devoted to the coordinate free differential calculus on manifolds and to the notion of differentiable bundles. In Chapters III and IV there is given the explanation of the notion of connections on principal bundles and holonomy groups. In Chapters V and VI the attention is focused on derivation laws on the associated vector bundles determined by the connection on principal bundle and to the applications in holomorphic connection theory.

Péter T. Nagy (Szeged)

J. P. LaSalle, The Stability and Control of Discrete Processes (Applied Mathematical Sciences, 62), V+150 pages, Springer-Verlag, New York—Berlin—Heidelberg—London—Paris—Tokyo, 1986.

The book is concerned with systems whose development in time can be described by difference equations. The system is observed at any integer point of time, and it is assumed that the state of the system at time $n+1$ is completely determined by its state at time n. This means, that $x(n+1)=T(x(n))$, where $x\in\mathbf{R}^m$ is the state variable and the function $T\colon \mathbf{R}^m\to\mathbf{R}^m$ is given. Therefore, if the initial state $x(0)$ is known, then the future of the system can be computed. However, not only computing problems arise. For example, if $\bar{x}$ is a periodic point or equilibrium (i.e. $T(\bar{x})=\bar{x}$) then it is important to know whether or not it is stable. This means that $x(n)$ remains arbitrarily close to $\bar{x}$ for all n if $x(0)$ is sufficiently close to $\bar{x}$. As is known, the stability theory for the continuous processes (for the differential equations) has been developed by A. M. Lyapunov. LaSalle has established the corresponding theory for difference equations. During this extension a great number of deep questions were to be solved, and the new theory is interesting and useful not only for those dealing with discrete processes but also for mathematicians interested in differential equations.

The second part of the book is devoted to the control system $(x(n+1)=Ax(n)+f(n))$, where the matrix A is given, f is the control function. This model often appears in controlling vehicles, economy, illnesses, epidemics, populations, floods, crime, manufacturing processes, etc. The book is concluded by the stabilization by feedback.

The book was published posthumously with the assistance of Kenneth Meyer, one of the students of LaSalle.

Anyway, this monograph also has the characteristic feature of every LaSalle's book and paper: it gives a very clear and plastic presentation of a sophisticated theory, which is enjoyable and useful equally for students, users of mathematics and mathematicians.

László Hatvani (Szeged)

Tamás Matolcsi, A Concept of Mathematical Physics (Models in Mechanics), 335 pages, Akadémiai Kiadó, Budapest, 1986.

This is a continuation of the author's monograph "A Concept of Mathematical Physics, Models for Space-Time" published in 1984. The notations and results of that monograph are used and referred to throughout this volume.

The author sets forward his program in the introduction: "The modelling of some sort of physical phenomena means a construction of a *category*. The objects of the category are the models and we require that there be no morphisms between the models of different physical phenomena, there be morphisms between models of similar phenomena and two models be isomorphic if and only if the modelled phenomena are physically identical."

In this book he presents mathematical models of mechanical phenomena. The models of classical and quantum mechanics (nonrelativistic and special relativistic) presented here are based on a consistent application of the basic principles of *covariance* and *relativity*. The construction of mechanical models takes up the first half of the book, the second half is devoted to mathematical tools. Among the topics touched upon in the second part of the book are the following: probability theory on subset lattices and Hilbert lattices, star algebras, elements from functional analysis and from the theory of group representations, representations of space-time groups, basic notions concerning symplectic manifolds and Poisson brackets.

In this monograph the material is treated from a uniform viewpoint of principle. This book is not an easy reading but it is well worth studying for everybody interested in its subject.

L. Gy. Fehér (Szeged)

Kazuo Murota, Systems Analysis by Graphs and Matroids, Structural Solvability and Controllability, (Algorithms and Combinatorics, Volume 3), 281 pages, Springer-Verlag, Berlin—Heidelberg—New York—London—Paris—Tokyo, 1987.

This monograph is devoted to the study of the structural analysis of a system of linear/nonlinear equations and the structural controllability of a linear time-invariant dynamical system. The outline of the contents of this book is as follows:

In the first Chapter mathematical preliminaries are given. Basic results in graph theory and matroid theory are mentioned and some useful relevant theorems as the Dulmage—Mendelsohn decomposition of bipartite graphs are shown. This chapter presents some results on the submodular functions as well.

Chapter two is devoted to a graph-theoretic method for the structural analysis of a system of equations. First the structural solvability of a system of equations is formulated. The L-decomposition and the M-decomposition of graphs are introduced in connection with Menger-type linkings, together with their applications to the hierarchial decomposition of a system of equations into smaller subsystems.

Chapter 3 presents graph-theoretic conditions to the structural controllability of a linear dynamical system expressed in the descriptor: $F \cdot dx/dt = Ax + Bu$. Some known results on controllability condition of a descriptor system are mentioned, too. Various descriptions of a dynamical system are compared from the viewpoint of structural analysis.

Physical observations are made for providing the physical basis for the more elaborate and faithful mathematical models adapted in the second half of the book. It is explained in Chapter 4 that two different kinds are to be distinguished among the nonvanishing numbers characterizing real-word systems. Algebraic implications motivate the introduction of "mixed matrix" and "physical matrix".

In Chapter 5 a matroid-theoretic method is developed for the structural analysis of a system of equations. The rank of a mixed matrix is characterized, and an efficient algorithm for computing is described. Matroidal conditions are given to the structural solvability under the refined formulation.

In the last Chapter a structural controllability of a dynamical system is investigated. The dynamical degree is characterized in connection with the independent-flow problem. Relations to other works are mentioned.

G. Galambos (Szeged)

Stefan Pokorski, Gauge Field Theories (Cambridge Monographs on Mathematical Physics), XIV+394 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1987.

This new volume in the authorative series Cambridge Monographs on Mathematical Physics deals with physical and technical aspects of gauge theories.

The author first presents an overview of the standard $SU(3)\times SU(2)\times U(1)$ model, then he gives a short introduction to (path integral formulation of) perturbative quantum field theory and Feynman rules for Yang—Mills theories. In the following there is a careful discussion of the renormalization program. Separate chapters are devoted to quantum electrodynamics, renormalization group techniques and quantum chromodynamics. The book contains a detailed examination of global and gauge symmetries and their breaking schemes. The important topics of chiral symmetry, its breaking and chiral anomalies are also treated in detail. A fair amount of space is given to questions concerning scale invariance and low energy effective Lagrangians. The last chapter contains a discussion of basic elements of supersymmetric field theory.

The author presented here an extraordinarily wealthy material on theoretical methods and computational techniques of gauge field theories underlying our present understanding of elementary particle phenomena. The book is clearly written and practically self-contained, the reader is only assumed to have some familiarity with standard quantum field theory in its canonical formulation. Consequently, this book is warmly recommended to every research worker and graduate student interested in modern developments of gauge theories.

L. Gy. Fehér (Szeged)

George Pólya, The Pólya Picture Album: Encounters of a Mathematician. Edited by G. L. Alexanderson, 160 pages, Birkhäuser, Boston—Basel, 1987.

Imagine Albert Einstein, "young and good looking, not the Einstein we usually see", and young Lisi Hurwitz, whom you don't know, playing the violin as a duet and Adolf Hurwitz whom of course you know playfully conducting with a drumstick. This is the cover photo of this most enjoyable selected personal picture album. Then picture yourself to be conducted and guided through his album by Uncle George Pólya himself, at his best humour, describing the people or the occasion you see, relating the pictures to each other, and telling stories and anecdotes most charmingly, with intelligence and wit, and with obvious fondness towards all these people even if the story has a mild edge. This is exactly what you get in this book, a guided tour through the Pólya album by the late Professor Pólya, a nice afternoon in Palo Alto, California. His words were taped and transcribed. Therefore, the advanture is very intimate. Most, but not all, of the stories from Pólya's famous lecture "Some mathematicians I have known" [*Amer. Math. Monthly* 76(1969), 746—753] are told again, some of them almost verbatim (he must have told them many times), but there are quite a number of new ones, new at least to the reviewer, like the one about the absent-mindedness of Paul Lévy, or Pringsheim's remark that "Rosenthal was just a special case of Blumenthal".

Of necessity, the book is rather Magyar. The editor's care in using proper Hungarian first names and especially in accenting without an error deserves special mention. All the more so that such a care, an elementary courtesy, seems to have died out with the generations of the Pólyas.

The nicest things are of course the pictures themselves. One notices that quite a few of the photos in the illustrated history of the International Mathematical Congresses by D. A. Albers, G. L. Alexanderson and C. Reid [Springer-Verlag, New York, 1987; a review of which is in these *Acta* 51 (1987), p. 503] were in fact taken from Pólya's album. The present book has an introduction (pp. 7—8), a really intelligent biographical sketch of Pólya by the editor (pp. 9—22), and the photos with Pólya's accompanying remarks take the pages 23—155. A useful index of names completes the album. Any decent mathematics library will want to have a copy of it. It would still be better just to leave a copy in the coffee lounge or mail room of the Department of Mathematics.

Slips of the memory make narratives more authentic. The following nice little contradiction (or is it really a contradiction?) was left without remark by the editor. Probably this was intentional, and if so, then rightly so. On page 78 Pólya says: "And here are the Nevanlinnas and myself. These pictures were taken in Switzerland the year Nevanlinna came to take Weyl's place at the ETH, when Weyl went to Princeton to the Institute." On the other hand, on page 131 he says: "Ernst Völlm, myself, and Heinz Hopf in Switzerland, 1949. Hopf had replaced Weyl at the ETH when Weyl went to the Institute at Princeton." Was Nevanlinna declined in Zürich? Did he just go there to take the place but did not like it? Or, who took Weyl's place?

Sándor Csörgő (Szeged)

Lothar Sachs, A Guide to Statistical Methods and to the Pertinent Literature. Literatur zur Angewandten Statistik, XI+212 pages, Springer-Verlag, Berlin—Heidelberg—New York—London—Paris—Tokyo, 1986.

About 5500 statistical key words and phrases are arranged in alphabetical order, a smaller portion of which is in German. To each entry reference numbers are assigned which represent 1449 papers and books from the statistical literature listed also in alphabetical order. The orientation is very much toward applications. Although the book cannot compete with recent encyclopedic works, it may prove to be useful to practicing applied statisticians and to research workers from many fields who use statistical methods as a quick and handy guide.

Sándor Csörgő (Szeged)

Robert I. Soare, Recursive Enumerable Sets and Degrees (Perspectives in Mathematical Logic), XVIII+437 pages, Springer-Verlag, Berlin—Heidelberg—New York—London—Paris—Tokyo, 1987.

The study of computable functions and computably generated (or recursively enumerable) sets of numbers goes back to the 1930's when Gödel proved his Incompleteness Theorem and Church, Gödel, Kleene, Post and Turing formulated several versions of computability. Since then recursion theory has become one of the basic parts of mathematical logic.

A classical topic, initiated by Post, deals with the classification of sets of integers into "degrees" on the basis of how difficult it is to compute them. Two sets are said to belong to the same class called degree or degree of unsolvability if they are "equally difficult to compute" and degrees are partially ordered by the relation "is more difficult to compute than". Degree theory studies this structure.

This book mainly deals with the degree theory of r.e. degrees, i.e. degrees that contain an r.e. set (and is complemented by M. Lerman: Degrees of Unsolvability (1983), also in the Omega Series).

Part A contains an introduction to recursion theory (computable functions, r.e. sets, reducibilities, complete and creative sets, the recursion theorem, the jump operator, the arithmetical hierarchy, etc.).

The latter parts contain more advanced results.

Part B describes Post's initial problem (are there more than two r.e. degrees?), the initial results of Post and Kleene—Post, simple, hypersimple, hyperhypersimple sets and the solution of Friedberg and Muchnik to Post's problem explaining the fundamental finite injury priority method.

Part C explains the infinite injury priority method and gives deeper results about the upper semi-lattice of r.e. degrees such as the Density Theorem of Sacks, the theorem of Lachlan and Yates about the existence of minimal pairs, a theorem of Lachlan about nonbranching degrees and many others. It also discusses another important structure on r.e. sets: the lattice of r.e. sets formed under inclusion, proving e.g. splitting theorems and the existence of maximal sets. The relationship between the two structures is also considered (e.g. the connections between high degrees and maximal sets). The final chapter deals with index sets, e.g. the Index Set Theorem of Yates.

Part D contains more recent results which already lead toward current research. The topics include promptly simple degrees, priority arguments even more powerful (and more complicated) than the previous ones (leading to a proof of Zachlan's Nonbounding Theorem) and Soare's theorem about the automorphisms of the lattice of r.e. sets. The last chapter contains most recent work such as the unsolvability results of Herrmann, Harrington and Shelah (without proofs) and a valuable collection of open problems.

The book, written by one of the main researchers of the field, gives a complete account of the theory of r.e. degrees. Without requiring any preliminaries, the author set up and realized the aim to "bring the reader to the frontiers of current research" which is even more to be appreciated considering the high stage of development of the field. The definitions, results and proofs are always clearly motivated and explained before the formal presentation; the proofs are described with remarkable clarity and conciseness.

The book is highly recommended to everyone interested in logic. It also provides a useful background to computer scientists, in particular to theoretical computer scientists. Reading the book, one can agree with the author who points out similarities between the beauty of this field of mathematics and the art of the Renaissance. It can be added that his book reflects this beauty.

Zoltán Fülöp—György Turán (Szeged)

Richard J. Trudeau, The Non-Euclidean Revolution, XIII+269 pages with 257 Illustrations, Birkhäuser, Boston—Basel—Stuttgart, 1987.

There is a more than 2000-year-old controversy whether the Euclidean geometry is the true description of the physical world. This philosophical and mathematical debate climaxed in the first half of the last century with the invention of the non-Euclidean geometry. As a result of this "new" geometry, from the second half of the 19th century mathematicians and scientists changed the way they viewed their subject. This was a real scientific revolution. R. J. Trudeau considers it as significant as the Copernican revolution in astronomy, the Newtonian revolution in physics or as the Darwinian revolution in biology.

According to the author's aim this book proceeds on three levels. On the first this is a book on plane geometry (both Euclidean and hyperbolic) with extra material on history and philosophy. On the second this is a book on a scientific revolution, and on the third level this book is about the possibility of significant, absolute certain knowledge about the world.

To read this very interesting and enjoyable book only a sound knowledge of high school (secondary school) geometry is needed. In the first chapter we can read on the origin of the deductive geometry, and on introduction to the axiomatic method. The second chapter deals with Euclidean geometry. The short Chapter 3, entitled Geometry and the Diamond Theory of Truth contains philosophical material. In Chapter 4 we can read about the attempts to prove or disprove Postulate 5 of Euclid. The next two chapters deal with the possibility of the non-Euclidean geometry and the hyperbolic geometry. In Chapter 7 we can read about consistency questions. The last chapter deals with the question of truth. Almost every chapter ends with exercises and notes.

This well organized material is warmly recommended to the wide mathematical community, especially to the teachers of mathematics. We share Felix Klein's view on the non-Euclidean geometry (it can be read in the Introduction written by H. S. M. Coxeter), who described it as "one of the few parts of mathematics which is talked about in wide circles, so that any teacher may be asked about it at any moment."

Lajos Klukovits (Szeged)

J. Wloka, Partial Differential Equations, XI+518 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1987.

This is the English translation of the successful textbook in German on the abstract theory of partial differential equations. A modern approach to this theory needs many sophisticated concepts and methods, so it is the cardinal problem of writing a self-contained text on it to find a good proportion of the cited and detailed prerequisites from the functional analysis. The book establishes a good balance in this respect. The reader should be familiar with the language and basic theorems of functional analysis relevant for analysis, but the less familiar material, such as the theory of Fredholm operators, Gelfand triples, abstract Green solution operators, the Schauder fixed point theorem and Bochner integral are thoroughly considered in separate sections.

The first chapter is an excellent introduction to the theory of distribution and Sobolev spaces working with the Fourier transformation. The second and third chapters give the principal part of the book. In the second chapter the Lopatinskiĭ—Šapiro condition and theorems on the index of elliptic boundary value problems are treated. It is a good choice that the L. Š. condition is formulated as an initial value problem for ordinary differential equations and not algebraically as a "covering condition". The third chapter is devoted to the strongly elliptic differential operators and the method of variations. In the fourth and fifth chapter those parabolic and hyperbolic equations are considered, respectively, for which the right-hand side, i.e. the derivatives with respect to the spatial variables is an elliptic differential operator. The sixth chapter gives a brief account on the difference method for the numerical solution of the elliptic equations and the wave equation.

This well-written book is recommended to graduate students, physicists and mathematicians interested in differential equations and mathematical physics.

László Hatvani (Szeged)

H. P. Yap, Some Topics in Graph Theory (London Mathematical Society Lecture Note Series, 108) 230 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1986.

The author of this book gave an optional course on Graph Theory to Fourth Year Honours students of the Department of Mathematics, National University of Singapore in the academic year 1982/83. This book has grown out from these lectures. It is not only suitable for using as a supplement to a course text at advanced undergraduate or postgraduate level but very useful to researchers in Graph Theory, too. The book consists of five chapters.

Each main part gives an up-to-date account of a particular topic in Graph Theory which is very active in current research. After the introduction and basic terminology the four main topics are Edge-colourings of Graphs (Chapter 2), Symmetries in Graphs (Chapter 3), Packing of Graphs (Chapter 4) and Computational Complexity of Graph Properties (Chapter 5).

In Chapter 2 after a few basic and important theorems for chromatic index Dr. Yap gives several properties of "so-called critical graphs". The author produces several methods for constructing critical graphs and counterexamples to the Critical Graph Conjecture. The main results of this chapter have been proved by Vizing, Fiorini, Yap, Gol'dberg and others.

"The investigation of symmetries of a given mathematical structure has always yield the most powerful results" wrote E. Artin. Chapter 3 studies various general properties of vertex — or edge — transitiv graphs and their automorphism groups. The author write Weiss' elegant proof of Tutte's famous theorem on S-transitiv cubic graphs. There are several theorems for Cayley graphs, and the author discusses some progress made towards the resolution of Lovász' question which asks whether or not every connected Cayley graph is Hamiltonian.

Packing of graphs is a NP-hard problem for arbitrary graphs, but for trees there exist polynomial time algorithms. The author presents several results for trees and small size graphs. The proof or disproof of Tree Packing Conjecture, Ringel's, Erdős and Sós', Bollobás and Eldridge's Conjecture wait for research workers.

A graph property is "elusive" if it cannot be found without all information of a hypothetical graph.

The connectedness and planarity are elusive properties. The main object of last chapter is to introduce a Two Person Game to tackle the problem whether or not a graph property is elusive.

Each chapter contains numerous examples, exercises and open problems for the reader.

Zoltán Blázsik (Szeged)

INDEX — TARTALOM

ACTA SCIENTIARUM MATHEMATICARUM
SZEGED (HUNGARIA), ARADI VÉRTANÚK TERE 1
On peut s'abonner à l'entreprise de commerce des livres et journaux „Kultúra" (1061 Budapest, I., Fő utca 32)

ISSN 0324-6523 Acta Univ. Szeged
ISSN 0001 6969 Acta Sci. Math.

INDEX. 26024

87-5433—Szegedi Nyomda — Felelős vezető: Surányi Tibor igazgató

Felelős szerkesztő és kiadó: Leindler László
A kézirat a nyomdába érkezett: 1987. november 11.
Megjelenés: 1988. május

Példányszám: 1000. Terjedelem: 20,3 (A/5) ív
Készült monószedéssel, íves magasnyomással,
az MSZ 5601-24 és az MSZ 5602-55 szabvány szerint