ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 39
FASC. 3—4

SZEGED, 1977

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

39. KÖTET
FASC. 3—4

SZEGED, 1977

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

$\boxed{\text{GÉZA FODOR}}$

(1927—1977)

It is a great loss to Hungarian science, to the University of Szeged, and to *Acta Scientiarum Mathematicarum*, that Professor Géza Fodor, member of the editorial board of these *Acta*, prematurely died on September 28, 1977.

He was born in Szeged on May 6, 1927. After having graduated in 1950 from the Faculty of Sciences of the University of Szeged as a high school teacher of mathematics and physics, he first worked in the Department of Theoretical Physics of the same University. From 1951 to 1954 he was an aspirant (post-graduate student) under the supervision of Professor B. Szőkefalvi-Nagy. In 1954 he got his Candidate's degree in Mathematics, in combinatorial set theory, a domain which was to remain his main interest during his lifetime. After 1954 he worked first as a research associate, and then as an associate professor, at the Chair of Professor L. Kalmár in Szeged. In 1967, after having got his Academy Doctor degree, he was appointed full professor of mathematics; he led the Department of Set Theory and Mathematical Logic from its creation in 1971.

Géza Fodor was elected corresponding member of the Hungarian Academy of Sciences in 1973. He was also awarded two state medals. During the period 1973—76 he acted as Rector of the University of Szeged, and for many years he was also a member of the Hungarian Socialist Workers' Party's University Executive Committee.

Professor Fodor had a charming personality, both colleagues and students liked him much.

The mathematical abilities of Géza Fodor were particularly impressive. He published 40 papers full of new ideas. Many of his results were pioneering work in this subject.

Fodor's most important contribution to set theory is his fundamental theorem concerning stationary sets which was published in these *Acta,* vol. **17** (1956) and which reads as follows:

For all stationary $A \subset \varkappa$ ($\varkappa > \omega$) and for all f regressive on A there exists a $\varrho < \varkappa$ such that $f^{-1}(\{\varrho\})$ is stationary as well. In terms of present day set theory this theorem states that the $\varkappa$-complete ideal of the nonstationary sets is normal. Nowadays this result appears in all textbooks, since it features in all important modern branches of set theory developed after 1960.

Professor Fodor initiated a general theory of stationary sets and formulated quite a few hard problems in this topic. It was Fodor who conjectured that for all regular $\varkappa > \omega$ every stationary subset of $\varkappa$ can be split into the union of $\varkappa$ disjoint stationary sets. (This was proved by R. M. Solovay in 1967.)

Despite his grave illness he carried on his duties to the last minute.

We cherish the memory of Géza Fodor, the scholar, the teacher, and the man.

The Editors

Acta Sci. Math., 39 (1977), 205—231

Tensor operations on characteristic functions of C_0 contractions

H. BERCOVICI and D. VOICULESCU

By the results of [14], [15] and [1] every contraction T of class C_0 acting on a separable Hilbert space is quasi-similar to a unique Jordan operator. If T has finite defect indices then its Jordan model also shares this property and B. Sz.-Nagy and C. Foiaş proved in [14] that the determinant of the characteristic function of T and of the Jordan model coincide in this case.

Also in the case of finite defect indices, from the work of E. A. Nordgren and B. Moore ([10] and [8]; cf. also [16]) it is known that the inner functions appearing in the Jordan model of T can be computed from the minors of the determinant of the characteristic function of T.

It is an immediate problem to find characterizations for the inner functions in the Jordan model of a general C_0 contraction, and to look for special characterizations in the case of weak contractions of class C_0 ([13], chapter VIII) when the characteristic function has a determinant.

Also, the determinant being a representation of the unitary group on a finite-dimensional space, more generally we may perform on the characteristic function of a contraction tensor operations of the type associated to irreducible representations of unitary groups, and ask about the properties of the operators having these functions as characteristic functions.

In the first part of this paper we consider tensor operations corresponding to irreducible representations of unitary groups applied to characteristic functions of operators of class C_0, the main result being that these operations preserve the quasisimilarity of the associated operators, provided the given operators have equal defect indices. This assertion is also adapted for the case of unequal defect indices, using impure characteristic functions.

As a corollary we characterize the inner functions in the Jordan model of a C_0 contraction by means of the smallest scalar inner multiples of the exterior

Received October 10, 1976, in augmented form March 29, 1977.

powers of the characteristic function. We also obtain estimates for the defect operator of a C_0 contraction in terms of the Jordan model.

In the second part of the paper we construct higher order algebraic adjoints of the characteristic function of a weak contraction. This enables us, using the results of the first part, to extend the above mentioned result of E. A. Nordgren and B. Moore to the case of weak contractions of class C_0.

We also prove that the determinant of the characteristic function of such a contraction is an inner function.

Using the results of the first part concerning defect operators, we prove that a C_0 contraction is a weak contraction, if and only if its Jordan model is a weak contraction. This extends a result of L. E. Isaev [5] on dissipative operators, which via Cayley transform (see [13] ch. IX) shows that a C_0 contraction with Jordan model $S(m_a)$, $m_a(\lambda)=\exp\left(-a(1+\lambda)/(1-\lambda)\right)$ $(a>0)$, is a weak contraction.

Part I

§ 1. Notation and preliminaries

1. We shall consider separable (finite or infinite dimensional) Hilbert spaces over the complex field **C**.

We shall denote by $\mathfrak{H}$, $\mathfrak{K}$, ... Hilbert spaces; $\langle .,. \rangle$ will denote the scalar product in any such space. If $\mathfrak{Y}$ is a subspace of $\mathfrak{H}$ we denote by $P_{\mathfrak{Y}}$ the orthogonal projection of $\mathfrak{H}$ onto $\mathfrak{Y}$ and by $\mathfrak{Y}^{\perp}$ or $\mathfrak{H}\ominus\mathfrak{Y}$ the orthogonal complement of $\mathfrak{Y}$. $(M)^-$ denotes the norm-closure of the subset $M\subset\mathfrak{H}$. If $\{Y_\alpha\}_{\alpha\in A}$ is a family of subsets of $\mathfrak{H}$, $\bigvee_{\alpha\in A} Y_\alpha$ will denote the closed linear span of $\bigcup_{\alpha\in A} Y_\alpha$. $X\vee Y$ will denote the closed linear span of $X\cup Y$.

If $\mathfrak{H}$ and $\mathfrak{K}$ are Hilbert spaces we shall denote by $\mathfrak{H}\otimes\mathfrak{K}$ their tensor product, which is also a Hilbert space. Recall that

$$\langle f\otimes g, f'\otimes g'\rangle = \langle f, f'\rangle \langle g, g'\rangle \quad \text{for} \quad f, g\in\mathfrak{H}, f' g'\in\mathfrak{H}'. \tag{1.1}$$

$\mathfrak{H}^{\otimes n}$ will denote the tensor product $\mathfrak{H}\otimes\mathfrak{H}\otimes\ldots\otimes\mathfrak{H}$ (n times).

We denote by $\mathscr{L}(\mathfrak{H}, \mathfrak{K})$ the linear space of all linear bounded operators $X\colon \mathfrak{H}\to\mathfrak{K}$, $\mathscr{L}(\mathfrak{H})=\mathscr{L}(\mathfrak{H}, \mathfrak{H})$. If S is any subset of $\mathscr{L}(\mathfrak{H})$, $(S)'$ denotes the commutant of S. $\mathscr{U}(\mathfrak{H})$ denotes the group of unitary operators on $\mathfrak{H}$.

If $T\in\mathscr{L}(\mathfrak{H})$, the operator $\Gamma_n(T)\in\mathscr{L}(\mathfrak{H}^{\otimes n})$ is determined by

$$\Gamma_n(T)(h_1\otimes h_2\otimes\ldots\otimes h_n) = Th_1\otimes Th_2\otimes\ldots\otimes Th_n, \quad h_j\in\mathfrak{H} \quad (1\leqq j\leqq n). \tag{1.2}$$

The map Γ_n is multiplicative, commutes with the $*$-operation and restricted to $\mathscr{U}(\mathfrak{H})$ is a unitary representation.

2. Let us recall that H^∞ is the Banach algebra of bounded analytic functions in the unit disc $D=\{z\in\mathbf{C}\,|\,|z|<1\}$. We denote by H_i^∞ the set of inner functions in H^∞, that is $m\in H_i^\infty$ if and only if m has (dt-)almost everywhere radial limits $m(e^{it})$ of modulus one. We shall abuse notation sometimes, writing $m=m'$ for two inner functions such that m/m' is a constant (of modulus one).

If $\{f_\alpha\}_{\alpha\in A}$ is a family of H^∞-functions, not all 0, we denote by $\bigwedge_{\alpha\in A} f_\alpha$ the greatest common inner divisor of the functions f_α.

Consider also the Hardy space H^2 and, for a Hilbert space $\mathfrak{H}$, the vector-valued Hardy space $H^2(\mathfrak{H})$ which can be identified with $\mathfrak{H}\otimes H^2$.

If $T\in\mathscr{L}(\mathfrak{H})$ and $S\in\mathscr{L}(H^2)$ we shall consider $T\otimes S$ as an operator on $H^2(\mathfrak{H})$. For $f\in H^\infty(\mathfrak{H}), g\in H^\infty(\mathfrak{K})$ we shall denote (somewhat ambiguously) by $f\otimes g$ the element of $H^2(\mathfrak{H}\otimes\mathfrak{K})$ defined by

$$(f\otimes g)(z)=f(z)\otimes g(z),\quad z\in D. \tag{1.3}$$

For any two Hilbert spaces $\mathfrak{H}$, $\mathfrak{K}$ the operator-valued Hardy space $H^\infty(\mathscr{L}(\mathfrak{H},\mathfrak{K}))$ is the set of all bounded, $\mathscr{L}(\mathfrak{H},\mathfrak{K})$-valued analytic functions in the unit disc.

A function $\Theta\in H^\infty(\mathscr{L}(\mathfrak{H},\mathfrak{K}))$ is contractive if $\|\Theta(z)\|\leq 1, z\in D$. Any function $\Theta\in H^\infty(\mathscr{L}(\mathfrak{H},\mathfrak{K}))$ may be considered as an element of $\mathscr{L}(H^2(\mathfrak{H}), H^2(\mathfrak{K}))$ that commutes with scalar H^∞-multiplications.

We say that two functions

$$\Theta_i\in H^\infty(\mathscr{L}(\mathfrak{H}_i,\mathfrak{K}_i))\quad (i=1,2)$$

coincide if there are unitary operators $U\colon \mathfrak{H}_1\to\mathfrak{H}_2$, $V\colon \mathfrak{K}_1\to\mathfrak{K}_2$ such that $\Theta_2(\lambda)U= =V\Theta_1(\lambda)$ for all $\lambda\in D$.

A function $\Theta\in H^\infty(\mathscr{L}(\mathfrak{H},\mathfrak{K}))$ is *inner* if it is isometric as an element of $\mathscr{L}(H^2(\mathfrak{H}), H^2(\mathfrak{K}))$. Θ is $*$-*inner* if the function $\Theta^\sim$ defined by

$$\Theta^\sim(z)=\Theta(\bar z)^*,\quad z\in D \tag{1.4}$$

is inner. Θ is *two-sided inner* if it is simultaneously inner and $*$-inner. We denote by $H_i^\infty(\mathscr{L}(\mathfrak{H},\mathfrak{K}))$ the set of two-sided inner functions in $H^\infty(\mathscr{L}(\mathfrak{H},\mathfrak{K}))$.

For any $\Theta\in H^\infty(\mathscr{L}(\mathfrak{H}))$ we denote by $\Gamma_n(\Theta)$ the element of $H^\infty(\mathscr{L}(\mathfrak{H}^{\otimes n}))$ defined by

$$(\Gamma_n(\Theta))(z)=\Gamma_n(\Theta(z)),\quad z\in D. \tag{1.5}$$

If $\Theta\in H_i^\infty(\mathscr{L}(\mathfrak{H}))$ then $\Gamma_n(\Theta)\in H_i^\infty(\mathscr{L}(\mathfrak{H}^{\otimes n}))$.

3. For any $\Theta\in H_i^\infty(\mathscr{L}(\mathfrak{H}))$ we define $S(\Theta)$ as the operator acting on

$$\mathfrak{H}(\Theta)=H^2(\mathfrak{H})\ominus\Theta H^2(\mathfrak{H}) \tag{1.6}$$

and defined by

$$(S(\Theta)^*u)(z)=z^{-1}(u(z)-u(0)),\quad z\in D,\quad u\in\mathfrak{H}(\Theta). \tag{1.7}$$

If Θ is pure then it coincides with the characteristic function of $S(\Theta)$ and in this case dim $\mathfrak{H}$ equals the defect indices of $S(\Theta)$ [13]. Recall that, for a contraction $T\in\mathcal{L}(\mathfrak{K})$, the *defect operators* are $D_T=(I-T^*T)^{1/2}$, $D_{T^*}=(I-TT^*)^{1/2}$ and the *defect indices* $\mathfrak{d}_T$, $\mathfrak{d}_{T^*}$ are the ranks of D_T and D_{T^*}, respectively.

Let μ_T denote the *multiplicity* of T, i.e. the least cardinal of cyclic sets for T.

We shall need the lifting of commutants theorem of [13] in the following form.

If $\Theta\in H_i^\infty(\mathcal{L}(\mathfrak{H}))$, $\Theta'\in H_i^\infty(\mathcal{L}(\mathfrak{H}'))$ and $X\in\mathcal{L}(\mathfrak{H}(\Theta), \mathfrak{H}(\Theta'))$ satisfy the relation

$$S(\Theta')X = XS(\Theta)$$

then there is an $A\in H^\infty(\mathcal{L}(\mathfrak{H}, \mathfrak{H}'))$ such that

$$A\Theta H^2(\mathfrak{H}) \subset \Theta' H^2(\mathfrak{H}') \quad \text{and} \tag{1.8}$$

$$Xh = P_{\mathfrak{H}(\Theta')}Ah, \quad h\in\mathfrak{H}(\Theta). \tag{1.9}$$

The operator X is one-to-one if and only if, for $h\in H^2(\mathfrak{H})$,

$$h\in\Theta H^2(\mathfrak{H}) \Leftrightarrow Ah\in\Theta' H^2(\mathfrak{H}'), \tag{1.10}$$

and has dense range if and only if

$$AH^2(\mathfrak{H}) \vee \Theta' H^2(\mathfrak{H}') = H^2(\mathfrak{H}'). \tag{1.11}$$

Let us recall that X is called a *quasi-affinity* if it is one-to-one and has dense range.

The operator $S(\Theta)$ is of class C_0 if and only if Θ has a scalar multiple, that is, if

$$\Theta H^2(\mathfrak{H}) \supset mH^2(\mathfrak{H}) \tag{1.12}$$

for some $m\in H_i^\infty$. The minimal function of $T=S(\Theta)$ is then the greatest common inner divisor m_T of the functions m satisfying (1.12) [13].

A Jordan operator is an operator $S(\Theta)$ determined by a function of the form

$$\Theta = \begin{bmatrix} m_1 & & 0 \\ & m_2 & \\ & & \cdot \\ 0 & & & \cdot \\ & & & & \cdot \end{bmatrix}$$

where $m_j\in H_i^\infty$ and m_{j+1} divides m_j for each j. We shall denote it also by $S(M)$, $M=\{m_j\}_{j=1}^\infty$. By the results of [14], [15], [1] every C_0 contraction acting on a separable Hilbert space is quasisimilar to a unique Jordan model $S(M)$.

4. For a finite group G we shall denote by $C^*(G)$ the C^*-algebra of G [2], and by $\hat{G}$ the set of all (equivalence classes) of irreducible unitary representations of G. The elements of $C^*(G)$ will be written in the form $\sum_{g\in G} c_g\, g$ where $c_g\in\mathbf{C}$, so that for any unitary representation π of G the associated representation of $C^*(G)$ is given by

$$\pi\left(\sum_{g\in G} c_g g\right) = \sum_{g\in G} c_g \pi(g).$$

Let $\mathfrak{S}_n$ be the group of permutations of the set $\{1, 2, ..., n\}$. The group $\mathfrak{S}_{n-1}$ will be identified with the subgroup of $\mathfrak{S}_n$ consisting of those permutations of $\mathfrak{S}_n$ that leave n fixed and $C^*(\mathfrak{S}_{n-1})$ will be considered as a sub-algebra of $C^*(\mathfrak{S}_n)$.

$\hat{\mathfrak{S}}_n$ is known to be in one-to-one correspondence with signatures $\tau=(t_1\geqq\ldots\geqq t_n)$, t_j non-negative integers, $\sum_{j=1}^{n} t_j=n$, and the corresponding minimal central projections p_τ of $C^*(\mathfrak{S}_n)$ are given by the central Young symmetrizers [18], [6], [9]. It is known [17], Ch. V, § 18, that an irreducible representation of signature $\tau=(t_1\geqq t_2\geqq\ldots\geqq t_n)$ restricted to $\mathfrak{S}_{n-1}$ contains the irreducible representation of signature $\tau'=(t_1'\geqq t_2'\geqq\ldots$ $\ldots\geqq t_{n-1}')$ if and only if

$$t_1 \geqq t_1' \geqq t_2 \geqq t_2' \geqq \ldots \geqq t_{n-1} \geqq t_{n-1}' \geqq t_n \tag{1.13}$$

(this will be written $\tau'\prec\tau$) and that the multiplicity of τ' is one in this case.

Consider now a Hilbert space $\mathfrak{K}$. On $\mathfrak{K}^{\otimes n}$ there is a unitary representation π_n of $\mathfrak{S}_n$ given by

$$\pi_n(\sigma)(k_1\otimes\ldots\otimes k_n) = k_{\sigma^{-1}(1)}\otimes\ldots\otimes k_{\sigma^{-1}(n)}, \quad \sigma\in\mathfrak{S}_n. \tag{1.14}$$

By one of the basic results of HERMANN WEYL ([18], [6], see also [11], [7] for the adaptation to the case when $\dim\mathfrak{K}$ is infinite) we have

$$(\Gamma_n(\mathscr{U}(\mathfrak{K})))' = (\Gamma_n(\mathscr{L}(\mathfrak{K})))' = \pi_n(C^*(\mathfrak{S}_n)). \tag{1.15}$$

The irreducible representations of $\mathscr{U}(\mathfrak{K})$ which will be considered are also labelled by signatures, so we shall first make a convention. A signature will be a decreasing sequence $\tau=(t_1\geqq t_2\geqq\ldots)$ of nonnegative integers, of finite or infinite length $l(\tau)$. By $\iota(\tau)$ we shall denote the number of nonzero elements among the t_j's and $|\tau|$ will stand for $\sum_{j=1}^{l(\tau)} t_j$.

Thus for instance the set $\hat{\mathfrak{S}}_n$ is in a one-to-one correspondence with those signatures τ for which $l(\tau)=|\tau|=n$. Two signatures $\tau=(t_1\geqq t_2\geqq\ldots)$ and $\tau'=$ $=(t_1'\geqq t_2'\geqq\ldots)$ are *essentially equivalent* if $\iota(\tau)=\iota(\tau')$ and $t_j=t_j'$ for $j=1, 2, ..., \iota(\tau)$.

For a signature τ with $l(\tau)=\dim\mathfrak{K}$, $|\tau|<\infty$, there corresponds an irreducible representation ϱ_τ of $\mathscr{U}(\mathfrak{K})$ on a Hilbert space $\mathfrak{K}^\tau$ (these are the irreducible representations of "positive" signatures; cf. [18], [6] for the case $\dim\mathfrak{K}<\infty$ and [11] for the extension to the case $\dim\mathfrak{K}=\infty$).

The representation ϱ_τ can be defined as follows: consider $\tilde{\tau}$, the signature of length $|\tau|$ essentially equivalent to τ, and let $q_{\tilde{\tau}}$ be any minimal projection in $C^*(\mathfrak{S}_{|\tau|})$ such that $q_{\tilde{\tau}}\leqq p_{\tilde{\tau}}$. Then ϱ_τ is defined as the restriction of $\Gamma_{|\tau|}$ to $\pi_{|\tau|}(q_{\tilde{\tau}})\mathfrak{K}^{\otimes|\tau|}$. Clearly

ϱ_τ extends to a multiplicative homomorphism of the multiplicative semigroup $\mathscr{L}(\mathfrak{K})$ which is holomorphic. Also clearly the restriction of $\Gamma_{|\tau|}$ to $\pi_{|\tau|}(p_\tau)\mathfrak{K}^{\otimes|\tau|}$ is a finite multiple of ϱ_τ.

Another classical fact we need is that for τ with $l(\tau)=|\tau|=n$ we have $\pi_n(p_\tau)\neq 0$ if and only if $\lambda(\tau)\leqq \dim \mathfrak{K}$.

§ 2. Tensor operations on operator-valued functions

Let $\mathfrak{K}$ be a Hilbert space. For any $k\in\mathfrak{K}$ we shall consider the map $T_k\colon \mathfrak{K}^{\otimes n}\to\mathfrak{K}^{\otimes(n+1)}$ defined by

$$T_k(k_1\otimes k_2\otimes\ldots\otimes k_n) = k_1\otimes\ldots\otimes k_n\otimes k. \tag{2.1}$$

Clearly T_k is proportional to an isometry and

$$T_k^*(k_1\otimes\ldots\otimes k_{n+1}) = \langle k_{n+1}, k\rangle k_1\otimes\ldots\otimes k_n. \tag{2.2}$$

Lemma 2.1. *Consider two signatures $\tau'\prec\tau$, $l(\tau')=|\tau'|=n$, $l(\tau)=|\tau|=n+1$ such that $\lambda(\tau)\leqq\dim\mathfrak{K}$. Then we have:*

$$\bigvee_{k\in\mathfrak{K}} \pi_n(p_{\tau'})T_k^*\pi_{n+1}(p_\tau)\quad \mathfrak{K}^{\otimes(n+1)} = \pi_n(p_{\tau'})\mathfrak{K}^{\otimes n}. \tag{2.3}$$

Proof. Let us denote by $\mathfrak{F}$ the space on the left hand side of (2.3). Then $\mathfrak{F}$ is $\pi_n(\mathfrak{S}_n)$-invariant and $\Gamma_n(\mathscr{U}(\mathfrak{K}))$-invariant.

Indeed, for $\sigma\in\mathfrak{S}_n$ we have

$$\pi_n(p_{\tau'})T_k^*\pi_{n+1}(p_\tau)\pi_{n+1}(\sigma) = \pi_n(\sigma)\pi_n(p_{\tau'})T_k^*\pi_{n+1}(p_\tau),$$

since $p_{\tau'}, p_\tau$ commute with $C^*(\mathfrak{S}_n)$ and $T_k\pi_n(\sigma)=\pi_{n+1}(\sigma)T_k$. Also,

$$\mathfrak{K}^{\otimes n}\ominus\mathfrak{F} = \bigcap_{k\in\mathfrak{K}} \operatorname{Ker}[\pi_{n+1}(p_\tau)T_k\pi_n(p_{\tau'})]$$

and for any $U\in\mathscr{U}(\mathfrak{K})$ we have

$$\Gamma_n(U)\operatorname{Ker}[\pi_{n+1}(p_\tau)T_k\pi_n(p_{\tau'})] = \operatorname{Ker}[\pi_{n+1}(p_\tau)T_{Uk}\pi_n(p_{\tau'})]$$

so that $\mathfrak{K}^{\otimes n}\ominus\mathfrak{F}$ is invariant for $\Gamma_n(\mathscr{U}(\mathfrak{K}))$ and hence so is $\mathfrak{F}$.

Therefore $P_{\mathfrak{F}}\in(\pi_n(C^*(\mathfrak{S}_n))\cup\Gamma_n(\mathscr{U}(\mathfrak{K})))'$ and $P_{\mathfrak{F}}\leqq\pi_n(p_{\tau'})$. Hence by Hermann Weyl's theorem and because of the minimality of $p_{\tau'}$ in the center of $C^*(\mathfrak{S}_n)$ either $P_{\mathfrak{F}}=0$ or $P_{\mathfrak{F}}=\pi_n(p_{\tau'})$. So it will be sufficient to prove that $\mathfrak{F}\neq\{0\}$.

Observe that $\pi_n(p_{\tau'})T_k^*\pi_{n+1}(p_\tau)=T_k^*\pi_{n+1}(p_{\tau'}p_\tau)$. On the other hand, p_τ is the central support of $p_{\tau'}p_\tau$ in $C^*(\mathfrak{S}_{n+1})$ as explained in the next paragraph. Thus, from $\pi_{n+1}(p_\tau)\neq 0$ we infer $\pi_{n+1}(p_{\tau'}p_\tau)\neq 0$. Now $\bigcap_{k\in K}\operatorname{Ker} T_k^*=\{0\}$ so we can find $k\in\mathfrak{K}$ such that $T_k^*\pi_{n+1}(p_{\tau'}p_\tau)\neq 0$.

If ϱ is an irreducible representation of the finite-dimensional C^*-algebra A, there is a minimal central projection p of A such that $\ker \varrho=(1-p)A$. Let $A_1\subset A_2$ be finite dimensional C^*-algebras with $1_{A_2}\in A_1$, ϱ_i irreducible representations of A_i, and p_i the corresponding minimal central projection of A_i $(i=1,2)$. Then $\varrho_2|A_1$ contains ϱ_1 if and only if $p_1p_2\neq 0$. Indeed, if $\varrho_2|A_1$ contains ϱ_1 we obviously have $\ker(\varrho_2|A_1)\subset\ker\varrho_1$, so that $p_1p_2\neq 0$ (since $p_1\notin\ker\varrho_1$). Conversely, if $p_1p_2\neq 0$ the two-sided ideal $J=\{x\in A_1;\ p_1p_2x=0\}$ of A_1 contains $\ker\varrho_1$ and $p_1\notin J$. Since ϱ_1 is irreducible and A_1 is finite-dimensional, $\ker\varrho_1$ is a maximal ideal of A_1, so that $J=\ker\varrho_1$. It follows that $\ker(\varrho_2|A_1)\subset\ker\varrho_1$ and this in turn implies that $\varrho_2|A_1$ contains ϱ_1.

This completes the proof.

Lemma 2.2. *Consider two signatures* $\tau'\prec\tau$, $l(\tau')=|\tau'|=n$, $l(\tau)=|\tau|=n+1$, *such that* $\iota(\tau)\leqq\dim\mathfrak{K}$ *and let* $\Theta\in H^\infty(\mathscr{L}(\mathfrak{K}))$. *For any* $k\in\mathfrak{K}$ *we have:*

$$(2.4)\qquad \big((\pi_n(p_{\tau'})T_k^*\pi_{n+1}(p_\tau))\otimes I_{H^2}\big)\Gamma_{n+1}(\Theta)H^2(\mathfrak{K}^{\otimes(n+1)})\subset$$
$$\subset\big((\pi_n(p_{\tau'})\otimes I_{H^2})\Gamma_n(\Theta)H^2(\mathfrak{K}^{\otimes n})\big)^-.$$

Proof. Clearly both terms of (2.4) are invariant with respect to multiplication operators by scalar H^∞-functions. Hence it is easily seen that it will be enough to prove that a function of the form

$$z\to\pi_n(p_{\tau'})T_k^*\pi_{n+1}(p_\tau)\big(\Theta(z)k_1\otimes\ldots\otimes\Theta(z)k_{n+1}\big)$$

is in

$$(\pi_n(p_{\tau'})\otimes I_{H^2})\Gamma_n(\Theta)H^2(\mathfrak{K}^{\otimes n}).$$

Writing $p_\tau=\sum_{\sigma\in\mathfrak{S}_{n+1}}c_\sigma\sigma$ the assertion becomes obvious from the following computation:

$$\pi_n(p_{\tau'})T_k^*\pi_{n+1}(p_\tau)\big(\Theta(z)k_1\otimes\ldots\otimes\Theta(z)k_{n+1}\big)=$$
$$=\pi_n(p_{\tau'})T_k^*\sum_{\sigma\in\mathfrak{S}_{n+1}}c_\sigma\big(\Theta(z)k_{\sigma^{-1}(1)}\otimes\ldots\otimes\Theta(z)k_{\sigma^{-1}(n+1)}\big)=$$
$$=\sum_{\sigma\in\mathfrak{S}_{n+1}}c_\sigma\langle\Theta(z)k_{\sigma^{-1}(n+1)},k\rangle\pi_n(p_{\tau'})\Gamma_n(\Theta(z))\big(k_{\sigma^{-1}(1)}\otimes\ldots\otimes k_{\sigma^{-1}(n)}\big).$$

Let us now consider $\Theta\in H^\infty(\mathscr{L}(\mathfrak{K}))$ and let τ be a signature with $|\tau|<\infty$ and $l(\tau)=\dim\mathfrak{K}$. Consider also $\tilde\tau$, the signature of length $|\tau|$ essentially equivalent to τ. We define an inner function $d^\tau(\Theta)$ by

$$(2.5)\qquad d^\tau(\Theta)=\wedge\{m\in H_i^\infty\,|\,mH^2(\mathfrak{K}^\tau)\subset(\varrho_\tau(\Theta)H^2(\mathfrak{K}^\tau))^-\}$$

(by convention we put $\wedge\varnothing=0$, $\varnothing$-the empty set).

Remark that in case Θ is an inner function, $\varrho_\tau(\Theta)$ is still an inner function and $d^\tau(\Theta)$ is the minimal function of $S(\varrho_\tau(\Theta))$ in case $\varrho_\tau(\Theta)$ has a scalar multiple and zero otherwise. In case τ is of the form $(1,1,\ldots,1,0,\ldots)$ with j nonzero terms, that is, ϱ_τ is the representation in antisymmetric tensors of degree j, we shall use the notation $d_j(\Theta)$ for $d^\tau(\Theta)$.

Since the restriction of $\Gamma_{|\tau|}$ to $\pi_{|\tau|}(p_{\tilde\tau})\mathfrak{K}^{\otimes|\tau|}$ is a multiple of ϱ_τ, we have

$$(2.6)\qquad d^\tau(\Theta) = \wedge\{m\in H_i^\infty | mH^2(\pi_{|\tau|}(p_{\tilde\tau})\mathfrak{K}^{\otimes|\tau|}) \subset \\ \subset (\Gamma_{|\tau|}(\Theta)H^2(\pi_{|\tau|}(p_{\tilde\tau})\mathfrak{K}^{\otimes|\tau|}))^-\}.$$

For the next lemma let τ', τ be signatures with $|\tau'|=n$, $|\tau|=n+1$ (n finite), $\ell(\tau')=\ell(\tau)=\dim\mathfrak{K}$ and such that denoting by $\tilde\tau'$ and $\tilde\tau$ the signatures of length n, $n+1$, essentially equivalent to τ', τ, respectively, we have

Lemma 2.3. *For Θ in $H^\infty(\mathcal{L}(\mathfrak{K}))$ and τ', τ as above, $d^{\tau'}(\Theta)$ divides $d^\tau(\Theta)$.*

Proof. Consider $m\in H_i^\infty$ such that

$$mH^2(\pi_{n+1}(p_{\tilde\tau})\mathfrak{K}^{\otimes(n+1)}) \subset (\Gamma_{n+1}(\Theta)H^2(\pi_{n+1}(p_{\tilde\tau})\mathfrak{K}^{\otimes(n+1)}))^-.$$

It follows from Lemma 2.2. that

$$m\Big(\bigvee_{k\in\mathfrak{K}}((\pi_n(p_{\tilde\tau'})T_k^*\pi_{n+1}(p_{\tilde\tau}))\otimes I_{H^2})H^2(\mathfrak{K}^{n+1})\Big)^- \subset \\ \subset ((\pi_n(p_{\tilde\tau'})\otimes I_{H^2})\Gamma_n(\Theta)H^2(\mathfrak{K}^{\otimes n}))^- = (\Gamma_n(\Theta)H^2(\pi_n(p_{\tilde\tau'})\mathfrak{K}^{\otimes n}))^-$$

and hence by Lemma 2.1

$$mH^2(\pi_n(p_{\tilde\tau'})\mathfrak{K}^{\otimes n}) \subset (\Gamma_n(\Theta)H^2(\pi_n(p_{\tilde\tau'})\mathfrak{K}^{\otimes n}))^-$$

so that by (2.6) $d^{\tau'}$ divides m. Q.E.D.

Let us also record the following simple fact for further use.

Remark 2.4. Let $\mathfrak{X}_i$, $\mathfrak{Y}_i$ $(i=1,2)$ be Hilbert spaces, $A_i\in H^\infty(\mathcal{L}(\mathfrak{X}_i, \mathfrak{Y}_i))$, $B\in H^\infty(\mathcal{L}(\mathfrak{X}_1\otimes\mathfrak{X}_2, \mathfrak{Y}_1\otimes\mathfrak{Y}_2))$, $B(z)=A_1(z)\otimes A_2(z)$ $(z\in D)$ and suppose $f_i\in(A_iH^2(\mathfrak{X}_i))^-\cap \cap H^\infty(\mathfrak{Y}_i)$. Then we have $f_1\otimes f_2\in(BH^2(\mathfrak{X}_1\otimes\mathfrak{X}_2))^-$. Indeed, consider $h_i^{(n)}\in H^\infty(\mathfrak{X}_i)$ such that

$$\lim_{n\to\infty}\|A_ih_i^{(n)}-f_i\| = 0 \quad\text{in}\quad H^2(\mathfrak{Y}_i).$$

Then in

$$H^2(\mathfrak{Y}_1\otimes\mathfrak{Y}_2)$$

we have

$$\lim_{n\to\infty}\|B(h_1^{(n)}\otimes h_2^{(m)})-f_1\otimes A_2h_2^{(m)}\| = 0$$

and

$$\lim_{m\to\infty}\|f_1\otimes f_2-f_1\otimes A_2h_2^{(m)}\| = 0$$

which is the desired result.

For the following theorem consider $\Theta\in H^\infty(\mathcal{L}(\mathfrak{K}))$, $\Theta'\in H^\infty(\mathcal{L}(\mathfrak{K}'))$ and suppose there are $A\in H^\infty(\mathcal{L}(\mathfrak{K}, \mathfrak{K}'))$, $B\in H^\infty(\mathcal{L}(\mathfrak{K}', \mathfrak{K}))$ such that the following set of relations holds

$$\begin{cases} A\Theta H^2(\mathfrak{K}) \subset (\Theta' H^2(\mathfrak{K}'))^-, \\ B\Theta' H^2(\mathfrak{K}') \subset (\Theta H^2(\mathfrak{K}))^-, \\ BAH^2(\mathfrak{K})\vee \Theta H^2(\mathfrak{K}) = H^2(\mathfrak{K}). \end{cases} \tag{2.7}$$

Theorem 2.5. *Let Θ, Θ', A, B be as before and suppose* (2.7) *holds. Let further τ, τ' be essentially equivalent signatures with $l(\tau)=\dim\mathfrak{K}$, $l(\tau')=\dim\mathfrak{K}'$, $|\tau|<\infty$ and $\iota(\tau)=\iota(\tau')\leqq\min(\dim\mathfrak{K}, \dim\mathfrak{K}')$. Then $d^\tau(\Theta)$ divides $d^{\tau'}(\Theta')$.*

Proof. If $d^{\tau'}(\Theta')=0$, the assertion of the theorem is obvious, so assume $d^{\tau'}(\Theta')=m\in H_i^\infty$. Let $\tilde{\tau}$ denote the signature of length $n=|\tau|$ that is essentially equivalent to τ.

Consider $f_1, f_2, \dots, f_n\in H^\infty(\mathfrak{K})$, $g_1, g_2, \dots, g_n\in H^\infty(\mathfrak{K})$ and

$$s = (\pi_n(p_{\tilde{\tau}})\otimes I_{H^2})\big((BAf_1+\Theta g_1)\otimes\dots\otimes(BAf_n+\Theta g_n)\big). \tag{2.8}$$

Using (2.7) it is easily seen that the elements s form a total subset of $H^2(\pi_n(p_{\tilde{\tau}})\mathfrak{K}^{\otimes n})$, so that it will be sufficient to prove that

$$ms\in\big(\Gamma_n(\Theta)H^2(\pi_n(p_{\tilde{\tau}})\mathfrak{K}^{\otimes n})\big)^-. \tag{2.9}$$

Now, s is a finite sum of elements of the form

$$r = \big((\pi_n(p_{\tilde{\tau}})\pi_n(\sigma))\otimes I_{H^2}\big)(BAf_1'\otimes\dots\otimes BAf_j'\otimes\Theta g_1'\otimes\dots\otimes\Theta g_{n-j}') \tag{2.10}$$

where $0\leqq j\leqq n$, $\sigma\in\mathfrak{S}_n$ and f_i', g_i' are some of the f' and g. Thus to prove (2.9) it will be enough to show that

$$mr\in\big(\Gamma_n(\Theta)H^2(\pi_n(p_{\tilde{\tau}})\mathfrak{K}^{\otimes n})\big)^-. \tag{2.11}$$

Because $\sum_{\gamma\in\hat{\mathfrak{S}}_j} p_\gamma=1$ and $\mathfrak{S}_j$ is considered as a subgroup of $\mathfrak{S}_n$ $(j\leqq n)$, we have $\sum_{\gamma\in\hat{\mathfrak{S}}_j} p_{\tilde{\tau}}p_\gamma=p_{\tilde{\tau}}$ and $p_{\tilde{\tau}}p_\gamma\neq 0$ if and only if the restriction of the representation of signature $\tilde{\tau}$ to $\mathfrak{S}_j$ contains the representation of signature γ. So, $p_{\tilde{\tau}}p_\gamma\neq 0$ if and only if there are $\gamma_k\in\hat{\mathfrak{S}}_k$ $(j<k<n)$ such that

$$\gamma\prec\gamma_{j+1}\prec\dots\prec\gamma_{n-1}\prec\tilde{\tau}. \tag{2.12}$$

Hence denoting by $\check{\gamma}$ the signature of length $\dim\mathfrak{K}'$ that is essentially equivalent to γ, using Lemma 2.3 several times we conclude that $d^{\check{\gamma}}(\Theta')$ divides $d^{\tau'}(\Theta')=m$.

Now we have:

$$mr =$$

$$= (\pi_n(p_{\tilde{\tau}})\pi_n(\sigma)\otimes I_{H^2})\sum_{\substack{\gamma\in\hat{\mathfrak{S}}_j\\ p_{\tilde{\tau}}p_\gamma\neq 0}}\big(m\,(\pi_j(p_\gamma)\otimes I_{H^2})\,(BAf_1'\otimes\dots\otimes BAf_j')\big)\otimes(\Theta g_1'\otimes\dots\otimes\Theta g_{n-j}').$$

To end the proof it will be sufficient to show that

$$m(\pi_j(p_\gamma)\otimes I_{H^2})(BAf_1'\otimes\ldots\otimes BAf_j') \quad \text{is in} \quad (\Gamma_j(\Theta)H^2(\mathfrak{K}^{\otimes j}))^-,$$

because then using Remark 2.4 we will have that mr is in

$$((\pi_n(p_\tau)\pi_n(\sigma)\otimes I_{H^2})\Gamma_n(\Theta)H^2(\mathfrak{K}^{\otimes n}))^- = (\Gamma_n(\Theta)H^2(\pi_n(p_\tau)\mathfrak{K}^{\otimes n}))^-$$

which is the desired result.

Now further $m(\pi_j(p_\gamma)\otimes I_{H^2})(Af_1'\otimes\ldots\otimes Af_j')$ is in $d^\gamma\,(\Theta')H^2(\pi_j(p_\gamma)\mathfrak{K}'^{\otimes j})$, since $d^\gamma(\Theta')$ divides m, and hence is in $(\Gamma_j(\Theta')H^2(\pi_j(p_\gamma)\mathfrak{K}'^{\otimes j}))^-\subset(\Gamma_j(\Theta')H^2(\mathfrak{K}'^{\otimes j}))^-$. Thus it will be sufficient to prove that

$$(\Gamma_j(B)\Gamma_j(\Theta')H^2(\mathfrak{K}'^{\otimes j}))^-\subset(\Gamma_j(\Theta)H^2(\mathfrak{K}^{\otimes j}))^-$$

in order that

$$m(\pi_j(p_\gamma)\otimes I_{H^2})(BAf_1'\otimes\ldots\otimes BAf_j')\in(\Gamma_j(\Theta)H^2(\mathfrak{K}^{\otimes j}))^-.$$

To this end remark that the elements of the form $B\Theta'h_1\otimes\ldots\otimes B\Theta'h_j$ with $h_i\in H^\infty(\mathfrak{K}')$ are total in $(\Gamma_j(B)\Gamma_j(\Theta')H^2(\mathfrak{K}'^{\otimes j}))^-$ and

$$B\Theta'h_1\otimes\ldots\otimes B\Theta'h_j\in(\Gamma_j(\Theta)H^2(\mathfrak{K}^{\otimes j}))^-$$

because fo (2.7) and Remark 2.4. Q.E.D.

§ 3. Applications to quasi-similar C_0 operators

The following Proposition is an easy application of Theorem 2.5.

Proposition 3.1. *Let $\Theta\in H_i^\infty(\mathscr{L}(\mathfrak{K}))$, $\Theta'\in H_i^\infty(\mathscr{L}(\mathfrak{K}'))$ and let τ,τ' be essentially equivalent signatures with $l(\tau)=\dim\mathfrak{K}$, $l(\tau')=\dim\mathfrak{K}'$ and $\iota(\tau)=\iota(\tau')\leqq\leqq\min(\dim\mathfrak{K},\dim\mathfrak{K}')$. If $S(\Theta)$ and $S(\Theta')$ are quasi-similar, we have*

$$d^\tau(\Theta)=d^{\tau'}(\Theta'). \tag{3.1}$$

Proof. Let X and Y be two quasi-affinities such that $S(\Theta')X=XS(\Theta)$ and $S(\Theta)Y=YS(\Theta')$. From the lifting theorem (see (1.8—11)) it follows that we can find $A\in H^\infty(\mathscr{L}(\mathfrak{K},\mathfrak{K}'))$ and $B\in H^\infty(\mathscr{L}(\mathfrak{K}',\mathfrak{K}))$ such that

$$X=P_{\mathfrak{H}(\Theta')}A|\mathfrak{H}(\Theta),\quad Y=P_{\mathfrak{H}(\Theta)}B|\mathfrak{H}(\Theta'), \tag{3.2}$$

$$A\Theta H^2(\mathfrak{K})\subset\Theta'H^2(\mathfrak{K}'),\quad B\Theta'H^2(\mathfrak{K}')\subset\Theta H^2(\mathfrak{K}) \tag{3.3}$$

and

$$ABH^2(\mathfrak{K}')\vee\Theta'H^2(\mathfrak{K}')=H^2(\mathfrak{K}'),\quad BAH^2(\mathfrak{K})\vee\Theta H^2(\mathfrak{K})=H^2(\mathfrak{K}) \tag{3.4}$$

so that the assumptions of Theorem 2.5 are satisfied. It follows that $d^\tau(\Theta)$ divides $d^{\tau'}(\Theta')$ and $d^{\tau'}(\Theta')$ divides $d^\tau(\Theta)$ and this proves (3.1). Q.E.D.

Let T be any operator unitarily equivalent to some $S(\Theta)$ with a pure $\Theta \in H_i^\infty(\mathscr{L}(\mathfrak{K}))$. It is easy to see that the functions $d^\tau(\Theta)$ and $d_j(\Theta)$ depend only on T and not on the particular function Θ, so we shall denote them by $d^\tau(T)$ and $d_j(T)$, respectively.

Corollary 3.2. *If T and T' are two quasisimilar C_0 operators and $\mathfrak{d}_T = \mathfrak{d}_{T'}$, then $d^\tau(T) = d^\tau(T')$ for each τ with $l(\tau) = \mathfrak{d}_T$.*

Proof. T and T' are unitarily equivalent to $S(\Theta)$ and $S(\Theta')$, respectively, where $\Theta \in H_i^\infty(\mathscr{L}(\mathfrak{K}))$, $\Theta' \in H_i^\infty(\mathscr{L}(\mathfrak{K}'))$ with $\dim \mathfrak{K} = \dim \mathfrak{K}' = \mathfrak{d}_T$. The corollary obviously follows from Proposition 3.1. Q.E.D.

Consider now a C_0 operator T with Jordan model $S = S(m_1) \oplus S(m_2) \oplus \ldots$. If $\mathfrak{d}_S < \mathfrak{d}_T$ we shall put $m_j = 1$ for $\mathfrak{d}_S < j \leqq \mathfrak{d}_T$. So we have

$$S = \bigoplus_{j=1}^{\mathfrak{d}_T} S(m_j). \tag{3.5}$$

Corollary 3.3. *For any C_0 operator T and any signature $\tau = (t_1 \geqq t_2 \geqq \ldots)$, $|\tau| < \infty$, $l(\tau) = \mathfrak{d}_T$, we have*

$$d^\tau(T) = m_1^{t_1} m_2^{t_2}, \ldots, m_n^{t_n}, \quad n = \iota(\tau). \tag{3.6}$$

Proof. We have only to apply Proposition 3.1 to Θ coinciding with the characteristic function of T and to

$$\Theta' = \operatorname{diag}(m_1, m_2, \ldots) \in H_i^\infty(\mathscr{L}(\mathfrak{K}')) \quad \text{with} \quad \dim \mathfrak{K}' = \mathfrak{d}_T.$$

Since $\tau = (t_1 \geqq t_2 \geqq \ldots)$ represents the highest weight to the representation ϱ_τ (see [18], [6] to the finite-dimensional and [1] for the infinite-dimensional case) it is immediate that:

$$d^\tau(\Theta') = m_1^{t_1}, \ldots, m_n^{t_n}.$$ Q.E.D.

Corollary 3.4. *For any C_0 operator T, the functions m_j appearing in the Jordan model can be computed as*

$$m_j = d_j(T)/d_{j-1}(T), \quad 1 \leqq j \leqq \mathfrak{d}_T \quad \text{where} \quad d_0(T) = 1. \tag{3.7}$$

Proof. The preceding Corollary gives for $\tau_j = (1, \ldots, 1, 0, \ldots)$ (with j nonzero terms)

$$d_j(T) = d^{\tau_j}(T) = m_1 \ldots m_j, \quad j \leqq \mathfrak{d}_T$$

so relation (3.7) becomes obvious. Q.E.D.

Since the quasisimilarity class of a C_0 operator is determined by the Jordan model, Corollary 3.4 shows that a C_0 operator T is determined up to quasisimilarity by the least inner multiples of the exterior powers of any function coinciding with the characteristic function of T. This enables us to prove the following theorem.

Theorem 3.5. *Let $\Theta \in H_i^\infty(\mathscr{L}(\mathfrak{K}))$, $\Theta' \in H_i^\infty(\mathscr{L}(\mathfrak{K}'))$ be such that $d_1(\Theta) \neq 0$ and* $\dim \mathfrak{K} = \dim \mathfrak{K}'$. *If $S(\Theta)$ and $S(\Theta')$ are quasisimilar then $S(\varrho_\tau(\Theta))$ and $S(\varrho_\tau(\Theta'))$ are quasisimilar for each signature τ such that $\ell(\tau) = \dim \mathfrak{K}$, $|\tau| < \infty$.*

Proof. By Corollary 3.4 we have only to show that $d_j(\varrho_\tau(\Theta)) = d_j(\varrho_\tau(\Theta'))$ for each $j \leq \dim \mathfrak{K}^\tau$. Let $\tau_j = (1, 1, \ldots, 1, 0, \ldots)$ (with j nonzero terms), $\ell(\tau_j) = \dim \mathfrak{K}^\tau$.

The representation $\varrho_{\tau_j} \circ \varrho_\tau$ of $\mathscr{U}(\mathfrak{K})$ is a subrepresentation of the representation of $\mathscr{U}(\mathfrak{K})$ on $\mathfrak{K}^{\otimes j|\tau|}$ and hence a finite direct sum of representations $\varrho_{\tau'}$, with $\ell(\tau') = \dim \mathfrak{K}$, $|\tau'| < \infty$:

$$\varrho_{\tau_j} \circ \varrho_\tau = \bigoplus_{\tau'} \varrho_{\tau'}. \tag{3.8}$$

From (3.8) it follows then that

$$\varrho_{\tau_j}(\varrho_\tau(\Theta)) = \bigoplus_{\tau'} \varrho_{\tau'}(\Theta), \quad \varrho_{\tau_j}(\varrho_\tau(\Theta')) = \bigoplus_{\tau'} \varrho_{\tau'}(\Theta')$$

and hence $d_j(\varrho_\tau(\Theta))$ is the least inner common multiple of the $d^{\tau'}(\Theta)$ and $d_j(\varrho_\tau(\Theta'))$ the least inner common multiple of the $d^{\tau'}(\Theta')$. Since $d^{\tau'}(\Theta) = d^{\tau'}(\Theta')$ by Proposition 3.1, we infer that $d_j(\varrho_\tau(\Theta)) = d_j(\varrho_\tau(\Theta'))$. Q.E.D.

§ 4. Defect operators of C_0 contractions

For an operator $A \in \mathscr{L}(\mathfrak{K})$ and a closed subspace $\mathfrak{M} \subset \mathfrak{K}$ we consider

$$\gamma[A, \mathfrak{M}] = \inf_{\substack{k \in \mathfrak{M} \\ \|k\| = 1}} \|Ak\|, \quad \gamma_j(A) = \sup_{\operatorname{codim} \mathfrak{M} = j-1} \gamma[A, \mathfrak{M}].$$

As is known from the minimax principle, $\gamma_j(A)$ $(1 \leq j \leq \dim \mathfrak{K})$ are eigenvalues of $(A^*A)^{1/2}$ in increasing order. In case $\dim \mathfrak{K} < \infty$ all eigenvalues of $(A^*A)^{1/2}$ repeated according to their multiplicity appear in the sequence of the $\gamma_j(A)$. In case $\dim \mathfrak{K} = \infty$, $\gamma_1(A)$ is the least eigenvalue of $(A^*A)^{1/2}$, discrete eigenvalues smaller than the least essential eigenvalue appear in increasing order repeated according to their multiplicity and the sequence becomes stationary if the least essential eigenvalue of $(A^*A)^{1/2}$ is reached.

For the next two lemmas, τ_j denotes the signature

$$\tau_j = (1, \ldots, 1, 0 \ldots), \quad \ell(\tau_j) = \dim \mathfrak{K}, \quad \iota(\tau_j) = j.$$

Lemma 4.1. *Let $A \in \mathscr{L}(\mathfrak{K})$ and τ_j be as above. Then we have:*

$$\gamma_1(\varrho_{\tau_j}(A)) = \gamma_1(A)\gamma_2(A) \ldots \gamma_j(A). \tag{4.1}$$

Proof. Remark first that applying ϱ_{τ_j} to the polar decomposition of A we get the polar decomposition of $\varrho_{\tau_j}(A)$, so we can suppose A is positive. Moreover, in view of the minimax definition of γ_j, we have $|\gamma_j(A)-\gamma_j(B)|\leqq\|A-B\|$, and thus by continuity it will be sufficient to consider the case when $A\geqq 0$ has finite spectrum.

In this case, ϱ_{τ_j} being the representation in antisymmetric tensors of degree j, $\varrho_{\tau_j}(A)$ has finite spectrum, the eigenvalues being products $\lambda_1\ldots\lambda_j$ of eigenvalues of A, a given eigenvalue appearing in such a product at most a number of times equal to its multiplicity. Clearly $\gamma_1(A)\ldots\gamma_j(A)$ is then the least eigenvalue of $\varrho_{\tau_j}(A)$.

Q.E.D.

Lemma 4.2. *Let T be a C_0 operator, let $\Theta\in H_i^\infty(\mathscr{L}(\mathfrak{K}))$ coincide with the characteristic function of T and let $\{m_j\}_{j=1}^{\mathfrak{d}_T}$ be inner functions for the Jordan model of T with $m_j\equiv 1$ for $\mu_T<j\leqq\mathfrak{d}_T$. Then we have*

$$\gamma_1(\Theta(\lambda))\ldots\gamma_j(\Theta(\lambda))\geqq|m_1(\lambda)\ldots m_j(\lambda)| \tag{4.2}$$

where $1\leqq j\leqq\mathfrak{d}_T$ and $\lambda\in D$.

Proof. In view of Corollary 3.3, $m_1\ldots m_j$ is the least inner multiple of $\varrho_{\tau_j}(\Theta)\in\in H_i^\infty(\mathscr{L}(\mathfrak{K}^{\tau_j}))$. Hence there is a contractive function $\Omega\in H^\infty(\mathscr{L}(\mathfrak{K}^{\tau_j}))$ such that

$$\Omega(\lambda)\varrho_{\tau_j}(\Theta(\lambda))=m_1(\lambda)\ldots m_j(\lambda)I_{\mathfrak{K}^{\tau_j}}.$$

Since $\|\Omega(\lambda)\|\leqq 1$ this clearly implies

$$\gamma_1(\varrho_{\tau_j}(\Theta(\lambda)))\geqq|m_1(\lambda)\ldots m_j(\lambda)|$$

and by Lemma 4.1

$$\gamma_1(\varrho_{\tau_j}(\Theta(\lambda)))\leqq\gamma_1(\Theta(\lambda))\ldots\gamma_j(\Theta(\lambda)),$$

which gives the desired inequality. Q.E.D.

Proposition 4.3. *Let T be a C_0 operator acting on $\mathfrak{H}$ and $\{m_j\}_{j=1}^\infty$ inner functions for the Jordan model of T, with $m_j\equiv 1$ in case $\mu_T<j$.*

a) *If* $\sum_{j=1}^\infty(1-|m_j(0)|)<\infty$, *then* $\operatorname{tr}(I-T^*T)<\infty$.

b) *If* $\lim_{j\to\infty}|m_j(0)|=1$, *then $I-T^*T$ is compact.*

Proof. a) The assumptions are that the Jordan model $S=S(m_1)\oplus S(m_2)\oplus\ldots$ is a weak contraction ([13] ch. VIII) since $\operatorname{tr}(I-S^*S)=\sum_{j=1}^\infty(1-|m_j(0)|^2)\leqq\leqq 2\cdot\sum_{j=1}^\infty(1-|m_j(0)|)<\infty$. As usual for weak contractions there will be no loss of generality to assume that $m_j(0)\neq 0$ (one uses a conformal automorphism of the unit disc as in [13] ch. VIII). Thus the infinite product $\prod_{j=1}^\infty|m_j(0)|$ converges to some $c>0$. Hence by Lemma 4.2 for Θ the characteristic function of T, we infer that

$$\prod_{1\leqq j\leqq\mathfrak{d}_T}\gamma_j(\Theta(0))>0.$$

Since in case $\mathfrak{d}_T=\infty$ this implies $\lim_{j\to\infty} \gamma_j(\Theta(0))=1$, it follows that

$$\operatorname{tr}(I_{\mathfrak{D}_T}-\Theta(0)^*\Theta(0)) = \sum_{1\leq j\leq \mathfrak{d}_T} (1-\gamma_j(\Theta(0))^2)$$

and

$$\sum_{1\leq j\leq \mathfrak{d}_T} (1-\gamma_j(\Theta(0))^2) < \infty$$

since

$$\prod_{1\leq j\leq \mathfrak{d}_T} \gamma_j(\Theta(0)) > 0. \quad \text{But} \quad I_{\mathfrak{D}_T}-\Theta(0)^*\Theta(0) = D_T^2|\mathfrak{D}_T,$$

so that $\operatorname{tr}(I-T^*T)<\infty$.

b) The proof is quite similar to that of a), so we can be brief in details. Again we may suppose T is invertible and hence $m_j(0)\neq 0$. Then $\lim_{j\to\infty} |m_j(0)|=1$ gives

$$\lim_{j\to\infty} |m_1(0)\dots m_j(0)|^{1/j} = 1.$$

Using Lemma 4.2 this implies

$$\lim_{j\to\infty} (\gamma_1(\Theta(0))\dots\gamma_j(\Theta(0)))^{1/j} = 1$$

so that $\lim_{j\to\infty} \gamma_j(\Theta(0))=1$ which gives that $I-T^*T$ is compact.

Q.E.D.

Remark 4.4. As we shall see in § 8 the converse of 4.3 a) is also true. For 4.3 b) the converse is in general false. An example can be constructed as follows.

Let μ be a finite non-negative measure on $[0, 2\pi]$, singular with respect to Lebesgue measure and without atoms. Consider the inner functions

$$m_{j,n}(\lambda) = \exp\left[-\int_{2\pi(j-1)/n}^{2\pi j/n} \frac{e^{it}+\lambda}{e^{it}-\lambda}\,d\mu(t)\right], \quad 1\leq j\leq n$$

and the operators

$$T=\bigoplus_{n=1}^{\infty}\left(\bigoplus_{j=1}^{n} S(m_{j,n})\right), \quad S=S(m_{1,1})\oplus S(m_{1,1})\oplus\dots.$$

Then S is the Jordan model of T, $I-T^*T$ is compact and $|m_{1,1}(0)|$, $|m_{1,1}(0)|$, ... tends to $|m_{1,1}(0)|<1$.

Proposition 4.5. *Let T be a C_0 operator, let $\{m_j\}_{j=1}^{\infty}$ be inner functions for the Jordan model of T ($m_j\equiv 1$ in case $\mu_T<j$) and let $\Theta\in H_i^\infty(\mathscr{L}(\mathfrak{K}))$ coincide with the characteristic function of T. Suppose moreover $m_1(0)\neq 0$ and $n\in\mathbf{N}$ is such that $|m_n(0)|<$ $<\lim_{j\to\infty} |m_j(0)|$. Then the following conditions are equivalent:*

(i) $|m_1(0)\dots m_n(0)|=\gamma_1(\Theta(0))\dots\gamma_n(\Theta(0))$,

(ii) *T is unitarily equivalent to $T_1\oplus T_2$, where $\mathfrak{d}_{T_1}=n$ and T_1, T_2 are quasisimilar to $S(m_1)\oplus\dots\oplus S(m_n)$ and respectively to $S(m_{n+1})\oplus S(m_{n+2})\oplus\dots$.*

Proof. (i)$\Rightarrow$(ii). The condition $0\neq|m_n(0)|<\lim_{j\to\infty}|m_j(0)|$ implies that $\gamma_n(\Theta(0))$ is less than the least essential eigenvalue of $(\Theta(0)^*\Theta(0))^{1/2}$, for otherwise we would have $\gamma_n(\Theta(0))=\gamma_{n+1}(\Theta(0))=\dots$ which in view of Lemma 4.2 would imply $\lim_{j\to\infty}|m_j(0)|\leqq\leqq\gamma_n(\Theta(0))$ and hence $|m_n(0)|<\gamma_n(\Theta(0))$ which when combined with (i) would give $|m_1(0)\dots m_{n-1}(0)|>\gamma_1(\Theta(0))\dots\gamma_{n-1}(\Theta(0))$, contradicting Lemma 4.2. Thus replacing Θ by some equivalent inner operator-valued function in $H_i^\infty(\mathcal{L}(\mathfrak{K}))$ we may assume there is an orthonormal set $\{e_1,\dots,e_n\}$ in $\mathfrak{K}$ such that $\Theta(0)e_j=\gamma_j(\Theta(0))e_j$ for $1\leqq j\leqq n$. Consider $f=\pi_n(p_{\tau_n})(e_1\oplus\dots\oplus e_n)$. Then

$$\varrho_{\tau_n}(\Theta(0))f=\gamma_1(\Theta(0))\dots\gamma_n(\Theta(0))f$$

and since $p_{\tau_n}=(n!)^{-1}\sum_{\sigma\in\mathfrak{S}_n}\varepsilon(\sigma)\sigma$ ($\varepsilon(\sigma)$ is the sign of the permutation σ), we have $f\neq0$. But $\Omega\varrho_{\tau_n}(\Theta)=m_1\dots m_nI_{\mathfrak{K}^{\tau_n}}$ for some contractive Ω, and we infer $\|\Omega(0)f\|=\|f\|$ so that $\Omega(\lambda)f=\mu f$ for some constant μ, $|\mu|=1$. This in turn implies $\varrho_{\tau_n}(\Theta(\lambda))f==\mu^{-1}m_1(\lambda)\dots m_n(\lambda)f$ for all $\lambda\in D$. In view of the known properties of p_{τ_n} this last equality implies that $\mathfrak{B}=\mathbf{C}e_1+\dots+\mathbf{C}e_n$ is invariant for $\Theta(\lambda)$ for all $\lambda\in D$. Since Θ is two-sided inner we infer that $\mathfrak{B}$ is a reducing subspace for $\Theta(\lambda)$, $\lambda\in D$. Hence $\Theta=\Theta_1\oplus\Theta_2$ where $\Theta_1=\Theta|\mathfrak{B}$, $\Theta_2=\Theta|\mathfrak{K}\ominus\mathfrak{B}$.

Thus we define $T_i=S(\Theta_i)$ for $i=1,2$ and clearly T is unitarily equivalent to $T_1\oplus T_2$ and $\mathfrak{d}_{T_1}=n$. Remark also that $\varrho_{\tau_n}(\Theta_1)$ coincides with $m_1\dots m_n$. Let $S(m_1')\oplus\dots\oplus S(m_n')$ and $S(m_1'')\oplus S(m_2'')\oplus\dots$ be the Jordan models of T_1 and T_2 (we do not exclude the possibility that some m_j' or m_j'' be 1). Then we have:

$$m_1\dots m_n=m_1'\dots m_n'=\bigvee_{k=0}^{n}m_1'\dots m_k'm_1''\dots m_{n-k}'' \tag{4.3}$$

(use for instance Proposition 3.1 with $\tau=\tau_n$). From 4.3 we infer that $m_1'\dots m_{n-1}'m_1''$ divides $m_1'\dots m_n'$ and hence m_1'' divides m_n'. Thus $S(m_1')\oplus\dots\oplus S(m_n')\oplus S(m_1'')\oplus\oplus S(m_2'')\oplus\dots$ is the Jordan model of $T_1\oplus T_2$ and hence $m_j'=m_j$, $m_k''=m_{n+k}$ ($1\leqq j\leqq n$, $k=1,2,\dots$). This ends the proof of (i)$\Rightarrow$(ii).

(ii)$\Rightarrow$(i). Let Θ_1, Θ_2 coincide with the characteristic functions of T_1, T_2. Then $\varrho_{\tau_n}(\Theta_1)$ coincides with $m_1\dots m_n$ so that $\gamma_1(\Theta_1(0))\dots\gamma_n(\Theta_1(0))=\gamma_1(\tau_n(\Theta_1(0)))==|m_1(0)\dots m_n(0)|$ (use Proposition 3.1 for instance and then Lemma 4.1). Now clearly $\gamma_j(\Theta_1(0))\geqq\gamma_j(\Theta(0))$ and hence $\gamma_1(\Theta(0))\dots\gamma_n(\Theta(0))\leqq|m_1(0)\dots m_n(0)|$ which in view of Lemma 4.2 gives $\gamma_1(\Theta(0))\dots\gamma_n(\Theta(0))=|m_1(0)\dots m_n(0)|$.

Q.E.D

Remark 4.6. If T is a contraction and Θ is its characteristic function then $\gamma_j(\Theta(0))=\gamma_j(T)$. Thus, let T be a C_0 contraction with Jordan model $S(m_1)\oplus\oplus S(m_2)\oplus\dots$ such that $m_1(0)\neq0$ and $\lim_{j\to\infty}|m_j(0)|=1$. Proposition 4.5 shows that the Jordan model of T can be characterized within the class $\mathcal{T}$ of contractions, which are quasisimilar with T by its extremal properties. Indeed, define $\mathscr{T}_n$ recurrently, by $\mathscr{T}_0=\mathscr{T}$ and $\mathscr{T}_{n+1}=\{T'\in\mathscr{T}_n \mid \gamma_{n+1}(T')=\inf\limits_{S\in\mathscr{T}_n}\gamma_{n+1}(S)\}$. Then the only member up to unitary equivalence of $\bigcap\limits_{n=0}^{\infty}\mathscr{T}_n$ is the Jordan model of T.

Part II

§ 5. Preliminaries

1. We begin with a short review of the properties of infinite determinants (see [4], ch. IV, § 1), in order to discuss (in the next section) minors of such determinants.

Let $\mathfrak{K}$ be a complex separable Hilbert space and $\mathscr{C}_1(\mathfrak{K})$ the ideal of nuclear operators, endowed with the trace-norm

$$\|X\|_1 = \operatorname{tr}|X|,\ |X| = (X^*X)^{1/2} \quad (X\in\mathscr{C}_1(\mathfrak{K})). \tag{5.1}$$

Consider $X\in I+\mathscr{C}_1(\mathfrak{K})$ and let $\{\lambda_j(X)\}_{j=1}^{\infty}$ be the eigenvalues of X (repeated according to their multiplicities). We have

$$\sum_{j=1}^{\infty}|1-\lambda_j(X)| \leqq \operatorname{tr}|I-X| < \infty$$

and it follows that the infinite product defining the determinant

$$\det(X) = \prod_{j=1}^{\infty}\lambda_j(X) \tag{5.2}$$

converges absolutely. Moreover, $\det(I+Y)$ as a function of $Y\in\mathscr{C}_1(\mathfrak{K})$ is analytic (in particular continuous on the Banach space $\mathscr{C}_1(\mathfrak{K})$). This follows from [4], Ch. IV, Corollary 1.1 and property 8° on p. 207, combined with Proposition 2 on p. 11 of [3].

Also for $\{e_j\}_{j=1}^{\infty}$ an orthonormal basis of $\mathfrak{K}$ and $X\in I+\mathscr{C}_1(\mathfrak{K})$, we have

$$\det(X) = \lim_{N\to\infty}\det[\langle Xe_i, e_j\rangle]_{1\leqq i,\, j\leqq N} \tag{5.3}$$

(cf. [4], property 2° on p. 203).

Furthermore, for $X, X'\in I+\mathscr{C}_1(\mathfrak{K})$ we have (cf. the proof of property 7° on p. 206 of [4]):

$$\det(XX') = \det(X)\det(X'). \tag{5.4}$$

In view of (5.2) the following assertions are easily seen to be true: a) if $X\in I+ +\mathscr{C}_1(\mathfrak{K})$ is unitary then $|\det(X)|=1$; b) if $X\in I+\mathscr{C}_1(\mathfrak{K})$ is a contraction then

$|\det(X)|\leqq 1$; c) $X\in I+\mathscr{C}_1(\mathfrak{K})$ is invertible if and only if $\det(X)\neq 0$; d) the determinant is invariant under similarities.

2. For any Hilbert space $\mathfrak{K}$ we shall indicate by "$\rightharpoonup$" the weak convergence in $\mathfrak{K}$ and in $\mathscr{L}(\mathfrak{K})$. In order to avoid antilinear mappings we shall consider the dual space $\mathfrak{K}^d$. If $T\in\mathscr{L}(\mathfrak{K})$, the dual operator is denoted by T^d $(T^d\in\mathscr{L}(\mathfrak{K}^d))$. $(\mathfrak{K}^d)^d$ can be identified in the usual way with $\mathfrak{K}$.

3. For any Hilbert space $\mathfrak{K}$ and $n\geqq 0$ we shall denote by $\mathfrak{K}^{\wedge n}$ the n-th exterior power of $\mathfrak{K}$. For $n=0$ this is just the complex field $\mathbf{C}$ and in general $\mathfrak{K}^{\wedge n}$ coincides with $\mathfrak{K}^{\tau_n}$ for $\tau_n=(1,1,1,\dots,1,0,\dots)$, $\ell(\tau_n)=\dim\mathfrak{K}$, $\lambda(\tau_n)=n$ (cf. § 1.4). $\mathfrak{K}^{\wedge n}$ is generated by vectors of the form

$$(5.5)\qquad k_1\wedge k_2\wedge\dots\wedge k_n=(n!)^{-1/2}\sum_{\sigma\in\mathfrak{S}_n}\varepsilon(\sigma)k_{\sigma(1)}\otimes\dots\otimes k_{\sigma(n)},\quad k_j\in\mathfrak{K}\quad(1\leqq j\leqq n),$$

where $\varepsilon(\sigma)$ is the sign of the permutation σ.

The factor $(n!)^{-1/2}$ has been chosen so that $\|e_1\wedge\dots\wedge e_n\|=1$ for any orthonormal system $\{e_1, e_2,\dots, e_n\}$.

For n, m two positive integers there is a bilinear map

$$\wedge:\mathfrak{K}^{\wedge n}\times\mathfrak{K}^{\wedge m}\to\mathfrak{K}^{\wedge(m+n)}$$

such that $(k_1\wedge k_2\wedge\dots\wedge k_n)\wedge(k_{n+1}\wedge\dots\wedge k_{n+m})=k_1\wedge\dots\wedge k_{n+m}$. For each $A\in\mathscr{L}(\mathfrak{K})$ we shall denote $\varrho_{\tau_n}(A)$ as $A^{\wedge n}$, so that

$$(5.6)\qquad A^{\wedge n}(k_1\wedge\dots\wedge k_n)=Ak_1\wedge\dots\wedge Ak_n.$$

Let $\mathfrak{K}$ now be a Hilbert space of finite dimension n. If $\{e_1,\dots,e_n\}$ is an orthonormal basis of $\mathfrak{K}$, we can define a bilinear form

$$B:\mathfrak{K}^{\wedge k}\times\mathfrak{K}^{\wedge(n-k)}\to\mathbf{C}$$

by the formula

$$(5.7)\qquad B(h,g)=\langle h\wedge g, e_1\wedge\dots\wedge e_n\rangle.$$

Choosing in $\mathfrak{K}^{\wedge j}$ the usual orthonormal basis

$$\{e_{i_1}\wedge\dots\wedge e_{i_j}\,|\,1\leqq i_1<i_2<\dots<i_j\leqq n\}$$

it is easy to see that the mapping

$$(5.8)\qquad C:\mathfrak{K}^{\wedge(n-k)}\to(\mathfrak{K}^{\wedge k})^d$$

given by $C(g)(h)=B(h,g)$ for $g\in\mathfrak{K}^{\wedge(n-k)}$, $h\in\mathfrak{K}^{\wedge k}$ is a linear isometry. If $A\in\mathscr{L}(\mathfrak{K})$ we have

$$(5.9)\qquad B(A^{\wedge k}h, A^{\wedge(n-k)}g)=\det(A)\,B(h,g)$$

because $A^{\wedge n} = \det(A) I_{\mathfrak{K}^{\wedge n}}$. Let us define

$$F = CA^{\wedge(n-k)}C^{-1} \in \mathscr{L}((\mathfrak{K}^{\wedge k})^d) \tag{5.10}$$

and

$$A^{Adk} = F^d \in \mathscr{L}(\mathfrak{K}^{\wedge k}). \tag{5.11}$$

We have $B(A^{Adk}h, g) = C(g)(A^{Adk}h) = (F(C(g)))(h) = (C(A^{\wedge(n-k)}g))(h) = B(h, A^{\wedge(n-k)}g)$ and since C is isometric,

$$\|A^{\wedge(n-k)}\| = \|F\| = \|A^{Adk}\|. \tag{5.12}$$

Also, as B is nondegenerate we have

$$A^{Adk}A^{\wedge k} = \det(A) I_{\mathfrak{K}^{\wedge k}}. \tag{5.13}$$

It is obvious by the definition of A^{Adk} that

$$(A_1A_2)^{Adk} = A_2^{Adk}A_1^{Adk},\ A_1, A_2 \in \mathscr{L}(\mathfrak{K}), \tag{5.14}$$

and it can be shown that

$$(A^*)^{Adk} = (A^{Adk})^*. \tag{5.15}$$

Moreover, for invertible A we infer from (5.13) that

$$A^{\wedge k}A^{Adk} = \det(A) I_{\mathfrak{K}^{\wedge k}} \tag{5.16}$$

and by continuity it follows that (5.16) always holds.

For $\{f_1, f_2, \dots, f_k\}$ an orthonormal system in $\mathfrak{K}$ we shall show that

$$\langle A^{Adk}(f_1\wedge\dots\wedge f_k), f_1\wedge\dots\wedge f_k\rangle = \det(P+(I-P)A(I-P)) \tag{5.17}$$

where P denotes the orthogonal projection onto the linear span of $\{f_1, f_2, \dots, f_k\}$. Completing the system $\{f_1, \dots, f_k\}$ to an orthonormal basis $\{f_1, \dots, f_n\}$, we have

$$\begin{aligned}
&\langle A^{Adk}(f_1\wedge\dots\wedge f_k), f_1\wedge\dots\wedge f_k\rangle = \\
&= \langle (A^{Adk}(f_1\wedge\dots\wedge f_k))\wedge f_{k+1}\wedge\dots\wedge f_n, f_1\wedge\dots\wedge f_n\rangle = \\
&= B(A^{Adk}(f_1\wedge\dots\wedge f_k), f_{k+1}\wedge\dots\wedge f_n)\cdot\langle f_1\wedge\dots\wedge f_n, e_1\wedge\dots\wedge e_n\rangle^{-1} = \\
&= B(f_1\wedge\dots\wedge f_k, A^{\wedge(n-k)}(f_{k+1}\wedge\dots\wedge f_n))\cdot\langle f_1\wedge\dots\wedge f_n, e_1\wedge\dots\wedge e_n\rangle^{-1} = \\
&= \langle f_1\wedge\dots\wedge f_k\wedge A^{\wedge(n-k)}(f_{k+1}\wedge\dots\wedge f_n), f_1\wedge\dots\wedge f_n\rangle = \\
&= \langle (P+A(I-P))^{\wedge n}(f_1\wedge\dots\wedge f_n), f_1\wedge\dots\wedge f_n\rangle = \\
&= \det(P+A(I-P)) = \det(P+(I-P)A(I-P)).
\end{aligned}$$

Formulas (5.14), (5.16), (5.17) show that A^{Adk} does not depend on the particular choice of the orthonormal basis $\{e_1, \dots, e_n\}$.

Let us now suppose that A is a positive operator with eigenvalues $\lambda_1 \geqq \lambda_2 \geqq \dots \geqq \lambda_n$ and the corresponding eigenvectors $f_1, f_2, \dots, f_n$. Then $A^{\wedge(n-k)}$ is positive with eigenvalues

$$\lambda_{i_1}\lambda_{i_2}\dots\lambda_{i_{n-k}} \quad (1 \leqq i_1 < i_2 < \dots < i_{n-k} \leqq n).$$

It follows that

$$\|A^{Adk}\| = \|A^{\wedge(n-k)}\| = \lambda_1\dots\lambda_{n-k} \leqq (1+|\lambda_1-1|)(1+|\lambda_2-1|)\dots(1+|\lambda_{n-k}-1|) \leqq$$
$$\leqq \exp(|\lambda_1-1|)\exp(|\lambda_2-1|)\dots\exp(|\lambda_{n-k}-1|) \leqq \exp(\mathrm{Tr}|A-I|).$$

Now for any $T\in\mathscr{L}(\mathfrak{K})$, we have

$$\|T^*T-I\|_1 \leqq (1+\|T-I\|_1)^2-1 \quad \text{and} \quad \|(T^*T)^{1/2}-I\|_1 \leqq \|T^*T-I\|_1$$

as can be seen by comparing the eigenvalues of these operators. Therefore,

$$\|(T^*T)^{1/2}-I\|_1 \leqq (1+\|T-I\|_1)^2-1.$$

In particular, for the polar decomposition $T=UA$ of T $\left(A=|T|=(T^*T)^{1/2}\right)$ it follows that:

(5.18)
$$\|T^{Adk}\| = \|A^{Adk}U^{Adk}\| \leqq \|A^{Adk}\| \leqq \exp(\mathrm{Tr}|A-I|) \leqq$$
$$\leqq \exp\left((1+\|T-I\|_1)^2-1\right).$$

§ 6. Infinite dimensional adjoints and minors

Let us now consider $\mathfrak{K}$ an infinite dimensional Hilbert space and $A\in\mathscr{L}(\mathfrak{K})$ so that $\mathrm{rank}\,(I-A)<\infty$. From the preceding considerations we easily infer the existence of an operator $A^{Adk}\in\mathscr{L}(\mathfrak{K}^{\wedge k})$ satisfying

(6.1)
$$A^{Adk}A^{\wedge k} = A^{\wedge k}A^{Adk} = \det(A)I_{\mathfrak{K}^{\wedge k}};$$

(6.2)
$$\langle A^{Adk}(f_1\wedge\dots\wedge f_k), f_1\wedge\dots\wedge f_k\rangle = \det\left(P+(I-P)A(I-P)\right),$$

for P the orthogonal projection onto the linear span of the orthonormal system $\{f_1, \dots, f_k\}$;

(6.3)
$$\|A^{Adk}\| \leqq \exp\left((1+\|A-I\|_1)^2-1\right).$$

Also for $A_1, A_2\in\mathscr{L}(\mathfrak{K})$ with rank $(I-A_j)<\infty$, $j=1, 2$, we have

(6.4)
$$(A_1A_2)^{Adk} = A_2^{Adk}A_1^{Adk}.$$

Let $A\in\mathscr{L}(\mathfrak{K})$ now be such that $I-A\in\mathscr{C}_1(\mathfrak{K})$ and let A_n be such that rank $(I-A_n)<\infty$ and $\lim_{n\to\infty}\|A-A_n\|_1=0$.

Using the fact that the function

$$\mathscr{C}_1(\mathfrak{K})\ni X \to \det(I+X)$$

is continuous, it follows from (6.2—3) that the sequence A_n^{Adk} converges weakly. The limit, which will be denoted by A^{Adk}, satisfies (6.2—3). Because $A_n^{\wedge k}$ converges to $A^{\wedge k}$ in norm and $\det(A_n)\to\det(A)$ we also obtain property (6.1) for A^{Adk}. Using now (6.2—3) it follows that:

$$A, A_n\in I+\mathscr{C}_1(\mathfrak{R}) \quad \text{and} \quad \|A_n-A\|_1\to 0 \quad \text{imply} \quad A_n^{Adk}\rightharpoonup A^{Adk}. \tag{6.5}$$

Property (6.4) for $A_1, A_2\in I+\mathscr{C}_1(\mathfrak{R})$ follows from (6.1), provided A_1, A_2 are invertible, and can be extended using (6.5) to the case when only A_1 is invertible. Using (6.5) once again it follows that (6.4) holds in the general case.

We have shown in § 5.1 that the function $Y\to\det(I+Y)$ is analytic on the Banach space $\mathscr{C}_1(\mathfrak{R})$. Using (6.2—3) we infer that for $\xi,\eta\in\mathfrak{R}^{\wedge k}$ the mapping

$$\mathbf{C}\ni\lambda\to\langle(I+X+\lambda Y)^{Adk}\xi,\eta\rangle$$

is analytic when $X, Y\in\mathscr{C}_1(\mathfrak{R})$.

From this fact and from (6.3), using [3], Proposition 2 it follows that

$$\mathscr{C}_1(\mathfrak{R})\ni X\to\langle(I+X)^{Adk}\xi,\eta\rangle$$

for $\xi,\eta\in\mathfrak{R}^{\wedge k}$ is analytic.

This again implies the following stronger fact: *the mapping*

$$\mathscr{C}_1(\mathfrak{R})\ni X\to(I+X)^{Adk}\in\mathscr{L}(\mathfrak{R}^{\wedge k}) \tag{6.6}$$

is analytic (in particular continuous with respect to the norm topologies).

Let us also remark that for any contraction $A\in I+\mathscr{C}_1(\mathfrak{R})$ the adjoints A^{Adk} are contractions. This is obvious if $\dim\mathfrak{R}=n<\infty$ (since in this case $\|A^{Adk}\|=\|A^{\wedge(n-k)}\|$) and follows in the general case by a simple limit argument.

We are now going to define the minors of an infinite determinant. Let $\mathfrak{M}$ and $\mathfrak{N}$ be two closed subspaces of $\mathfrak{R}$, $P_{\mathfrak{M}}$ and $P_{\mathfrak{N}}$ the corresponding projections, and suppose there is a unitary operator $U\in I+\mathscr{C}_1(\mathfrak{R})$ such that $U\mathfrak{M}=\mathfrak{N}$. Then for $A\in I+\mathscr{C}_1(\mathfrak{R})$ the minor of $\det(A)$ corresponding to the triple $(\mathfrak{M},\mathfrak{N},U)$ is

$$\det(UP_{\mathfrak{M}}A|\mathfrak{N}). \tag{6.7}$$

The definition makes sense because it is easily seen that $UP_{\mathfrak{M}}A|\mathfrak{N}\in I_{\mathfrak{N}}+\mathscr{C}_1(\mathfrak{N})$. In case $\mathfrak{N}$ (and hence $\mathfrak{M}$ also) is of finite codimension in $\mathfrak{R}$, we shall say that $\det(UP_{\mathfrak{M}}A|\mathfrak{N})$ is a minor of corank $\dim\mathfrak{M}^{\perp}$.

Let $\det(UP_{\mathfrak{M}}A|\mathfrak{N})$ be a minor of corank k of A. Then, by (6.2)

$$\begin{aligned}\det(UP_{\mathfrak{M}}A|\mathfrak{N}) &= \det(P_{\mathfrak{N}}UAP_{\mathfrak{N}}+(I-P_{\mathfrak{N}})) = \\ &= \langle(UA)^{Adk}(e_1\wedge\ldots\wedge e_k), e_1\wedge\ldots\wedge e_k\rangle\end{aligned} \tag{6.8}$$

for $\{e_1,\ldots,e_k\}$ an orthonormal basis of $\mathfrak{R}\ominus\mathfrak{N}$. Thus the minors of corank k of A coincide with some matrix elements of $(UA)^{Adk}=A^{Adk}U^{Adk}$.

§ 7. Determinants of contractive analytic functions

Let $\Theta\in H^\infty(\mathscr{L}(\mathfrak{K}))$ be a contractive function (here $\mathfrak{K}$ denotes as usual a separable Hilbert space). Let us suppose that $I-\Theta(\lambda)$ is nuclear for $\lambda\in D$ and let $\{e_n\}_{n=1}^\infty$ be an orthonormal basis of $\mathfrak{K}$. The functions

$$d_n(\lambda) = \det\left[\langle\Theta(\lambda)e_i, e_j\rangle\right]_{1\leqq i,j\leqq n} = \det\left(P_n\Theta(\lambda)P_n+(1-P_n)\right)$$

(here P_n denotes the orthogonal projection onto the linear span of $\{e_1, \ldots, e_n\}$) are analytic,

$$(7.1)\qquad |d_n(\lambda)| \leqq 1,$$

and

$$(7.2)\qquad \lim_{n\to\infty} d_n(\lambda) = \det\left(\Theta(\lambda)\right).$$

From (7.1) and (7.2) we infer, by the Vitali—Montel theorem, that $\det\left(\Theta(\lambda)\right)$ is an analytic function. A similar argument shows that the functions $\lambda\to\left(\Theta(\lambda)\right)^{Adk}$ are analytic and contractive (cf. § 6) and that

$$(7.3)\qquad \Theta^{\wedge k}\Theta^{Adk} = \Theta^{Adk}\Theta^{\wedge k} = \det(\Theta)I_{\mathfrak{K}^{\wedge k}}$$

In particular, if $\Theta(\lambda)$ is invertible for some $\lambda\in D$, it follows that Θ has a scalar multiple (cf. [13], ch. V, § 6).

In case $\mathfrak{M}, \mathfrak{N}$ are subspaces of $\mathfrak{K}$ of finite codimension and $U\in I+\mathscr{C}_1(\mathfrak{K})$ is a unitary operator such that $U\mathfrak{M}=\mathfrak{N}$, the function $\lambda\to\det\left(UP_{\mathfrak{M}}\Theta(\lambda)|\mathfrak{N}\right)$ is analytic and of modulus $\leqq 1$. We call such a function a minor of Θ of corank $\dim\mathfrak{M}^\perp$.

Let us denote by $\delta_r(\Theta)$ the greatest common inner divisor of the minors of corank r of Θ $(r=0, 1, 2, \ldots)$. For $r=0$, $\delta_0(\Theta)$ coincides with the inner factor of $\det(\Theta(\lambda))$. From (6.8) it follows that $\delta_r(\Theta)$ coincides with the greatest common inner divisor of the matrix elements of Θ^{Adr}.

Lemma 7.1. *$\delta_{r+1}(\Theta)$ divides $\delta_r(\Theta)$ for each r.*

Proof. We have to prove that $\delta_{r+1}(\Theta)$ divides each minor of corank r of Θ. Clearly it suffices to prove that $\delta_1(\Theta)$ divides $\det(\Theta)$ or, equivalently,

$$\det(\Theta)H^2(\mathfrak{K}) \subset \delta_1(\Theta)H^2(\mathfrak{K}).$$

But this easily follows from the relation $\Theta\Theta^{Ad1}=\det(\Theta)I_{\mathfrak{K}}$. Indeed, $\Theta^{Ad1}H^2(\mathfrak{K})\subset \subset\delta_1(\Theta)H^2(\mathfrak{K})$ and, since Θ is analytic,

$$\det(\Theta)H^2(\mathfrak{K}) = \Theta\Theta^{Ad1}H^2(\mathfrak{K}) \subset \Theta\delta_1(\Theta)H^2(\mathfrak{K}) \subset \delta_1(\Theta)H^2(\mathfrak{K}).$$

Lemma 7.2. *The greatest common inner divisor of the functions $\delta_j(\Theta)$ $(j=1, 2, \ldots)$ is* 1.

Proof. Let us denote by m the greatest common inner divisor of the family $\{\delta_j(\Theta)\}_0^\infty$ and let $\{e_j\}_{j=1}^\infty$, be an orthonormal basis of $\mathfrak{K}$. Since $\Theta^{Adr}H^2(\mathfrak{K}^{\wedge r})\subset \subset mH^2(\mathfrak{K}^{\wedge r})$ for each r, we have

$$|m(0)| \geqq |\langle \Theta(0)^{Adr}(e_1\wedge e_2\wedge\ldots\wedge e_r), e_1\wedge e_2\wedge\ldots\wedge e_r\rangle| = = |\det((I-P_r)\Theta(0)(I-P_r)+P_r)|,$$

where P_r denotes as usual the orthogonal projection onto the linear span of $\{e_1,\ldots,e_r\}$. We infer

$$|m(0)| \geqq \limsup_{r\to\infty} |\det((I-P_r)\Theta(0)(I-P_r)+P_r)| = 1$$

and the lemma follows.

Let us also note the relations

$$\delta_j(\Theta^\sim) = \delta_j(\Theta)^\sim \quad (j=1,2,\ldots) \tag{7.4}$$

which hold for each function Θ of the type considered in this section.

§ 8. Weak contractions

Let us recall that a contraction T acting on a Hilbert space $\mathfrak{H}$ is a weak contraction if its spectrum does not cover the unit disk D and $I-T^*T$ is a nuclear operator. T is a weak contraction if and only if T^* is a weak contraction.

If a weak contraction T is of class C_{00} (that is $T^n\to 0$ and $T^{*n}\to 0$ strongly as $n\to\infty$), then T is of class C_0 and acts on a necessarily separable space. The proof of this fact goes as follows (cf. [13], Ch. VIII, § 1).

If we put

$$T_\lambda = (T-\lambda I)(I-\bar\lambda T)^{-1}, \quad \lambda\in D \tag{8.1}$$

we have

$$I-T_\lambda^*T_\lambda = X_\lambda^*(I-T^*T)X_\lambda,\ X_\lambda = (1-|\lambda|^2)^{1/2}(I-\bar\lambda T)^{-1}. \tag{8.2}$$

So T is a weak contraction if and only if T_λ is a weak contraction. Moreover, we have $(T_\lambda)_{-\lambda}=T$. Therefore we may suppose without loss of generality that T is invertible. Let $\{\mu_j\}_1^n$ $(n\leqq\aleph_0)$ be the eigenvalues of $(I-T^*T)|\mathfrak{D}_T$, $\mathfrak{D}_T=((I-T^*T)H)^-$ (multiple eigenvalues repeated according to their multiplicities). We have $\mu_j\neq 1$ because $\ker T=\{0\}$.

Let $\{\varphi_j\}_1^n$ be an orthonormal basis of $\mathfrak{D}_T$ such that $(I-T^*T)\varphi_j=\mu_j\varphi_j$. It is easy to verify that the system $\{\psi_j\}_1^n$, where $\psi_j=(1-\mu_j)^{-1/2}T\varphi_j$, is an orthonormal basis of $\mathfrak{D}_{T^*}$ and that we have also $T^*\psi_j=(1-\mu_j)^{1/2}\varphi_j$.

Let us denote by U the unitary operator determined by

$$U:\mathfrak{D}_T\to\mathfrak{D}_{T^*},\ U\varphi_j=-\psi_j. \tag{8.3}$$

Then the operator $(U+T)\mathfrak{D}_T$ is nuclear. Indeed,

$$(U+T)h = \sum_{j=1}^{n}\left((1-\mu_j)^{1/2}-1\right)(h,\varphi_j)\psi_j, \quad h\in\mathfrak{D}_T$$

and from the relations

$$\lim_{\mu\to 0}\mu^{-1}\left(1-(1-\mu)^{1/2}\right) = 1/2, \quad \sum_{j=1}^{n}\mu_j<\infty$$

we infer

$$\sum_{j=1}^{n}\left(1-(1-\mu_j)^{1/2}\right)<\infty.$$

Furthermore, if $\Theta_T\in H_i^\infty(\mathscr{L}(\mathfrak{D}_T,\mathfrak{D}_{T^*}))$ is the characteristic function of T, $U-\Theta_T(\lambda)$ is nuclear for $\lambda\in D$. Indeed,

$$U-\Theta_T(\lambda) = (U+T)|\mathfrak{D}_T-\lambda D_{T^*}(I-\lambda T^*)^{-1}D_T|\mathfrak{D}_T \quad (D_T=(I-T^*T)^{1/2})$$

and since D_T and D_{T^*} are Hilbert—Schmidt operators because T is a weak contraction, $\lambda D_{T^*}(I-\lambda T^*)^{-1}D_T$ is nuclear. Thus the function $\Theta\in H_i^\infty(\mathscr{L}(\mathfrak{D}_T))$ defined by $\Theta(\lambda)=U^*\Theta_T(\lambda)$ coincides with Θ_T and $I-\Theta(\lambda)$ is nuclear for $\lambda\in D$.

Let us put

$$d_T(\lambda) = \det(\Theta(\lambda)), \quad \delta_j(T)=\delta_j(\Theta), \quad (j=0,1,2,\ldots). \tag{8.4}$$

We have $d_T(0)=\prod_{j=1}^{n}(1-\mu_j)^{1/2}\neq 0$ and from (7.3) (with $k=1$) it follows that d_T is a scalar multiple of Θ. As in [13], Theorem VI. 5.2 we obtain

Lemma 8.1. *Each weak contraction T of class C_{00} is a C_0 contraction and its minimal function coincides with $\delta_0(T)/\delta_1(T)$.*

Let us remark that we have a converse: suppose $\Theta\in H_i^\infty(\mathscr{L}(\mathfrak{K}))$ is such that $\Theta(\lambda)\in I+\mathscr{C}_1(\mathfrak{K})$, $\lambda\in D$, and $\det(\Theta)\not\equiv 0$. Since $\det(\Theta)$ is then a scalar multiple of Θ (by (7.3) with $k=1$), it follows that Θ coincides with the characteristic function of an operator T of class C_0 and from [13], Ch. IV § 1 it follows that $\operatorname{tr}(I-T^*T)=$ $=\operatorname{tr}\left(I-\Theta(0)^*\Theta(0)\right)<\infty$ so that T is a weak contraction. Let us also note that the relations

$$d_{T^*}=d_T^{\sim}, \quad \delta_j(T^*)=\delta_j(T)^{\sim} \quad (j=0,1,\ldots) \tag{8.5}$$

hold for each weak contraction T.

Proposition 8.2. *Let T be a weak C_0 contraction acting on the Hilbert space $\mathfrak{H}$ and let $T=\begin{bmatrix}T_1 & X\\ 0 & T_2\end{bmatrix}$, $\mathfrak{H}=\mathfrak{H}_1\oplus\mathfrak{H}_2$ be the triangularization associated with the T-invariant subspace $\mathfrak{H}_1$. Then T_1 and T_2 are weak C_0 contractions and we have*

$$d_T=d_{T_1}d_{T_2}, \quad \delta_0(T)=\delta_0(T_1)\delta_0(T_2).$$

Proof. We may suppose without loss of generality that T is invertible, thus $m_T(0)\neq 0$. By [13], Proposition III. 6.1, T_1 and T_2 are C_0 operators and m_{T_1}, m_{T_2} are divisors of m. It follows that $m_{T_1}(0)\neq 0$, $m_{T_2}(0)\neq 0$ so that T_1 and T_2 are invertible. Moreover, we have

$$I_{\mathfrak{H}_1}-T_1^*T_1 = P_{\mathfrak{H}_1}(I-T^*T)|\mathfrak{H}_1, \quad I_{\mathfrak{H}_2}-T_2T_2^* = P_{\mathfrak{H}_2}(I-TT^*)|\mathfrak{H}_2,$$

thus T_1 and T_2 are weak contractions.

By [13] Theorem VII.1.1 and Proposition VII.2.1, we can associate with the invariant subspace $\mathfrak{H}_1$ a regular factorization

$$\Theta_T(\lambda) = \Theta_2(\lambda)\Theta_1(\lambda) \tag{8.6}$$

such that the characteristic functions $\Theta_{T_1}(\lambda)$, $\Theta_{T_2}(\lambda)$ coincide with the pure parts of $\Theta_1(\lambda)$, $\Theta_2(\lambda)$, respectively. Then we have

$$\Theta_j(\lambda) = U_j'\begin{bmatrix}\Theta_{T_j}(\lambda) & 0\\ 0 & I_j\end{bmatrix}U_j'' \tag{8.7}$$

where U_j', U_j'' are unitary operators and I_j denotes the identity operator on some Hilbert space ($j=1, 2$). Now, from the consideration preceding Lemma 8.1, it follows that $I-U_j^{0*}\Theta_{T_j}(\lambda)$ is nuclear and $d_{T_j}(\lambda)=\det\left(U_j^{0*}\Theta_{T_j}(\lambda)\right)$ for some unitary operators U_j^0 ($j=1, 2$). With the notation

$$U_j = U_j'\begin{bmatrix}U_j^0 & 0\\ 0 & I_j\end{bmatrix}U_j''$$

we see that $I-U_j^*\Theta_j(\lambda)$ is nuclear and

$$d_{T_j}(\lambda) = \det\left(U_j^*\Theta_j(\lambda)\right). \tag{8.8}$$

Using (8.6) and (8.7) we obtain

$$U^*\Theta_T(\lambda) = U^*U_2U_1[U_1^*(U_2^*\Theta(\lambda))U_1](U_1^*\Theta_1(\lambda))\cdot \tag{8.9}$$

From this relation it follows that $I_{\mathfrak{D}_T}-U^*U_2U_1$ is a nuclear operator such that $\det(U^*U_1U_2)$ exists. Using (8.8-9) and (5.4) we then obtain

$$\begin{aligned} d_T(\lambda) &= \det(U^*U_2U_1)\det\left(U_1^*(U_2^*\Theta_2(\lambda))U_1\right)\det\left(U_1^*\Theta_1(\lambda)\right) = \\ &= \det(U^*U_2U_1)\det\left(U_2^*\Theta_2(\lambda)\right)\det\left(U_1^*\Theta_1(\lambda)\right) = \\ &= \det(U^*U_2U_1)\,d_{T_2}(\lambda)\,d_{T_1}(\lambda). \end{aligned}$$

The relation $\delta_0(T)=\delta_0(T_1)\delta_0(T_2)$ follows by taking the inner factors in the last obtained relations. The proposition is proved.

Remark 8.3. This proposition is a generalization of [13], Lemma IX. 3.1.

Lemma 8.4. *A Jordan operator $S(M)$, $M=\{m_j\}_1^\infty$, is a weak contraction if and only if $\sum_{j=1}^{\infty}(1-|m_j(0)|)<\infty$. In this case we have $d_{S(M)}=\delta_0(S(M))=\prod_{j=1}^{\infty} m_j$, where $\prod_{j=1}^{\infty} m_j$ means the limit of some converging subsequence of $\{m_1 m_2 \dots m_n\}_{n=1}^\infty$.*

Proof. For any inner function $m\in H^\infty$ we have

$$(I_{\mathfrak{H}(m)}-S(m)S(m)^*)h = P_{\mathfrak{H}(m)}(I-UU^*)h = (h, c_0)P_{\mathfrak{H}(m)}c_0 =$$
$$= (h, c_0)(1-\overline{m(0)}\,m)$$

$(h\in\mathfrak{H}(m))$, where U denotes the unilateral shift on H^2 and c_0 is the constant functions $c_0\equiv 1$. Thus $I-S(m)S(m)^*$ is of rank one and has norm $(1-\overline{m(0)}m, c_0)=1-|m(0)|^2$. It follows that the trace norm of $I-S(M)S(M)^*$ equals $\sum_{j=1}^{\infty}(1-|m_j(0)|^2)$. We have only to remark that

$$1-|m_j(0)|\leqq 1-|m_j(0)|^2\leqq 2(1-|m_j(0)|).$$

The equality $d_{S(M)}=\prod_{j=1}^{\infty} m_j$ obviously follows from the special form of the characteristic function of $S(M)$. So it remains only to prove that $\prod_{j=1}^{\infty} m_j$ is an inner function. To see this, let us remark that $\prod_{j=1}^{\infty} m_j$ and $\prod_{j=n}^{\infty} m_j$ have the same outer factor, such that this outer factor must be 1 because $\left|\prod_{j=n}^{\infty} m_j(\lambda)\right|\to 1$ for each $\lambda\in D$. The lemma is proved.

From now on T will denote a weak C_0 contraction acting on $\mathfrak{H}$, $\Theta\in H_i^\infty(\mathcal{L}(\mathfrak{K}))$ will denote a function coinciding with the characteristic function of T and $\Theta(\lambda)\in I+\mathscr{C}_1(\mathfrak{K})$, $\lambda\in D$. We shall also denote by $S(M)$, $M=\{m_j\}_1^\infty$, the Jordan model of T. From the relation

$$\Theta^{\wedge r}\Theta^{Adr}=\Theta^{Adr}\Theta^{\wedge r}=d_T\cdot I_{\mathfrak{K}^{\wedge r}}, \quad \text{see (7.3)},$$

we infer, because $\Theta^{\wedge r}$ is two-sided inner, that $\delta_0(T)/\delta_r(T)$ is the least inner scalar multiple of $\Theta^{\wedge r}$. Thus we have

$$d_r(T)=\delta_0(T)/\delta_r(T). \tag{8.10}$$

Theorem 8.5. *A C_0 contraction T is a weak contraction if and only if its Jordan model $S(M)$, $M=\{m_j\}_1^\infty$, is a weak contraction.*

Proof. That T is a weak contraction if $S(M)$ is so follows from Proposition 4.3, via Lemma 8.4. So let us assume that T is a weak contraction. Then, by Corollary 3.3.

and relation (8.10) it follows that $m_1 m_2 \dots m_r$ divides $\delta_0(T)$ for each r. If we suppose T is invertible, we have $\delta_0(T)(0) \neq 0$ and from the inequality

$$|m_1(0) \dots m_r(0)| \geqq |\delta_0(T)(0)|$$

it follows that the infinite product $\prod_{j=1}^{\infty} |m_j(0)|$ converges. Therefore $\sum_{j=1}^{\infty} (1-|m_j(0)|) < \infty$ and our theorem follows by Lemma 8.4.

Proposition 8.6. *For each weak C_0 contraction T, the determinant function d_T is an inner function.*

Proof. Let us write the inner-outer decomposition of d_T

(8.11) $$d_T = d_i d_o.$$

Because d_i is a scalar multiple of $\Theta^{\wedge k}$, there exists a contractive function $\Omega^{(k)} \in H^\infty(\mathscr{L}(\mathfrak{K}^{\wedge k}))$ such that

(8.12) $$\Omega^{(k)} \Theta^{\wedge k} = \Theta^{\wedge k} \Omega^{(k)} = d_i I_{\mathfrak{K}^{\wedge k}}.$$

Then, by (7.3) and (8.12) we have

$$\Theta^k (d_o \Omega^{(k)} - \Theta^{Adk}) = 0$$

so that (Θ^k being inner)

(8.13) $$\Theta^{Adk} = d_o \Omega^{(k)}.$$

Let $\{e_i\}_{i=1}^{\infty}$ be an orthonormal basis of $\mathfrak{K}$ and denote by P_n the orthogonal projection onto the linear span of $\{e_1, \dots, e_n\}$. By (8.13) we have

$$\langle \Theta^{Adk}(e_1 \wedge \dots \wedge e_k), e_1 \wedge \dots \wedge e_k \rangle = d_o \langle \Omega^{(k)}(e_1 \wedge \dots \wedge e_k), e_1 \wedge \dots \wedge e_k \rangle$$

and therefore

$$|d_o(0)| \geqq \limsup_{k \to \infty} |\langle \Theta(0)^{Adk}(e_1 \wedge \dots \wedge e_k), e_1 \wedge \dots \wedge e_k \rangle| =$$

$$= \limsup_{k \to \infty} |\det((I-P_k)\Theta(0)(I-P_k)+P_k)| = 1.$$

It follows that $|d_o(0)| = 1$ so that $|d_o| \equiv 1$. The proposition follows.

We are now able to prove that the determinant function of a weak C_0 contraction is a quasi-similarity invariant.

Theorem 8.7. *For each C_0 contraction T with Jordan model $S(M)$, $M = \{m_j\}_1^{\infty}$, we have*

(8.14) $$m_j = \delta_{j-1}(T)/\delta_j(T);$$

(8.15) $$d_T = d_{S(M)} = \prod_{j=1}^{\infty} m_j.$$

Proof. From (8.10) it follows that $\delta_{j-1}(T)/\delta_j(T)=d_j(T)/d_{j-1}(T)$ so the relation (8.14) obviously follows from Corollary 3.4.

For the second relation let us write (8.10) under the form

$$d_T \doteq \delta_0(T) = m_1 m_2 \dots m_n \cdot \delta_n(T) \tag{8.16}$$

(cf. Corollary 3.4). From Lemma 7.2 and Lemma 1 of [12] it follows that d_T coincides with the least common inner multiple of the family $\{m_1 m_2 \dots m_n\}_{n=1}^{\infty}$, which coincides with $d_{S(M)}$.

The theorem follows.

References

[1] H. Bercovici, C. Foiaş, B. Sz.-Nagy, Compléments à l'étude des opérateurs de classe C_0. III, *Acta Sci. Math.*, **37** (1975), 313—322.

[2] J. Dixmier, *Les C*-algèbres et leurs représentations*, Gauthier-Villars (Paris, 1969).

[3] A. Douady, Le problème des modules pour les sous-espaces analytiques compacts d'un espace analytique donné, *Ann. Inst. Fourier*, **16**: **1** (1966), 1—98.

[4] И. Ц. Гохберг, М. Г. Крейн, *Введение в теорию линейных несамосопряженных операторов в гильбертовом пространстве*, Наука (Москва, 1965).

[5] Л. Е. Исаев, *Докл. АН СССР*, **178** (1968) 783—785.

[6] Д. П. Желобенко, *Компактные группы Ли и их представления*, Наука (Москва, 1970).

[7] А. А. Кириллов, Представления безконечномерной унитарной группы, *Докл. Акад Наук*, **212** (1973), 288—290.

[8] B. Moore III, E. A. Nordgren, On quasisimilarity and quasiequivalence, *Acta Sci. Math.*, **34** (1973), 311—316.

[9] М. А. Наймарк, *Теория представлений групп*, Наука (Москва, 1976).

[10] E. A. Nordgren, On quasiequivalence of matrices over H^∞, *Acta Sci. Math.*, **34** (1973), 301—310.

[11] I. E. Segal, The structure of a class of representations of the unitary group on a Hilbert space, *Proc. Amer. Math. Soc.*, **8** (1957), 197—203.

[12] B. Sz.-Nagy, Diagonalization of matrices over H^∞, *Acta Sci. Math.*, **38** (1976), 223—238.

[13] B. Sz.-Nagy, C. Foiaş, *Harmonic analysis of operators on Hilbert space*, North Holland /Akadémiai Kiadó (Amsterdam/ Budapest, 1970).

[14] B. Sz.-Nagy, C. Foiaş, Modèle de Jordan pour une classe d'opérateurs de l'espace de Hilbert, *Acta Sci. Math.*, **31** (1970), 91—115.

[15] B. Sz.-Nagy, C. Foiaş, Compléments à l'étude des opérateurs de classe C_0, *Acta Sci. Math.*, **31** (1970), 287—296.

[16] J. Szűcs, Diagonalization theorems for matrices over certain domains, *Acta Sci. Math.*, **36** (1974), 193—201.

[17] H. Weyl, *Gruppentheorie und Quantenmechanik*, 2nd ed. (Leipzig, 1931).

[18] H. Weyl, *The classical groups* (Princeton, 1939).

[19] P. Y. Wu, Jordan model for weak contractions, *Acta Sci. Math.*, to appear.

H. BERCOVICI
BD. BĂLCESCU 32
BUCUREŞTI, ROMANIA

D. VOICULESCU
INCREST
CALEA VICTORIEI 114
SECTOR 1
BUCUREŞTI 22, ROMANIA

Acta Sci. Math., 39 (1977), 233—243

Normal dilations and operator approximations

R. H. BOULDIN and D. D. ROGERS

§ 1. Preliminaries

This paper continues the research presented in [2]; the earlier results are refined and extended in several directions. Consideration is given to best approximation by self-adjoint operators as well as by non-negative operators. A best approximation from the first set is a "self-adjoint approximant" and from the second set is a "positive approximant". For elementary facts about positive approximants the reader is directed to [8] and [5]; for self-adjoint approximants, check [8] and [6]. A general reference for terms not explained is [9].

For a given operator each set of approximants is convex; the main results of this paper identify broad classes of operators for which each of these sets of approximants is infinite-dimensional. (For a discussion of the dimension of a convex set see [16, p. 7].) Moreover, the constructive proofs of these results develop concrete techniques for obtaining approximants for a given operator.

In [8] HALMOS showed that for any (bounded linear) operator $A=B+iC$ ($B=B^*$, $C=C^*$) a positive approximant is $B+(\delta^2-C^2)^{1/2}$ where $\delta=\delta(A)$ is the distance from A to the non-negative operators; this positive approximant, denoted P_0, is referred to as the "Halmos positive approximant". Halmos also showed that B is a self-adjoint approximant for A, or equivalently the distance from A to the self-adjoint operators is $\|C\|$.

The work in this paper exploits a fundamental relation between an operator T and various normal dilations of T. Before establishing this relation we recall the following two lemmas which state some previously known facts in a form appropriate to this work. These facts are proved in [12] and [10], respectively.

Received August 4, 1976.

1.1. Lemma. *If N is a normal element of a C^*-algebra and if A is any element, then*

$$\|A-N\| \geqq h(\sigma(A), \sigma(N))$$

where $h(M_1, M_2)=\sup \{\text{dist}(m_1, M_2): m_1 \in M_1\}$.

1.2. Lemma. *For any normal operator N the following formula holds*:

$$\|N-P_A(N)\| = h(\sigma(N), A)$$

where $P_A(N)$ is a best approximation for N from the normal operators with spectrum in the nonempty closed set A, denoted $\mathcal{N}(A)$.

The notation of the preceding lemmas is continued in the next theorem.

1.3. Theorem. *Assume A is a closed convex subset of the real line. If T is an operator on H with normal dilation N on $K \supset H$ such that $\sigma(N) \subset \sigma(T)$, then*

$$\|T-P_A(T)\| = h(\sigma(T), A) = h(\sigma(N), A) = \|N-P_A(N)\|.$$

Furthermore, provided Q is the orthogonal projection of K onto H, $QP_A(N)Q|H$ is a best approximation for T from $\mathcal{N}(A)$.

Proof. It follows from the hypothesis and the two preceding lemmas that

$$\|N-P_A(N)\| = h(\sigma(N), A) \leqq h(\sigma(T), A) \leqq \|T-P_A(T)\|.$$

Because $QLQ|H$ belongs to $\mathcal{N}(A)$ on H for any $L \in \mathcal{N}(A)$ on K, the following inequality holds

$$\|T-P_A(T)\| \leqq \|T-QP_A(N)Q\| = \|Q)(N-P_A(N)Q\| \leqq \|N-P_A(N)\|.$$

The inequalities prove the theorem.

We use Theorem 1.3 in each of the next four sections. In sections § 2 and § 3 it is assumed that T is subnormal with minimal normal extension N, and in sections § 4 and § 5 it is assumed that T is a Toeplitz operator with N the corresponding Laurent operator. It should be noted that other general hypotheses guarantee that $\sigma(N) \subset \sigma(T)$ — for example, if $\sigma(T)$ is a spectral set for T —, then such a normal dilation N exists.

§ 2. Positive approximants of a subnormal operator

In the next theorem and throughout the remainder of this section the symbol T will denote a subnormal operator defined on H and the symbol N will denote a normal operator defined on K that is the minimal normal extension of T. Also, N equals $B+iC$ where B and C are self-adjoint operators.

2.1. Theorem. *For any subnormal operator T one has $\|T-P_\Lambda(T)\|=\|N-P_\Lambda(N)\|$ where N is the minimal normal extension of T and Λ is a nonempty, closed, convex subset of the real line. Moreover, the compression of any $P_\Lambda(N)$ to H is a best approximation for T from $\mathcal{N}(\Lambda)$.*

Proof. Recall $\sigma(T)$ differs from $\sigma(N)$ only by filling in some holes (see [9, Problems 157, 158]). Thus, the above theorem is a special case of Theorem 1.3.

A curious consequence of the preceding theorem is that the norm of the imaginary part of the subnormal operator T, denoted $\|\mathrm{im}\, T\|$, equals the norm of the imaginary part of the minimal normal extension, denoted $\|\mathrm{im}\, N\|$. The first norm is the distance $\|T-P_{\mathbf{R}}(T)\|$ and the second norm is the distance $\|N-P_{\mathbf{R}}(N)\|$.

Let T be a subnormal operator on H. There is a subspace H_1 which reduces T to a normal operator and is maximal with respect to this property. Moreover, the orthogonal complement of H_1, denoted $H_1^\perp$, includes no subspace M invariant under T such that $T|M$ is normal. (See Proposition 1.1 of [1], for example.) Thus, T is the direct sum of a normal operator and a completely nonnormal operator. Since positive approximants of a normal operator are studied extensively in [3], attention is now concentrated on completely nonnormal subnormal operators.

Let Γ denote that set of z such that the distance from z to $[0, \infty)$ is exactly $\delta(T)$ and $\mathrm{re}\, z$ does not exceed $\|T\|$.

2.2. Lemma. *Let T be a completely nonnormal subnormal operator. Then T has infinitely many distinct approximate eigenvalues, say $\{z_1, z_2, \ldots\}$, such that $\{z_1, z_2, \ldots\}$ does not intersect Γ.*

Proof. If $\sigma(T)$ were contained in Γ, then it would follow that T is normal (see [15, Corollary 2] or [11, Theorem 1]). Thus, $\sigma(T)$ and the topological boundary of $\sigma(T)$, denoted bdry $\sigma(T)$, contains some z_0 such that $z_0 \notin \Gamma$. If z_0 were isolated from its complement in $\sigma(T)$, then it would follow that z_0 is an eigenvalue for T and the corresponding eigenspace reduces T to a normal operator (see [14, Theorem 2 and Lemma 6]). Thus, z_0 must be an accumulation point for $\sigma(T)$ and it follows that bdry $\sigma(T)$ contains an infinite number of points off the contour Γ. Since bdry $\sigma(T)$ consists of approximate eigenvalues, the lemma is proved.

In the next lemma and throughout the remainder of this section the symbol P_0 will denote the Halmos positive approximant of the normal operator $N=B+iC$; thus, P_0 is $B+(\delta^2-C^2)^{1/2}$. It should not be confused with the Halmos positive approximant for the subnormal operator T.

2.3. Lemma. *Let $E(\cdot)$ denote the spectral measure for the normal operator N and let K_0 denote the subspace $(P_0K)^- \cap ((\delta^2-C^2)K)^-$.*

(i) *The subspaces K_0 and $E(\Gamma^c)K$ are equal.*

(ii) *If D is a compact set not intersecting Γ, then $E(D)K$ reduces (δ^2-C^2) and P_0 to invertible operators.*

Proof. The first statement follows from Lemma 2.1 of [4].

It follows from Lemma 1.2 that $|(\mathrm{re}\, z)_- + i(\mathrm{im}\, z)| \leqq \delta(N)$ for every z in $\sigma(N)$ where x_- denotes the maximum of $\{-x, 0\}$. Since D and Γ are both compact, there is a positive distance between them. It follows that there is a positive number γ such that $|(\mathrm{re}\, z)_- + i(\mathrm{im}\, z)| \leqq \delta - \gamma$ for every z in the intersection of D and $\sigma(N)$. Consequently the sets $\{\delta^2 - (\mathrm{im}\, z)^2 : z \in D \cap \sigma(N)\}$ and $\{\mathrm{re}\, z + (\delta^2 - (\mathrm{im}\, z)^2)^{1/2} : z \in D \cap \sigma(N)\}$ are bounded away from zero, and these sets are the spectra of $\delta^2 - C^2$ and P_0 restricted to $E(D)K$, respectively.

2.4. Theorem. *Let T be a completely nonnormal subnormal operator defined on H with minimal normal extension N defined on K. Then the real dimension of the convex set $\mathscr{P}(T)$ of positive approximants of T, denoted* $\dim \mathscr{P}(T)$, *is infinite.*

Proof. Let z_1 and z_2 be approximate eigenvalues of T off the contour Γ. Let α_j and β_j be real numbers such that $z_j = \alpha_j + i\beta_j$ and let $\{e_{j1}, e_{j2}, \dots\}$ be a normalized approximate eigenvector for T corresponding to z_j. Because H is invariant under N, z_j is an approximate eigenvalue for N and $\{e_{j1}, e_{j2}, \dots\}$ is a corresponding approximate eigenvector for N.

Let D be a compact set not intersecting Γ and containing $\{z_1, z_2\}$ in its interior. Define τ by the equation

$$2\tau = \inf \{\mathrm{re}\, z + (\delta^2 - (\mathrm{im}\, z)^2)^{1/2}, (\delta^2 - (\mathrm{im}\, z)^2)^{1/2} : z \in D \cap \sigma(N)\}$$

and note that the proof of (ii) of Lemma 2.3 implies that τ is positive. The functional calculus for N readily shows that $\lim \{\|f_{jk} - e_{jk}\| : k = 1, 2, \dots\}$ is zero, where $f_{jk} = E(D)e_{jk}$ for $j = 1, 2$. It follows that $\lim \{\|(I-Q)f_{jk}\| : k = 1, 2, \dots\}$ is zero, where Q is the orthogonal projection of K onto H. Replace the original sequences with subsequences if necessary so that $\{f_{1n}, f_{2n}\}$ is linearly independent for each n.

By definition the operator $A(\varrho; n)$ is zero on $(E(D)K)^\perp$, on $K(n) = \mathrm{span}\, \{f_{1n}, f_{2n}\}$ it is the matrix

$$\begin{pmatrix} \tau & \varrho \\ \varrho & \tau \end{pmatrix}$$

and on $E(D)K \ominus K(n)$ it is τI. It will be shown that $P_0 - A(\varrho; n)$ is a positive approximant of N for ϱ in an interval $(0, \varrho_0)$ for all n sufficiently large. If $(N - P_0)|E(D)K$ is written as a matrix relative to $K(n) \oplus (E(D)K \ominus K(n))$, then the nondiagonal entries converge to zero in operator norm by the choice of z_1 and z_2. Thus, it suffices to show that both $\|(N - P_0 + A(\varrho; n))|E(D)K \ominus K(n)\|$ and $\|(N - P_0 + A(\varrho; n))|K(n)\|$ are strictly less than ϱ for appropriate ϱ and n. The first inequality follows from (ii) of Lemma 2.3 and the choice of τ, and the second inequality is proved in the next paragraph.

Define $R(n)$ to be $(N-P_0)|K(n)$ minus the diagonal operator with entries $-(\delta^2-\beta_1^2)^{1/2}+i\beta_1$, $-(\delta^2-\beta_2^2)^{1/2}+i\beta_2$, respectively, and note that $(N-P_0+ + A(\varrho;n))|K(n)$ equals

$$\begin{pmatrix} \tau-(\delta^2-\beta_1^2)^{1/2}+i\beta_1 & 0 \\ 0 & \tau-(\delta^2-\beta_2^2)^{1/2}+i\beta_2 \end{pmatrix}+\begin{pmatrix} 0 & \varrho \\ \varrho & 0 \end{pmatrix}+R(n).$$

By the choice of τ, the norm of the first operator is strictly less than δ and the norms of the remaining two operators can be made arbitrarily small by the choice of ϱ and n, respectively. Thus, $P_0-A(\varrho;n)$ is a positive approximant of N.

Let m be any positive integer; a distinguished set of m positive approximants for N will be constructed. Let $\{z_1, \dots, z_{m+1}\}$ be a set of $m+1$ distinct approximate eigenvalues of T. For each pair $\{z_1, z_j\}$, the preceding construction results in a positive approximant $P_0-A(\varrho;n;j)$ for $j=2, \dots, m+1$.

By Theorem 2.1, $Q(P_0-A(\varrho;n;j))|H$ is a positive approximant for T, where Q is the orthogonal projection of K onto H. Recall from the second paragraph of this proof that $\lim\,\{\|f_{jk}-e_{jk}\| : k=1,2,\dots\}$ is zero for $j=1, \dots, n+1$. It follows that $\{Qf_{j1}, Qf_{j2}, \dots\}$ is an approximate eigenvector for T corresponding to z_j for $j=1, \dots, m+1$. The linear independence of $\{Qf_{1n}, \dots, Qf_{m+1n}\}$ for all n sufficiently large is clear. In order to show that the dimension of $\mathscr{P}(T)$ is at least m it suffices to show the linear independence of

$$\{QA(\varrho;\ n;\ 2)|H, \dots, QA(\varrho;\ n;\ m+1)|H\};$$

thus, consider the matrix of $QA(\varrho;n;j)|H$ compressed to span $\{Qf_{1n}, \dots, Qf_{m+1n}\}$ relative to $\{Qf_{1n}, \dots, Qf_{m+1n}\}$. Make an appropriate choice for ϱ, and note that it is determined by $z_1, z_2, \dots, z_{m+1}, \tau$. Because each entry in the matrix for $QA(\varrho;n;j)|H$ relative to $\{Qf_{1n}, \dots, Qf_{m+1n}\}$ converges to the corresponding entry in the matrix of the compression of $A(\varrho;n;j)$ relative to $\{f_{1n}, \dots, f_{m+1n}\}$ as $n\to\infty$, it is not difficult to show that the first set of matrices are linearly independent for appropriately large n.

In fact, choose n so large that each entry in the matrix of the compression of $QA(\varrho;n;j)|H$ differs from the corresponding entry of the matrix of $A(\varrho;n;j)$ by less than ϱ/m. Denote those matrices by $M_1, \dots, M_m$ and assume that $c_1, \dots, c_m$ are real constants such that

$$0 = c_1M_1+\dots+c_mM_m.$$

By considering each entry in the first row, one obtains m equations of similar form, and each equation implies an inequation of the form

$$|c_j|\varrho < (\varrho/m)\sum_{k=1}^{m}|c_k|.$$

Adding up these inequalities results in a contradiction, which proves the theorem.

Recall the standard decomposition of a subnormal operator that was discussed prior to Lemma 2.2. If T is the orthogonal direct sum $T_1\oplus T_2$, then it is clear that $\delta(T)$ is the maximum of $\{\delta(T_1), \delta(T_2)\}$. Consequently, unless $\delta(T_1)$ equals $\delta(T_2)$ there is much arbitrariness in the approximation of T. For example, if $\delta(T_2)$ exceeds $\delta(T_1)$, then $(P_1-A)\oplus P_2$ is a positive approximant for T, where P_2 is any positive approximant for T_2, P_1 is any positive approximant for T_1 and A is any nonnegative operator dominated by P_1 and having norm dominated by $\delta(T_2)-\delta(T_1)$.

It should be noted that the construction carried out for the minimal normal extension N in the proof of Theorem 2.4 proves the following corollary.

2.5. Corollary. *If the spectrum of the normal operator N has an accumulation point not on the contour Γ consisting of all z with distance to* $[0, \infty)$ *exactly equal to $\delta(N)$, then the dimension of $\mathcal{P}(N)$ is infinite.*

A convex set, for example a disc in the plane, may have uncountably many extreme points, and the only implication about the dimension of the convex set is that it exceeds one. On the other hand, conclusions about the dimension of a convex set have immediate nontrivial implications about the number of extreme points.

2.6. Corollary. *If T is a completely nonnormal subnormal operator, then $\mathcal{P}(T)$ has an infinite number of extreme points.*

A consequence of some results of T. Sekiguchi in [13] is that $\mathcal{P}(T)$ has uncountably many extreme points.

§ 3. Self-adjoint approximants of a subnormal operator

Recall that $E(\cdot)$ denotes the spectral measure of the normal operator $N=B+iC$ with $B=B^*$, $C=C^*$, defined on the Hilbert space K.

3.1. Lemma. *If D is a compact set not intersecting the set $\Sigma=\{z\colon (\operatorname{im} z)^2= =\|C\|^2,\ |z|\leqq\|N\|\}$, then $E(D)K$ reduces $(\|C\|^2-C^2)^{1/2}$ to an invertible operator.*

Proof. Clearly $|\operatorname{im} z|$ does not exceed $\|C\|$ for any z in $\sigma(N)$. Since D and Σ are both compact, there is a positive distance between them. It follows that there is a positive number v such that $|\operatorname{im} z|\leqq\|C\|-\gamma$ for every z in the intersection of D and $\sigma(N)$. Consequently the set $\{(\|C\|^2-(\operatorname{im} z)^2)^{1/2}\colon z\in D\cap\sigma(N)\}$ is bounded away from zero, and this set is the spectrum of $(\|C\|^2-C^2)^{1/2}$ restricted to $E(D)K$.

Since self-adjoint approximants of a normal operator are studied in [6] and [10], attention is now concentrated on completely nonnormal subnormal operators.

3.2. Theorem. *Let T be a completely nonnormal subnormal operator defined on H with minimal normal extension N defined on K. Then the real dimension of the convex set $\mathcal{S}(T)$ of self-adjoint approximants of T, denoted* $\dim \mathcal{S}(T)$, *is infinite.*

Proof. This proof uses the same techniques as the proof of Theorem 2.4 with a few modifications which will be indicated. Choose $\{z_1, z_2\}$ as in the earlier proof and let D be a compact set not intersecting Σ and containing $\{z_1, z_2\}$ in its interior. Define τ by the equation

$$2\tau = \inf\{(\|C\|^2-(\operatorname{im} z)^2)^{1/2} : z\in D\cap\sigma(N)\}$$

and note that the proof of Lemma 3.1 implies that τ is positive. Proceed with the construction in the proof of Theorem 2.4.

It will be shown that $B_0-A(\varrho; n)$ is a self-adjoint approximant of N for ϱ in an interval $(0, \varrho_0)$ for all n sufficiently large where henceforth, B_0 denotes $B+$ $+(\|C\|^2-C^2)^{1/2}$. (Recall that Theorem 1 of [6] shows that B_0 dominates every self-adjoint approximant of N and Proposition 2 of [6] shows that B_0 is a self-adjoint approximant.) As in the earlier proof, it suffices to show that $\|(N-B_0+A(\varrho; n))|$ $|E(D)K\ominus K(n)\|$ and $\|(N-B_0+A(\varrho; n))|K(n)\|$ are strictly less than $\|C\|$ for appropriate ϱ and n. The first inequality follows from the choice of τ and Lemma 3.1, and the second inequality in the next paragraph.

Define $R(n)$ to be $(N-B_0)|K(n)$ minus the diagonal operator with entries $-(\|C\|^2-\beta_1^2)^{1/2}+i\beta_1$, $-(\|C\|^2-\beta_2^2)^{1/2}+i\beta_2$ respectively. The remainder of the proof of this theorem proceeds by conspicuous analogy to the proof of Theorem 2.4. The resulting self-adjoint approximants of T are $Q(B_0-A(\varrho; n; j))|H$.

Recall the discussion immediately subsequent to Theorem 2.4. Analogously, unless $\|\operatorname{im} T_1\|$ equals $\|\operatorname{im} T_2\|$ there is much arbitrariness in the self-adjoint approximation of $T=T_1\oplus T_2$. For example, if $\|\operatorname{im} T_1\|$ exceeds $\|\operatorname{im} T_2\|$ and R_j is a self-adjoint approximant for T_j, with $j=1, 2$, then $R_1\oplus(R_2-A)$ is a self-adjoint approximant for T provided A is any self-adjoint operator whose norm is dominated by $\|\operatorname{im} T_1\|-\|\operatorname{im} T_2\|$.

The discussion prior to Corollary 2.5 and the discussion prior to Corollary 2.6 indicate the methods used to prove the next two results.

3.3. Corollary. *If the spectrum of the normal operator $N=B+iC$ has an accumulation point not contained in the set Σ consisting of all z with distance to $(-\infty, \infty)$ equal to $\|C\|$, then the dimension of $\mathcal{S}(N)$ is infinite.*

3.4. Corollary. *If T is a completely nonnormal subnormal operator, then $\mathcal{S}(T)$ has an infinite number of extreme points.*

§ 4. Positive approximants of Toeplitz operators

Notation. For $\delta=\delta(T)$, let $\Gamma=\{z$ in $\mathbf{C}$: $\operatorname{dist}(z,[0,\infty))=\delta\}$. Let μ denote Lebesgue measure on the unit circle Δ, normalized so that $\mu(\Delta)=1$. For $p=2$ or $p=\infty$, we denote by $L^p(\Delta)$ the usual Lebesgue spaces. If φ is in $L^\infty(\Delta)$, then the definitions of the Laurent operator L_φ and Toeplitz operator T_φ are as in [9]. By [9, Problem 196] $\sigma(L_\varphi)\subset\sigma(T_\varphi)$.

From Theorem 1.3 it follows that $\delta(T_\varphi)=\delta(L_\varphi)=h(\text{ess range } \varphi,[0,\infty))$ for all φ in $L^\infty(\Delta)$. We examine next the dimension of the convex set $\mathscr{P}(L_\varphi)$ $[\mathscr{P}(T_\varphi)]$ of positive approximants of Laurent [Toeplitz] operators.

4.1. Theorem. *Let φ be in $L^\infty(\Delta)$.*

(i) *If $\mu(\varphi^{-1}(\Gamma))<1$, then both $\mathscr{P}(L_\varphi)$ and $\mathscr{P}(T_\varphi)$ are infinite-dimensional.*

(ii) *If $\mu(\varphi^{-1}(\Gamma))=1$, then L_φ has a unique positive approximant; T_φ has a unique positive approximant if and only if* im φ *is constant.*

Proof. (i) Notice $\delta(T)>0$ in this case. Thus Γ is nondegenerate and the spectra of L_φ and T_φ lie inside Γ. Define the sets $F_k=\{\zeta\colon \operatorname{dist}(\zeta,[0,\infty))\leqq\delta(1-1/k)\}$. Because $\mu(\varphi^{-1}(\Gamma))<1$, there exists $k\geqq 1$ such that $\mu(\varphi^{-1}(F_k))>0$. Fix such a k and write $\varphi^{-1}(F_k)=\bigcup_{j=1}^{\infty} S_j$, where $\{S_j\}$ is a pairwise disjoint collection of measurable sets, each having non-zero measure. Define non-negative functions $p(j)$ in $L^\infty(\Delta)$ by

$$p(j)(z)=\begin{cases}(\operatorname{re}\varphi(z))_+ & z \text{ in } S_j\\ \operatorname{re}\varphi(z)+(\delta^2-(\operatorname{im}\varphi(z))^2)^{1/2} & \text{otherwise}\end{cases}$$

for $j=1,2,\ldots$. It is straightforward to verify that $\|L_{p(j)}-L_\varphi\|=\delta$. Hence each $L_{p(j)}$ is a positive approximant of L_φ, and $T_{p(j)}$ is a positive approximant of T_φ.

We next show that both $\mathscr{P}(L_\varphi)$ and $\mathscr{P}(T_\varphi)$ are infinite-dimensional by proving that

$$\{L_{p(j)}-L_{\operatorname{re}\varphi+(\delta^2-(\operatorname{im}\varphi)^2)^{1/2}}\colon j=1,2,\ldots\}$$

is linearly independent; this also shows that

$$\{T_{p(j)}-T_{\operatorname{re}\varphi+(\delta^2-(\operatorname{im}\varphi)^2)^{1/2}}\colon j=1,2,\ldots\}$$

is linearly independent since [9, Problem 196] Toeplitz operators and Laurent operators defined by the same function in $L^\infty(\Delta)$ have the same norm.

If $c_1,\ldots,c_n$ are real numbers such that $\sum_{j=1}^{n} c_j(L_{p(j)}-L_{\operatorname{re}\varphi+(\delta^2-(\operatorname{im}\varphi)^2)^{1/2}})=0$, then choose r with $1\leqq r\leqq n$ and apply this linear combination to the characteristic function of S_r, which is in $L^2(\Delta)$. This clearly yields a function that is zero off S_r, and for z in S_r it is

$$c_r(-(\operatorname{re}\varphi(z))_- -(\delta^2-(\operatorname{im}\varphi(z))^2)^{1/2})=0.$$

For $\varphi(z)$ in F_k, however, $-(\operatorname{re}\varphi(z))_- -(\delta^2-(\operatorname{im}\varphi(z))^2)^{1/2}$ is bounded below in absolute value by a strictly positive constant that depends only on k (which is fixed). Because $\mu(S_r)>0$, this proves $c_r=0$.

(ii) If $\mu(\varphi^{-1}(\Gamma))=1$, then the essential range of φ is included in Γ. Hence [5, Theorem 5.6] L_φ has a unique positive approximant.

If $\operatorname{im}\varphi$ is constant, then T_φ is normal with spectrum included in Γ, and so it also has [5, Theorem 5.6] a unique positive approximant.

If $\operatorname{im}\varphi$ is not a constant, then the Halmos positive approximant $T_{\operatorname{re}\varphi}+(\delta^2-(T_{\operatorname{im}\varphi})^2)^{1/2}$ and $T_{\operatorname{re}\varphi+(\delta^2-(\operatorname{im}\varphi)^2)^{1/2}}$ are two distinct positive approximants of T_φ. Proof: that both are positive approximants is straightforward to verify. If they were equal, then it would follow that $\delta^2=(T_{\operatorname{im}\varphi})^2+(T_{(\delta^2-(\operatorname{im}\varphi)^2)^{1/2}})^2$. To show this is impossible, let $e_k(z)=z^k$, $k=0, 1, 2, \ldots$ be the usual orthonormal basis of $H^2(\Delta)$; with respect to this basis Toeplitz operators have matrices that are constant along each diagonal [9, Problem 194]. Hence there exists $k>0$ such that $\langle T_{\operatorname{im}\varphi}e_k, e_0\rangle\neq 0$ because $T_{\operatorname{im}\varphi}$ is self-adjoint and not a scalar. Notice that for a self-adjoint Toeplitz operator the fact that the entries in its corresponding matrix are constant along diagonals implies that the sum of the squares of each entry in a given column is exactly one term plus the same sum for the adjacent column on the left. Thus,

$$\begin{aligned}\delta^2 &= \langle T^2_{\operatorname{im}\varphi}e_k, e_k\rangle+\langle T^2_{(\delta^2-(\operatorname{im}\varphi)^2)^{1/2}}e_k, e_k\rangle = \|T_{\operatorname{im}\varphi}e_k\|^2+\|T_{(\delta^2-(\operatorname{im}\varphi)^2)^{1/2}}e_k\|^2\\ &\geqq \|T_{\operatorname{im}\varphi}e_0\|^2+|\langle T_{\operatorname{im}\varphi}e_k, e_0\rangle|^2+\|T_{(\delta^2-(\operatorname{im}\varphi)^2)^{1/2}}e_0\|^2 >\\ &> \langle T^2_{\operatorname{im}\varphi}e_0, e_0\rangle+\langle T^2_{(\delta^2-(\operatorname{im}\varphi)^2)^{1/2}}e_0, e_0\rangle = \delta,\end{aligned}$$

a contradiction.

This proves Theorem 4.1.

The previous theorem shows that the Halmos approximant of T_φ is distinct from the compression to $H^2(\Delta)$ of the Halmos approximant of L_φ (if $\operatorname{im}\varphi$ is not a scalar). The former, of course, always dominates the latter [5, Theorem 4.2]. The next result gives one more comparison of these two operators.

4.2. Theorem. *If* $\operatorname{im}\varphi$ *is continuous, then* $T_{\operatorname{re}\varphi}+(\delta^2-(T_{\operatorname{im}\varphi})^2)^{1/2}$ *is a compact perturbation of* $T_{\operatorname{re}\varphi+(\delta^2-(\operatorname{im}\varphi)^2)^{1/2}}$.

Proof. Let π denote the canonical homomorphism [7, p. 127] onto the Calkin algebra. By [7, Proposition 7.22], if ξ is a continuous function on Δ, then $(T_\xi)^2-T_{\xi^2}$ is compact, *i.e.* $\pi((T_\xi)^2)=\pi(T_{\xi^2})$. If ξ is also non-negative, then $\pi(T_\xi)=\pi(T_{\xi^{1/2}})^2$, and hence $\pi(T_{\xi^{1/2}})=(\pi(T_\xi))^{1/2}$. Thus, since $\pi(p^{1/2})=\pi(p)^{1/2}$ for all $p\geqq 0$, it follows that

$$\begin{aligned}\pi((\delta^2-(T_{\operatorname{im}\varphi})^2)^{1/2}) &= (\pi(\delta^2-(T_{\operatorname{im}\varphi})^2))^{1/2} = (\pi(\delta^2)-\pi(T_{(\operatorname{im}\varphi)^2}))^{1/2} =\\ &= (\pi(T_{\delta^2-(\operatorname{im}\varphi)^2}))^{1/2} = \pi(T_{(\delta^2-(\operatorname{im}\varphi))^{1/2}}).\end{aligned}$$

Q.E.D.

§ 5. Self-adjoint approximants of Toeplitz operators

The results of the previous section on positive approximation have analogues for self-adjoint approximation. Of course, both the distance from L_φ to the self-adjoint operators on $L^2(\Delta)$ and the distance from T_φ to the self-adjoint operators on $H^2(\Delta)$ are $\|\mathrm{im}\,\varphi\|_\infty$ [8] for any φ in $L^\infty(\Delta)$. We now examine the dimension of the convex set $\mathscr{S}(L_\varphi)$ $[\mathscr{S}(T_\varphi)]$ of self-adjoint approximants of Laurent [Toeplitz] operators. We use from § 3 the definition $\Sigma=\{z\colon (\mathrm{im}\,z)^2=\|\mathrm{im}\,\varphi\|_\infty^2,\ |z|\leqq\|\varphi\|_\infty\}$.

5.1. Theorem. *Let φ be in $L^\infty(\Delta)$.*

(i) *If $\mu(\varphi^{-1}(\Sigma))<1$, then both $\mathscr{S}(L_\varphi)$ and $\mathscr{S}(T_\varphi)$ are infinite-dimensional.*

(ii) *If $\mu(\varphi^{-1}(\Sigma))=1$, then L_φ has a unique self-adjoint approximant; the Toeplitz operator T_φ has a unique self-adjoint approximant if and only if* $\mathrm{im}\,\varphi$ *is constant.*

Proof. (i) Notice that in this case $\mathrm{im}\,\varphi$ is not identically zero. For $k=1, 2, 3, \dots$ define $F_k=\{\zeta\colon |\mathrm{im}\,\zeta|\leqq(1-1/k)^{1/2}\|\mathrm{im}\,\varphi\|_\infty\}$. Because $\mu\varphi^{-1}(F_k)<1$, there exists $k\geqq 1$ such that $\mu(\varphi^{-1}(F_k))>0$. Fix such a k and write $\varphi^{-1}(F_k)=\bigcup_{j=1}^{\infty} S_j$ where $\{S_j\}$ is a pairwise disjoint collection of measurable sets of non-zero measure. Define the real-valued functions $s(j)$ in $L^\infty(\Delta)$ by

$$s(j)(z)=\begin{cases}\mathrm{re}\,\varphi(z) & z \text{ in } S_j\\ \mathrm{re}\,\varphi(z)+(\|\mathrm{im}\,\varphi\|^2-(\mathrm{im}\,\varphi(z))^2)^{1/2} & \text{otherwise.}\end{cases}$$

It is again straightforward to verify that $\|L_{s(j)}-L_\varphi\|=\|\mathrm{im}\,\varphi\|_\infty$. Hence each $L_{s(j)}$ is a self-adjoint approximant of L_φ and each $T_{s(j)}$ is a self-adjoint approximant of T_φ

We prove that both $\mathscr{S}(L_\varphi)$ and $\mathscr{S}(T_\varphi)$ are infinite-dimensional by proving that

$$\{L_{s(j)}-L_{\mathrm{re}\,\varphi+(\|\mathrm{im}\,\varphi\|^2-(\mathrm{im}\,\varphi)^2)^{1/2}}\colon\quad j=1,2,\dots\}$$

is linearly independent, which also proves that the corresponding Toeplitz operators are linearly independent. If $\sum_{j=1}^{n} c_j(L_{s(j)}-L_{\mathrm{re}\,\varphi+(\|\mathrm{im}\,\varphi\|_\infty^2-(\mathrm{im}\,\varphi)^2)^{1/2}})=0$, then choose r with $1\leqq r\leqq n$ and apply this linear combination to the characteristic function of S_r, which is in $L^2(\Delta)$. This yields a function that is zero off S_r and for z in S_r it is $-c_r(\|\mathrm{im}\,\varphi\|_\infty^2-(\mathrm{im}\,\varphi(z))^2)^{1/2}$. For $\varphi(z)$ in F_k, however, $(\|\mathrm{im}\,\varphi\|_\infty^2-(\mathrm{im}\,\varphi)z))^2)$ is bounded below by $\frac{1}{k}\,\|\mathrm{im}\,\varphi\|_\infty^2$, which is independent of r. Because $\mu(S_r)>0$, this proves $c_r=0$.

(ii) If $\mu(\varphi^{-1}(\Sigma))=1$, then the essential range of φ is included in Σ. Hence [6] L_φ has a unique self-adjoint approximant.

If $\mathrm{im}\,\varphi$ is constant, then T_φ is normal with spectrum included in Σ, so it also has a unique self-adjoint approximant. If $\mathrm{im}\,\varphi$ is not a constant, then $T_{\mathrm{re}\,\varphi}+$

$+(\|\mathrm{im}\,\varphi\|_\infty^2-(T_{\mathrm{im}\,\varphi})^2)^{1/2}$ and $T_{\mathrm{re}\,\varphi+(\|\mathrm{im}\,\varphi\|_\infty^2-(\mathrm{im}\,\varphi)^2)^{1/2}}$ are two distinct self-adjoint approximants of T_φ. The proofs of these last two assertions are entirely analogous to those given in Theorem 4.1 and are hence omitted. This completes the proof of Theorem 5.1.

It follows from [6, Theorem 1] that

$$T_{\mathrm{re}\,\varphi}+(\|\mathrm{im}\,\varphi\|_\infty^2-(T_{\mathrm{im}\,\varphi})^2)^{1/2} \geqq T_{\mathrm{re}\,\varphi+(\|\mathrm{im}\,\varphi\|_\infty^2-(\mathrm{im}\,\varphi)^2)^{1/2}}.$$

The following comparison of these two operators can be proved as in Theorem 4.2.

5.2. Theorem. *If* $\mathrm{im}\,\varphi$ *is continuous, then* $T_{\mathrm{re}\,\varphi}+(\|\mathrm{im}\,\varphi\|_\infty^2-(T_{\mathrm{im}\,\varphi})^2)^{1/2}$ *is a compact perturbation of* $T_{\mathrm{re}\,\varphi+(\|\mathrm{im}\,\varphi\|_\infty^2-(\mathrm{im}\,\varphi)^2)^{1/2}}$.

Remark. It is not known whether $\mathcal{S}(T_\varphi)$ and $\mathcal{P}(T_\varphi)$ must be either zero-dimensional or infinite-dimensional for each φ in $L^\infty(\Delta)$.

References

[1] M. B. Abrahamse and R. G. Douglas, A class of subnormal operators related to multiply-connected domains, *Advances in Math.*, **19** (1976), 106—148.

[2] R. H. Bouldin, Approximating the shift with Teoplitz operators, *Math. Z.*, **143** (1975), 203—208.

[3] R. H. Bouldin, The convex structure of the set of positive approximants for a given operator, *Acta Sci Math.*, **37** (1975), 177—190.

[4] R. H. Bouldin, Operator approximation in two different norms, *to appear*.

[5] R. H. Bouldin, Positive approximants, *Trans. Amer. Math. Soc.*, **177** (1973), 391—403.

[6] C. K. Chui, P. W. Smith and J. D. Ward, Aprroximation with restricted spectra, *Math. Z.*, **144** (1975), 289—297.

[7] R. G. Douglas, *Banach Algebra Techniques in Operator Theory*, Academic Press (New York, 1972).

[8] P. R. Halmos, Positive approximants of operators, *Indiana U. Math. J.*, **21** (1971/72), 951—960.

[9] P. R. Halmos, *A Hilbert Space Problem Book*, Van Nostrand (Princeton, 1967).

[10] P. R. Halmos, Spectral approximants of normal operators, *Proc. Edinburgh Math. Soc.*, **19** (1974), 51—58.

[11] C. R. Putnam, An inequality for the area of hyponormal spectra, *Math. Z.*, **116** (1970), 323—330.

[12] D. D. Rogers, *Thesis*, Indiana University, 1975.

[13] T. Sekiguchi, Positive approximants of normal operator, *preprint*.

[14] J. G. Stampfli, Hyponormal operators, *Pacific J. Math.*, **12** (1972), 1453—1458.

[15] J. G. Stampfli, Minimal range theorems for operators with thin spectra, *Pacific J. Math.*, **23** (1967), 601—612.

[16] F. A. Valentine, *Convex Sets*, McGraw—Hill (New York, 1964).

Acta Sci. Math., 39 (1977), 245—245

Distributive lattices whose prime ideals are principal

V. R. CHANDRAN and HARRY LAKSER

A well known theorem of I. S. COHEN [1] states that if R is a commutative ring with 1, and every prime ideal in R is principal, then every ideal in R is principal. In this note, the analogue of this theorem is proved for distributive lattices with 1.

Theorem. *Let L be a distributive lattice with* 1. *If every prime ideal in L is principal, then every ideal is principal.*

Proof. Suppose the theorem is false. Let $\mathscr{C}$ denote the non-empty collection of non-principal ideals of L. It is clear that $\mathscr{C}$ is closed under the formation of unions of chains in $\mathscr{C}$. So, by Zorn's lemma we get a maximal element M in $\mathscr{C}$ which is not principal. Since L is principal, $L \neq M$. So, there exist elements $a, b \notin M$ such that $a \wedge b \in M$. Now as M is a maximal element in $\mathscr{C}$, $M \vee (a]$ and $M \vee (b]$ are principal ideals, and, by distributivity $M = (M \vee (a]) \wedge (M \vee (b])$ contradicting that M is not principal. Hence the result.

Obviously, the condition of distributivity cannot be dropped from the Theorem as stated. However, G. Grätzer pointed out that the proof of our theorem carries over to meet-irreducible ideals of general lattices. In a distributive lattice, prime ideals are exactly the meet-irreducible ideals. So we have the following result.

Theorem. *Let L be a lattice with* 1. *If in L every meet-irreducible ideal is principal, then every ideal in L is principal.*

Reference

[1] I. S. COHEN, Rings with restricted minimum condition, *Duke Math. J.*, **17** (1950), 27—42.

DEPARTMENT OF MATHEMATICS AND ASTRONOMY
THE UNIVERSITY OF MANITOBA
WINNIPEG, MANITOBA
CANADA R3T 2N2

Received August 30, 1976.
Research of both authors supported by National Research Council of Canada.

Acta Sci. Math., 39 (1977), 247—250

Eigenvectors of unitary ϱ-dilations

E. DURSZT

Let T be a linear bounded operator on a Hilbert space H and ϱ a positive number. We say that U is a unitary ϱ-dilation of T, if U is a unitary operator on a Hilbert space $K \supset H$ and

$$T^n h = \varrho P U^n h \quad \text{for all} \quad h \in H \quad \text{and for} \quad n = 1, 2, \ldots,$$

where P (as always in the following) denotes the orthogonal projection of K onto H. Clearly U is a unitary ϱ-dilation of T if and only if U^{-1} is a unitary ϱ-dilation of T^*. U is called to be minimal, if

$$K = \bigvee_{n=-\infty}^{\infty} U^n H.$$

$\mathscr{C}_\varrho$ denotes the class of those operators which have unitary ϱ-dilations.

Unitary ϱ-dilations were introduced and operators of classes $\mathscr{C}_\varrho$ were characterized by Sz.-Nagy and Foiaş [6]. Spectral properties of unitary ϱ-dilations were studied in [1], [5], [3]. In this Note we prove two theorems about eigenvalues and eigenvectors, generalizations to the unitary ϱ-dilation case of facts known for the unitary 1-dilations ([7], Ch. 2, Proposition 6.1).

In what follows we fix a positive number ϱ, a minimal unitary ϱ-dilation U of T, and introduce the following notations:

$$(1) \qquad K_+ = \bigvee_{n=0}^{\infty} U^n H, \quad U_+ = U|K_+; \quad K_- = \bigvee_{n=0}^{\infty} U^{-n} H, \quad U_- = U^{-1}|K_-;$$

$$L_+ = K_+ \ominus UK_+, \quad L_- = K_- \ominus U^{-1}K_-;$$

$$(2) \qquad R_+ = \bigcap_{n=0}^{\infty} U^n K_+, \quad R_- = \bigcap_{n=0}^{\infty} U^{-n} K_-;$$

$$(3) \qquad R_0 = R_+ \cap R_-.$$

Received May 2, 1977.

Clearly, L_+ and L_- are wandering subspaces for the isometric operators U_+ and U_-, respectively. Moreover,

$$K_+ = \left(\bigoplus_{n=0}^{\infty} U^n L_+\right) \oplus R_+, \quad K_- = \left(\bigoplus_{n=0}^{\infty} U^{-n} L_-\right) \oplus R_-$$

are the corresponding Wold decompositions. By the minimality of U this implies that

$$K = \left(\bigoplus_{n=-\infty}^{\infty} U^n L_+\right) \oplus R_+, \quad K = \left(\bigoplus_{n=-\infty}^{\infty} U^{-n} L_-\right) \oplus R_-. \tag{4}$$

Finally we denote by P_+, P_-, and P_0 the orthogonal projections of K onto R_+, R_-, and R_0, respectively.

Theorem A. (Mlak [4]) $U^n T^{*n} h \rightharpoonup P_+ h, \quad U^{-n} T^n h \rightharpoonup P_- h \quad (\forall h \in H)$ *(weak convergence).*

Theorem 1. *If*

$$Ug = \varepsilon g \tag{5}$$

for some $g \in K$ and complex number ε, then

$$TPg = \varepsilon Pg, \quad T^* Pg = \bar{\varepsilon} Pg, \quad \text{and} \quad g \in R_0.$$

Proof. Let Q_n denote the orthogonal projection of g onto $U^n L_+$ ($n=0$, $\pm 1, \pm 2, \ldots$), then by (4)

$$\sum_{n=-\infty}^{\infty} \|Q_n g\|^2 \leqq \|g\|^2, \quad UQ_n g = Q_{n+1} Ug, \tag{6}$$

and so by (5)

$$\|Q_n g\| = \|Q_{n+1} Ug\| = \|Q_{n+1} \varepsilon g\| = \|Q_{n+1} g\| \quad (n = 0, \pm 1, \pm 2, \ldots).$$

By (6) this implies $Q_n g = 0$ for all n, thus by (4)

$$g \in R_+. \tag{7}$$

For $n = 1, 2, \ldots$ and $h \in H$ we have

$$TP(U^n h) = \frac{1}{\varrho} T^{n+1} h = PU(U^n h).$$

Since by (1) and (2)

$$R_+ = \bigcap_{n=0}^{\infty} \bigvee_{m=n}^{\infty} U^m H \subset \bigvee_{m=1}^{\infty} U^m H,$$

we conclude that

$$TPf = PUf \quad \text{for all} \quad f \in R_+. \tag{8}$$

Thus by (5) and (7)

$$TPg = PUg = \varepsilon Pg,$$

and this is the first statement of our theorem.

Repeating the above argument with T^*, U^{-1} and $\bar{\varepsilon}$ in place of T, U and ε, respectively, we get $g \in R_-$, $T^*Pg = \bar{\varepsilon}Pg$, and by (3) and (7), $g \in R_0$. This concludes the proof.

Theorem 2. *If*

$$Th = \varepsilon h \tag{9}$$

for an $h \in H$ *and a complex number* ε, $|\varepsilon| = 1$, *then*

$$UP_0 h = \varepsilon P_0 h, \quad h = \varrho PP_0 h, \quad T^* h = \bar{\varepsilon} h.$$

Proof. Theorem A implies for each $g \in K$

$$(P_- h, g) = \lim_n (U^{-n} T^n h, g) = \lim_n (U^{-n+1} T^{n-1} \varepsilon h, Ug) = (P_- \varepsilon h, Ug).$$

Consequently,

$$UP_- h = \varepsilon P_- h. \tag{10}$$

By Theorem 1 this implies $P_- h \in R_0$, and thus by (3) and the definitions of P_- and P_0,

$$P_0 h = P_- h. \tag{11}$$

This fact and (10) prove the first statement of our theorem. Using again Theorem 1, (10) implies

$$T^* PP_- h = \bar{\varepsilon} PP_- h. \tag{12}$$

Now by (11), Theorem A and (9) we have

$$\varrho(h, P_0 h) = \varrho(h, P_- h) = \varrho \lim_n (h, U^{-n} T^n h) = \varrho \lim_n (U^n h, \varepsilon^n h) =$$

$$= \lim_n (T^n h, \varepsilon^n h) = \lim_n (\varepsilon^n h, \varepsilon^n h) = \|h\|^2,$$

i.e.

$$\varrho(h, P_0 h) = \|h\|^2. \tag{13}$$

Again by (11), Theorem A, and (12),

$$\varrho \|PP_0 h\|^2 = \varrho(P_- h, PP_- h) = \varrho \lim_n (U^{-n} T^n h, PP_- h) = \varrho \lim_n (\varepsilon^n h, U^n PP_- h) =$$

$$= \lim_n (\varepsilon^n h, T^n PP_- h) = \lim_n (\varepsilon^n h, \varepsilon^n PP_- h) = (h, P_- h) = (h, P_0 h).$$

This fact and (13) imply that

$$\varrho \|PP_0 h) = \|h\|. \tag{14}$$

Now by (13) and (14)

$$\|h-\varrho PP_0 h\|^2 = \|h\|^2 - 2\varrho \operatorname{Re}(h, PP_0 h) + \varrho^2 \|PP_0 h\|^2 = 0,$$

and consequently $h=\varrho PP_0 h$. This fact, (11), and (12) imply $T^*h=\bar{\varepsilon}h$, and the proof is complete.

In connection with Theorem 2 let us recall that G. Eckstein [2] has proved the following statement. If $T\in\mathscr{C}_\varrho$ for some positive ϱ and if the complex number ε of modulus 1 is an approximate eigenvalue of T, then $\bar{\varepsilon}$ is an approximate eigenvalue of T^* with the same approximate proper vectors.

References

[1] E. Durszt, On the spectrum of unitary ϱ-dilations, *Acta Sci. Math.*, **28** (1967), 299—304.
[2] G. Eckstein, Sur les opérateurs de classe $\mathscr{C}_\varrho$, *Acta Sci. Math.*, **33** (1972), 349—352.
[3] W. Mlak, Spectal properties of Q-dilations, *Bull. Acad. Polon. Sci.* (sér. math., astr., phys.), **17** (1969), 397—400.
[4] W. Mlak, Convergence properties of operators of class $\mathscr{C}_\varrho$, *Acta Sci. Math.*, **33** (1972), 353—354.
[5] A. Rácz, Sur les transformations de classe $\mathscr{C}_\varrho$ dans l'espace de Hilbert, *Acta Sci. Math.*, **28** (1967), 305—310.
[6] B. Sz.-Nagy—C. Foiaş, On certain classes of power-bounded operators in Hilbert space, *Acta Sci. Math.*, **27** (1966), 17—25.
[7] B. Sz.-Nagy—C. Foiaş, *Harmonic Analysis of Operators on Hilbert Space*, North Holland and Akadémiai Kiadó (Amsterdam—London and Budapest, 1970).

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1.
6720 SZEGED, HUNGARY

Acta Sci. Math., 39 (1977), 251—254

On the spectrum of contractions of class $C_{\cdot 1}$

G. ECKSTEIN

1. In this paper we shall consider (bounded) operators in complex separable Hilbert spaces. We shall use notations from [8], and **Z** will denote the integers, **N** the natural numbers, **C** the field of complex numbers. We denote the open unit disc by D, the unit circle by C, and the annulus $\{\lambda \in \mathbf{C}: 1/2 \leqq |\lambda| \leqq 1\}$ by K. For a contraction $T \in \mathscr{L}(\mathfrak{H})$ we denote by $\sigma(T)$ its spectrum, by $\sigma_p(T)$ its point spectrum, $D_T = (I - T^*T)^{1/2}$ denotes the defect operator, $\mathfrak{D}_T = \overline{D_T \mathfrak{H}}$ the defect space, and $\mathfrak{d}_T = \dim \mathfrak{D}_T$ the defect number of T.

B. SZ.-NAGY and C. FOIAŞ call the contraction T of *class* $C_{\cdot 1}$ if $T^{*n}x \not\to 0$ for all $x \in \mathfrak{H}$, $x \neq 0$ (see [8], Ch. II. Section 4). In [8] Ch. VII, 6.3, or [8], Th. 2* they prove that if $T \in C_{\cdot 1}$ and δ_{T^*} is finite then $\sigma(T) = \bar{D}$ or $\sigma(T) \subseteqq C$. Moreover, in the first case, $\sigma_p(T) \supseteqq D$ and $T \notin C_{11}$, while in the second, $T \in C_{11}$. In the case $\mathfrak{d}_{T^*} = \infty$ it is posible that $T \in C_{11}$ and $\sigma(T) = \bar{D}$ (see [8], Ch. VI, Section 4).

This raises the following questions:

a) If $T \in C_{01}$, does it follow that $\sigma(T) \cap D \neq \emptyset$?

b) If $T \in C_{01}$ and $\sigma(T) \cap D \neq \emptyset$, then does it follow that $\sigma_p(T) \neq \emptyset$?

c) If $T \in C_{\cdot 1}$, does it follow that $\sigma(T) = \bar{D}$ or $\sigma(T) \subseteqq C$?

d) If $T \in C_{\cdot 1}$ and $\sigma_p(T) \cap D \neq \emptyset$, does it follow that $\sigma_p(T) \supset D$?

e) If $T \in C_{\cdot 1}$ and $1 \notin \sigma(T)$, does it follow that $\sigma(T) \subset C$?

GILFEATHER [2] gave a negative answer to a) and b). Using weighted shifts he proved that

a) there is an operator $T \in C_{01}$ with $\sigma(T) = C$, and

b) there is an operator $T \in C_{01}$ with $\sigma(T) = \bar{D}$ and $\sigma_p(T) = \emptyset$.

The aim of this note is to give a negative answer to c) and d).

2. Theorem 1. *There exists $T \in C_{01}$ with $\sigma(T) = K$.*

Received June 23, 1973.

Proof. Let $\mathfrak{H}$ be a space with orthonormal basis $\{\varphi_n\}_{n\in\mathbf{Z}}$ and let T be the weighted shift in $\mathfrak{H}$ defined by

$$T\varphi_n = w_n\varphi_{n+1} \quad (n\in\mathbf{Z}) \quad \text{where} \quad w_n = 1 \quad \text{for} \quad n \leqq 0 \quad \text{and } 1/2 \text{ for} \quad n > 0.$$

T is a contraction ($\|T\| = \sup|w_n| = 1$, see [4]). It is of class C_{01} since $\prod_{n>0} w_n$ diverges and $\prod_{n<0} w_n$ converges (see [2]). The spectrum of T is K, since

$$1/2 = \lim_{n\to\infty} \inf_{k\in\mathbf{Z}} \{(w_k w_{k+1}\dots w_{k+n-1})^{1/n}\}, \quad 1 = \lim_{n\to\infty} \sup_{k\in\mathbf{Z}} \{(w_k w_{k+1}\dots w_{k+n-1})^{1/n}\}$$

(see [3], [6] and [4]).

We shall see now that the alternative of problem c) does not occur even if $T\in C_{11}$.

Theorem 2. *There exists an operator $T\in C_{11}$ with $\sigma(T)=K$.*

Proof. Let $\mathfrak{H}$ be a space with orthonormal basis $\{\varphi_{ij}\}_{(i,j)\in\mathbf{N}\times\mathbf{Z}}$ and let $T\in L(\mathfrak{H})$ be defined by

$$T\varphi_{ij} = w_{ij}\varphi_{i,j+1} \quad (i\in\mathbf{N}, j\in\mathbf{Z}),$$

where

$$w_{ij} = 1 \quad \text{for} \quad j\notin[0,i] \quad \text{and } 1/2 \text{ for} \quad j\in[0,i].$$

One can verify that $T\in C_{11}$ and $0\notin\sigma(T)$. Taking $h_n = \sum_{k=0}^{n-1} n^{-1/2}\varphi_{nk}\in\mathfrak{H}$ we have $\|h_n\|=1$ and $\|\mathrm{Th}_n-(1/2)h_n\| = \|(1/2)n^{-1/2}\varphi_{nn}-(1/2)n^{-1/2}\varphi_{n0}\| = (2n)^{-1/2}\to 0$; hence $1/2\in\sigma(T)$. We have $\sigma(T)\neq\bar{D}$ and $\sigma(T)\not\subset C$. Consider the unitary operators $S_t\in\mathscr{L}(\mathfrak{H})$ defined by

$$S_t\varphi_{mn} = \exp(-\mathrm{int})\varphi_{mn}.$$

We have $S_t^{-1}TS_t = \exp(\mathrm{it})\,T$ from which we deduce the circular symmetry of $\sigma(T)$. Moreover, by condition $T\in C_{11}$, the spectrum of T has no components far from C (since then there would exist a non-trivial subspace $\mathfrak{H}_0$ of $\mathfrak{H}$, with $TH_0\subset\mathfrak{H}_0$ and $\sigma(T|\mathfrak{H}_0)\subset D$, so that $T^n h_0\to 0$ for $h_0\in\mathfrak{H}_0$). Since $\|T^{-1}\|=2$ and since, by [6], $\sigma(T)$ is an annulus, it follows that $\sigma(T)=K$.

3. In this section we shall give a class of contractions for which the alternative of c) is true. We shall use the functional model introduced by Sz.-Nagy and Foiaş (see [8], Ch. V and VI). For a contraction $T\in\mathscr{L}(\mathfrak{H})$ we have:

$$\sigma_p(T)\cap D = \{\lambda\in D: \Theta_T(\lambda) \quad \text{is not injective}\},$$

$$\sigma(T)\cap D = \{\lambda\in D: \Theta_T(\lambda) \quad \text{is not invertible}\},$$

where $\Theta_T(\lambda)\colon \mathfrak{D}_T\to\mathfrak{D}_{T^*}$ is the characteristic function

$$\Theta_T(\lambda) = [-T+\lambda D_{T^*}(I-\lambda T^*)^{-1}D_T]|\mathfrak{D}_T \quad (\lambda\in D).$$

Since T maps $\mathfrak{H}\ominus\mathfrak{D}_T$ unitarily on $\mathfrak{H}\ominus\mathfrak{D}_{T^*}$ we can here replace $\Theta_T(\lambda)$ by

$$T(\lambda)=-T+\lambda D_{T^*}(I-\lambda T^*)^{-1}D_T.$$

Suppose that $\sigma_p(T^*)\cap D=\emptyset$ and that $D_{T^*}(I-\lambda T^*)^{-1}D_T$ is compact for each $\lambda\in D$. If $\lambda_0\in D\setminus\sigma(T)$, then $T(\lambda_0)$ is invertible, hence it is Fredholm of index 0 (that is $T(\lambda_0)\mathfrak{H}$ is closed and dim Ker $T(\lambda_0)=$dim Ker $T^*(\lambda_0)<\infty$). For $\lambda\in D$ we have

$$T(\lambda)=T(\lambda_0)+[T(\lambda)-T(\lambda_0)].$$

Since $T(\lambda)-T(\lambda_0)$ is compact, we deduce that $T(\lambda)$ is Fredholm of index 0 (see [1] or [5]). But Ker $T^*(\lambda)=\{0\}$ since $\sigma_p(T^*)\cap D=\emptyset$ hence $T(\lambda)$ is invertible. We have proved the following

Proposition. *If $T\in\mathcal{L}(\mathfrak{H})$ is a contraction with $\sigma_p(T^*)=\emptyset$ and if $\Theta_T(\lambda)-\Theta_T(0)$ is compact for each $\lambda\in D$ then $\sigma(T)=\bar{D}$ or $\sigma(T)\subset C$.*

Remark. The hypothesis of this proposition is fulfilled in particular if $T\in C_{\cdot 1}$ with D_T or D_{T^*} compact.

We shall see that even under the hypothesis of the proposition, problem d) has a negative answer.

Theorem 3. *There exists an operator $T\in C_{\cdot 1}$ with $\Theta_T(\lambda)-\Theta_T(0)$ compact and $\sigma_p(T)=\{0\}$.*

Proof. Let $\mathfrak{E}$ be a Hilbert space with orthonormal basis $\{e_n\}_{n\geq 0}$, $\mathfrak{E}_1$ the subspace of $\mathfrak{E}$ generated by $\{e_n\}_{n\geq 1}$, and let $S\in\mathcal{L}(\mathfrak{E})$ be the operator defined by

$$e_0\mapsto 0,\quad e_n\mapsto (1/n)\,e_{n-1}\quad (n>0).$$

Let $F\in\mathcal{L}(\mathfrak{E})$ be the compact operator defined by

$$Fe_0=f=\sum_{k=1}^{\infty}\frac{1}{k+1}e_k,\quad F\mathfrak{E}_1=\{0\}.$$

We have $\overline{S\mathfrak{E}}=\overline{S\mathfrak{E}_1}=\mathfrak{E}$ and $f\notin S\mathfrak{E}$. Consider the analytic contractive function $\{\mathfrak{E},\mathfrak{E},\Theta(\lambda)\}$ defined by

$$\Theta(\lambda)=(\|S\|+\|F\|)^{-1}(S+\lambda F).$$

As $F|\mathfrak{E}_1=0$, we have

$$\overline{\Theta H^2(\mathfrak{E})}\supset\overline{\Theta H^2(\mathfrak{E}_1)}=\overline{SH^2(\mathfrak{E}_1)}=H^2(\mathfrak{E}),$$

that is, $\Theta(\lambda)$ is an outer function. If $\lambda\in D\setminus\{0\}$ and $\Theta(\lambda)x=0$, then $Sx=-\lambda Fx$, hence $Sx=0$ and $Fx=0$. But from the first equality it follows that $x=\alpha e_0$, and from the second that $\alpha=0$, hence $\Theta(\lambda)$ is injective for each $\lambda\in D\setminus\{0\}$. Constructing the contraction $\mathbf{T}$ (see [8], Ch. VI. 3) we obtain a contraction of class $C_{\cdot 1}$ with $\sigma_p(\mathbf{T})=\{0\}$ and $\Theta_{\mathbf{T}}(\lambda)-\Theta_{\mathbf{T}}(0)$ compact.

References

[1] S. Berberian, The Weyl spectrum of an operator, *Indiana Math. J.*, **20** (1970), 529—544.
[2] F. Gilfeather, Weighted bilateral shifts of class C_{01}, *Acta Sci. Math.*, **32** (1971), 251—254.
[3] R. Gellar, Operators commuting with a weighted shift, *Proc. Amer. Math. Soc.*, **23** (1969), 538—545.
[4] P. Halmos, *A Hilbert Space Problem Book*, Van Nostrand (Princeton, 1967).
[5] R. Palais, *Seminar on Atiyah-Singer theorem*, Princeton University Press (Princeton, 1965).
[6] S. Ridge, Approximate point spectrum of a weighted shift, *Transactions Amer. Math. Soc.*, **147** (1970), 349—356.
[7] B. Sz.-Nagy—C. Foiaş, Sur les contractions de l'espace de Hilbert. IX, *Acta Sci. Math.*, **25** (1964), 283—316, Corrections et compléments, *ibidem*, **26** (1965), 193—196.
[8] B. Sz.-Nagy—C. Foiaş, *Analyse harmonique des opérateurs de l'espace de Hilbert* (Paris—Budapest, 1967).

UNIVERSITY OF TIMIŞOARA
BUL. V. PÂRVAN 4.
TIMIŞOARA, ROMANIA

Acta Sci. Math., 39 (1977), 255—263

On modules over Dedekind rings

S. FEIGELSTOCK

1. A ring in this paper always signifies an integral domain, and will be denoted by R. M will denote a unitary R-module.

In section 2 some properties of abelian groups will be generalized to R-modules. In most cases, R will be taken to be a Dedekind ring. The results of section 2 will be utilized in section 3 to obtain information about A-high submodules of M, A a submodule of M. In section 4, the results of section 2 will be employed in determining the structure of the tensor product of R-modules for several types of modules over a Dedekind ring.

2. Definition 1. M is said to be a *divisible* R-module, if $rM=M$ for all $0\neq r\in R$.

Definition 2. Let P be a prime ideal in R. M is said to be *P-divisible*, if $PM=M$

Definition 3. Let N be a submodule of M. N is said to be a *pure submodule* of M if for all $r\in R$ and for all $m\in M$, if $rm\in N$, then there exists an $n\in N$ such that $rm=rn$.

Definition 4. Let N be a submodule of M. N is said to be an *ideal pure submodule* of M if for every ideal I in R, $N\cap IM=IN$.

Ideal purity clearly implies purity.

Notation. Let $m\in M$, $\operatorname{ord}(m)=\{r\in R|rm=0\}$.

Definition 5. Let P be a prime ideal in R. M is said to be a *P-primary module* if for every $m\in M$ there exists a positive integer $k(m)$ such that $P^{k(m)}\subseteq \operatorname{ord}(m)$.

Definition 6. Let P be a prime ideal in R. A submodule N of M is said to be *P-pure* in M if $N\cap P^kM=P^kN$ for every positive integer k.

Lemma 1. *Let R be a Dedekind ring, P a prime ideal in R, and M a P-primary R-module. Then for every prime ideal Q in R, $Q\neq P$, $QM=M$.*

Received October 16. 1975.

Proof. Q is a maximal ideal in R, hence $Q \nsubseteq P$. Therefore $QM=M$ by [2, Lemma 4].

Lemma 2. *Let R be a Dedekind ring, P a prime ideal in R, M a P-primary R-module, and N a submodule of M. If N is P-pure in M then N is ideal pure in M.*

Proof. Let I be an ideal in R. Then $I=\prod Q^{k(Q)}$, Q running over the set of prime ideals in R, $k(Q)$ being a non-negative integer, $k(Q)=0$ for all but finitely many Q [8, p. 274].

By Lemma 1, $IM=P^{k(P)}M$, and $IN=P^{k(P)}N$. Therefore, $N\cap IM=N\cap P^{k(P)}M= =P^{k(P)}N=IN$.

Definition 7. An exact sequence $0\to L\xrightarrow{\varphi} M\xrightarrow{\psi} N\to 0$ of R-modules is said to be *(ideal) pure exact* if im φ is an (ideal) pure submodule of M.

Lemma 3. *Let*

$$(*) \qquad 0\to L\xrightarrow{\varphi} M\xrightarrow{\psi} N\to 0$$

be an exact sequence of R-modules:

a) *If $(*)$ is pure exact then the sequence*

$$0\to L/rL\xrightarrow{\bar\varphi} M/rM\xrightarrow{\bar\psi} N/rN\to 0$$

is exact for every $r\in R$. $\bar\varphi$ and $\bar\psi$ are defined in the natural way.

b) *If $(*)$ is ideal pure exact, then the sequence*

$$0\to L/IL\xrightarrow{\bar\varphi} M/IM\xrightarrow{\bar\psi} N/IN\to 0$$

is exact for every ideal I in R. $\bar\varphi$ and $\bar\psi$ are defined in the natural way.

Proof. Same as for abelian groups [3, Theorem 29.1].

The following are known facts concerning modules over a Dedekind ring:

Proposition 1. (Steinitz [7]) *Let R be a Dedekind ring and let M be a finitely generated R-module, then M is a direct sum of cyclic modules, and rank one torsion free modules.*

Proposition 2. (Kaplansky [5, Theorem 1].) *Let R be a Dedekind ring, and let M be a finitely generated R-module, then $M\cong M_t\oplus(M/M_t)$, M_t the torsion part of M.*

Proposition 3. [5, p. 332] *Let R be a Dedekind ring, and let M be a torsion module. Then M is a direct sum of P-primary modules.*

Definition 8. Let S be a subset of M. S is said to be an *independent set* in M if for every positive integer k and for all $r_j \in R$, $m_j \in S$ $(1 \leqq j \leqq k)$,

$$\sum_{j=1}^{k} r_j m_j = 0 \quad \text{implies} \quad r_j m_j = 0 \quad (1 \leqq j \leqq k).$$

Definition 9. Let S be a subset of M and let P be a prime ideal in R. S is said to be a *P-independent set* in M if for all positive integers k, l, and for all $r_j \in R$, $m_j \in S$ $(1 \leqq j \leqq k)$,

$$\sum_{j=1}^{k} r_j m_j \in P^l M \quad \text{implies} \quad r_j \in P^l \quad (1 \leqq j \leqq k).$$

Lemma 4. *Let R be a Dedekind ring, P a prime ideal in R, and S a P-independent set in M. Then S is independent.*

Proof. Let $r_j \in R$, $m_j \in S$ $(1 \leqq j \leqq k)$ for k a positive integer. Suppose $\sum_{j=1}^{k} r_j m_j = 0$. Then $\sum_{j=1}^{k} r_j m_j \in P^e M$ for every positive integer e. S is P-independent, hence $r_j \in P^e$ for every positive integer e $(1 \leqq j \leqq k)$. However, R is Noetherian, so that $r_j = 0$ $(1 \leqq j \leqq k)$ and S is therefore independent.

Lemma 5. *Let P be a prime ideal in R, and let S be a P-independent subset of M. $\langle S \rangle$, the submodule of M generated by S, is P-pure in M.*

Proof. Let $x \in \langle S \rangle \cap P^e M$, e a positive integer. Then $x = \sum_{j=1}^{k} r_j m_j$, $r_j \in R$, $m_j \in S$ $(1 \leqq j \leqq k)$ and $x \in P^e M$. S is P-independent, so that $r_j \in P^e$ $(1 \leqq j \leqq k)$. Therefore $x \in P^e \langle S \rangle$.

It has been observed [5, p. 332] that if R is a Dedekind ring, P a prime ideal in R, and M a P-primary R-module, then M may be viewed as an R_p-module (R_p the localization of R at the prime P). This may be done in the following manner.

Let $r/s \in R_p$, $r \in R$, $s \in R - P$, and let $m \in M$. $s \notin P$, hence there exists an $m' \in M$ such that $m = sm'$. Define $(r/s)\, m = rm'$. It is easily verified that this action of R_p on M gives M the structure of an R_p-module.

Lemma 6. *Let R be a Dedekind ring, P a prime ideal in R, $S \subseteqq M$. S is a (maximal) P-independent set in M iff S is a (maximal) PR_p-independent set in M.*

Proof. 1) Suppose S is P-independent. Let $r_j \in R$, $s_j \in R - P$, $m_j \in S$ $(1 \leqq j \leqq k)$, k a positive integer, and suppose that $x = \sum_{j=1}^{k} (r_j/s_j) m_j \in (PR_p)^e M$, e a positive integer. $(PR_p)^e$ is a principal ideal: $(PR_p)^e = \langle r/s \rangle$, $r \in P^e$, $s \in R - P$. Therefore $x = (r/s)m$,

$m \in M$, and

$$\left(s \cdot \prod_{i=1}^{k} s_i\right) x = \sum_{j=1}^{k} \left(s \cdot \prod_{\substack{i=1 \\ i \neq j}}^{k} s_i\right) r_j m_j = \left(r \cdot \prod_{i=1}^{k} s_i\right) m \in P^e M.$$

S is P-independent, hence $\left(s \cdot \prod_{\substack{i=1 \\ i \neq j}}^{k} s_i\right) r_j \in P^e \, (1 \leqq j \leqq k)$. However, $s, s_i \notin P \, (1 \leqq i \leqq k)$ so that $r_j \in P^e$, and hence $r_j/s_j \in (PR_p)^e$ $(1 \leqq j \leqq k)$.

Suppose that S is a maximal P-independent set in M. Let $0 \neq m \in M$. Then there exist $r_j \in R$ $(0 \leqq j \leqq k)$ and $m_j \in S$ $(1 \leqq j \leqq k)$, such that $r_0 m \neq 0$, $r_j m_j \neq 0$ $(1 \leqq j \leqq k)$,

$$r_0 m + \sum_{j=1}^{k} r_j m_j \in P^e M, \quad e \text{ a positive integer, but } r_0 \notin P^e,$$

$$r_0 m + \sum_{j=1}^{k} r_j m_j \in (PR_p)^e M \quad (\text{we are here identifying } (r/1) \in R_p \text{ with } r \in R).$$

Suppose $r_0 \in (PR_p)^e$. Then $r_0 = r/s$, $r \in P^e$, $s \in R - P$. Then $r_0 s \in P^e$. However $s \notin P$ and hence $r_0 \in P^e$: a contradiction.

2) Suppose that S is a PR_p-independent subset of M. Let $r_j \in R$, $m_j \in S$ $(1 \leqq j \leqq k)$, and suppose that $x = \sum_{j=1}^{k} r_j m_j \in P^e M$, e a positive integer. Then $x \in (PR_p)^e M$, and hence $r_j \in (PR_p)^e$ $(1 \leqq j \leqq k)$. As was the case with r_0 above, this implies that $r_j \in P^e$ $(1 \leqq k \leqq j)$.

Suppose that S is a maximal PR_p-independent set in M. Let $0 \neq m \in M$. Then there exist $r_j \in R$, $s_j \in R - P$ $(0 \leqq j \leqq k)$, and $m_j \in S$, $(1 \leqq j \leqq k)$ such that $(r_0/s_0) m \neq 0$, $(r_j/s_j) m_j \neq 0$ $(1 \leqq j \leqq k)$.

$x = (r_0/s_0) m + \sum_{j=1}^{k} (r_j/s_j) m_j \in (PR_p)^e M$, e a positive integer, but $r_0/s_0 \notin (PR_p)^e$. This implies $r_0 \notin P^e$.

$$x = (r/s) m', \quad r \in P^e, \quad s \in R - P, \quad m' \in M.$$

Therefore, $\left(s \cdot \prod_{i=0}^{k} s_i\right) x \in P^e$, but $\left(s \cdot \prod_{i=1}^{k} s_i\right) r_0 \notin P^e$. This implies that S is maximal P-independent in M.

Lemma 7. *Let R be a Dedekind ring, P a prime ideal in R, and M a P-primary R-module. M is P-divisible iff M is PR_p divisible.*

Proof. 1) Suppose that M is P-divisible. Obviously, $P \subseteqq PR_p$. Hence, $M = PM \subseteqq PR_p M \subseteqq M$, and M is PR_p-divisible.

2) Suppose that M is PR_p-divisible. Let $m \in M$. There exist $p \in P$, $s \in R - P$, and $m' \in M$, such that $(p/s) m' = m$. Hence $sm \in PM$, but $s \notin P$. This implies that $m \in PM$.

Theorem 1. *Let R be a Dedekind ring, P a prime ideal in R, and M a P-primary R-module. Let S be a maximal P-independent set in M. Then $M/\langle S\rangle$ is P-divisible.*

Proof. Let $0\neq m\in M$. By Lemma 6, there exist $r_j\in R, s_j\in R-P$ $(0\leqq j\leqq k)$ and $m_j\in S$ $(1\leqq j\leqq k)$ such that

$$x=(r_0/s_0)\,m+\sum_{j=1}^{k}(r_j/s_j)\,m_j\in(PR_p)^e M,$$

e a positive integer, $(r_0/s_0)m\neq 0$, and $(r_j/s_j)m_j\neq 0$ $(1\leqq j\leqq k)$, $(r_0/s_0)\notin(PR_p)^e$, $(r_0/s_0)\in\in(PR_p)^{e'}$ for $0\leqq e'<e$. This implies that $(r_j/s_j)\in(PR_p)^{e'}$ $(1\leqq j\leqq k)$. $(PR_p)^{e'}$ is a principal ideal in R_p; hence $(PR_p)^{e'}=(r/s)$, $r\in P^{e'}$, $s\in R-P$. Therefore $x=(r/s)m'$, $m'\in M$, and $(r_j/s_j)=(r/s)\,(r_j'/s_j')$, $r_j'\in R$, $s_j'\in R-P$ $(0\leqq j\leqq k)$, $r_0\in R-P$.

This yields that $(r/s)t=0$, where

$$t=(r_0'/s_0')\,m-m'+\sum_{j=1}^{k}(r_j'/s_j')\,m_j=0.$$

Put $\bar{y}=y+\langle S\rangle$ for $y\in M$. Clearly, $\bar{y}\in PR_p(M/\langle S\rangle)$ holds for every $y\in M$ for which $(PR_p)^0\subseteqq\operatorname{ord}(y)$.

Suppose that $\bar{y}\in PR_p(M/\langle S\rangle)$ for every $y\in M$ for which $(PR_p)^{k'}\subseteqq\operatorname{ord}(y)$ $(0\leqq k'<e)$. Then $\bar{t}\in PR_p(M/\langle S\rangle)$. This implies that $(r_0'/s_0')m-m'\in PR_p(M/\langle S\rangle)$. However, $m'\in(PR_p)^{e-e'}M\subseteqq(PR_p)M$, so that $(r_0'/s_0')\bar{m}\in(PR_p)\,(M/\langle S\rangle)$. $(r_0'/s_0')\notin PR_p$, so that $\bar{m}\in(PR_p)\,(M/\langle S\rangle)$. $M/\langle S\rangle$ is therefore PR_p-divisible, and hence P-divisible by Lemma 7.

Lemma 8. *Let R be a ring for which every finitely generated ideal is principal. Let I be an ideal in R, and let A and B be R-modules. Then $I(A+B)=IA+IB$.*

Proof. Clearly $I(A+B)\subseteqq IA+IB$. Let $x\in IA+IB$. Then

$$x=\sum_{j=1}^{k}(i_ja_j+i_j'b_j),\quad i_j,i_j'\in I,\quad a_j\in A,\quad b_j\in B\quad(1\leqq j\leqq k).$$

The ideal $\langle i_j, i_j'\,|1\leqq j\leqq k\rangle=\langle i\rangle$, $i\in I$. Therefore $i_j=r_ji$, $i_j'=r_j'i$; $r_j,r_j'\in R$ $(1\leqq j\leqq k)$. Hence $x=i\left(\sum_{j=1}^{k}r_ja_j+\sum_{j=1}^{k}r_j'b_j\right)\in I(A+B)$.

3. Definition 10. Let A be a submodule of M. A submodule B of M is said to be *A-high* if $A\cap B=0$, and if for every submodule C of M, $B\underset{\text{properly}}{\subset}C$, implies that $A\cap C\neq 0$.

Lemma 9. *Let A be a submodule of M, B an A-high submodule of M, and $N=A\oplus B$. Then M/N is a torsion module.*

Proof. Let $m\in M$, $m\notin N$. Then there exists a non-zero $a\in A\cap\langle B,m\rangle$. Let $a=b+rm$, $b\in B$, $r\in R$. If $r=0$, then $a\in B$, contradicting the fact that $A\cap B=0$. Therefore $r\bar{m}=\bar{0}=N$.

Lemma 10. *Let R be a Dedekind ring, and let M, A, B, and N be as in Lemma 9. Let P be a prime ideal in R, and let $m \in M$. If $Pm \subseteqq B$, then $m \in N$.*

Proof. If $m \in B$, then $m \in N$. Suppose $m \notin B$. Then there exists a non-zero $a \in A \cap \langle B, m\rangle$, $a = b + rm$, $b \in B$, $r \in R$. Since $Pm \subseteqq B$, and $A \cap B = 0$, we have that $r \notin P$. However, P is a maximal ideal in R, so that there exist $p \in P$, and $u, v \in R$, such that $ur + vp = 1$. $m = rum + vpm = u(a - b) + vpm \in N$.

Theorem 2. *Let R be a principal ideal ring, and let M, A, B, and N be as above. Let π_A be the projection of N onto A. $M = N$ iff for every $m \in M$, and for every prime ideal P in R, $Pm \subseteqq N$ implies that $\pi_A(Pm) \subseteqq PA$.*

Proof. 1) Suppose that for every $m \in M$, and for every prime ideal P in R, $Pm \subseteqq N$ implies that $\pi_A(Pm) \subseteqq PA$. Let $m \in M$, and suppose that $Pm \subseteqq N$, P a proper prime ideal in R. Then $Pm \subseteqq PA \oplus B$. $P = \langle p \rangle$, $p \in P$, hence there exist $a \in A$, and $b \in B$ such that $pm = pa + b$, or $p(m - a) = b$. By Lemma 10, $m - a \in N$, and hence $m \in N$.

We have shown that for every $m \in M$, $m \notin N$, $P \subseteqq \operatorname{ord}(\overline{m})$ for every prime ideal P in R. By Lemma 9, M/N is a torsion module. A contradiction.

2) Suppose $M = N = A \oplus B$, and let P be a prime ideal in R. By Lemma 8, $PM = PA \oplus PB$, and hence $\pi_A(PM) = PA$.

Corollary. *Let R be a Dedekind ring, and let M, A, B and N be as above. If M is a torsion module, then the statement of Theorem 2 remains true.*

Proof. By Proposition 3 we may consider M to be a P-primary module. M is then an R_p module, R_p a principal ideal ring. We may therefore employ Theorem 2.

Notation: Let I be an ideal in R. Then $M[I] = \{m \in M \mid I \subseteqq \operatorname{ord}(m)\}$.

Theorem 3. *Let R be a Dedekind ring, M, A, B, and N as above, and let P be a prime ideal in R. Then $(M/N)[P] \cong [\langle PM, B\rangle \cap A]/PA$.*

Proof. If R is a principal ideal ring, then the theorem may be proved as in the case of abelian groups [4]. In the general case $(M/N)[P]$ is a P-primary module, and hence an R_p-module, so that the theorem remains true.

Several results concerning abelian groups may be generalized to modules over a Dedekind ring R as a result of Theorem 3; see [4]. For example, Kulikov's theorem stating that a bounded pure subgroup of an abelian group is a direct summand [3, Theorem 27.5] can thus be generalized. This result has already been obtained by Kaplansky [5, Theorem 5] in a different manner.

4. Notation. The tensor product $\otimes_R$ will be denoted by $\otimes$.

Lemma 11. *Let $\langle m\rangle$ be a cyclic R-module, N an arbitrary R-module. Then* $\langle m\rangle \otimes N \cong N/\mathrm{ord}\,(m)\cdot N$.

Proof. Same as for abelian groups [3, p. 255].

Theorem 4. *Let R be a Dedekind ring, and let*

$$(*)\qquad 0 \to A \xrightarrow{\varphi} B \xrightarrow{\psi} C \to 0$$

be an ideal pure exact sequence of R-modules. Then for every R-module M, the sequence

$$(**)\qquad 0 \to A\otimes \xrightarrow{\varphi\otimes 1_M} B\otimes M \xrightarrow{\psi\otimes 1_M} C\otimes M \to 0$$

is exact.

Proof. 1) Let M be a torsion module, and let M' be a finitely generated submodule of M. By Proposition 1, M' is a direct sum of cyclic modules. The sequence

$$(***)\qquad 0 \to A\otimes M' \xrightarrow{\varphi\otimes 1_M} B\otimes M' \xrightarrow{\psi\otimes 1_M} C\otimes M' \to 0$$

is therefore exact by Lemma 11 and Lemma 3.

For every R-module L,

$$L\otimes M \cong \varinjlim \{L\otimes M' \mid M' \text{ a finitely generated submodule of } M\}$$

so that $(**)$ is exact by [6, Theorem 2.13].

2) Let M be a torsion free module. Then M is flat [6, Theorem 4.23] so tha- $(**)$ is exact.

3) Let M be an arbitrary R-module, and let M' be a finitely generated subt module of M. Proposition 2 together with 1) and 2) yield that $(***)$ is exact. We may proceed as in 1) to obtain that $(**)$ is exact.

Lemma 12. *Let R be a Dedekind ring, and let J be an injective R-module. Then for every R-module M,* $M\otimes J \cong (M/M_t)\otimes J$.

Proof. The sequence

$$0 \to M_t \to M \to M/M_t \to 0$$

is exact, hence the sequence

$$M_t\otimes J \to M\otimes J \to (M/M_t)\otimes J \to 0$$

is exact. J is divisible so that $M_t\otimes J=0$. Therefore, $M\otimes J\cong (M/M_t)\otimes J$.

Let S be a maximal independent subset of an R-module M, and let $S_0=\{x\in \in S \mid x \text{is torsion free}\}$. It is easy to verify that the cardinality of S_0, $|S_0|$, is independent of the choice of S. We may therefore give the following

Definition 11. Let M, S, and S_0 be as above. Then $r_0(M)=|S_0|$ is called the *torsion free rank* of M.

Theorem 5. *Let R be a Dedekind ring, let J be a torsion free injective R-module, and let M be an arbitrary R-module. Then* $M\otimes J\cong \sum_{r_0(M)} \oplus J$.

Proof. By Lemma 12 we may assume that M is torsion free. Let S be a maximal independent subset of M. The sequence

$$0 \to \langle S\rangle \to M \to M/\langle S\rangle \to 0$$

is exact. J is flat so that the sequence

$$0 \to \langle S\rangle \otimes J \to M\otimes J \to (M/\langle S\rangle)\otimes J \to 0$$

is exact. $M/\langle S\rangle$ is a torsion module, and J is divisible. Hence $(M/\langle S\rangle)\otimes J=0$, and $M\otimes J\cong\langle S\rangle\otimes J\cong \sum_{r_0(M)} \oplus J$.

Corollary. *Let R be a Dedekind ring, K the quotient field of R, M and N torsion free R-modules. Then there exist embeddings,*

$$\sum_{r_0(M)r_0(N)} R \to M\otimes N, \quad \text{and} \quad M\otimes N \to \sum_{r_0(M)r_0(N)} K.$$

Proof. Let S be a maximal independent subset of M, and let T be a maximal independent subset of N. Then

$$\langle S\rangle \cong \sum_{r_0(M)} \oplus R, \quad \text{and} \quad \langle T\rangle \cong \sum_{r_0(N)} \oplus R.$$

The sequence

$$0 \to \langle S\rangle \otimes T \to M\otimes N$$

is exact ([1] Theorem 3, and [2] Lemma 6), and

$$\langle S\rangle \otimes \langle T\rangle \cong \sum_{r_0(M)r_0(N)} \oplus R.$$

N is flat, hence there exists an exact sequence $0\to N \xrightarrow{\varphi} N\otimes K$. However, M is also flat, so that the sequence

$$0 \to M\otimes N \xrightarrow{1_M\otimes\varphi} (M\otimes N)\otimes K$$

is exact. By Theorem 5

$$(M\otimes N)\otimes K \cong \sum_{r_0(M)r_0(N)} \oplus K.$$

Lemma 13. *Let M be a P-primary module, and let N be a P-divisible module. Then* $M\otimes N=0$.

Proof. Let $m\in M$, $n\in N$, and let e be a positive integer such that $P^e\subseteqq \operatorname{ord}(m)$. Since $N=P^eN$, there exist $r_i\in P^e$, $n_i\in N$, $1\leqq i\leqq k$ such that

$$n=\sum_{i=1}^{k} r_i n_i, \quad m\otimes n=\sum_{i=1}^{k} r_i m\otimes n_i=0.$$

Theorem 6. *Let R be a Dedekind ring, P a prime ideal in R, M and N P-primary R-modules, and S a maximal P-independent subset of M. Then $M\otimes N\cong\langle S\rangle\otimes N$.*

Proof. The sequence

$$0\to\langle S\rangle\to M\to M/\langle S\rangle\to 0$$

is ideal pure exact by Lemmata 5 and 2. By Theorem 4, the sequence

$$0\to\langle S\rangle\otimes N\to M\otimes N\to (M/\langle S\rangle)\otimes N\to 0$$

is exact. By Theorem 1, $M/\langle S\rangle$ is P-divisible. Hence by Lemma 13, $(M/\langle S\rangle)\otimes N=0$, and $M\otimes N\cong\langle S\rangle\otimes N$.

Corollary. *Let R be a Dedekind ring, and let M and N be torsion R-modules. Then $M\otimes N$ is a direct sum of cyclic modules.*

Proof. By Proposition 3, we may assume that M and N are P-primary modules. Let S be a maximal P-independent subset of M, and let T be a maximal P-independent subset of N. By Theorem 6, $M\otimes N\cong\langle S\rangle\otimes\langle T\rangle$. Proposition 1 and Lemma 11 yield that $S\otimes T$ is a direct sum of cyclic modules.

References

[1] J. Dieudonné, Sur les produits tensoriels, *Ann. Sci. Ecole Norm. Sup.*, **64** (1947), 101—117.

[2] S. Feigelstock, On the tensor product of exact sequences of modules over a Dedeking ring, *Archiv der Math.*, **25** (1973), 591—601.

[3] L. Fuchs, *Infinite Abelian Groups*, Academic Press (New York and London) vol. 1 (1970); vol. 2 (1973).

[4] L. Fuchs, On a useful lemma for Abelian groups, *Acta Sci. Math.*, **17** (1956), 134—138.

[5] I. Kaplansky, Modules over Dedekind rings, *Trans. Amer. Math. Soc.*, **72** (1952), 327—340.

[6] J. Rotman, *Notes on Homological Algebra*, Van Nostrand (New York and Toronto, 1970).

[7] E. Steinitz, Rechteckige Systeme und Moduln in algebraischen Zahlkörpern, *Math. Ann.*, **71** (1912), 328—354.

[8] O. Zariski—P. Samuel, *Commutative Algebra*, vol. 1, Van Nostrand (Princeton, New Jersey, 1958).

BAR-ILAN UNIVERSITY
RAMAT-GAN ISRAEL

Acta Sci. Math., 39 (1977), 265—272

Congruence-equalities and Mal'cev conditions in regular equational classes

H. PETER GUMM

FREESE and NATION have shown in [1] that there is no lattice equality holding in all congruence lattices of semilattices. It follows easily that this result remains true if one replaces the variety of semilattices by any variety defined by a set of regular equations. On the other hand not every algebraic lattice is the congruence lattice of a semilattice, see HALL [4] and PAPERT [5]. WILLE has introduced in [9] the notion of a congruence equality using the binary term $\circ$ (relational product) in addition to the binary terms $\vee$ (join) and $\wedge$ (meet). We are going to show in this paper that the result of Freese and Nation is also true for a certain class of congruence-equalities in $\wedge$, $\vee$ and $\circ$, and on the other hand we provide congruence-equalities which are nontrivial and which do hold in semilattices. This also gives us examples of congruence-equalities which do not imply any lattice equation.

Two such congruence equalities are characterized in terms of Mal'cev conditions and it turns out that they are within the class of regular varieties equivalent to the Mal'cev conditions

$$\exists p(p(x,x)=x, p(x,y)=p(y,x)), \quad \text{resp.} \quad \exists p(p(x,x,x)=x, p(x,y,z)=p(z,x,y)).$$

Finally we characterize the above Mal'cev conditions within the class of all varieties in terms of fixed points of involutions similar to [3]. For basic facts and notations used in this paper see GRÄTZER [2]. For the notion of equivalence see, e. g., TAYLOR [8].

1. Regular varieties

1.1. Definition. (PŁONKA [7]) An equation $p=q$ is called *regular* if the set of variables and constants appearing in p is the same as that in q. A *variety* is *regular* if it can be defined by a set of regular equations.

Received May 26, 1976.

1.2. Example. The variety of *semilattices* is a regular variety. The defining equations are: $x \cdot x = x$, $x \cdot y = y \cdot x$, $x \cdot (y \cdot z) = (x \cdot y) \cdot z$.

Next we formulate two basic lemmas. The first can be easily proved, and the second was essentially proved in [10].

1.3. Lemma. *Let $\Delta = (n_i | i \in I)$ be a type with corresponding function symbols f_i, $i \in I$. Let $\mathbf{2} := \{0, 1\}$ be a two-element set and define an algebra $\mathbf{2}_\Delta$ by setting*

$$f_i(x_1, \dots, x_{n_i}) := \begin{cases} 1 & \text{if} \quad x_1 = x_2 = \dots = x_{n_i} = 1 \\ 0 & \text{otherwise.} \end{cases}$$

If there are 0-ary function symbols define them to be 0. Let $\mathbf{SL}_\Delta$ be the variety generated by $\mathbf{2}_\Delta$. Then,

(i) $\mathbf{SL}_\Delta$ *is equivalent to* $\mathbf{SL}_{(2)}$, *the variety of all semilattices iff $n_i \geqq 2$ for some i, and $n_i \neq 0$ for all i.*

(ii) $\mathbf{SL}_\Delta$ *is equivalent to* $\mathbf{SL}_{(0,2)}$, *the variety of all 0-semilattices, iff $n_i \geqq 2$ for some i and $n_i = 0$ for some i.*

(iii) $\mathbf{SL}_\Delta$ *is equivalent to* $\mathbf{\Omega}_\Delta$, *the variety of pointed sets iff $n_i \leqq 1$ for all i, and $n_i = 0$ for some i.*

(iv) $\mathbf{SL}_\Delta$ *is equivalent to the variety of sets otherwise.*

1.4. Lemma. [10] *Let $\mathfrak{B}$ be a variety of type Δ, containing no nullary operation. Then $\mathfrak{B}$ is regular if and only if $\mathfrak{B}$ contains $\mathbf{SL}_\Delta$ as a subvariety. If Δ contains a 0-ary operation, the only if part is still true.*

2. Congruence equalities

Congruence equalities were introduced by Wille [9].

2.1. Definition. A *congruence equality* is an expression $\alpha = \beta$ where α and β are terms in variables and the binary polynomial symbols $\wedge$, $\vee$ and $\circ$. A congruence-equality $\alpha = \beta$ is said to be *congruence-valid* in an algebra $\mathfrak{A}$ if for any interpretation of the variables occurring in $\alpha = \beta$ by congruences of $\mathfrak{A}$ the equation holds if we interpret $\wedge$ as meet, $\circ$ as relational product and $\vee$ as *relational join,* that means: If σ and τ are binary relations on A, we define: $\sigma \vee \tau := \bigcup_{n \in \mathbf{N}} \{\underbrace{\sigma \circ \tau \circ \sigma \circ \dots \circ \tau}_{\text{n-times}} | n \in \mathbf{N}\}$.

We have to be careful because if γ, θ are congruences then $\gamma \circ \theta$ need not be a congruence. If σ and τ happen to be congruences, then $\sigma \vee \tau$ is the join of σ and τ.

We call a congruence-equality *trivial* if it holds in each partition lattice. We say that $\alpha = \beta$ is *congruence-valid* in a variety $\mathfrak{B}$ if it is congruence-valid for each algebra $\mathfrak{A} \in \mathfrak{B}$.

Now it is obvious what we mean by a *congruence-inequality* $\alpha\leqq\beta$ and in fact we can replace each congruence-equality $\alpha=\beta$ by the congruence-inequalities $\alpha\leqq\beta$ and $\alpha\geqq\beta$. Clearly, a congruence inequality $\alpha\leqq\beta$ which holds in a variety $\mathfrak{B}$ will hold in each variety $\mathfrak{B}'$ which is equivalent to $\mathfrak{B}$ as well.

For the proof of our first theorem we need the following simple lemma:

2.2. Lemma. *Let $\alpha\leqq\beta$ be a nontrivial congruence-inequality. Then there exists a finite set X, such that $\alpha\leqq\beta$ fails to hold in $\pi(X)$, the partition lattice of X.*

Proof. The proof essentially uses the ideas of theorem 6.15 in WILLE [9]. $\alpha\leqq\beta$ is nontrivial, thus there exists a set X such that $\alpha\leqq\beta$ does not hold for the partitions of X. Let $x_1, \ldots, x_n$ be the variables occurring in $\alpha\leqq\beta$. Let $\mathbf{i}$ be an interpretation map assigning to x_i, $0<i\leqq n$, the partition θ_i of X such that for a certain pair (a, b) we have $(a, b)\in\mathbf{i}(\alpha)$ and $(a, b)\notin\mathbf{i}(\beta)$.

Let γ now be an arbitrary expression in $\wedge$, $\vee$ and $\circ$ and the variables amongst $\{x_1, \ldots, x_n\}$. Let x, y be arbitrary elements of X. Define recursively:

1) If γ is a variable,

$$R^{\gamma}_{(x,y)} := \begin{cases} \{x, y\} & \text{if } (x, y)\in\mathbf{i}(\gamma) \\ \emptyset & \text{otherwise.} \end{cases}$$

2) If $\gamma=\sigma\circ\tau$,

$$R^{\gamma}_{(x,y)} := \begin{cases} R^{\sigma}_{(x,z)}\cup R^{\tau}_{(z,y)} & \text{for some } z \text{ with } (x, z)\in\mathbf{i}(\sigma) \text{ and } (z, y)\in\mathbf{i}(\tau) \\ \emptyset & \text{if } (x, y)\notin\mathbf{i}(\gamma). \end{cases}$$

3) If $\gamma=\sigma\vee\tau$,

$$R^{\gamma}_{(x,y)} := \begin{cases} R^{\sigma}_{(x,z_1)}\cup R^{\tau}_{(z_1,z_2)}\cup\ldots\cup R^{\tau}_{(z_n,y)} & \text{for some } z_1, \ldots, z_n \text{ with} \\ & x\,\mathbf{i}(\sigma)\,z_1\,\mathbf{i}(\tau)\,z_2\ldots z_n\,\mathbf{i}(\tau)\,y \\ \emptyset & \text{if } (x, y)\notin\mathbf{i}(\gamma). \end{cases}$$

4) If $\gamma=\sigma\wedge\tau$,

$$R^{\gamma}_{(x,y)} := \begin{cases} R^{\sigma}_{(x,y)}\cup R^{\tau}_{(x,y)} & \text{if } (x, y)\in\mathbf{i}(\gamma) \\ \emptyset & \text{otherwise.} \end{cases}$$

Then $X_0 := R^{\alpha}_{(a,b)}$ is finite and nonempty. Define $\theta^0_i := \theta_i\cap X_0\times X_0$ and $\mathbf{i}_0\colon x_i\to\theta^0_i$, $0<i\leqq n$. Then clearly by the construction we have $(a, b)\in\mathbf{i}_0(\alpha)$ and $(a, b)\notin\mathbf{i}_0(\beta)$. Thus $\alpha=\beta$ does not hold for the partitions of the finite set X_0.

2.3. Theorem. *Let $\alpha\leqq\beta$ a nontrivial congruence-inequality where α is arbitrary and β is of the form $\sigma_1\wedge\sigma_2\wedge\ldots\wedge\sigma_k$ where each σ_i is a term in $\vee$ and $\circ$. Then each regular variety contains a finite algebra where $\alpha\leqq\beta$ is not congruence-valid.*

Proof. If a congruence-inequality holds in a variety $\mathfrak{B}$ then it obviously holds in each subvariety of $\mathfrak{B}$ and in each variety which is equivalent to $\mathfrak{B}$. By lemma 1.5

we need to prove our statement only for $\mathbf{SL}_\Delta$. As the variety of sets and the variety of pointed sets do not fulfil any nontrivial congruence equality we need in view of lemma 1.4 only consider $\mathbf{SL}_{(2)}$ and $\mathbf{SL}_{(0,2)}$, semilattices and 0-semilattices. Let now X be a set, π a partition of X and $\mathbf{FSL}\,(X)$ (resp. $\mathbf{FSL}_0\,(X)$) be the free semilattice resp. 0-semilattice generated by X. Let θ_π be the congruence generated by π in $\mathbf{FSL}\,(X)$ (resp. $\mathbf{FSL}_0\,(X)$) and let p and q be elements of $\mathbf{FSL}\,(X)$ (resp. $\mathbf{FSL}_0\,(X)$). We assume that p and q are in reduced normal form. Then we have $p\theta_\pi q$ if and only if for each variable x in p there is a variable y in q such that $x\pi y$ and vice versa.

By a repeated use of this argument one obtains that for a set $\pi_1, \ldots, \pi_n$ of partitions of X and $x, y \in X$ we have:

$$(*) \qquad x\theta_{\pi_1}\circ \ldots \circ\theta_{\pi_n}y \quad \text{if and only if} \quad x\pi_1\circ \ldots \circ\pi_n y.$$

Now let $\alpha \leqq \beta$ a congruence-inequality of the form required in our theorem. Then there exists by lemma 2.2 a finite set X and partitions $\pi_1, \ldots, \pi_n$ of X and an interpretation $\mathbf{i}$ assigning the variables $x_1, \ldots, x_n$ of $\alpha \leqq \beta$ to the partitions $\pi_1, \ldots, \pi_n$ such that for some $x, y \in X$ we have $(x, y) \in \mathbf{i}(\alpha)$ and $(x, y) \notin \mathbf{i}(\beta)$.

Take now $\mathbf{FSL}\,(X)$ resp. $\mathbf{FSL}_0\,(X)$ and define $\bar{\mathbf{i}}: x_i \mapsto \theta_{\pi_i}$. Of course we still have $(x, y) \in \bar{\mathbf{i}}(\alpha)$, but by $(*)$ we have $(x, y) \notin \bar{\mathbf{i}}(\beta)$. Thus $\alpha \leqq \beta$ does not hold in $\mathbf{FSL}\,(X)$ nor in $\mathbf{FSL}_0\,(X)$; and both are finite algebras, which concludes the proof.

2.4. Definition. A variety is *n-permutable* iff the congruence-inequality $\theta_1\circ\theta_2\circ\ldots\circ\theta_2 \subseteqq \theta_2\circ\theta_1\circ\ldots\circ\theta_1$, with n factors on each side, holds in $\mathfrak{V}$.

2.5. Corollary. *Regular varieties are not n-permutable for any n.*

Now we are going to show that we cannot drop the assumption on the form of β.

3. Mal'cev conditions

For basic facts concerning Mal'cev conditions see e.g. Taylor [8].

3.1. Definition. A strong Mal'cev condition is an expression of second order logic of the form $\exists p_1, \ldots, p_n(\Sigma)$ where Σ is a finite conjunction of equations universally quantified in individual variables, containing the function variables $p_1, \ldots, p_n$. A strong Mal'cev condition $\mathbf{M} := \exists p_1, \ldots, p_n(\Sigma)$ holds in a variety $\mathfrak{V}$ (shortly $\mathfrak{V} \vdash \mathbf{M}$) iff there exist polynomials $p_1, \ldots, p_n$ in the language of $\mathfrak{V}$ such that Σ holds in $\mathfrak{V}$.

3.2. Definition. An *involution* is an automorphism of order two.

3.3. Theorem. *For an arbitrary variety $\mathfrak{V}$ the following are equivalent:*

(i) *The strong Mal'cev condition $\exists p\big(p(x, x) = x \wedge p(x, y) = p(y, x)\big)$ holds in $\mathfrak{V}$.*

(ii) *If φ is an involution of an algebra $\mathfrak{A} \in \mathfrak{V}$ then for each $x \in \mathfrak{A}$ there exists a fixed point y of φ such that $(x, \varphi x) \in \theta$ implies $(x, y) \in \theta$ for arbitrary congruences θ of $\mathfrak{A}$.*

A similar theorem with automorphisms of order n holds for the Mal'cev condition $\exists p(p(x, \ldots, x)=x \wedge p(x_1, \ldots, x_n)=p(x_2, \ldots, x_n, x_1))$.

Proof. (i)$\to$(ii): Assume (i) and let φ be an involution of $\mathfrak{A}\in\mathfrak{B}$. Take $x\in\mathfrak{A}$. Then define $y:=p(x, \varphi x)$. We have: $\varphi(y)=\varphi(p(x, \varphi x))=p(\varphi(x), \varphi^2(x))= =p(\varphi(x), x)=p(x, \varphi(x))=y$. Thus y is a fixed point of φ. Assume $(x, \varphi x)\in\theta$. Then $x=p(x, x)\theta p(x, \varphi x)=y$. Thus $(x, y)\in\theta$.

(ii)$\to$(i) Let $\mathbf{F}_{\mathfrak{B}}(x, y)$ be the free algebra in $\mathfrak{B}$ generated by the two distinct elements x and y. Then the map $\varphi: x\to y, y\to x$ extends uniquely to a homomorphism φ of $\mathbf{F}_{\mathfrak{B}}(x, y)$ which is moreover an involution. For x we then have an element $z\in\mathbf{F}_{\mathfrak{B}}(x, y)$ which is a fixed point of φ. Here $z=p(x, y)$ for some polynomial p and $\varphi z=z$, thus $\varphi p(x, y)=p(x, y)$. As $\varphi p(x, y)=\varphi p(x, \varphi x)=p(\varphi x, \varphi^2 x)=p(y, x)$ we conclude $p(x, y)=p(y, x)$. Now $(x, \varphi x)\in\theta_{(x,y)}$, the smallest congruence which collapses x and y. By (ii) we have: $(x, z)\in\theta_{(x,y)}$ which means $(x, p(x, y))\in\theta_{(x,y)}$ and thus $p(x, x)=x$. Hence, $p(x, y)=p(y, x)$ and $p(x, x)=x$ holds in the variety $\mathfrak{B}$. Wille [9] and Pixley [6] have shown that in a variety each congruence-inequality in $\wedge$, $\vee$ and $\circ$ is equivalent to a countable conjunction of countable disjunctions of strong Mal'cev conditions.

Let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{g}$ be the following congruence inequalities:

$$\mathbf{e}_1: \theta_0 \wedge (\theta_1\circ\theta_2) \wedge (\theta_3\circ\theta_4) \leqq \theta_1\circ\{(\theta_2\circ\theta_3) \wedge \{[(\theta_1\circ\theta_3) \wedge (\theta_2\circ\theta_4)]\circ\theta_0\}\}\circ\theta_4,$$

$$\mathbf{e}_2: (\theta_1\circ\theta_2) \wedge (\theta_3\circ\theta_4) \leqq \theta_1\circ\{(\theta_2\circ\theta_3) \wedge \{[(\theta_1\circ\theta_3) \wedge (\theta_2\circ\theta_4)]\circ[(\theta_1\circ\theta_2) \wedge (\theta_3\circ\theta_4)]\}\}\circ\theta_4.$$

($\mathbf{e}_2$ is obtained by replacing θ_0 in $\mathbf{e}_1$ by $(\theta_1\circ\theta_2)\wedge(\theta_3\circ\theta_4)$.

$$\mathbf{g}: \theta_0 \wedge \{\theta_1\circ[\theta_2 \wedge (\theta_3\circ\theta_4)]\} \wedge \{[\theta_5 \wedge (\theta_6\circ\theta_7)]\circ\theta_8\} \subseteqq$$
$$\subseteqq \theta_1\circ\theta_6\circ\{(\theta_0\circ\theta_3\circ\theta_7) \wedge \{\theta_5\circ\theta_2\circ[(\theta_6\circ\theta_1\circ\theta_3) \wedge (\theta_7\circ\theta_8\circ\theta_4)]\}\}\circ\theta_4\circ\theta_8.$$

Then we have the following theorems:

3.4. Theorem. *For a regular variety the following are equivalent:*

(i) $\mathbf{e}_1$ *is congruence-valid in* $\mathfrak{B}$.

(ii) $\mathbf{e}_2$ *is congruence-valid in* $\mathfrak{B}$.

(iii) *The strong Mal'cev condition* $\exists p\,(p(x, x)=x\wedge p(x, y)=p(y, x))$ *holds in* $\mathfrak{B}$.

3.5. Theorem. *For a regular variety t.f.a.e.:*

(i) $\mathbf{g}$ *is congruence-valid in* $\mathfrak{B}$.

(ii) *The strong Mal'cev condition* $\exists p\,(p(x, x, x)=x, p(x, y, z)=p(y, z, x))$ *holds in* $\mathfrak{B}$.

We prove only the first theorem, the proof of the second is essentially the same but needs a little bit more of computation.

Proof. (iii)$\to$(i): Assume in $\mathfrak{B}$ there exists an idempotent and commutative binary polynomial p. Take $(x, y)\in\theta_0\wedge(\theta_1\circ\theta_2)\wedge(\theta_3\circ\theta_4)$. Then there exist a and b

such that $x\theta_0 y$, $x\theta_1 a\theta_2 y$, $x\theta_3 b\theta_4 y$. Using p we get:

$$x = p(x,x)\theta_1 p(a,x)\theta_2 p(y,x)\theta_3 p(y,b)\theta_4 p(y,y) = y \quad \text{and}$$
$$p_1(a,x)\,[(\theta_1\circ\theta_3)\wedge(\theta_2\circ\theta_4)]\,p(b,x)\theta_0 p(b,y)$$

As $p(y,b)=p(b,y)$ we get:

$$(x,y)\in\theta_1\circ\{(\theta_2\circ\theta_3)\wedge\{[(\theta_1\circ\theta_3)\wedge(\theta_2\circ\theta_4)]\circ\theta_0\}\}\circ\theta_4.$$

(i)$\to$(ii) is trivial. Only in the next step will we use regularity.

(ii)$\to$(iii): First we use Wille's algorithm to write down the Mal'cev condition for $\mathbf{e}_2$. We get, that in the class of all varieties $\mathbf{e}_2$ is equivalent to the following strong Mal'cev condition: $\exists p_1, p_2, \ldots, p_8$ with

(1) $x = p_1(x,x,v,y)$,

(2) $p_1(x,y,v,y) = p_2(x,y,v,y)$

(3) $p_2(x,u,x,y) = p_3(x,u,x,y)$,

(4) $p_3(x,u,y,y) = y$,

(5) $p_1(x,x,v,y) = p_5(x,x,v,y)$,

(6) $p_5(x,u,x,y) = p_4(x,u,x,y)$,

(7) $p_4(x,x,v,y) = p_7(x,x,v,y)$,

(8) $p_7(x,y,v,y) = p_3(x,y,v,y)$,

(9) $p_1(x,y,v,y) = p_6(x,y,v,y)$,

(10) $p_6(x,u,y,y) = p_4(x,u,y,y)$,

(11) $p_4(x,u,x,y) = p_8(x,u,x,y)$,

(12) $p_8(x,u,y,y) = p_3(x,u,y,y)$.

Now if this Mal'cev condition holds in a regular variety, each of its equations must be regular. We can thus conclude: From (1) it follows that p_1 depends only on the first two places, therefore in (2) p_2 can depend at most on the first, second and fourth place. From (4) it follows that p_3 depends at most on the last two places thus p_2 depends at most on the first, third and fourth place. Together with the above then p_2 depends at most on the first and fourth place. Thus we can replace (1) to (4) in a regular variety by

(1′) $x = p_1(x,x)$,

(2′) $p_1(x,y) = p_2(x,y)$,

(3′) $p_2(x,y) = p_3(x,y)$,

(4′) $p_3(y,y) = y$.

Carrying these cancellations out in (5) up to (12) we finally obtain: $\exists p_1, \ldots, p_8$ with (1′) to (4′) and

(5′) $$p_1(x, x) = p_5(x),$$
(6′) $$p_5(x) = p_4(x, x),$$
(7′) $$p_4(x, v) = p_7(x, v),$$
(8′) $$p_7(y, v) = p_3(v, y),$$
(9′) $$p_1(x, y) = p_6(x, y),$$
(10′) $$p_6(x, y) = p_4(x, y),$$
(11′) $$p_4(x, x) = p_8(x),$$
(12′) $$p_8(y) = p_3(y, y).$$

Now let us have a look at p_1. By (1′) we get $p_1(x, x) = x$ and we obtain

$$p_1(x, y) = p_2(x, y) = p_3(x, y) = p_7(y, x) = p_4(y, x) = p_6(y, x) = p_1(y, x).$$

Thus we have: $\exists p$ with $p(x, x) = x \wedge p(x, y) = p(y, x)$.

This finishes the proof.

4. Applications

We consider the equational classes of groupoids defined by subsets of the following set $\mathbf{\Sigma}$ of regular equations

$$\mathbf{\Sigma} := \{x(yz) = (xy)z, xy = yx, xx = x\},$$

and define

$\mathfrak{B}_1 = \mathbf{Mod}\,(x(yz) = (xy)z)$	semigroups,
$\mathfrak{B}_2 = \mathbf{Mod}\,(xy = yx)$	commutative groupoids,
$\mathfrak{B}_3 = \mathbf{Mod}\,(xx = x)$	idempotent groupoids,
$\mathfrak{B}_4 = \mathbf{Mod}\,(x(yz) = (xy)z,\ xy = yx)$	commutative semigroups,
$\mathfrak{B}_5 = \mathbf{Mod}\,(x(yz) = (xy)z,\ xx = x)$	idempotent semigroups,
$\mathfrak{B}_6 = \mathbf{Mod}\,(xy = yx,\ xx = x)$	commutative, idempotent groupoids,
$\mathfrak{B}_7 = \mathbf{Mod}\,(x(yz) = (xy)z,\ xy = yx,\ xx = x)$	semilattices.

As projections: $\pi_i^n(x_1, \ldots, x_n) := x_i$ are idempotent and associative we have that the variety of sets is contained up to polynomial equivalence as a subvariety in $\mathfrak{B}_1, \mathfrak{B}_3, \mathfrak{B}_5$. Furthermore, the variety of pointed sets is up to equivalence contained in $\mathfrak{B}_2$ and in $\mathfrak{B}_4$, so $\mathfrak{B}_1, \mathfrak{B}_2, \mathfrak{B}_3, \mathfrak{B}_4$ and $\mathfrak{B}_5$ do not fulfil any nontrivial congruence inequalities.

We are going to show now that we can separate the remaining varieties by congruence inequalities.

4.1. Theorem. *The congruence inequalities* $\mathbf{e}_1$ *and* $\mathbf{e}_2$ *are nontrivial and hold in commutative, idempotent groupoids. The congruence inequality* $\mathbf{g}$ *holds in semilattices but not in commutative idempotent groupoids.*

Proof. The first part of the theorem is a direct consequence of theorem 3.2. Theorem 3.3. implies that $\mathbf{g}$ holds in semilattices. Assume $\mathbf{g}$ holds in commutative idempotent groupoids.

In [3] we characterized the strong Mal'cev condition $\exists p(p(x, y, z)=p(y, z, x))$ and it was shown that it is equivalent to the statement that every automorphism φ of order 3 has a fixed point.

So in order to show that $\mathbf{g}$ does not hold for all commutative idempotent groupoids we only have to find a commutative idempotent groupoid $\mathcal{G}$ and an automorphism $\varphi: G \to G$ of order 3 which has no fixed point.

Take $\mathcal{G}=(\{0, 1, 2\}, \cdot)$ with $\cdot$ defined as $x \cdot y := 2x+2y \pmod 3$. Take the map $\varphi: G \to G$ with $\varphi(x) := x+1 \pmod 3$. φ is an automorphism of order 3 but φ has no fixed point. This finishes the proof. Notice that $\mathbf{g}$ happens to hold in $\mathcal{G}$ because $\mathcal{G}$ is simple.

4.2. Corollary. *The congruence inequalities* $\mathbf{e}_1, \mathbf{e}_2$ *and* $\mathbf{g}$ *do not imply any lattice inequality.*

Proof. Freese and Nation have shown that there is no lattice inequality holding for the congruence lattices of semilattices, but $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{g}$ are congruence-valid in semilattices.

The author wishes to acknowledge with many thanks the helpful discussions with Professor R. Padmanabhan, who also provided the commutative idempotent groupoid used in the proof of theorem 4.1.

References

[1] R. Freese and J. B. Nation, Congruence lattices of semilattices, *Pacific J. Math.*, **49** (1973), 51—58.

[2] G. Grätzer, *Universal algebra*, Van Nostrand (Princeton, N. J., 1968).

[3] H. P. Gumm, Mal'cev conditions in sums of varieties and a new Mal'cev condition, *Algebra Universalis*, **5** (1975), 56—64.

[4] T. E. Hall, On the lattice of congruences on a semilattice, *Australian Math. Soc.*, **12** (1971), 456—460.

[5] D. Papert, Congruence relations in semilattices, *London Math. Soc.*, **39** (1964) 723—729.

[6] A. F. Pixley, Local Mal'cev conditions, *Canad. Math. Bull.*, **15** (1972), 559—568.

[7] J. Płonka, On a method of construction of abstract algebras, *Fund. Math.*, **61** (1967), 183—189.

[8] W. Taylor, Characterizing Mal'cev conditions, *Algebra Universalis*, **3** (1973), 351—397.

[9] R. Wille, *Kongruenzklassengeometrien*, Lecture notes in Mathematics 113, Springer Verlag (Berlin, 1970).

[10] B. Jónsson, E. Nelson, Relatively free products in regular varieties, *Alg. Univ.*, **4** (1974), 14—19.

UNIVERSITY OF MANITOBA
WINNIPEG, MANITOBA
CANADA

Acta Sci. Math., 39 (1977), 273—289

On the problem of the choice of first approximants in a two-sided iteration method

J. HEGEDŰS

Introduction

In the papers [2], [3] a two-sided iteration method is worked out for the general homogeneous boundary value and initial value problems of non-linear differential equations of order n under the assumption of the contractivity of the corresponding integral operator A. More precisely, two sequences of functions were constructed that approximated, together with their derivatives, from above and from below arbitrarily precisely and uniformly on the considered segment [0, 1] the solution of the boundary value problem and its derivatives. However, the construction of the approximants just mentioned depends on the assumption of the existence of first approximants z_1, w_1 that are in a well-defined relation with the second, and, which is even worse, with the third pair of approximants.

In this note we shall prove that under the only assumption of contractivity the first pair of approximants exists.

Let us consider the boundary value problem

$$(0.1)\qquad \begin{cases} y^{(n)}(x) = f[y] \equiv f(x, y(x), \ldots, y^{(n-1)}(x)) \quad (0 \leqq x \leqq 1, n \geqq 2), \\ L_i y = \sum_{k=0}^{n-1} \left(a_{ik} y^{(k)}(0) + b_{ik} y^{(k)}(1)\right) = 0 \quad (i = 0, \ldots, n-1) \end{cases}$$

with the given function

$$f(x, u_0, \ldots, u_{n-1}) : [0, 1] \times \mathscr{R} \times \ldots \times \mathscr{R} \to \mathscr{R}, \quad \left|\frac{\partial f}{\partial u_i}\right| \leqq N \quad (i = 0, \ldots, n-1)$$

that is continuous and continuously differentiable with respect to each u_i ($i=0, \ldots, n-1$), and with coefficients a_{ik}, b_{ik} such that the problem

$$y^{(n)}(x) = 0 \quad (0 \leqq x \leqq 1;\ n \geqq 2), \quad L_i y = 0 \quad (i = 0, \ldots, n-1)$$

has the only solution $y=0$.

Received November 1, 1975.

We remark that the general problem on $[a, b]$ with homogeneous boundary value condition leads to problem (0.1).

Problem (0.1) is equivalent to the integral equation

$$y = y(x) = Ay \equiv \int_0^1 G(x,t) f[y(t)]\,dt, \tag{0.2}$$

(where G is Green's function) in the space $\mathcal{M}$ of $n-1$ times continuously differentiable functions defined on $[0, 1]$ and satisfying the boundary value restriction. Let us introduce an ordering and a metric in $\mathcal{M}$ by the formulas

$$\begin{cases} z \leqq w \Leftrightarrow z^{(i)}(x) \leqq w^{(i)}(x) \quad (0 \leqq x \leqq 1;\ i = 0, \ldots, n-1), \\ \varrho(z, w) = \|z-w\| = \sum\limits_{i=0}^{n-1} \max\limits_{[0,1]} |z^{(i)}(x) - w^{(i)}(x)|. \end{cases} \tag{0.3}$$

Let us suppose that the condition

$$N \sum_{i=0}^{n-1} \max_{[0,1]} \int_0^1 \left| \frac{\partial^i G(x,t)}{\partial x^i} \right| dt = \theta < 1 \tag{0.4}$$

is satisfied. This means that the operator A is strongly contractive in $\mathcal{M}$. In the paper [3] for problem (0.2) and for given $\varepsilon > 0$ we constructed an auxiliary function (minorant) $\tilde{G}(x, t)$, which is $n-1$ times differentiable with respect to x in each of the sectors

$$x_j \leqq x \leqq x_{j+1}, \quad 0 \leqq t \leqq x; \quad x_j \leqq x \leqq x_{j+1}, \quad x \leqq t \leqq 1 \quad (j = 0, \ldots, m-1),$$

where $\max\limits_{j=0,\ldots,m-1} (x_{j+1} - x_j)$ is sufficiently small (cf. the concluding part of the proof of the theorems),

$$x_0 = 0 < x_1 < \ldots < x_m = 1$$

are numbers, and along the straight lines $x = x_j$, $x = x_{j+1}$; $x = t$ the function $\tilde{G}$ or some of its derivatives with respect to x may be multivalued (they can have points of discontinuity of the first kind), moreover, the inequalities

$$N \sum_{i=0}^{n-1} \max_{[0,1]} \int_0^1 \left| \frac{\partial^i \tilde{G}(x,t)}{\partial x^i} \right| dt = \theta_1 \leqq \theta + \varepsilon < 1 \tag{0.5}$$

$$\frac{\partial^i \tilde{G}(x,t)}{\partial x^i} \leqq - \left| \frac{\partial^i G(x,t)}{\partial x^i} \right| \quad (i = 0, \ldots, n-1) \tag{0.6}$$

are satisfied.

Finally, we introduce the linear space $\tilde{\mathcal{M}}$ of $n-1$ times continuously differentiable functions defined on the segments $[x_j, x_{j+1}]$ $(j = 0, \ldots, m-1)$. We consider

$\tilde{\mathcal{M}}$ with the following ordering and metric

$$(0.7)\qquad z \leqq w \Leftrightarrow \begin{cases} z^{(i)}(x) \leqq w^{(i)}(x) & \left(x \in \bigcup_{j=0}^{m-1} (x_j, x_{j+1})\right), \\ z^{(i)}(x_j+0) \leqq w^{(i)}(x_j+0), \\ z^{(i)}(x_{j+1}-0) \leqq w^{(i)}(x_{j+1}-0) \\ (i = 0, \dots, n-1;\ j = 0, \dots, m-1), \end{cases}$$

$$(0.8)\qquad \varrho(z, w) = \|z-w\| = \sum_{i=0}^{n=1} \max_{[0,1]} |z^{(i)}(x)-w^{(i)}(x)|.$$

It is obvious that $\tilde{\mathcal{M}} \supset \mathcal{M}$ and that the ordering and metric of $\tilde{\mathcal{M}}$ are extensions of those of $\mathcal{M}$, thus our above notation is justified. It is also obvious that for the operator $\tilde{A}$ (the extension of A to $\tilde{\mathcal{M}}$), we have

$$(0.9)\qquad \tilde{A}z = \int_0^1 G(x, t) f[z(t)]\,dt = \sum_{j=0}^{m-1} \int_{x_j}^{x_{j+1}} G(x, t) f[z(t)]\,dt$$

and for any $z, w \in \tilde{\mathcal{M}}$

$$(0.10)\qquad \varrho(\tilde{A}z, \tilde{A}w) \leqq \theta\varrho(z, w).$$

Let us introduce the notation

$$\Delta_{z,w}(t) = \sum_{i=0}^{n-1} (z^{(i)}(t) - w^{(i)}(t)), \quad \tilde{B}(z, w) = N\int_0^1 \tilde{G}(x, t)\Delta_{z,w}(t)\,dt.$$

Let us define the operators $E, F\colon \tilde{\mathcal{M}} \times \tilde{\mathcal{M}} \to \tilde{\mathcal{M}}$ in the following way

$$E(z, w) = \frac{1}{2}(\tilde{A}z + \tilde{A}w) + \frac{1}{2}\tilde{B}(z, w),$$

$$F(z, w) = \frac{1}{2}(\tilde{A}z + \tilde{A}w) - \frac{1}{2}\tilde{B}(z, w).$$

Consider the iteration process

$$(0)\qquad z_{p+1} = E(z_p, w_p), \quad w_{p+1} = F(z_p, w_p) \quad (p = 1, 2, \dots).$$

We remark that E is non-increasing in z and non-decreasing in w; F is non-decreasing in z and non-increasing in w.

In the so-called monotone case of problem (0.1), i.e., when

$$\frac{\partial^i G(x, t)}{\partial x^i} \leqq 0 \quad (0 \leqq x, t \leqq 1; \quad i = 0, \dots, n-1),$$

or when the derivatives of G are non-negative, $\tilde{G}$ can be taken to be equal to $-|G|$ and $\tilde{\mathcal{M}}$ can be taken to be equal to $\mathcal{M}$ and one can consider in it the approximation according to the rula (0).

I.

Let us consider first the case of problem (0.1) when the following strong assumption of contractivity is satisfied:

$$\text{(i)} \qquad N\sum_{i=0}^{n-1}\max_{[0,1]}\int_0^1\left|\frac{\partial^i G(x,t)}{\partial x^i}\right|dt = \tilde{\theta} < \frac{1}{2}.$$

It is obvious that in this case there exists a minorant with the required properties of smoothness and for which we have

$$(1.1) \qquad N\sum_{i=0}^{n-1}\max_{[0,1]}\int_0^1\left|\frac{\partial^i \tilde{G}(x,t)}{\partial x^i}\right|dt = \tilde{\tilde{\theta}} < \frac{1}{2}.$$

Lemma 1.1. *Under the assumption of* (i) *there exist* $z_1, w_1 \in \tilde{\mathcal{M}}$ *such that*

$$(1.2) \qquad z_2 \leqq z_1, \quad w_1 \leqq w_2.$$

Proof. Let us take a positive element φ from $\tilde{\mathcal{M}}$, i.e., a function $\varphi(x)$ for which

$$\varphi^{(i)}(x) > 0 \quad (i = 0, \ldots, n-1;\ 0 \leqq x \leqq 1).$$

The assertion of our lemma follows from the fact that by virtue of (1.1) the system of equations

$$(1.3) \qquad E(z_1, w_1) = z_1 - \varphi, \quad F(z_1, w_1) = w_1 + \varphi$$

has a solution in $\tilde{\mathcal{M}} \times \tilde{\mathcal{M}}$.

Theorem 1.1. *Under the assumption of* (i) *there exist* $z_1, w_1 \in \tilde{\mathcal{M}}$ *for which*

$$z_2 \leqq z_1, \quad z_3 \leqq z_1; \quad w_1 \leqq w_2, \quad w_1 \leqq w_3.$$

Proof. The elements $z_1, w_1 \in \tilde{\mathcal{M}}$ can be taken as in Lemma 1.1. On account of the monotonicity of E, F we then have

$$z_2 \leqq z_1, \quad z_2 \leqq z_3; \quad w_1 \leqq w_2, \quad w_3 \leqq w_2.$$

It remains to prove that the function φ can be chosen so that the inequalities

$$(1.4) \qquad z_3 \leqq z_1, \quad w_1 \leqq w_3$$

are also satisfied. These inequalities are equivalent to the inequalities

$$(1.5) \qquad \begin{cases} 0 \leqq \varphi + [z_1 - \varphi - E(z_1 - \varphi, w_1 + \varphi)] \\ -\varphi + [w_1 + \varphi - F(z_1 - \varphi, w_1 + \varphi)] \leqq 0. \end{cases}$$

We remark that by Lagrange's formula

$$f[z_1-\varphi]-f[z_1] = -\sum_{i=0}^{n-1} \frac{\partial f}{\partial u_i}\bigg|_{P(x)} \varphi^{(i)},$$

$$f[w_1+\varphi]-f[w_1] = \sum_{i=0}^{n-1} \frac{\partial f}{\partial u_i}\bigg|_{Q(x)} \varphi^{(i)},$$

where we denoted by $\frac{\partial f}{\partial u_i}\Big|_{P(x)}, \frac{\partial f}{\partial u_i}\Big|_{Q(x)}$ the values of these derivatives at the corresponding points of the $n+1$ dimensional space (in the sequel the letters P, Q shall be often omitted). By using these formulas it is easy to prove that the inequalities (1.5), understood in the sense of the ordering of $\tilde{\mathcal{M}}$, are equivalent to the inequalities

(1.6)
$$\varphi^{(r)}(x) \geqq \frac{1}{2}\int_0^1 \sum_{i=0}^{n-1}\left[\pm\frac{\partial^r G(x,t)}{\partial x^r}\left(\frac{\partial f}{\partial u_i}\bigg|_P - \frac{\partial f}{\partial u_i}\bigg|_Q\right) - 2N\frac{\partial^r \tilde{G}(x,t)}{\partial x^r}\right]\varphi^{(i)}(t)\,dt$$
$$(r = 0, \dots, n-1; \quad 0 \leqq x \leqq 1).$$

Let us rewrite inequality (1.1) in the form

(1.7)
$$\sum_{i=0}^{n-1} \max_{0\leqq x\leqq 1} \int_0^1 -2N\frac{\partial^i \tilde{G}(x,t)}{\partial x^i}\,dt = 2\tilde{\tilde{\theta}} = \theta^* < 1.$$

Let us denote by $\theta^*\lambda_i$ the i-th term in this sum. We have

(1.8)
$$\lambda_0, \dots, \lambda_{n-1} > 0; \quad \sum_{i=0}^{n-1} \lambda_i = 1$$

(cf. the construction of $\tilde{G}$ in [3]). It is easy to verify that the function

$$\varphi(x) = \begin{cases} \delta \sum_{i=0}^{n-1} \frac{\lambda_i}{i!} x^i & (0 \leqq x \leqq a), \\ \delta \sum_{i=0}^{n-1} \frac{\lambda_i}{i!} (x-a_j)^i & (a_j = j\cdot a \leqq x \leqq a_{j+1} = (j+1)a), \end{cases}$$

where $\delta > 0$ is an arbitrary constant, $a_0 = 0$; $j = 0, \dots, k$; $(k+1)a = 1$, for sufficiently small $a > 0$ satisfies inequality (1.6). This follows from (1.7), (1.8) and from the fact that for small $a > 0$ the ratio of the maximum and minimum of $\varphi^{(i)}$ is close to the unit for every $i = 0, \dots, n-1$.

Remark 1.1. Since the change to a finer division of the segment [0, 1] (cf. [3]) in the construction of $\tilde{G}$ and $\tilde{\mathcal{M}}$ does not bother those properties of $\tilde{G}$ and $\tilde{\mathcal{M}}$ that are needed by us, the initial division

$$x_0 = 0 < x_1 < \dots < x_m = 1$$

of the segment [0, 1] can be chosen to be equal to

$$x_0 = 0, \quad x_1 = a, \ldots, x_m = (k+1)a = 1 \quad (k+1 = m),$$

where a denotes the number in the proof of Theorem 1.1.

Remark 1.2. To our regret we did not succeed in proving the existence of the first pair (z_1, w_1) from $\mathcal{M}\times\mathcal{M}$ in an analogous way as in Theorem 1.1, since the ratios of the maximums and minimums corresponding to the function $\varphi\in\mathcal{M}$ cannot be made arbitrarily close to the unit.

Remark 1.3. We note that instead of the accurate solution of system (1.3) one may take its approximate solution $(z^{(1)}, w^{(1)})$ obtained, for example, by the method of successive approximation starting with an arbitary pair $(z, w)\in\tilde{\mathcal{M}}\times\tilde{\mathcal{M}}$. Hence from the practical point of view Theorem 1.1. may be useful. An analogous statement concerns all the following theorems of this paper.

II.

In connection with the negative statement in Remark 1.2 in this section we shall elaborate another method of the construction of the pair (z_1, w_1) from $\mathcal{M}\times\mathcal{M}$. This method relies on the usage of defect functions (cf. [1], [2], [3]). Actually, we shall consider only the monotone case of problem (0.1), i.e., when

$$\frac{\partial^i G(x,t)}{\partial x^i} \leqq 0 \quad (i = 0, \ldots, n-1;\ 0 \leqq x, t \leqq 1)$$

besides, we suppose that at all points

$$\frac{\partial f}{\partial u_i} \geqq 0 \quad (i = 0, \ldots, n-1).$$

The case when all partial derivatives of f are non-positive is entirely obvious (cf. [2]).

In this case the approximants z_p, w_p may be constructed independently of each other according to the rule

$$\text{(A)} \qquad z_{p+1} = Az_p, \quad w_{p+1} = Aw_p \quad (p = 1, 2, \ldots).$$

Theorem 2.1. *Under the weak assumption of contractivity* (0.4) *there exist elements* $z_1, w_1\in\mathcal{M}$ *for which in the sense of the ordering of* $\mathcal{M}$ *we have*

$$\text{(2.1)} \qquad z_2 \leqq z_1, \quad z_3 \leqq z_1; \quad w_1 \leqq w_2, \quad w_1 \leqq w_3.$$

Proof. We remark that the rule (*A*) is equivalent to the rule

$$\text{(A}'\text{)}\qquad \begin{cases} z_{p+1}(x) = z_p(x) - \sigma_p(x), \quad w_{p+1}(x) = w_p(x) - \eta_p(x), \\ \sigma_p(x) = \int\limits_0^1 G(x,t)\alpha_p(t)\,dt, \quad \eta_p(x) = \int\limits_0^1 G(x,t)\beta_p(t)\,dt, \\ \alpha_p(x) = z_p^{(n)}(x) - f[z_p], \quad \beta_p(x) = w_p^{(n)}(x) - f[w_p] \\ (0 \leqq x \leqq 1;\ p = 1, 2, \ldots). \end{cases}$$

Consequently, we shall seek $z_1, w_1 \in \mathcal{M}$ for which

(2.2)
$$\alpha_1(x) < 0, \quad \alpha_1(x) + \alpha_2(x) \leqq 0; \quad \beta_1(x) > 0, \quad \beta_1(x) + \beta_2(x) \geqq 0 \quad (0 \leqq x \leqq 1).$$

To attain the negativeness of α_1 it is enough to take the solution $z_1 \in \mathcal{M}$ of the equation

$$z^{(n)}(x) - f[z(x)] = -c = \alpha_1.$$

This equation is solvable in $\mathcal{M}$ by virtue of the contractivity assumption (0.4). For this function z_1 the second inequality in (2.2) is also satisfied as because of the rule (A′) we have

$$\alpha_1(x) + \alpha_2(x) = z_1^{(n)}(x) - f[z_1] + z_1^{(n)}(x) - \sigma_1^{(n)}(x) - f[z_2].$$

Thus by using Lagrange's formula for $f[z_1] - f[z_2]$ we arrive at the inequality

$$\alpha_1(x) + \alpha_2(x) = \alpha_1(x) + \sum_{i=0}^{n-1} \left.\frac{\partial f}{\partial u_i}\right| \sigma_1^{(i)}(x) \leqq 0.$$

This inequality is equivalent to the inequality

$$1 \geqq \sum_{i=0}^{n-1} \left|\frac{\partial f}{\partial u_i}\right| \int\limits_0^1 \left|\frac{\partial^i G(x,t)}{\partial x^i}\right| dt$$

the validity of which follows from the contractivity condition (0.4).

The existence of w_1 may be proved in a similar way.

Remark 2.1. In the case when

$$\frac{\partial^i G(x,t)}{\partial x^i} \geqq 0, \quad \frac{\partial f}{\partial u_i} \leqq 0 \quad (i = 0, \ldots, n-1;\ 0 \leqq x, t \leqq 1)$$

we may also prove the existence of $z_1, w_1 \in \mathcal{M}$.

III.

Under the strong assumption of contractivity (i) let us now consider the monotone case of problem (0.1), i.e.,

$$\frac{\partial^i G(x,t)}{\partial x^i} \leqq 0 \quad (0 \leqq x, t \leqq 1;\ i = 0, \ldots, n-1)$$

however, the derivatives of f are not all of the same sign, and themselves may change sign, too. The approximants z_p, w_p shall now be constructed in $\mathcal{M}$ according to the rule (0), i.e., we have

$$\text{(B)} \quad \begin{cases} z_{p+1} = E(z_p, w_p) = \dfrac{1}{2}\displaystyle\int_0^1 G(x,t)\big(f[z_p(t)] + f[w_p(t)]\big)dt + \\ \qquad + \dfrac{N}{2}\displaystyle\int_0^1 G(x,t)\sum_{i=0}^{n-1}\big(z_p^{(i)}(t) - w_p^{(i)}(t)\big)dt, \\ w_{p+1} = F(z_p, w_p) = \dfrac{1}{2}\displaystyle\int_0^1 G(x,t)\big(f[z_p(t)] + f[w_p(t)]\big)dt - \\ \qquad - \dfrac{N}{2}\displaystyle\int_0^1 G(x,t)\sum_{i=0}^{n-1}\big(z_p^{(i)}(t) - w_p^{(i)}(t)\big)dt. \end{cases}$$

Theorem 3.1. *Under the assumption of* (i) *in the case considered there exist elements* $z_1, w_1 \in \mathcal{M}$ *for which*

$$\text{(3.1)} \qquad z_2 \leqq z_1, \quad z_3 \leqq z_1; \quad w_1 \leqq w_2, \quad w_1 \leqq w_3$$

in the sense of the ordering of $\mathcal{M}$.

Proof. We note that rule (B) is equivalent to the rule

$$\text{(B')} \quad \begin{cases} z_{p+1}(x) = z_p(x) - \sigma_p(x), \quad w_{p+1}(x) = w_p(x) - \eta_p(x), \\ \sigma_p(x) = \displaystyle\int_0^1 G(x,t)\alpha_p(t)dt, \quad \eta_p(x) = \displaystyle\int_0^1 G(x,t)\beta_p(t)dt, \\ \alpha_p(x) = z_p^{(n)}(x) - \dfrac{1}{2}f[z_p] - \dfrac{1}{2}f[w_p] - \dfrac{N}{2}\displaystyle\sum_{i=0}^{n-1}\big(z_p^{(i)}(x) - w_p^{(i)}(x)\big), \\ \beta_p(x) = w_p^{(n)}(x) - \dfrac{1}{2}f[z_p] - \dfrac{1}{2}f[w_p] + \dfrac{N}{2}\displaystyle\sum_{i=0}^{n-1}\big(z_p^{(i)}(x) - w_p^{(i)}(x)\big) \\ (0 \leqq x \leqq 1;\ p = 1, 2, \ldots). \end{cases}$$

On the other hand, the inequalities (3.1) are equivalent to the inequalities

$$\text{(3.2)} \quad \begin{cases} \alpha_1(x) \leqq 0, \quad \alpha_1(x) + \alpha_2(x) \leqq 0; \quad \beta_1(x) \geqq 0, \quad \beta_1(x) + \beta_2(x) \geqq 0 \\ (0 \leqq x \leqq 1). \end{cases}$$

The solution $(z_1, w_1)\in\mathcal{M}\times\mathcal{M}$ of the system

$$\tag{3.3} \begin{cases} z^{(n)}(x)-\dfrac{1}{2}f[z]-\dfrac{1}{2}f[w]-\dfrac{N}{2}\sum\limits_{i=0}^{n-1}\left(z^{(i)}(x)-w^{(i)}(x)\right)=-c,\\ w^{(n)}(x)-\dfrac{1}{2}f[z]-\dfrac{1}{2}f[w]+\dfrac{N}{2}\sum\limits_{i=0}^{n-1}\left(z^{(i)}(x)-w^{(i)}(x)\right)=c, \end{cases}$$

where $c>0$ is a constant, exists by virtue of (i) and obviously satisfies the first and the third inequalities under (3.2). Moreover, z_1, w_1 also satisfies the second and the fourth inequalities in (3.2). Indeed, by the rule (B′) and by using Lagrange's formula for the difference

$$f[z_1]-f[z_1-\sigma_1],\quad f[w_1]-f[w_1-\eta_1]$$

one can show that, for example, the second inequality in (3.2) is equivalent to the inequality

$$\alpha_1(x)=-c\leqq\frac{1}{2}\sum_{i=0}^{n-1}\left\{\sigma_1^{(i)}(x)\left[\left.\frac{\partial f}{\partial u_i}\right|+N\right]+\eta_1^{(i)}(x)\left[\left.\frac{\partial f}{\partial u_i}\right|-N\right]\right\}.$$

If we replace $\sigma_1^{(i)}$ and $\eta_1^{(i)}$ by their expressions via Green's function and α_1, β_1, we arrive at the inequality

$$\tag{3.4} 1\geqq\sum_{i=0}^{n-1}\frac{1}{2}\left[\left(\left.\frac{\partial f}{\partial u_i}\right|_P-\left.\frac{\partial f}{\partial u_i}\right|_Q\right)+2N\right]\int_0^1\left|\frac{\partial^i G(x,t)}{\partial x^i}\right|dt.$$

The validity of (3.4) follows from condition (i) and from the condition

$$\left|\frac{\partial f}{\partial u_i}\right|\leqq N\quad(i=0,\dots,n-1).$$

The fourth inequality under (3.2) may be proved in an analogous way.

Remark 3.1. The case

$$\frac{\partial^i G(x,t)}{\partial x^i}\geqq 0\quad(i=0,\dots,n-1;\quad 0\leqq x,t\leqq 1)$$

may be handled analogously.

IV.

Now we shall give still another method, concerning the general G and rule (0), for the construction of functions from $\tilde{\mathcal{M}}$ that satisfy the inequalities (1.2). This method shall be very useful later.

Let us take two arbitrary functions $(z_1, w_1)\in\tilde{\mathcal{M}}\times\tilde{\mathcal{M}}$. Let the operators E and F map (z_1, w_1) into (z_2, w_2). We shall seek the elements $Z_1, W_1\in\tilde{\mathcal{M}}$ that satisfy the inequalities (1.2) in the form

(4.1) $$Z_1 = z_1+\varphi_1, \quad W_1 = w_1+\eta_1.$$

In this case

$$Z_2(x) =$$
$$= \frac{1}{2}\int_0^1 \Big\{G(x,t)\,(f[z_1+\varphi_1]+f[w_1+\eta_1])+\tilde{G}(x,t)\,N\sum_{i=0}^{n-1}((z_1+\varphi_1)^{(i)}-(w_1+\eta_1)^{(i)})\Big\}\,dt,$$

$$W_2(x) =$$
$$= \frac{1}{2}\int_0^1 \Big\{G(x,t)\,(f[z_1+\varphi_1]+f[w_1+\eta_1])-\tilde{G}(x,t)\,N\sum_{i=0}^{n-1}((z_1+\varphi_1)^{(i)}-(w_1+\eta_1)^{(i)})\Big\}\,dt.$$

By Lagrange's formula we have

$$f[z_1+\varphi_1]=f[z_1]+\sum_{i=0}^{n-1}\varphi_1^{(i)}\frac{\partial f}{\partial u_i}\Big|_P, \quad f[w_1+\eta_1]=f[w_1]+\sum_{i=0}^{n-1}\eta_1^{(i)}\frac{\partial f}{\partial u_i}\Big|_Q.$$

Consequently, we have to satisfy the following inequalities in $\tilde{\mathcal{M}}$:

(4.2)
$$\begin{cases} \dfrac{1}{2}\displaystyle\int_0^1\sum_{i=0}^{n-1}\Big\{\varphi_1^{(i)}(t)\Big[G(x,t)\frac{\partial f}{\partial u_i}\Big|_P+\tilde{G}(x,t)N\Big]+\eta_1^{(i)}(t)\Big[G(x,t)\frac{\partial f}{\partial u_i}\Big|_Q-\tilde{G}(x,t)N\Big]\Big\}dt \leqq \\ \qquad\qquad \leqq \varphi_1(x)+z_1(x)-z_2(x), \\ \dfrac{1}{2}\displaystyle\int_0^1\sum_{i=0}^{n-1}\Big\{\varphi_1^{(i)}(t)\Big[G(x,t)\frac{\partial f}{\partial u_i}\Big|_P-\tilde{G}(x,t)N\Big]+\eta_1^{(i)}(t)\Big[G(x,t)\frac{\partial f}{\partial u_i}\Big|_Q+\tilde{G}(x,t)N\Big]\Big\}dt \geqq \\ \qquad\qquad \geqq \eta_1(x)+w_1(x)-w_2(x). \end{cases}$$

For the given z_1, z_2, w_1, w_2 we obviously can choose φ_1, η_1 from $\tilde{\mathcal{M}}$ such that the inequalities

(4.3)
$$\left.\begin{cases} \varphi_1(x) \geqq |z_1(x)-z_2(x)|, \dots, \varphi_1^{(n-1)}(x) \geqq |z_1^{(n-1)}(x)-z_2^{(n-1)}(x)|, \\ \eta_1(x) \leqq -|w_1(x)-w_2(x)|, \dots, \eta_1^{(n-1)}(x) \leqq -|w_1^{(n-1)}(x)-w_2^{(n-1)}(x)| \end{cases}\right\} \quad (0 \leqq x \leqq 1)$$

are satisfied. Hence the left and right sides (and then their derivatives, too) of the inequalities (4.2) either coincide or have different signs.

Thus we have proved the following lemma.

Lemma 4.1. *For any elements z_1, w_1 of $\tilde{\mathcal{M}}$ the functions*

$$Z_1(x) = z_1(x)+\varphi_1(x), \quad W_1(x) = w_1(x)+\eta_1(x)$$

together with the auxiliary functions φ_1, η_1 *satisfying* (4.3) *also satisfy the inequalities* (1.2):

$$Z_2 = E(Z_1, W_1) \leqq Z_1, \quad W_1 \leqq W_2 = F(Z_1, W_1).$$

Consequence 4.1. Because of what we said in the concluding part of the introduction, Lemma 4.1 gives a method of construction for $z_1, w_1; \varphi_1, \eta_1$ $(\in \mathcal{M})$ such that with the functions $Z_1 = z_1 + \varphi_1$, $W_1 = w_1 + \eta_1$ we have the inequalities:

$$Z_2 = E(Z_1, W_1) \leqq Z_1, \quad W_1 \leqq W_2 = F(Z_1, W_1).$$

In this case for the given z_1, w_1 we have to choose $\varphi_1 > 0$, $\eta_1 < 0$ from $\mathcal{M}$ in such a way that

(4.3′) $$\varphi_1 \geqq z_1 - z_2, z_2 - z_1; \quad \eta_1 \leqq w_1 - w_2, w_2 - w_1$$

is satisfied in the sense of the ordering of $\mathcal{M}$. Let us take two numbers $K<0$, $L>0$ for which

$$-K \geqq \max_{0 \leqq x \leqq 1} |z_1^{(n)}(x) - x_2^{(n)}(x)|, \quad L \geqq \max_{0 \leqq x \leqq 1} |w_1^{(n)}(x) - w_2^{(n)}(x)|.$$

Then the functions

$$\varphi_1(x) = \int_0^1 KG(x,t)\,dt, \quad \eta_1(x) = \int_0^1 LG(x,t)\,dt$$

satisfy (4.3′).

Now we shall prove the existence of the first approximating pair $(z_1, w_1) \in \mathcal{M} \times \mathcal{M}$ in the monotone case of problem (0.1) under the weak assumption of contractivity (0.4), i.e., when

$$N \sum_{i=0}^{n-1} \max_{0 \leqq x \leqq 1} \int_0^1 \left| \frac{\partial^i G(x,t)}{\partial x^i} \right| dt = \theta < 1,$$

where $N>0$ majorizes the modules of the partial derivatives of f.

Let us take an arbitrary negative number d and consider the problem

(4.4) $$\varphi_1^{(n)}(x) - N \sum_{i=0}^{n-1} \varphi_1^{(i)}(x) = d, \quad \varphi_1 \in \mathcal{M}.$$

Lemma 4.2. *The solution of problem* (4.4) *exists and is unique. It satisfies the inequality*

(4.5) $$\varphi_1^{(n)}(x) \leqq d(1-\theta) \quad (0 \leqq x \leqq 1).$$

Proof. The existence and uniqueness of the solution follows from the contractivity condition (0.4). For the proof of (4.5) let us consider the problem

(4.6) $$u^{(n)}(x) = N \sum_{i=0}^{n-1} u^{(i)}(x), \quad u \in \mathcal{M}.$$

This problem has the only trivial solution $u=0$ in virtue of (0.4). With respect to the equation (4.6) the solution φ_1 of problem (4.4) and the corresponding φ_2 have defect functions (cf. for example [1], section 4.)

$$\tilde{\alpha}_1(x) = d, \quad \tilde{\alpha}_2(x) = dN \sum_{i=0}^{n-1} \int_0^1 \frac{\partial^i G(x,t)}{\partial x^i}\, dt.$$

Consequently, (0.4) and Theorem 4.1 of [1] imply that

$$\left.\begin{aligned} \varphi_1^{(n)}(x) - \varphi_3^{(n)}(x) = \tilde{\alpha}_1(x) + \tilde{\alpha}_2(x) \leqq d(1-\theta), \\ \varphi_1^{(n)}(x) \leqq \varphi_3^{(n)}(x) \leqq 0 \end{aligned}\right\} \quad (0 \leqq x \leqq 1).$$

Hence we have

$$\varphi_1^{(n)}(x) \leqq d(1-\theta) \quad (0 \leqq x \leqq 1).$$

Theorem 4.1. *In the monotone case of problem* (0.1) *under the contractivity assumption* (0.4) *there exist elements* $Z_1, W_1 \in \mathcal{M}$ *such that*

$$Z_2 = E(Z_1, W_1) \leqq Z_1, \quad W_1 \leqq W_2 = F(Z_1, W_1), \tag{4.7}$$

$$Z_3 = E(Z_2, W_2) \leqq Z_1, \quad W_1 \leqq W_3 = F(Z_2, W_2). \tag{4.8}$$

Proof. Let us take an arbitrary positive number $\varepsilon > 0$ and two functions z_1, w_1 from $\mathcal{M}$ so close to the solution y of problem (0.1) (z_1 and w_1 can be constructed by means of successive approximation) as to satisfy the inequalities

$$|z_1^{(n)}(x) - z_2^{(n)}(x)| < \varepsilon, \quad |w_1^{(n)}(x) - w_2^{(n)}(x)| < \varepsilon \quad (0 \leqq x \leqq 1).$$

Let us now seek Z_1, W_1 as in Lemma 4.1. in the form

$$Z_1 = z_1 + \varphi_1, \quad W_1 = w_1 + \eta_1$$

with unknown functions $\varphi_1, -\eta_1 \geqq 0$ from $\mathcal{M}$ (in the sense of the ordering of $\mathcal{M}$). In order that (4.7) be satisfied, φ_1 and η_1 have to satisfy (4.3′) and for such φ_1, η_1 the fulfilment of (4.8) is equivalent to the assertion that the positive functions

$$\hat{\varphi} = Z_1 - Z_2, \quad \hat{\eta} = W_2 - W_1$$

occurring in Lemma 4.1 satisfy the inequalities (1.6). These inequalities (in this case $\tilde{G} = G$, $\tilde{\mathcal{M}} = \mathcal{M}$) are equivalent to the inequalities

$$\begin{aligned} \hat{\alpha}_1(x) \leqq \frac{1}{2} \sum_{i=0}^{n-1} \Bigg[\left(\frac{\partial f}{\partial u_i}\bigg|_P + N \right) \int_0^1 -\frac{\partial^i G(x,t)}{\partial x^i}\, \hat{\alpha}_1(t)\, dt + \\ + \left(\frac{\partial f}{\partial u_i}\bigg|_Q - N \right) \int_0^1 -\frac{\partial^i G(x,t)}{\partial x^i}\, \hat{\beta}_1(t)\, dt \Bigg], \\ \hat{\beta}_1(x) \geqq \frac{1}{2} \sum_{i=0}^{n-1} \Bigg[\left(\frac{\partial f}{\partial u_i}\bigg|_P - N \right) \int_0^1 -\frac{\partial^i G(x,t)}{\partial x^i}\, \hat{\alpha}_1(t)\, dt + \\ + \left(\frac{\partial f}{\partial u_i}\bigg|_Q + N \right) \int_0^1 -\frac{\partial^i G(x,t)}{\partial x^i}\, \hat{\beta}_1(t)\, dt \Bigg], \end{aligned} \tag{4.9}$$

where we have

$$\hat{\alpha}_1 = Z_1^{(n)} - Z_2^{(n)} \leqq 0, \quad \hat{\beta}_1 = W_1^{(n)} - W_2^{(n)} \geqq 0$$

as because of (4.3′) $Z_1 \geqq Z_2$ and $W_1 \leqq W_2$.

The derivatives $\left.\frac{\partial f}{\partial u_i}\right|_P, \left.\frac{\partial f}{\partial u_i}\right|_Q$ in (4.9) are taken from the formulas

$$f[z_1] - f[z_2] = \sum_{i=0}^{n-1} \left.\frac{\partial f}{\partial u_i}\right|_P (z_1 - z_2)^{(i)},$$

$$f[w_1] - f[w_2] = \sum_{i=0}^{n-1} \left.\frac{\partial f}{\partial u_i}\right|_Q (w_1 - w_2)^{(i)}.$$

By using the formulas of Lemma 4.1 and taking $\varphi_1 = -\eta_1 > 0$ for the sake of simplicity we obtain

(4.10)

$$\hat{\alpha}_1(x) = z_1^{(n)}(x) - z_2^{(n)}(x) + \left(\varphi_1^{(n)}(x) - N \sum_{i=0}^{n-1} \varphi_1^{(i)}(x)\right) - \frac{1}{2} \sum_{i=0}^{n-1} \varphi_1^{(i)}(x) \left(\left.\frac{\partial f}{\partial u_i}\right|_P - \left.\frac{\partial f}{\partial u_i}\right|_Q\right),$$

$$\hat{\beta}_1(x) = w_1^{(n)}(x) - w_2^{(n)}(x) - \left(\varphi_1^{(n)}(x) - N \sum_{i=0}^{n-1} \varphi_1^{(i)}(x)\right) - \frac{1}{2} \sum_{i=0}^{n-1} \varphi_1^{(i)}(x) \left(\left.\frac{\partial f}{\partial u_i}\right|_P - \left.\frac{\partial f}{\partial u_i}\right|_Q\right).$$

In Theorem 3.1 everything went well essentially because we succeeded in finding z_1, w_1 such that the defect functions α_1, β_1 turned out to be constants ($-c$ resp. c), hence the necessary properties of z_1, w_1 immediately followed from the contractivity assumption.

The circumstances are similar now. For the $\varepsilon > 0$, z_1, w_1 already chosen we choose a number $d < 0$ in such a way that with the solution φ_1 of problem (4.4) the corresponding inequalities (4.3′) be satisfied. To achieve this it is enough (cf. Consequence 4.1) to take $d = -\sqrt{\varepsilon}(1-\theta)^{-1}$ (in the sequel we suppose that $0 < \varepsilon < 1$) as because of (4.5) the inequality

$$\varphi_1^{(n)}(x) \leqq d(1-\theta) \leqq -\varepsilon < -\max_{0 \leqq x \leqq 1} |z_1^{(n)}(x) - z_2^{(n)}(x)|$$

implies inequality (4.3′) for φ_1. One can prove analogously (taking $\eta_1 = -\varphi_1$ for the sake of simplicity) that for $\eta_1 = -\varphi_1$ the corresponding inequalities (4.3′) are also satisfied. The inequalities (4.7) are also satisfied for any $0 < \varepsilon < 1$.

In contrast with Section III now we may use the continuity of the partial derivatives of f, i.e., that

$$h(\varepsilon) = \max_{i,x} \left| \left.\frac{\partial f}{\partial u_i}\right|_{P(x)} - \left.\frac{\partial f}{\partial u_i}\right|_{Q(x)} \right| \to 0 \quad (\varepsilon \to 0).$$

From equality (4.4) it is easy to obtain that

$$\sum_{i=0}^{n-1} \varphi_1^{(i)}(x) \leqq \frac{-d\theta}{N} \quad (0 \leqq x \leqq 1),$$

whence in (4.10) we get that

$$\max_{0\leqq x\leqq 1}\hat{\alpha}_1(x) \geqq -\varepsilon-\frac{\sqrt{\varepsilon}}{1-\theta}-\frac{1}{2}|h(\varepsilon)|\frac{\theta\sqrt{\varepsilon}}{N(1-\theta)} = -\sqrt{\varepsilon}\left(\frac{1}{1-\theta}+\circ(\varepsilon)\right)-\varepsilon,$$

$$\max_{0\leqq x\leqq 1}\hat{\alpha}_1(x) \leqq -\sqrt{\varepsilon}\left(\frac{1}{1-\theta}+\circ(\varepsilon)\right)+\varepsilon$$

and analogously

$$\sqrt{\varepsilon}\left(\frac{1}{1-\theta}+\circ(\varepsilon)\right)-\varepsilon \leqq \hat{\beta}_1 \leqq \sqrt{\varepsilon}\left(\frac{1}{1-\theta}+\circ(\varepsilon)\right)+\varepsilon.$$

i.e., although $\hat{\alpha}_1, \hat{\beta}_1$ are not constant, the ratio of the maximum and minimum of each of $\hat{\alpha}_1, \hat{\beta}_1$ tends to $+1$ as $\varepsilon\to 0$ and the ratio of the maxima of $\hat{\alpha}_1$ and $\hat{\beta}_1$ tends to -1 as $\varepsilon\to 0$. Consequently, there exists $\varepsilon>0$ for which (4.9) is satisfied. To complete the proof one has to use condition (0.4), and the continuity of the partial derivatives of f.

V.

Finally let us consider the general G under the assumption of (0.4). The space $\tilde{\mathcal{M}}$, the functions $\tilde{G}, \tilde{\varphi}$ will be assumed to correspond to a sufficiently fine division of [0,1]. Let us take the function $\tilde{\varphi}=\varphi$ from Section I with undetermined $\delta>0$. Besides, in this case

$$\sum_{i=0}^{n-1}\max_{0\leqq x\leqq 1}\int_0^1 -N\frac{\partial^i\tilde{G}(x,t)}{\partial x^i}\,dt \equiv \sum_{i=0}^{n-1}\lambda_i\theta_1 = \theta_1 < 1.$$

Let us consider the problem

$$\text{(5.1)}\quad \varphi_1(x)-A_0\varphi_1 \equiv \varphi_1(x)-N\int_0^1\tilde{G}(x,t)\left(\sum_{i=0}^{n-1}\varphi_1^{(i)}(t)\right)dt = \tilde{\varphi}(x),\quad \varphi_1\in\tilde{\mathcal{M}}.$$

Lemma 5.1. *The solution φ_1 of problem* (5.1) *exists, is unique, and satisfies the inequality (in the sense of the ordering of $\tilde{\mathcal{M}}$)*

$$\text{(5.2)}\qquad \varphi_1 \geqq c\tilde{\varphi}$$

with a general constant $c>0$ for all fine enough divisions of the segment [0, 1].

Proof. The existence and uniqueness of the solution φ_1 of problem (5.1) follow from (0.4). To prove (5.2) let us consider the problem

$$\text{(5.3)}\qquad u(x)-A_0u = 0,\quad u\in\tilde{\mathcal{M}},$$

which in virtue of (0.4) has the only solution $u=0$. With respect to equation (5.3) for φ_1 we have the following inequality in $\tilde{\mathcal{M}}$:

$$\text{(5.4)}\qquad \varphi_1 = \tilde{\varphi}+A_0\varphi_1 \geqq A_0\varphi_1.$$

Moreover,

$$\varphi_1-A_0^2\varphi_1=\tilde{\varphi}+N\int_0^1\tilde{G}(x,t)\sum_{i=0}^{n-1}\tilde{\varphi}^{(i)}(t)\,dt,$$

from which, as $\tilde{\varphi}\geqq 0$ (in the sense of the ordering of $\tilde{\mathcal{M}}$), we obtain for $0\leqq x\leqq 1$, $r=0,\ldots,n-1$ that

$$\varphi_1^{(r)}(x)-(A_0^2\varphi_1)^{(r)}\geqq\tilde{\varphi}^{(r)}(x)-\left\{N\max_{0\leqq x\leqq 1}\int_0^1\left|\frac{\partial^r\tilde{G}(x,t)}{\partial x^r}\right|dt\right\}\max_{0\leqq\xi\leqq 1}\sum_{i=0}^{n-1}\tilde{\varphi}^{(i)}(\xi).$$

By taking a fine enough division of [0, 1] from this we obtain that

$$\varphi_1^{(r)}(x)-(A_0^2\varphi_1)^{(r)}\geqq\delta\lambda_r(1-\theta_2)\quad(0<\theta_2<1). \tag{5.5}$$

This, together with Theorem 2 of [2], implies that

$$0\leqq A_0^2\varphi_1\leqq\varphi_1;\quad\varphi_1\geqq(1-\theta_3)\tilde{\varphi}\quad(\theta_2<\theta_3<1). \tag{5.6}$$

Here θ_2 and θ_3 are absolute constants for all fine enough divisions of the segment [0, 1]. The lemma is proved.

Theorem 5.1. *For problem* (0.1) *under condition* (0.4) *there exist* $Z_1, W_1\in\tilde{\mathcal{M}}$ *for which the inequalities*

$$Z_2=E(Z_1,W_1)\leqq Z_1,\quad W_1\leqq W_2=F(Z_1,W_1), \tag{5.7}$$

$$Z_3=E(Z_2,W_2)\leqq Z_1,\quad W_1\leqq W_3=F(Z_2,W_2) \tag{5.8}$$

are satisfied in $\tilde{\mathcal{M}}$.

Proof. We are going to seek Z_1, W_1 in the form

$$Z_1=z_1+\varphi_1,\quad W_1=w_1+\eta_1$$

(cf. Lemma 4.1). Given an arbitrary positive number $\varepsilon>0$, let us choose $z_1, w_1\in\tilde{\mathcal{M}}$ so close to the solution y of problem (0.1) as to satisfy the inequalities

$$d_z=\max_{i,x}|z_1^{(i)}(x)-z_2^{(i)}(x)|<\varepsilon,\quad d_w=\max_{i,x}|w_1^{(i)}(x)-w_2^{(i)}(x)|<\varepsilon$$

with $z_2=E(z_1,w_1)$, $w_2=F(z_1,w_1)$. The maxima are taken over $i=0,\ldots,n-1$ and $0\leqq x\leqq 1$.

Let us now take the solution φ_1 of problem (5.1) with such a $\tilde{\varphi}$ where $\delta=\sqrt{\varepsilon}$. Then for small enough $\varepsilon>0$ we have, in virtue of (5.2), that

$$\varphi_1\geqq c\tilde{\varphi}\geqq\varepsilon\geqq d_z,d_w,$$

i.e., with this φ_1 and with $\eta_1=-\varphi_1$ on account of Lemma 4.1 we obtain that (5.7) is satisfied, i.e.,

$$\varphi^*=Z_1-Z_2\geqq 0,\quad\eta^*=W_2-W_1\geqq 0.$$

Owing to the formulas of Lemma 4.1 for $\eta_1 = -\varphi_1$ we obtain that

$$
(5.9)\quad
\begin{aligned}
\varphi^*(x) &= z_1(x) - z_2(x) + \tilde{\varphi}(x) - \frac{1}{2}\int_0^1 G(x,t)\left[\sum_{i=0}^{n-1} \varphi_1^{(i)}(t)\left(\frac{\partial f}{\partial u_i}\bigg|_P - \frac{\partial f}{\partial u_i}\bigg|_Q\right)\right] dt,\\
\eta^*(x) &= w_2(x) - w_1(x) + \tilde{\varphi}(x) + \frac{1}{2}\int_0^1 G(x,t)\left[\sum_{i=0}^{n-1} \varphi_1^{(i)}(t)\left(\frac{\partial f}{\partial u_i}\bigg|_P - \frac{\partial f}{\partial u_i}\bigg|_Q\right)\right] dt,
\end{aligned}
$$

where the partial derivatives of f are taken from the formulas

$$
f[z_1] - f[z_2] = \sum_{i=0}^{n-1} \frac{\partial f}{\partial u_i}\bigg|_P \left(z_1^{(i)}(x) - z_2^{(i)}(x)\right),
$$

$$
f[w_1] - f[w_2] = \sum_{i=0}^{n-1} \frac{\partial f}{\partial u_i}\bigg|_Q \left(w_1^{(i)}(x) - w_2^{(i)}(x)\right).
$$

Consequently, as $\varepsilon \to 0$ we get that

$$
h(\varepsilon) = \max_{\substack{i=0,\dots,n-1\\ 0 \leqq x \leqq 1}} \left| \frac{\partial f}{\partial u_i}\bigg|_{P(x)} - \frac{\partial f}{\partial u_i}\bigg|_{Q(x)} \right| \to 0.
$$

Expressing (5.8) by $\varphi^* = Z_1 - Z_2, -\eta^* = W_1 - W_2$ as well as (1.4) by $\varphi = z_1 - z_2, -\varphi = w_1 - w_2$ in Theorem 1.1, we get inequalities analogous to (1.6). The latter inequalities are surely satisfied if

$$
(5.10)\quad
\begin{cases}
1 \geqq \dfrac{-N}{2}\displaystyle\int_0^1 \frac{\partial^r \tilde{G}(x,t)}{\partial x^r} \sum_{i=0}^{n-1}\left(\frac{\max \varphi^{*(i)}}{\min \varphi^{*(r)}} + \frac{\max \eta^{*(i)}}{\min \varphi^{*(r)}}\right) dt + \\
+\dfrac{1}{2}\displaystyle\int_0^1 \left|\frac{\partial^r G(x,t)}{\partial x^r}\right| \sum_{i=0}^{n-1} \max_{0 \leqq \xi \leqq 1} \left| -\frac{\partial f}{\partial u_i}\bigg|_P \frac{\varphi^{*(i)}(\xi)}{\min \varphi^{*(r)}} + \frac{\partial f}{\partial u_i}\bigg|_Q \frac{\eta^{*(i)}(\xi)}{\min \varphi^{*(r)}} \right| dt \\
(r = 0, \dots, n-1;\ 0 \leqq x \leqq 1)
\end{cases}
$$

and the same inequality with the rôles of φ^* and η^* interchanged, are satisfied.

From (5.2) it is easy to derive that

$$
\frac{1}{2}\int_0^1 \left|\frac{\partial^r \tilde{G}(x,t)}{\partial x^r}\right| \sum_{i=0}^{n-1} \varphi_1^{(i)}(t)\,dt \leqq \frac{1-c}{2N}\tilde{\varphi}^{(r)}(x) \quad (r = 0, \dots, n-1;\ 0 \leqq x \leqq 1).
$$

Consequently, as

$$
\max_{0 \leqq x \leqq 1} \tilde{\varphi}^{(i)}(x) \leqq \sqrt{\varepsilon}\left(\lambda_i + k(\varepsilon)\right) \quad \left(i = 0, \dots, n-1;\ k(\varepsilon) \to 0(\varepsilon \to 0)\right)
$$

we obtain that

$$
\max_{0 \leqq t \leqq 1} \varphi^{*(i)}(t) \leqq \varepsilon + \sqrt{\varepsilon}\left(\lambda_i + k(\varepsilon)\right) + \frac{1-c}{2N}\sqrt{\varepsilon}\left(\lambda_i + k(\varepsilon)\right) h(\varepsilon) =
$$

$$
= \sqrt{\varepsilon}\left(\lambda_i + a(\varepsilon)\right).
$$

Analogously,

$$\min_{0\le t\le 1} \varphi^{*(i)}(t) \geqq \sqrt{\varepsilon}\,(\lambda_i+b(\varepsilon)), \quad \max_{0\le t\le 1} \eta^{*(i)}(t) \leqq \sqrt{\varepsilon}\,(\lambda_i+c(\varepsilon)),$$

$$\min_{0\le t\le 1} \eta^{*(i)}(t) \geqq \sqrt{\varepsilon}(\lambda_i+d(\varepsilon)),$$

where

$$a(\varepsilon), b(\varepsilon), c(\varepsilon), d(\varepsilon) \to 0 \quad (\varepsilon \to 0).$$

From this it follows that the ratios in (5.10) tend to $\frac{\lambda_i}{\lambda_r}$ as $\varepsilon\to 0$. Hence the second integral under (5.10) tends to zero and the sum under the first integral to the number

$$\sum_{i=0}^{n-1} \frac{2\lambda_i}{\lambda_r} = \frac{2}{\lambda_r}.$$

Thus for $0\leqq x\leqq 1$ the right hand side of (5.10) tends to a limit not greater than $\theta_1<1$, as $\varepsilon\to 0$. Consequently, for small $\varepsilon>0$ the inequalities (5.8) are also satisfied. This finishes the proof of our theorem.

We remark that Theorem 4.1 guarantees the existence of the required pair $Z_1, W_1\in\mathcal{M}$ (cf. [2]) for the initial value problem

$$(5.11)\qquad \begin{cases} y^{(n)}(x) = f[y] \equiv f(x, y(x), \ldots, y^{(n-1)}(x)) & (0\leqq x\leqq 1;\ n\geqq 1), \\ y(0) = \ldots = y^{(n-1)}(0) = 0 \end{cases}$$

with a function f having the same properties as in problem (0.1) and under the assumption of contractivity (0.4). The general initial value problem on the segment $[a, b]$ can be reduced to problem (5.11).

Analogously as above one can prove the existence of the first pair of approximants for the solution of a boundary value problem that belongs to a wide class of partial differential equations or equations with delayed argument (cf., for example, the references part of [1]).

References

[1] Ю. И. Ковач, J. Hegedűs, Об одном двустороннем итерационном методе решения краевой задачи с запаздыванием, *Acta Sci. Math.*, **36** (1974), 69—89.

[2] J. Hegedűs, On a two-sided iterative method, *Colloquia Mathematica Societatis János Bolyai*, 15, *Differential Equations*, Keszthely (Hungary), 1975.

[3] J. Hegedűs, О двусторонних приближениях решений некоторых дифференциальных и операторных уравнений, *Publicationes Math., Debrecen*, to appear.

Acta Sci. Math., 39 (1977), 291—301

The radical in ring varieties

A. A. ISKANDER

All rings considered here are associative and not necessarily with 1. A ring variety is a class of rings closed under subrings, homomorphic images and Cartesian products; equivalently, it is the class of all rings satisfying a set of polynomial identities. In the present paper, we study syntactic and semantic properties and also structure of varieties in which the Jacobson radical of every member is 1) nil, 2) nilpotent, or 3) a direct summand. 1) is equivalent to a one variable identity: $x^n+x^{n+1}\cdot h(x)=0$; and also the variety is locally nilpotent by finite. 2) is equivalent to an n variable identity: $x_1...x_n+f(x_1, ..., x_n)=0$, where every term in $f(x_1, ..., x_n)$ is of degree larger than n. Also, every variety satisfying 2) is generated by a finitely generated ring and is finitely based. 3) is equivalent to a finite set of two variable identities.

For an account of the variety theory, the reader may consult [1, 2, 3, 5, 12, 13, 14]. Script letters will denote classes or varieties of rings; the corresponding Latin letters will denote their T-ideals of identities. We denote the free associative ring on $\{x_1, x_2, ..., x_n, ...\}$ by F. The join of varieties will be denoted by $\vee$. Var $\mathcal{K}$ will mean the variety generated by $\mathcal{K}$. $\langle f, g, h, ...\rangle$ will mean the variety of all rings satisfying the identities $0=f=g=h= ...$.

A residually finite ring is a ring in which every nonzero element does not belong to an ideal of finite index. A T-ideal is an ideal closed under all endomorphisms.

1. Definition 1. A ring variety is called *locally nilpotent by finite,* if every finitely generated member possesses a nilpotent ideal of finite index. A *locally finite variety* is a variety in which every finitely generated member is finite.

In [7], it is shown that locally finite ring varieties are precisely varieties satisfying $cx=0$, $x^n+x^{n+1}f(x)=0$, for some positive integers c and n and for some polynomial $f(x)$ with integral coefficients. For locally nilpotent by finite varieties, we have

Received February 12, 1976.

Theorem 1. *The following conditions on a ring variety $\mathcal{V}$ are equivalent:*

1) *$\mathcal{V}$ is locally nilpotent by finite.*
2) *every member of $\mathcal{V}$ generated by one element is nilpotent by finite.*
3) *$\mathcal{V}$ satisfies $x^n+x^{n+1}f(x)=0$, for some positive integer n and $f(x)\in\mathbf{Z}[x]$.*
4) *the Jacobson radical of every member of $\mathcal{V}$ is nil.*

2. We will first state and prove some lemmas.

Lemma 2. *Let A be an algebra over a field K. Then either A is nil, or A contains a copy of K or of $xK[x]$.*

Proof. Let $a\in A$, $a\neq 0$. The subalgebra of A generated by a is isomorphic to $xK[x]/I$, where I is an ideal of $xK[x]$ and hence principal; i.e., $I=g(x)K[x]$. If $g(x)=0$, then A contains a copy of $xK[x]$. If $g(x)\neq 0$, we can assume that $g(x)=x^s+ +x^{s+1}h(x)$. If a is not nilpotent, then $a^s=-a^{s+1}h(a)=a^{s+1}q(a)$. Hence $a^s=a^{2s}(q(a))^s$ and $(aq(a))^s$ is an idempotent element of A that generates a subalgebra of A isomorphic to K.

Lemma 3. *The following conditions on a ring variety $\mathcal{V}$ are equivalent:*

i) *$x^m+x^{m+1}h(x)\in V$ for some $m>0$, $h(x)\in\mathbf{Z}[x]$.*

ii) *$\Sigma\{a_kx^k:1\leqq k\leqq s\}\in V$ for some $s>0$ and* g.c.d.$(a_1,\ldots,a_s)=1$.

iii) *$V(x\mathbf{Z}[x])$ the T-ideal of one variable identities of $\mathcal{V}$ is not contained in $px\mathbf{Z}[x]$ for any prime p.*

Proof. It is clear that i)⇒ii)⇒iii). Let $\mathcal{V}$ satisfy iii). By Hilbert's basis theorem, $I=V(x\mathbf{Z}[x])$ is finitely generated, say by $g_1(x), g_2(x), \ldots, g_k(x)$. For any prime p one of the coefficients of $g_1(x),\ldots,g_k(x)$ is not divisible by p. Let u be an integer larger than all the degrees of $g_1(x),\ldots,g_k(x)$. Then

$$g(x)=\Sigma\{x^{iu}g_i(x):1\leqq i\leqq k\}\in I$$

satisfies ii).

Let $\mathcal{V}$ satisfy ii). Hence every member of $\mathcal{V}$ satisfies $rx^t=x^{t+1}f(x)$ for some positive integers r and t and some $f(x)\in\mathbf{Z}[x]$. Thus $\mathcal{V}$ satisfies

$$r^2x^t=rx^{t+1}f(x)=x^{t+2}(f(x))^2,\ \ldots,\ r^tx^t=x^{2t}(f(x))^t,\quad r^{2t}x^t=x^{3t}(f(x))^{2t}.$$

and substituting r^4x for x, we get

$$r^{6t}x^t=(r^{6t}x^t)^2g(x).$$

Let $A\in\mathcal{V}$ and denote by B the Jacobson radical of A. For every $b\in B$, $r^{6t}b^tg(b)$ is an idempotent element of B. Hence $r^{6t}b^tg(b)=0$ and B satisfies $r^{6t}x^t=0$. Let C be the ideal of B generated by all b^t, $b\in B$. Then C satisfies $r^{6t}x=0$. We will show that C satisfies $x^n=0$ where n depends only on the polynomial in ii), and consequently B satisfies $x^{nt}=0$. C is the direct sum of a finite number of rings of prime power characteristic. Hence, for our purpose, we can assume that C satisfies $p^k x=0$ $(k<r^{6t})$. C/pC is an algebra over the field $\mathbf{Z}_p$. $C/pC \in V$ and hence satisfies $\Sigma\{a_i x^i : 1 \leqq i \leqq s\}=0$ where g.c.d. $(a_1, \ldots, a_s)=1$. Hence C/pC cannot contain any copies of $x\mathbf{Z}_p[x]$. If C/pC were not nil, then by Lemma 2, C/pC would contain a copy of Z_p and hence an idempotent different from 0. But pC is nilpotent $(pC)^k=0$, and this nonzero idempotent can be lifted to an idempotent in C, contradicting the fact that B does not have any nonzero idempotents. Thus C/pC is nil by ii)

$$x^u = x^{u+1}h(x) \quad \text{where} \quad 1 \leqq u \leqq s,$$

and hence $x^u = x^{u+v}(h(x))^v$ for any $v>0$. If $a \in C/pC$ satisfies $a^v=0$, then $a^u=0=a^s$. Thus C satisfies $x^{sk}=0$.

Now A/B is a subdirect sum of primitive rings satisfying an identity of type ii). By Kaplansky's theorem [8] every such primitive ring is of dimension $\leqq (\frac{1}{2}s)^2$ over their centers. These centers satisfy the same identity of type ii), and hence they can be only a finite number of finite fields. Thus all these fields satisfy an identity of the type $x+x^2h(x)=0$, and hence all these primitive rings satisfy $(x+x^2h(x))^u=0$ for some $u>0$ depending only on the identity $\Sigma\{a_i x^i : 1 \leqq i \leqq s\}=0$ [9]. Thus A satisfies $(x+x^2h(x))^{uskt}=0$; i.e., $\mathcal{V}$ satisfies i).

Lemma 4. *The following two conditions on a ring variety $\mathcal{V}$ are equivalent.*

i) *if $A \in \mathcal{V}$, and if A is nil, then $x^e=0$ in A.*

ii) *$\mathcal{V}$ satisfies $x^e + x^{e+1}h(x)=0$ for some $h(x) \in \mathbf{Z}[x]$.*

Proof. ii)$\Rightarrow$i) since $x^e+x^{e+1}h(x)=0$ implies $x^e=-x^{e+1}h(x)=x^{e+1}H(x)= =x^{e+k}(H(x))^k$ for all $k>1$. Thus if $A\,(\in\mathcal{V})$ is nil, then $a^e=0$ for all $a \in A$.

Conversely, the subvariety $\mathcal{U}$ of $\mathcal{V}$ of all rings satisfying $x^{e+1}=0$, satisfies $x^e=0$ by i). Hence $x_1^e \in V+(x_1^{e+1})$ where (x_1^{e+1}) is the T-ideal of identities generated by x_1^{e+1}.

Thus $x_1^e = q(x_1)+x_1^{e+1}f(x_1)$, $q(x_1) \in V$; i.e., $q(x_1)=x_1^e - x_1^{e+1}f(x_1) \in V$.

Lemma 5. *The following conditions on a ring variety $\mathcal{V}$ are equivalent:*

i) *every nilpotent member of $\mathcal{V}$ satisfies $x_1 \ldots x_e=0$.*

ii) *$\mathcal{V}$ satisfies $x_1 \ldots x_e + f(x_1, \ldots, x_e)=0$ for some multinomial f, all terms of which are of degree larger than e in $x_1, \ldots, x_e$.*

Proof. ii) implies that the product of any e elements of a ring $A \in \mathcal{V}$ belongs to A^{e+1}, i.e., $A^e \subseteqq A^{e+1}$. Hence $A^e = A^{e+k}$ for all $k>0$. Thus if $A^s=0$ for some s, $A^e=A^{e+s}=0$.

Conversely, if $\mathcal{V}$ satisfies i) and $U=V+F^{e+1}$, then $x_1 \ldots x_e \in U$; i.e., $x_1 \ldots x_e = q+g$ where $g \in F^{e+1}$ and $q \in V$. We can assume that both q and g involve only $x_1, \ldots, x_e$; i.e., $q = x_1 \ldots x_e - g \in V$ where $g \in F^{e+1}$.

3. Proof of Theorem 1. It is obvious that 1)⇒2). Let $F_1(\mathscr{V})$ — the free member of $\mathscr{V}$ of rank 1 — be nilpotent by finite; i.e., there is an ideal I of $F_1(\mathscr{V})$ such that $F_1(\mathscr{V})/I$ is finite and $I^n=0$ for some $n>0$. Then $F_1(\mathscr{V})/I$ satisfies $(x+x^2h(x))^m=0$ for $m>0$, $h(x)\in\mathbf{Z}[x]$ [7, 9]. Hence $F_1(\mathscr{V})$ satisfies $(x+x^2h(x))^{mn}=0$. Thus $(x_1+x_1^2h(x))^{mn}\in V$; i.e., $\mathscr{V}$ satisfies 3).

Let $\mathscr{V}$ satisfy 3) and let $A\in\mathscr{V}$. Let B be the Jacobson radical of A. For any $b\in B$, $b^n=-b^{n+1}f(b)=b^{n+1}g(b)=b^{2n}(g(b))^n$. Thus $(bg(b))^n$ is an idempotent of B. Hence $(bg(b))^n=0$ and $b^n=b^n(bg(b))^n=0$; i.e., B is nil.

Let $\mathscr{V}$ satisfy 4), and let $A\in\mathscr{V}$ be finitely generated. $V(x\mathbf{Z}[x])$ is not contained in $px\mathbf{Z}[x]$; otherwise, $x\mathbf{Z}_p[x]$ would belong to $\mathscr{V}$, and hence $\mathscr{V}$ contains all commutative rings of characteristic p. The Jacobson radical of such rings may not be nil. Thus, by Lemma 3, $\mathscr{V}$ satisfies $x^n+x^{n+1}h(x)=0$ for some $n>0$ and $h(x)\in\mathbf{Z}[x]$. By Lemma 4 the Jacobson radical B of A satisfies $x^n=0$ since B is nil. A/B is a finitely generated semisimple ring satisfying $x^n+x^{n+1}h(x)=0$. Hence A/B is the subdirect sum of matrix rings over a finite number of finite fields. Thus A/B satisfies $cx=0$ for some positive integer c. Hence A/B satisfies $cx=0$, $x^n+x^{n+1}h(x)=0$. Thus A/B is finite by [7] as it is finitely generated. Now B is a subring of finite index in the finitely generated ring A. By a result of Lewin [10], B is a finitely generated ring. As B also satisfies $x^n=0$, by a result of Kaplansky [8], B is nilpotent; i.e., A possesses a nilpotent ideal of finite index concluding the proof of Theorem 1.

4. Let c, d, e be integers, $c\geqq 0$, $d>0$, $e>0$.

Definition 2. $\mathscr{C}(c, d, e)$ is the class of all rings A with the properties:

1) $cx=0$ for all $x\in A$.
2) Let B be a homomorphic image of a subring of A. Then
   a) if B is nilpotent, then $B^e=0$,
   b) if B is not nilpotent and B is primitive, then B is a finite simple ring of order dividing d.

For $c>0$, the class $\mathscr{C}(c, d, e)$ was defined by Kruse [9] in analogy to the corresponding definition for groups [13]. In [9] it is shown that

Proposition 6 [9]. *$\mathscr{C}(c, d, e)$, for $c>0$, is a variety generated by a finite ring.*

It is clear that the following also holds:

Proposition 7. *$\mathscr{C}(c, d, e)=\mathscr{C}(0, d, e)\cap\langle cx\rangle$. $\mathscr{C}(0, d, e)$ is closed under subrings and homomorphic images.*

Definition 4. $\mathscr{D}(c, d, e)$ is the class of all rings A with the properties:

1) $cx=0$ for all $x\in A$.
2) Let B be a homomorphic image of a subring of A. Then

a) if B is nil, B satisfies $x^e=0$,
b) if B is not nilpotent and B is primitive, then B is a finite simple ring of order dividing d.

It is clear that $\mathscr{C}(c, d, e) \leqq \mathscr{D}(c, d, e)$ for all $c \geqq 0$ and

Proposition 8. *$\mathscr{D}(c, d, e)=\mathscr{D}(0, d, e) \cap \langle cx \rangle$ and $\mathscr{D}(c, d, e)$ is closed under subrings and homomorphic images.*

In [4] EVERETT defines a ring C to be an extension of a ring A by a ring B if C possesses an ideal isomorphic to A whose factor is isomorphic to B.

Definition 5 [6, 11, 13]. Let $\mathscr{U}, \mathscr{V}$ be classes of rings. $\mathscr{U} \cdot \mathscr{V}$ is the class of all rings that are extensions of a ring of $\mathscr{U}$ by a ring of $\mathscr{V}$.

In [6, 13] it is shown that

Proposition 9. *If $\mathscr{U}, \mathscr{V}$ are varieties, then $\mathscr{U} \cdot \mathscr{V}$ is a variety satisfying $f(g_1, \ldots, g_n)=0$ for all $f(x_1, \ldots, x_n) \in U$, $g_1, \ldots, g_n \in V$.*

We will need two more results.

Proposition 10 [9]. *A finitely generated nilpotent by finite ring is residually finite.*

Proposition (HIGMAN [13]). *If a locally finite variety is generated by a family $\mathscr{K}$ of finite rings, then every finite member of the variety is a homomorphic image of a subring of a finite direct sum of members of $\mathscr{K}$.*

Higman's result was stated for groups. It also holds for rings.

We will show that all the $\mathscr{C}$ and $\mathscr{D}$ classes introduced here are actually varieties.

5. Theorem 12. *$\mathscr{D}(c, d, e)$ is a locally finite variety for all $c>0$.*

Claim 1: If $c>0$ and $A \in \mathscr{D}(c, d, e)$, then A satisfies $\left(x^e+x^{e+1}h(x)\right)^e=0$ where $x+x^2h(x)$ is an identity satisfied by all finite fields of order dividing d.

Let $c=p_1^{k_1} \times \ldots \times p_s^{k_s}$ be the prime factorization of c. Then $A=A_1 \times A_2 \times \ldots \times A_s$ where A_i is of prime power characteristic. It is sufficient to establish the claim for $c=p^k$. $A/pA \in \mathscr{D}(c, d, e)$. $(pA)^k=0$ and $pA \in \mathscr{D}(c, d, e)$; hence pA satisfies $x^e=0$. Let $a \in A/pA$. The subring $[a]$ of A/pA generated by a belongs to $\mathscr{D}(c, d, e)$. It is isomorphic to $x\mathbf{Z}_p[x]/I$, $I \neq 0$ since $x\mathbf{Z}_p[x] \notin \mathscr{D}(c, d, e)$. Thus $I=\left(x^r+x^{r+1}g(x)\right)\mathbf{Z}_p[x]$ for some $r>0$, $g(x) \in \mathbf{Z}[x]$. $I=x^r\mathbf{Z}_p[x] \cap \left(x+x^2g(x)\right)\mathbf{Z}_p[x]$. Hence $[a]$ is isomorphic to a subdirect sum of $x\mathbf{Z}_p[x]/x^r\mathbf{Z}_p[x]$ and $x\mathbf{Z}_p[x]/\left(x+x^2g(x)\right)\mathbf{Z}_p[x]$; i.e., $[a]$ is isomorphic to a subdirect sum of a nil ring and a finite number of finite fields all belonging to $\mathscr{D}(c, d, e)$. Let $x+x^2h(x)$ be an identity satisfied by all fields of order dividing d. Then $[a]$ satisfies $x^e+x^{e+1}h(x)=0$ since every nil member of $\mathscr{D}(c, d, e)$ satisfies $x^e=0$. Thus A/pA satisfies $x^e+x^{e+1}h(x)=0$. Hence A satisfies $\left(x^r+x^{e+1}h(x)\right)^e=0$.

Let r be the largest square free integer dividing d. All finite simple rings of order dividing d belong to $\mathscr{C}(r, d, d)$ since all are finite matrix rings over finite fields.

Claim 2: $\mathscr{D}(c, d, e) \subseteq \langle x^e \rangle \cdot \mathscr{C}(r, d, d)$ for all $c > 0$.

Let $A \in \mathscr{D}(c, d, e)$. Then by Claim 1 and Theorem 1, Rad A is nil. Hence Rad $A \in \mathscr{D}(c, d, e)$ belongs to $\langle x^e \rangle$. $A/\mathrm{Rad}\, A$ is a subdirect sum of primitive rings belonging to $\mathscr{D}(c, d, e)$. All these primitive rings are finite simple rings of order dividing d. Hence they belong to $\mathscr{C}(r, d, d)$ which is a variety by Proposition 6.

Claim 3: $\mathrm{Var}\,(\mathscr{D}(c, d, e)) \subseteq \mathscr{D}(c, d, ed)$.

Since by Claim 2 $\mathscr{D}(c, d, e) \subseteq \langle x^e \rangle \cdot \mathscr{C}(r, d, d)$ and by Proposition 9 $\langle x^e \rangle \cdot \mathscr{C}(r, d, d)$ is a variety, $\mathrm{Var}\,(\mathscr{D}(c, d, e)) \subseteq \langle x^e \rangle \cdot \mathscr{C}(r, d, d)$. Let $A \in \mathrm{Var}\,(\mathscr{D}(c, d, e))$. Then A satisfies $cx = 0$ and also $A \in \langle x^e \rangle \cdot \mathscr{C}(r, d, d)$. So, there is an ideal B of A such that B satisfies $x^e = 0$ and $A/B \in \mathscr{C}(r, d, d)$. If A is nil, then A/B is nil and hence satisfies $x^d = 0$. Thus A satisfies $x^{ed} = 0$. If A is primitive, A satisfies $(x^m + x^{m+1} f(x))^e = 0$ (by Claim 1 and Proposition 9). Hence A is finite dimensional over its center (by KAPLANSKY's theorem [8]). Thus A is a simple ring that is not nil. Hence $A \in \mathscr{C}(r, d, d)$; i.e., A is a finite simple ring of order dividing d.

Claim 4: Let $A \in \mathrm{Var}\,(\mathscr{D}(c, d, e))$ be nil. Then A satisfies $x^e = 0$.

$\mathscr{V} = \mathrm{Var}\,(\mathscr{D}(c, d, e))$ satisfies $cx = 0 = x^m + x^{m+1}h(x)$. Hence by [7] $\mathscr{V}$ is locally finite. Thus $\mathscr{V}$ is generated by its finite members belonging to $\mathscr{D}(c, d, e)$. It will be sufficient to establish the claim for the case A is finite. By Higman's Proposition 11, $A = T/I$ where T is a finite subdirect sum of finite rings from $\mathscr{D}(c, d, e)$. Thus T is finite, and hence its Jacobson radical R is nilpotent. T/R is generated by idempotents that can be lifted to a set of idempotents B such that T is generated by B and R. Since T/I is nil, $B \subseteq I$. Thus $T/I \cong R/R \cap I$. As $\mathscr{D}(c, d, e)$ is closed under subrings, R is a subdirect sum of members of $\mathscr{D}(c, d, e)$. Since R is nil, all these rings are nil, and hence satisfy $x^e = 0$. Thus R, and consequently A, satisfy $x^e = 0$. This argument is similar to an argument of KRUSE [9].

By Claims 3 and 4, $\mathscr{D}(c, d, e) \supseteq \mathrm{Var}\,(\mathscr{D}(c, d, e))$. Hence $\mathscr{D}(c, d, e)$ is a variety. Since every nil member of $\mathscr{D}(c, d, e)$ satisfies $x^e = 0$, by Lemma 4 $\mathscr{D}(c, d, e)$ satisfies $x^e + x^{e+1}h(x) = 0$ for some $h \in Z[x]$. By [7], $\mathscr{D}(c, d, e)$ is a locally finite variety.

6. Theorem 13. *$\mathscr{D}(0, d, e)$ is a variety; moreover, $\mathscr{D}(0, d, e) = \bigvee \{\mathscr{D}(c, d, e) : c > 1\}$.*

Denote by $\mathscr{K}$ the join of all $\mathscr{D}(c, d, e)$, $c > 1$.

Claim 5: $\mathscr{K}$ satisfies $(x^e + x^{e+1}h(x))^e = 0$ where $x + x^2h(x) = 0$ is an identity satisfied by all finite fields of order dividing d.

This is immediate from Claim 1.

Claim 6: $\mathscr{D}(0, d, e) \subseteq (\mathscr{K} \cdot \langle x^e \rangle) \cdot \mathscr{K}$.

Let $A \in \mathscr{D}(0, d, e)$, $B = \bigcap \{cA : c > 1\}$, and let C be the torsion ideal of B. It is clear that A/B, $C \in \mathscr{K}$ and $B/C \in \mathscr{D}(0, d, e)$ is an algebra over the field Q of rational numbers. By Lemma 2, if B/C is not nil, it contains a copy of Q or a copy of $xQ[x]$ in contradiction to the statement $B/C \in \mathscr{D}(0, d, e)$. Thus B/C is nil and hence satisfies $x^e = 0$.

Claim 7: $\mathscr{D}(0, d, e) \subseteq \mathscr{K}$.

From Claims 5 and 6 and by Proposition 9, $\mathscr{D}(0, d, e)$ satisfies an identity of the type $x^{e^3} + x^{e^3+1}g(x) = 0$. By Theorem 1 $\mathrm{Var}(\mathscr{D}(0, d, e))$ is locally nilpotent by finite. Hence every finitely generated member of $\mathscr{D}(0, d, e)$ is nilpotent by finite, and by Proposition 10, is residually finite. Thus every finitely generated member of $\mathscr{D}(0, d, e)$ is a subdirect sum of finite members of $\mathscr{D}(0, d, e)$ and hence is also a subdirect sum of finite members of $\mathscr{K}$. Thus $\mathscr{D}(0, d, e) \subseteq \mathscr{K}$.

Claim 8: $\mathrm{Var}(\mathscr{D}(0, d, e)) = \mathscr{K} \subseteq \langle x^e \rangle \mathscr{C}(r, d, d)$.

It is clear that $\mathscr{D}(c, d, e) \subseteq \mathscr{D}(0, d, e)$ for all $c > 1$. Hence $\mathscr{K} \subseteq \mathrm{Var}(\mathscr{D}(0, d, e)) \subseteq \mathscr{K}$ (by Claim 7). Also by Claim 2, $\mathscr{D}(c, d, e) \subseteq \langle x^e \rangle \cdot \mathscr{C}(r, d, d)$.

Claim 9: $\mathrm{Var}(\mathscr{D}(0, d, e)) \subseteq \mathscr{D}(0, d, ed)$.

The proof is the same as in Claim 3.

Claim 10: $\mathscr{K}$ satisfies $(x + x^2 f(x))^e = 0$ where $x + x^2 f(x)$ is an identity satisfied by all fields of order dividing d.

Consider A — the free member of rank 1 of $\mathscr{D}(c, d, e)$. If R is the Jacobson radical of A, R is nilpotent and A/R is a finite semisimple commutative ring. Thus A/R is the direct sum of finite fields belonging to $\mathscr{D}(c, d, e)$. Hence A/R satisfies $x + x^2 f(x) = 0$ where f depends only on d and R satisfies $x^e = 0$. Hence $\mathscr{D}(c, d, e)$ satisfies $(x + x^2 f(x))^e = 0$, and so does $\mathscr{K}$.

From Claims 8, 9 and 10 and by Lemma 4, $\mathrm{Var}(\mathscr{D}(0, d, e)) \subseteq \mathscr{D}(0, d, e)$, concluding the proof of Theorem 13.

7. Theorem 14. *A variety is locally finite iff it is contained in $\mathscr{D}(c, d, e)$ for some positive integers c, d, e.*

By Theorem 12, $\mathscr{D}(c, d, e)$ $(c > 0)$ is a locally finite variety. Hence every subvariety of $\mathscr{D}(c, d, e)$ is locally finite. Conversely, if $\mathscr{V}$ is a locally finite variety, it satisfies $cx = 0 = x^e + x^{e+1}h(x)$ [7]. If d is the least common multiple of the orders of all nonnilpotent finite simple rings satisfying $x^e + x^{e+1}h(x) = 0$, then $\mathscr{V} \subseteq \mathscr{D}(c, d, e)$.

Theorem 15. *A variety is locally nilpotent by finite iff it is contained in $\mathscr{D}(0, d, e)$ for some positive integers d, e.*

A subvariety of $\mathscr{D}(0, d, e)$ satisfies $x^e + x^{e+1}g(x) = 0$ and hence is locally nilpotent by finite (by Theorem 1). Conversely, if a variety is locally nilpotent by finite, it satisfies $x^e+x^{e+1}h(x)=0$, and hence it is contained in $\mathcal{D}(0, d, e)$ where d is the least common multiple of the orders of all finite simple rings satisfying $x^e+x^{e+1}h(x)=0$.

8. Theorem 16. *$\mathcal{C}(0, d, e)$ is a variety; moreover, $\mathcal{C}(0, d, e)=\bigvee\{\mathcal{C}(c, d, e): c>1\}$.*

Claim 11: $\mathcal{C}(0, d, e)\leqq\langle x_1...x_e\rangle\cdot\mathcal{C}(r, d, d)$.

Since $\mathcal{C}(0, d, e)\leqq\mathcal{D}(0, d, e)$, by Theorem 15, $\mathcal{C}(0, d, e)$ is locally nilpotent by finite. Hence, if $A\in\mathcal{C}(0, d, e)$, then the Jacobson radical R of A is nil satisfying $x^e=0$. Thus by KAPLANSKY's theorem [8], R is locally nilpotent. Hence R satisfies $x_1...x_e=0$ since $R\in\mathcal{C}(0, d, e)$. A/R is a subdirect sum of primitive rings belonging to $\mathcal{D}(0, d, e)$ and hence to $\mathcal{C}(r, d, d)$.

Claim 12: $\mathcal{C}(0, d, e)\leqq\bigvee\{\mathcal{C}(c, d, e): c>1\}$.

Let $A\in\mathcal{C}(0, d, e)$ be finitely generated. Hence A is nilpotent by finite, and by Proposition 10, A is residually finite. Hence A is a subdirect sum of finite rings belonging to $\mathcal{C}(0, d, e)$. These finite rings belong to $\bigcup\{\mathcal{C}(c, d, e): c>1\}$. Thus $A\in\bigvee\{\mathcal{C}(c, d, e): c>1\}$. Hence $\mathcal{C}(0, d, e)\leqq\bigvee\{\mathcal{C}(c, d, e): c>1\}$.

Claim 13: $\mathcal{C}(c, d, e)$ satisfies $r^e x_1...x_e=0$ for all $c\geqq 0$.

By Claim 11 and by Proposition 9, $\mathcal{C}(0, d, e)$ satisfies $(rx_1)(rx_2)...(rx_e)=0$. Also $\mathcal{C}(0, d, e)\leqq\bigvee\{\mathcal{C}(c, d, e): c>1\}$.

Claim 14: If $\mathcal{C}(r^e, d, e)$ satisfies an identity in $x_1, ..., x_n$, for some $n\geqq e$, where every term in the identity involves precisely all $x_1, ..., x_n$, then the same identity holds in $\mathcal{C}(c, d, e)$ for all $c>1$.

Denote by V_c the T-ideal of the variety $\mathcal{C}(c, d, e)$. It is clear that $\mathcal{C}(c_1, d, e)\leqq\leqq\mathcal{C}(c_2, d, e)$ if $c_1|c_2$. Thus, if $s=$g.c.d. (c, r^e), then

$$V_s = V_c+sF = V_c+cF+sF = V_c+cF+r^eF = V_c+r^eF.$$

If $n\geqq e$ and $g(x_1, ..., x_n)\in V_{r^e}$ is as described in Claim 14, then $g\in V_{r^e}\leqq V_s=V_c+r^eF$. Thus there are $v\in V_c$ and $f\in F$ such that $g=v+r^ef$. By substituting 0 for all variables outside $Y\subseteq\{x_1, x_2, ...\}$ we get equality between the sum of terms involving variables from Y only in g and $v+r^ef$. Hence we can assume that every term in v and f involves precisely $x_1, ..., x_n$. Thus $f\in F^n\subseteq F^e$. But $V_c\geqq r^eF^e$. Thus $r^ef\in V_c$, and $g=v+r^ef\in V_c$.

Claim 15: $\mathcal{C}(0, d, e)$ satisfies $x_1...x_e+f(x_1, ..., x_e)=0$ where $f\in F^{e+1}$.

By Lemma 5 $\mathcal{C}(r^e, d, e)$ satisfies $x_1...x_e+f(x_1, ..., x_e)=0$ since every nilpotent member of $\mathcal{C}(r^e, d, e)$ satisfies $x_1...x_e=0$. By Claim 14, $\mathcal{C}(c, d, e)$ satisfies $x_1...x_e+$ $+f(x_1, ..., x_e)=0$. Hence Claim 15 follows from Claim 12.

Var $(\mathcal{C}(0, d, e))$ satisfies $x_1...x_e+f=0$, $f\in F^{e+1}$. Hence every nilpotent member of Var $(\mathcal{C}(0, d, e))$ satisfies $x_1...x_e=0$. If A is primitive and not nilpotent, and $A\in$Var$(\mathcal{C}(0, d, e))$, then $A\in\mathcal{D}(0, d, e)$. Hence A is a finite simple ring of order dividing d. Thus

$$\mathrm{Var}\,(\mathscr{C}(0, d, e)) \leqq \mathscr{C}(0, d, e)).$$

But $\mathrm{Var}\,(\mathscr{C}(0, d, e)) = \bigvee \{\mathscr{C}(c, d, e) : c > 1\}$. This concludes the proof of Theorem 16.

9. The following is a generalization of KRUSE's theorem [9] that the identities of a finite ring are finitely based.

Theorem 17. *If $\mathscr{V} \leqq \mathscr{C}(0, d, e)$ then the identities of $\mathscr{V}$ are finitely based.*

Denote by F_k the free associative ring on $\{x_1, \dots, x_k\}$. $\mathscr{V}^{(k)}$ the variety defined by the k-variable identities of V is finitely based for any k. Since $F_k \cap V$ — the T-ideal of V in F_k — determines the variety $\mathscr{V}^{(k)}$ $F_k/F_k \cap V = F_k(V)$ is nilpotent by finite. Hence $V \cap F_k \leqq I$ for some ideal I of F_k of finite index and $I/V \cap F_k$ is nilpotent. By LEWIN's result [10], I is finitely generated, and so $F_k \cap V$ is finitely generated.

$\mathscr{W} = \mathscr{V} \cap \langle r^e x \rangle$ is a subvariety of $\mathscr{C}(r^e, d, e)$. Hence by [7] $\mathscr{W}$ is generated by a finite ring and by KRUSE's theorem [9], $\mathscr{W}$ is finitely based; i.e., $W = V + r^e F$ is finitely generated as a T-ideal. Thus all identities of $\mathscr{W}$ are consequences of $r^e x, v_1, \dots, v_n$ where $v_1, \dots, v_n$ can be chosen in V. Let $v \in V$ involve precisely $x_1, \dots, x_m$, $m \geqq e$. Then $v \equiv r^e f + w$ where $f \in F$ and w is a consequence of $v_1, \dots, v_n$. Comparing the terms involving the same set of variables, we get

$$v \equiv r^e f' + w', \quad 0 \equiv r^e f'' + w''$$

where f', w' are the sums of all terms of f and w involving precisely $x_1, \dots, x_m$, $f'' \equiv f - f'$, $w'' \equiv w - w'$. But w', w'' are also consequences of $v_1, \dots, v_n$ and $r^e f' \in V$ since $f' \in F^e$. $r^e f'' \equiv -w''$. Hence $v \equiv r^e f' - w'' + w$. Thus v is a consequence of $r^e x_1 \dots x_e$, $v_1, \dots, v_n$. Thus $V \cap F_{e-1} \cup \{r^e x_1 \dots x_e, v_1, \dots, v_n\}$ is a basis for V. Hence $\mathscr{V}$ is finitely based.

Theorem 18. *The following conditions on a variety $\mathscr{V}$ are equivalent:*

1) *$\mathscr{V} \leqq \mathscr{C}(0, d, e)$ for some positive integers d, e.*

2) *$\mathscr{V}$ satisfies $x_1 \dots x_n + f(x_1, \dots, x_n) = 0$ for some $f \in F^{n+1}$ and some $n > 0$.*

3) *the Jacobson radical of every member is nilpotent.*

We have established that 1)$\Rightarrow$2)$\Rightarrow$3). If $\mathscr{V}$ satisfies 3), then by Theorem 1, $\mathscr{V}$ is locally nilpotent by finite. Thus by Theorem 15, $\mathscr{V} \leqq \mathscr{D}(0, d, e)$ for some positive integers d and e. By Claim 8, $\mathscr{V} \leqq \langle x^e \rangle \cdot \mathscr{C}(r, d, d)$. Let A be the free ring of $\mathscr{V}$ of rank ω. Then $A/\mathrm{Rad}\, A \in \mathscr{C}(r, d, d)$ and $\mathrm{Rad}\, A$ is nilpotent say $(\mathrm{Rad}\, A)^m = 0$. Thus $A \in \langle x_1 \dots x_m \rangle \cdot \mathscr{C}(r, d, d)$. Since A generates $\mathscr{V}$

$$\mathscr{V} \leqq \langle x_1 \dots x_m \rangle \cdot \mathscr{C}(r, d, d) \leqq \mathscr{C}(0, d, md).$$

Corollary 1. *If a variety satisfies $x_1 \dots x_n + f(x_1, \dots, x_n) = 0$, where $f \in F^{n+1}$ and $n > 0$, then it is finitely based.*

Corollary 2. *A variety is generated by a finite ring iff it satisfies* $cx=0$ *and* $x_1 \dots x_n + f(x_1, \dots, x_n) = 0$ *for some* $n>0$, $c>0$, *and* $f \in F^{n+1}$.

This is immediate from Theorem 18 and the observation that a finite ring belongs to $\mathscr{C}(c, d, e)$ for some $c, d, e>0$.

10. We now come to the condition that the Jacobson radical be a direct summand.

Theorem 19. *The following conditions on a ring variety are equivalent:*

1) *the Jacobson radical of every finitely generated member is a direct summand;*
2) *the Jacobson radical of every member generated by two elements is a direct summand.*

Proof. It is obvious that 1)$\Rightarrow$2).

If $\mathscr{V}$ is a variety such that $V(x\mathbf{Z}[x]) \leqq px\mathbf{Z}[x]$, for some prime p, then $\mathscr{V}$ contains $x\mathbf{Z}_p[x]$ and hence all commutative rings of characteristic p. The ring $\{(x, y): x, y \in \mathbf{Z}_p\}$ with component-wise addition and $(x, y)(z, t) = (xt+yz, yt)$ is commutative of characteristic p; its Jacobson radical is $\{(x, 0): x \in \mathbf{Z}_p\}$, and the radical is not a direct summand. However, this ring is generated by $(1, 1)$. Thus, if $\mathscr{V}$ satisfies 2), $V(x\mathbf{Z}[x])$ is not contained in $px\mathbf{Z}[x]$ for any prime p. By Lemma 3, $\mathscr{V}$ satisfies $x^n + x^{n+1}f(x) = 0$, and hence by Theorem 1, the radical of every member of $\mathscr{V}$ is nil; moreover, it satisfies $x^n = 0$. Let A be a finitely generated member of $\mathscr{V}$, and let R be the Jacobson radical of A. Hence A/R is a finite semisimple ring. So A/R has 1. As R is nil, 1 can be lifted to an idempotent $c \in A$. Hence $A = cA + R$. Let $b \in R$. Then the subring B generated by b and c belongs to $\mathscr{V}$, and hence its radical C is a direct summand. But the radical C of B contains $R \cap B$. The projection of B onto C sends idempotents to 0. Thus $cb = bc = 0$. Hence $A(cA) = (cA+R)(cA) = (cA)(cA) + R(cA) = c(AcA) \subseteqq cA$. Hence cA is a two sided ideal of A. $R \cap cA = 0$ and $R + cA = A$. Hence $A = R \oplus cA$.

From Theorem 19, the condition 1), equivalent to 2), is equivalent to $\mathscr{V}^{(2)}$, the variety of all rings satisfying the two variable identities of $\mathscr{V}$, satisfies condition 1) of Theorem 19. But the identities of $\mathscr{V}^{(2)}$ are finitely based. Thus condition 2) is equivalent to a finite set of two variable identities. The following shows that this cannot be improved.

Theorem 20. *Let* $\mathscr{V}$ *be a ring variety for which the condition that every finitely generated member of* $\mathscr{V}$ *has the radical as a direct summand is equivalent to a set of one variable identities. Then* $\mathscr{V}$ *satisfies* $x^e = 0$ *or* $x + x^2h(x) = 0$.

In the varieties $\langle x^e \rangle$, every ring is radical. In the varieties $\langle x + x^2h(x) \rangle$, the radical is 0.

If $\mathscr{V}$ is a variety in which the radical of every ring generated by one element is a direct summand, then $V(x\mathbf{Z}[x])$ is not containedi n $px\mathbf{Z}[x]$ for any prime p; otherwise, all commutative rings of characteristic p belong to $\mathscr{V}$. In the proof of Theorem 19, we have shown that there is a ring of characteristic p generated by one element and its radical is not a direct summand. Thus, by Lemma 3, $\mathcal{V}$ satisfies $x^e+x^{e+1}h(x)=0$. For some prime p, $\mathcal{V}$ contains the minimal variety $\langle px, xy\rangle$ [15]; otherwise, $\mathcal{V}$ is a variety not containing $\langle px, xy\rangle$ for any prime p; and by [6], this is equivalent to the validity of $x+x^2f(x)=0$ in $\mathcal{V}$ for some $f(x)\neq 0$. Also for some prime q, $\mathcal{V}$ contains the prime field of q elements $\mathbf{Z}_q$. This is true since by Theorem 15, $\mathcal{V}\leqq\mathcal{D}(0, d, e)$, where d is the least common multiple of all orders of finite simple nonnilpotent rings belonging to $\mathcal{V}$. If $\mathcal{V}$ does not contain any nonnilpotent simple finite ring, then $\mathcal{V}\leqq\mathcal{D}(0, 1, e)=\langle x^e\rangle$. Thus, if $\mathcal{V}$ does not satisfy $x^e=0$, then $\mathcal{V}$ contains a nonnilpotent finite simple ring whose center is $\mathbf{Z}_q$ for some prime q. The variety $\mathcal{U}=\langle qx, xy-x^qy, xy-xy^q\rangle$ is contained in $\langle qx, xy\rangle\vee\langle qx, x-x^q\rangle\leqq\mathcal{V}$. The one variable identities of $\mathcal{U}$ are all consequences of $qx=x^2-x^{q+1}=0$. The ring $\{(a, b)\colon a, b\in\mathbf{Z}_q\}$ with $(a, b)+(c, d)=(a+c, b+d)$, $(a, b)(c, d)=(ac, ad)$ satisfies $qx=0=x^2-x^{q+1}$. Its radical is $\{(0, b)\colon b\in\mathbf{Z}_p\}$, and the radical is not a direct summand. Thus, if $\mathcal{V}$ is a variety not satisfying $x^e=0$ or $x+x^2h(x)=0$ for any $e>0$ or $h(x)\neq 0$, then the condition that the radical of every finitely generated member is a direct summand is not equivalent to any set of one variable identities.

References

[1] G. Birkhoff, *Lattice Theory* (3rd ed), AMS Colloq. Publ. (Providence, 1967).
[2] P. M. Cohn, *Universal Algebra*, Harper & Row (New York, 1965).
[3] P. M. Cohn, *Free Rings and their Relations*, Academic Press (London and New York, 1971).
[4] C. J. Everett, An extension theory for rings, *Amer. J. Math.*, **64** (1942), 363—370.
[5] G. Grätzer, *Universal Algebra*, Van Nostrand (Princeton, N. J., 1968).
[6] A. A. Iskander, Product of ring varieties and attainability, *Trans. Amer. Math. Soc.*, **193** (1974), 231—238.
[7] A. A. Iskander, Locally finite ring varieties, *Proc. Amer. Math. Soc.*, to appear.
[8] I. Kaplansky, Rings with a polynomial identity, *Bull. Amer. Math. Soc.*, **54** (1948), 575—580.
[9] R. L. Kruse, Identities satisfied by a finite ring, *J. Algebra*, **26** (1973), 298—318.
[10] J. Lewin, Subrings of finite index in finitely generated rings, *J. Algebra*, **5** (1967), 84—88.
[11] A. I. Mal'cev, On class multiplication of algebraic systems, *Sibirsk. Math. Ž.*, **8**: 2 (1967), 346—355., (Russian).
[12] A. I. Mal'cev, *Algebraic Systems*, Nauka (Moscow, 1970) (Russian); English transl., Die Grundlehren der math. Wissenschaften, Band 192, Springer-Verlag (Berlin and New York, 1973).
[13] Hanna Neumann, *Varieties of Groups*, Ergebnisse der Math. und ihrer Grenzgebiete, vol. 37, Springer-Verlag (New York, 1967).
[14] J. M. Osborn, Varieties of algebras, *Advances in Math.*, **8** (1972), 163—369.
[15] A. Tarski, Equationally complete rings and relation algebras, *Nederl. Akad. Wetensch. Proc.* Ser. A. **59** = *Indag. Math.*, **18** (1956), 39—46.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTHWESTERN LOUISIANA
LAFAYETTE, LOUISIANA 70501, USA

Acta Sci. Math., 39 (1977), 303—312

Theory of radicals for hereditarily artinian rings

A. WIDIGER and R. WIEGANDT

I. Introduction

As is well known, a radical-semisimple theory can be built up in every universal class of rings, but in most of the cases the universal class considered is that of all associative rings or at least another *variety* of rings. So it seems to be reasonable to have a better look at radical-semisimple theory developed in a class which is *not a variety*. In the present paper we shall consider the class **K** of all artinian rings with artinian Jacobson radical i.e. of every artinian ring each ideal of which is artinian. The rings of **K** will be called *hereditarily artinian rings*. The class **K** does not admit infinite direct sums and subrings of **K**-rings are not necessarily **K**-rings, though **K** is hereditary, homomorphically closed, and closed under extensions either. The structure of **K**-rings has been described in [2]. In the present paper we shall develop the general radical and semisimple theory in the category **K**, we shall characterize the radical and semisimple classes by certain algebraic properties and shall give explicitely all the radical and semisimple classes of **K**. Among others it will be proved that a subclass of **K** is a semisimple class iff it is hereditary and closed under extensions, further a subclass **R** is a radical class iff **R** is homomorphically closed, closed under extensions, and contains the zeroring $Z(p^\infty)$ whenever $Z(p)\in\mathbf{R}$. Also all N-radicals in **K** are determined. Since **K** is not a variety, connections among algebraic properties are different from those in a variety. For instance, in **K** every hereditary radical class is a homomorphically closed semisimple class, but the converse statement is not true. Let us remind that in a ring variety the situation is just the opposite: homomorphically closed semisimple classes are always subvarieties and strongly hereditary radical classes, but not conversely (cf. e.g. [5] Theorem 34.1 and Corollary 32.2).

Received January 19, 1976.

2. Preliminaries

A subclass $\mathbf{R}$ of $\mathbf{K}$ is called a *radical class* if it satisfies the following conditions:

(R_1) if $A \in \mathbf{R}$, then every non-zero homomorphic image B of A has a non-zero accessible subring C in $\mathbf{R}$;

(R_2) if every non-zero homomorphic image B of a ring $A \in \mathbf{K}$ has a non-zero accessible subring C in $\mathbf{R}$, then $A \in \mathbf{R}$.

Dualizing the definition of radical classes we get that of semisimple classes. A subclass $\mathbf{S}$ of $\mathbf{K}$ is said to be a *semisimple class,* if it satisfies the conditions:

(S_1) if $A \in \mathbf{S}$, then every non-zero accessible subring B of A has a non-zero homomorphic image C in $\mathbf{S}$;

(S_2) if every non-zero accessible subring B of a ring $A \in \mathbf{K}$ has a non-zero homomorphic image C in $\mathbf{S}$, then $A \in \mathbf{S}$.

Let us consider the class functions $\mathscr{U}$ and $\mathscr{S}$ defined by

$$\mathscr{U}\mathbf{P} = \{A \in \mathbf{K} \mid A \text{ has no non-zero homomorphic image in } \mathbf{P}\}$$

and

$$\mathscr{S}\mathbf{Q} = \{A \in \mathbf{K} \mid A \text{ has no non-zero ideal in } \mathbf{Q}\}.$$

As has been shown in [1] Theorem 1 in the framework of a more general theory, the class functions $\mathscr{U}$ and $\mathscr{S}$ establish a Galois-correspondence between hereditary and homomorphically closed subclasses of $\mathbf{K}$, and the closed subclasses are exactly the semisimple and radical classes, respectively. Thus, if $\mathbf{P}$ is a hereditary subclass of $\mathbf{K}$ and $\mathbf{Q}$ is a homomorphically closed subclass of $\mathbf{K}$, then $\mathscr{U}\mathbf{P}$ is a radical class and $\mathscr{S}\mathbf{Q}$ is a semisimple class, moreover, for each radical class $\mathbf{R}$ and semisimple class $\mathbf{S}$ we have $\mathbf{R} = \mathscr{U}\mathscr{S}\mathbf{R}$ and $\mathbf{S} = \mathscr{S}\mathscr{U}\mathbf{S}$. *Replacing the notion "accessible subring" by that of "ideal" in conditions* (R_1), (R_2), (S_1) *and* (S_2) *we get again the definitions of radical and semisimple classes* (cf. for instance [5] Theorems 10.6 and 10.7). It is easy to check that for every radical class $\mathbf{R}_A$ and semisimple class $\mathbf{S}_A$ of the class $\mathbf{A}$ of all associative rings, the intersections $\mathbf{R}_A \cap \mathbf{K}$ and $\mathbf{S}_A \cap \mathbf{K}$ yield radical and semisimple classes in $\mathbf{K}$, respectively. Further, if $\mathbf{R}$ is a radical class in $\mathbf{K}$ and $\mathscr{L}\mathbf{R}$ denotes the lower radical class of $\mathbf{A}$ generated by $\mathbf{R}$ that is

$$\mathscr{L}\mathbf{R} = \left\{ A \in \mathbf{A} \;\middle|\; \begin{array}{l} \text{every non-zero homomorphic image of } A \\ \text{has a non-zero accessible subring in } \mathbf{R} \end{array} \right\}$$

then $\mathbf{R} = \mathscr{L}\mathbf{R} \cap \mathbf{K}$ holds. Similarly, if $\mathbf{S}$ is a semisimple class of $\mathbf{K}$ and

$$\mathscr{M}\mathbf{S} = \left\{ A \in \mathbf{S} \;\middle|\; \begin{array}{l} \text{every non-zero accessible subring of } A \\ \text{has a non-zero homomorphic image in } \mathbf{S} \end{array} \right\}$$

is the semisimple class of $\mathbf{A}$ generated by $\mathbf{S}$, then $\mathbf{S} = \mathscr{M}\mathbf{S} \cap \mathbf{K}$ is valid.

All these considerations are valid not only for **K**, but for any hereditary and homomorphically closed subclass of **A**.

For further details of the radical theory we refer to [5].

Let A be a ring. We shall denote the additive group of A by $(A, +)$, the Jacobson radical of A by $\mathbf{J}(A)$, and the ring of all $n \times n$ matrices over A by A_n, respectively. The symbol $\oplus$ stands for ring theoretic direct sum. For any prime p, $Z(p^k)$ denotes the cyclic group of order p^k and also the zeroring on this group, $Z(p^\infty)$ the quasicyclic p-group and also the zeroring on this group.

In the following we use the results of [2]. The main result of [2] (Satz 3) states that A is hereditarily artinian iff

$$A = S_{n_1}^{(1)} \oplus \ldots \oplus S_{n_k}^{(k)} \oplus A^* \tag{D}$$

where each $S_{n_i}^{(i)}$ is the matrix ring over an infinite division ring $S^{(i)}$ and A^* is a strong artinian ring, i.e. $(A^*, +)$ satisfies the minimum condition on subgroups. As is well known, A^* is a torsion ring and

$$A^* = A(p_1) \oplus \ldots \oplus A(p_l)$$

where the $A(p_i)$-s, the so called p_i-components of A^*, are p_i-rings for distinct primes p_i. $(A(p_i), +)$ is a direct sum of a finite group and finitely many copies of $Z(p^\infty)$ lying in the annihilator of $A(p_i)$ (and thus of A).

3. Semisimple classes of K

A class **H** is said to be *hereditary*, if $I \triangleleft A \in \mathbf{H}$ implies $I \in \mathbf{H}$. We say that a class **E** is *closed under extensions*, if $B \triangleleft A$, $B \in \mathbf{E}$ and $A/B \in \mathbf{E}$ implies $A \in \mathbf{E}$.

Theorem 1. *A subclass* **S** *of* **K** *is a semisimple class in* **K** *iff*

(i) **S** *is hereditary, and*

(ii) **S** *is closed under extensions.*

Proof. In view of [5] Theorem 30.1 it suffices to show that conditions (i) and (ii) imply (S_1) and (S_2). The validity of (S_1) follows immediately from (i).

Next, take a ring $A \in \mathbf{K}$ satisfying the requirements of condition (S_2). We shall prove that $A \in \mathbf{K}$. For this end let us consider the ideal

$$I = \bigcap_\alpha (M_\alpha \triangleleft A \mid A/M_\alpha \in \mathbf{S}).$$

Since A is artinian, the ideal I can be represented as a finite intersection $I = \bigcap_{i=1}^{n} M_i$. By induction we exhibit $A/I \in \mathbf{S}$. For $n=1$ the statement is trivial. Assuming

$A/\bigcap_{i=1}^{n-1} M_i \in \mathbf{S}$ we get

$$(*) \qquad A/\bigcap_{i=1}^{n} M_i/(M_n/\bigcap_{i=1}^{n} M_i) \cong A/M_n \in \mathbf{S}$$

and

$$M_n/\bigcap_{i=1}^{n} M_i \cong \left(M_n + \bigcap_{i=1}^{n=1} M_i\right)\Big/\bigcap_{i=1}^{n-1} M_i \lhd A/\bigcap_{i=1}^{n-1} M_i.$$

By (i) and the hypothesis it follows that $M_n/\bigcap_{i=1}^{n-1} M_i \in \mathbf{S}$, hence $(*)$ and (ii) imply $A/I \in \mathbf{S}$.

The proof will be done if we show $I=0$. Suppose $I \neq 0$ and consider the ideal

$$J = \bigcap_{\beta} (K_\beta \lhd I \mid I/K_\beta \in \mathbf{S})$$

of I. We want to see that J is an ideal of A, too. Taking into account that $\mathbf{K}$ is hereditary and so I also is artinian, by arguments similar to those used in proving $A/I \in \mathbf{S}$, we get $I/J \in \mathbf{S}$. Choose an arbitrary element $a \in A$ and define the mapping

$$\varphi : J \to (aJ+J)/J$$

by $\varphi(x) = ax+J$ for all $x \in J$. By [5] Proposition 5.1 φ maps J homomorphically onto the ideal $(aJ+J)/J$ of I/J. Since $I/J \in \mathbf{S}$, condition (i) implies

$$J/\operatorname{Ker}\varphi \cong (aJ+J)/J \in \mathbf{S}$$

where

$$\operatorname{Ker}\varphi = \{y \in J \mid ay \in J\}.$$

We claim that $\operatorname{Ker}\varphi$ is an ideal of I, too. Suppose that $y \in \operatorname{Ker}\varphi$ and $i \in I$. Then

$$a(iy) = (ai)y \in J \quad \text{and} \quad a(yi) = (ay)i \in J,$$

since $y \in \operatorname{Ker}\varphi$. Thus $\operatorname{Ker}\varphi$ is an ideal of I. Now

$$I/\operatorname{Ker}\varphi/(J/\operatorname{Ker}\varphi) \cong I/J \in \mathbf{S}$$

holds and since $J/\operatorname{Ker}\varphi \in \mathbf{S}$, condition (ii) implies $I/\operatorname{Ker}\varphi \in \mathbf{S}$. Hence $J = \bigcap_\beta K_\beta \subset \operatorname{Ker}\varphi$ and it follows that $(aJ+J)/J \cong J/\operatorname{Ker}\varphi = 0$. Thus $aJ \subset J$ holds for every $a \in A$. We get similarly $Ja \in A$, and so J is an ideal of A. Moreover, applying (ii) by $I/J \in \mathbf{S}$ and

$$A/J/(I/J) \cong A/I \in \mathbf{S}$$

we have $A/J \in \mathbf{S}$. Hence $I \subset J$ follows. But on the other hand by the assumption upon A the non-zero ideal I of A has a non-zero homomorphic image I/K_β in $\mathbf{S}$, hence $J \subset I$ and $J \neq I$ follows, a contradiction. Thus $I=0$ and the proof is complete.

The proof of Theorem 1 is a modified version of the proof given in [3] for characterizing semisimple classes of associative or alternative rings as hereditary classes being closed under extensions and subdirect sums. Working in the category $\mathbf{K}$, we could eliminate the requirement of being closed under subdirect sums.

Theorem 2. *Let* $\mathbf{S}$ *be a semisimple class of* $\mathbf{K}$. *Then there exist*

(1) *a set* $P(\mathbf{S})$ *of primes and for each prime* $p_i \in P(\mathbf{S})$ *a semisimple class* $\mathbf{L}_{p_i}(\mathbf{S})$ *of strong artinian* p_i*-rings containing all finite nilpotent* p_i*-rings,*

(2) *a class* $\mathbf{H}_1(\mathbf{S})$ *of matrix rings over infinite division rings,*

(3) *a class* $\mathbf{H}_2(\mathbf{S})$ *of matrix rings over finite fields of characteristic* $\notin P(\mathbf{S})$ *such that* $A \in \mathbf{S}$ *iff* A *is a direct sum of rings taken from* $\cup (\mathbf{L}_{p_i}(\mathbf{S}) | p_i \in P(\mathbf{S})) \cup \mathbf{H}_1(\mathbf{S}) \cup \cup \mathbf{H}_2(\mathbf{S})$.

Conversely, if P *is a set of primes,* $\mathbf{L}_{p_i}(p_i \in P)$, $\mathbf{H}_1$, $\mathbf{H}_2$ *classes of rings as given in* (1), (2), (3), *then*

$$\mathbf{S}' = \{A \mid A \text{ is a finite direct sum of rings from } (\cup \mathbf{L}_{p_i}) \cup \mathbf{H}_1 \cup \mathbf{H}_2\}$$

is a semisimple class of $\mathbf{K}$. *Moreover if* $P = P(\mathbf{S})$, $\mathbf{L}_{p_i} = \mathbf{L}_{p_i}(\mathbf{S})$ *for every prime* $p_i \in P$ *and* $\mathbf{H}_1 = \mathbf{H}_1(\mathbf{S})$, $\mathbf{H}_2 = \mathbf{H}_2(\mathbf{S})$, *then* $\mathbf{S} = \mathbf{S}'$.

Proof. Assume that $\mathbf{S}$ is a semisimple class, that is $\mathbf{S}$ satisfies (i) and (ii). Define

$$\mathbf{H}_1(\mathbf{S}) = \left\{\begin{array}{l}\text{all matrix rings over infinite division rings}\\ \text{occurring in the decomposition } (D) \text{ of any ring of } \mathbf{S}\end{array}\right\}$$

Also define $P(\mathbf{S})$ by $p_i \in P(\mathbf{S})$ iff there exists a ring $A \in \mathbf{S}$ with $\mathbf{J}(A(p_i)) \neq 0$. By (i) also $\mathbf{J}(A(p_i)) \in \mathbf{S}$ holds and for any $p_i \in P(\mathbf{S})$. (i) and (ii) easily yield that $\mathbf{S}$ contains every finite nilpotent p_i-ring. (If $Z(p_i^\infty)$ is contained in a ring of $\mathbf{S}$, then clearly $\mathbf{S}$ contains all nilpotent artinian p_i-rings.) Let us consider the classes

$$\mathbf{H}_2(\mathbf{S}) = \left\{\begin{array}{l}\text{all matrix rings over finite fields of characteristic}\\ \notin P(\mathbf{S}) \text{ occurring as a direct summand of any ring of } \mathbf{S}\end{array}\right\}$$

and

$$\mathbf{L}_{p_i}(\mathbf{S}) = \left\{\begin{array}{l}\text{all } p_i\text{-rings, } p_i \in P(\mathbf{S}), \text{ occurring as a direct summand}\\ \text{of the strong artinian part of any ring of } \mathbf{S}\end{array}\right\}$$

Since $\mathbf{S}$ satisfies (i) and (ii), so does every $\mathbf{L}_{p_i}(\mathbf{S})$.

Conversely, suppose that P, $\mathbf{L}_{p_i}(p_i \in P)$, $\mathbf{H}_1$, $\mathbf{H}_2$ are given as required. We have to show that the class $\mathbf{S}'$ defined above has properties (i) and (ii). The class $\mathbf{K}$ is closed under extensions (this follows easily from the fact that $\mathbf{K}$ is hereditarily artinian). Hence an extension A of a ring of $\mathbf{S}'$ with a ring of $\mathbf{S}'$ is contained in $\mathbf{K}$ and so the main result of [2] is applicable to A. Thus as each $\mathbf{L}_{p_i}$ satisfies (ii),we obtain $A \in \mathbf{S}'$. Again by the main result of [2] the class $\mathbf{S}'$ satisfies also (i).

The last statement of the Theorem is obvious.

In Theorem 2 $P(\mathbf{S})$, $\mathbf{H}_1$ or $\mathbf{H}_2$ may be empty. Further, let us remark that the semisimple classes of $\mathbf{K}$ are described only up to semisimple classes $\mathbf{L}_p$ of strong artinian p-rings containing all finite nilpotent p-rings. For homomorphically closed semisimple classes the characterization will be more explicit.

If A is a ring of $\mathbf{K}$, then by the Wedderburn—Artin Structure Theorem the ring $A/\mathbf{J}(A)$ is a direct sum of infinite simple rings and a uniquely determined finite ring F. The ring F will be called the *finite part* of $A/\mathbf{J}(A)$.

Proposition 3. *A semisimple class* $\mathbf{S}$ *of* $\mathbf{K}$ *is homomorphically closed iff it contains also the finite part of* $A/\mathbf{J}(A)$, *whenever* $A\in\mathbf{S}$.

Proof. Let $\mathbf{S}$ be a semisimple class. If $\mathbf{S}$ is homomorphically closed and $A\in\mathbf{S}$, then the finite part of $A/\mathbf{J}(A)$ is clearly a homomorphic image of A and therefore contained in $\mathbf{S}$.

Conversely, assume that $\mathbf{S}$ satisfies the condition imposed in the Proposition. From Theorem 2 it follows that for each $p_i\in P(\mathbf{S})$ all finite nilpotent p_i-rings are in $\mathbf{S}$ and if $Z(p_i^\infty)\in\mathbf{S}$, then all artinian nilpotent p_i-rings are in $\mathbf{S}$. Hence all factor rings of $\mathbf{J}(A)$ are in $\mathbf{S}$. Thus, if I is an ideal of A, we have

$$(I+\mathbf{J}(A))/I \cong \mathbf{J}(A)/I\cap\mathbf{J}(A)\in\mathbf{S}.$$

Applying Theorem 2 and the assumption that the finite part of $A/\mathbf{J}(A)$ is in $\mathbf{S}$, it follows that $A/\mathbf{J}(A)\in\mathbf{S}$. Since

$$A/I/((I+\mathbf{J}(A))/I)\cong A/(I+\mathbf{J}(A)),$$

we get $A/I\in\mathbf{S}$ by (ii) whenever $A/(I+\mathbf{J}(A))\in\mathbf{S}$. But $A/(I+\mathbf{J}(A))$, as a factor ring of $A/\mathbf{J}(A)$, is isomorphic to a direct summand of $A/\mathbf{J}(A)$ and therefore by (i) contained in $\mathbf{S}$.

Theorem 4. *A homomorphically closed semisimple class* $\mathbf{S}$ *is uniquely determined by two sets* P_1, P_2 *of primes with* $P_2\subset P_1$ *and a class* $\mathbf{H}$ *of matrix rings over division rings in the following way:* $A\in\mathbf{S}$ *iff* $A/\mathbf{J}(A)$ *is a direct sum of rings of* $\mathbf{H}$, *and* $\mathbf{J}(A)$ *is a direct sum of nilpotent artinian* p_i*-rings for* $p_i\in P_1$ *and* $\mathbf{J}(A)$ *does not contain* $Z(p_i^\infty)$ *if* $p_i\in P_2$.

Proof. Choosing $P_1=P(\mathbf{S})$ the proof follows immediately from Theorem 2 and Proposition 3.

4. Radical classes in K

Firstly we shall characterize the radical classes of $\mathbf{K}$ by

Theorem 5. *A subclass* $\mathbf{R}$ *of* $\mathbf{K}$ *is a radical class in* $\mathbf{K}$ *iff*

(a) $\mathbf{R}$ *is homomorphically closed,*

(b) $\mathbf{R}$ *is closed under extensions,*

(c) *if* $Z(p)\in\mathbf{R}$ *for a prime* p, *then also* $Z(p^\infty)\in\mathbf{R}$.

Proof. Let $\mathbf{R}$ be a radical class in $\mathbf{K}$. Then (a) and (b) follows as in Theorems 3.2 and 3.3 in [5]. If $Z(p)\in\mathbf{R}$ for a prime p, then by (R_2) it follows immediately that $Z(p^\infty)\in\mathbf{R}$.

Conversely, let $\mathbf{R}$ be a subclass of $\mathbf{K}$ satisfying conditions (a), (b), (c). Condition (a) implies trivially the validity of (R_1). To show that $\mathbf{R}$ is a radical class it suffices to show that if every non-zero homomorphic image of a ring $A \in \mathbf{K}$ has a non-zero ideal in $\mathbf{R}$, then $A \in \mathbf{R}$ also holds. In view of decomposition (D) we may confine ourselves to the case $A = A(p)$ where $A(p)$ is a strong artinian p-ring. We claim that the maximal divisible ideal D of A is in $\mathbf{R}$. For $D=0$ the assertion is trivial. If $D \neq 0$, then considering the so called kernel A_0 of A (cf. [2] Satz 1), A_0 is a finite image ideal of A contained in D, furthermore, $A/A_0 \cong D$. Hence D, as a homomorphic of A, contains a non-zero ideal in $\mathbf{R}$. Now (a), (c) and (b) imply $Z(p) \in \mathbf{R}$, $Z(p^\infty) \in \mathbf{R}$ and $D \in \mathbf{R}$.

Applying (a), (b) and the second isomorphism theorem, the sum of two $\mathbf{R}$-ideals of A/D is again in $\mathbf{R}$. Hence A/D being finite, it contains a maximal $\mathbf{R}$-ideal J/D. If $J \neq A$ then by assumption A/J has a non-zero $\mathbf{R}$-ideal K/J. Thus by (b) and

$$K/D/(J/D) \cong K/J$$

we obtain $K/D \in \mathbf{R}$ contradicting the maximality of J/D. Hence $A = J$ and by $A/D \in \mathbf{R}$, $D \in \mathbf{R}$ condition (b) infers $A \in \mathbf{R}$. Thus (R_2) holds.

Corollary 6. *Every hereditary radical class in* $\mathbf{K}$ *is a homomorphically closed semisimple class.*

This is clear by Theorems 1 and 5.

The converse statemant of Corollary 6 is, however, false. *A homomorphically closed semisimple class in* $\mathbf{K}$ *need not be a radical class.* Take, for instance, the class $\mathbf{M}$ of all finite nilpotent p-rings for a fixed prime p. $\mathbf{M}$ is a homomorphically closed semisimple class. But $\mathbf{M}$ fails to be a radical class, for $Z(p^\infty) \notin \mathbf{M}$. Thus a class $\mathbf{P} \subset \mathbf{K}$ which is homomorphically closed and closed under extensions, is not necessarily a radical class i.e. condition (c) is necessary.

Theorem 7. *Let* $\mathbf{R}$ *be a radical class in* $\mathbf{K}$. *Then there exist*

(1) *a set* $Q(\mathbf{R})$ *of primes and for every* $p_i \in Q(\mathbf{R})$ *a radical class* $\mathbf{M}_{p_i}(\mathbf{R})$ *of strong artinian* p_i*-rings which does not contain non-zero finite nilpotent rings;*

(2) *a class* $\mathbf{H}(\mathbf{R})$ *of matrix rings over division rings containing all such rings which are in* $\mathbf{M}_{p_i}(\mathbf{R})$, $p_i \in Q(\mathbf{R})$; *such that* $A \in \mathbf{R}$ *iff* $A/\mathbf{J}(A)$ *is a finite direct sum of rings of* $\mathbf{H}(\mathbf{R})$ *and* $A(p_i) \in \mathbf{M}_{p_i}(\mathbf{R})$ *for every* $p_i \in Q(\mathbf{R})$.

Conversely, if Q *is a set of primes,* $\mathbf{M}_{p_i}$ *and* $\mathbf{H}$ *are classes as required in* (1) *and* (2), *respectively, then*

$$\mathbf{R}' = \left\{ A \in \mathbf{K} \,\middle|\, \begin{array}{l} A/\mathbf{J}(A) \text{ is a finite direct sum of rings} \\ \text{of } \mathbf{H} \text{ and } A(p_i) \in \mathbf{M}_{p_i},\ p_i \in Q \end{array} \right\}$$

is a radical class. Moreover, if $Q = Q(\mathbf{R})$, $\mathbf{M}_{p_i} = \mathbf{M}_{p_i}(\mathbf{R})$, $(p_i \in Q)$, $\mathbf{H} = \mathbf{H}(\mathbf{R})$, *then* $\mathbf{R}' = \mathbf{R}$.

Proof. Let $\mathbf{R}$ be a radical class in $\mathbf{K}$. Define $Q(\mathbf{R})$ by $p_i \in Q(\mathbf{R})$ iff $\mathbf{R}$ does not contain non-zero finite nilpotent p_i-rings, and define

$$\mathbf{H}(\mathbf{R}) = \left\{\begin{array}{l}\text{all matrix rings over division rings occurring as direct summands}\\ \text{in any factor ring } A/\mathbf{J}(A),\ A \in \mathbf{R}\end{array}\right\},$$

$$\mathbf{M}_{p_i}(\mathbf{R}) = \left\{\begin{array}{l}\text{all } p_i\text{-rings occurring as direct components in the}\\ \text{decomposition (D) of any ring } A \in \mathbf{R} \text{ for } p_i \in Q(\mathbf{R})\end{array}\right\}.$$

Since by Theorem 5 $\mathbf{R}$ is homomorphically closed and closed under extensions, so is each $\mathbf{M}_{p_i}(\mathbf{R})$, $p_i \in Q(\mathbf{R})$. The classes $\mathbf{M}_{p_i}(\mathbf{R})$ satisfy condition (c) of Theorem 5 trivially.

Conversely, if Q is a set of primes, $\mathbf{M}_{p_i}$, $p_i \in Q$, are radical classes of p_i-rings as demanded in (1), $\mathbf{H}$ is a class of matrix rings over division rings, then define $\mathbf{R}'$ as above. Since each $\mathbf{M}_{p_i}$ is homomorphically closed, also $\mathbf{R}'$ is homomorphically closed. Condition (c) is trivially fulfilled by the definition of $\mathbf{R}'$. Let $I \lhd A$ be such that I, $A/I \in \mathbf{R}'$. Since A is artinian,

$$\mathbf{J}(A/I) = (I+\mathbf{J}(A))/I$$

holds and consequently

$$A/(I+\mathbf{J}(A)) \cong A/I/((I+\mathbf{J}(A))/I) = A/I/(\mathbf{J}(A/I))$$

is a direct sum of rings of $\mathbf{H}$, since $A/I \in \mathbf{R}'$. $A/\mathbf{J}(A)$ contains $(I+\mathbf{J}(A))/\mathbf{J}(A) \cong I/\mathbf{J}(I)$ as a direct summand. The simple direct summands of $I/\mathbf{J}(I)$ are contained in $\mathbf{H}$, since $I \in \mathbf{R}'$. The other simple direct summands of $A/\mathbf{J}(A)$ are in $\mathbf{H}$, for

$$A/\mathbf{J}(A)/((I+\mathbf{J}(A))/\mathbf{J}(A)) \cong A/(I+\mathbf{J}(A)).$$

Moreover, the classes $\mathbf{M}_{p_i}$, $p_i \in Q$, are closed under extensions, therefore $\mathbf{R}'$ is closed under extensions, too. Hence Theorem 5 yields that $\mathbf{R}'$ is a radical class of $\mathbf{K}$.

The last statement is obvious.

In Theorem 7, analogously to Theorem 2, the radical classes are determined only up to radical classes of strong artinian p_i-rings. Nevertheless, the hereditary radical classes are fully described by

Theorem 8. *A hereditary radical class* $\mathbf{R}$ *in* $\mathbf{K}$ *is uniquely determined by a set* P *of primes and a class* $\mathbf{H}$ *of matrix rings over division rings in the following way:* $A \in \mathbf{R}$ *iff* $A/\mathbf{J}(A)$ *is a direct sum of rings of* $\mathbf{H}$, *and* $\mathbf{J}(A)$ *is a finite direct sum of nilpotent artinian* p_i*-rings for* $p_i \in P$.

The statement follows immediately from Theorems 4 and 5 and from Corollary 6.

In view of Theorems 2, 5 and Corollary 6 we have

Corollary 9. *If the subclass* $\mathbf{C}$ *of* $\mathbf{K}$ *is hereditary, closed under extensions and does not contain non-zero zerorings, then* $\mathbf{C}$ *is a homomorphically closed semisimple class as well as a hereditary radical class.*

Of course, *not every radical class in* $\mathbf{K}$ *is hereditary.* We give an example for a non-hereditary radical class of $\mathbf{K}$ which does not contain zero-rings $Z(p^\infty)$. Let

N denote the ring of integers, and p a fixed prime, and consider the class

$$\mathbf{R} = \{\text{finite direct sums of copies of } N/(p) \text{ and } N/(p^2)\}.$$

One can easily check that $\mathbf{R}$ is a radical class of $\mathbf{K}$, for $\mathbf{R}$ satisfies conditions (a), (b) and (c) of Theorem 5. On the other hand $\mathbf{R}$ is not hereditary, because $(p)/(p^2) \triangleleft N/(p^2)$ and $(p)/(p^2) \cong Z(p) \notin \mathbf{R}$ hold. This example is the special case $Q=\{p\}$, $\mathbf{M}_p=\mathbf{R}$, $\mathbf{H}=\{N/(p)\}$ of Theorem 7.

Let us mention that in view of Theorems 4 and 8 *homomorphically closed semisimple classes and hereditary radical classes of* $\mathbf{K}$ *are not necessarily subring-hereditary classes* (e.g. whenever $\mathbf{H}$ is not subring-hereditary). We recall the fact that in a ring variety the homomorphically closed semisimple classes are always subvarieties and hence subring-hereditary classes.

5. N-radicals in K

A radical $\mathbf{R}$ is called *left hereditary,* if every left ideal of an $\mathbf{R}$-ring is also in $\mathbf{R}$. A radical $\mathbf{R}$ is said to be a *left strong radical,* if the radical $\mathbf{R}(A)$ of any ring A contains every $\mathbf{R}$-left ideal of A. Following SANDS [4], a radical $\mathbf{N}$ is called an *N-radical,* if it is left hereditary, left strong and it contains every zero-ring.

Let $\mathbf{R}$ be a left strong radical containing all zero-rings, and containing a division ring D. Take the $n\times n$ matrix ring D_n over D. Denote by ε_{ij} the matrix having $1\in D$ at the i-th j-th place and 0 at every other place. Then $D_n=\sum_{i=1}^{n} D_n\varepsilon_{ii}$, further $\varepsilon_{ii}D_n\varepsilon_{ii}\cong \cong D$. Moreover, the mapping.

$$\varphi : D_n\varepsilon_{ii} \to \varepsilon_{ii}D\varepsilon_{ii} \cong D$$

defined by

$$\varphi(x\varepsilon_{ii}) = \varepsilon_{ii}x\varepsilon_{ii} \quad (x\in D_n)$$

is a ring homomorphism onto $\varepsilon_{ii}D_n\varepsilon_{ii}$. Hence

$$D \cong D_n\varepsilon_{ii}/\operatorname{Ker}\varphi$$

and

$$\operatorname{Ker}\varphi = \{x\varepsilon_{ii}\in D_n\varepsilon_{ii} \mid \varepsilon_{ii}x\varepsilon_{ii} = 0\}$$

is a zero-ring, since $(x\varepsilon_{ii})(y\varepsilon_{ii})=x(\varepsilon_{ii}y\varepsilon_{ii})=0$ holds for every $x\varepsilon_{ii}, y\varepsilon_{ii}\in\operatorname{Ker}\varphi$. Since $\mathbf{R}$ contains all zero-rings and $D\in\mathbf{R}$, so the extension property of $\mathbf{R}$ implies $D_n\varepsilon_{ii}\in\mathbf{R}$. Taking into account that $\mathbf{R}$ is a left strong radical, we get $D_n\in\mathbf{R}$. Hence we arrived at

Proposition 10. *Let* $\mathbf{R}$ *be a left strong radical containing all zero-rings. If a division ring* D *is contained in* $\mathbf{R}$, *then every matrix ring* D_n *over* D *is also in* $\mathbf{R}$.

Next, assume that $\mathbf{R}$ is a left hereditary radical and D_n is an $n\times n$ matrix ring over a division ring D. Suppose $D_n\in\mathbf{R}$. Since $D_n=\sum_{i=1}^{n} D_n\varepsilon_{ii}$, and $\mathbf{R}$ is left hereditary, we have $D_n\varepsilon_{ii} \in \mathbf{R}$. Consider again the homomorphism φ defined above. φ maps $D_n\varepsilon_{ii}$ onto $\varepsilon_{ii}D\varepsilon_{ii} \cong D$ homomorphically. Since $\mathbf{R}$ is homomorphically closed, we have $D \in \mathbf{R}$. Thus we obtained

Proposition 11. *Let* $\mathbf{R}$ *be a left hereditary radical. If a matrix ring* D_n *over a division ring* D *is in* $\mathbf{R}$, *then also* $D \in \mathbf{R}$ *holds.*

Applying Theorem 8 and Propositions 10 and 11 we obtain a full description of N-radicals in $\mathbf{K}$.

Theorem 12. *To any N-radical* $\mathbf{N}$ *of* $\mathbf{K}$ *there belongs a class* $\mathbf{D}$ *of all division rings* D *with* $D \in \mathbf{N}$. *Conversely, any class* $\mathbf{D}$ *of division rings determines an N-radical;* $\mathbf{N}$ *is the lower radical of all zero-rings of the class* $\mathbf{K}$ *and of all matrix rings over division ring from* $\mathbf{D}$.

Comparing Theorem 12 with the situation in the variety of associative rings, there is a remarkable difference. As has been shown in [6], if an N-radical $\mathbf{N}$ of associative rings contains the Brown—McCoy radical (the upper radical of all simple rings with unity), then $\mathbf{N}$ contains all division rings. The Brown—McCoy radical fails to be left hereditary, so it is not an N-radical. (In the category $\mathbf{K}$, the Brown—McCoy radical coincides with Baer's lower radical, and it is an N-radical.) The well-known N-radicals of associative rings are Baer's lower radical, the Levitzki and the Jacobson radical and some peculiar radicals constructed artificially.

Finally we give a characterization of the class of Jacobson radical rings (i.e. all nilpotent rings) in $\mathbf{K}$.

Corollary 13. *If* $\mathbf{J}$ *is a left hereditary radical class of* $\mathbf{K}$ *such that* $\mathbf{J}$ *contains* $Z(p)$ *for every prime* p *and* $\mathbf{J}$ *does not contain division rings, then* $\mathbf{J}$ *is the class of all nilpotent rings in* $\mathbf{K}$.

References

[1] E. Fried—R. Wiegandt, Connectednesses and disconnectednesses of graphs, *Algebra Universalis,* **5** (1975), 411—428.

[2] A. Kertész—A. Widiger, Artinsche Ringe mit artinschem Radikal, *J. reine angew. Math.,* **242** (1970), 8—15.

[3] L. C. A. van Leeuwen—C. Roos—R. Wiegandt, Characterizations of semisimple classes, *J. Austral. Math. Soc.,* **23** (1977), 172—182.

[4] A. D. Sands, Radicals and Morita contexts, *J. Algebra,* **24** (1973), 335—345.

[5] R. Wiegandt, *Radical and semisimple classes of rings,* Queen's papers in pure and appl. math. No. 37 (Kingston, 1974).

[6] R. Wiegandt, On *N*-radicals, *J. Nat. Sci. and Math. Lahore,* **13** (1973), 255—262.

ALFRED WIDIGER
MARTIN-LUTHER-UNIVERSITÄT
SEKTION MATHEMATIK
UNIVERSITÄTSPLATZ 6
DDR-401 HALLE

RICHARD WIEGANDT
MATHEMATICAL INSTITUTE
HUNGARIAN ACADEMY OF SCIENCES
REÁLTANODA U. 13–15.
H-1053 BUDAPEST

Acta Sci. Math., 39 (1977), 313—318

Complements of radicals in the class of hereditarily artinian rings

L. C. A. VAN LEEUWEN

1. Introduction

In [4], WIDIGER and WIEGANDT developed a theory of radicals for the class **K** of hereditarily artinian rings, i.e. the class **K** of all artinian rings with artinian Jacobson radical. It is remarked in their paper, that since **K** is not a variety, connections among algebraic properties are different from those in a variety. For instance, in **K** every hereditary radical class is a homomorphically closed semi-simple class, but the converse statement is not true. Other phenomena of this type are considered in this paper. It will be proved that any radical **R** in **K**, which contains **J** (the Jacobson radical) has a uniquely determined complement, which differs from the situation in a ring variety. This complement is a subidempotent radical (see [1], [2]). It is also shown that any hypernilpotent or subidempotent radical in **K** can be obtained as the upper radical, lower radical resp. of a suitable class of simple prime rings. The notation in this paper is that of [4]. For the definitions of radical class, semi-simple class etc. we refer to that paper and to [5].

I would like to thank A. WIDIGER and R. WIEGANDT for the preprint of their paper [4] and for their courtesy on permitting me to use results of [4].

2. Hypernilpotent radicals

Let **R** be a radical class in **K**, such that **R** contains all nilpotent rings in **K**. The class of nilpotent rings in **K** coincides with the Jacobson—radical class in **K**, so $\mathbf{R}\supseteq\mathbf{J}$, where **J** is the Jacobson radical. Then any **R**-semi-simple ring is a **J**-semi-simple ring. Since any ring in **K** is artinian, an **R**-semi-simple ring is a **J**-semi-simple

Received February 5, 1976.

artinian ring, hence it is a finite direct sum of matrix rings over division rings by the Wedderburn—Artin theorem.

Lemma 1. *$\mathbf{R}$ is a radical class in $\mathbf{K}$, such that $\mathbf{R} \supseteqq \mathbf{J}$. Let $\mathbf{T}=\{R \in \mathbf{K}: R$ is a simple $\mathbf{R}$-semi-simple ring$\}$. Then $\mathbf{R}=\mathscr{U}\mathbf{T}$, the upper radical determined by the class $\mathbf{T}$.*

Proof. Since all rings in $\mathbf{T}$ are $\mathbf{R}$-semi-simple, it is clear that $\mathbf{R} \subseteqq \mathscr{U}\mathbf{T}$. Now if $\mathbf{R} \in \mathscr{U}\mathbf{T}$, then R has no non-zero homomorphic image in $\mathbf{T}$. We claim that R has no non-zero homomorphic image in $\mathscr{S}\mathbf{R}$, the class of $\mathbf{R}$-semi-simple rings. Indeed, if $0 \neq R/I \in \mathscr{S}\mathbf{R}$, then R/I is a finite direct sum of simple rings, which must be $\mathbf{R}$-semi-simple, since $\mathscr{S}\mathbf{R}$ is hereditary. Hence R/I, and also R, can be mapped onto a non-zero ring in $\mathbf{T}$, which is impossible. Therefore, $R \in \mathscr{U}\mathscr{S}\mathbf{R}=\mathbf{R}$ and $\mathscr{U}\mathbf{T} \subseteqq \mathbf{R}$.

Lemma 2. *Let $\mathbf{R}$ and $\mathbf{T}$ be as in Lemma 1. Let $\mathbf{D}=\{R \in \mathbf{K}: R$ is a finite direct sum of rings in $\mathbf{T}\}$. Then $\mathbf{D}$ is the class of $\mathbf{R}$-semi-simple rings. Moreover, $\mathbf{D}$ is a radical class.*

Proof. First we show that $\mathbf{D}=\mathscr{S}\mathbf{R}$, the class of $\mathbf{R}$-semi-simple rings. Since each ring in $\mathbf{T}$ is $\mathbf{R}$-semi-simple, $\mathbf{D} \subseteqq \mathscr{S}\mathbf{R}$. Conversely, if $R \in \mathscr{S}\mathbf{R}$, then R is a finite direct sum of simple rings. Each of these simple rings is in $\mathbf{T}$, hence $R \in \mathbf{D}$, so $\mathscr{S}\mathbf{R} \subseteqq \mathbf{D}$ and $\mathbf{D}=\mathscr{S}\mathbf{R}$. Next we show that $\mathbf{D}$ is a radical class. If $R \in \mathbf{D}$, then $R=S_1 \oplus \dots$ $\dots \oplus S_k$, S_i simple ring in $\mathbf{D}$ ($i=1, \dots, k$). A homomorphic image of R has the same form, hence $\mathbf{D}$ is homomorphically closed. Also, if $R/I \in \mathbf{D}$ and $I \in \mathbf{D}$ for some ideal I of R, then I has a unity, hence it is a direct summand of R, say $R=I \oplus J$. Now $R/I \cong J \in \mathbf{D}$, so $I \oplus J \in \mathbf{D}$ or $R \in \mathbf{D}$. Then $\mathbf{D}$ is closed under extensions. This shows that $\mathbf{D}$ satisfies conditions (a) and (b) of theorem 5 ([4]). Condition (c) of that theorem is vacuous, since $Z(p) \notin \mathscr{S}\mathbf{R}=\mathbf{D}$ for a prime number p ($\mathbf{R}$ contains all nilpotent rings). So $\mathbf{D}$ is a radical class ([4]).

Remark 1. Lemma 1 and the first statement of lemma 2 can be proved without any assumption about the radical class $\mathbf{R}$. However, for an arbitrary radical class $\mathbf{R}$, the class $\mathbf{D}$ may fail to be a radical class. This is a consequence of the fact that a homomorphically closed semi-simple class $\mathbf{D}$ in $\mathbf{K}$ need not be a radical class. For a counterexample, see the remark after corollary 6 in [4].

Lemma 3. *Let $\mathbf{R}$, $\mathbf{T}$ and $\mathbf{D}$ be as in Lemmas 1 and 2. Then $\mathbf{D}$ is the complement of $\mathbf{R}$.*

Proof. Let $R \in \mathbf{D} \cap \mathbf{R}$. Then, as a ring of $\mathbf{D}$, R is a finite direct sum of rings in $\mathbf{T}$. On the other hand, $R \in \mathbf{R}=\mathscr{U}\mathbf{T}$ (Lemma 1), so R has no non-zero homomorphic rings in $\mathbf{T}$. Hence $R=(0)$.

Next let $R\in\mathscr{S}\mathbf{D}\cap\mathscr{S}\mathbf{R}$. Then $R\in\mathscr{S}\mathbf{R}$ implies that $R\in\mathbf{D}$ (Lemma 2), hence $R\in\mathscr{S}\mathbf{D}\cap\mathbf{D}=(0)$. Hence $\mathbf{D}$ is a complement of $\mathbf{R}$. Also, since any ring in $\mathbf{T}$ is a simple ring with unity and $\mathbf{R}=\mathscr{U}\mathbf{T}$, it follows that $\mathbf{R}$ is hereditary. The lattice of all hereditary radicals is distributive ([3], Cor. 16, p. 212), so within this lattice a complement of $\mathbf{R}$ is uniquely determined. Since $\mathbf{D}$ is a hereditary radical, $\mathbf{D}$ is the complement of $\mathbf{R}$. Summarizing our results we get

Theorem 1. *For any radical* $\mathbf{R}\supseteq\mathbf{J}$ *in the category* $\mathbf{K}$ *there exists a uniquely determined complement* $\mathbf{D}$*, where* $\mathbf{D}$ *is the class of all finite direct sums of all simple* $\mathbf{R}$*-semi-simple rings.* $\mathbf{D}$ *is also a semi-simple class, in fact* $\mathbf{D}$ *is the class of all* $\mathbf{R}$*-semi-simple rings. Moreover,* $\mathbf{R}$ *is the upper radical determined by the class of all simple* $\mathbf{R}$*-semi-simple rings.*

Remark 2. In [1] it is shown that for any hereditary radical $\mathbf{R}$ there exists a radical $\mathbf{R}'$ which is a complement of $\mathbf{R}$ and $\mathbf{R}'$ is the upper radical determined by the class of all subdirectly irreducible rings with $\mathbf{R}$-radical heart. The $\mathbf{R}$-radical rings are the strongly $\mathbf{R}$-semi-simple rings ([1], Theorem 2).

This result holds in the category of all associative rings. Our theorem 1 reveals that in the subcategory $\mathbf{K}$ a much stronger results holds. Not only is the complement $\mathbf{D}$ of $\mathbf{R}$ uniquely determined, but $\mathbf{D}$ is also a semi-simple class ($\mathbf{R}\supseteq\mathbf{J}$), i.e. the class of $\mathbf{R}$-semi-simple rings. Since the class $\mathbf{D}$ is homomorphically closed, the strongly $\mathbf{R}$-semi-simple rings are all semi-simple rings. The complement $\mathbf{R}$ of $\mathbf{D}$ is the upper radical determined by the class $\mathbf{T}$ of simple $\mathbf{R}$-semi-simple rings (=simple $\mathbf{D}$-radical rings). This class $\mathbf{T}$ of simple rings is, in general, a subclass of the class of all subdirectly irreducible rings with $\mathbf{D}$-radical hearts. However, they determine the same upper radical $\mathbf{R}$.

Examples. 1. Let $\mathbf{R}=\mathbf{J}$, the Jacobson radical. Then $\mathbf{D}$ is the class of all finite direct sums of simple $\mathbf{J}$-semi-simple rings i.e. finite direct sums of all matrix rings over division rings.

2. Let $\mathbf{R}$ be the class of all strong artinian rings, i.e. all rings $R\in\mathbf{K}$, where $(R,+)$ has d.c.c. for subgroups. It can easily be seen that $\mathbf{R}$ is a radical class, which we call $\mathbf{R}_S$. The complement of $\mathbf{R}_S$ is the class of finite direct sums of simple $\mathbf{R}_S$-semi-simple rings, i.e. finite direct sums of all matrix rings over infinite division rings.

3. Let $\mathbf{R}$ be the class of all torsion radical rings, i.e. all rings $R\in\mathbf{K}$ where $(R,+)$ is a torsion group. This is a radical class, which we call $\mathbf{R}_T$. The complement of $\mathbf{R}_T$ is the class of finite direct sums of simple $\mathbf{R}_T$-semi-simple rings, i.e. finite direct sums of simple torsion-free rings. These simple torsion-free rings are matrix rings over (infinite) torsion-free division rings.

Remark 3. Any radical $\mathbf{R}\supseteq\mathbf{J}$ is hypernilpotent i.e. $\mathbf{R}$ contains all nilpotent rings and $\mathbf{R}$ is hereditary, (Lemma 1). By corollary 6 ([4]), $\mathbf{R}$ is a homomorphically closed semi-simple class. Let $\mathbf{D}$ be the complement of $\mathbf{R}$, then $\mathbf{D}=\mathscr{S}\mathbf{R}$ (Lemma 2). One might conjecture, that $\mathbf{R}=\mathscr{S}\mathbf{D}$. This is not true in general. However, $\mathbf{R}\subseteqq\mathscr{S}\mathbf{D}$. Indeed, if $R\in\mathbf{R}$ then $R\in\mathscr{U}\mathbf{T}$, so R has no non-zero homomorphic image in $\mathbf{T}$. If $\mathbf{D}(R)=S_1\oplus\ldots\oplus S_k\neq 0$, then $\mathbf{D}(R)$ is a non-zero ideal in R and $\mathbf{D}(R)$ has a unity, so it is a direct summand of R. At least one of the $S_i\neq 0$, say $S_j\neq 0$, and S_j is a direct summand of R. Now $S_j\in\mathbf{T}$ and R could be mapped homomorphically onto $0\neq S_j\in\mathbf{T}$, which is a contradiction. Hence $\mathbf{D}(R)=0$, and $R\in\mathscr{S}\mathbf{D}$. That $\mathbf{R}\neq\mathscr{S}\mathbf{D}$ may be seen by taking $\mathbf{R}=\mathbf{J}$. The ring Z_4 of integers mod 4 is not $\mathbf{J}$-radical, so $Z_4\notin\mathbf{R}$. If $\mathbf{D}$ is the complement of $\mathbf{J}$, then $Z_4\in\mathscr{S}\mathbf{D}$, however.

3. Subidempotent radicals

Definition. A ring R will be called *hereditarily idempotent* if every ideal of R is idempotent. A hereditary radical $\mathbf{R}$ will be called a subidempotent radical if $\mathbf{R}$-radical rings are hereditarily idempotent rings (cf. [1]).

Examples. In the category $\mathbf{K}$ the complements of hypernilpotent radicals are subidempotent.

Lemma 4. *$\mathbf{E}$ is a subidempotent radical. Let $\mathbf{P}=\{R\in\mathbf{K}: R$ is a simple $\mathbf{E}$-radical ring$\}$. Then $\mathbf{E}=\mathscr{L}\mathbf{P}$ the lower radical determined by the class $\mathbf{P}$.*

Proof. Since every ring in $\mathbf{P}$ is $\mathbf{E}$-radical, it is clear that $\mathscr{L}\mathbf{P}\subseteqq\mathbf{E}$. Next let $R\in\mathbf{E}$. Then R is a hereditarily idempotent ring. Hence any ideal of R is idempotent. However $\mathbf{J}(R)$ is nilpotent, so $\mathbf{J}(R)=(0)$. Then R is a finite direct sum of matrix rings over division rings. Each of the direct summands is a simple ring and, since $\mathbf{E}$ is hereditary, a simple $\mathbf{E}$-radical ring. A non-zero homomorphic image of R is in $\mathbf{E}$ since $\mathbf{E}$ is homomorphically closed. Such an image is again a finite direct sum of simple $\mathbf{E}$-radical rings, hence it has a non-zero ideal in $\mathbf{P}$. Then $R\in\mathbf{P}_2$. Since $\mathbf{P}$ is a homomorphically closed class of idempotent rings, $\mathscr{L}\mathbf{P}=\mathbf{P}_2$ ([5], Corollary 12.6), so $R\in\mathscr{L}\mathbf{P}$. Therefore $\mathbf{E}\subseteqq\mathscr{L}\mathbf{P}$.

Lemma 5. *Let $\mathbf{E}$ and $\mathbf{P}$ be as in Lemma 4. Let $\mathbf{Q}=\{R\in\mathbf{K}: R$ is a finite direct sum of rings in $\mathbf{P}\}$. Then $\mathbf{Q}$ is the class of $\mathbf{E}$-radical rings. Moreover, $\mathbf{Q}=\mathbf{E}$ is a semi-simple class, in fact, $\mathbf{Q}$ is the class of $\mathscr{U}\mathbf{P}$-semi-simple rings.*

Proof. Since every ring in $\mathbf{P}$ is $\mathbf{E}$-radical, it is clear that any ring in $\mathbf{Q}$ is $\mathbf{E}$-radical. From the proof of Lemma 4 it follows that if $R\in\mathbf{E}$, R is a finite direct sum of simple $\mathbf{E}$-radical rings, i.e. $R\in\mathbf{Q}$. This shows that $\mathbf{Q}=\mathbf{E}$.

Since $\mathbf{Q}=\mathbf{E}$ is hereditary and closed under extensions, it follows that $\mathbf{Q}$ is a semi-simple class ([4], Theorem 1). Now we show that $\mathbf{Q}$ is the class of $\mathscr{U}\mathbf{P}$-semi-simple rings. Let $R \in \mathbf{Q}$ then R is a finite direct sum of rings in $\mathbf{P}$, each of which is $\mathscr{U}\mathbf{P}$-semi-simple, hence R is $\mathscr{U}\mathbf{P}$-semi-simple. Conversely, assume that R is $\mathscr{U}\mathbf{P}$-semi-simple. Any ring in $\mathbf{P}$ is a simple prime ring, hence a $\mathbf{J}$-semi-simple ring, so $\mathbf{P} \subseteq \mathscr{S}\mathbf{J}$, which implies $\mathscr{U}\mathscr{S}\mathbf{J} \subseteq \mathscr{U}\mathbf{P}$ or $\mathbf{J} \subseteq \mathscr{U}\mathbf{P}$. Then R is a $\mathbf{J}$-semi-simple ring and a finite direct sum of simple $\mathbf{J}$-semi-simple rings, i.e. simple $\mathscr{U}\mathbf{P}$-semi-simple rings. But a simple $\mathscr{U}\mathbf{P}$-semi-simple ring is a simple ring in $\mathbf{P}$. Hence R is a finite direct sum of rings in $\mathbf{P}$ or $R \in \mathbf{Q}$. Therefore $\mathbf{Q}$ is the class of $\mathscr{U}\mathbf{P}$-semi-simple rings.

Lemma 6. *Let* $\mathbf{E}$, $\mathbf{P}$ *and* $\mathbf{Q}$ *be as in Lemmas 4 and 5. Then* $\mathscr{U}\mathbf{P}$ *is the complement of* $\mathbf{E}$.

Proof. Let $R \in \mathbf{E} \cap \mathscr{U}\mathbf{P}$. Then $R \in \mathbf{Q}$ (Lemma 5), so R is a finite direct sum of rings in $\mathbf{P}$. But $R \in \mathscr{U}\mathbf{P}$ implies that R cannot be mapped homomorphically onto a non-zero ring in $\mathbf{P}$. Hence $R=(0)$. Next, let $R \in \mathscr{S}\mathbf{E} \cap \mathscr{S}\mathscr{U}\mathbf{P}$. Since $\mathbf{Q}=\mathscr{S}\mathscr{U}\mathbf{P}$ (Lemma 5), it follows that $R \in \mathbf{Q}$. Also $\mathbf{E}=\mathbf{Q}$, so $R \in \mathscr{S}\mathbf{Q}$. Then $R \in \mathbf{Q} \cap \mathscr{S}\mathbf{Q}$ implies $R=(0)$.

This shows that $\mathscr{U}\mathbf{P}$ is a complement of $\mathbf{E}$. Each ring in $\mathbf{P}$ is a simple $\mathbf{E}$-radical ring and a simple $\mathbf{J}$-semi-simple ring (proof of Lemma 4). So such a ring is a simple ring with unity and $\mathscr{U}\mathbf{P}$ is a hereditary radical. It follows that $\mathscr{U}\mathbf{P}$ is the complement of $\mathbf{E}$. In the proof of Lemma 5 we have seen that $\mathbf{J} \subseteq \mathscr{U}\mathbf{P}$, so $\mathscr{U}\mathbf{P}$ is a hypernilpotent radical. Summarizing the results we get

Theorem 2. *Let* $\mathbf{E}$ *be an arbitrary subidempotent radical in the category* $\mathbf{K}$. *Then* $\mathbf{E}=\mathscr{L}\mathbf{P}$, *where* $\mathbf{P}$ *is the class of simple* $\mathbf{E}$*-radical rings. Any ring in* $\mathbf{E}$ *is a finite direct sum of rings in* $\mathbf{P}$. *Also* $\mathbf{E}$ *is a semi-simple class, i.e. the class of* $\mathscr{U}\mathbf{P}$*-semi-simple rings. The radical* $\mathscr{U}\mathbf{P}$ *is hypernilpotent and the complement of* $\mathbf{E}$.

Remark 4. It can easily be seen that using the notation of Lemmas 1, 2 and 3, the complement $\mathbf{D}$ of $\mathbf{R}$ equals $\mathscr{L}\mathbf{T}$, the lower radical determined by $\mathbf{T}$. Indeed, $\mathbf{D}$ is a subidempotent radical and $\mathbf{T}$ is the class of simple $\mathbf{R}$-semi-simple rings i.e. simple $\mathbf{D}$-radical rings (Lemma 2). Now apply Lemma 4.

By theorem 2 of [1] the class $\mathbf{D}$ can also be characterized as the upper radical determined by the class of all subdirectly irreducible rings with $\mathbf{R}$-radical hearts.

Comparing our results with those of theorem 4 of [1] it turns out that, contrary to the general situation in the category of associative rings, any radical $\mathbf{R} \supseteq \mathbf{J}$ is a dual radical, i.e. the complement of $\mathbf{D}$ is $\mathbf{R}$, if $\mathbf{D}$ is the complement of $\mathbf{R}$. Here $\mathbf{R}$ is a dual hypernilpotent radical, while $\mathbf{D}$ is a dual subidempotent radical.

The radical $\mathbf{R}$ resp. $\mathbf{D}$ is the upper radical resp. lower radical determined by the same class $\mathbf{T}$, i.e. the class $\mathbf{T}$ of simple $\mathbf{R}$-semi-simple rings or simple $\mathbf{D}$-radical rings. In the next section we investigate radicals, determined by a class of simple prime rings.

4. Simple prime rings in K

Let $\mathbf{M}$ be an arbitrary non-empty class of simple prime rings in $\mathbf{K}$. Then $\mathbf{M}$ is a class of simple rings with unity. Let $\mathbf{Q}=\{R\in\mathbf{K}: R$ is a finite direct sum of rings from $\mathbf{M}\}$. Then $\mathbf{Q}$ is homomorphically closed, closed under extensions and has no non-zero nilpotent rings. Hence $\mathbf{Q}$ is a radical class in $\mathbf{K}$ ([4] Theorem 5). All-rings in $\mathbf{M}$ are $\mathbf{Q}$-radical, hence $\mathscr{L}\mathbf{M}\subseteqq\mathbf{Q}$. But if $R\in\mathbf{Q}$, then $0\neq R/I\in\mathbf{Q}$ for any ideal I in R, so R/I has a non-zero ideal in $\mathbf{M}$ which implies $R\in\mathbf{M}_2$. Since $\mathbf{M}$ is a class of idempotent rings $\mathscr{L}\mathbf{M}=\mathbf{M}_2$, hence $R\in\mathscr{L}\mathbf{M}$. Therefore $\mathbf{Q}=\mathscr{L}\mathbf{M}$.

Since $\mathbf{Q}$ is hereditary and closed under extensions, $\mathbf{Q}$ is a semi-simple class ([4], Theorem 1). From the proof of Lemma 5 it follows that $\mathbf{Q}=\mathscr{S}\mathscr{U}\mathbf{M}$. Also both $\mathscr{U}\mathbf{M}$ and $\mathscr{L}\mathbf{M}$ are hereditary radicals, since $\mathbf{M}$ is a hereditary class. From $\mathbf{Q}=\mathscr{L}\mathbf{M}=\mathscr{S}\mathscr{U}\mathbf{M}$ it follows directly that $\mathscr{U}\mathbf{M}$ and $\mathscr{L}\mathbf{M}$ are complements.
This shows:

Theorem 3. *Let $\mathbf{M}$ be an arbitrary non-empty class of simple prime rings in $\mathbf{K}$. Then both $\mathscr{U}\mathbf{M}$ and $\mathscr{L}\mathbf{M}$ are hereditary radicals, where $\mathscr{U}\mathbf{M}$ is hypernilpotent and $\mathscr{L}\mathbf{M}$ is subidempotent. In addition, $\mathscr{L}\mathbf{M}=\mathscr{S}\mathscr{U}\mathbf{M}$ and $\mathscr{L}\mathbf{M}$ and $\mathscr{U}\mathbf{M}$ are complementary radicals.*

From Lemmas 1 and 4 it follows that any hypernilpotent (subidempotent) radical $\mathbf{R}(\mathbf{E})$ is the upper radical (lower radical), determined by a class $\mathbf{T}(\mathbf{P})$ of simple prime rings.

Remark 5. Finally we want to compare our results with theorem 10 of [1], the so-called duality theorem for radicals. It is said there that all dual hypernilpotent and dual subidempotent radicals can be obtained both as upper radicals determined by certain classes of subdirectly irreducible rings with idempotent hearts. In our case any hypernilpotent or subidempotent radical is dual and the hypernilpotent radicals are upper radicals, while the subidempotent ones are lower radicals. Both are determined by classes of simple prime rings, which are matrix rings over division rings.

References

[1] V. A. Andrunakievitch, Radicals of associative rings I, *Mat. Sbornik*, **44** (1958), 179—212 (Russian).

[2] V. A. Andrunakievitch, Radicals of associative rings II, *Mat. Sbornik*, **55** (1961), 329—346 (Russian).

[3] R. L. Snider, Lattices of radicals, *Pac. J. Math.* **40** (1972), 207—220.

[4] A. Widiger and R. Wiegandt, Theory of radicals for hereditarily artinian rings (preprint).

[5] R. Wiegandt, Radical and semi-simple classes of rings, *Queen's papers in pure and applied mathematics*, No. 37 (Kingston, Ont., Canada 1974).

MATHEMATICS INSTITUTE
UNIVERSITY OF GRONINGEN
GRONINGEN, G.P.O. 800
THE NETHERLANDS

Acta Sci. Math., 39 (1977), 319—328

Change of the sum of digits by multiplication

I. KÁTAI

1. Introduction

Let N be a natural number and set $\mathscr{A}_N=\{0, 1, ..., 2^N-1\}$. Every $n\in\mathscr{A}_N$ can be written in the form

$$n = \sum_{i=0}^{N-1} \varepsilon_i \cdot 2^i, \tag{1.1}$$

where $\varepsilon_i=0$ or 1 ($i=0, 1, ..., N-1$). This representation is unique. Let $\alpha(n)$ denote the sum of the digits of n, i.e.

$$\alpha(n) = \sum_{i=0}^{N-1} \varepsilon_i. \tag{1.2}$$

Let $M_N(x)$ denote the number of those $n\in\mathscr{A}_N$ for which

$$\frac{\alpha(n)-N/2}{\sqrt{N}/2} < x.$$

Using the central limit theorem of probability theory in the simplest form, we have that

$$2^{-N}M_N(x) \to \Phi(x) \quad (N\to\infty)$$

for every real x, where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\,dt. \tag{1.3}$$

Furthermore, we have

$$2^{-N} \sum_{n\in A_N} \left(\alpha(n)-\frac{N}{2}\right)^2 = \frac{N}{4}. \tag{1.4}$$

It seems to be interesting to consider the distribution of the difference $\alpha(hn)-\alpha(n)$, $n\in\mathscr{A}_N$ for fixed h. This question is trivial for $h=2$, since $\alpha(2n)-\alpha(n)=0$.

Received October 25, 1976.

We shall be dealing with the case $h=3$. Let

$$\Delta(n) = \alpha(3n)-\alpha(n). \tag{1.5}$$

Our main result is the following

Theorem 1. *Let $K_N(x)$ denote the number of $n\in\mathcal{A}_N$ for which*

$$\frac{\sqrt{3}\,\Delta(n)}{\sqrt{N}} < x. \tag{1.6}$$

Then, for every real number x, we have

$$2^{-N}K_N(x) \to \Phi(x) \quad (N\to\infty). \tag{1.7}$$

We can deduce a more precise result, with a remainder term, but now we do not try to give the best one.

This and similar results may have some importance in the probabilistic treatment of rounding errors in numerical analysis.

2. The splitting of the binary representation of integers

We define the sets $\mathfrak{M}_k$ as follows. Let $\mathfrak{M}_0=\{0\}$. The sets $\mathfrak{M}_k$ contain those integers m_k for which $2^{k-1}\leqq m_k<2^k$ and the binary representation of which does not contain two consecutive zeros. Let m_k denote a general element of $\mathfrak{M}_k$, and A_k the number of its elements. It is obvious that $A_0=1$, $A_1=1$, $A_2=2$. We shall show that

$$A_k = A_{k-1}+A_{k-2} \quad (k\geqq 2). \tag{2.1}$$

Indeed, m_k can be written as

$$m_k = 2^{k-1}+m_{k-1} \quad \text{or} \quad m_k = 2^{k-1}+m_{k-2},$$

whence (2.1) immediately follows.

Let

$$F(z) = \sum_{k=0}^{\infty} A_k z^{k+2}. \tag{2.2}$$

By an easy calculation we get

$$F(z) = \frac{z^2}{1-z-z^2}. \tag{2.3}$$

Let $\mathfrak{M}=\sum_{k=0}^{\infty}\mathfrak{M}_k$. Assume that $N\geqq 2$. Then for every $n\in A_N$ there exists a unique element $m_{l_1}\in\mathfrak{M}$ for which

$$n = m_{l_1}+2^{l_1+2}u, \quad u\in A_{N-l_1-2}. \tag{2.4}$$

For $N=0$ or 1 we use the representation $n=m_{l_1}$. Repeating this, we get

(2.5) $$n = m_{l_1}+2^{l_1+2}\{m_{l_2}+2^{l_2+2}\{\dots\{m_{l_{\nu(n)-1}}+2^{l_{\nu(n)-1}+2}\{m_{l_{\nu(n)}}\}\}\dots\}\}.$$

So, for every n we order the sequence of elements of $\mathfrak{M}$. It is obvious that

(2.6) $$l_1+\dots+l_{\nu(n)}+2(\nu(m)-1) = N-1 \quad \text{or} \quad N.$$

Furthermore, from (2.4) we have

$$3n = 3m_{l_1}+2^{l_1+2}\cdot(3u), \quad 3m_{l_1} < 2^{l_1+2},$$

and so

(2.7) $$\Delta(n) = \Delta(m_{l_1})+\Delta(n).$$

Hence we have

(2.8) $$\Delta(n) = \sum_{j=1}^{\nu(n)} \Delta(m_{l_j}).$$

3. The distribution of the number of $\mathfrak{M}$-components

Now we consider the number of those integers $n\in\mathscr{A}_N$ for which $\nu(n)=H$. For the sake of brevity we use the notation

(3.1) $$t_j = l_j+2.$$

So we write (2.6) in the form

(3.2) $$t_1+\dots+t_H = N+2-\delta, \quad \delta = 0 \quad \text{or} \quad 1.$$

Let $2^N\beta_H(\delta)$ denote the number of $n\in\mathscr{A}_N$ for which $\nu(n)=H$ and (3.2) holds. Since

$$\int_{-1/2}^{1/2} e^{2\pi\theta}\,d\theta = \begin{cases} 1 & \text{if} \quad n=0 \\ 0 & \text{if} \quad n\neq 0, \quad \text{integer,} \end{cases}$$

we get

(3.3) $$\beta_H(\delta) = 2^{2-\delta}\int_{-1/2}^{1/2} F\left(\frac{z}{2}\right)^H \cdot z^{-(N+2-\delta)}\,d\theta, \quad z=e^{2\pi i\theta}.$$

First we integrate in the neighbourhood of $\theta=0$. By taking $\omega=z-1$, $|\omega|\leqq\frac{1}{2}$, we have

$$\ln F\left(\frac{z}{2}\right) = \ln\frac{z^2}{4-2z-z^2} = \ln\frac{\omega^2+2\omega+1}{1-4\omega-\omega^2} = 6\omega+8\omega^2+O(\omega^3),$$

$$\ln z = \ln(1+\omega) = \omega-\frac{\omega^2}{2}+O(\omega^3).$$

Let

$$A_H = 6H-(N+2-\delta), \quad B_H = 8H+\frac{N+2-\delta}{2}. \tag{3.4}$$

Then

$$g(z) \overset{\text{def}}{=} F\left(\frac{z}{2}\right)^H z^{-(N+2-\delta)} = \exp\left(A_H\omega+B_H\omega^2+O(N\omega^3)\right).$$

Observing that $\omega=2\pi i\theta-4\pi^2\theta^2+0(\theta^3)$, $\omega^2=-4\pi^2\theta^2+O(\theta^3)$, we get

$$g(z) = \exp\left(2\pi i\cdot A_H\theta-4\pi^2(A_H+\beta_H)\theta^2\right)\exp\left(O(N\theta^3)\right). \tag{3.5}$$

Let

$$\mathcal{I}_1 = \int_{-\Lambda}^{\Lambda} g(z)\,d\theta, \quad \mathcal{I}_2 = \int_{\Lambda<|\theta|\le\frac{1}{2}} g(z)\,d\theta, \tag{3.6}$$

where we choose Λ so that $N\Lambda^2\to\infty$, $N\Lambda^3\to 0$.

From (3.5) we get

$$\mathcal{I}_1 = \mathcal{I}_3+O(K), \tag{3.7}$$

where

$$\mathcal{I}_3 = \int_{-\Lambda}^{\Lambda} \exp\left(2\pi i A_H\theta-4\pi^2(A_H+B_H)\theta^2\right)d\theta, \tag{3.8}$$

$$K = \int_{-\Lambda}^{\Lambda} \exp\left(-4\pi^2(A_H+B_H\theta^2\right)N\theta^3\,d\theta. \tag{3.9}$$

In what follows we assume that $|A_H|\ll N^{2/3}$. Then $H=\frac{N}{6}+O(N^{2/3})$, and so

$$B_H = \frac{11}{6}N+O(N^{2/3}).$$

Hence for $\mathcal{K}$ we easily get that

$$\mathcal{K}\ll\frac{N}{B_H^2}\ll\frac{1}{N}. \tag{3.10}$$

To estimate $\mathcal{I}_3$, we use the following

Lemma 1. *Let*

$$J(A, B, \Lambda) = \int_{-\Lambda}^{\Lambda} \exp(iA\varphi-B\varphi^2)\,d\varphi,$$

A, B, Λ real numbers, $B>0$, $\Lambda>0$. Then

$$J(A, B, \Lambda) = \exp\left(-\frac{A^2}{4B}\right)\cdot\sqrt{\frac{\pi}{B}}+O\left(B^{-3/2}|A|\exp(-\Lambda^2B)\right)+$$

$$+O\left(B^{-1/2}\exp\left(-\frac{A^2}{4B}-\Lambda^2B\right)\right).$$

Proof.

$$J(A, B, \Lambda) = B^{-1/2} \int_{-\Lambda\sqrt{B}}^{\Lambda\sqrt{B}} \exp\left(i\frac{A}{\sqrt{B}}\tau - \tau^2\right) d\tau =$$

$$= \frac{\exp\left(-\frac{A^2}{4B}\right)}{i\sqrt{B}} \int_{-\Lambda\sqrt{B}}^{\Lambda\sqrt{B}} \exp\left(\left(\frac{A}{2\sqrt{B}} + i\tau\right)^2\right) d\left(\tau i + \frac{A}{2\sqrt{B}}\right) = \frac{\exp\left(-\frac{A^2}{4B}\right)}{i\sqrt{B}} \int_L e^{z^2} dz,$$

where L denotes the segment $\left[\frac{A}{2\sqrt{B}} - \Lambda\sqrt{B}, \frac{A}{2\sqrt{B}} + \Lambda\sqrt{B}\right]$. Transforming the integral to the imaginary axis, by an easy estimation we get the desired result.

Applying this Lemma by choosing $A = 2\pi A_H$, $B = 4\pi(A_H + B_H)$, $\Lambda = N^{-1/3-\varepsilon}$ ($\varepsilon > 0$), we get

$$\mathscr{I}_1 = \mathscr{I}_3 + O(\mathscr{K}) = \frac{\exp\left(-\frac{A_H^2}{4(A_H+B_H)}\right)}{2\sqrt{\pi}\sqrt{A_H+B_H}} + O(1/N). \tag{3.11}$$

Now we estimate $\mathscr{I}_2$. By taking $Y = \cos 2\pi\,\theta$, $Y = 1 - t$, we get

$$|4 - 2z - z^2|^2 = 1 + 44t - 16t^2.$$

So in $\Lambda \leqq |\theta| \leqq \frac{1}{2}$ we get

$$|4 - 2z - z^2|^{-H} \leqq \left(1 + 44\pi^2\Lambda^2(1-\varepsilon)\right)^{-H/2} \leqq \exp(-c_1 H\Lambda^2),$$

$c_1 > 0$ constant. Consequently

$$\mathscr{I}_2 \ll \exp(-c_1 N\Lambda^2). \tag{3.12}$$

Finally, taking into account (3.3), (3.6), (3.7), (3.10), (3.11), (3.12), we get

$$\beta_H \overset{\text{def}}{=} \beta_H(0) + \beta_H(1) = \frac{1}{\sigma\sqrt{2\pi N}} \exp\left(-\frac{(H - N/6)^2}{2\sigma^2 N}\right) + O(1/N),$$

where $\sigma = \frac{1}{6} \cdot \sqrt{\frac{11}{3}}$.

So we have proved:

Theorem 2. *Let $2^N\beta_H$ denote the number of those $n \in \mathscr{A}_N$ for which $v(n) = H$. Then*

$$\beta_H = \frac{1}{\sigma\sqrt{2\pi N}} \exp\left(-\frac{(H - N/6)^2}{2\sigma^2 N}\right) + O(1/N)$$

where $\sigma = \frac{1}{6} \cdot \sqrt{\frac{11}{3}}$, uniformly in H.

4. The function $H(\tau)$

Let

$$S_l(\tau) = \sum_{m_l \in \mathfrak{M}_l} e^{i\tau\Delta(m_l)}, \quad H(\tau) + \sum_{l=0}^{\infty} 2^{-l-2} S_l(\tau). \tag{4.1}$$

It is obvious that $S_0(\tau)=1$, $S_1(\tau)=e^{i\tau}$, $S_2(\tau)=1+e^{i\tau}$. Let $\mathfrak{M}_k=\mathfrak{M}_k^1 \cup \mathfrak{M}_k^0$, where $\mathfrak{M}_k^0$ contains the even numbers of $\mathfrak{M}_k$, and $\mathfrak{M}_k^1$ the odd numbers. Let

$$S_k^{(i)}(\tau) = \sum_{m_k \in \mathfrak{M}_k^i} e^{i\tau\Delta(m_k)}.$$

First of all we observe that

$$S_k^{(0)}(\tau) = S_{k-1}^{(1)}(\tau)$$

for $k \geqq 2$. Then

$$H(\tau) = \frac{1}{4} + \frac{3e^{i\tau}}{16} + \frac{3}{8} \cdot \sum_{l=2}^{\infty} \frac{S_l^{(1)}(\tau)}{2^l}. \tag{4.2}$$

Now we compute $S_k^{(1)}(\tau)$. The general form of the binary representation of n_A is the following one:

$$n = \overbrace{\boxed{}}^{A}\;\underbrace{\boxed{0\,|\,1\,|\;\dots\;|\,0\,|\,1}}_{2h \text{ places}},$$

where A is one of the following types:

1) $A = \boxed{1}$

2) $A = \boxed{1\,|\,1}$

3) $A = \boxed{1\,||\,1\,|\,1}$

Case 1) holds for odd k only. If $k=2t+1$ and $m_k=101\dots01$, then, obviously

$$\Delta(101\dots01) = t+1. \tag{4.3}$$

In the other cases we say that $m_k \in \mathfrak{M}_k^1$ is of $\mathscr{B}_{h,r}$ type if in A there exist exactly r zeros. In case 2) $k-2h=2$ (k even, $r=0$). In case 3) $k-2h \geqq 3$. It is easy to see that the number of elements of type $\mathscr{B}_{h,r}$ is

$$\binom{k-2h-r-2}{r}.$$

We observe that for $m_k \in \mathcal{B}_{h,r}$, $\Delta(m_k) = h - r$. We consider $3m_k$ as the sum of $2m_k$ and m_k. See the following figure.

m_k			1	0…	…	0	… 1 1	0 1 0 1 … 0 1
$2m_k$		1				1	… 1 0	1 0 1 0 … 1 0
$3m_k$							1	1 ………… 1 1
$\Delta(m_k)$		0		-1…	…		-1	$+h$

Now we take

$$\sum_{l=2}^{\infty} 2^{-l} S_l^{(1)}(\tau) = \sum_{t=1}^{\infty} 2^{-2t} S_{2t}^{(1)}(\tau) + \sum_{t=1}^{\infty} 2^{-2t-1} S_{2t+1}^{(1)}(\tau) = \Sigma_A + \Sigma_B.$$

We have

$$\Sigma_A =$$

$$= \sum_{t=1}^{\infty} 2^{-2t} \cdot \sum_{h,r} \binom{2t-2h-r-2}{r} e^{i\tau(h-r)} = \sum_{r=0}^{\infty} \sum_{h=0}^{\infty} e^{i\tau(h-r)} \sum_{v=1}^{\infty} 2^{-2(h+r+v)} \cdot \binom{r+2v-2}{r} =$$

$$= \left(\sum_{h=0}^{\infty} 2^{-2h} e^{i\tau h}\right) \cdot \frac{1}{4} \left\{ \sum_{r=0}^{\infty} 2^{-2r} e^{-i\tau r} \cdot \sum_{v=0}^{\infty} \binom{r+2v}{r} \cdot 2^{-2v} \right\} = \Sigma_0 \cdot \frac{1}{4} \cdot \{\Sigma_1\}.$$

We observe that

$$\Sigma_1 = \sum_{s=0}^{\infty} \sum_{2v \leqq r} \binom{s}{2v} \left(\frac{1}{2}\right)^{2v} \cdot \left(\frac{e^{-i\tau}}{2^2}\right)^r = \sum_{s=0}^{\infty} \frac{1}{2} \left\{ \left(\frac{1}{2} + \frac{e^{-i\tau}}{4}\right)^s + \left(-\frac{1}{2} + \frac{e^{-i\tau}}{4}\right)^s \right\} =$$

$$= \frac{2}{2-e^{-i\tau}} + \frac{2}{3-e^{-i\tau}}.$$

Furthermore,

$$\Sigma_0 = \frac{1}{1-e^{i\tau}/4}.$$

So we have

$$\Sigma_A = \frac{1}{4-e^{i\tau}} \left\{ \frac{2}{2-e^{-i\tau}} + \frac{2}{3-e^{-i\tau}} \right\}.$$

In the sum Σ_B the extraordinary case (4.3) occurs. We get

$$\Sigma_B = \Sigma_E + \Sigma_C,$$

where

$$\Sigma_E = \sum_{t=0}^{\infty} 2^{-2t-1} \cdot e^{i\tau(t+1)} = \frac{e^{2i\tau}}{2(4-e^{i\tau})}, \quad \Sigma_C = \sum_{h,r} e^{i\tau(h-r)}$$

We hawe, similarly as for Σ_A,

$$\Sigma_C = \frac{1}{4-e^{i\tau}} \left\{ \frac{2}{2-e^{-i\tau}} - \frac{2}{3-e^{-i\tau}} \right\}.$$

Summing up, we have

$$\sum_{l=2}^{\infty} 2^{-l} S_l^{(1)}(\tau) = \frac{1}{4-e^{i\tau}} \left\{ \frac{e^{2i\tau}}{2} + \frac{4}{2-e^{-i\tau}} \right\}.$$

So we have from (4.2) that

$$H(\tau) = \frac{1}{4} + \frac{3}{16} \cdot e^{i\tau} + \frac{3}{16} \cdot \frac{e^{2i\tau}}{4-e^{i\tau}} + \frac{3}{2} \cdot \frac{1}{(2-e^{-i\tau})(4-e^{i\tau})}.$$

Differentiating two times we can deduce that $H(0)=1$, $H'(0)=0$, $H''(0)=-2$.

5. Proof of Theorem 1

From (2.1) we have that

(5.1) $$A_k = c_1 \theta_1^k + c_2 \theta_2^k,$$

where

(5.2) $$\theta_1 = \frac{1+\sqrt{5}}{2}, \quad \theta_2 = \frac{1-\sqrt{5}}{2}, \quad c_1 = \frac{5+\sqrt{5}}{10}, \quad c_2 = \frac{5-\sqrt{5}}{10}.$$

Lemma 2. *Let $C(N, l_0)$ be the number of those $n \in A_N$ the longest component of which is greater than l_0. Then*

(5.3) $$2^{-N} C(N, l_0) \ll N \cdot \left(\frac{\theta_1}{2}\right)^{l_0}.$$

Proof. Assume that the longest $\mathfrak{M}$ component of n is $l\,(\geqq l_0)$. Then for a suitable integer t we have $n = h + 2^{t+2}u + 2^{t+l+4}v$, where

(5.4) $$h < 2^t, \quad v < 2^{N-t-l-4}, \quad u \in \mathfrak{M}_l.$$

The number of n satisfying (5.4) is $\ll A_l \cdot 2^t \cdot 2^{N-t-l-4}$. Summing up for t, and l, we have

$$2^{-N} C(N, l_0) \leqq N \sum_{l \geqq l_0} 2^{-l-2} A_l \ll N \left(\frac{\theta_1}{2}\right)^{l_0}.$$

Lemma 3. *Let $H_1 = \frac{N}{6} - \varrho(N)\sqrt{N}$, where $\varrho(N) \to \infty\ (N \to \infty)$, and*

(5.5) $$S = \sum_{t_1 + \ldots + t_{H_1} \geqq N} \prod_{j=1}^{H_1} \frac{A_{t_j - 2}}{2^{t_j}}.$$

Then

(5.6) $$s \ll e^{-\sqrt{\varrho(N)}}.$$

Proof. First we observe that for $c>1$

$$S \leqq c^{-N} \sum_{t_1,\dots,t_{H_1}} \prod_{j=1}^{H_1} A_{t_j-2}\cdot\left(\frac{c}{2}\right)^{t_j} = c^{-N}F\left(\frac{c}{2}\right)^{H_1}.$$

We take $c=1+\delta$, $\delta\to 0$. Then, by repeating the estimation that we used for the deduction of (3.5), we get

$$c^{-N}F^{H_1}\left(\frac{c}{2}\right) = \exp\left[(6H_1-N)\delta+(8H_1+N/2)\delta^2+O(N\delta^3)\right].$$

By choosing $\delta=\left(N\varrho(N)\right)^{-1/2}$, we get (5.6).

Let

(5.7) $$\varphi(\tau) = 2^{-N}\sum_{n\in\mathscr{A}_N} e^{i\tau\Delta(n)},$$

(5.8) $$A(n, H_1) = \sum_{j=1}^{H_1} \Delta(m_{l_j}),$$

(5.9) $$B(n, H_1) = \Delta(n)-A(n, H_1).$$

We take $A(n, H_1)=A\left(n, \nu(n)\right)$, when $\nu(n)<H_1$. Let

(5.10) $$\varphi_0(\tau) = 2^{-N}\sum_{n\in\mathscr{A}_N} e^{i\tau A(n,H_1)}.$$

First we consider $\varphi_0(\tau)$. It is obvious that

(5.11) $$\varphi_0(\tau) = \sum_{t_1+\dots+t_{H_1}\leqq N}\prod_{j=1}^{H_1}\frac{S_{l_j}(\tau)}{2^{t_j}}.$$

From Lemma 3 we get

(5.12) $$\varphi_0(\tau) = H(\tau)^{H_1}+O(S) = H(\tau)^{H_1}+O\left(e^{-\sqrt{\varrho(N)}}\right).$$

Now we estimate the difference $\varphi(\tau)-\varphi_0(\tau)$. Let $\mathscr{A}$ denote the set of those integers $n\in\mathscr{A}_N$ for which

$$t_1+t_2+\dots+t_{H_1}\leqq N-\varrho(N)(\log N)^2\sqrt{N}.$$

Let $\mathscr{B}=\mathscr{A}_N\setminus\mathscr{A}$. We show that $\mathscr{A}$ has at most $O(2^N/N)$ elements. From Theorem 2 it follows easily that the number of those $n\in\mathscr{A}_N$ for which

$$\nu(n)\geqq\frac{N}{6}+\varrho(N)\sqrt{N}$$

is $O(2^N/N)$. For the remaining elements of $\mathscr{A}$ we get

$$\varrho(N)(\log N)^2\sqrt{N}\leqq t_{H_1+1}+\dots+t_{\nu(n)}\leqq$$
$$\leqq \max(l_j+2)\cdot\left(\nu(n)-H_1\right)\leqq 4(\max l_j)\cdot\varrho(N)\sqrt{N},$$

i.e.

$$\max l_j \geqq \frac{1}{4}(\log N)^2. \tag{5.13}$$

From Lemma 2 we have that the number of $n \in \mathcal{A}_N$ satisfying (5.13) is smaller than $O(2^N/N)$. We have

$$|\varphi_0(\tau)-\varphi(\tau)| \ll 2^{-N} \sum_{n\in\mathcal{A}} 1 + 2^{-N}|\tau| \sum_{n\in\mathcal{B}} |B(n, H_1)| = \tag{5.14}$$

$$= 2^{-N}(\Sigma_1 + |\tau| \cdot \Sigma_2).$$

It is obvious that

$$\Sigma_1 \ll 2^N/N.$$

We can write $n\in\mathcal{B}$ in the following form:

$$n = M+2^s l, \quad t_1+\ldots+t_{H_1} = s, \quad l\in\mathcal{A}_{N-s-2},$$

where M has the components $m_{t_1}, \ldots, m_{t_{H_1}}$.

Let $\mathcal{D}_M$ denote the set of these elements. Then, by (1.4), applying the Cauchy inequality, we have

$$\sum_{n\in\mathcal{D}_M} |B(n, H_1)| \ll \sum_{j<2^{N-s}} |\Delta(j)| \ll 2^{N-s}\sqrt{N-s}.$$

Observing that

$$N-s \leqq \varrho(N)(\log^2 N)\cdot\sqrt{N}$$

for $n\in\mathcal{B}$, we get that

$$\Sigma_2 \ll 2^N \cdot N^{1/4}\sqrt{\varrho(N)}\cdot\log N.$$

So we get that

$$|\varphi_0(\tau)-\varphi(\tau)| \ll \frac{1}{N} + |\tau| N^{1/4}\sqrt{\varrho(N)}\cdot\log N. \tag{5.15}$$

Consequently,

$$\varphi(\tau) = H(\tau)^{H_1} + O\left(e^{-\sqrt{\varrho(N)}}\right) + O(1/N) + O\,|\tau|\, N^{1/4}\sqrt{\varrho(N)}\cdot\log N).$$

Observing that $H(0)=1$, $H'(1)=0$, $H''(0)=-2$, we get

$$H(\tau) = 1-\tau^2+O(\tau^3) = \exp\left(-\tau^2+O(\tau^3)\right).$$

By taking $\tau = \varkappa/N$, we have

$$\varphi(\varkappa/\sqrt{N}) = \exp(-\varkappa^2/6)+o(1) \quad (N\to\infty)$$

uniformly for every $\varkappa$ in an arbitrary bounded interval. But $\exp(-\varkappa^2/6)$ is the characteristic function of the normal distribution function with zero mean, and variance $1/\sqrt{3}$. Using the well-known theorem of probability theory on the convergence of characteristic functions, we get Theorem 1 immediately.

MATH. DEPARTMENT
EÖTVÖS LORÁND UNIVERSITY
H-1445 BUDAPEST, PF. 323.

Acta Sci. Math., 39 (1977), 329—340

О вложении полугрупп в полугруппу, в которой исходные полугруппы являются независимыми

Л. МЕДЕШИ *

§ 1. Введение

Понятие независимости подполугрупп было введено Е. С. Ляпиным в работах [5], [7]. Это понятие является естественным перенесением в теорию полугрупп условия, которому в теории групп удовлетворяют компоненты свободного произведения групп. Однако в отличие от теории групп, независимые подполугруппы могут пересекаться по любой полугруппе. Настоящая работа посвящена изучению следующей проблемы: При каких условиях полугрупповая амальгама (определение см. ниже) вкладывается в полугруппу так, чтобы в этой полугруппе данные полугруппы, которые составляют амальгаму, являлись независимыми подполугруппами? Основная теорема работ [5], [7] Е. С. Ляпина (которая приведена ниже) решает эту проблему в случае амальгамы двух полугрупп. В настоящей работе упомянутая теорема Е. С. Ляпина обобщается на случай произвольной амальгамы.

В теории полугрупп амальгамы являются наиболее изученными частичными группоидами. Проблема о вложении амальгамы в полугруппу представляется слишком трудной и далекой от разрешимости. Для некоторых частных случаев она исследовалась Хауи в работах [2], [3], [4]. Результаты Хауи существенно отличны от результатов настоящей работы.

Работа [8], Е. С. Ляпина была посвящена независимым полугрупповым продолжениям частичных группоидов. Проблема, рассматриваемая в настоящей статье является частью проблемы независимых полугрупповых продолжений частичных группоидов. Следуя работам [6], [8] мы введем следующие определения.

Поступило 7. апреля 1977.

* L. Megyesi

Определение. Пусть $\{A_\xi\}_{\xi\in\mathscr{I}}$ — некоторая совокупность полугрупп ($\mathscr{I}$ — некоторое множество индексов), причем никакая A_α не содержится в объединении остальных A_ξ ($\xi\in\mathscr{I}\setminus\langle\alpha\rangle$) и пусть действия в полугруппах A_ξ ($\xi\in\mathscr{I}$) согласованы между собой, т. е. из $ab=c$ в полугруппе A_ϱ и $ab=c'$ в полугруппе A_σ ($a, b, c\in A_\varrho$; $a, b, c'\in A_\sigma$) всегда следует $c=c'$. Обозначим через A объединение множеств A_ξ ($\xi\in\mathscr{I}$) и определим в A частичное действие следующим образом: $ab=c$ ($a, b, c\in A$) имеет место тогда и только тогда, если существует такой элемент $\eta\in\mathscr{I}$, что $a, b, c\in A_\eta$ и в полугруппе A_η $ab=c$. Определенный на множестве A частичный группоид называется *амальгамой* полугрупп $\{A_\xi\}_{\xi\in\mathscr{I}}$.

Определение. Будем говорить, что частичное действие в частичном группоиде C *слабо ассоциативно*, если из того, что xy, yz, $(xy)z$, $x(yz)$ ($x, y, z\in C$) в группоиде C определены, следует $(xy)z=x(yz)$.

Определение. Непустое подмножество K полугруппы A называется *разделяющейся единично идеальной* (сокращенно *р. е. и.*) *подполугруппой*, если

а) $K=L\cup M^l\cup M^r\cup T$, где $L=\langle x|\ x\in K$ и $bx=xb=b$ при $b\in A\setminus K\rangle$, $M^l=\langle x|\ x\in K$ и $xb=b$, $bx\in K$ при $b\in A\setminus K\rangle$, $M^r=\langle x|\ x\in K$ и $bx=b$, $xb\in K$ при $b\in A\setminus K\rangle$, $T=\langle x|\ x\in K$ и bx, $xb\in K$ при $b\in A\setminus K\rangle$.

б) подполугруппа K является связкой подполугрупп со следующей таблицей умножения:

(*)

	L	M^l	M^r	T
L	L	M^l	M^r	T
M^l	M^l	M^l	T	T
M^r	M^r	T	M^r	T
T	T	T	T	T

Эту связку будем называть *$*$-связкой*, и подполугруппы L, M^l, M^r, T — *L-, M^l-, M^r-, T-компонентамы* подполугруппы K в A.

Теорема 1. (Е. С. Ляпин [7]) *Пусть A есть слабо ассоциативная амальгама двух полугрупп A_α и A_β. Для того, чтобы существовала такая полугруппа B, которая содержит амальгаму A и в которой A_α и A_β являются независимыми подполугруппами, необходимо и достаточно выполнение следующих условий:*

А) $K_{\alpha\beta}=A_\alpha\cap A_\beta$ *пусто, или является р. е. и. подполугруппой и для A_α и для A_β.*

Б) *Для всяких $y_\alpha\in F_{\alpha\beta}=A_\alpha\setminus K_{\alpha\beta}$, $y_\beta\in F_{\beta\alpha}=A_\beta\setminus K_{\alpha\beta}$, $x\in K_{\alpha\beta}$ из $y_\alpha x=y_\alpha$ следует $xy_\beta=y_\beta$, из $xy_\alpha=y_\alpha$ следует $y_\beta x=y_\beta$ и наоборот.*

Замечание. Из определения р. е. и. подполугрупп вытекает, что в предыдущей теореме можно заменить условие Б) следующим:

Б') L- и T-компоненты р. е. и. подполугруппы $K_{\alpha\beta}$ в A_α и в A_β совпадают, а M^l-компонента р. е. и. подполугруппы $K_{\alpha\beta}$ в A_α равняется M^r-компоненте в A_β и аналогично M^r-компонента в A_α равняется M^r-компоненте в A_β.

§ 2. Амальгама полугрупп, каждая пара которых удовлетворяет условиям теоремы Е. С. Ляпина

В этом параграфе рассматривается такая амальгама A полугрупп A_ξ $(\xi \in \mathscr{I})$ ($\mathscr{I}$ — некоторое множество индексов), что каждая пара A_α, A_β $(\alpha, \beta \in \mathscr{I})$ удовлетворяет условиям А), Б), (или А), Б')) теоремы Е. С. Ляпина. Отметим, что в этом параграфе не требуется слабой ассоциативности.

Обозначения. Пусть $C_\alpha = \bigcup\limits_{\xi \in \mathscr{I} \setminus \langle\alpha\rangle} K_{\alpha\xi}$; $B_\alpha = A_\alpha \setminus C_\alpha$. По определению амальгамы $B_\alpha \neq \emptyset$, для каждого $\alpha \in \mathscr{I}$ Обозначим через $\overline{w}$ для каждого элемента $w \in A$ множество всех индексов полугрупп, которые содержат w. Пусть $H = \bigcup\limits_{\xi \in \mathscr{I}} C_\xi$ т. е. множество всех элементов, которые содержатся не только в одной из A_ξ $(\xi \in \mathscr{I})$.

Определения. Назовем элемент $x \in H$ *L-элементом*, если для каждого $\alpha \in \bar{x}$ имеет место $xb = bx = b$ для всяких $b \in B_\alpha$. Элемент t называется *T-элементом*, если для каждых $\alpha, \beta \in \bar{t}$ имеет место tb, $bt \in K_{\alpha\beta}$ для всяких $b \in B_\beta \cup B_\beta$. Элемент u называется *M-элементом*, если $\bar{u}$ состоит только из двух индексов: α, β и либо

а) $ub_\alpha = b_\alpha$, $b_\alpha u \in K_{\alpha\beta}$, $b_\beta u = b_\beta$, $ub_\beta \in K_{\alpha\beta}$ для всякого $b_\alpha \in B_\alpha$, $b_\beta \in B_\beta$; либо

б) $b_\alpha u = b_\alpha$, $ub_\alpha \in K_{\alpha\beta}$, $ub_\beta = b_\beta$, $b_\beta u \in K_{\alpha\beta}$ для всякого $b_\alpha \in B_\alpha$, $b_\beta \in B_\beta$.

Для M-элемента u обозначим через $\hat{u}$ множество $\bar{u}$, упорядоченное следующим образом: $\hat{u} = (\alpha, \beta)$, если для u имеет место случай а) и $\hat{u} = (\beta, \alpha)$, если выполнено б).

Докажем несколько свойств, которые покажут как устроена амальгама A.

Свойство α. *Множество $H \subset A$ содержит только L-, M- и T-элементы.*

Доказательство. Достаточно доказать, что элемент $z \in K_{\alpha\beta}$ является или L-, или M-, или T-элементом. Возможны следующие случаи:

1. z — элемент L-компоненты р. е. и. подполугруппы $K_{\alpha\beta}$ в A_α. Тогда $zb = bz = b$ для каждого $b \in B_\alpha$. Пусть $\xi \in \bar{z}$. Из условия *Б)* следует, что $zb = bz = b$ для каждого $b \in B_\xi$. Значит z является L-элементом.

2. z — элемент T-компоненты р. е. и. подполугруппы $K_{\alpha\beta}$ в A_α. Тогда zb, $bz\in K_{\alpha\beta}$ для каждого $b\in B_\alpha$. Пусть $\xi\in\bar{z}$. Из условия *Б')* следует, что T-компоненты $K_{\alpha\xi}$ в A_α и в A_ξ совпадают. Очевидно, что z содержится в T-компоненте $K_{\alpha\xi}$ в A_α (как видели $zb\neq b$, $bz\neq b$ для всех $b\in B_\alpha$). Поэтому z содержится в T-компоненте $K_{\alpha\xi}$ и в A_ξ, т. е. zb, $bz\in K_{\alpha\xi}$ для всех $b\in B_\xi$. Значит z — T-элемент.

3. z — элемент M^l-компоненты р. е. и. подполугруппы $K_{\alpha\beta}$ в A_α. Для того, чтобы показать что z — M-элемент, достаточно доказать, что z содержится только в двух полугруппах из A_ξ $(\xi\in\mathscr{I})$, именно в A_α и в A_β. Предположим противное, пусть $z\in A_\alpha\cap A_\beta\cap A_\gamma\cap\ldots$. Так как z в A_α является элементом M^l-компоненты подполугруппы $K_{\alpha\beta}$, то имеет место $zb_\alpha=b_\alpha$, $b_\alpha z\in K_{\alpha\beta}$ для всех $b_\alpha\in B_\alpha$. Тогда по условию *Б')* z в A_β является элементом M^r-компоненты, т. е. $b_\beta z=b_\beta$, $zb_\beta\in K_{\alpha\beta}$ для всех $b_\beta\in B_\beta$. Отсюда и из условия Б') следует, что z — элемент M^r-компоненты полугруппы $K_{\alpha\gamma}$ в A_γ, т. е. $b_\gamma z=b_\gamma$, $zb_\gamma\in K_{\alpha\gamma}$ для всех $b_\gamma\in B_\gamma$ и одновременно z является элементом M^l-компоненты полугруппы $K_{\beta\gamma}$ в A_γ, т. е. $zb_\gamma=b_\gamma$, $b_\gamma z\in K_{\beta\gamma}$ для всех $b_\gamma\in B_\gamma$, что неверно. Аналочигно доказывается, что z является M-элементом, если он содержится в M^r-компоненте р. е. и. подполугруппы $K_{\alpha\beta}$ в A_α. На основе предыдущего рассуждения легко видеть:

Следствие. Каждый L-элемент является элементом L-компоненты во всех пересечениях $K_{\alpha\beta}$, которые содержат его, каждый T-элемент является элементом T-компоненты во всех пересечениях $K_{\alpha\beta}$, которые содержат его и каждый M-элемент является элементом M^l- или M^r-компоненты р. е. и. подполугруппы $K_{\alpha\beta}$, которая содержит его. В этом параграфе это утверждение многократно используется без дополнительных ссылок на него.

Свойство β. *Пусть x_1, x_2 L-элементы и $\bar{x}_1\cap\bar{x}_2\neq\emptyset$. Тогда или $\bar{x}_1\subseteqq\bar{x}_2$ или $\bar{x}_2\subseteqq\bar{x}_1$. Если $\bar{x}_1\subset\bar{x}_2$ (но $\bar{x}_1\neq\bar{x}_2$), то $x_1x_2=x_2x_1=x_1$. Если $\bar{x}_1=\bar{x}_2$, то x_1x_2, x_2x_1 L-элементы и $x_1x_2\supseteqq x_1=x_2$; $x_2x_1\supseteqq x_1=x_2$.*

Доказательство. Пусть $\alpha\in\bar{x}_1\cap\bar{x}_2$. Если бы существовали $\beta\in\bar{x}_1$, $\beta\bar{\in}\bar{x}_2$ и $\gamma\bar{\in}\bar{x}_1$, $\gamma\in\bar{x}_2$, то $x_1\in K_{\alpha\beta}$, $x_2\in F_{\alpha\beta}$ и поэтому $x_1x_2=x_2x_1=x_2$ и аналогично $x_2\in K_{\alpha\gamma}$, $x_1\in F_{\alpha\gamma}$, $x_1x_2=x_2x_1=x_1$, имеем $x_1=x_2$. Получили противоречие. Пусть $\bar{x}_1\subset\bar{x}_2$ (но $\bar{x}_1\neq\bar{x}_2$) и пусть $\gamma\bar{\in}\bar{x}_1$, $\gamma\in\bar{x}_2$. Тогда $x_2\in K_{\alpha\gamma}$ и $x_1\in F_{\alpha\gamma}$ поэтому $x_1x_2=x_2x_1=x_1$.

Пусть теперь $\bar{x}_1=\bar{x}_2$. Так как x_1 и x_2 — элементы L-компоненты полугруппы $K_{\alpha\beta}$ для любых двух $\alpha,\beta\in\bar{x}_1=\bar{x}_2$, то согласно свойствам $*$-связки произведения x_1x_2, x_2x_1 также являются элементами L-компоненты $K_{\alpha\beta}$ (для любых двух $\alpha,\beta\in\bar{x}_1=\bar{x}_2$). Значит x_1x_2, x_2x_1 L-элементы и $\overline{x_1x_2}\supseteqq\bar{x}_1=\bar{x}_2$, $\overline{x_2x_1}\supseteqq\bar{x}_1=\bar{x}_2$.

Свойство γ. *Если u, v такие M-элементы, для которых $\bar{u}\cap\bar{v}\neq\emptyset$ но $\bar{u}\neq\bar{v}$, то у $\hat{u}$, $\hat{v}$ или первые или вторые члены совпадают.*

Доказательство. Пусть u, v $(u \neq v)$ M-элементы и пусть $\alpha \in \bar{u} \cap \bar{v}$ и $\bar{u} \neq \bar{v}$. Достаточно доказать, что случай $\hat{u}=(\alpha, \beta)$ $\hat{v}=(\gamma, \alpha)$ $(\beta \neq \gamma)$ не может иметь места. (Случай $\hat{u}=(\beta, \alpha)$ $\hat{v}=(\alpha, \gamma)$ доказывается аналогичным образом.) Действительно, если бы $\hat{u}=(\alpha, \beta)$, $\hat{v}=(\gamma, \alpha)$, то в полугруппе A_α элемент u является элементом M^l-компоненты полугруппы $K_{\alpha\beta}$ и $v \in F_{\alpha\beta}$ поэтому $uv=v$, аналогично в A_α v является элементом M^r-компоненты $K_{\alpha\gamma}$, $u \in F_{\alpha\gamma}$, поэтому $uv=u$. Следовательно $u=v$, что неверно.

Свойство δ. *Пусть u — M-элемент и x — L-элемент. Если $\bar{x} \cap \bar{u} \neq \emptyset$, то $\bar{u} \subseteqq \bar{x}$. Если $\bar{u} \subset \bar{x}$ (но $\bar{u} \neq \bar{x}$), то $xu=ux=u$. Если $\bar{u}=\bar{x}$, то ux, xu M-элементы и $\widehat{xu}=\widehat{ux}=\hat{u}$.*

Доказательство. Пусть u — M-элемент, x — L-элемент. Предположим противное первому утверждению. Пусть $\alpha \in \bar{x}$, $\beta \bar{\in} \bar{x}$. Так как u — элемент M^l-компоненты полугруппы $K_{\alpha\beta}$ в полугруппе A_α и $x \in F_{\alpha\beta}$ поэтому $ux=x$. $\bar{x}$ состоит по крайней мере из двух индексов, т. е. существует $\gamma \in \bar{x}$. Тогда $x \in K_{\alpha\gamma}$, $u \in F_{\alpha\gamma}$ (и следовательно $ux=u$). Отсюда $u=x$. Получили противоречие.

Пусть теперь $\bar{u} \subset \bar{x}$ (но $\bar{u} \neq \bar{x}$). Существует индекс γ такой, что $\gamma \in \bar{x}$, $\gamma \bar{\in} \bar{u}$. Пусть $\alpha \in \bar{u} \subset \bar{x}$. Элемент x содержится в L-компоненте полугруппы $K_{\alpha\beta}$ и $u \in F_{\alpha\gamma}$, поэтому $xu=ux=u$.

Пусть наконец, $\bar{u}=\bar{x}=\langle\alpha, \beta\rangle$. Тогда x, u являются элементами $K_{\alpha\beta}$. Из свойств $*$-связки следует, что произведения ux, xu являются M-элементами, и $\widehat{xu}=\widehat{ux}=\hat{u}$.

Свойство ε. *Пусть u, v — M-элементы. Если $\hat{u}=\hat{v}$, то uv, vu — M-элементы и $\widehat{uv}=\widehat{vu}=\hat{u}=\hat{v}$. Если $\bar{u}=\bar{v}$ (но $\hat{u} \neq \hat{v}$), то uv, vu T-элементы и $\overline{uv} \supseteqq \bar{u}=\bar{v}$, $\overline{vu} \supseteqq \bar{u}=\bar{v}$. Если $\bar{u} \cap \bar{v} \neq \emptyset$, $\bar{u} \neq \bar{v}$ и у $\hat{u}$, $\hat{v}$ первые члены совпадают, то $uv=v$, $vu=u$, если — вторые, то $uv=u$, $vu=v$.*

Доказательство. Если $\hat{u}=\hat{v}$, или если $\bar{u}=\bar{v}$ но $\hat{u} \neq \hat{v}$, то элементы u, v содержатся в одной и той же самой полугруппе $K_{\alpha\beta}$ и утверждение о том, что амальгама A обладает свойством ε непосредственно следует из свойств $*$-связки.

Пусть теперь u, v такие M-элементы, для которых $\hat{u}=(\alpha, \beta)$ $\hat{v}=(\alpha, \gamma)$ $(\beta \neq \gamma)$. Тогда u и v содержатся в M^l-компоненте для полугруппы $K_{\alpha\beta}$ и $K_{\alpha\gamma}$ соотвтственно, и так как $u \in F_{\alpha\gamma}$, $v \in F_{\alpha\beta}$ то $uv=v$, $vu=u$.

(Доказательство аналогично в случае $\hat{u}=(\beta, \alpha)$ $\hat{v}=(\gamma, \alpha)$ $(\beta \neq \gamma)$.)

Свойство ζ. *Пусть x — L-элемент и t — T-элемент. Если $\bar{x} \cap \bar{t} \neq \emptyset$ и $\bar{x} \nsubseteq \bar{t}$, то $xt=tx=t$. Если $\bar{x} \subseteqq \bar{t}$, то xt, tx — T-элементы и $\overline{xt} \supseteqq \bar{t}$, $\overline{tx} \supseteqq \bar{t}$.*

Доказательство. Пусть $\alpha \in \bar{x} \cap \bar{t}$ и $\beta \in \bar{x}$, $\beta \bar{\in} \bar{t}$. Тогда x содержится в L-компоненте полугруппы $K_{\alpha\beta}$ и $t \in F_{\alpha\beta}$ следовательно $xt=tx=t$. Пусть теперь

$\bar{x} \subseteqq \bar{t}$. В случае любых двух $\alpha, \beta \in \bar{x}$ имеем $x, t \in K_{\alpha\beta}$, и из свойств $*$-связки следует, что xt, tx являются T-элементами. Одновременно получается, что $\overline{xt}$, $\overline{tx} \supseteqq \bar{x}$. Если существует $\gamma \in \bar{t}$, $\gamma \bar{\in} \bar{x}$ то так как t содержится в T-компоненте $K_{\alpha\gamma}$ и $x \in F_{\alpha\gamma}$, имеем xt, $tx \in K_{\alpha\gamma}$, причем xt, tx содержатся в T-компоненте, так как они — T-элементы. Значит $\overline{xt} \supseteqq \bar{t}$, $\overline{tx} \supseteqq \bar{t}$.

Свойство η. *Пусть u — M-элемент и t — T-элемент. Пусть $\hat{u} = (\alpha, \beta)$. Если $\alpha \in \bar{t}$, $\beta \bar{\in} \bar{t}$, то $ut = t$ и tu является T-элементом, для которого $\overline{tu} \supseteqq \bar{t} \cup \bar{u}$. Аналогично: если $\alpha \bar{\in} \bar{t}$, $\beta \in \bar{t}$, то $tu = t$ и ut — T-элемент, для которого $\overline{ut} \supseteqq \bar{t} \cup \bar{u}$. Если $\bar{u} \subseteqq \bar{t}$, то tu, ut являются T-элементами, для которых $\overline{tu} \supseteqq \bar{t}$, $\overline{ut} \supseteqq \bar{t}$.*

Доказательство. Пусть $\hat{u} = (\alpha, \beta)$ и $\alpha \in \bar{t}$, $\beta \bar{\in} \bar{t}$. Тогда u содержится в M^l-компоненте полугруппы $K_{\alpha\beta}$ в A_α и $t \in F_{\alpha\beta}$. Таким образом $ut = t$ и $\overline{tu} \supseteqq \bar{u}$. Для каждого $\gamma \in \bar{t}$, $\gamma \bar{\in} \bar{u}$ имеет место, что $tu \in K_{\alpha\gamma}$. Следовательно $\overline{tu} \supseteqq \bar{t}$. Получается: $\overline{tu} \supseteqq \bar{t} \cup \bar{u}$. Отсюда следует, что tu содержится в не меньше чем трех полугруппах из A_ξ $(\xi \in \mathscr{I})$. Поэтому tu не является M-элементом. Очевидно, что tu может быть только T-элементом.

Второе утверждение доказывается аналогично. Пусть теперь $\bar{u} \subseteqq \bar{t}$, $\bar{u} = \langle \alpha, \beta \rangle$. Полугруппа $K_{\alpha\beta}$ — $*$-связка, $u, t \in K_{\alpha\beta}$. Из свойств $*$-связки следует, что ut, tu — T-элементы и $\overline{ut}$, $\overline{tu} \supseteqq \bar{t}$. Утверждение, что $\overline{ut}$, $\overline{tu} \supseteqq \bar{t}$ вытекает из того, что t содержится в T-компоненте всякой полугруппы $K_{\alpha\gamma}$, где $\gamma \in \bar{t}$, $\gamma \bar{\in} \bar{u}$.

Свойство ϑ. *Пусть t_1, t_2 — T-элементы. Если $\bar{t}_1 \cap \bar{t}_2 \neq \emptyset$, то $t_1 t_2$ и $t_2 t_1$ — T-элементы и $\bar{t}_1 \cup \bar{t}_2 \subseteqq \overline{t_1 t_2}$, $\bar{t}_1 \cup \bar{t}_2 \subseteqq \overline{t_2 t_1}$.*

Доказательство. Пусть $\alpha \in \bar{t}_1 \cap \bar{t}_2$. Если $\bar{t}_1 = \bar{t}_2$, то утверждение следует из свойств $*$-связки. Пусть $\beta \in \bar{t}_1$, $\beta \bar{\in} \bar{t}_2$. Тогда $t_1 \in K_{\alpha\beta}$, $t_2 \in F_{\alpha\beta}$ и поэтому $t_1 t_2$, $t_2 t_1 \in K_{\alpha\beta}$. Продолжая этот процесс, получается $\overline{t_1 t_2} \supseteqq \bar{t}_1 \cup \bar{t}_2$, $\overline{t_2 t_1} \supseteqq \bar{t}_1 \cup \bar{t}_2$. Отсюда следует, что $t_1 t_2$ (аналогично $t_2 t_1$) только тогда содержится только в двух полугруппах из A_ξ $(\xi \in \mathscr{I})$, когда t_1 и t_2 содержатся только в этих полугруппах. В этом случае из свойств $*$-связки следует, что $t_1 t_2$ и $t_2 t_1$ являются T-элементами. Если $t_1 t_2$, аналогично $t_2 t_1$ содержится не меньше чем в трех полугруппах из A_ξ $(\xi \in \mathscr{I})$, то очевидно, что оно не является M-элементом, а может быть только T-элементом.

Теорема 2. *Пусть A — амальгама полугрупп A_ξ $(\xi \in \mathscr{I})$ ($\mathscr{I}$ — некоторое множество индексов). Амальгама A удовлетворяет условиям* А) *и* Б) *тогда и только тогда, когда она обладает свойствами $\alpha - \vartheta$.*

Доказательство. Выше уже было доказано, что если амальгама удовлетворяет условиям А) и Б) то свойства $\alpha - \vartheta$ выполнены.

Предположим теперь, что для A имеют место $\alpha - \vartheta$. Пусть $K_{\alpha\beta} = A_\alpha \cap A_\beta$ $(\alpha, \beta \in \mathscr{I})$. Чтобы доказать А) достаточно показать, что $K_{\alpha\beta}$ — р. е. и. подполугруппа в A_α. Тот факт, что $K_{\alpha\beta}$ — полугруппа, следует из определения амальгамы (это также легко видеть из свойств $\alpha-\vartheta$).

Будем доказывать, что L-элементы, которые содержатся в $K_{\alpha\beta}$ образуют L-компоненту в $K_{\alpha\beta}$, M-элементы u, для которых $\hat{u}=(\alpha, \beta)$ M^l-компоненту, элементы u, для которых $\hat{u}=(\beta, \alpha)$ образуют M^r-компоненту и T-элементы, содержащиеся в $K_{\alpha\beta}$ T-компоненту полугруппу $K_{\alpha\beta}$. Из этих фактов следует $A)$.

1. Пусть x — L-элемент в $K_{\alpha\beta}$. Докажем, что

$$xy = yx = y \tag{1}$$

для каждого $y \in F_{\alpha\beta}$. Если $y \in L_\alpha$, то (1) следует из определения L-элемента. Пусть $y \in C_\alpha \setminus K_{\alpha\beta}$. Очевидно, что $\bar{x} \subsetneqq \bar{y}$. Если y — L-элемент, то из свойства β получается, что $\overline{xy} \subseteqq \bar{x}$, но очевидно $\bar{x} \neq \bar{y}$, таким образом из β следует и (1). Рассмотрим случай, когда y — M-элемент. Из δ следует, что $\bar{y} \subseteqq \bar{x}$, но $\bar{x} \neq \bar{y}$ и снова используя δ получаем, что выполняется (1). Пусть y — T-элемент. Так как $\bar{x} \subsetneqq \bar{y}$ из ζ вытекает (1).

2. Пусть u — M-элемент в $K_{\alpha\beta}$, для которого $\hat{u}=(\alpha, \beta)$. Будем доказывать что

$$uy = y \quad yu \in K_{\alpha\beta} \tag{2}$$

для каждого $y \in F_{\alpha\beta}$. Если $y \in B_\alpha$, то (2) следует из определения M-элемента. Пусть $y \in C_\alpha \setminus K_{\alpha\beta}$. Согласно свойству δ все L-элементы из C_α, если они существуют, содержатся в $K_{\alpha\beta}$. Далее $\bar{u} \neq \bar{y}$ и из δ следует (2).

Предположим, что y — M-элемент. Свойство γ показывает, что в упорядоченных парах $\hat{u}$ и $\hat{y}$ первые члены совпадают (они равняются α). Из ε следует $uy=y$, $yu=u$ т. е. выполняется (2). Пусть y — T-элемент. Свойство η показывает, что (2) имеет место.

Доказательство аналогично, если u такой M-элемент, для которого $\hat{u}=(\beta, \alpha)$.

3. Пусть t — T-элемент в $K_{\alpha\beta}$. Будем доказывать, что

$$ty, yt \in K_{\alpha\beta} \tag{3}$$

для всех $y \in F_{\alpha\beta}$. Если $y \in B_\alpha$, то это утверждение следует из определения T-элемента. Пусть $y \in C_\alpha \setminus K_{\alpha\beta}$. Если y — L-элемент, то согласно ζ или $\bar{x} \subsetneqq \bar{t}$ и $yt=ty=t$ или $\bar{x} \subseteqq \bar{t}$ и тогда $\overline{xt} \supseteqq \bar{t}$, $\overline{tx} \supseteqq \bar{t}$ откуда получаем (3). Рассмотрим случай, когда y — M-элемент, и $\hat{y}=(\alpha, \gamma)$ $(\gamma \neq \beta)$. Если $\gamma \bar{\in} \bar{t}$, то из свойства η следует, что $yt=t$ и ty — T-элемент, для которого $\overline{ty} \supseteqq \bar{t} \cup \bar{y}$, т. е. $ty \in K_{\alpha\beta}$. Если $\gamma \in \bar{t}$ то из η также вытекает выполнение (3), так как $\overline{ty} \supseteqq \bar{t}$, $\overline{yt} \supseteqq \bar{t}$. Аналогично рассматривается случай, когда y такой M-элемент, для которого $\hat{y}=(\gamma, \alpha)$ $(\gamma \neq \beta)$.

Пусть y — T-элемент в $C_\alpha \setminus K_{\alpha\beta}$. Из свойства ϑ следует, что $\overline{yt} \supseteqq \bar{t} \cup \bar{y}$, $\overline{ty} \supseteqq \bar{t} \cup \bar{y}$, т. е. ty, $yt \in K_{\alpha\beta}$.

Теперь докажем выполнение условия *Б)*.

Пусть $x \in K_{\alpha\beta}$, $y_\alpha \in F_{\alpha\beta}$, $y_\beta \in F_{\beta\alpha}$. Если $y_\alpha x = y_\alpha$, то x содержится в L-компоненте $K_{\alpha\beta}$ или в M^r-компоненте $K_{\alpha\beta}$ в A_α. В первом случае x — L-элемент, который по предыдущему доказательству находится в L-компоненте $K_{\alpha\beta}$ и в A_β, т. е. $xy_\beta = y_\beta$. Во втором случае x — M-элемент, для которого $\hat{x} = (\beta, \alpha)$ и по предыдущему доказательству он содержится в M^l-компоненте $K_{\alpha\beta}$ в A_β поэтому $xy_\beta = y_\beta$. Дальнейшие случаи *Б)* доказываются аналогично.

Теорема доказана.

Замечание. По теореме 2 имеем, что условия А), Б) и набор сложных свойств $\alpha - \vartheta$ эквивалентны. Заметим, что свойствами $\alpha - \vartheta$ бывает удобно пользоваться, что покажет доказательство теоремы следующего параграфа. Кроме того с помощью их можно сделать вывод о строении амальгамы.

§ 3. Обобщение теоремы Е. С. Ляпина

Теорема 3. *Пусть A — амальгама полугрупп $\{A_\xi\}_{\xi \in \mathscr{I}}$ ($\mathscr{I}$ — некоторое множество индексов). Амальгама A тогда и только тогда погружаема в надполугруппу B, так чтобы все полугруппы A_ξ ($\zeta \in \mathscr{I}$) являлись в B независимыми подполугруппами, когда каждая пара A_α, A_β ($\alpha, \beta \in \mathscr{I}$) может быть вложена в некоторую полугруппу таким образом, чтобы A_α, A_β в ней являлись независимыми подполугруппами.*

Эквивалентная форма теоремы 3. *Пусть A — слабо ассоциативная амальгама полугрупп $\{A_\xi\}_{\xi \in \mathscr{I}}$.*

Амальгама A тогда и только тогда погружаема в надполугруппу B так, чтобы все полугруппы A_ξ ($\xi \in \mathscr{I}$) являлись в B независимыми подполугруппами, когда A удовлетворяет условиям A), Б).

Замечание. Эквивалентность двух форм теоремы очевидна по упомянутой теореме Е. С. Ляпина.

Также очевидно, что выполнение слабой ассоциативности необходимо для того, чтобы амальгама была вложима в некоторую полугруппу.

1. Доказательство достаточности мы не приводим, ибо оно не отличается существенно от доказательства теоремы Е. С. Ляпина. Надо лишь провести некоторые очевидные изменения (например, такое: *B* состоит из последовательностей вида $u = (x_1, \ldots, x_n)$ ($n = 1, 2, \ldots$) где каждый x_i является эле-

ментом хотя бы одной полугруппы из амальгамы, причем $\bar{x}_i \cap \bar{x}_{i+1} = \emptyset$ для всех $i = 1, 2, \ldots, n-1$).

2. Доказательство необходимости. Пусть дана полугруппа B, которая содержит систему подполугрупп $A' = \{A_\xi\}_{\xi \in \mathscr{I}}$, причем подполугруппы A_ξ являются в B независимыми. Используется теорема 2. Будем доказывать, что амальгама A полугрупп A_ξ $(\xi \in \mathscr{I})$ овладает свойствами $\alpha - \vartheta$. Предварительно докажем следующую лемму.

Лемма. *Пусть $x, y, z \in A' \subseteqq B$. Если $\bar{x} \cap \bar{y} \neq \emptyset$, $\bar{y} \cap \bar{z} \neq \emptyset$, то имеет место одно из следующих двух утверждений:*

$$\overline{xy} \cap \bar{z} = \emptyset, \quad \bar{x} \cap \overline{yz} = \emptyset \quad \text{и} \quad xy = x, \quad yz = z \tag{4}$$

$$\overline{xy} \cap \bar{z} \neq \emptyset, \quad \bar{x} \cap \overline{yz} \neq \emptyset. \tag{5}$$

Доказательство леммы. Если $\bar{x} \cap \bar{y} \cap \bar{z} \neq \emptyset$ то очевидно, что выполняется (5) так как x, y, z содержатся в одной полугруппе из A_ξ $(\xi \in \mathscr{I})$. Пусть $\bar{x} \cap \bar{y} \cap \bar{z} = \emptyset$. В полугруппе B имеет место следующее равенство $(xy)z = x(yz)$. По условию существуют $\alpha \in \bar{x} \cap \bar{y}$, $\beta \in \bar{y} \cap \bar{z}$, т. е. $x, y \in A_\alpha$, $y, z \in A_\beta$, таким образом $u = xy \in A_\alpha$; $v = yz \in A_\beta$. Если бы одно из слов uz, xv являлось приведенным относительно A' (определение понятия приведенного слова и соотношения в работе [7]) а другое нет, например, $\bar{u} \cap \bar{z} = \emptyset$ а $\gamma \in \bar{x} \cap \bar{v}$, т. е. $xv = s \in A_\gamma$, то приведенное соотношение $uz = s \in A_\gamma$ противоречило бы тому, что система подполугрупп A_ξ $(\xi \in \mathscr{I})$ в B независима. Получили, что или оба слова приведены относительно A' т. е. $\bar{u} \cap \bar{z} = \emptyset$, $\bar{x} \cap \bar{v} = \emptyset$ и приведенное соотношение $uz = xv$ тривиально $u = x$, $z = v$, значит имеет место (4), или ни одно из них не является приведенным, т. е. $\bar{u} \cap \bar{z} \neq \emptyset$, $\bar{x} \cap \bar{v} \neq \emptyset$ значит имеет место (5).

Перейдем к доказательству свойств $\alpha - \vartheta$. В этом доказательстве постоянно без дополнительных оговорок употребляются обозначения $b_\alpha \in B_\alpha$, $b_\beta \in B_\beta$, ... и т. д.

α) Пусть $y \in H$; $\alpha, \beta \in \bar{y}$ (т. е. $y \in K_{\alpha\beta}$). Воспользуемся леммой: первый случай $\overline{b_\alpha y} \cap \bar{b}_\beta = \emptyset$, $\bar{b}_\alpha \cap \overline{y b_\beta} = \emptyset$ и $b_\alpha y = b_\alpha$, $y b_\beta = b_\beta$. Тогда для каждого $\xi \in \bar{y}$ так как $\langle\alpha\rangle = \bar{b}_\alpha = \overline{b_\alpha y}$ и $\bar{b}_\xi = \langle\xi\rangle$ имеет место $\overline{b_\alpha y} \cap \bar{b}_\xi = \emptyset$ и по лемме $\bar{b}_\alpha \cap \overline{y b_\xi} = \emptyset$ $y b_\xi = b_\xi$. Аналогично доказывается, что из $y b_\alpha = b_\alpha$ следует, что $b_\xi y = b_\xi$, в случае всякого $\xi \in \bar{y}$.

Второй случай: $\overline{b_\alpha y} \cap \bar{b}_\beta \neq \emptyset$, $\bar{b}_\beta \cap \overline{y b_\beta} \neq \emptyset$, т. е. $y b_\beta \in K_{\alpha\beta}$. Тогда $\overline{b_\alpha y} \cap \bar{b}_\xi \neq \emptyset$ и $\bar{b}_\alpha \cap \overline{y b_\xi} \neq \emptyset$ (иначе имело бы место $b_\alpha y = b_\alpha$), т. е. $y b_\xi \in K_{\alpha\xi}$ для всех $\xi \in \bar{y}$, $b_\xi \in B_\xi$. Аналогично из $y b_\alpha \in K_{\alpha\beta}$ следует $b_\xi y \in K_{\alpha\xi}$ для всех $\xi \in \bar{y}$. Из этих рассуждений вытекают следующие:

1. Если $b_\alpha y = b_\alpha$, $y b_\alpha = b_\alpha$ то y является L-элементом.
2. Если $b_\alpha y \in K_{\alpha\beta}$, $y b_\alpha \in K_{\alpha\beta}$ то y является T-элементом.

3. Если $b_\alpha y \in K_{\alpha\beta}$, $yb_\alpha = b_\alpha$ и $\bar{y}$ состоит только из α, β то y является M-элементом, $\hat{y} = (\alpha, \beta)$.

4. Аналогично, если $b_\alpha y = b_\alpha$, $yb_\alpha \in K_{\alpha\beta}$, то y — M-элемент, $\hat{y} = (\beta, \alpha)$.

β) Пусть x_1, x_2 — L-элементы, $\alpha \in \bar{x}_1 \cap \bar{x}_2$, $\bar{x}_1 \neq \bar{x}_2$. Если существовали бы $\beta \in \bar{x}_1, \beta \bar{\in} \bar{x}_2$, $\gamma \bar{\in} \bar{x}_1$, $\gamma \in \bar{x}_2$ то $b_\beta x_1 = b_\beta$, $\langle\beta\rangle = \bar{b}_\beta = \overline{b_\beta x_1}$, итак $\overline{b_\beta x_1} \cap \bar{x}_2 = \emptyset$, по лемме $\bar{b}_\beta \cap \overline{x_1 x_2} = \emptyset$ и $x_1 x_2 = x_2$. Аналогично, $x_2 b_\gamma = b_\gamma$, $\bar{x}_1 \cap \overline{x_2 b_\gamma} = \emptyset$, по лемме $x_1 x_2 = x_1$. Значит $x_1 = x_2$, что неверно. Показали, что или $\bar{x}_1 \subseteqq \bar{x}_2$, или $\bar{x}_2 \subseteqq \bar{x}_1$. Пусть $\bar{x}_1 \subset \bar{x}_2$ (но $\bar{x}_1 \neq \bar{x}_2$) именно $\gamma \bar{\in} \bar{x}_1$, $\gamma \in \bar{x}_2$. Используя лемму для троек b_γ, x_2, x_1 и x_1, x_2, b_γ, получаем $x_1 x_2 = x_2 x_1 = x_1$. Пусть теперь $\bar{x}_1 = \bar{x}_2$. Очевидно, что произведения $x_1 x_2$, $x_2 x_1$ являются L-элементами, и так как они содержатся во всех полугруппах, в которых содержатся и x_1, и x_2, то

$$\overline{x_1 x_2} \supseteqq \bar{x}_1 = \bar{x}_2, \quad \overline{x_2 x_1} \supseteqq \bar{x}_1 = \bar{x}_2$$

γ) Пусть u, v $(u \neq v)$ — M-элементы. Предположим, что $\hat{u} = (\alpha, \beta)$, $\hat{v} = (\gamma, \alpha)$ $(\beta \neq \gamma)$. Используя лемму для элементов b_β, u, v получаем $uv = v$ и для элементов u, v, b_γ получаем $uv = u$. Следовательно $u = v$, что неверно.

δ) Пусть u — M-элемент, $\hat{u} = (\alpha, \beta)$ и x — L-элемент. Предположим, что $\alpha \in \bar{x}$, $\beta \bar{\in} \bar{x}$ и $\gamma \in \bar{x}$, $\gamma \neq \beta$. Используя лемму сначала для элементов b_β, u, x получаем $ux = x$, потом для u, x, b_γ получаем $ux = u$. Отсюда $u = x$, что неверно. Пусть теперь $\bar{u} \subset \bar{x}$ (но $\bar{u} \neq \bar{x}$), т. е. пусть $\gamma \in \bar{x}$, $\gamma \bar{\in} \bar{u}$. Из леммы для троек b_γ, x, u и u, x, b_γ следует $ux = xu = u$. Пусть наконец, $\bar{u} = \bar{x} = \langle\alpha, \beta\rangle$. Очевидно, что произведения xu, ux содержатся как в A_α, так и в A_β. Из определения M- и L-элементов следует, что xu, ux являются M-элементами и имеет место $\widehat{xu} = \widehat{ux} = \hat{u}$.

ε) Два первых утверждения свойства ε непосредственно следуют из определения M-элементов.

Рассмотрим третье. Пусть u, v — M-элементы, для которых $\hat{u} = (\alpha, \beta)$, $\hat{v} = (\alpha, \gamma)$ $\beta \neq \gamma$. Используем лемму для троек b_β, u, v и b_γ, v, u. Получается $uv = v$, $vu = u$. Доказательство аналогично в случае: $\hat{u} = (\beta, \alpha)$, $\hat{v} = (\gamma, \alpha)$ $(\beta \neq \gamma)$.

ζ) Пусть x — L-элемент, t — T-элемент. Предположим сначала, что $\alpha \in \bar{x} \cap \bar{t}$ и $\beta \in \bar{x}$, $\beta \bar{\in} \bar{t}$. Из леммы для троек b_β, x, t и t, x, b_β следует $xt = tx = t$. Пусть теперь $\bar{x} \subseteqq \bar{t}$. Из определений L- и T-элементов следует, что произведения xt, tx являются T-элементами, и одновременно очевидно, что $\overline{xt}, \overline{tx} \supseteqq \bar{x}$. Если существует $\gamma \in \bar{t}$, $\gamma \bar{\in} \bar{x}$, то используя лемму для троек b_γ, t, x и для x, t, b_γ получаем, что $\gamma \in \overline{xt}$, $\gamma \in \overline{tx}$, следовательно $\overline{xt} \supseteqq \bar{t}$, $\overline{tx} \supseteqq \bar{t}$.

η) Пусть u — M-элемент и t — T-элемент, и пусть $\hat{u} = (\alpha, \beta)$. Пусть сначала $\alpha \in \bar{t}$, $\beta \bar{\in} \bar{t}$. По лемме для элементов b_β, u, t имеем, что $ut = t$, а для элементов t, u, b_β, что $\overline{tu}\supseteqq\bar{u}$. Применяя лемму для всех троек b_ξ, t, u где $\xi\in\bar{t}$, получаем $\overline{tu}\supseteqq\bar{t}$. Значит $\overline{tu}\supseteqq\bar{t}\cup\bar{u}$. tu содержится по крайней мере в трех полугруппах из A_ξ $(\xi\in\mathscr{I})$, очевидно, что он T-элемент.

Второе утверждение из $\eta)$ доказывается аналогично.

Пусть теперь $\bar{u}\subseteqq\bar{t}$, $\bar{u}=\langle\alpha, \beta\rangle$

Из определения M- и T-элементов следует, что ut, tu — T-элементы. Очевидно, что $\overline{tu}, \overline{ut}\supseteqq\bar{u}$. Если существует $\xi\in\bar{t}$, $\xi\bar{\in}\bar{u}$ то из леммы для троек b_ξ, t, u и u, t, b_ξ следует, что $\xi\in\overline{tu}, \overline{ut}$ следовательно $\bar{t}\subseteqq\overline{tu}$, $\bar{t}\subseteqq\overline{ut}$.

ϑ) Пусть t_1, t_2 — T-элементы $\alpha\in\bar{t}_1\cap\bar{t}_2$. Очевидно, что $\overline{t_1t_2}\supseteqq\bar{t}_1\cap\bar{t}_2$, $\overline{t_2t_1}\supseteqq\bar{t}_1\cap\bar{t}_2$.

Пусть $\beta\in\bar{t}_1$, $\beta\bar{\in}\bar{t}_2$. Тогда применяя лемму для троек b_β, t_1, t_2 и t_2, t_1, b_β имеем $\beta\in\overline{t_1t_2}$, $\beta\in\overline{t_2t_1}$. Продолжая этот процесс получаем, что $\overline{t_1t_2}\supseteqq\bar{t}_1\cup\bar{t}_2$, $\overline{t_2t_1}\supseteqq\bar{t}_1\cup\bar{t}_2$. Если $\overline{t_1t_2}$ или $\overline{t_1t_2}$ состоит только из двух индексов α, β, т. е. $\langle\alpha, \beta\rangle=$ $=\bar{t}_1=\bar{t}_2$, то из определения следует, что t_1t_2, t_2t_1 являются T-элементами. Если t_1t_2 или t_2t_1 содержатся по крайней мере в трех полугруппах из A_ξ $(\xi\in\mathscr{I})$, то, очевидно, что они — T-элементы.

Теорема доказана.

Замечание. 1. Из того, что система полугрупп $A'=\{A_\xi\}_{\xi\in\mathscr{I}}$ является в полугруппе B независимой не следует, что каждая пара A_α, A_β является независимой. Действительно, существуют слова, приведенные относительно $\langle A_\alpha, A_\beta\rangle$, но не являющиеся приведенными относительно A'. Например, слово xy, где $x\in F_{\alpha\beta}$, $y\in F_{\beta\alpha}$, приведено относительно $\langle A_\alpha, A_\beta\rangle$, но возможно, что оно неприведено относительно A', если есть такая A_γ $(\gamma\in\mathscr{I})$ которая содержит и x, и y.

2. Рассмотрим следующий частный случай. Пусть A слабо ассоциативная амальгама не меньше трех полугрупп A_ξ $(\xi\in\mathscr{I})$ удовлетворяющая А), Б), в которой пересечение каждой пары полугрупп одно и то же, т. е. $A_\alpha\cap A_\beta=K$ для всех $\alpha, \beta\in\mathscr{I}$, $\alpha\neq\beta$. Тогда K содержит только L- и T-элементы. Множество T-элементов в K и для всех A_ξ $(\xi\in\mathscr{I})$ образует идеал (см. [9] пункт I. 4).

Утвеждения, вытекающие из вышедоказанных теорем являются обобщениями результатов статьи А. Грилле, М. Петрича [1].

Литература

[1] P. A. GRILLET, M. PETRICH, Free products of semigroups amalgamating an ideal, *J. London Math. Soc.*, (2) **2** (1970), 389—392.

[2] J. M. HOWIE, Embedding theorems with amalgamation for semigroups, *Proc. London Math. Soc.*, (3) **12** (1962), 511—534.

[3] J. M. HOWIE, Subsemigroups of amalgamated free products of semigroups, *Proc. London Math. Soc.*, (3) **13** (1963), 672—686.

[4] J. M. HOWIE, Embedding theorem with amalgamation for cancellative semigroups, *Proc. Glasgow Math. Assoc.*, **6** (1963), 19—26.

[5] Е. С. Ляпин, Независимость подполугрупп полугруппы, *ДАН СССР*, **185** (1969), 1229—1231.

[6] Е. С. Ляпин, Единично идеальные подполугруппы полугруппы, *Матем. записки Уральского ун-та*, **7** (1970), 119—128.

[7] Е. С. Ляпин, Пересечение независимых подполугрупп полугруппы, *Изв. ВУЗ-ов, Матем.*, **4** (1970), 67—73.

[8] Е. С. Ляпин, Независимые полугрупповые продолжения частичных группоидов, *Современная алгебра*, **2** (1974), 55—71.

[9] Л. Медеши, Расширение полугрупп при помощи разделяющейся единично идеальной подполугруппы, *Современная алгебра*, **2** (1974), 72—97.

BOLYAI INTÉZET
ARADI VÉRTANÚK TERE 1
H-6720 SZEGED

Acta Sci. Math., 39 (1977), 341—351

Similarity and interpolation between projectors

Z. V. KOVARIK

1. Introduction

If two bounded linear projectors E_0, E_1 on a Banach space X are "close" in some sense, are they similar? Can we connect them by a projector-valued continuous path $t \to E_t (0 \leqq t \leqq 1)$? If $\|E_1 - E_0\| < 1$ for some operator norm then [6, I. 4.6, with the footnote and Problem 4.13] gives a positive answer to both questions. For pairs of orthoprojectors on a Hilbert space, the papers [1] and [2, Sec. 3] give a complete set of unitary invariants, with an extensive bibliography on the history of the subject, and the latter work (formula 1.18) expresses a particular unitary U (called direct rotation) such that $E_1 = UE_0U^{-1}$ in the form

$$U = \exp(J\theta), \quad \theta \geqq 0, \ J \text{ normal}, \quad J^3 = -J,$$

which offers a path $E_t = \exp(tJ\theta)E_0 \exp(-tJ\theta)$.

We shall give a similar expression for Banach space projectors in (12), except that we do not try to separate J from θ in $J\theta = -iW$, essentially because square roots of arbitrary operators are generally unavailable.

Differentiable paths between finite decompositions of identity into projectors and their relation to similarity are studied in [6, II. 4.5].

We base our exposition on the concept of an approximate projector (Sec. 2) from which we derive an expression of the bisector $E_{1/2}$ of E_0 and E_1. Bisections are then repeated, leading to $E_{1/4}$, $E_{3/4}$ etc., until analytic operational calculus is applicable to help extend the domain of E_t to all $t \in [0, 1]$. Throughout, we use that holomorphic branch of the natural logarithm whose value at 1 is 0. Accordingly, we define $z^{1/2} = \exp(1/2 \ln z)$ etc.

Received July 14, 1976.

This research was sponsored in part by the National Research Council of Canada, Grant #1861, and by the Summer Research Institute at the University of Victoria.

2. Approximate projectors and involutions

We call $P\in B(X)$ an *approximate projector* if $\|P-P^2\|<1/4$, and $Q\in B(X)$ an *approximate involution* if $\|1-Q^2\|<1$. The correspondence between approximate projectors and involutions is obviously $Q=2P-1$.

The spectrum $\sigma(P)$ of an approximate projector lies, by the Spectral Mapping Theorem, inside the Bernoulli lemniscate L: $|z(1-z)|=1/4$. Let us compute the spectral projector E of P corresponding to the part of $\sigma(P)$ inside the right loop L_+: $|z(1-z)|=1/4$, $\operatorname{Re}(z)\geqq 1/2$. We have (orienting L_+ correctly)

$$E=(2\pi i)^{-1}\int_{L_+}(z-P)^{-1}dz.$$

Denote $P-P^2=R$ and substitute $z-z^2=v$, hence $z=\frac{1}{2}(1+\sqrt{1-4v})$, $dz=-(1-4v)^{-1/2}dv$, and we can verify

$$(z-P)^{-1}=(1-z-P)(v-R)^{-1}.$$

The path L_+ transforms into the positively oriented circle C: $|v|=1/4$. By virtue of the compactness of $\sigma(P)$, we can deform homotopically C into a smaller circle C_1 to avoid the singular point $v=1/4$. Altogether,

$$E=(2\pi i)^{-1}\int_{C_1}[\tfrac{1}{2}+\tfrac{1}{2}(2P-1)(1-4v)^{-1/2}](v-R)^{-1}dv,$$

i.e.

$$E=\tfrac{1}{2}+\tfrac{1}{2}(2P-1)(1-4R)^{-1/2}=P+\tfrac{1}{2}(2P-1)[(1-4R)^{-1/2}-1]. \tag{1}$$

Note that $(1-4R)^{-1/2}$ can be obtained by evaluating the MacLaurin series for $(1-4v)^{-1/2}$ at R in place of v.

For the involution $T=2E-1$, (1) simplifies into

$$T=Q(1-(1-Q^2))^{-1/2}\quad\text{where}\quad Q=2P-1. \tag{2}$$

Since the power series above has positive coefficients, we have the following estimate where $r=\|P-P^2\|$:

$$\|E-P\|\leqq\tfrac{1}{2}\|2P-1\|[(1-4r)^{-1/2}-1]. \tag{3}$$

Intuitively, the closer P is to being a projector, the closer it is to E.

Taking first two terms in the power series, we derive an iterative scheme for computing E, namely:

(4) Set $P_0=P$, and given P_k, compute $R_k=P_k-P_k^2$. Stop if $\|E-P_k\|$ from (3) is satisfactorily small. Else, compute $P_{k+1}=P_k+(2P_k-1)R_k\equiv P_k^2+2P_kR_k$ and return to testing P_{k+1}.

We can prove that $R_{k+1}=R_k^2(3+4R_k)$, hence from $\|R_k\|<1/4$ there follows $4\|R_{k+1}\|\leqq(4\|R_k\|)^2$ (quadratic convergence), and that (1) applied to P_k yields the same E.

3. The bisector of two close projectors

For the sake of exposition, we shall assume that E_0 and E_1 are two projectors satisfying $\|E_1-E_0\|<1$. Their mean $P=\frac{1}{2}(E_0+E_1)$ need not be a projector but it is an approximate projector in the sense of the previous section. Indeed, we can verify directly that $P-P^2=\frac{1}{4}(E_1-E_0)^2$, hence $\|P-P^2\|<\frac{1}{4}$, and using (1), we can define

(5) $$E_{1/2}=\tfrac{1}{2}(E_0+E_1)+\tfrac{1}{2}(E_0+E_1-1)\left[(1-(E_1-E_0)^2)^{-1/2}-1\right]$$

and call $E_{1/2}$ the *bisector* of E_0 and E_1. For the associated involution $T_{1/2}=2E_{1/2}-1$, we obtain

(6) $$T_{1/2}=(E_0+E_1-1)[1-(E_1-E_0)^2]^{-1/2}=\tfrac{1}{2}(T_0+T_1)[1-\tfrac{1}{4}(T_1-T_0)^2]^{-1/2}.$$

4. The trigonometry of projectors

With a pair of projectors E_0, E_1 and their involutions $T_i=2E_i-1$ $(i=0, 1)$, we associate the following operators (compare [1]):

(7) $S_1=(E_1-E_0)^2$, the *separation* of (E_0, E_1);
$C_1=(E_0+E_1-1)^2$, the *closeness* of (E_0, E_1);
$V_1=T_0T_1$;

and we can verify the following properties:

(8) (i) $C_1=\frac{1}{4}(T_0+T_1)^2=\frac{1}{4}(2+V_1+V_1^{-1})$; $S_1=\frac{1}{4}(T_1-T_0)^2=\frac{1}{4}(2-V_1-V_1^{-1})$;

(ii) $C_1+S_1=1$;

(iii) both C_1 and S_1 commute with $\{E_0, E_1\}$;

(iv) C_1 and S_1 are symmetric functions of E_0 and E_1;

(v) $C_1E_0=E_0E_1E_0$ and $C_1E_1=E_1E_0E_1$.

We can think of S_1 as the operator analogue of $\sin^2\theta$ and C_1 as $\cos^2\theta$ where θ is the non-obtuse angle between the ranges of E_0 and E_1. If E_i are one-dimensional ortho-projectors on the Euclidean plane, the analogy is perfect (we identify number 1 with the identity operator if convenient), and higher dimensional pairs of orthoprojectors will essentially decompose into direct sums or integrals of planar pairs, as shown in [1]. Continuing in this analogy, we may give V_1 the meaning of a generalization of the square of the direct rotation that moves the range of E_0 onto the range of E_1. However, the convenient property $0 \leqq S_1 \leqq 1$ and the "angle" $\arcsin(S_1^{1/2})$, which would commute with both E_0 and E_1, lose their meaning in general Banach spaces, as we can see from Example 1. We fill in this gap partially by developing an oblique way of expressing $\sin^2(t\theta)$ for $0 \leqq t \leqq 1$ in terms of $\sin^2\theta$ without having to evaluate θ itself.

5. An auxiliary function

For $z \in (0, 1)$ and t complex, we define

$$f_t(z) = \sin^2(t \arcsin \sqrt{z}). \tag{9}$$

To extend the domain of f, we observe that the function $g(z) = \frac{1}{\sqrt{z}} \arcsin \sqrt{z}$ $(z \neq 0)$, $g(0) = 1$ has a MacLaurin series with radius of convergence equal to 1, and $h(z) = \sin^2 \sqrt{z}$ has a MacLaurin series convergent for all z. With g and h extended by means of their expansions, we can extend $f_t(z)$ as

$$f_t(z) = h\big(t^2 z g^2(z)\big) \tag{10}$$

for all $|z| < 1$, so that f_t is holomorphic. To guarantee uniqueness of further continuations of f_t, we consider a simple smooth curve Γ connecting 1 with ∞ while missing 0 and define

$$\Delta = \mathbf{C} \setminus \operatorname{range}(\Gamma),$$

so that Δ is a simply connected domain in which f_t is arbitrarily continuable. By the Monodromy Theorem ([5], VI. 6.3), f_t can be continued to a holomorphic function on Δ. We retain, with Γ fixed, the notation of (9) for this holomorphic function.

We remark that for $\Gamma = \Gamma_0 = [1, +\infty]$, the function g from above can be extended from $0 < z < 1$ to $\Delta_0 = \mathbf{C} \setminus \Gamma_0$ by means of the series

$$g(z) = \frac{2}{\sqrt{z}} \arctan \frac{\sqrt{z}}{1 + \sqrt{1-z}} = 2 \sum_{n=0}^{\infty} \frac{(-z)^n}{(2n+1)\big(1+\sqrt{1-z}\big)^{2n+1}}$$

since $|z(1+\sqrt{1-z})^{-2}| < 1$ iff $z \in \Delta_0$. Thus $f_t(z)$ is described by (10) on all of Δ_0.

Here is a list of properties of f_t which will be useful later.

(11) (i) For sufficiently small $|\theta|$, if $z = \sin^2\theta$ then $f_t(z) = \sin^2(t\theta)$;

(ii) $f_0(z)=0, f_1(z)=z, f_2(z)=4z(1-z)$ (in general, f_t is a polynomial in z for every integer t),

(iii) $f_{1/2}(z) = \frac{1}{2}(1-\sqrt{1-z}$; note that $\sqrt{1-z}$ is well-defined on Δ;

(iv) for all s, t and for $z\in\Delta$, $|z|$ sufficiently small, $f_s(f_t(z))=f_{st}(z)$; this gives an unambiguous continuation of $f_s \circ f_t$ to all of Δ; the composition is correct in all of Δ if s is an integer;

(v) if $t=2^{-n}$, $n\geqq 0$ integer, then $f_t(z)\neq 1$ $(z\in\Delta)$; if $n\geqq 1$ then $f_t(z)\neq 1/2$. (This follows by induction.)

6. Similarity and interpolation

For a set $M\subset B(X)$, denote by $\mathscr{A}\{M\}$ the norm-closed sub-algebra of $\mathscr{B}(X)$ generated by M and the identity and closed under the inversion if defined. Recall also the notation in (7).

Theorem 1. *Let E_0, E_1 be projectors in $\mathscr{B}(X)$, $E_0\neq E_1$.*

(i) *If the number 1 lies in the unbounded component of the complement of $\sigma(S_1)$ then there exists an involution $T_{1/2}$ in $\mathscr{A}\{E_0, E_1\}$ such that $E_1=T_{1/2}E_0T_{1/2}$ and there exists $W\in\mathscr{A}\{V_1\}$ such that the projector-valued path*

$$t \to E_t = e^{-itW}E_0e^{itW}, \quad 0\leqq t\leqq 1, \tag{12}$$

connects E_0 with E_1. Moreover, $V_1=e^{2iW}$ and $T_tW=-WT_t$ where $T_t=2E_t-1$.

(ii) *(Poor man's path): If E_0+E_1-1 is invertible then there exists Z which is a product of two involutions from $\mathscr{A}\{E_0, E_1\}$ such that $E_1=Z^{-1}E_0Z$, and there exists a projector-valued path $t\to E_t$, $0\leqq t\leqq 1$, connecting E_0 with E_1 and consisting of two straight line segments.*

Remarks.

(a) The condition $\|E_1-E_0\|<1$ clearly implies the assumption (i) of the theorem, for then $\|S_1\|<1$.

(b) Assumption (i) implies assumption (ii) because $C_1=(E_0+E_1-1)^2=1-S_1$ does not have 0 in its spectrum (recall (7) and (8) (ii)). We shall see a counterexample demonstrating that (ii) does not imply (i).

(c) In (ii), the inequality $E_0\neq E_1$ implies that E_0 does not commute with E_1. In fact, more than that is true: If $E_0E_1=E_1E_0$ and (E_0+E_1-1) is either left- or right-cancellable then $E_0=E_1$. Indeed, look at the identities:

$$(E_0+E_1-1)(E_0-E_1) = (E_1E_0-E_0E_1) = (E_1-E_0)(E_0+E_1-1).$$

(d) The operator W from (i) is an operator analogue of the angle between E_0 and E_1. Indeed, from (8i) and the equation $V_1=e^{2iW}$ there follows that $\sin^2 W = S_1$, except that W need not be in $\mathcal{A}\{S_1\}$ because $T_0W=-WT_0$ while S_1 commutes with T_0.

Proof. (Theorem 1). In part (i), let us choose a simple smooth curve Γ which connects 1 with ∞ within the complement of $\sigma(S_1)\cup\{0\}$. As in section 5, we can construct the function f relative to Γ. For the construction of W, as well as for future use, we define

$$(13)\qquad S_t=f_t(S_1)\quad \text{for all (complex) } t.$$

According to (11v), for $t=2^{-n}$ ($n\geqq 0$ integer) the operator $1-2S_{t/2}$ is invertible, and since $S_t=f_2(S_{t/2})=4S_{t/2}(1-S_{t/2})$ by (11ii and 11iv), we have

$$(14)\qquad (1-2S_{t/2})^2=1-S_t.$$

All S_t commute with T_0, T_1 by (8iii).

By induction, we will now construct involutions T_t for $t=2^{-n}$, $n\geqq 0$ integer. They too will commute with all S_t. T_1 is already given and $1/4(T_0+T_1)^2=1-S_1$ by (8). Assume that for $t=2^{-n}$, T_t has been defined and

$$(15)\qquad T_t^2=1,\tfrac{1}{4}(T_0+T_t)^2=1-S_t$$

holds. Along with T_t, we consider $V_t=T_0T_t$, and define

$$(16)\qquad T_{t/2}=\tfrac{1}{2}(T_0+T_t)(1-2S_{t/2})^{-1}.$$

Note that $1-2S_{t/2}$ is invertible. Using (14) and (15), we verify mechanically that $T_{t/2}^2=1$ and $\frac{1}{4}(T_0+T_{t/2})^2=1-S_{t/2}$. For the associated V_t, we have

$$(17)\qquad V_{t/2}=\tfrac{1}{2}(1+V_t)(1-2S_{t/2})^{-1}=T_0T_{t/2}=T_{t/2}T_t,$$

so that

$$(18)\qquad V_{t/2}^2=T_0T_{t/2}T_{t/2}T_t=V_t.$$

Thus, T_t and V_t are constructed for all $t=2^{-n}$ and, in addition, $V_t\in\mathcal{A}\{V_1\}$ and $V_{2^{-n}}^{2^n}=V_1$. Also, for $t=1$, (17) implies

$$T_1=T_{1/2}T_0T_{1/2},\quad \text{so that}\quad E_1=T_{1/2}E_0T_{1/2},$$

as claimed in the theorem.

The next task is to show that $\lim_{n\to\infty} V_{2^{-n}}=1$ in norm. Indeed, writing $V_{2^{-n}}=1+K_n$, we re-write (17) as

$$K_{n+1}=\tfrac{1}{2}(1-2S_{t/2})^{-1}K_n+2S_{t/2}(1-2S_{t/2})^{-1}\quad \text{with}\quad t=2^{-n}.$$

Since $\lim_{t\to 0} S_t = \lim_{t\to 0} S_t(1-2S_t)^{-1}=0$ in norm, we have for all sufficiently large n,

$$\|\tfrac{1}{2}(1-2S_t)^{-1}\| \leqq \frac{2}{3}, \quad \text{and} \quad \|K_{n+1}\| \leqq \frac{2}{3}\|K_n\|+p_n \quad \text{where} \quad \lim_{n\to\infty} p_n = 0.$$

This implies $\lim_{n\to\infty} \|K_n\|=0$, as claimed.

We select an n_0 such that $\|V_{2^{-n}}-1\|<1$ for all $n\geqq n_0$ and define $W_n=-i2^{n-1}\cdot \ln V_{2^{-n}}$. Due to (18), W_n is independent of $n\geqq n_0$, and we claim that the common value is the desired W. The relations $V_1=e^{2iW}$ and $E\in\mathcal{A}\{V_1\}$ are immediate. Further, (17) implies

$$T_t = V_{t/2}^{-1}T_0V_{t/2} = e^{-itW}T_0e^{itW} \tag{19}$$

for $t=2^{-n}$, $n\geqq 0$ integer. We can therefore use (19) to define T_t and $V_t=T_0T_t$ for all t, so that T_t are involutions, $E_t=e^{-itW}E_0e^{itW}$ are projectors and the equation is correct for both $t=0$ and $t=1$.

For $n\geqq n_0$, $t=2^{-n}$, we also have $T_0V_tT_0=V_t^{-1}$ (since $T_0V_t=T_t$ is an involution), and taking logarithms on both sides, we obtain $T_0WT_0=-W$, or $T_0W=-WT_0$, hence

$$T_tW+WT_t = e^{-itW}(T_0W+WT_0)e^{itW} = 0,$$

completing the proof of part (i).

Proof of part (ii): Since E_0+E_1-1 has an inverse in $\mathcal{B}(X)$, so does $C_1=(E_0+E_1-1)^2$, and we claim:

$F=E_0C_1^{-1}E_1$ is a projector, $(E_1-F)^2=0=(F-E_0)^2$, $E_0F=FE_1=F$, $FE_0=E_0$, $E_1F=E_1$. Indeed, by (8 iii and v), $F^2=C_1^{-2}(E_0E_1E_0)E_1=C_1^{-2}C_1E_0E_1=F$, and the remaining statements follow similarly.

Consequently, E_0+F-1 and E_1+F-1 are involutions, and we can set $Z=(E_0+F-1)(E_1+F-1)$ which makes $E_1=Z^{-1}E_0Z$. The straight line segment from E_0 to F consists of projectors since

$$(E_0+t(F-E_0))^2 = E_0+tE_0(F-E_0)+t(F-E_0)E_0+t^2(F-E_0)^2 = E_0+t(F-E_0)$$

by the above equations, and similarly the line segment from F to E_1 consists of projectors. The proof is complete.

More remarks. (a) In part (i), the projectors E_t move from E_0 to E_1 at a constant angular velocity in the sense that for s and t sufficiently close, the angle operator between E_s and E_t is

$$\frac{1}{2i}\ln(T_sT_t) = \frac{1}{2i}\ln e^{2i(t-s)W} = (t-s)W,$$

and the separation is $(E_t-E_s)^2=S_{t-s}$, as we can verify from the relations (19) defining T_t and from $T_0W=-WT_0$ which implies $T_0e^{-itW}=e^{-itW}T_0$. It may be interesting that

$$E_sE_t-E_tE_s=\tfrac{1}{2}i\sin 2(t-s)W.$$

(b) From $T_0WT_0=-W$ it follows that the spectrum of W is centrally symmetric, including its fine structure. For example, if ω is an eigenvalue of W with an eigenvector x then T_0x is an eigenvector corresponding to $(-\omega)$.

(c) If $S_1=0$, as between E_0 and F in the proof of part (ii) of the theorem, then the angle operator $W=-i(E_0E_1-E_1E_0)$, but $(E_0E_1-E_1E_0)^2=0$ and hence $W^2=0$. In this case, the path (12) becomes the straight line segment from E_0 and E_1.

We can describe those projector pairs (E_0, E_1) for which $S_1=0$ as follows. E_0 defines a direct decomposition $X=\text{Range } E_0\oplus\text{Ker } E_0$, so that E_0 is represented by $1\oplus 0$. Then E_1 can be represented by

$$\begin{bmatrix}1 & A\\ B & 0\end{bmatrix} \quad\text{where}\quad AB=0 \quad\text{and}\quad BA=0,$$

as we can verify by writing $E_1^2=E_1$ and $(E_1-E_0)^2=0$ in components.

(d) The possible non-uniqueness of the part Γ in part (i) may cause non-uniqueness of the path (12). It would be interesting to find conditions on the pair (E_0, E_1) which would allow to characterize at least $T_{1/2}$ or $V_{1/2}$ before it is constructed, in a similar way as its unitary counterpart for self adjoint pairs was described in [2, Prop. 3.3] and [1, Theorem 4.1].

(e) The projector $F=E_0C_1^{-1}E_1$ can be obtained as the spectral projector of E_0E_1 corresponding to the complement of $\{0\}$. Indeed, using property $(E_0E_1)^2=$ $=C_1E_0E_1$ from (8v), we can verify a partial fraction decomposition

$$(\lambda-E_0E_1)^{-1}=\lambda^{-1}(1-F)+(\lambda-C_1)^{-1}F$$

and obtain $1-F$ as $(2\pi i)^{-1}\int_\Gamma(\lambda-E_0E_1)^{-1}d\lambda$ where Γ is a small circle around 0.

F is the unique projector which shares its range with E_0 and its nullspace with E_1. An equivalent construction of F is implied in [6, Problem I. 4.12], namely

$$F=(1-E_1+E_0E_1)E_1(1-E_1+E_0E_1)^{-1}.$$

7. Examples

Example 1. For every complex λ_0, there exist two 2×2 idempotent matrices e_0, e_1 such that $(e_1-e_0)^2=\lambda_0 1$ and their Euclidean bound norms are

(20) $$\|e_0\|=\|e_1\|=[\tfrac{1}{2}(1+|\lambda_0|+|\lambda_0-1|)]^{1/2}.$$

For a construction, take any δ satisfying $(\delta-\delta^{-1})^2=-4\lambda_0$, and define

$$e_0=1/2\begin{bmatrix}1 & \delta^{-1}\\ \delta & 1\end{bmatrix} \quad\text{and}\quad e_1=1/2\begin{bmatrix}1 & \delta\\ \delta^{-1} & 1\end{bmatrix}.$$

It can be proved that for every pair of idempotent 2×2 matrices with a prescribed separation λ_0, the maximum of their norms is not less than the quantity in (20).

Example 2. Given a non-empty compact set K in the plane, there exist two projectors E_0 and E_1 on the space l_2 for which the spectrum of $(E_1-E_0)^2$ is K, and its point spectrum is dense in K. These projectors can be built as direct sums of examples of type 1 over a sequence $\{\lambda_0^{(n)}\}$ whose range is dense in K. Boundedness is guaranted by (20).

A rich source of examples is afforded by a pair (φ,ψ) of formal expressions in x with values in a commutative algebra with identity such that

$$\varphi(-x)=-\varphi(x) \quad\text{and}\quad \varphi^2(x)+\psi(x)\psi(-x)=1, \tag{21}$$

by means of which we can define an involution T:

$$Tf(x)=\varphi(x)f(x)+\psi(x)f(-x). \tag{22}$$

The expressions f, φ and ψ may be functions defined on a centrally symmetric set in the plane, or formal power series with complex coefficients, or elements of suitable subspaces and quotient spaces of the above, and direct sums thereof.

Example 3. A pair E_0, E_1 of projectors on a Banach space for which $\sigma(S_1)$ is a prescribed non-empty compact set K, and every interior point λ of K belongs to the residual spectrum of S_1, i.e. $\lambda-S_1$ is one-to-one but its range is not dense.

Here we first construct $L=\{\lambda \text{ complex}: \lambda^2\in K\}$, take $X=H_\infty(L)$, the space of functions continuous on L and holomorphic in the interior of L, with supremum norm, and define

$$\begin{aligned} T_0f(x)&=-xf(x)+(1+x)f(-x),\\ T_1f(x)&=xf(x)+(1+x)f(-x), \quad x\in L,\quad f\in X.\end{aligned} \tag{23}$$

The conditions (21) are met, $E_i=(1+T_i)/2$ as usual $(i=0, 1)$, and $S_1f(x)=\frac{1}{4}(T_1-T_0)^2f(x)=x^2f(x)$. Consequently $\sigma(S_1)=K$ with residual spectrum as claimed. Starting with an arbitrary nonempty compact centrally symmetric set L, we have an example of $\sigma(E_1-E_0)=L$.

Example 4. If $X=L^2(D)$ where D is the unit disc with the restricted Lebesgue measure and if T_0, T_1 are as in (23) then $\sigma(S_1)=D$ and every $\lambda\in D$ is in the continuous spectrum of S_1 (i.e. $\lambda-S_1$ has a densely defined unbounded inverse).

Example 5. The closeness operator C_1 can be invertible but can fail to have a square root in $\mathscr{A}\{S_1\}$ so that a bisector $E_{1/2}$ from (5) cannot exist.

In Example 3, take

$$L = \{\lambda \text{ complex}: |\lambda| \leqq 2, \quad |\lambda - 1| \geqq \tfrac{1}{2}, \quad |\lambda+1| \geqq \tfrac{1}{2}\}.$$

so that from (23) there follows $C_1 f(x) = (1-x^2)f(x)$. The function $(1-x^2)$ has a reciprocal but not a square root in L, and every operator in $\mathscr{A}\{S_1\}$ is of the form (22) with $\psi = 0$.

Example 6. Every even and every odd non-constant polynomial p can be the minimal polynomial of $(E_1 - E_0)$ for suitable E_0 and E_1 acting on a space of dimension degree (p). Consequently, every nilpotent matrix can be written as the difference of two idempotent matrices.

Indeed, take $X = \mathbf{C}[x]/(p)$, the polynomials modulo p. The operators from (23), well-defined on $\mathbf{C}[x]$, leave (p) invariant if p is either even or odd, hence they induce involutions $T_i^{(p)}$ $(i=0, 1)$ on X; note that $\dim X = \text{degree}\,(p)$. For the corresponding projectors, we have $(E_1^{(p)} - E_0^{(p)})[f(x)] = [xf(x)]$, so that the minimal polynomial of $(E_1^{(p)} - E_0^{(p)})$ is indeed p.

The proof for nilpotent matrices can be reduced to the case of one Jordan cell in the normal form and then reconstructed via direct sums. But every Jordan cell of dimension k represents multiplication by x in $\mathbf{C}[x]/(x^k)$, which is the previous case.

Example. 7. Take a non-void compact centrally symmetric set L in the plane and define in $H_\infty(L)$: $T_0 f(x) = f(-x)$ and $T_1 f(x) = e^{-2ix} f(-x)$. It is obvious that $T_t f(x) = e^{-2tix} f(-x)$ $(0 \leqq t \leqq 1)$ defines a path from T_0 to T_1 consisting of involutions. Although we can define $Wf(x) = xf(x)$, so that $T_0 T_1 = e^{2iW}$, $T_t W = -WT_t$ and (12) holds, we cannot always obtain W by a procedure described in Theorem 1. Observe that $S_1 f(x) = (\sin^2 x) f(x)$, and the proof breaks down if L contains points $(2k+1)\pi/2$, k integer.

8. Further connections between similarity and interpolation

Theorem 2. *If there exists a continuous projector-valued path $t \to E_t$, $0 \leqq t \leqq 1$, then E_0 and E_1 are similar via a finite product of involutions. On the other hand, if two bounded linear operators A_0, A_1 on a complex Hilbert space H are similar (or unitarily equivalent) then there exists a continuous path $t \to A_t$, $0 \leqq t \leqq 1$, consisting of operators similar (or unitarily equivalent) to A_0.*

Proof. For the first part, we use uniform continuity of E to subdivide [0, 1] into $0 = t_0 < t_1 < \ldots < t_p = 1$ so that $\|E_{t_{k+1}} - E_{t_k}\| < 1$, $k = 0, \ldots, p-1$, and then we apply Theorem 1 p times.

For the rest, write $A_1=T^{-1}A_0T$ and decompose $T=RV$ with R positive definite and V unitary. Evidently $t\to R^t$, $0\leqq t\leqq 1$, connects 1 with R while $t\to R^{-1}$ connects 1 with R^{-1}. Next, V can be written as the product of at most four self-adjoint involutions if $\dim H=\infty$, as shown in [3]. Every involution $Q=2E-1$ can in turn be connected with 1 by $U_t=E+e^{i\pi t}(1-E)$, $0\leqq t\leqq 1$, so that U_t are unitary. If $\dim H<\infty$, the unitary group of H is known to be arcwise connected. Thus, 1 can be connected with T be a continuous path consisting of invertible operators, and if T is unitary then $R=1$. Hence the conclusion.

Application. Proposition 2 in [4] states that if a pathwise connected subset $\mathcal{T}\subseteqq B(X)$ has the property that the union S of $\sigma(T)$ over $T\in\mathcal{T}$ has a bounded separated subset M then the spectral invariant subspaces of the operators T corresponding to M are mutually homeomorphic and homotopic. Using Theorem 2, we can strengthen the conclusion by asserting the similarity of the corresponding spectral projectors.

Acknowledgements. The author is thankful to Professor Chandler Davis and to the referee for their criticism which helped improve the text and correct a false statement in the example section.

References

[1] C. Davis, Separation of two linear subspaces, *Acta Sci. Math.*, **19** (1958), 172—187.
[2] C. Davis and W. M. Kahan, The rotation of eigenvectors by a perturbation. III, *SIAM J. Numer. Anal.*, (1970).
[3] P. R. Halmos and S. Kakutani, Products of Symmetries, *Bull. Amer. Math. Soc.*, **64** (1958), 77—78.
[4] Z. V. Kovarik, Spectrum localization in Banach spaces, I., *Lin. Alg. and its Appl.*, **8** (1974), 225—236.
[5] S. Saks and A. Zygmund, *Analytic functions*, Polish Scientific Publishers (Warsaw, 1965).
[6] T. Kato, *Perturbation theory for linear operators*, Springer-Verlag (New York, 1966).

MCMASTER UNIVERSITY
HAMILTON, ONTARIO, CANADA

Acta Sci. Math., 39 (9177), 353—366

Moment inequalities for the maximum of partial sums of random fields

F. MÓRICZ

§ 1. Introduction and preliminaries

Let $\{\zeta_{kl}\}$ $(k, l=1, 2, \ldots)$ be a random field. It is not assumed that the random variables (in abbreviation: rv's) ζ_{kl} are mutually independent or identically distributed. Set

$$S(b, m;\ c, n) = \sum_{k=b+1}^{b+m} \sum_{l=c+1}^{c+n} \zeta_{kl}$$

and

$$M(b, m;\ c, n) = \max_{1 \leqq p \leqq m} \max_{1 \leqq q \leqq n} |S(b, p;\ c, q)|,$$

where $b, c \geqq 0$, and $m, n \geqq 1$ are integers.

The subject of this paper is to provide bounds on $E\big(M^\gamma(b, m;\ c, n)\big)$ in terms of given bounds on $E|S(b, m;\ c, n)|^\gamma$, where γ is a given positive exponent. We emphasize that the only restrictions on the dependence will be those imposed by the assumed bounds for $E|S(b, m;\ c, n)|^\gamma$. These assumed bounds are guaranteed under a suitable dependence restriction, e. g., martingale difference, multiplicativity of finite order, orthogonality, mixing condition, or the like.

Bounds on $E\big(M^\gamma(b, m;\ c, n)\big)$ are of use in deriving convergence properties of $S(m, n)=S(0, m;\ 0, n)$ as $m, n \to \infty$, probability inequalities for $M(b, m;\ c, n)$, and tightness criteria for certain sequences of random functions (see [3]). To develop such results under various dependence restrictions, the theorems of this paper reduce the problem of placing appropriate bounds on $E\big(M^\gamma(b, m;\ c, n)\big)$ to the much easier problem of placing appropriate bounds on $E|S(b, m;\ c, n)|^\gamma$. Various applications of our theorems, for example to obtain strong laws of large numbers, will be presented in a subsequent paper [8].

Received July 6, 1976.

The proofs are based on the "bisection" technique, which goes back to Rademacher and Menšov; see, e.g., BILLINGSLEY [3, pp. 87—103]. The treatment is similar to [6]. The results obtained can be considered as extensions of those in [6] from sequences $\{\xi_j\}$ of rv's to random fields $\{\zeta_{kl}\}$.

In the following, $f(b, m; c, n)$ will denote a non-negative function depending on the joint distribution function (in abbreviation: df) of $\{\zeta_{kl}: k=b+1, \ldots, b+m; l=c+1, \ldots, c+n\}$, and possessing the following two properties of a rather general nature:

$$f(b, h; c, n)+f(b+h, m-h; c, n) \leqq f(b, m; c, n) \tag{1.1}$$

and

$$f(b, m; c, i)+f(b, m; c+i, n-i) \leqq f(b, m; c, n) \tag{1.2}$$

for all $b \geqq 0$, $1 \leqq h < m$ and $c \geqq 0$, $1 \leqq i < n$. In other words, condition (1.1) means that $f(b, m; c, n)$ as a function of the interval $(b+1, b+m)$ is "superadditive" for fixed c and n, while (1.2) expresses the superadditivity in $(c+1, c+n)$ for fixed b and m. Examples are $f(b, m; c, n)=m^{\beta_1}n^{\beta_2}$ with $\beta_1, \beta_2 \geqq 1$ or $f(b, m; c, n)= = \sum_{k=b+1}^{b+m} \sum_{l=c+1}^{c+n} \sigma_{kl}^2$, in the latter case assuming the existence of the finite variances σ_{kl}^2 of the rv's ζ_{kl}.

The upper bound on $E|S(b, m; c, n)|^\gamma$ will be considered in the general form

$$E|S(b, m; c, n)|^\gamma \leqq f^\alpha(b, m; c, n), \tag{1.3}$$

where α and γ are given numbers, $\alpha \geqq 1$, $\gamma > 0$, and $f(b, m; c, n)$ satisfies (1.1)—(1.2).

The treatment of the case $0 < \gamma \leqq 1$ is quite simple. In fact, applying the well-known inequality

$$E|\xi+\eta|^\gamma \leqq E|\xi|^\gamma + E|\eta|^\gamma \quad (0 < \gamma \leqq 1),$$

we obviously have

$$E\big(M^\gamma(b, m; c, n)\big) \leqq \sum_{k=b+1}^{b+m} \sum_{l=c+1}^{c+n} E|\zeta_{kl}|^\gamma \leqq \sum_{k=b+1}^{b+m} \sum_{l=c+1}^{c+n} f^\alpha(k-1, 1; l-1, 1),$$

provided (1.3) holds for all $b, c \geqq 0$ and $m, n \geqq 1$. Now using (1.1)—(1.2) and the elementary inequality

$$\sum_{i=1}^{r} u_i^\alpha \leqq \Big(\sum_{i=1}^{r} u_i\Big)^\alpha, \quad \text{where} \quad \alpha \geqq 1 \quad \text{and} \quad u_i \geqq 0 \quad (i = 1, 2, \ldots, r),$$

we arrive at the following result.

Theorem 1. *Suppose that there exists a non-negative function $f(b, m; c, n)$ satisfying* (1.1)—(1.2) *such that* (1.3) *holds for all $b, c \geqq 0$ and $m, n \geqq 1$, where $\alpha \geqq 1$*

and $0<\gamma\leqq 1$. *Then we have*

$$E\big(M^{\gamma}(b, m;\ c, n)\big) \leqq f^{\alpha}(b, m;\ c, n)$$

for all $b, c\geqq 0$ *and* $m, n\geqq 1$.

Hence the case $\alpha\geqq 1$ and $\gamma>1$ is of interest. The subcases (i) $\alpha>1$ and (ii) $\alpha=1$ are discussed in the subsequent sections. Section 4 is devoted to the estimation of the maxima of square sums and spherical sums, respectively. In the last section we point out that the results can be extended in a natural way to the general multi-parameter case from the two-parameter case, and there is no need to restrict the theorems proved to finite measures.

Throughout the paper, $C, C_1, \ldots$ will denote positive constants, not necessarily the same at different occurrences.

§ 2. An asymptotically optimal inequality in the case $\alpha>1$

Theorem 2 below provides a bound on $E\big(M^{\gamma}(b, m;\ c, n)\big)$ which is asymptotically optimal as $m, n\to\infty$ in the sense that it is of the same order of magnitude as the bound assumed on $E|S(b, m;\ c, n)|^{\gamma}$.

Theorem 2. *Suppose that there exists a non-negative function* $f(b, m;\ c, n)$ *satisfying* (1.1)—(1.2) *such that* (1.3) *holds for all* $b, c\geqq 0$ *and* $m, n\geqq 1$, *where* $\alpha>1$ *and* $\gamma>1$. *Then we have*

$$E\big(M^{\gamma}(b, m;\ c, n)\big) \leqq C_{2,\alpha,\gamma} f^{\alpha}(b, m;\ c, n) \tag{2.1}$$

for all $b, c\geqq 0$ *and* $m, n\geqq 1$.

Although its specific value will have no importance for us, the constant $C_{2,\alpha,\gamma}$ may be taken as

$$C_{2,\alpha,\gamma} = (1-2^{(1-\alpha)/\gamma})^{-2\gamma}. \tag{2.2}$$

Before proving Theorem 2, let us introduce the following "striped" maxima that are the maxima of partial sums taken with respect to only p or q:

$$M_1(b, m;\ c, n) = \max_{1\leqq p\leqq m} |S(b, p;\ c, n)|$$

and

$$M_2(b, m;\ c, n) = \max_{1\leqq q\leqq n} |S(b, m;\ c, q)|,$$

where $b, c\geqq 0$ and $m, n\geqq 1$ are integers.

We need the following auxiliary result in the sequel.

Lemma 1. *Let* $\alpha>1$ *and* $\gamma>1$. *Suppose that there exists a non-negative function* $f(b, m; c, n)$ *satisfying* (1.3) *for all* $b, c\geqq 0$ *and* $m, n\geqq 1$. *If* (1.2) *holds, then we have*

$$E\big(M_2^\gamma(b, m;\ c, n)\big) \leqq C_{1,\alpha,\gamma} f^\alpha(b, m;\ c, n) \tag{2.3}$$

for all $b, c\geqq 0$ *and* $m, n\geqq 1$; *here* $C_{1,\alpha,\gamma}$ *may be taken as*

$$C_{1,\alpha,\gamma} = (1-2^{(1-\alpha)/\gamma})^{-\gamma}. \tag{2.4}$$

An analogous result is true for $M_1(b, m; c, n)$ under the assumptions (1.1) and (1.3).

This lemma can be obtained by a simultaneous application to all possible fixed values of $b\geqq 0$ and $m\geqq 1$ of a recent result [6, Theorem 1] in the case when

$$\xi_l = \sum_{k=b+1}^{b+m} \zeta_{kl}, \quad g(c, n) = f(b, m;\ c, n), \quad \text{and} \quad M_{c,n} = M_2(b, m;\ c, n),$$

where $l=c+1, \dots, c+n$ (the notations are the same as in the cited paper).

Proof of Theorem 2. The proof will be done in a similar way as that of [6, Theorem 1]. We are going to find a constant $C\geqq C_{1,\alpha,\gamma}$, depending only on α and γ, for which the inequality

$$E\big(M^\gamma(b, k;\ c, n)\big) \leqq Cf^\alpha(b, k;\ c, n) \tag{2.5}$$

holds for all $b, c\geqq 0$ and $k, n\geqq 1$.

The proof of (2.5) goes by induction on k. If $k=1$, then (2.5) is a consequence of Lemma 1, since we have

$$M(b, 1;\ c, n) = M_2(b, 1;\ c, n)$$

for all $b, c\geqq 0$ and $n\geqq 1$.

Now assume as induction hypothesis that (2.5) holds for all $k<m$ (and for all $b, c\geqq 0$, $n\geqq 1$) and prove it for $k=m$ (and for all $b, c\geqq 0$, $n\geqq 1$).

If for certain $b, c\geqq 0$ and $m, n\geqq 1$ we have $f(b, m; c, n)=0$, then by (1.1)—(1.2) we also have $f(b, k; c, l)=0$, and hence $S(b, k; c, l)=0$ a.s. for $k=1, 2, \dots, m$; $l=1, 2, \dots, n$. Thus $M(b, m; c, n)=0$ a.s., and (2.5) is clearly satisfied.

From now on we assume that $f(b, m; c, n)\neq 0$. Since $f(b, m; c, n)$ is a non-decreasing function in m for any fixed $b, c\geqq 0$ and $n\geqq 1$, there exists an integer h, $1\leqq h\leqq m$, such that

$$f(b, h-1;\ c, n) \leqq \frac{1}{2} f(b, m;\ c, n) < f(b, h;\ c, n), \tag{2.6}$$

where $f(b, h-1; c, n)$ on the left is 0, if $h=1$. Then (1.1) implies

$$f(b+h, m-h;\ c, n) \leqq f(b, m;\ c, n) - f(b, h;\ c, n) < \frac{1}{2} f(b, m;\ c, n). \tag{2.7}$$

Now, for $1 \leqq p < h$ and $1 \leqq q \leqq n$, we have

$$|S(b, p;\ c, q)| \leqq M(b, h-1;\ c, n),$$

and, for $h \leqq p \leqq m$ and $1 \leqq q \leqq n$,

$$|S(b, p;\ c, q)| \leqq M_2(b, h;\ c, n) + M(b+h, m-h;\ c, n).$$

In the last two inequalities we tacitly assume that for either $h=1$ or $h=m$

$$M(b, 0;\ c, n) = M(b+m, 0;\ c, n) = 0.$$

Therefore,

$$M(b, m;\ c, n) \leqq M_2(b, h;\ c, n) + \{M^\gamma(b, h-1;\ c, n) + M^\gamma(b+h, m-h;\ c, n)\}^{1/\gamma}$$

and, by Minkowski's inequality,

$$\text{(2.8)}\quad \{E(M^\gamma(b, m;\ c, n))\}^{1/\gamma} \leqq \{E(M_2^\gamma(b, h;\ c, n))\}^{1/\gamma} + \\ + \{E(M^\gamma(b, h-1;\ c, n)) + E(M^\gamma(b+h, m-h;\ c, n))\}^{1/\gamma}.$$

Applying the induction hypothesis to $M(b, h-1;\ c, n)$, we get that

$$\text{(2.9)}\qquad E(M^\gamma(b, h-1;\ c, n)) \leqq Cf^\alpha(b, h-1;\ c, n) \leqq \frac{C}{2^\alpha} f^\alpha(b, m;\ c, n),$$

the right-most inequality following from (2.6). Applying again the induction hypothesis this time to $M(b+h, m-h;\ c, n)$ and using (2.7), we find that

$$\text{(2.10)}\qquad E(M^\gamma(b+h, m-h;\ c, n)) < \frac{C}{2^\alpha} f^\alpha(b, m;\ c, n).$$

Finally, by (2.3),

$$\text{(2.11)}\qquad E(M_2^\gamma(b, h;\ c, n)) \leqq C_{1,\alpha,\gamma} f^\alpha(b, h;\ c, n) \leqq C_{1,\alpha,\gamma} f^\alpha(b, m;\ c, n).$$

Combining inequalities (2.9)—(2.11) with (2.8), we obtain that

$$\{E(M^\gamma(b, m;\ c, n))\}^{1/\gamma} \leqq (C_{1,\alpha,\gamma}^{1/\gamma} + 2^{(1-\alpha)/\gamma} C^{1/\gamma}) f^{\alpha/\gamma}(b, m;\ c, n).$$

If C is large enough, then it follows that

$$\{E(M^\gamma(b, m;\ c, n))\}^{1/\gamma} \leqq C^{1/\gamma} f^{\alpha/\gamma}(b, m;\ c, n),$$

which proves (2.5) for $k=m$. This completes the induction step and the proof of Theorem 2.

The smallest C satisfying

$$C_{1,\alpha,\gamma}^{1/\gamma} + 2^{(1-\alpha)/\gamma} C^{1/\gamma} \leqq C^{1/\gamma}$$

is given by

$$C = C_{2,\alpha,\gamma} = C_{1,\alpha,\gamma}(1 - 2^{(1-\alpha)/\gamma})^{-\gamma}.$$

By (2.4) this provides (2.2).

Since the stress of this paper is mostly on the method of proving Theorem 2 (Theorem 3 etc. later on), we shall not exhibit the full strength of Theorem 2 and mention only one consequence.

Corollary 1. *Let $\gamma>2$. Suppose that we have*

$$E|S(b,m;\ c,n)|^\gamma \leqq C\left(\sum_{k=b+1}^{b+m}\ \sum_{l=c+1}^{c+n}\sigma_{kl}^2\right)^{\gamma/2}$$

for all $b, c\geqq 0$ and $m, n\geqq 1$, where the σ_{kl}^2 are the finite variances of the rv's ζ_{kl}. Then we have

$$E\big(M^\gamma(b,m;\ c,n)\big) \leqq C(1-2^{(2-\gamma)/2\gamma})^{-2\gamma}\left(\sum_{k=b+1}^{b+m}\ \sum_{l=c+1}^{c+n}\sigma_{kl}^2\right)^{\gamma/2}$$

for all $b, c\geqq 0$ and $m, n\geqq 1$.

The corresponding result for sequences $\{\xi_j\}$ of rv's was established by Erdős [4] and S. B. Stečkin. To be more precise, this result was proved by Erdős for the special case when $\gamma=4$ and $\{\xi_j\}$ is a lacunary sequence of trigonometric functions, while the general case was an oral communication of Stečkin (cf. Gapoškin [5, pp. 29—31]).

§ 3. A generalization of the Rademacher-Menšov inequality in the case $\alpha=1$

Let us proceed to the study of the case $\alpha=1$. Then a factor $(\log 2m)^\gamma\ (\log 2n)^\gamma$ will occur instead of the constant $C_{2,\alpha,\gamma}$ on the right-hand side of (2.1). Here and in the sequel all logarithms are of base 2.

Theorem 3. *Suppose that there exists a non-negative function $f(b,m;\ c,n)$ satisfying* (1.1)—(1.2) *such that* (1.3) *holds for all $b, c\geqq 0$ and $m, n\geqq 1$, where $\alpha=1$ and $\gamma>1$. Then we have*

$$E\big(M^\gamma(b,m;\ c,n)\big) \leqq (\log 2m)^\gamma(\log 2n)^\gamma f(b,m;\ c,n) \tag{3.1}$$

for all $b, c\geqq 0$ and $m, n\geqq 1$.

This is a special case of the following more general result. Before its formulation, let us introduce two recursive definitions. Let $\varkappa(m)$ and $\lambda(n)$ be positive, non-decreasing functions of the natural numbers m and n, respectively. Set, for $m=1$ and $n=1$,

$$K(1)=\varkappa(1) \quad \text{and} \quad \Lambda(1)=\lambda(1),$$

and set, for $m\geqq 2$ and $n\geqq 2$,

$$K(m)=\varkappa(h)+K(h-1), \quad h=\left[\frac{1}{2}(m+2)\right]; \tag{3.2}$$

$$\Lambda(n)=\lambda(i)+\Lambda(i-1), \qquad i=\left[\frac{1}{2}(n+2)\right];$$

here [.] denotes integral part. It is obvious that both $K(m)$ and $\Lambda(n)$ are positive and non-decreasing functions of $m, n=1, 2, \dots$. Further, from (3.2) it follows that if $2^p \leqq m < 2^{p+1}$ with $p \geqq 0$, then

(3.3) $$K(m) \leqq K(2^{p+1}-1) = \sum_{k=0}^{p} \varkappa(2^k);$$

similarly, if $2^q \leqq n < 2^{q+1}$ with $q \geqq 0$, then

$$\Lambda(n) \leqq \sum_{l=0}^{q} \lambda(2^l).$$

Theorem 4. *Suppose that there exist positive, non-decreasing functions $\varkappa(m)$ and $\lambda(n)$, and a non-negative function $f(b, m; c, n)$ satisfying* (1.1)—(1.2) *such that*

(3.4) $$E|S(b, m; c, n)|^\gamma \leqq \varkappa^\gamma(m)\lambda^\gamma(n) f(b, m; c, n)$$

for all $b, c \geqq 0$ and $m, n \geqq 1$, where $\gamma \geqq 1$. Let $K(m)$ and $\Lambda(n)$ be defined by (3.2). *Then we have*

(3.5) $$E\big(M^\gamma(b, m; c, n)\big) \leqq K^\gamma(m)\Lambda^\gamma(n) f(b, m; c, n)$$

for all $b, c \geqq 0$ and $m, n \geqq 1$.

We note that if $\varkappa(m) \equiv 1$ and $\lambda(n) \equiv 1$, then $K(m) \leqq \log 2m$ and $\Lambda(n) \leqq \log 2n$. These follow from the inequalities $1+\log 2(h-1) \leqq \log 2m$ and $1+\log 2(i-1) \leqq \leqq \log 2n$, which are true owing to $m \geqq 2h-2$ and $n \geqq 2i-2$. Consequently, Theorem 3 is a particular case of Theorem 4.

If the rv's ζ_{kl} are mutually orthogonal, i.e.,

(3.6) $$E(\zeta_{ij}\zeta_{kl}) = 0 \quad \text{unless} \quad i = k \quad \text{and} \quad j = l,$$

and if

(3.7) $$E\zeta_{kl}^2 = \sigma_{kl}^2,$$

then obviously

$$E\big(S^2(b, m; c, n)\big) = \sum_{k=b+1}^{b+m} \sum_{l=c+1}^{c+n} \sigma_{kl}^2$$

for all $b, c \geqq 0$ and $m, n \geqq 1$. Hence Theorem 3 implies

Corollary 2. (The two-parameter version of the Rademacher—Menšov inequality) *Under the conditions* (3.6) *and* (3.7) *we have*

$$E\big(M^2(b, m; c, n)\big) \leqq (\log 2m)^2 (\log 2n)^2 \sum_{k=b+1}^{b+m} \sum_{l=c+1}^{c+n} \sigma_{kl}^2$$

for all $b, c \geqq 0$ and $m, n \geqq 1$.

This result was firstly achieved by AGNEW [1]. As for the one-dimensional Rademacher—Menšov inequality see, e.g., RÉVÉSZ [9, p. 83].

If $\varkappa(m)=m^{\beta_1}$ and $\lambda(n)=n^{\beta_2}$ with some positive β_1 and β_2, then by (3.3) we have $K(m)\leqq(2m)^{\beta_1}/(2^{\beta_1}-1)$ and $\Lambda(n)\leqq(2n)^{\beta_2}/(2^{\beta_2}-1)$. Thus in this case we can guarantee again a bound on $E\big(M^\gamma(b,m;\,c,n)\big)$ of the same order of magnitude as the bound assumed on $E|S(b,m;\,c,n)|^\gamma$.

Corollary 3. *Suppose that there exists a non-negative function $f(b,m;\,c,n)$ satisfying* (1.1)—(1.2) *such that*

$$E|S(b,m;\;c,n)|^\gamma \leqq m^{\gamma\beta_1}n^{\gamma\beta_2}f(b,m;\;c,n)$$

for all $b,c\geqq 0$ and $m,n\geqq 1$, where $\beta_1,\beta_2>0$ and $\gamma\geqq 1$. Then we have

$$E\big(M^\gamma(b,m;\;c,n)\big)\leqq\frac{2^{\gamma(\beta_1+\beta_2)}m^{\gamma\beta_1}n^{\gamma\beta_2}}{(2^{\beta_1}-1)^\gamma(2^{\beta_2}-1)^\gamma}f(b,m;\;c,n)$$

for all $b,c\geqq 0$ and $m,n\geqq 1$.

Before proving Theorem 4 we recall the following one-parameter maximal inequality concerning $M_2(b,m;\,c,n)$ defined in § 2.

Lemma 2. *Let $\gamma\geqq 1$. Suppose that there exist positive, non-decreasing functions $\varkappa(m)$ and $\lambda(n)$, and a non-negative function $f(b,m;\,c,n)$ satisfying* (3.4) *for all $b,c\geqq 0$ and $m,n\geqq 1$. If* (1.2) *holds, then we have*

$$E\big(M_2^\gamma(b,m;\;c,n)\big)\leqq\varkappa^\gamma(m)\Lambda^\gamma(n)f(b,m;\;c,n) \tag{3.8}$$

for all $b,c\geqq 0$ and $m,n\geqq 1$, where $\Lambda(n)$ is defined by (3.2).

An analogous result is valid for $M_1(b,m;\,c,n)$ under the assumptions (1.1) and (3.4).

Lemma 2 immediately follows from [6, Theorem 4] (cf. the reasoning after Lemma 1).

Proof of Theorem 4. The proof goes by induction on m. If $m=1$, then (3.5) is a consequence of Lemma 2 owing to

$$K(1)=\varkappa(1)\quad\text{and}\quad M(b,1;\;c,n)=M_2(b,1;\;c,n).$$

Let $m>1$ be given and let h be the integral part of $(m+2)/2$. Then $m=2h-1$ or $m=2h-2$. Let $b,c\geqq 0$ and $n\geqq 1$ be arbitrary integers.

In the same way as in the proof of Theorem 2 we arrive at (2.8). Now suppose that the conclusion (3.5) to be proved is true for all $h<m$. Then we obtain

$$E\big(M^\gamma(b,h-1;\;c,n)\big)\leqq K^\gamma(h-1)\Lambda^\gamma(n)f(b,h-1;\;c,n)$$

and

$$E\big(M^\gamma(b+h,m-h;\;c,n)\big)\leqq K^\gamma(m-h)\Lambda^\gamma(n)f(b+h,m-h;\;c,n)\leqq$$
$$\leqq K^\gamma(h-1)\Lambda^\gamma(n)f(b+h,m-h;\;c,n),$$

since $m \leqq 2h-1$ and the function $K(m)$ is non-decreasing. (In case $m=2$ we have $h=2$, and the second inequality becomes trivial by agreeing that $M(b+2, 0;\ c, n)=0$ and $K(0)=0$.) Putting these two inequalities together, by (1.1) we find that

$$E(M^\gamma(b, h-1;\ c, n))+E(M^\gamma(b+h, m-h;\ c, n)) \leqq \leqq K^\gamma(h-1)A^\gamma(n)f(b, m;\ c, n). \tag{3.9}$$

By virtue of (3.8)

$$E(M_2^\gamma(b, h;\ c, n)) \leqq \varkappa^\gamma(h)A^\gamma(n)f(b, h;\ c, n) \leqq \varkappa^\gamma(h)A^\gamma(n)f(b, m;\ c, n). \tag{3.10}$$

Collecting inequalities (2.8) and (3.9)—(3.10), we get that

$$\{E(M^\gamma(b, m;\ c, n))\}^{1/\gamma} \leqq (\varkappa(h)+K(h-1))A(n)f^{1/\gamma}(b, m;\ c, n).$$

Taking into account the definition (3.2) of $K(m)$, the last inequality gives the wanted (3.5). Thus the proof of Theorem 4 is complete.

§ 4. The maxima of square sums and spherical sums

In this previous sections we established moment inequalities for the maximum of the rectangular sums $S(b, p;\ c, q)$ as p and q run, independently of each other, over the values $1, 2, \ldots, m$ and $1, 2, \ldots, n$, respectively. The situation becomes simpler if $p=q$ or, more generally, if p and q are connected with each other in a certain way.

Let $Q_1 \subset Q_2 \subset \ldots$ be an arbitrary sequence of finite regions in the positive quadrant R^2_+ of the real plane R^2 such that $\bigcup_{r=1}^{\infty} Q_r$ contains infinitely many points with integer coordinates (but not necessarily coincides with R^2_+). Set

$$T(a, r) = \sum_{(k, l) \in Q_{a+r} \setminus Q_a} \zeta_{kl}$$

and

$$N(a, r) = \max_{1 \leqq s \leqq r} |T(a, s)|,$$

where $a \geqq 0$ and $r \geqq 1$ are integers, $Q_0=\emptyset$.

The assumed bounds on $E|T(a, r)|^\gamma$ will be of the form

$$E|T(a, r)|^\gamma \leqq g^\alpha(a, r), \tag{4.1}$$

where α and γ are given numbers, $\alpha \geqq 1$, $\gamma > 0$, and $g(a, r)$ is a non-negative function with the property

$$g(a, s)+g(a+s, r-s) \leqq g(a, r) \tag{4.2}$$

for all $a \geqq 0$ and $1 \leqq s < r$. Our goal is to deduce an upper bound on $E(N^\gamma(a, r))$.

The one-parameter version of Theorems 1 and 2 reads as follows.

Theorem 5. *Suppose that there exists a non-negative function $g(a, r)$ satisfying* (4.2) *such that* (4.1) *holds for all $a \geqq 0$ and $r \geqq 1$, where either $\alpha \geqq 1$ and $0 < \gamma \leqq 1$ or $\alpha > 1$ and $\gamma > 1$. Then we have*

$$E(N^\gamma(a, r)) \leqq C_{1,\alpha,\gamma} g^\alpha(a, r) \tag{4.3}$$

for all $a \geqq 0$ and $r \geqq 1$.

We remark that the constant $C_{1,\alpha,\gamma}$ in (4.3) may be taken as

$$C_{1,\alpha,\gamma} = \begin{cases} 1 & \text{if } \alpha \geqq 1 \text{ and } 0 < \gamma \leqq 1, \\ (1 - 2^{(1-\alpha)/\gamma})^{-\gamma} & \text{if } \alpha > 1 \text{ and } \gamma > 1. \end{cases}$$

By setting $\xi_j = \sum_{(k,l) \in Q_j \setminus Q_{j-1}} \zeta_{kl}$ $(j = 1, 2, \ldots)$, Theorem 5 follows immediately from [6, Theorems 1 and 2]. On the other hand, if we apply [6, Theorem 4] to this sequence $\{\xi_j\}$, we get the following one-parameter version of the present Theorem 4.

Theorem 6. *Suppose that there exist a positive, non-decreasing function $\varkappa(r)$, and a non-negative function $g(a, r)$ satisfying* (4.2) *such that*

$$E|T(a, r)|^\gamma \leqq \varkappa^\gamma(r) g(a, r)$$

holds for all $a \geqq 0$ and $r \geqq 1$, where $\gamma \geqq 1$. Let $K(r)$ be defined by (3.2). *Then we have*

$$E(N^\gamma(a, r)) \leqq K^\gamma(r) g(a, r)$$

for all $a \geqq 0$ and $r \geqq 1$.

Let us consider two interesting special cases for the choice of $\{Q_r\}$, which provide (i) the square sums, among others, and (ii) the spherical sums.

Case (i). Let $m = m(r)$ and $n = n(r)$, where

$$1 \leqq m(1) \leqq m(2) \leqq \ldots \quad \text{and} \quad 1 \leqq n(1) \leqq n(2) \leqq \ldots$$

are two sequences of integers such that $\max\{m(r), n(r)\} \to \infty$ as $r \to \infty$, and let $Q_r = \{(k, l) \colon k \leqq m(r) \text{ and } l \leqq n(r)\}$. It is convenient to put $m(0) = n(0) = 0$ and $Q_0 = \emptyset$. Now

$$T(a, r) = \left\{ \sum_{k=1}^{m(a+r)} \sum_{l=1}^{n(a+r)} - \sum_{k=1}^{m(a)} \sum_{l=1}^{n(a)} \right\} \zeta_{kl}.$$

In particular, if $m(r) = n(r) = r$, then the $T(0, r) = S(0, r; 0, r)$ give back the square sums. The case $m(r) = n(r) = 2^r$ is also of interest.

We mention that if $f(b, m; c, n)$ is a non-negative function satisfying (1.1)—(1.2), then $g(a, r)$ defined by

$$g(a, r) = f(0, m(a+r); 0, n(a+r)) - f(0, m(a); 0, n(a))$$

is also non-negative and satisfies (4.2).

It is worth stating Theorem 6 explicitly in the special case of mutually orthogonal ζ_{kl}. Then $\gamma=2$, $\varkappa(r)\equiv 1$, and with $\sigma_{kl}^2=E(\zeta_{kl}^2)$ we have

$$E\big(T^2(a,r)\big)=\left\{\sum_{k=1}^{m(a+r)}\sum_{l=1}^{n(a+r)}-\sum_{k=1}^{m(a)}\sum_{l=1}^{n(a)}\right\}\sigma_{kl}^2$$

for all $a\geqq 0$ and $r\geqq 1$.

Corollary 4. *Let $\{m(r)\}$ and $\{n(r)\}$ be two non-decreasing sequences of positive integers. Under the conditions* (3.6) *and* (3.7) *we have*

$$E\big(N^2(a,r)\big)\leqq(\log 2r)^2\left\{\sum_{k=1}^{m(a+r)}\sum_{l=1}^{n(a+r)}-\sum_{k=1}^{m(a)}\sum_{l=1}^{n(a)}\right\}\sigma_{kl}^2$$

for all $a\geqq 0$ and $r\geqq 1$.

Case (ii). The spherical sums are defined with the aid of $Q_r=\{(k,l): k^2+l^2\leqq r\}$ $(r=1,2,\ldots;\ Q_1=\emptyset)$, i.e., now

$$T(a,r)=\sum_{a<k^2+l^2\leqq a+r}\zeta_{kl}.$$

The case of orthogonal ζ_{kl} is again of interest in itself.

Corollary 5. *Under the conditions* (3.6) *and* (3.7) *we have*

$$E\big(N^2(a,r)\big)\leqq(\log 2r)^2\sum_{a<k^2+l^2\leqq a+r}\sigma_{kl}^2$$

for all $a\geqq 1$ and $r\geqq 1$.

Corollaries 4 and 5 were proved earlier in [7, Corollary 3 and Theorem 4].

§ 5. Generalizations to multiparameter case

Let Z^d denote the set of all d-tuples of non-negative integers, and let Z_+^d denote the set of all d-tuples of positive integers, where $d\geqq 1$ is a fixed integer. The points in Z^d are denoted by $\mathbf{k},\mathbf{m}$ etc., or sometimes, when necessary, more explicitly by $(k_1,k_2,\ldots,k_d)$, $(m_1,m_2,\ldots,m_d)$ etc. Two d-tuples $\mathbf{k}$ and $\mathbf{m}$ are said to be distinct if for at least one j we have $k_j\neq m_j$ $(1\leqq j\leqq d)$. Z^d is partially ordered by agreeing that $\mathbf{k}\leqq\mathbf{m}$ iff $k_j\leqq m_j$ for each j, $1\leqq j\leqq d$. We write $\mathbf{0}$ and $\mathbf{1}$ respectively for the points $(0,0,\ldots,0)$ and $(1,1,\ldots,1)$ in Z^d.

Let $\{\zeta_{\mathbf{k}}\}=\{\zeta_{\mathbf{k}}:\mathbf{k}\in Z_+^d\}$ be a random field, i.e. a collection of rv's indexed by the set Z_+^d. Put

$$S(\mathbf{b},\mathbf{m})=\sum_{\mathbf{b}+\mathbf{1}\leqq\mathbf{k}\leqq\mathbf{b}+\mathbf{m}}\zeta_{\mathbf{k}}=\sum_{k_1=b_1+1}^{b_1+m_1}\cdots\sum_{k_d=b_d+1}^{b_d+m_d}\zeta_{k_1,\ldots,k_d}$$

and

$$M(\mathbf{b},\mathbf{m})=\max_{\mathbf{1}\leqq\mathbf{k}\leqq\mathbf{m}}|S(\mathbf{b},\mathbf{k})|=\max_{1\leqq k_1\leqq m_1}\cdots\max_{1\leqq k_d\leqq m_d}|S(\mathbf{b},\mathbf{k})|,$$

where $\mathbf{b}\in Z^d$, $\mathbf{m}\in Z_+^d$, and $\mathbf{b}+\mathbf{1}$, $\mathbf{b}+\mathbf{m}$ are the usual coordinatewise sums.

To formulate the generalizations of the above Theorems 1—6, let $f(\mathbf{b},\mathbf{m})$ denote a non-negative function depending on the joint df of $\{\zeta_{\mathbf{k}}:\mathbf{b}+\mathbf{1}\leqq\mathbf{k}\leqq\mathbf{b}+\mathbf{m}\}$ with the following property. Set, for $1\leqq j\leqq d$,

$$\mathbf{b}_j=(b_1,\ldots,b_{j-1},b_{j+1},\ldots,b_d)\in Z^{d-1}$$

and

$$\mathbf{m}_j=(m_1,\ldots,m_{j-1},m_{j+1},\ldots,m_d)\in Z_+^{d-1},$$

where $\mathbf{b}=(b_1,b_2,\ldots,b_d)\in Z^d$ and $\mathbf{m}=(m_1,m_2,\ldots,m_d)\in Z_+^d$. We require that the inequality

$$f(\mathbf{b}_j,\mathbf{m}_j;b_j,h_j)+f(\mathbf{b}_j,\mathbf{m}_j;b_j+h_j,m_j-h_j)\leqq f(\mathbf{b}_j,\mathbf{m}_j;b_j,m_j)=f(\mathbf{b},\mathbf{m}) \tag{5.1}$$

holds true for all $\mathbf{b}\in Z^d$, $\mathbf{m}\in Z_+^d$, $1\leqq h_j<m_j$, and $1\leqq j\leqq d$.

Inequality (5.1) expresses that $f(\mathbf{b},\mathbf{m})$ as a function of the interval (b_j+1,b_j+m_j) is superadditive for any fixed values of $b_1,m_1,\ldots,b_{j-1},m_{j-1},b_{j+1},m_{j+1},\ldots,b_d,m_d$. Examples are $f(\mathbf{b},\mathbf{m})=m_1^{\beta_1}m_2^{\beta_2}\ldots m_d^{\beta_d}$ with $\beta_j\geqq 1$ for each j, $1\leqq j\leqq d$; or $f(\mathbf{b},\mathbf{m})=\sum_{\mathbf{b}+\mathbf{1}\leqq\mathbf{k}\leqq\mathbf{b}+\mathbf{m}}\sigma_{\mathbf{k}}^2$, in the latter case assuming the existence of the finite variances $\sigma_{\mathbf{k}}^2$ of the rv's $\zeta_{\mathbf{k}}$.

Theorem 7. *Suppose that there exists a non-negative function* $f(\mathbf{b},\mathbf{m})$ *satisfying* (5.1) *such that*

$$E|S(\mathbf{b},\mathbf{m})|^\gamma\leqq f^\alpha(\mathbf{b},\mathbf{m})$$

holds for all $\mathbf{b}\in Z^d$ *and* $\mathbf{m}\in Z_+^d$, *where either* $\alpha\geqq 1$ *and* $0<\gamma\leqq 1$ *or* $\alpha>1$ *and* $\gamma>1$. *Then we have*

$$E\big(M^\gamma(\mathbf{b},\mathbf{m})\big)\leqq C_{d,\alpha,\gamma}f^\alpha(\mathbf{b},\mathbf{m})$$

for all $\mathbf{b}\in Z^d$ *and* $\mathbf{m}\in Z_+^d$.

Here the constant $C_{d,\alpha,\gamma}$ may be chosen as follows: $C_{d,\alpha,\gamma}=1$ if $\alpha\geqq 1$ and $0<\gamma\leqq 1$, and

$$C_{d,\alpha,\gamma}=C_{1,\alpha,\gamma}^d=(1-2^{(1-\alpha)/\gamma})^{-d\gamma}$$

if $\alpha>1$ and $\gamma>1$.

In connection with Theorem 7 we note that Bickel and Wichura [2] proved a fine but not comparable result, providing a multiparameter extension of Billingsley's main fluctuation inequality [3, Theorem 12.5]. Roughly speaking, they obtain an asymptotically optimal inequality on $P\{M(\mathbf{b},\mathbf{m})\geqq\lambda\}$ in terms of assumed bounds on $P\{|S(\mathbf{b},\mathbf{m})|\geqq\lambda\}$, where $\mathbf{b}\in Z^d$, $\mathbf{m}\in Z_+^d$, and λ is a positive number.

For each j, $1\leqq j\leqq d$, let $\lambda_j(m_j)$ be a positive and non-decreasing function of the natural number m_j. Define $\Lambda_j(m_j)$ by the recurrence relation (3.2), that is, for $m_j=1$ set $\Lambda_j(1)=\lambda_j(1)$, and for $m_j\geqq 2$ set

$$\Lambda_j(m_j)=\lambda_j(h_j)+\Lambda_j(h_j-1),\quad\text{where}\quad h_j=\left[\frac{1}{2}(m_j+2)\right]. \tag{5.2}$$

Theorem 8. *Suppose that there exist positive, non-decreasing functions* $\lambda_j(m_j)$ *for* $j=1, 2, \ldots, d$, *and a non-negative function* $f(\mathbf{b}, \mathbf{m})$ *satisfying* (5.1) *such that*

$$E|S(\mathbf{b}, \mathbf{m})|^\gamma \leqq \prod_{j=1}^{d} \lambda_j^\gamma(m_j) f(\mathbf{b}, \mathbf{m}) \tag{5.3}$$

holds for all $\mathbf{b} \in Z^d$ *and* $\mathbf{m} \in Z_+^d$, *where* $\gamma \geqq 1$. *Let* $\Lambda_j(m_j)$ *be defined by* (5.2) *for* $j=1, 2, \ldots, d$. *Then we have*

$$E(M^\gamma(\mathbf{b}, \mathbf{m})) \leqq \prod_{j=1}^{d} \Lambda_j^\gamma(m_j) f(\mathbf{b}, \mathbf{m}) \tag{5.4}$$

for all $\mathbf{b} \in Z^d$ *and* $\mathbf{m} \in Z_+^d$.

The proof of Theorems 7 and 8 may be carried out by induction on d in the same manner as we did it from $d=1$ to $d=2$ in the case of Theorems 2 and 4. The simplest case $d=1$ was proved in [6].

As is well-known, the random field $\{\zeta_\mathbf{k}\}$ is said to be orthogonal if

$$E(\zeta_\mathbf{k} \zeta_\mathbf{l}) = 0 \quad \text{if} \quad \mathbf{k} \neq \mathbf{l}. \tag{5.5}$$

Setting

$$E(\zeta_\mathbf{k}^2) = \sigma_\mathbf{k}^2, \tag{5.6}$$

for orthogonal $\zeta_\mathbf{k}$ we obviously have

$$E\big(S^2(\mathbf{b}, \mathbf{m})\big) = \sum_{\mathbf{b}+\mathbf{1} \leqq \mathbf{k} \leqq \mathbf{b}+\mathbf{m}} \sigma_\mathbf{k}^2.$$

This is a particular case of the condition (5.3) with $\gamma=2$, $\lambda_j(m_j) \equiv 1$ for each j, $1 \leqq j \leqq d$, and $f(\mathbf{b}, \mathbf{m}) = \sum_{\mathbf{b}+\mathbf{1} \leqq \mathbf{k} \leqq \mathbf{b}+\mathbf{m}} \sigma_\mathbf{k}^2$ (the latter is even additive). Now $\Lambda_j(m_j) = \log 2m_j$ for each j and Theorem 8 provides the following

Corollary 6. (The d-parameter version of the Rademacher—Menšov inequality) *Under the conditions* (5.5) *and* (5.6) *we have*

$$E\big(M^2(\mathbf{b}, \mathbf{m})\big) \leqq \prod_{j=1}^{d} (\log 2m_j)^2 \sum_{\mathbf{b}+\mathbf{1} \leqq \mathbf{k} \leqq \mathbf{b}+\mathbf{m}} \sigma_\mathbf{k}^2$$

for all $\mathbf{b} \in Z^d$ *and* $\mathbf{m} \in Z_+^d$.

Similar generalizations from 2 to d of Theorems 5 and 6 are valid, too. Instead of stating them explicitly, we formulate a useful consequence for orthogonal $\zeta_\mathbf{k}$. Let $Q_1\, Q_2 \ldots$ be an arbitrary sequence of finite regions in Z_+^d such that $\bigcup_{r=1}^{\infty} Q_r$ is not bounded (but may not coincide with Z_+^d). Set

$$T(a, r) = \sum_{\mathbf{k} \in Q_{a+r} \setminus Q_a} \zeta_\mathbf{k}$$

and

$$N(a, r) = \max_{1 \leqq s \leqq r} |T(a, s)|,$$

where $a \geqq 0$ and $r \geqq 1$ are integers, $Q_0 = \emptyset$.

Corollary 7. *Under the conditions* (5.5) *and* (5.6) *we have*

$$E(N^2(a,r)) \leqq (\log 2r)^2 \sum_{k\in Q_{a+r}\setminus Q_a} \sigma_k^2 \tag{5.7}$$

for all $a\geqq 0$ *and* $r\geqq 1$.

We note that in the more general setting when the coordinates m_j of $\mathbf{m}\in Z_+^d$, $1\leqq j\leqq d$, depend on an e-dimensional parameter $\mathbf{r}=(r_1, r_2, ..., r_e)\in Z_+^e$, where $1\leqq e<d$, the following result can be achieved for orthogonal $\zeta_\mathbf{k}$. Let $\{Q_\mathbf{r}: \mathbf{r}\in Z^e\}$ be an arbitrary collection of finite regions in Z_+^d such that $Q_\mathbf{0}=\emptyset$, $Q_\mathbf{s}\subset Q_\mathbf{r}$ if $\mathbf{s}\leqq\mathbf{r}$, and $\bigcup_{\mathbf{r}\geqq 0} Q_\mathbf{r}$ is not bounded in Z_+^d, where $\mathbf{r}, \mathbf{s}\in Z^e$. Set

$$T(\mathbf{a},\mathbf{r}) = \sum_{\mathbf{k}\in Q_{\mathbf{a}+\mathbf{r}}\setminus Q_\mathbf{a}} \zeta_\mathbf{k}$$

and

$$N(\mathbf{a},\mathbf{r}) = \max_{1\leqq \mathbf{s}\leqq \mathbf{r}} |T(\mathbf{a},\mathbf{s})| = \max_{1\leqq s_1\leqq r_1} \dots \max_{1\leqq s_e\leqq r_e} |T(\mathbf{a},\mathbf{s})|,$$

where $\mathbf{a}\in Z^e$ and $\mathbf{r}\in Z_+^e$. Then, under (5.5) and (5.6), we have

$$E(N^2(\mathbf{a},\mathbf{r})) \leqq \prod_{i=1}^{e} (\log 2r_i)^2 \sum_{\mathbf{k}\in Q_{\mathbf{a}+\mathbf{r}}\setminus Q_\mathbf{a}} \sigma_\mathbf{k}^2 \tag{5.8}$$

for all $\mathbf{a}\in Z^e$ and $\mathbf{r}\in Z_+^e$.

In case $e=1$, (5.8) reduces to (5.7).

Finally, we mention that viewing our proofs, it is striking that we use no full power of a probability space. In fact, Minkowski's inequality was applied only, which is available in any measure space $(X, \mathcal{A}, \mu)$. Hence our theorems are true in $(X, \mathcal{A}, \mu)$, too, taking integrals over X with respect to μ in place of the expectations on the left-hand sides of the corresponding inequalities.

References

[1] R. P. Agnew, On double orthogonal series, *Proc. London Math. Soc.*, **33** (1932), 420—434.
[2] P. J. Bickel and M. J. Wichura, Convergence criteria for multiparameter stochastic processes and some applications, *Ann. Math. Statist.*, **42** (1971), 1656—1670.
[3] P. Billingsley, *Convergence of probability measures*, Wiley (New York, 1968).
[4] P. Erdős, On the convergence of trigonometric series, *J. Math. and Phys. (Massachusetts Inst. Tech.)*, **22** (1943), 37—39.
[5] В. Ф. Гапошкин, Лакунарные ряды и независимые функции, *Успехи матем. наук*, **21** (6) (1966), 3—82.
[6] F. Móricz, Moment inequalities and the strong laws of large numbers, *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **35** (1976), 299—314.
[7] F. Móricz, On the convergence of double orthogonal series, *Analysis Math.*, **2** (1976), 287—304.
[8] F. Móricz, Multiparameter strong laws of large numbers, *Acta Sci. Math.* (to appear).
[9] P. Révész, *The laws of large numbers*, Academic Press (New York, 1968).

BOLYAI INSTITUTE, UNIV. SZEGED
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Acta Sci. Math., 39 (1977), 367—389

On an extension of semigroups

M. B. SZENDREI

1. Since the appearance of N. R. REILLY's paper [13] in 1966 a number of structure theorems has been proved for regular semigroups. In the paper [13] it is proved that a semigroup is a $\mathcal{D}$-simple regular ω-semigroup if and only if it is isomorphic to a Bruck semigroup over a group ([12]). This result was generalized by B. P. KOČIN ([4]) and W. D. MUNN ([9]) by showing that a semigroup is a simple regular ω-semigroup if and only if it is a Bruck semigroup over a finite chain of groups. The structure of a 0-$\mathcal{D}$-simple orthodox semigroup the subsemigroup of idempotents of which is isomorphic to the direct product of a descending ω-chain and a rectangular 0-band whose non-zero idempotents form a subsemigroup, has been described by G. LALLEMENT and M. PETRICH in [6].

In order to generalize these constructions we define the concept of the (0-) extension of a semigroup Σ by a semigroup S. The sets of nonzero elements of S and Σ will be denoted by S_0 and Σ_ω, their zero elements by o and ω, respectively. Let $S_0^{(2)}$ be the subset of $S_0 \times S_0$ consisting of all those pairs (s, t) of elements for which $st \in S_0$. Let C be a cancellative monoid. Its identity element will be denoted by 1. Let $f, g\colon S_0^{(2)} \to C$ be a pair of functions with the following properties:

$$f_{r,s} f_{rs,t} = f_{r,st}, \tag{1}$$

$$g_{r,s} f_{rs,t} = f_{s,t} g_{r,st}, \tag{2}$$

$$g_{rs,t} = g_{s,t} g_{r,st} \tag{3}$$

whenever $rst \in S_0$. Moreover, let a homomorphism $\varkappa$ of C into the endomorphism monoid of Σ be given.

Definition. Define a multiplication on the set $S_0 \times \Sigma_\omega \cup 0$ by

$$(s, \sigma)(t, \tau) = \begin{cases} (st, \sigma(f_{s,t}\varkappa)\tau(g_{s,t}\varkappa)) & \text{if } st \in S_0 \text{ and } \sigma(f_{s,t}\varkappa)\tau(g_{s,t}\varkappa) \in \Sigma_\omega \\ 0 & \text{otherwise,} \end{cases}$$

$$(s, \sigma)0 = 0(s, \sigma) = 0 \cdot 0 = 0.$$

Received February 5, 1976.

The groupoid obtained in this way is a *semigroup*, denoted by $\mathscr{S}^0(S, \Sigma, C, f, g, \varkappa)$ and called a *0-extension of Σ by S over C.*

If none of S and Σ has zero elements then $\mathscr{S}^0(S, \Sigma, C, f, g, \varkappa)\backslash 0$ is a *semigroup*. This will be denoted by $\mathscr{S}(S, \Sigma, C, f, g, \varkappa)$ and called the *extension of Σ by S over C.*

For example, the Bruck semigroup $\mathscr{B}(\Sigma, \pi)$ over the monoid Σ is the extension $\mathscr{S}(B, \Sigma, N^0, f^*, g^*, \varkappa)$ of Σ by the bicyclic semigroup B, where N^0 is the additive monoid of nonnegative integers, $B \cong N^0 \times N^0$ with the multiplication defined by

$$(m, n)(p, q) = \big(m+p-\min(n, p), n+q-\min(n, p)\big),$$

$f^*, g^*\colon B \times B \to N^0$ are defined as follows:

$$f^*_{(m,n),(p,q)} = p-\min(n, p), \quad g^*_{(m,n),(p,q)} = n-\min(n, p),$$

and $\varkappa$ is the homomorphism of N^0 into the monoid of endomorphisms of Σ mapping k into π^k. Note that the functions f^* and g^* have the properties (1)—(3).

It is clear that it suffices to investigate the properties of the semigroup $\mathscr{S}^0(S, \Sigma, C, f, g, \varkappa)$ because the properties of $\mathscr{S}(S, \Sigma, C, f, g, \varkappa)$ can be deduced from those of $\mathscr{S}^0(S, \Sigma, C, f, g, \varkappa)$.

Define an equivalence relation $\mathscr{C}$ on $\mathscr{S}^0(S, \Sigma, C, f, g, \varkappa)$ such that $0\mathscr{C}0$ and $(r, \varrho)\mathscr{C}(s, \sigma)$ if and only if $r=s$. The relation $\mathscr{C}$ is a 0-congruence in the sense that if $(r, \varrho)\mathscr{C}(s, \sigma)$ and $(r', \varrho')\mathscr{C}(s', \sigma')$ then $(r, \varrho)(r', \varrho') \neq 0$, $(s, \sigma)(s', \sigma') \neq 0$ imply $(r, \varrho)(r', \varrho')\mathscr{C}(s, \sigma)(s', \sigma')$.

The pair of functions $f, g\colon S_0^{(2)} \to C$ is said to be *trivial* if $S_0^{(2)}f = S_0^{(2)}g = 1$. In this case $\mathscr{S}^0(S, \Sigma, C, f, g, \varkappa)$ is the 0-direct product of S and Σ. Note that the semidirect product of Σ by S introduced by K. Krohn and J. Rhodes in [5] can be considered to be an extension of Σ by S over the free monoid F_S^e generated by the set S where $f, g\colon S \times S \to F_S^e$ are defined as follows: $(S \times S)g = 1$ while f depends on its second variable only and is a homomorphism.

The constructions used in [15] and [16] by R. J. Warne to describe the structure of $\mathscr{D}$-simple and simple regular I-semigroups, are extensions of a group and of a finite chain of groups, respectively, by the extended bicyclic semigroup if and only if they have trivial distinguished elements. Construction I applied in [1] and [2] by J. E. Ault and M. Petrich to give the structure of 0-simple ω-regular semigroups, is a 0-extension of a finite chain of groups by the 0-$\mathscr{D}$-simple ω-regular semigroup with trivial $\mathscr{H}$-equivalence if and only if the maximal idempotents belong to the same $\mathscr{D}$-class.

The aim of this paper is to investigate the properties of (0-) extensions. In section 2 we deal with functions $f, g\colon S_0^{(2)} \to C$ satisfying (1)—(3). The main result of this section is Theorem 2.3 characterizing these functions when S has an identity element and C is a monoid embeddable in a group. In Theorem 2.4 a necessary and sufficient condition is given for f and g enabling us to extend their definitions to S^e. Finally, applying Theorem 2.3, we describe the structure of those 0-$\mathcal{D}$-simple semigroups with identity which admit f, g of a special type. In section 3 we prove criteria for $\mathcal{S}^0(S, \Sigma, C, f, g, \varkappa)$ to be regular or inverse. We investigate Green's relations, ideals and homomorphisms of 0-extensions. We introduce a concept of equivalent 0-extensions and give conditions for 0-extensions to be equivalent. These results are essentially independent of the results of section 2. Theorem 2.3 is needed only in Theorems 3.9 and 3.11.

For brevity, if we consider functions $f, g\colon S_0^{(2)} \to C$ or a 0-extension $\mathcal{S}^0(S, \Sigma, C, f, g, \varkappa)$ we always assume conditions (1)—(3) to be satisfied. We shall write lower case Roman letters for the elements of S, in particular e for its identity element, and lower case Greek letters to denote the elements of Σ, in particular ε to denote its identity element.

S_0 together with the multiplication in S restricted to S_0 is a partial semigroup. By a right [left] ideal of S_0 we mean a non-empty subset R [L] of S_0 with the property that $r \in R$ [$l \in L$] implies $rs \in R$ [$sl \in L$] for any element s of S_0 whenever the product is defined. Analogously, a homomorphism of S_0 into a semigroup T is a mapping $\varphi\colon S_0 \to T$ such that for all elements s, t in S_0 we have $(st)\varphi = s\varphi \cdot t\varphi$ provided st is defined.

For convenience, we use the expressions "if $s=o$" and "if $s \neq o$" also in the case when S has no zero element. If this is the case then $s=o$ is false, $s \neq o$ is true for every s in S.

The results and notations of [3] will be used without any comments.

2. In this section we investigate the properties of functions $f, g\colon S_0^{(2)} \to C$ satisfying conditions (1)—(3).

Lemma 2.1. *If x, s are elements of S_0 such that $sx=s$, then $f_{s,x}=1$, and if $xs=s$, then $g_{x,s}=1$.*

Proof. Assume that $sx=s$. Applying conditions (1) and (2) we get

$$f_{s,s} f_{s^2,x} = f_{s,s} \quad \text{and} \quad g_{s,s} f_{s^2,x} = f_{s,x} g_{s,s}.$$

Since C is a cancellative monoid, $f_{s^2,x}=1$ follows from the first equality and $f_{s,x}=1$ from the second one. The second half of the lemma follows by duality.

Lemma 2.2. *Let s, s', t, t' be elements of S_0 such that $s \mathcal{L} s'$, $t \mathcal{R} t'$ and $st \neq o$. Then we have*

(i) $f_{s,t} \mathcal{R} f_{s,t'}$ *and* $g_{s,t} \mathcal{L} g_{s',t}$;

(ii) *if the group of units of C is trivial, then* $f_{s,t}=f_{s',t'}$ *and* $g_{s,t}=g_{s',t'}$.

Proof. Suppose $t\neq t'$. Then there exist u, v in S such that $tu=t'$ and $t'v=t$. Clearly, $st'\neq o$. Condition (1) implies the following equalities:

$$f_{s,t'}=f_{s,tu}=f_{s,t}f_{st,u}, \quad f_{st,uv}=f_{st,u}f_{stu,v}.$$

For $st\,uv=st$ Lemma 2.1 shows that $f_{st,uv}=1$. Hence $f_{st,u}$ has an inverse in C, which implies that $f_{s,t}\mathcal{R}f_{s,t'}$. If the group of units of C is trivial, then the second equality implies that $f_{st,u}=1$ and the first one that $f_{s,t'}=f_{s,t}$. Moreover, if $s\neq s'$, then $xs=s'$, $ys'=s$ for some x and y in S. We have $yxs=s$, so it follows from Lemma 2.1 that $g_{yx,s}=1$. (3) implies

$$g_{yx,s}=g_{x,s}g_{y,xs},$$

which gives $g_{x,s}=1$. Analogously, one can show that $g_{x,st'}=1$. By condition (2) we have

$$g_{x,s}f_{s',t'}=f_{s,t'}g_{x,st'},$$

that is, $f_{s',t'}=f_{s,t'}=f_{s,t}$. The proof for g is similar.

In what follows we assume that C can be embedded in a group. It is well known that if this is the case then C can be embedded in the group of right quotients which will be denoted by C^*. Let us identify C with its image under this embedding. If two functions $\chi_1, \chi_2\colon S_0\to C$ are given, let $\chi_1/\chi_2\colon S_0\to C^*$ be the mapping defined by $s\chi_1/\chi_2=s\chi_1(s\chi_2)^{-1}$. The next theorem characterizes the functions $f, g\colon S_0^{(2)}\to C$ by functions of one variable provided S has an identity element.

Theorem 2.3. (i) *Let S be an arbitrary semigroup and $\chi_1, \chi_2\colon S_0\to C$ two functions such that $R_I=\{s\in S_0 | s\chi_1\in I\}$ is a right ideal in S_0, $L_I=\{s\in S_0 | s\chi_2\in I\}$ is a left ideal in S_0 for every right ideal I of C, moreover, the mapping $\varphi=\chi_1/\chi_2\colon S_0\to C^*$ is a homomorphism. Then the functions $f, g\colon S_0^{(2)}\to C$ defined by*

$$(5)\qquad f_{s,t}=(s\chi_1)^{-1}(st)\chi_1 \quad \text{and} \quad g_{s,t}=(t\chi_2)^{-1}(st)\chi_2$$

satisfy conditions (1)—(3).

(ii) *If S has an identity element e then for all $f, g\colon S_0^{(2)}\to C$ with properties* (1)—(3) *there exists a unique pair of functions $\chi_1, \chi_2\colon S_0\to C$ with $e\chi_1=e\chi_2=1$ such that* (5) *holds. They are*

$$(6)\qquad s\chi_1=f_{e,s} \quad \text{and} \quad s\chi_2=g_{s,e}.$$

Furthermore, these functions satisfy the conditions required in (i).

Proof. Since the facts that $R_{s\chi_1C}$ is a right ideal and $L_{t\chi_2C}$ is a left ideal of S_0 ensure $f_{s,t}\in C$ and $g_{s,t}\in C$ for every $(s, t)\in S_0^{(2)}$, (i) can be checked by simple calculation.

In proving (ii) suppose S has an identity e and $e\chi_1=e\chi_2=1$. Then (5) implies (6). Because of (i), it is sufficient to show that (5) holds and the conditions required in (i) are satisfied by the functions χ_1, χ_2 defined by (6). Clearly, (5) is an immediate consequence of (1) and (3). On the other hand, (2) and $st \neq o$ imply

(7) $$f_{s,t} g_{s,t}^{-1} = g_{s,e}^{-1} f_{e,t}.$$

Applying (5) and (6), this yields

$$(st)\chi_1\big((st)\chi_2\big)^{-1} = s\chi_1(s\chi_2)^{-1} t\chi_1(t\chi_2)^{-1},$$

that is, that χ_1/χ_2 is a homomorphism. Finally, if I is a right ideal of C and $s \in R_I$, then $(st)\chi_1 = s\chi_1 f_{s,t} \in I$ for every t provided that $st \neq o$. Hence R_I is a right ideal. Dually, L_I is a left ideal.

We have seen that the pair of functions f, g can be simply characterized if S has an identity element. Now it is natural to raise the problem of finding conditions under which the definition of f and g can be extended to S^e. Before treating this question we introduce some notations.

Let S be a semigroup. Denote the right and left annihilator ideals of S by Z_r and Z_l, respectively. If S does not contain a zero element, then $Z_r = Z_l = \square$. Further, $h\colon S_0^{(2)} \to C^*$ will denote the mapping defined by $h_{s,t} = f_{s,t} g_{s,t}^{-1}$ provided $f, g\colon S_0^{(2)} \to C$ are defined and $C \subseteqq C^*$.

Theorem 2.4. *Suppose the semigroup S has the properties that $Z_r = Z_l$ (which will be denoted by Z) and for any elements s, s', t, t' in S, the relations $st, st', s't \neq o$ imply $s't' \neq o$. Let $f, g\colon S_0^{(2)} \to C$ be given, where C is a monoid embeddable in a group. The definition of f, g can be extended to S^e if and only if*

(a) *for each element q in $S \setminus Z$*

$$J_q = \bigcap_{\substack{s,t \\ st,\, sq \neq o}} (C h_{s,t}^{-1} \cap C) h_{s,q} \cap C$$

is not empty, and for arbitrary $p, q, s, t \in S \setminus Z$

(b) $$h_{s,t} h_{p,t}^{-1} h_{p,q} h_{s,q}^{-1} = 1$$

provided $st, pt, pq, sq \neq o$.

Remark. The definition of f, g can be extended to S^e if we require (a) and (b) to hold only for the elements p and q of some subsets P and Q of $S \setminus Z$, respectively, where P and Q have the following property: For each s, t, t' not contained in Z we have $sq \neq o$ for some q in Q and $pt, pt' \neq o$ for some p in P.

Proof. If f and g are defined on S^e then, applying the foregoing results, we have

$$h_{s,t} h_{p,t}^{-1} h_{p,q} h_{s,q}^{-1} = (g_{s,e}^{-1} f_{e,t})(f_{e,t}^{-1} g_{p,e})(g_{p,e}^{-1} f_{e,q})(f_{e,q}^{-1} g_{s,e}) = 1$$

for every p, q, s, t in $S\setminus Z$ with $st, pt, pq, sq\neq o$. Furthermore, if $q\in S\setminus Z$, then there exists an element s such that $sq\neq o$. If $st\neq o$, then we have

$$f_{e,q}=g_{s,e}h_{s,q}=g_{s,e}h_{s,t}h_{p,t}^{-1}h_{p,q}=f_{e,t}h_{p,t}^{-1}h_{p,q}=f_{e,t}h_{s,t}^{-1}h_{s,q}.$$

Hence $f_{e,q}\in J_q$, and the proof of necessity is complete.

As for sufficiency we prove the stronger statement formulated in the Remark. Suppose that (a) and (b) hold for some subsets P and Q of $S\setminus Z$. We define a relation $\sim$ on S by writing $s\sim s'$ if and only if $s=s'$ or ts and $ts'\neq o$ for some t in S. Clearly, this relation is reflexive and symmetric. If $s\sim s'$ and $s'\sim s''$, then ts, ts', $t's', t's''\neq o$ for some t and t' in S. But then $t's\neq o$, that is, $s\sim s''$. Hence $\sim$ is an equivalence relation. We restrict this relation to Q and choose an element q^0 from each equivalence class of Q and an element c_{q^0} from J_{q^0}. If $q\sim q^0$ and $q\neq q^0$, then, by the definition of P, we have $pq^0, pq\neq o$ for some p in P. Now we define c_q by the equality $c_q=c_{q^0}h_{p,q^0}^{-1}h_{p,q}$. Since $c_{q^0}\in J_{q^0}$ and C is cancellative, there exists a unique element c in C such that $c_{q^0}=ch_{p,q}^{-1}h_{p,q^0}$ and $ch_{p,q}^{-1}=c_{q^0}h_{p,q^0}^{-1}\in C$. Clearly, $c=c_q$ and hence $c_q\in C$. Let s be an element of $S\setminus Z$ such that $sq^0\neq o$. Since $pq, pq^0\neq o$, we have $sq\neq o$ and (b) implies $h_{p,q^0}^{-1}h_{p,q}h_{s,q}^{-1}=h_{s,q^0}^{-1}$. Thus we have

$$c_qh_{s,q}^{-1}=c_{q^0}h_{p,q^0}^{-1}h_{p,q}h_{s,q}^{-1}=c_{q^0}h_{s,q^0}^{-1}.$$

Hence $c_q\in J_q$. Relation (b) ensures that c_q is welldefined. Let s, t be elements of S not contained in Z. Then, on the one hand, there exists an element q in Q such that $sq\neq o$ and, on the other hand, there is an element p in P such that $pt\neq o$ and hence an element q' such that $pq'\neq o$. Let us define $f_{e,t}, g_{s,e}$ to be the uniquely determined elements of C such that

$$f_{e,t}h_{p,t}^{-1}h_{p,q'}=c_{q'} \quad\text{and}\quad g_{s,e}h_{s,q}=c_q.$$

(b) implies that $f_{e,t}$ and $g_{s,e}$ are well defined. If $z\in Z$, then $f_{e,z}$ and $g_{z,e}$ can be arbitrarily defined. By Theorem 2.3 it suffices to check (5) for the mappings defined by (6) and to check (7). Let s, t be elements of S with $st\neq o$. Then $pst\neq o$ for some p in P and $pq'\neq o$ for some q' in Q. Clearly, $ps\neq o$ and we have

$$f_{e,s}^{-1}f_{e,st}=h_{p,s}^{-1}h_{p,q'}c_{q'}^{-1}c_{q'}h_{p,q'}^{-1}h_{p,st}=g_{p,s}f_{p,s}^{-1}f_{p,st}g_{p,st}^{-1}=$$
$$=g_{p,s}f_{ps,t}g_{p,st}^{-1}=f_{s,t}.$$

In the last two equalities conditions (1) and (2) are applied. Analogously, we have $g_{t,e}^{-1}g_{st,e}=g_{s,t}$ if $st\neq o$. Finally, if $st\neq o$, then $sq\neq o$ for a q in Q and hence $pt, pq\neq o$ for some p in P. Applying (b), we have

$$g_{s,e}^{-1}f_{e,t}=h_{s,q}c_q^{-1}c_qh_{p,q}^{-1}h_{p,t}=h_{s,t},$$

as was to be proved.

It is easy to see that for any elements p, q, s, t of S and x, y, u, v of S^e

$$h_{s,t}h_{p,t}^{-1}h_{p,q}h_{s,q}^{-1} = h_{us,tv}h_{xp,tv}^{-1}h_{xp,qy}h_{us,qy}^{-1}$$

provided $ustv, xptv, xpqy, usqy \neq o$. One has only to observe that

$$h_{s,t} = f_{s,t}g_{s,t}^{-1} = f_{s,tv}f_{st,v}^{-1}f_{st,v}g_{s,tv}^{-1}f_{t,v}^{-1} = h_{s,tv}f_{t,v}^{-1}$$

and dually

$$h_{s,t} = g_{u,s}h_{us,t}.$$

A subset M of $S\setminus Z_r$ will be called *left* 0*-reversible* if for any pair s, s' of elements of S the existence of elements m in M and t in S with $st, sm, s't, s'm \neq o$ implies the existence of an element x in $tS \cap mS$ such that sx and $s'x \neq o$. It follows by straightforward calculation that in this case

$$h_{s,t}h_{s',t}^{-1}h_{s',m}h_{s,m} = h_{s,x}h_{s',x}^{-1}h_{s',x}h_{s,x} = 1.$$

Hence Theorem 2.4 implies the following

Corollary 2.5. *Suppose $Z=Z_r=Z_l$ holds in the semigroup S. Assume, furthermore, that S has the property that for any elements s, s', t, t' the relations $st, st', s't \neq o$ imply $s't' \neq o$ and S contains a left* 0*-reversible subset M such that for every element s of S not belonging to Z the set M has an element m with $sm \neq o$. Then the definition of $f, g\colon S_0^{(2)} \to C$ can be extended to S^e if and only if for each element m of M*

$$J_m = \bigcap_{\substack{s,t \\ st,\, sm \neq o}} (Ch_{s,t}^{-1} \cap C)h_{s,m} \cap C$$

is not empty.

The assumption of Corollary 2.5 is satisfied for example if S is an inverse semigroup in which the semilattice of idempotents is an orthogonal sum of semilattices. M can be chosen to be the set of idempotents. If S has no zero element then in Theorem 2.4 P and Q can be chosen to be singletons. In this case the assumption of Corollary 2.5 means that S contains a left reversible element m.

Evidently, condition (a) of Theorem 2.4 is satisfied if C is a group.

The following example shows that there exist functions $f, g\colon S_0^{(2)} \to C$ which cannot be extended to S^e, while considered as functions $f, g\colon S_0^{(2)} \to C^*$ they can be extended. Let S be the extended bicyclic semigroup defined by R. J. Warne in [14]. We denote the set of integers by I. S is the set $I \times I$ equipped with the multiplication

$$(i, j)(k, l) = \big(i+k-\min(j, k), l+j-\min(j, k)\big).$$

Clearly, $f, g\colon S \times S \to N^0$ defined by $f_{(i,j),(k,l)} = k-\min(j, k)$, $g_{(i,j),(k,l)} = j-\min(j,k)$ satisfy (1)—(3), while $J_{(k_0,l_0)}$ is empty for every (k_0, l_0). On the other hand, S is an inverse semigroup without zero, hence the definition of f, g can be extended to S^e if negative integers are allowed to be used.

Now we determine all the pairs of functions $f, g\colon F_X\times F_X\to C$ which can be defined on the free semigroup freely generated by its subset X.

Theorem 2.6. *Let C be a monoid embeddable in a group and let $\check{\chi}_1, \check{\chi}_2\colon F_X\to C$ be two functions such that $X\check{\chi}_1=X\check{\chi}_2=1$, $R_I=\{s\in F_X | s\check{\chi}_1\in I\}$ is a right ideal, $L_I=\{s\in F_X | s\check{\chi}_2\in I\}$ is a left ideal of F_X for every right ideal I of C and the mapping $\check{\varphi}=\check{\chi}_1/\check{\chi}_2$ satisfies the following condition: for all $s=x_1\dots x_n\in F_X\setminus X$, where $x_i\in X$ $(i=1,\dots,n)$, we have*

$$s\check{\varphi}=(x_1, x_2)\check{\varphi}(x_2x_3)\check{\varphi}\dots(x_{n-1}x_n)\check{\varphi}. \tag{8}$$

Then $f, g\colon F_X\times F_X\to C$ defined by

$$f_{s,t}=(s\check{\chi}_1)^{-1}(st)\check{\chi}_1, \quad g_{s,t}=(t\check{\chi}_2)^{-1}(st)\check{\chi}_2 \tag{9}$$

have the properties (1)—(3). *Conversely, for any $f, g\colon F_X\times F_X\to C$ satisfying* (1)—(3) *there exists a unique pair of functions $\check{\chi}_1, \check{\chi}_2\colon F_X\to C$ with the above properties. These functions are defined on $F_X\setminus X$ by*

$$s\check{\chi}_1=f_{x_1, x_2\dots x_n}, \quad s\check{\chi}_2=g_{x_1\dots x_{n-1}, x_n} \tag{10}$$

where $x_1,\dots,x_n\in X$ and $s=x_1\dots x_n$.

Proof. Since $R_{s\check{\chi}_1C}$ is a right ideal and $L_{t\check{\chi}_2C}$ is a left ideal of F_X, we have $f_{s,t}$ and $g_{s,t}\in C$ for all s, t in F_X. The first statement of the theorem can be verified by calculation.

Now let f, g be given with properties (1)—(3). Relations (9) ensure that the only functions $\check{\chi}_1, \check{\chi}_2$ with $X\check{\chi}_1=X\check{\chi}_2=1$ are the ones defined by (10). All we need to prove is that these functions have the required properties. (9) is implied immediately by (1) and (3). If I is a right ideal of C, $s\in R_I$, $t\in F_X$, then $(st)\check{\chi}_1=s\check{\chi}_1f_{s,t}\in I$, that is, R_I is a right ideal of F_X. Dually, L_I is a left ideal of F_X. Let $x_1,\dots,x_n\in X$, where $n\geqq 3$. Applying (9) and (10), relation (2) implies that

$$f_{x_1, x_2\dots x_n}g^{-1}_{x_1\dots x_{n-1}, x_n}=f_{x_1,x_2}g^{-1}_{x_1,x_2}f_{x_2,x_3\dots x_n}g^{-1}_{x_2\dots x_{n-1},x_n},$$

that is, we have

$$(x_1\dots x_n)\check{\varphi}=(x_1x_2)\check{\varphi}(x_2\dots x_{n-1})\check{\varphi}.$$

By induction on n one can show that (8) holds, which completes the proof.

Now it is easy to construct a pair of functions on a free semigroup such that its definition cannot be extended to the free monoid generated by the same set. Let X be the two-element set $\{x, y\}$, C the cancellative monoid of non-negative integers with the usual addition. Define $\check{\chi}_1$ in the following way: let $x^2\check{\chi}_1=y^2\check{\chi}_1=0$, $(xy)\check{\chi}_1=(yx)\check{\chi}_1=1$ and $(x_1\dots x_n)\check{\chi}_1=(x_1x_2)\check{\chi}_1+\dots+(x_{n-1}x_n)\check{\chi}_1$ if $n\geqq 3$ and $x_1,\dots,x_n\in\{x, y\}$. Let $\check{\chi}_2$ be identically 0. Obviously, these functions have the required properties enabling us to define $f, g\colon F_X\times F_X\to C$ by (5). However, $h_{y,y}-h_{x,y}+h_{x,x}-h_{y,x}=-2$.

In what follows we prove a structure theorem for 0-$\mathscr{D}$-simple semigroups with identity on which a nontrivial pair of functions $f, g\colon S_0^{(2)}\to N^0$ is defined where N^0 denotes the additive semigroup of nonnegative integers. The operation in N^0 will be denoted by $+$. Clearly, N^{0*} is the infinite cyclic group.

Lemma 2.7. *Let the semigroup S have an identity e and two elements a, b such that $ba=e$. If $f_{e,a}=n$ $(n\in N^0)$, then for all nonnegative integers k and m*

$$f_{e,a^k b^m} = kn, \quad g_{a^k b^m, e} = mn.$$

Proof. Since $ba=e$, we have $b^k a^k=e$ for all k in N^0. Hence $a^k\mathscr{L}e$ and $b^k\mathscr{R}e$. This implies by Lemmas 2.2 (ii) and 2.1 that

$$f_{e,b^k}=f_{e,e}=0 \quad\text{and}\quad g_{a^k,e}=g_{e,e}=0.$$

Using the homomorphism φ defined in Theorem 2.3 we have

$$f_{e,a^k}=f_{e,a^k}-g_{a^k,e}=a_k\varphi = k(a\varphi)=k(f_{e,a}-g_{a,e})=kf_{e,a}=kn.$$

On the other hand, we have

$$0=e\varphi=(ba)\varphi=b\varphi+a\varphi=-g_{b,e}+f_{e,a}$$

whence $g_{b,e}=n$. In the same way as above one can prove that $g_{b^k,e}=kn$. Since for all k, m in N^0 we have $a^k\mathscr{R}a^k b^m\mathscr{L}b^m$, Lemma 2.2 (ii) ensures that $f_{e,a^k b^m}= =f_{e,a^k}=kn$ and $g_{a^k b^m,e}=g_{b^m,e}=mn$.

An immediate consequence of this lemma is

Corollary 2.8. *The functions $f, g\colon B\times B\to N^0$ definable on the bicyclic semigroup B are exactly the constant multiples of f^* and g^** (see § **1.**).

Let S be a semigroup with identity e and zero element 0 on which a nontrivial pair of functions $f, g\colon S_0^{(2)}\to N^0$ is given. Let

$$F_i=\{s\in S_0 | f_{e,s}=i\}, \quad G_i=\{s\in S_0 | g_{s,e}=i\}$$

for all i in N^0. $\{F_i | i\in N^0\}$ and $\{G_i | i\in N^0\}$ are partitions of S_0. The equivalence relations induced by them will be denoted by $\mathscr{F}$ and $\mathscr{G}$, respectively. Let $\mathscr{K}=\mathscr{F}\cap\mathscr{G}$. Clearly, its equivalence classes are the sets $K_{i,j}=F_i\cap G_j$.

We remark that $R_{k+N^0}=\bigcup_{i=k}^{\infty} F_i$ and $L_{k+N^0}=\bigcup_{i=k}^{\infty} G_i$ where R_{k+N^0} and L_{k+N^0} denote the right and left ideals of S_0 respectively used in Theorem 2.3. Since φ defined in the same theorem is a homomorphism, $\bigcup_{i=k}^{\infty} K_{i,i}\cup 0$ is a subsemigroup of S for every k in N^0. Lemma 2.2 (ii) implies that $\mathscr{R}\subseteq\mathscr{F}$ and $\mathscr{L}\subseteq\mathscr{G}$. Hence if S is 0-$\mathscr{D}$-simple, then the following holds: $hr=r\,[rh=r]$ for all r in $F_i[G_j]$, whenever $hk=k[kh=k]$ for all k in $K_{i,j}$. These facts will be used without reference. To prove Theorem 2.10 we need

Lemma 2.9. *Assume that nontrivial functions $f, g\colon S_0^{(2)}\to N^0$ are given on a 0-$\mathscr{D}$-simple semigroup S with identity e such that the subsemigroup $\bigcup_{i=1}^{\infty} K_{i,i}\cup 0$ has an identity element e_1. Let $e_1\in K_{n,n}$. If $e_1=ab$ with $ba=e$, then $a^m b^m$ is the identity element of $\bigcup_{i=mn}^{\infty} K_{i,i}\cup 0$.*

Proof. We prove by induction on m that $r\in \bigcup_{i=mn}^{(m+1)n-1} K_{i,0}$ implies $a^m b^m r=r$. Clearly, this holds for $m=0$ and $r\in K_{n,0}$ implies $abr=r$. Suppose that $r\in \bigcup_{i=mn}^{(m+1)n-1} K_{i,0}$ implies $a^m b^m r=r$ for all m smaller than m' $(m'\geqq 1)$ and $r\in K_{m'n,0}$ implies $a^{m'}b^{m'}r=r$. Now let $r\in K_{j,0}$, where $m'n<j\leqq (m'+1)n$. Since ab is an identity element of $\bigcup_{i=1}^{\infty} K_{i,i}$, we have $abr=r$. Hence $br\mathscr{L}r$, that is, $br\in G_0$. On the other hand, we have

$$(br)\varphi = b\varphi + r\varphi = -n+j,$$

whence $br\in K_{-n+j,0}$. By assumption, $(a^{m'-1}b^{m'-1})br=br$, that is,

$$r = abr = a\,a^{m'-1}b^{m'-1}br = a^{m'}b^{m'}r.$$

Moreover, if $j=(m'+1)n$, i.e. $-n+j=m'n$, then $br=(a^{m'}b^{m'})br$ and

$$r = abr = a^{m'+1}b^{m'+1}r.$$

This completes the proof of the fact that $a^m b^m$ is a left identity element in $\bigcup_{i=mn}^{\infty} K_{i,i}\cup 0$. Dually, it follows that it is also a right identity.

Theorem 2.10. *Let a nontrivial pair of functions $f, g\colon S_0^{(2)}\to N^0$ be given on the 0-$\mathscr{D}$-simple semigroup S with identity e such that the subsemigroup $\bigcup_{i=1}^{\infty} K_{i,i}\cup 0$ has an identity element e_1. Assume that $e_1\in K_{n,n}$. Then*

(i) *the ranges of f and g are the set of multiples of n.*

(ii) *If, moreover, $K_{0,0}^0=K_{0,0}\cup 0$ is a subsemigroup of S and $e_1K_{0,0}\subseteq K_{n,n}$, then $S\cong \mathscr{S}^0(B, K_{0,0}^0, N^0, f^*, g^*, \varkappa)$, where B denotes the bicyclic semigroup and the endomorphism $\pi=1\varkappa$ preserves e.*

Proof. Since $e_1\in F_n$, the number n is the least positive integer with $F_n\neq\square$. The semigroup S is 0-$\mathscr{D}$-simple, hence $e\mathscr{L}a\mathscr{R}e_1$ for some a and since S is regular, there exists an inverse b of a such that $ba=e$, $ab=e_1$. Suppose that in contrast to (i), $F_p\neq\square$ for some p, where $n\nmid p$. Let d be the greatest common divisor of n and p. Then $up-vn=d$ for some positive integers u, v. Let c be an element of F_p such that $e\mathscr{L}c$. By Lemma 2.7 we have $c^u\in K_{up,0}$, and since $up>vn$, Lemma 2.9 implies $a^v b^v c^u=c^u$. Hence $b^v c^u\mathscr{L}c^u\mathscr{L}e$. However, we have

$$(b^v c^u)\varphi = v(b\varphi)+u(c\varphi) = -vn+up = d,$$

whence it follows that $b^v c^u \in K_{d,0}$ with $d<n$, a contradiction. On the other hand, $a^m \in F_{mn}$ by Lemma 2.7, which proves (i) for f. Dually, one can show (i) for g.

Turning to (ii), we first show that all the elements of S_0 can be uniquely represented in the form $a^k h b^m$, where $h \in K_{0,0}$. Let $s \in K_{kn,mn}$. Since S is 0-$\mathscr{D}$-simple, Green's lemma ensures that $s = a_k h' b_m$ for some a_k, b_m and h' such that $e \mathscr{L} a_k \mathscr{R} s$, $e \mathscr{R} b_m \mathscr{L} s$ and $h' \mathscr{H} e$. Applying Lemma 2.9 we obtain

$$s = a^k b^k s a^m b^m = a^k (b^k a_k) h' (b_m a^m) b^m.$$

Since $a_k \in K_{kn,0}$, the equality $a^k b^k a_k = a_k$ holds. Hence $b^k a_k \mathscr{L} a_k$, that is, $b^k a_k \in G_0$. Moreover, $(b^k a_k)\varphi = -kn + kn = 0$, whence $b^k a_k \in K_{0,0}$. The fact that $b_m a^m \in K_{0,0}$ can be proved similarly. By assumption $K^0_{0,0}$ is a subsemigroup of S. This implies that $h = (b^k a_k) h' (b_m a^m) \in K^0_{0,0}$ and $h \neq 0$ because $s \neq 0$. We have obtained that $s = a^k h b^m$. To prove uniqueness suppose that we have

$$a^k h b^m = a^{k'} h' b^{m'}$$

for some h, h' in $K_{0,0}$. Since

$$(a^k h b^m)\varphi = (k-m)n, \quad (a^{k'} h' b^{m'})\varphi = (k'-m')n,$$

we have $k-m = k'-m'$, that is, $k-k' = m-m' = r$. Without loss of generality we can assume that r is nonnegative. Multiplying the equality above by $b^{k'}$ on the left and by $a^{m'}$ on the right it yields $a^r h b^r = h'$. Hence $h' a^r b^r = h'$. Should $r>0$ hold, then $a^r b^r$ would belong to L_{1+N^0} implying $h' \in L_{1+N^0}$. Since $h' \in K_{0,0}$, we have $r=0$ and $h = h'$, as was to be proved.

Let h be any element of $K_{0,0}$. Since $(bh)\varphi = b\varphi + h\varphi = -n$ we have either $bh \in K_{0,n}$ or $bh \in L_{n+1+N^0}$. The latter would imply $abh = e_1 h \in L_{n+1+N^0}$, contrary to the assumption $e_1 h \in K_{n,n}$. Hence $bh \in K_{0,n}$ and by the foregoing $bh = \hat{h} b$ for the unique element $\hat{h} = bha$ of $K_{0,0}$. Similarly, we have $ha \in K_{n,0}$ and hence $ha = (ab)ha = a\hat{h}$. If $h=0$, then $\hat{h}=0$ is the unique element such that $bh = \hat{h} b$ and $ha = a\hat{h}$. The mapping π sending h into $\hat{h}$ is an endomorphism of $K^0_{0,0}$ as we have

$$[(gh)\pi] b = b(gh) = (bg)h = (g\pi) bh = (g\pi)(h\pi) b,$$

whence $(gh)\pi = (g\pi)(h\pi)$. Obviously, $e\pi = e$. Using these results we obtain the product of any two elements of S_0 in the form

$$\begin{aligned}(a^m g b^n)(a^p h b^q) &= a^m g b^{n-r} a^{p-r} h b^q = a^m g a^{p-r} b^{n-r} h b^q = \\ &= a^{m+p-r} (g\pi^{p-r} h\pi^{n-r}) b^{n+q-r},\end{aligned}$$

where $r = \min(n, p)$. In the second step we made use of the equality $b^{n-r} a^{p-r} = a^{p-r} b^{n-r}$ implied by the fact that at least one of the exponents equals 0. This implies that the mapping Φ defined by

$$s\Phi = \begin{cases} (m, h, n) & \text{if} \quad s \neq 0 \quad \text{and} \quad s = a^m h b^n, \quad h \in K_{0,0} \\ 0 & \text{if} \quad s = 0 \end{cases}$$

is an isomorphism of S onto $\mathscr{S}^0(B, K^0_{0,0}, C, f^*, g^*, \varkappa)$, where $\varkappa$ is the homomorphism of N^0 into the endomorphism monoid of $K^0_{0,0}$ mapping k into π^k.

Corollary 2.11. *In a semigroup S satisfying the conditions of Theorem 2.10 the relations $\mathscr{F}$, $\mathscr{G}$ and $\mathscr{K}$ are 0-congruences.*

Now we construct a $\mathscr{D}$-simple semigroup S with identity and a pair of functions $f, g\colon S\times S\to N^0$ with range N^0 which fail to have the property that the subsemigroup $\bigcup_{i=1}^{\infty} K_{i,i}$ has an identity element. We shall use the notions and results of W. D. Munn's paper [8].

The descending ω-chain as a meet semilattice is isomorphic to the semilattice $N^0_\wedge$ with underlying set N^0 and operation defined by

$$m\wedge n = \max(m, n).$$

Let E be the direct product $N^0_\wedge \times N^0_\wedge$. The semilattice E is uniform and has a greatest element $(0, 0)$. The set T_E of all isomorphisms of a principal ideal of E onto another one considered as partial mappings of E together with the usual multiplication is a $\mathscr{D}$-simple inverse semigroup with the semilattice of idempotents isomorphic to E. The principal ideal of E generated by (m, n) will be denoted by $[m, n]$. For each pair of elements (m, n), (p, q) of E there exist two isomorphisms $\alpha^+_{(m,n),(p,q)}$ and $\alpha^-_{(m,n),(p,q)}$ of $[m, n]$ onto $[p, q]$ defined by

$$(m+i, n+j)\alpha^+_{(m,n)(p,q)} = (p+i, q+j), \quad (m+i, n+j)\alpha^-_{(m,n)(p,q)} = (p+j, q+i),$$

where $i, j\geqq 0$. Let us define the functions χ_1 and $\chi_2\colon T_E\to N^0$ as follows:

$$\alpha^\eta_{(m,n),(p,q)}\chi_1 = m+n, \quad \alpha^\eta_{(m,n),(p,q)}\chi_2 = p+q,$$

where η may be $+$ as well as $-$. Since $\alpha^\eta_{(m,n),(p,q)}T_E\subseteq \alpha^{\eta'}_{(m',n'),(p',q')}T_E$ if and only if $(m,n)\leqq(m',n')$ in E, the set $R_{k+N^0}=\{\beta\in T_E|\beta\chi_1\in k+N^0\}$ is a right ideal of T_E. Dually, $L_{k+N^0}=\{\beta\in T_E|\beta\chi_2\in k+N^0\}$ is a left ideal of T_E. Furthermore, denoting χ_1/χ_2 by φ, we have

$$\begin{aligned}
&(\alpha^\eta_{(m,n),(p,q)}\alpha^{\eta'}_{(m',n'),(p',q')})\varphi = \\
&= \alpha^{\eta\eta'}_{(p\wedge m', q\wedge n')\alpha^\eta_{(p,q),(m,n)},(p\wedge m', q\wedge n')\alpha^{\eta'}_{(m',n'),(p',q')}}\varphi = \\
&= (m+n+(p\wedge m'-p)+(q\wedge n'-q))-(p'+q'+(p\wedge m'-m')+(q\wedge n'-n')) = \\
&= (m+n)-(p+q)+(m'+n')-(p'+q') = \\
&= \alpha^\eta_{(m,n),(p,q)}\varphi+\alpha^{\eta'}_{(m',n'),(p',q')}\varphi.
\end{aligned}$$

Hence φ is a homomorphism. Theorem 2.3 implies that $f, g\colon T_E\times T_E\to N^0$ defined by (5) have the desired properties (1)—(3). There are two idempotents in $K_{1,1}$ which are dual atoms in the semilattice of idempotents: $\alpha^+_{(0,1),(0,1)}$, $\alpha^+_{(1,0),(1,0)}$. Consequently, the subsemigroup $\bigcup_{i=1}^{\infty} K_{i,i}$ has no identity element.

3. In the present section we deal with the properties of 0-extensions $\mathscr{S}^0(S, \Sigma, C, f, g, \varkappa)$ of a semigroup Σ by S. We state a proposition on the 0-congruence induced by the 0-extension. Necessary and sufficient conditions are given for $\mathscr{S}^0(S, \Sigma, C, f, g, \varkappa)$ to have an identity, to be a regular or an inverse semigroup. We investigate its Green's relations and ideals, too. The homomorphisms of a semigroup $\mathscr{S}^0(S, \Sigma, C, f, g, \varkappa)$ into $\mathscr{S}^0(\bar{S}, \bar{\Sigma}, \bar{C}, \bar{f}, \bar{g}, \bar{\varkappa})$ are investigated in some special cases. We introduce a concept of equivalence between the 0-extensions of a semigroup Σ by another one denoted by S and deal with the equivalent 0-extensions. This section is mostly independent of section 2, the main result of which, Theorem 2.3, is used in Theorems 3.9 and 3.11 only.

Let the semigroups S and Σ be given. Consider a 0-extension $\mathscr{S}^0(S, \Sigma, C, f, g, \varkappa)$ of Σ by S over the cancellative monoid C. For brevity, denote $\mathscr{S}^0(S, \Sigma, C, f, g, \varkappa)$ by $\mathbf{S}$.

The 0-congruence induced by the 0-extension $\mathbf{S}$ will be denoted by $\mathscr{C}$. Its congruence classes are $C_r = \{(r, \sigma) | \sigma \in \Sigma_\omega\}$ and $C_0 = \{0\}$. Denote $C_r \cup 0$ by C_r^0.

Proposition 3.1. (i) *All congruence classes C_i^0 with 0 adjointed corresponding to nonzero idempotents of S are subsemigroups of* $\mathbf{S}$ *isomorphic to Σ^ω.*

(ii) *If Σ has an idempotent element ι preserved by all the endomorphisms in $\{f_{s,t}, g_{s,t} | s, t \in S, st \neq o\}\varkappa$ then*

$$\{(s, \iota) | s \in S_0\} \cup 0$$

is a subsemigroup of $\mathbf{S}$ *isomorphic to S^0.*

Proof. Since, by Lemma 2.1, $f_{i,i} = g_{i,i} = 1$ we have

$$(i, \varrho)(i, \sigma) = \begin{cases} (i, \varrho\sigma) & \text{if} \quad \varrho\sigma \neq \omega \\ 0 & \text{otherwise} \end{cases}$$

and

$$(i, \varrho)0 = 0(i, \varrho) = 0 \cdot 0 = 0.$$

Hence C_i^0 is isomorphic to Σ^ω. As for (ii) if ι has the required property then

$$(s, \iota)(t, \iota) = \begin{cases} (st, \iota) & \text{if} \quad st \neq o \\ 0 & \text{otherwise} \end{cases}$$

and

$$(s, \iota)0 = 0(s, \iota) = 0 \cdot 0 = 0.$$

Hence $\{(s, \iota) | s \neq o\} \cup 0$ is a subsemigroup ismorphic to S^0.

Proposition 3.2. (i) *An element* (i, ι) *of* $\mathbf{S}$ *is idempotent if and only if i and ι are idempotents in S and Σ, respectively.*

(ii) *The element* (e, ε) *of* $\mathbf{S}$ *is the identity of* $\mathbf{S}$ *if and only if* e, ε *are the identities of* S *and* Σ*, respectively, and the endomorphisms of* Σ *contained in* $\{f_{s,t}, g_{s,t} | s, t \in S, st \neq o\}\varkappa$ *preserve* ε.

Proof. Using the definition of $\mathbf{S}$, the element (i, ι) is idempotent if and only if

$$i^2 = i \quad \text{and} \quad \iota(f_{i,i}\varkappa)\iota(g_{i,i}\varkappa) = \iota.$$

By Lemma 2.1, $i^2=i$ implies $f_{i,i}=g_{i,i}=1$. Thus the above condition is equivalent to the following one: $i^2=i$ and $\iota^2=\iota$. Similarly, (e, ε) is an identity if and only if for any s, σ

$$se = es = s \quad \text{and} \quad \sigma(f_{s,e}\varkappa)\varepsilon(g_{s,e}\varkappa) = \varepsilon(f_{e,s}\varkappa)\sigma(g_{e,s}\varkappa) = \sigma.$$

Lemma 2.1 ensures $f_{s,e}=g_{e,s}=1$, so that the latter equality says that

$$\sigma\varepsilon(g_{s,e}\varkappa) = \varepsilon(f_{e,s}\varkappa)\sigma = \sigma$$

for any s, σ. Taking $s=e$ this yields that ε is the identity of Σ. But then $\varepsilon(f_{e,s}\varkappa)= =\varepsilon(g_{s,e})=\varepsilon$ for all $s\neq o$. Let s, t be any elements of S such that $st\neq o$. Applying (1) and the fact that $\varkappa$ is a homomorphism we have

$$\varepsilon = \varepsilon(f_{e,st}\varkappa) = \varepsilon(f_{e,s}\varkappa)(f_{s,t}\varkappa) = \varepsilon(f_{s,t}\varkappa),$$

and similarly, by (3), we have $\varepsilon=\varepsilon(g_{s,t}\varkappa)$. Conversely, if e and ε are identities of S and Σ and $\varkappa$ has the desired property, then (e, ε) is clearly an identity.

Theorem 3.3. (i) *The semigroup* $\mathbf{S}$ *is regular if and only if both* S *and* Σ *are regular.*

(ii) $\mathbf{S}$ *is an inverse semigroup if and only if both* S *and* Σ *are inverse semigroups.*

Proof. We show that two elements (r, ϱ) and (s, σ) of $\mathbf{S}$ are inverses of each other if and only if r, s and $\varrho, \sigma(f_{s,r}g_{r,sr}\varkappa)$ are inverses of each other in S and Σ, respectively, where $f_{s,r}g_{r,sr}\varkappa$ is an automorphism of Σ. This proves the theorem. By definition (r, ϱ) and (s, σ) are inverses of each other if and only if

$$rsr = r, \quad srs = s, \tag{11}$$

$$\varrho(f_{r,sr}\varkappa)\sigma(f_{s,r}g_{r,sr}\varkappa)\varrho(g_{rs,r}\varkappa) = \varrho, \quad \sigma(f_{s,rs}\varkappa)\varrho(f_{r,s}g_{s,rs}\varkappa)\sigma(g_{sr,s}\varkappa) = \sigma. \tag{12}$$

Using (2), (1) and (3), (11) implies that

$$f_{r,s}g_{s,rs}f_{s,r}g_{r,sr} = f_{r,s}f_{rs,r}g_{s,r}g_{r,sr} = f_{r,sr}g_{rs,r}.$$

It follows from Lemma 2.1 that $f_{r,sr}=g_{rs,r}=f_{s,rs}=g_{sr,s}=1$. Hence $f_{r,sr}\varkappa= =g_{rs,r}\varkappa=f_{s,rs}\varkappa=g_{sr,s}\varkappa$, the identity automorphisms of Σ and $f_{r,s}g_{s,rs}\varkappa$, $f_{s,r}g_{r,sr}\varkappa$ are automorphisms of Σ, inverses of each other. Thus conditions (11), (12) are equivalent to (11) and

(13) $\varrho\sigma(f_{s,r}g_{r,sr}\varkappa)\varrho = \varrho, \quad \sigma(f_{s,r}g_{r,sr}\varkappa)\varrho\sigma(f_{s,r}g_{r,sr}\varkappa) = \sigma(f_{s,r}g_{r,sr}\varkappa),$

as was to be proved.

As for Green's relations and ideals of the semigroup $\mathrm{S}=\mathscr{S}^0(S, \Sigma, C, f, g, \varkappa)$ in general one cannot say more than the definitions of them. Therefore we deal with the case when S is regular.

Lemma 3.4. *Let S be a regular semigroup. The principal left [right] ideal of* S *generated by (s, σ) is contained in the one generated by (r, ϱ) if and only if $(r, \varrho)=$ $=(s, \sigma)$ or the following conditions are satisfied:*

(a) $s\in Sr[s\in rS]$,

(b) $\sigma\in\Sigma\varrho\pi[\sigma\in g\pi\Sigma]$ *where* $\pi = g_{x,r}\varkappa(xr = s)\,[\pi = f_{r,x}(rx = s)]$

is an endomorphism of Σ depending only on r and s.

Proof. First we note that $xr=s$ and $x'r=s$ imply $g_{x,r}=g_{x',r}$. Indeed, if i is an idempotent in the $\mathscr{L}$-class containing r, then $ri=r$ and $si=s$. Thus by (3) we have

$$g_{r,i}g_{x,r} = g_{s,i} \quad \text{and} \quad g_{r,i}g_{x',r} = g_{s,i}.$$

Since C is cancellative, $g_{x,r}=g_{x',r}$.

By definition, $(s, \sigma)\in\mathrm{S}(r, \varrho)$ means that there exist elements x and ξ in S and Σ, respectively, such that

(14) $$xr = s, \quad \xi(f_{x,r}\varkappa)\varrho(g_{x,r}\varkappa) = \sigma.$$

Hence the necessity of (a), (b) is proved. Conversely, assume that (a), (b) hold, that is, there exist x and ξ' in S and Σ, respectively, such that

$$xr = s, \quad \xi'\varrho\pi = \sigma.$$

Since S is regular, x can be chosen to satisfy $x\mathscr{R}s$. If j is an idempotent belonging to the $\mathscr{R}$-class of s, then $j=sw$ for some w and the equality

$$f_{j,x}f_{x,r}f_{s,w} = f_{j,s}f_{s,w} = f_{j,j} = 1$$

follows from (1) and Lemma 2.1. Hence $f_{x,r}$ is in the group of units of C and $f_{x,r}\varkappa$ is an automorphism. Thus $\xi'=\xi(f_{x,r}\varkappa)$ for some ξ, that is, by (14), $(s, \sigma)\in\mathrm{S}(r, \varrho)$. The statement for right ideals can be proved dually.

The next theorem deals with Green's relations of S.

Theorem 3.5. *Let S be a regular semigroup. Two distinct elements (r, ϱ) and (s, σ) of S are $\mathscr{L}[\mathscr{R}]$-equivalent if and only if*

(a) $r\mathscr{L}s[r\mathscr{R}s]$ *in* S *and*

(b) $\sigma\in\Sigma\varrho\alpha,\ \varrho\alpha\in\Sigma\sigma[\sigma\in\varrho\alpha\Sigma,\ \varrho\alpha\in\sigma\Sigma]$ *where* $\alpha=g_{x,r}\varkappa\,(xr=s)$ $[\alpha=f_{r,x}\varkappa(rx=s)]$ *is an automorphism of Σ depending only on r and s.*

Note that if the group of units of C is trivial, then $g_{x,r}=1[f_{r,x}=1]$, whence α is the identity automorphism.

Proof. By Lemma 3.4. all we need to show is that if $xr=s, ys=r$ then $g_{x,r}\varkappa$ and $g_{y,s}\varkappa$ are automorphisms of Σ being inverses of each other. To prove this one has only to observe that we have

$$g_{x,r}g_{y,s}=g_{x,r}g_{y,xr}=g_{yx,r}=1$$

by (3) and Lemma 2.1.

An immediate consequence of this theorem is

Corollary 3.6. *Let S be regular and Σ have the property that $\sigma\in\sigma\Sigma\cap\Sigma\sigma$ for all elements σ in Σ.*

(i) *The distinct elements (r, ϱ) and (s, σ) of* **S** *are $\mathscr{D}$-equivalent if and only if $r\mathscr{D}s$ and there exists an element t in S such that $r\mathscr{L}t\mathscr{R}s$ and $\varrho\alpha\mathscr{D}\sigma\beta$, where $\alpha=g_{x,r}\varkappa$, $\beta=f_{s,y}\varkappa\,(xr=sy=t)$ are automorphisms of Σ depending only on r, s and t.*

(ii) *If both S and Σ are (0-)$\mathscr{D}$-simple, then* **S** *is also 0-$\mathscr{D}$-simple.*

(iii) *If the group of units of C is trivial, then* **S** *is 0-$\mathscr{D}$-simple if and only if S and Σ are (0-)$\mathscr{D}$-simple.*

To make the formulation of the theorem on the ideals of **S** easier we introduce the following notations. If **C** is a subset of **S** containing the 0 element let

$$C=\{s\in S|\exists\sigma\in\Sigma\;(s,\sigma)\in\mathbf{C}\}$$

provided S has no zero element and adjoin o to C if $o\in S$.

For all c in C different from o define Γ_c as

$$\Gamma_c=\{\sigma\in\Sigma|(c,\sigma)\in\mathbf{C}\}$$

if Σ has no zero element and adjoin ω to Γ_c if $\omega\in\Sigma$.

In particular, the subsets corresponding to the subsets of **S** denoted by **L**, **R** and **D** are denoted by L, $\Lambda_l (l\in L, l\neq o)$, R, $P_r\,(r\in R, r\neq o)$ and D, $\Delta_d\,(d\in D, d\neq o)$, respectively.

Theorem 3.7. *Let S be a regular semigroup. A subset* **L**[**R**] *of the semigroup* **S** *containing the* 0 *element is a left [right] ideal if and only if*

(a) *$L[R]$ is a left [right] ideal of S and*

(b) *for all elements l of $L[r$ of $R]$, l' of $Sl\,[r'$ of $rS]$ different from o and λ of $\Lambda_l[\varrho$ of $P_r]$ $\Sigma\lambda\pi\subseteq\Lambda_{l'}[\varrho\pi\Sigma\subseteq P_{r'}]$ holds where $\pi=g_{x,l}\varkappa(xl=l')[\pi=f_{r,x}\varkappa(rx=r')]$ is an endomorphism of Σ depending only on l and $l'[r$ and $r']$.*

The proof of this theorem is easy thanks to Lemma 3.4, therefore it is left to the reader.

Corollary 3.8. *Suppose the semigroup S is regular. The semigroup* $\mathbf{S}$ *is* 0*-simple if and only if S is* (0-)*simple and*

$$\Sigma = \bigcup_{\pi \in E_t^s} \Sigma\sigma\pi\Sigma \tag{16}$$

for all s, t and σ different from o and ω, respectively, where

$$E_t^s = \{(g_{x,s}f_{xs,y})\varkappa \mid x\mathcal{R}t\mathcal{L}y \quad and \quad xsy = t\}.$$

Proof. $\mathbf{S}$ is 0-simple if and only if for every element (s, σ) in it the ideal $\mathbf{D}=\mathbf{S}(s, \sigma)\mathbf{S}$ is equal to S itself. By the last theorem this means that S is (0-)simple and $\Delta_t = \Sigma$ for every $t \neq o$. So it is sufficient to prove that the right side of the equality (16) is equal to Δ_t. Theorem 3.7 (b) ensures that $\bigcup_{\pi \in E_t^s} \Sigma\sigma\pi\Sigma \subseteq \Delta_t$. Conversely, since the nonzero elements of $\mathbf{D}$ have the form $(x', \xi)(s, \sigma)(y', \eta)$, Δ_t is contained in the ideal $\bigcup_{\substack{x', y' \\ x'sy'=t}} \Sigma\sigma(g_{x',s}f_{x's,y'})\varkappa\Sigma$. Let i_1 and i_2 be idempotents such that $i_1\mathcal{R}t\mathcal{L}i_2$ and let $x=i_1x'$, $y=y'i_2$. Obviously, $x\mathcal{R}t\mathcal{L}y$ and $xsy=t$. Applying identities (1)—(3) and Lemma 2.1, we see that

$$g_{x',s}f_{x's,y'} = g_{x',s}f_{x's,y'}g_{i_1,x'sy'} = g_{x',s}g_{i_1,x's}f_{xs,y'} = g_{x,s}f_{xs,y'} =$$

$$= g_{x,s}f_{xs,y'}f_{xsy',i_2} = g_{x,s}f_{xs,y}.$$

Hence $(g_{x',s}f_{x's,y'})\varkappa \in E_t^s$ and $\Delta_t \subseteq \bigcup_{\pi \in E_t^s} \Sigma\sigma\pi\Sigma$, which completes the proof.

In what follows we deal with homomorphisms of a semigroup $\mathbf{S}=\mathcal{S}^0(S, \Sigma, C, f, g, \varkappa)$ into another one $\bar{\mathbf{S}}=\mathcal{S}^0(\bar{S}, \bar{\Sigma}, \bar{C}, \bar{f}, \bar{g}, \bar{\varkappa})$ in two special cases. In the first case we assume that $C\varkappa$ and $\bar{C}\bar{\varkappa}$ are contained in the group of automorphisms of Σ and $\bar{\Sigma}$, respectively. Then without loss of generality we can assume C and $\bar{C}$ to be groups. In the alternative case we suppose that S and $\bar{S}$ are inverse semigroups.

Theorem 3.9. *Suppose that C and $\bar{C}$ are groups and the definition of f, g and $\bar{f}$, $\bar{g}$ can be extended to S^e and $\bar{S}^{\bar{e}}$, respectively. Denote the suitable homomorphisms used in Theorem 2.3 by $\varphi\colon S_0^e \to C$ and $\bar{\varphi}\colon \bar{S}_{\bar{0}}^{\bar{e}} \to \bar{C}$. Suppose four mappings $m_1\colon S^0 \to \bar{S}^{\bar{0}}$, $\mu_1\colon \Sigma^\omega \to \bar{S}^{\bar{0}}$, $m_2\colon S^0 \to \bar{\Sigma}^{\bar{\omega}}$ and $\mu_2\colon \Sigma^\omega \to \bar{\Sigma}^{\bar{\omega}}$ are given with the following properties:*

(a) $om_1 = \omega\mu_1, \quad om_2 = \omega\mu_2, \quad \bar{o}m_1^{-1} = \bar{\omega}m_2^{-1}, \quad \bar{o}\mu_1^{-1} = \bar{\omega}\mu_2^{-1}$.

(b) *For any r, s, ϱ, σ in S and Σ, respectively,*

$$(rs)m_1 = rm_1 \cdot sm_1, \tag{17}$$

$$(\varrho\sigma)\mu_1 = \varrho\mu_1 \cdot \sigma\mu_1 \tag{18}$$

if $(rs)m_1 \neq \bar{o}$ *and* $(\varrho\sigma)\mu_1 \neq \bar{o}$, *furthermore*

$$(rs)m_2 = rm_2 \cdot sm_2(rm_1\bar{\varphi}\bar{\varkappa})^{-1}, \tag{19}$$

$$(\varrho\sigma)\mu_2 = \varrho\mu_2 \cdot \sigma\mu_2(\varrho\mu_1\bar{\varphi}\bar{\varkappa})^{-1} \tag{20}$$

whenever $(rs)m_1 \neq \bar{o}$ *or* $rm_1 \cdot sm_1 \neq \bar{o}$ *and* $(\varrho\sigma)\mu_1 \neq \bar{o}$ *or* $\varrho\mu_1 \cdot \sigma\mu_1 \neq \bar{o}$.

(c) *For any* r, ϱ *in* S *and* Σ, *respectively,*

$$rm_1 \cdot \varrho\mu_1 = \varrho(r\varphi\varkappa)^{-1}\mu_1 \cdot rm_1, \tag{21}$$

$$rm_2 \cdot \varrho\mu_2(rm_1\bar{\varphi}\bar{\varkappa})^{-1} = \varrho(r\varphi\varkappa)^{-1}\mu_2 \cdot rm_2\big(\varrho(r\varphi\varkappa)^{-1}\mu_1\bar{\varphi}\bar{\varkappa}\big)^{-1}$$

if all the four elements are different from $\bar{o}$ *and* $\bar{\omega}$, *respectively. In addition, the left hand sides differ from zero if and only if the right hand sides do so.*

Define a mapping $\Phi\colon \mathbf{S} \to \bar{\mathbf{S}}$ *in the following way. Let*

$$(r, \varrho)\varphi = \big(\varrho'\mu_1 \cdot rm_1, (\varrho'\mu_2 \cdot rm_2(\varrho'\mu_1\bar{\varphi}\bar{\varkappa})^{-1}\big)\big(f_{\bar{e},\varrho'\mu_1 \cdot rm_1}\bar{\varkappa}\big)\big),$$

where $\varrho' = \varrho(f_{e,r}\varkappa)^{-1}$, *when both components on the right are nonzero and* $(r,\varrho)\Phi = \bar{0}$ *otherwise. Further, let* $0\Phi = (om_1, om_2)$ *if* $om_1 \neq \bar{o}$ *and* $0\Phi = \bar{0}$ *otherwise. Then*

(i) *the mapping* Φ *is a homomorphism.*

(ii) Φ *is an isomorphism if and only if* $0\Phi = \bar{0}$ *and for all nonzero elements* $\bar{r}$, $\bar{\varrho}$ *of* $\bar{S}$ *and* $\bar{\Sigma}$, *respectively, there exist uniquely determined elements* r *and* ϱ *in* S *and* Σ *such that*

$$\bar{r} = \varrho\mu_1 \cdot rm_1 \quad \text{and} \quad \bar{\varrho} = \varrho\mu_2 \cdot rm_2(\varrho\mu_1\bar{\varphi}\bar{\varkappa})^{-1}.$$

(iii) *If the semigroups* S, $\bar{S}$, Σ *and* $\bar{\Sigma}$ *have identity elements, then all the homomorphisms of* $\mathbf{S}$ *into* $\bar{\mathbf{S}}$ *are of this form.*

Proof. It is not difficult to check statement (i) by computation. (ii) is implied immediately by the definition of Φ and the fact that the elements of $C\varkappa$ and $\bar{C}\bar{\varkappa}$ are automorphisms. Turning to (iii), consider the semigroups S, $\bar{S}$, Σ, $\bar{\Sigma}$ with identity. Since $f_{e,r}\varkappa$ is an automorphism, for any nonzero r, ϱ we have

$$(r, \varrho) = \big(e, \varrho(f_{e,r}\varkappa)^{-1}\big)(r, \varepsilon).$$

Hence all the nonzero elements of S can be uniquely written in the form $(e, \varrho)(r, \varepsilon)$, where $\{(e, \varrho) | \varrho \in \Sigma_\omega\} \cup 0$ and $\{(r, \varepsilon) | r \in S_0\} \cup 0$ are subsemigroups isomorphic to Σ^ω and S^0, respectively. Let $\Phi\colon \mathbf{S} \to \bar{\mathbf{S}}$ be a homomorphism. Define the mappings $m_1\colon S^0 \to \bar{S}^{\bar{o}}$, $\mu_1\colon \Sigma^\omega \to \bar{S}^{\bar{o}}$, $m_2\colon S^0 \to \bar{\Sigma}^{\bar{\omega}}$, $\mu_2\colon \Sigma^\omega \to \bar{\Sigma}^{\bar{\omega}}$ as follows. Let $om_1 = \omega\mu_1 = \bar{o}$, $om_2 = \omega\mu_2 = \bar{\omega}$ if $0\Phi = \bar{0}$ and $om_1 = \omega\mu_1 = \bar{r}$, $om_2 = \omega\mu_2 = \bar{\varrho}$ if $0\Phi = (\bar{e}, \bar{\varrho})(\bar{r}, \bar{\varepsilon})$. Let $rm_1 = \bar{o}$, $rm_2 = \bar{\omega}$ if $(r, \varepsilon)\Phi = \bar{0}$ and $\varrho\mu_1 = \bar{o}$, $\varrho\mu_2 = \bar{\omega}$ if $(e, \varrho)\Phi = \bar{0}$, respectively. In the opposite case, let

$$(r, \varepsilon)\Phi = (\bar{e}, rm_2)(rm_1, \bar{\varepsilon}) = \big(rm_1, rm_2(\bar{f}_{\bar{e},rm_1}\bar{\varkappa})\big),$$

$$(e, \varrho)\Phi = (\bar{e}, \varrho\mu_2)(\varrho\mu_1, \bar{\varepsilon}) = \big(\varrho\mu_1, \varrho\mu_2(\bar{f}_{\bar{e},\varrho\mu_1}\bar{\varkappa})\big).$$

Clearly, $om_1=\omega\mu_1$, $om_2=\omega\mu_2$, $\bar{o}m_1^{-1}=\bar{\omega}m_2^{-1}$ and $\bar{o}\mu_1^{-1}=\bar{\omega}\mu_2^{-1}$. Relations (1), (3) and the fact that $\bar{\varphi}$ is a homomorphism yield that for any $\bar{r}$, $\bar{s}$ with $\bar{r}\bar{s}\neq\bar{o}$ we have

$$\begin{gathered}\bar{f}_{\bar{e},\bar{r}}\bar{f}_{\bar{r},\bar{s}}=\bar{f}_{\bar{e},\bar{r}\bar{s}},\\ \bar{f}_{\bar{e},\bar{s}}\bar{g}_{\bar{r},\bar{s}}=\bar{f}_{\bar{e},\bar{s}}\bar{g}_{\bar{s},\bar{e}}^{-1}\bar{g}_{\bar{r}\bar{s},\bar{e}}=\bar{s}\bar{\varphi}(\bar{r}\bar{s}\bar{\varphi})^{-1}\bar{f}_{\bar{e},\bar{r}\bar{s}}=(\bar{r}\bar{\varphi})^{-1}\bar{f}_{\bar{e},\bar{r}\bar{s}}.\end{gathered}\tag{22}$$

Since $\varkappa$ is a homomorphism, we have

$$\begin{aligned}(r,\varepsilon)\Phi\cdot(s,\varepsilon)\Phi&=\big(rm_1\cdot sm_1,\, rm_2(\bar{f}_{\bar{e},rm_1}\bar{\varkappa})(\bar{f}_{rm_1,sm_1}\bar{\varkappa})\cdot sm_2(\bar{f}_{\bar{e},sm_1}\bar{\varkappa})(\bar{g}_{rm_1,sm_1}\bar{\varkappa})\big)=\\ &=\big(rm_1\cdot sm_1,\big(rm_2\cdot sm_2(rm_1\bar{\varphi}\bar{\varkappa})^{-1}\big)\bar{f}_{\bar{e},rm_1\cdot sm_1}\bar{\varkappa}\big)\end{aligned}$$

for every pair r,s, whenever both components on the right side are nonzero and $(r,\varepsilon)\Phi\cdot(s,\varepsilon)\Phi=\bar{0}$ in the opposite case. The same equality holds if (r,ε) or (s,ε) is replaced by 0, that is, if $r=o$ or $s=o$. But Φ is a homomorphism, $(r,\varepsilon)\Phi\cdot(s,\varepsilon)\Phi=$ $=(rs,\varepsilon)\Phi$, which proves that (17) and (19) hold under the conditions mentioned in the theorem. Investigating Φ restricted to the subsemigroup $\{(e,\varrho)|\varrho\in\Sigma_\omega\}\cup 0$ one can verify (18) and (20). Observe that for any r and ϱ

$$(r,\varepsilon)(e,\varrho)=\big(e,\varrho(r\varphi\varkappa)^{-1}\big)(r,\varepsilon).$$

Hence $(r,\varepsilon)\Phi(e,\varrho)\Phi=\big(e,\varrho(r\varphi\varkappa)^{-1}\big)\Phi\cdot(r,\varepsilon)\Phi$, that is, denoting $\varrho(r\varphi\varkappa)^{-1}\mu_1$ by $\bar{s}$ and $\varrho(r\varphi\varkappa)^{-1}\mu_2$ by $\bar{\sigma}$, we have

$$\big(rm_1\cdot\varrho\mu_1,\big(rm_2\cdot\varrho\mu_2(rm_1\bar{\varphi}\bar{\varkappa})^{-1}\big)(\bar{f}_{\bar{e},rm_1\cdot\varrho\mu_1}\bar{\varkappa})\big)=\big(\bar{s}\cdot rm_1,\big(\bar{\sigma}\cdot rm_2(\bar{s}\bar{\varphi}\bar{\varkappa})^{-1}(\bar{f}_{\bar{e},\bar{s}\cdot rm_1}\big)\big)$$

if all the components are nonzero. Moreover, if a component is zero on one side, then so is one on the other side. This is equivalent to condition (21), which completes the proof.

In the next theorem we use the notation $f_{r,s}^{-1}$ only if $f_{r,s}$ is contained in the group of units of C. If r is an element of an inverse semigroup the unique idempotent belonging to the $\mathscr{R}$-class containing r will be denoted by $[r]$.

Theorem 3.10. *Let S and $\bar{S}$ be inverse semigroups, E and $\bar{E}$ the semilattices of their idempotents, respectively. Assume that mappings $m_1\colon S^0\to\bar{S}^{\bar{0}}$, $m_2\colon S^0\to\bar{\Sigma}^{\bar{\omega}}$, $\mu_1^i\colon\Sigma^\omega\to\bar{S}^{\bar{0}}$, $\mu_2^i\colon\Sigma^\omega\to\bar{\Sigma}^{\bar{\omega}}$ $(i\in E_0)$ are given such that they have the following properties.*

(a) *For each i in E_0 we have $om_1=\omega\mu_1^i$, $om_2=\omega\mu_2^i$, $\bar{o}m_1^{-1}=\bar{\omega}m_2^{-1}$ and $\bar{o}(\mu_1^i)^{-1}=$ $=\bar{\omega}(\mu_2^i)^{-1}$.*

(b) *For any r,s,ϱ,σ in S and Σ, respectively, and i in E_0 we have*

$$(rs)m_1=rm_1\cdot sm_1,$$

$$(\varrho\sigma)\mu_1^i=\varrho\mu_1^i\cdot\sigma\mu_1^i.$$

if $(rs)m_1 \neq \bar{o}$ *and* $(\varrho\sigma)\mu_1^i \neq \bar{o}$. *Further we have*

$$(rs)m_2 = rm_2(\bar{f}_{[rm_1],[rm_1\cdot sm_1]}\bar{\varkappa}) \cdot sm_2((\bar{g}_{rm_1,[sm_1]}\bar{f}^{-1}_{[rm_1\cdot sm_1],rm_1[sm_1]})\bar{\varkappa}),$$
$$(\varrho\sigma)\mu_2^i = \varrho\mu_2^i(\bar{f}_{[\varrho\mu_1^i],[\varrho\mu_1^i\cdot\sigma\mu_1^i]}\bar{\varkappa}) \cdot \sigma\mu_2((\bar{g}_{\varrho\mu_1^i,[\sigma\mu_1^i]}\bar{f}^{-1}_{[\varrho\mu_1^i\cdot\sigma\mu_1^i],\varrho\mu_1^i[\sigma\mu_1^i]})\bar{\varkappa})$$

if $(rs)m_1 \neq \bar{o}$ *or* $rm_1\cdot sm_1 \neq \bar{o}$ *and* $(\varrho\sigma)\mu_1^i \neq \bar{o}$ *or* $\varrho\mu_1^i\cdot\sigma\mu_1^i \neq \bar{o}$.

(c) *For any* r, ϱ *in* S *and* Σ *and* i *in* E_0

$$rm_1\cdot\varrho\mu_1^i = \varrho((g_{r,i}f^{-1}_{[ri],ri})\varkappa)\mu_1^{[ri]}\cdot(ri)m_1,$$
$$rm_2(\bar{f}_{[rm_1],[rm_1\cdot\varrho\mu_1^i]}\bar{\varkappa})\cdot\varrho\mu_2^i((\bar{g}_{rm_1,[\varrho\mu_1^i]}\bar{f}^{-1}_{[rm_1\cdot\varrho\mu_1^i],rm_1[\varrho\mu_1^i]})\bar{\varkappa}) =$$
$$= \varrho'\mu_2^{[ri]}(\bar{f}_{[\varrho'\mu_1^{[ri]}],[\varrho'\mu_1^{[ri]}\cdot(ri)m_1]}\bar{\varkappa})\cdot(ri)m_2((\bar{g}_{\varrho'\mu_1^{[ri]},[(ri)m_1]}\bar{f}^{-1}_{[\varrho'\mu_1^{[ri]}\cdot(ri)m_1],\varrho'\mu_1^{[ri]}[(ri)m_1]})\bar{\varkappa})$$

if all the four elements are different from zero. Moreover, the left hand sides of the equalities differ from zero if and only if the right hand sides do so, too. In the second equality $\varrho' = \varrho((g_{r,i}f^{-1}_{[ri],ri})\varkappa)$. *Define a mapping* $\Phi\colon \mathbf{S}\to\bar{\mathbf{S}}$ *in the following way. Let*

$$(r,\varrho)\Phi = (\sigma\mu_1^{[r]}\cdot rm_1, (\sigma\mu_2^{[r]}(\bar{f}_{[\sigma\mu_1^{[r]}],[\sigma\mu_1^{[r]}\cdot rm_1]}\bar{\varkappa})\cdot$$
$$\cdot rm_2((\bar{g}_{\sigma\mu_1^{[r]},[rm_1]}\bar{f}^{-1}_{[\sigma\mu_1^{[r]}\cdot rm_1],\sigma\mu_1^{[r]}[rm_1]})\bar{\varkappa}))(\bar{f}_{[\sigma\mu_1^{[r]}\cdot rm_1],\sigma\mu_1^{[r]}\cdot rm_1}\bar{\varkappa})),$$

where $\sigma = \varrho(f^{-1}_{[r],r}\varkappa)$, *whenever both components are different from zero,* $(r,\varrho)\Phi = \bar{0}$ *otherwise. Further, put* $0\Phi = (om_1, om_2)$ *if* $om_1 \neq \bar{o}$ *and* $0\varphi = \bar{0}$ *otherwise. Then*

(i) *the mapping* Φ *is a homomorphism,*

(ii) *if the semigroups* Σ *and* $\bar{\Sigma}$ *have identity elements preserved by all endomorphisms* $f_{r,s}\varkappa$, $g_{r,s}\varkappa$ *and* $\bar{f}_{\bar{r},\bar{s}}\bar{\varkappa}$, $\bar{g}_{\bar{r},\bar{s}}\bar{\varkappa}$, *respectively, then all homomorphisms of* $\mathbf{S}$ *into* $\bar{\mathbf{S}}$ *are of this form.*

Proof. (i) can be verified by computation. If S is an inverse semigroup and Σ has an identity preserved by the endomorphisms $f_{r,s}\varkappa$ and $g_{r,s}\varkappa$, then all the nonzero elements (r,ϱ) of S can be uniquely written in the form $([r],\varrho')(r,\varepsilon)$ with $\varrho' = \varrho(f^{-1}_{[r],r}\varkappa)$ because $f_{[r],r}$ is in the group of units of C by Lemmas 2.2 and 2.1 and hence $f^{-1}_{[r],r}\varkappa$ is an automorphism. Since the proof of (ii) is similar to that of Theorem 3.9 (iii), it is left to the reader. We note only that by (1)—(3) we have

$$f_{[r],r}f_{r,s} = f_{[r],rs} = f_{[r],[rs]}f_{[rs],rs}.$$

Since $[rs] = rss^{-1}r^{-1}$, we have $[rs]r[s] = r[s]$ and

$$f_{[rs],r[s]}f_{[rs]r[s],s}f_{rs,s^{-1}r^{-1}} = f_{[rs],rs}f_{rs,s^{-1}r^{-1}} = f_{[rs],[rs]} = 1.$$

Hence $f_{[rs],r[s]}$ is in the group of units of C and we have

$$f_{[s],s}g_{r,s} = g_{r,[s]}f_{r[s],s} = g_{r,[s]}f^{-1}_{[rs],r[s]}f_{[rs],rs}.$$

Let S and Σ be two semigroups and consider two 0-extensions $\mathbf{S} = \mathcal{S}^0(S, \Sigma, C, f, g, \varkappa)$ and $\mathbf{S}' = \mathcal{S}^0(S, \Sigma, \bar{C}, \bar{f}, \bar{g}, \bar{\varkappa})$ of Σ by S.

Definition. The 0-extension S is said to be equivalent to $\bar{\mathrm{S}}'$ if for every s in S_0 there exists an automorphism ψ_s of Σ such that the mapping $\Psi\colon \mathrm{S}\to\mathrm{S}'$ defined by $0\Psi=\bar{0}$, $(s,\sigma)\Psi=(s,\sigma\psi_s)$ is an isomorphism.

This definition clearly determines an equivalence relation on the class of 0-extensions of Σ by S.

In the next theorems we investigate the equivalent 0-extensions.

Before formulating the first one we note that if the images of the functions $f\varkappa$ and $g\varkappa$ are contained in the group of automorphisms of Σ, then Theorem 2.3 applies to them provided S has an identity. The homomorphism used in this theorem will be denoted by $\varphi^{\varkappa}$.

Theorem 3.11. *Let S be a semigroup with identity and Σ a reductive semigroup. Assume that the images of $f\varkappa$ and $g\varkappa$ are in the group of automorphisms of Σ. The 0-extension S is equivalent to S' if and only if the images $\bar{f}\bar{\varkappa}$ and $\bar{g}\bar{\varkappa}$ are included in the group of automorphisms of Σ and $\bar{\varphi}^{\bar{\varkappa}}=\varphi^{\varkappa}\mathfrak{A}$ for some inner automorphism $\mathfrak{A}$ of the group of automorphisms of Σ.*

Proof. Suppose S and S' are equivalent. This means that for any r,s in S with $rs\neq o$ and ϱ,σ in Σ we have

$$(r,\varrho\psi_r)(s,\sigma\psi_s)=\big(rs,\big(\varrho(f_{r,s}\varkappa)\sigma(g_{r,s}\varkappa)\big)\psi_{rs}\big)$$

provided they are nonzero or else both of them are zero. In both cases we have

$$\varrho\psi_r(\bar{f}_{r,s}\bar{\varkappa})\sigma\psi_s(\bar{g}_{r,s}\bar{\varkappa})=\varrho(f_{r,s}\varkappa)\psi_{rs}\sigma(g_{r,s}\varkappa)\psi_{rs}.$$

For $r=e$ this yields

$$\varrho\psi_e(\bar{f}_{e,s}\bar{\varkappa})\sigma\psi_s=\varrho(f_{e,s}\varkappa)\psi_s\sigma\psi_s.$$

Σ is reductive and ψ_s is an automorphism. Hence for any $s\neq o$ we have

$$\psi_e(\bar{f}_{e,s}\bar{\varkappa})=(f_{e,s}\varkappa)\psi_s.$$

Dually, one can see that

$$g_{s,}\psi_{e,}(\bar{g}_{se}\bar{\varkappa})=({}_e\varkappa)\psi_s$$

for $s\neq o$. From these equalities it follows that $\bar{f}_{e,s}\bar{\varkappa}$ and $\bar{g}_{s,e}\bar{\varkappa}$ are automorphisms for all $s\neq o$, which implies by (1) and (3) that so are $\bar{f}_{r,s}\bar{\varkappa}$ and $\bar{g}_{r,s}\bar{\varkappa}$, where $rs\neq o$. Moreover, we have

$$s\bar{\varphi}^{\bar{\varkappa}}=\psi_e^{-1}(s\varphi^{\varkappa})\psi_e,$$

which completes the proof of the only if part. Conversely, suppose that the conditions of the theorem are satisfied. Denote the automorphism of Σ inducing $\mathfrak{A}$ by ψ. Define ψ_s by

$$\psi_s=(f_{e,s}\varkappa)^{-1}\psi(\bar{f}_{e,s}\bar{\varkappa}).$$

Making use of the equalities (22) and the fact that $s\bar{\varphi}^{\bar{\varkappa}}=\psi^{-1}(s\varphi^{\varkappa})\psi$ holds for every

$s \neq o$, one can obtain by computation

$$\psi_r(\bar{f}_{r,s}\bar{\varkappa}) = (f_{r,s}\varkappa)\psi_{rs}, \quad \psi_s(g_{r,s}\bar{\varkappa}) = (g_{r,s}\varkappa)\psi_{rs},$$

whenever $rs \neq o$.

Theorem 3.12. *Consider a regular semigroup S and a reductive semigroup $\sum$. The 0-extensions* S *and* S′ *are equivalent if and only if there exist automorphisms ψ_i of Σ indexed by the idempotents of S_0 such that for every pair i, i' of idempotents $i'i=i'$ and $i'i=i'$ imply*

$$(23) \qquad (f_{i,i'}\varkappa)\psi_{i'} = \psi_i(\bar{f}_{i,i'}\bar{\varkappa}) \quad \textit{and} \quad (g_{i',i}\varkappa)\psi_{i'} = \psi_i(\bar{g}_{i',i}\bar{\varkappa}),$$

respectively, and $i\mathcal{D}i'$ implies that

$$(24) \qquad (f_{i,r}^{-1}\varkappa)\psi_i(\bar{f}_{i,r}\bar{\varkappa}) = (g_{r,i'}^{-1}\varkappa)\psi_{i'}(\bar{g}_{r,i'}\bar{\varkappa})$$

for any r such that $i\mathcal{R}r\mathcal{L}i'$.

Proof. In the proof of the last theorem we saw that the 0-extensions are equivalent if and only if the equality

$$\varrho\psi_r(\bar{f}_{r,s}\bar{\varkappa})\sigma\psi_s(\bar{g}_{r,s}\bar{\varkappa}) = \varrho(f_{r,s}\varkappa)\psi_{rs}\sigma(g_{r,s}\varkappa)\psi_{rs}$$

holds for any r, s such that $rs \neq o$ and for arbitrary ϱ, σ in Σ. If $ir=r$, this implies by Lemma 2.1 that

$$\varrho\psi_i(\bar{f}_{i,r}\bar{\varkappa}) \cdot \sigma\psi_r = \varrho(f_{i,r}\varkappa)\psi_r \cdot \sigma\psi_r.$$

Since Σ is reductive we have

$$(f_{i,r}\varkappa)\psi_r = \psi_i(\bar{f}_{i,r}\bar{\varkappa}).$$

Dually, we can obtain that

$$(g_{r,i'}\varkappa)\psi_r = \psi_{i'}(\bar{g}_{r,i}\bar{\varkappa})$$

if $ri'=r$. In particular, this yields (23) if r is an idempotent. If $i\mathcal{R}r\mathcal{L}i'$, then, as it has been verified above, $f_{i,r}$ and $g_{r,i'}$ belong to the group of units of C. Hence $f_{i,r}\varkappa$ and $g_{r,i'}\varkappa$ are automorphisms and it follows from the foregoing that

$$(25) \qquad \psi_r = (f_{i,r}^{-1}\varkappa)\psi_i(\bar{f}_{i,r}\bar{\varkappa}) = (g_{r,i'}^{-1}\varkappa)\psi_{i'}(\bar{g}_{r,i'}\bar{\varkappa}).$$

Conversely, assume that the conditions of the theorem hold for some automorphisms ψ_i. Let r be an element of S_0 and i, i' idempotents such that $i\mathcal{R}r\mathcal{L}i'$. Define ψ_r by (25). Obviously, ψ_r is well defined. If j is an idempotent such that $ji=i$, then applying (1) and (23) we have

$$(f_{j,r}\varkappa)\psi_r = ((f_{j,r}f_{i,r}^{-1})\varkappa)\psi_i(\bar{f}_{i,r}\bar{\varkappa}) = (f_{j,i}\varkappa)\psi_i(\bar{f}_{i,r}\bar{\varkappa}) =$$
$$= \psi_j((\bar{f}_{j,i}\bar{f}_{i,r})\bar{\varkappa}) = \psi_j(\bar{f}_{j,r}\bar{\varkappa}).$$

Hence if r, s are elements of S such that $rs \neq o$, then denoting an idempotent in the $\mathcal{R}$-class of r by i, we have

$$\psi_r(\bar{f}_{r,s}\bar{\varkappa}) = (f_{i,r}^{-1}\varkappa)\psi_i((\bar{f}_{i,r}\bar{f}_{r,s})\bar{\varkappa}) = (f_{i,r}^{-1}\varkappa)\psi_i(\bar{f}_{i,rs}\bar{\varkappa}) =$$
$$= ((f_{i,r}^{-1}f_{i,rs})\varkappa)\psi_{rs} = (f_{r,s}\varkappa)\psi_{rs}.$$

Similarly, one can show that $\psi_s(\bar{g}_{r,s}\bar{\varkappa})=(g_{r,s}\varkappa)\psi_{rs}$. This completes the proof.

Note that if S is (0-)$\mathscr{D}$-simple and C has a trivial group of units, then (24) implies that all ψ_i are equal. Conversely, if all ψ_i coincide, then, denoting ψ_i $(i\in E_0)$ by ψ, (23) implies

$$\bar{f}_{i,r}\bar{\varkappa}=\psi^{-1}(f_{i,r}\varkappa)\psi \quad \text{and} \quad g_{r,j}\bar{\varkappa}=\psi^{-1}(g_{r,j}\varkappa)\psi,$$

whenever $ir=r$ and $rj=r$, respectively. Hence (24) holds trivially. Thus we have proved the following

Corollary 3.13. *If S is a (0-)$\mathscr{D}$-simple regular semigroup and Σ is reductive, then the 0-extensions* S *and* S′ *are equivalent if and only if there exists an automorphism ψ of Σ such that for all idempotents i, i' the equalities*

$$\bar{f}_{i,i'}\bar{\varkappa}=\psi^{-1}(f_{i,i'}\varkappa)\psi \quad \textit{and} \quad \bar{g}_{i',i}\bar{\varkappa}=\psi^{-1}(g_{i',i}\varkappa)\psi$$

are implied by $ii'=i'$ and $i'i=i'$, respectively.

References

[1] J. E. Ault and M. Petrich, The structure of ω-regular semigroups, *Bull. Amer. Math. Soc.*, **77** (1971), 196—199.

[2] J. E. Ault and M. Petrich, The structure of ω-regular semigroups, *J. reine angew. Math.*, **251** (1971), 110—141.

[3] A. H. Clifford and G. B. Preston, *The Algebraic Theory of Semigroups*, Vol. I, II., Math. Surveys No. 7, Amer. Math. Soc. (Providence, R. I., 1961, 1967).

[4] Б. П. Кочин, Строение инверсных идеално простых ω-полугрупп, *Вестник Ленинград. Унив.*, 1968, №. 7, 41—50.

[5] K. Krohn and J. Rhodes, Algebraic theory of machines. I. Prime decomposition theorem for finite semigroups and machines, *Trans. Amer. Math. Soc.*, **116** (1965), 450—464.

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

[7] W. D. Munn and N. R. Reilly, Congruences on a bisimple ω-semigroup, *Proc. Glasgow Math. Assoc.*, **7** (1966), 184—192.

[8] W. D. Munn, Uniform semilattices and bisimple inverse semigroups, *Quart. J. Math. Oxford Ser.*, (2) **17** (1966), 151—159.

[9] W. D. Munn, Regular ω-semigroups, *Glasgow Math. J.*, **9** (1968), 46—66.

[10] W. D. Munn, 0-bisimple inverse semigroups, *J. Algebra*, **15** (1970), 570—588.

[11] W. D. Munn, On simple inverse semigroups, *Semigroup Forum*, **1** (1970), 63—74.

[12] M. Petrich, *Structure of Regular Semigroups*, Pennsylvania State University, 1973.

[13] N. R. Reilly, Bisimple ω-semigroups, *Proc. Glasgow Math. Assoc.*, **7** (1966), 160—167.

[14] R. J. Warne, I-bisimple semigroups, *Trans. Amer. Math. Soc.*, **130** (1968), 367—386.

[15] R. J. Warne, I-regular semigroups, *Math. Japon.*, **15** (1970), 91—100.

[16] R. J. Warne, Some properties of simple I-regular semigroups, *Compositio Math.*, **22** (1970), 181—195.

BOLYAI INSTITUTE, UNIV. SZEGED
ARADI VÉRTANÚK TERE 1.
6720 SZEGED, HUNGARY

Acta Sci. Math., 38 (1977). 391—396

Generalization of the implicit function theorem and of Banach's open mapping theorem

T. SZILÁGYI

In the present paper we prove the existence of implicit functions (Theorem 1) and of "right-inverse" functions (Theorem 2) under very weak assumptions. In Theorem 3 we generalize the open mapping theorem of Banach to a non-linear case and in Theorem 4 we give a new proof of a known multiplier rule (see [4]). The proof of Theorem 1 is based on Banach's open mapping theorem (see for example [2]), on Nadler's fixed point theorem for multivalued contractions (see [3]), and on the Lagrange inequality (see for example [1]). Theorem 2 is a simple consequence of Theorem 1, Theorem 3 follows easily from Theorem 2, finally Theorem 4 is based on Theorem 2 and on the Banach—Hahn theorem.

Notations. If X and Y are Banach spaces, then the set of all linear continuous mappings from X into Y will be denoted by $L(X, Y)$.

For defining equations we use the symbol $:=$ on the left side of which we write the "quantity" (number, function, set, etc.) to be defined.

If (X, d) is a metric space, r a positive number and $x \in X$, then

$$S(x, r) := \{y \in X \mid d(x, y) < r\} \quad \textit{and} \quad B(x, r) := \{y \in X \mid d(x, y) \leqq r\}.$$

The dual of a Banach-space X will be denoted by X'.

In X and Y are Banach-spaces and $A \in L(X, Y)$, then

$$p(A) := \sup_{y \in Y \setminus \{0\}} \left\{ \|y\|^{-1} \cdot \inf \{\|x\| \mid x \in X, Ax = y\} \right\}.$$

Lemma 1. *If X and Y are Banach-spaces, $A \in L(X, Y)$ and* $\operatorname{Im} A = Y$, *then $p(A)$ is finite.*

Proof. The conditions of Banach's open mapping theorem are fulfilled, therefore there exists a positive r such that $B(O_Y, r)$ is contained in the A-image of

Received December 9, 1976.

$B(O_X, 1)$. Let us take an arbitrary $0 \neq y \in Y$, then

$$\inf \{\|x\| \, x \in X, Ax = y\} = \frac{\|y\|}{r} \inf\left\{\|x\| \,|\, x \in X, Ax = \frac{ry}{\|y\|}\right\} \leqq \frac{\|y\|}{r};$$

consequently $p(A) \leqq \dfrac{1}{r}$.

Lemma 2. *Let (X, d) be a complete metric space, $\bar{x} \in X$, $r>0$ and $\Phi\colon S(\bar{x}, r) \to 2^X$ such that*

a) *for all $x \in S(\bar{x}, r)$, $\Phi(x)$ is a non-empty closed subset of X,*

b) *for all $x_1, x_2 \in S(\bar{x}, r)$, the Hausdorff distance*

$$h(\Phi(x_1), \Phi(x_2)) := \max\{\sup_{x \in \Phi(x_1)} d(x, \Phi(x_2)), \sup_{x \in \Phi(x_2)} d(\Phi(x_1), x)\}$$

satisfies

$$h(\Phi(x_1), \Phi(x_2)) \leqq \frac{1}{2} d(x_1, x_2),$$

c) $d(\bar{x}, \Phi(\bar{x})) < \dfrac{r}{2}$.

Then there exists an $x \in S(\bar{x}, r)$ such that $x \in \Phi(x)$.

The proof of this lemma can be found in [3], and in [4].

Lemma 3. *Let X be a normed space, L a linear subspace of X, $u_1, u_2 \in X$; $M_i := u_i + L$ $(i=1, 2)$. Then the Hausdorff distance* (see Lemma 2) *of M_1 and M_2 equals* $\inf\{\|v_1 - v_2\| : v_1 \in M_1, v_2 \in M_2\}$.

Proof. Clearly,

$$\inf_{v_i \in M_i} \|v_1 - v_2\| \leqq \inf_{v_2 \in M_2} \|u_1 - v_2\| = d(u_1, M_2) \leqq \sup_{v_1 \in M_1} d(v_1, M_2) \leqq h(M_1, M_2).$$

If $v_1 \in M_1$ and $v_2 \in M_2$, then $v_i = u_i + y_i$ $(i=1, 2, y_i \in L)$, thus $u_2 + y_2 - y_1 \in M_2$ and

$$\|v_1 - v_2\| = \|u_1 - (u_2 + y_2 - y_1)\| \geqq d(u_1, M_2),$$

consequently,

$$\inf_{v_i \in M_i} \|v_1 - v_2\| = d(u_1, M_2).$$

Similar arguments show that for all $v_1 \in M_1$ and $v_2 \in M_2$

$$d(u_1, M_2) = d(v_1, M_2) = d(M_1, v_2),$$

therefore $d(u_1, M_2) = h(M_1, M_2)$.

Theorem 1. *Let X and Y be Banach spaces, T a topological space, $G \subset T \times X$ an open set, $(t_0, x_0) \in G$ and $F\colon G \to Y$ a function such that*

a) $F(t_0, x_0) = O_Y$, $\lim\limits_{t \to t_0} F(t, x_0) = O_Y$;

b) *for every* $(t, x)\in G$ *the function* $F(t, \cdot)$ *has a Fréchet-derivative at* x *denoted by* $D_2F(t, x)$,

c) *the function* $D_2F\colon G\to L(X, Y)$ *is continuous at* (t_0, x_0), *and*

d) $\operatorname{Im} D_2F(t_0, x_0)=Y$.

Then for every neighborhood $V\subset X$ *of* x_0 *there exist a neighborhood* $U\subset T$ *of* t_0 *and a function* $\varphi\colon U\to V$ *such that* $F(t, \varphi(t)) = O_Y$ *for all* $t\in U$.

Proof. Let us denote $A:=D_2F(t_0, x_0)$. By Lemma 1 we have $p(A)<+\infty$. By assumption c) there exist a neighborhood U_1 of t_0 and a positive number r such that $W:=S(x_0, r)\subset V$ and for all $(t, x)\in U_1\times W$

$$\|D_2F(t, x)-A\| < \frac{1}{2p(A)}.$$

Now we get from the Lagrange inequality for all (t, z_1) and $(t, z_2)\in U_1\times W$

$$\text{(1)}\qquad \|F(t, z_2)-F(t, z_1)-A(z_2-z_1)\| \leqq$$

$$\leqq \sup_{\lambda\in[0,1]} \|D_2F(t, \lambda z_1+(1-\lambda)z_2)-A\|\,\|z_2-z_1\| < \frac{\|z_2-z_1\|}{2p(A)}.$$

By assumption a) we may choose a neighborhood U of t_0 such that

$$\text{(2)}\qquad p(A)\|F(t, x_0)\| < \frac{r}{2} \quad \text{for all} \quad t\in U.$$

Let $t\in U$ be a fixed element. We shall show that the equation $F(t, x)=0$ has a solution $x\in W$. We apply Lemma 2 to the Banach space X, to the element $\bar{x}:=O_X$ and to the (multivalued) function

$$x\mapsto \Phi(x):=\{z\in X \mid Ax-Az = F(t, x_0+x)\}.$$

If $x\in S(0, \mathrm{r})$, assumption d) implies that $\Phi(x)$ is non-empty, the continuity of A implies that $\Phi(x)$ is closed. Moreover, since A is linear, $\Phi(x)$ is an affine subspace. Therefore (by Lemma 3) if $x_1, x_2\in S(O_X, r)$, then

$$h(\Phi(x_1), \Phi(x_2)) = \inf_{v_i\in\Phi(x_i)} \|v_1-v_2\| =$$

$$= \inf\{\|v_1-v_2\| \mid Av_i = Ax_i-F(t, x_0+x_i)\ i = 1, 2\}.$$

Since $\operatorname{Im} A=Y$, the latter infimum equals

$$\inf\{\|v\| \mid Av = A(x_1-x_2)-F(t, x_0+x_1)+F(t, x_0+x_2)\} \leqq$$

$$p(A)\|F(t, x_0+x_2)-F(t, x_0+x_1)-A(x_2-x_1)\| \leqq \frac{1}{2}\|x_1-x_2\|$$

(see (1)). From (2) it follows that condition c) in Lemma 2 is fulfilled, too:

$$d(0_X, \Phi(0_X)) = \inf\{\|z\| \mid Az = -F(t, x_0)\} \leqq p(A)\|F(t, x_0)\| < \frac{r}{2}.$$

Thus, by Lemma 2, there exists an element $g(t)\in S(O_X, r)$ such that $g(t(\in \Phi(g/t))$, consequently

$$0_Y = A(g(t)) - A(g(t)) = F(t, x_0 + g(t)).$$

Finally, for the element $\varphi(t) := x_0 + g(t) \in V$ we have

$$F(t, \varphi(t)) = 0_Y.$$

Theorem 2. *Suppose we are given two Banach spaces: X and Y, an open set $V\subset X$, an element $x_0\in V$ and a Fréchet-differentiable function $f\colon V\to Y$ for which*

a) *$f'\colon V\to L(X, Y)$ is continuous at x_0,*

b) $\operatorname{Im} f'(x_0) = Y$.

Then there exist a neighborhood U of the point $t_0 := f(x_0)$ and a function $\varphi\colon U\to V$ such that $f\circ\varphi = id_U$ (that is $f(\varphi(t)) = t$ for all $t\in U$); consequently, t_0 is an interior point of the range of f.

Proof. Let us define the function $F\colon Y\times V\to Y$ by letting

$$F(t, x) := f(x) - t.$$

Clearly, we can apply Theorem 1 whith $T := Y$ and $G := Y\times V$ and this gives the result to be proved.

Theorem 3. *If X and Y are Banach spaces, $g\colon X\to Y$ is a Fréchet-differentiable function, $g'\colon X\to L(X, Y)$ is continuous and* $\operatorname{Im} g'(x) = Y$ *in every point $x\in X$, then g is an open mapping.*

Proof. Let $V\subset X$ be an open set and $f := g|_V$. We must prove that the range R of f is an open set in Y. If $t_0\in R$, then there exists a point $x_0\in V$, for which $f(x_0) = t_0$. From Theorem 2 it follows that R is a neighborhood of t_0.

Theorem 4. *Let X and Z be Banach spaces, $W\subset X$ an open set, $g\colon W\to R$ and $G\colon W\to Z$ Fréchet-differentiable functions. If a point $x_0\in W$ affords a local minimum to g under the constraint $G(x) = O_Z$, g' and G' are continuous at x_0 and* $\operatorname{Im} G'(x_0)$ *is closed in Z, then there exist a real number λ and a continuous linear functional $l\in Z'$ such that*

(i) *at least one of them is different from* 0,

(ii) *for all $x\in X$, $\lambda g'(x_0)x + l(G'(x_0)x) = 0$.*

Proof. Let us choose an open set $V\subset X$ containing x_0 such that x_0 minimizes the function $g|_V$ under the constraint $G|_V = O_Z$ and let us denote $Y := R\times Z$; for all $x \in V$ $f(x) := (g(x), G(x))$. From our assumptions it follows that the function $f: V \to Y$ is Fréchet-differentiable, f' is continuous at x_0, and for all $x \in X$

$$f'(x_0)x = (g'(x_0)x, G'(x_0)x).$$

First we observe that $\operatorname{Im} f'(x_0) \neq Y$. Indeed, if $\operatorname{Im} f'(x_0)$ were the whole space Y, then we could apply Theorem 2: there would be points $x \in V$ with $G(x)=0$ and $g(x) < g(x_0)$, since $(g(x_0), G(x_0))$ would be an interior point to the range of f. But this is impossible, because x_0 is a solution of the minimum problem on V. Therefore $\operatorname{Im} f'(x_0)$ is a proper linear subspace of Y. If it is a closed subspace, then we can apply a known corollary of the Banach—Hahn theorem: there exists a $0 \neq \bar{l} \in Y'$ such that $\bar{l} \circ f'(x_0) = 0$; and since the continuous linear functional $\bar{l}$, defined on the product space $R \times Z$ is of the form

$$\bar{l}(t, z) = \lambda t + l(z)$$

(where $\lambda \in R$ and $l \in Z'$), in this case the proof is complete. If the subspace $\operatorname{Im} f'(x_0)$ is not closed, then there exist a sequence $(x_n) \subset X$ and an element $(r, z) \in Y \setminus \operatorname{Im} f'(x_0)$ such that

$$\lim g'(x_0)x_n = r \quad \text{and} \quad \lim G'(x_0)x_n = z.$$

Since $\operatorname{Im} G'(x_0)$ is closed, there is an element $u \in X$ such that $G'(x_0)u = z$. Now we observe that if $x \in \operatorname{Ker} G'(x_0)$, then $g'(x_0)x = 0$ (and consequently if $G'(x_0)u_1 = G'(x_0)u_2$, then $g'(x_0)u_1 = g'(x_0)u_2$). Indeed, if $g'(x_0)x$ were different from 0, then for the real number

$$t := \frac{r - g'(x_0)u}{g'(x_0)x}$$

we would get

$$g'(x_0)(u+tx) = r \quad \text{and} \quad G'(x_0)(u+tx) = z,$$

that is, $(r, z) \in \operatorname{Im} f'(x_0)$. Therefore we can define a functional l_1 on $\operatorname{Im} G'(x_0)$ in the following way: if $z \in \operatorname{Im} G'(x_0)$ and $u \in X$ such that $G'(x_0)u = z$, then

$$l_1(z) := g'(x_0)u.$$

Obviously, l_1 is linear and $g'(x_0) = l_1 \circ G'(x_0)$. Moreover, l_1 is continuous: if $U \subset R$ is any open set, then $(g'(x_0))^{-1}(U)$ is open since $g'(x_0)$ is continuous, and

$$l_1^{-1}(U) = G'(x_0)\,[(g'(x_0)^{-1}(U)]$$

is open in the Banach space $\operatorname{Im} G'(x_0)$, because $G'(x_0): X \to \operatorname{Im} G'(x_0)$ is an open mapping. By the Banach—Hahn theorem there is an extension $l \in Z'$ of l_1, it satisfies (ii) with $\lambda = -1$, as $\quad l \circ G'(x_0) = l \circ G'(x_0) = g'(x_0)$.

Remark. It is known (see [4]) that Theorem 4 implies various transversality conditions and Euler—Lagrange equations concerning the classical problems im the calculus of variations.

References

[1] J. Dieudonné, *Foundations of Modern Analysis*, Academic Press (New York—London, 1960).
[2] W. Rudin, *Functional Analysis*, McGraw-Hill Book Company (New York, 1973).
[3] S. B. Nadler, Jr., Multivalued contraction mappings, *Pacific J. Math.*, 30 (1969), 475—488.
[4] А. Д. Иоффе—В. М. Тихомиров, *Теория экстремальных задач*, Изд. Наука (Москва, 1974).

ELTE TTK ANALÍZIS II
1445 BUDAPEST 8.
PF. 323.

Acta Sci. Math., 38 (1977). 397—402

Endomorphism and subalgebra structure; a concrete characterization

N. SAUER and M. G. STONE

§ 1. Introduction

In [5] the following abstract structure problem is solved: For what semigroups S and what lattices L does there exist an algebra $\mathfrak{A}$ with $S \cong \mathrm{End}\, \mathfrak{A}$, the endomorphisms of $\mathfrak{A}$, and $L \cong \mathrm{Su}\, \mathfrak{A}$ the lattice of subalgebras of $\mathfrak{A}$? Here we provide a solution to the corresponding concrete representation problem, where isomorphism is replaced by equality. Thus let $S \subseteqq A^A$ be a given transformation monoid and $L \subseteqq 2^A$ a set lattice. It is well known that $L = \mathrm{Su}\, \mathfrak{A}$ for some algebra $\mathfrak{A}$ over the set A iff L is complete and compactly generated [2]; such lattices are called *algebraic*. In [4] necessary and sufficient conditions for $S = \mathrm{End}\, \mathfrak{A}$ for some algebra $\mathfrak{A}$ over the set A are given; such transformation semigroups are called *algebraic*. A similar characterization is given in [4] for semigroups of partial functions. We make use of the latter result by representing subalgebras with partial identity functions to derive a simultaneous characterization for S and L. Our characterization, like that for the endomorphisms alone involves the solutions to systems of linear equations.

If M is a set of partial functions on A to A with $\mathrm{id} \in M$ the identity function on A, a *system of linear equations* Σ *over* M is a set of functional equations each of the general form: $fx = y$, or $fx = g$ with $f, g \in M$, together with a specified solution variable X^Σ. An *assignment* α for Σ is a map from the variables of Σ to partial functions on A to A with a common domain. The assignment α *satisfies* Σ *at* $d \in A$. provided $f(\alpha x(d)) = \alpha y(d)$ whenever $fx = y \in \Sigma$ and $f(\alpha x(d)) = g(d)$ whenever $fx = g \in \Sigma$. The assignment α *satisfies* Σ *on* $D \subseteqq A$ iff α satisfies Σ at d for each $d \in D$. If X^Σ is the specified solution variable we say f *is a solution to* Σ *on* D provided there is

Received July 1, 1976.

This research was supported in part by NRC Operating grants A7213 and A8094 from the National Research Council of Canada.

an assignment α which satisfies Σ on D and $\alpha(X^{\Sigma})=f$. A solution f to Σ on D is *unique* provided $f \restriction D = h \restriction D$ whenever h is any solution to Σ on D. The *support* of a system Σ is the set of all $d \in A$ for which there exists a solution to Σ at d. We write $B = \mathrm{Spt}\,\Sigma$ if B is the support of Σ.

Denote by $\mathfrak{A}_M$ the algebra of all finitary operations which admit each $f \in M$ a a homomorphism. $\tilde{M}$ is the set of all partial endomorphisms of $\mathfrak{A}_M$ and $\overline{M}$ is the set of all (total) functions which are endomorphisms of $\mathfrak{A}_M$. As usual a partial function g is a homomorphism with respect to an operation P of rank ν provided $gP(x)$ is defined and equals $P(gx)$ whenever gx is defined for $x \in A^{\nu}$. A total function is one whose domain is all of A. We will use:

Proposition 1. *$g \in A^B$ belongs to $\tilde{M}$ iff $B \in \mathrm{Su}\,\mathfrak{A}_M$ and for each finite $D \subseteq B$ there is a system Σ over M with g a unique solution to Σ on D.*

Proof. Take $\mu = \aleph_0$ in Theorem 2 of [4]. □

§ 2. The subalgebras of $\mathfrak{A}_M$

We first establish some easy facts about the support of systems Σ over M.

Lemma 1. *If $C = \mathrm{Spt}\,\Sigma$ then there is an assignment α which satisfies Σ on C.*

Proof. For each $d \in C$ there is an assignment α_d which satisfies Σ at d. Define α for a variable x of Σ by:

$$\alpha x(d) = \begin{cases} \alpha_d x(d) & \text{if } d \in C \\ d & \text{otherwise.} \end{cases}$$

It is straightforward to verify that α satisfies Σ on C. □

Lemma 2. *If $C = \mathrm{Spt}\,\Sigma$ then there is a system Γ and an assignment β which satisfies Γ on C* and *$C = \mathrm{Spt}\,\Gamma$ and $\beta(X^{\Gamma}) = \mathrm{id} \restriction C$ is a unique solution to Γ on C.*

Proof. Let Γ have one additional new variable X^{Γ} not among those of Σ and let the equations of Γ consist of those of Σ together with the new equation $X^{\Gamma} = \mathrm{id}$. By Lemma 1 there is an assignment α which satisfies Σ on C. Let β extend α by assigning $\mathrm{id} \restriction C$ to X^{Γ}. Clearly β satisfies Γ on C and $C = \mathrm{Spt}\,\Gamma$. If g is any solution to Γ on C then for $d \in C$, $g(d) = d$ so $g \restriction C = \beta(X^{\Gamma}) = \mathrm{id} \restriction C$ thus $\mathrm{id} \restriction C$ is a unique solution to Γ on C. □

Lemma 3. *Let each $C \in \mathscr{F}$ be the support of some system Γ_C. Then $\bigcap_{C \in \mathscr{F}} C$ is also the support of some system Γ.*

Proof. Assume without loss of generality that each pair of systems Γ_C, Γ_D have no variables in common for $C \neq D$ and let X^Γ be a new variable distinct from all of those of the Γ_C. By Lemma 2 we may further assume that $\mathrm{id} \restriction C$ is a unique solution to Γ_c on C for each $C \in \mathscr{F}$. Form $\Gamma = \left(\bigcup_{C \in \mathscr{F}} \Gamma_C\right) \cup \{X^\Gamma = \mathrm{id}\}$. We claim $\bigcap_{C \cap \mathscr{F}} C = \mathrm{Spt}\, \Gamma$. If $d \in \mathrm{Spt}\, \Gamma$, say α satisfies Γ at d, then clearly α_C, the restriction of α to the variables of Γ_C, satisfies Γ_C at d for each $C \in \mathscr{F}$ so $d \in \bigcap_{C \in \mathscr{F}} C$. If on the other hand $d \in \bigcap_{C \in \mathscr{F}} C$ and α_C satisfies Γ_C at d then let $\alpha X = \alpha_C X \restriction \bigcap_{C \in \mathscr{F}} C$ for a variable X in Γ_C and let $\alpha X^\Gamma = \mathrm{id} \restriction \bigcap_{C \in \mathscr{F}} C$. Clearly α satisfies Γ on $\bigcap_{C \in \mathscr{F}} C$ so $d \in \mathrm{Spt}\, \Gamma$. Thus $\mathrm{Spt}\, \Gamma = \bigcap_{C \in \mathscr{F}} C$.

Lemma 4. *For $D \subseteqq A$ the operation defined by $\bar{D} = \bigcap_{D \subseteqq \mathrm{Spt}\, \Sigma} \mathrm{Spt}\, \Sigma$ is a closure operator.*

Proof. Clevarly $D \subseteqq \bar{D}$, and $[C \subseteqq D \Rightarrow \bar{C} \subseteqq \bar{D}]$. To show $\bar{\bar{D}} = \bar{D}$ it is only necessary to see then that $\bar{\bar{D}} \subseteqq \bar{D}$. By Lemma 3 there is some system Γ with $\bar{D} = \bigcap_{D \subseteqq \mathrm{Spt}\, \Sigma} \mathrm{Spt}\, \Sigma = \mathrm{Spt}\, \Gamma$. Clearly $\bar{\bar{D}} = \bigcap_{\bar{D} \subseteqq \mathrm{Spt}\, \Sigma} \mathrm{Spt}\, \Sigma = \mathrm{Spt}\, \Gamma = \bar{D}$.

Lemma 5. *For $D \subseteqq A$ the operation defined by $\tilde{D} = \bigcup_{C \text{ finite}, C \subseteqq D} \bar{C}$ is a closure operator.*

Proof. Since $d \in D \Rightarrow d \in \overline{\{d\}} \subseteqq \tilde{D}$ we have $D \subseteqq \tilde{D}$. Further $[C \subseteqq D \Rightarrow \tilde{C} \subseteqq \tilde{D}]$ since each finite subset of C is also a finite subset of D. To show $\tilde{\tilde{D}} = \tilde{D}$ it remains only to see $\tilde{\tilde{D}} \subseteqq \tilde{D}$. Suppose $\tilde{\tilde{D}} \nsubseteq \tilde{D}$; then there is some $a \in \tilde{\tilde{D}}$ with $a \notin \tilde{D}$. We will show this leads to a contradiction. Since $a \in \tilde{\tilde{D}}$ there is some $B \subseteqq \tilde{D}$, B finite, with $a \in \bar{B}$. Thus $a \in \bigcap_{B \subseteqq \mathrm{Spt}\, \Sigma} \mathrm{Spt}\, \Sigma$, and say $B = \{b_1, \ldots, b_n\}$. From $B \subseteqq \tilde{D}$ we have each $b_K \in \tilde{D}$, say $b_K \in \bar{C}_K$ for some $C_K \subseteqq D$, C_K finite, $K = 1, \ldots, n$. Then $C = \bigcup_{K=1}^{n} C_K$ is a finite subset of D, so $\bar{C} \subseteqq \tilde{D}$. Now $B \subseteqq \bigcup_{K=1}^{n} \bar{C}_K \subseteqq \overline{\bigcup_{K=1}^{n} C_K}$ so $a \in \bar{B} \subseteqq \overline{\overline{\bigcup_{K=1}^{n} C_K}} = \overline{\bigcup_{K=1}^{n} C_K} = \bar{C} \subseteqq \tilde{D}$. Thus $a \in \tilde{D}$, contrary to the original choice $a \notin \tilde{D}$, the desired contradiction.

We can now describe explicitly the subalgebras of $\mathfrak{A}_M$:

Theorem 1. $B \in \mathrm{Su}\, \mathfrak{A}_M$ *iff* $B = \bigcup_{D \text{ finite}, D \subseteqq B} \bar{D}$.

Proof. Let $B = \bigcup_{D \text{ finite}, D \subseteqq B} \bar{D}$. We first consider the case $B = \emptyset$. Thus for each $D \subseteqq B$, $D = \emptyset$ and $\bar{D} = \emptyset$ so $\bar{D} = \bigcap_{\emptyset \subseteqq \mathrm{Spt}\, \Sigma} \mathrm{Spt}\, \Sigma$. If $\mathfrak{A}_M$ has any nullary operations (constants) $a \in A$ we claim $a \in \mathrm{Spz}\, \Sigma$ for every system Σ. To see this let α be the assignment which associates with every variable the constant function $\mathrm{f}: A \to \{a\}$.

Note α satisfies arbitrary Σ at a, since each $g \in M$ must have a as a fixed point. Thus from $\bar{D}=\emptyset$ we conclude $\mathfrak{A}_M$ has no nullary operations, whence $B=\emptyset \in \in \mathrm{Su}\, \mathfrak{A}_M$. Now if $B \neq \emptyset$, fix an operation P of $\mathfrak{A}_M$ of rank n, and $a_1, a_2, \ldots, a_n \in B$. It suffices to show that $P(a_1, \ldots, a_n) \in \overline{\{a_1, \ldots, a_n\}}$ since $\overline{\{a_1, \ldots, a_n\}} \subseteq B$. Let $D = \{a_1, \ldots, a_n\}$. If $P(a_1, \ldots, a_n) \notin \bar{D}$ then there is some Σ with $D \subseteq \mathrm{Spt}\, \Sigma$ and $P(a_1, \ldots, a_n) \notin \mathrm{Spt}\, \Sigma$. By Lemma 1 we may assume that there is an assignment α which satisfies Σ on Spt Σ. We use α to produce an assignment α' which satisfies Σ at $d=P(a_1, \ldots, a_n)$ and thus obtain $P(a_1, \ldots, a_n) \in \mathrm{Spt}\, \Sigma$ contradicting the hypothesis that $P(a_1, \ldots, a_n) \notin \bar{D}$. For a variable x in Σ let $\alpha' x$ be defined by

$$\alpha' x(d) = \begin{cases} \alpha x(d) & \text{if } d \neq P(a_1, \ldots, a_n) \\ P(\alpha x(a_1), \ldots, \alpha x(a_n)) & \text{if } d = P(a_1, \ldots, a_n). \end{cases}$$

We claim α' satisfies Σ at $d=P(a_1, \ldots, a_n)$. To see this consider an equation $fx=y$ in Σ:

$$f\alpha' x(d) = P(f\alpha x(a_1), \ldots, f\alpha x(a_n)) = P(g(a_1), \ldots, g(a_n)) = gP(a_1, \ldots, a_n) = g(d).$$

Thus $d \in \mathrm{Spt}\, \Sigma$ and we must conclude $P(a_1, \ldots, a_n) \in \overline{\{a_1, \ldots, a_n\}}$, so $B = \bigcup_{D \text{ finite}, D \subseteq B} \bar{D} \Rightarrow \Rightarrow B \in \mathrm{Su}\, \mathfrak{A}_M$. To prove the converse, we suppose $B \in \mathrm{Su}\, \mathfrak{A}_M$. By Lemma 5 $B \subseteq \tilde{B} = \bigcup_{D \text{ finite}, D \subseteq B} \bar{D}$ so it remains only to show $B \supseteq \tilde{B}$. We proceed again by contradiction. Let $a \in \tilde{B}$ and suppose $a \notin B$. Since $a \in \tilde{B}$ there is some finite $C \subseteq B$ with $a \in \bar{C}$. From the first part of the theorem we know $\bar{C} \in \mathrm{Su}\, \mathfrak{A}_M$. In fact $\bar{C}$ is the subalgebra of $\mathfrak{A}_M$ generated by C, for if $D \in \mathrm{Su}\, \mathfrak{A}_M$ and $C \subseteq D$ then by the first part of our Theorem 1, $D = \bigcup_{G \text{ finite}, G \subseteq D} \bar{G}$ so $\bar{C} \subseteq D$. Thus $\bar{C}$ is the smallest subalgebra of $\mathfrak{A}_M$ which contains C. Now $a \in \bar{C}$ so $a = P(c_1, \ldots, c_n)$ for some operation P in $\mathfrak{A}_M$ and some sequence $c_1, \ldots, c_n$ from C. But $B \in \mathrm{Su}\, \mathfrak{A}_M$ so B is closed under P, and $C \subseteq B$. Thus $a = P(c_1, \ldots, c_n) \in B$. It follows that $B \supseteq \tilde{B}$ and thus $B \in \mathrm{Su}\, \mathfrak{A}_M \Rightarrow B = \bigcup_{D \text{ finite}, D \subseteq B} \bar{D}$. $\square$

§ 3. Characterization Theorems

For $L \subseteq 2^A$ let $\chi(L) = \{f \in A^B \mid B \in L, f = \mathrm{id} \restriction B\}$ be the set of *characteristic functions of* L. In general a function $f \in A^B$ will be called a *characteristic function* if $f = \mathrm{id} \restriction B$. Recall (§ 1) that when M is a set of partial functions, $\tilde{M}$ denotes the set of all partial endomorphisms of the algebra of all finitary operations which admit each $f \in M$ as a partial endomorphism. In what follows $\mathrm{id} \in S \subseteq A^A$, $L \subseteq 2^A$ and $M = S \cup \chi(L)$.

Theorem 2. $S = \mathrm{End}\, \mathfrak{A}$ *and* $L = \mathrm{Su}\, \mathfrak{A}$ *for some algebra* $\mathfrak{A}$ *iff* $\tilde{M}$ *contains no total functions other than* S *and* $\tilde{M}$ *contains no characteristic functions other than* $\chi(L)$.

Proof. Assume $S = \mathrm{End}\,\mathfrak{A}$ and $L = \mathrm{Su}\,\mathfrak{A}$ for some algebra $\mathfrak{A}$. Let $f \in A^A$, $f \notin S$. Since $S = \mathrm{End}\,\mathfrak{A}$ some operation of $\mathfrak{A}$ destroys f; this same operation must admit each map in $\chi(L)$ since $L = \mathrm{Su}\,\mathfrak{A}$ so the operation is among those of $\mathfrak{A}_M$. Thus the only total functions in $\tilde{M}$ are the members of S. Now let $g \in A^B$ with $g = \mathrm{id} \restriction B$, $B \notin L$, be a characteristic function. Then $B \notin \mathrm{Su}\,\mathfrak{A}$ so there is some operation P in $\mathfrak{A}$ and a finite sequence $b_1, \ldots, b_n \in B$ with $P(b_1, \ldots, b_n) \notin B$. This same operation P is again among the operations of $\mathfrak{A}_M$. But P does not admit g as a partial endomorphism since $gP(b_1, \ldots, b_n)$ is undefined. Thus the only characteristic functions in $\tilde{M}$ are members of $\chi(L)$. This proves one direction of the Theorem. To complete the proof let $M = S \cup \chi(L)$ and assume that $\tilde{M}$ contains no total functions other than S and no characteristic functions other than $\chi(L)$. Let $\mathfrak{A} = \mathfrak{A}_M$. Since $\tilde{M}$ is the set of all partial endomorphisms of $\mathfrak{A}_M$ we have $S = \mathrm{End}\,\mathfrak{A}$. Moreover if $B \in \mathrm{Su}\,\mathfrak{A}_M$ then $\mathrm{id} \restriction B$ is a partial endomorphism of $\mathfrak{A}_M$ so $\mathrm{id} \restriction B \in \tilde{M}$, thus $\mathrm{id} \restriction B \in \chi(L)$ so $B \in L$. Thus $L = \mathrm{Su}\,\mathfrak{A}$. □

We now combine Theorem 2 with Proposition 1 and Theorem 1 to obtain an equational condition for S and L to be jointly algebraic. The characterization theorem which follows says roughly that S must contain all functions which are unique solutions to systems of equations over $M = S \cup \chi(L)$ and that the support of every such system must belong to L. (For A finite the theorem says exactly that; the more general statement involves only additional "compactness" conditions which are "local" analogs of the above properties.)

Theorem 3. *$S = \mathrm{End}\,\mathfrak{A}$ and $L = \mathrm{Su}\,\mathfrak{A}$ for some algebra $\mathfrak{A}$ iff*

(1) $$\begin{bmatrix} \forall \text{ finite } D \subseteq A\ \exists \text{ system } \Sigma \text{ over } M \text{ with} \\ g \restriction D \text{ the unique solution to } \Sigma \text{ on } D \end{bmatrix} \Rightarrow g \in S.$$

and

(2) $$B = \bigcup_{D \text{ finite } D \subseteq B} \Big(\bigcap_{D \subseteq \mathrm{Spt}\,\Sigma,\, \Sigma \text{ over } M} \mathrm{spt}\,\Sigma \Big) \Rightarrow B \in L.$$

Proof. From Theorem 2 we know $S = \mathrm{End}\,\mathfrak{A}$ and $L = \mathrm{Su}\,\mathfrak{A}$ for some algebra $\mathfrak{A}$ iff

(i) $g \in A^A$ and $g \in \tilde{M} \Rightarrow g \in S$, and (ii) $id \restriction B \in \tilde{M} \Rightarrow B \in L$.

By Proposition 1 of § 1, (i) is equivalent to (1). Again by Proposition 1 of § 1, (ii) is equivalent to:

(ii′) $$\begin{bmatrix} B \in \mathrm{Su}\,\mathfrak{A}_M \text{ and } \forall \text{ finite } D \subseteq B\ \exists \text{ system } \Sigma \\ \text{over } M \text{ with } (\mathrm{id} \restriction B) \restriction D \text{ the unique solution to } \Sigma \text{ on } D \end{bmatrix} \Rightarrow B \in L.$$

Furthermore the system Σ: $\{X^{\Sigma} = \mathrm{id}\}$ has $\mathrm{id} \restriction D$ as a unique solution on each D, thus (ii′) is equivalent to: $[B \in \mathrm{Su}\, \mathfrak{A}_M \Rightarrow B \in L]$, and by Theorem 1 this is equivalent to:

$$(2) \qquad B = \bigcup_{D \text{ finite } D \subseteqq B} \Big(\bigcap_{D \subseteqq \mathrm{Spt}\, \Sigma,\, \Sigma \text{ over } M} \Big) \Rightarrow B \in L.$$

Thus $S = \mathrm{End}\, \mathfrak{A}$ and $L = \mathrm{Su}\, \mathfrak{A}$ iff (1) and (2) hold. □

For A finite Theorem 3 can be restated simply and completely as:

Corollary 1. *If A is finite and $\mathrm{id} \in S \subseteqq A^A$ and $L \subseteqq 2^A$ then $S = \mathrm{End}\, \mathfrak{A}$ and $L = \mathrm{Su}\, \mathfrak{A}$ for some algebra $\mathfrak{A}$ iff*

(1) *$g \in S$ whenever g is the unique solution to some system of equations with coefficients from $S \cup \chi(L)$, and*

(2) *$B \in L$ whenever B is the support of any system of equations with coefficients from $S \cup \chi(L)$.*

Bibliography

[1] M. Armbrust and J. Schmidt, Zum Cayleyschen Darstellungsatz, *Math. Annalen*, **154** (1964), 70—72.

[2] G. Birkhoff and O. Frink, Representations of lattices by sets, *Trans. Amer. Math. Soc.*, **64** (1948), 299—316.

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

[4] N. Sauer and M. G. Stone, The algebraic closure of a semigroup of functions (To appear).

[5] M. G. Stone, Subalgebra and automorphism structure in Universal Algebras, a concrete characterization, *Acta Sci. Math.*, **33** (1972), 45—48.

DEPT. MATHEMATICS AND STATISTICS
UNIVERSITY OF CALGARY
2920 24 AVE. N.W.
CALGARY, CANADA T2N 1N4

Acta Sci. Math., 38 (1977), 403—406

On normal subgroups of semigroups with identity element

MAGDA ROCKENBAUER[1]

In the theory of semigroups, normality of subsemigroups has been defined in several different ways. L. RÉDEI [3] has introduced this concept by the following two definitions;

D 1. The subsemigroup N of a semigroup S is called *left normal* if
(i) the partition $S=N\cup a_1N\cup a_2N\cup \ldots (a_1, a_2, \ldots \in S)$ is compatible, and
(ii) for each i and $n_1, n_2 \in N$, $a_in_1=a_in_2$ implies $n_1=n_2$.

Right normality is defined analogously.

D 2. The subsemigroup N of a semigroup S is called *normal,* if it is both right and left normal.

I. PEÁK [2] has modified these definitions by omitting condition (ii). Let us denote the modified definitions by D′ 1 and D′ 2, respectively.

The subgroup N of a semigroup S is called a *normal subgroup* of S if it is a normal subsemigroup in the sense of D 2 or D′ 2, respectively.

The following example shows that Theorem 1 of [2] is false.

Example. Let S be the semigroup of transformations of a set of cardinal 2 into itself.

The mistake in Peák's proof is in the part that (A) implies (B) where he used that N is right normal, too. Therefore, only the following modification of Peák's theorem holds true:

Theorem 1. *Let N be a subgroup of the semigroup S with identity element which contains the identity element of S. Then the following conditions are equivalent:*

Received June 13, 1974.

A) *N is normal in the sense of* D′ 2,

B) *for all $a \in S$, $aN = Na$ holds,*

C) *the set of the right cosets of N coincides with the set of the right cosets of N.*

The following Theorem 2 is from [2], but the proof for MN is not correct there.

Theorem 2. *Let S be a semigroup with identity element and let N and M be subgroups of S. If N and M are left normal in the sense of* D′ 1, *then MN is a left normal subgroup of S, and if S is also cancellative, then $M \cap N$ is a left normal subgroup of S in the sense of* D′ 1, *too.*

Theorem 2 can be generalized as follows:

Theorem 3. *Let S be a semigroup and N and M subsemigroups of S containing an identity element. If N and M are left normal in the sense of* D′ 1, *then MN is a left normal subsemigroup of S in the sense of* D′ 1, *and if S is also left cancellative and $M \cap N$ is non-empty, then $M \cap N$ is a left normal subsemigroup of S in the sense of* D′ 1, *too.*

Proof. It is well known that $M \cap N$ is a subsemigroup. If M and N are subgroups then $N \cap M$ is a subgroup. If

$$c \in a(M \cap N)b(M \cap N),$$

then

$$c \in (abM) \cap (abN),$$

and thus there exist an m in M and an n in N such that

$$c = abm = abn.$$

If S has an identity element, then, since N and M are left normal, N and M contain the identity element of S. Since S is left cancellative the last equation implies

$$c \in ab(M \cap N).$$

Let e be the identity element of N and let f be the identity element of M. Since M and N are left normal in the sense of D′ 1, $ef = e$ and $fe = f$ and MN with identity element f is a subsemigroup of S. fN is a left normal subsemigroup of MN in the sense of D′ 1.

If M and N are subgroups of S then fN and MN/fN are groups, therefore MN is a group.

MN is left normal, because if $c \in (aMN)(bMN)$ then $c = amnbm'n'$ holds for some $m, m' \in M$ and $n, n' \in N$. Thus

$$c \in (amN)(bm'N).$$

Since $amebm'e = amfebm'e = ambm'e$,

$$(amN)(bm'N) = ambm'N$$

holds. Since M is left normal too, we have

$$ambm'N = abm''N \subseteqq abMN$$

for $m'' \in M$, and, consequently

$$c \in abMN.$$

If $e=f$ and M and N are subgroups of S then $MN=NM$. It follows that the first assertion of Theorem 2 is true.

Remark 1. If we replace D′ 1 by D 1 in the second assertions of Theorems 2 and 3 then we can omit the condition that S be left cancellative. We introduce an equivalence relation (see Lyapin [1]):

Let S be a semigroup with identity element and N be a subgroup of S. We say that r is *ϱ_N-equivalent* to s, in symbols $r\varrho_N s$, if there exist elements n, m in N such that $rn=ms$.

The following assertion is a modification of an assertion of Peák [2], p. 349.

The partition corresponding to the equivalence relation ϱ_N coincides with the left (right) cosets of N if and only if N is left (right) normal in the sense of D′ 1.

Proof. Suppose that N is left normal in the sense of D′ 1. Any two elements of a left coset of N are ϱ_N-equivalent because $b \in aN$ implies the existence of an element n in N such that

$$an = b = eb, \quad \text{whence} \quad a\varrho_N b.$$

On the other hand, any element c that is ϱ_N-equivalent to a belongs to the left coset aN, because the partition

$$S = N \cup a_1 N \cup a_2 N \cup \ldots$$

is compatible.

Conversely, suppose that the partition corresponding to ϱ_N coincides with the left cosets of N. If $c \in N(aN)$ then $c\varrho_N a$. It follows that $c \in aN$. Since $e \in N$, we have $(bN)(aN)=baN$, as we wished to prove.

Peák has also made the following assertion:

Let N run over the set of all subgroups of a semigroup S with identity element, which are left normal in the sense of D′ 1 *and contain the identity element of S. Then either each or none of the factor semigroups S/N is a group.*

Proof. If N is left normal in the sense of $D'1$ and S/N is a group then S is a group.

References

[1] Е. С. Ляпин, Ядра гомоморфизмов ассоциативных систем, *Матем. сборник*, **20** (62) (1947), 497—514.

[2] I. PEÁK, Über gewisse spezielle kompatible Klasseneinteilungen von Halbgruppen, *Acta Sci. Math.*, **21** (1960), 346—349.

[3] L. RÉDEI, Die Verallgemeinerung der Schreierschen Erweiterungstheorie, *Acta Sci. Math.*, **14** (1952), 252—253.

LÁGYMÁNYOSI ÚT 14/A
H-1111 BUDAPEST

Bibliographie

Tom M. Apostol, Modular Functions and Dirichlet Series in Number Theory (Graduate Texts in Mathematics), X+198 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1976.

This textbook is the continuation of a book of the same author, which appeared in the Springer-Verlag series Undergraduate Texts in Mathematics under the title "Introduction to Analytic Number Theory". This second volume presupposes a background in number theory comparable to that provided in the first volume, together with a knowledge of the basic concepts of complex analysis.

The first three chapters provide an introduction to the theory of elliptic modular functions, which play a role in additive number theory analogous to that played by Dirichlet series in multiplicative number theory. Applications to the partition function are given in Ch. 5, while Chs. 4 and 6 contain, among others, Lehner's congruences for the Fourier coefficients of the modular function $j(\tau)$, and Hecke's theory of entire forms with multiplicative Fourier coefficients. Ch. 7 deals with the problem of approximating real numbers by rational numbers, including Kronecker's theorem with applications. The last chapter gives an account of Bohr's theory of equivalence of general Dirichlet series.

There are exercises at the end of each chapter. The book will certainly help the nonspecialist become acquainted with a fascinating part of mathematics and, at the same time, will provide an up-to-date background to every specialist in the field.

F. Móricz (Szeged)

J. Bergh and J. Löfström, Interpolation Spaces (An Introduction) (Grundlehren der mathematischen Wissenschaften, 223), X+207 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1976.

In recent years there has emerged a new field of study in functional analysis: the theory of interpolation spaces. Interpolation theory has been applied to other branches of analysis (e.g. partial differential equations, approximation theory, etc.), but it also has considerable interest in itself. This is the first attempt, as far as we know, to treat interpolation theory fairly comprehensively in book form.

The reader is supposed to be conversant with the elements of real (several variables) and complex (one variable) analysis, of Fourier series, and of functional analysis. Beyond elementary level the authors tried to supply proofs of the statements in the main text. Their general reference for elementary results is the first volume of the widely-known monograph of Dunford—Schwartz "Linear operators".

Ch. 1 presents the classical interpolation theorems of M. Riesz, with Thorin's proof, and of Marcinkiewicz, which provided the main impetus for the study of interpolation. The basic concepts are introduced in Ch. 2, where a few general results are discussed, e.g. the Aronszajn—Gagliardo theorem.

The authors treat two essentially different interpolation methods: the real method and the complex method. These two methods are modelled after the proofs of the Marcinkiewicz theorem and the Riesz—Thorin theorem, resp.. The real method is elaborated following Peetre in Ch. 3, the complex method following Calderon in Ch. 4.

The further three chapters contain applications of the above general methods. Ch. 5: Interpolation of L_p-Spaces, Ch. 6: Interpolation of Sobolev and Besov Spaces, Ch. 7: Applications to Approximation Theory.

In each chapter the penultimate section contains exercises, which extend and complement the results of the previous sections. Moreover, many important results and most of the applications can be found only as exercises. The last section of each chapter is devoted to notes and comments. These include historical sketches, various generalizations, related questions and references without aiming at completeness. There is a bibliography consisting of about 200 items.

The treatise provides a rich and up-to-date account of this fast-growing and important field, and it is warmly recommended to everyone who wants to learn, or do research in, interpolation theory.

F. Móricz (Szeged)

M. Braun, Differential equations and their applications (An introduction to applied mathematics, Applied Mathematical Sciences, Vol. 15) XIV+718 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1975.

Two main motives of the present-day development of the theory of differential equations can be emphasized. More and more interesting problems arise in the theory as in an independent branch of pure mathematics. On the other hand, in applications the number of processes that can be modelled mathematically by differential equations are constantly increasing. Accordingly, in the last years a great number of noticeable books have appeared emphasizing one or the other of the motives mentioned above. This book calls for the interest of both users of mathematics and mathematicians. The author writes in the preface: "the material is presented in a manner which is rigorous enough for our mathematics and applied mathematics majors, but yet intuitive and practical enough for our engineering, biology, economics, physics and geology majors." Numerous examples are given of how researchers have used differential equations to solve real life problems. Especially interesting are: the Van Meegeren art forgeries, population growth of various species, a model for the detection of diabetes, L. F. Richardson's mathematical theory of war, why the percentage of sharks caught in the Mediterranean Sea rose dramatically during World War I, the Tacoma Bridge disaster, and a model for the spread of epidemics.

There are many original interesting exercises at the end of each section, and complete Fortran and APL programs are given for every computer exercise in the text.

The titles of chapters are: First order differential equations, Second order differential equations, Systems of first order equations, Qualitative theory of differential equations, Separation of variables and Fourier series, Appendices.

The well-written self-contained book can be understood by anyone having attended a two-semester course in Calculus.

L. Hatvani—L. Pintér (Szeged)

Colloquio Internazionale sulle Teorie Combinatorie. I—II (Atti dei Convegni Lincei 17), 518+526 pages, Accademia Nazionale dei Lincei, Roma 1976.

This combinatorial colloquium, organized by the Accademia Nazionale dei Lincei with the collaboration of the American Mathematical Society, took place in Rome, September 3—15, 1973. The conference was dedicated to Professor Beniamino Segre on the occasion of his 70th birthday, and was chaired by Professor Segre. This is also reflected by the fact that the majority of papers delivered at the colloquium and reproduced in the volumes deal with the exciting and fast-developing field of finite geometries, block designs and matroids, to which field Professor Segre's contribution has been most important.

It would be impossible to list all of the 77 papers contained in the two extensive volumes of the Proceedings. First, there are many very useful survey papers: Hall writes about the "Construction of Combinatorial Designs", Richard Rado about "Partition Calculus", Beniamino Segre surveys "Incidence Structures and Galois Geometries", Buekenhout "Characterizations of Semi Quadrics", Bachman "Hjelmslev Groups", Seidel "2-graphs", Turán "Combinatorics, Partitions, Group Theory", just to mention some. Erdős has, as usual, a paper on "Problems and Results in Combinatorial Analysis". Besides, there are many papers which contain very significant new results and ideas, some of which have been known, and whose publication has been looked forward to, since the colloquium.

L. Lovász (Szeged)

E. T. Copson, Partial differential equations, VII+280 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1975.

A good survey on the theory of partial differential equations of the first order and of linear partial differential equations of the second order, using the methods of classical analysis. In spite of the advent of computers and the recent applications of the methods of functional analysis to the theory of partial differential equations, the classical theory retains its relevance in several important respects.

The book is well-organized. At the end of each chapter a number of interesting exercises help understanding. The titles of chapters show the treated topics of the theory: Partial differential equations of the first order, Characteristics of equations of the second order, Boundary value and initial value problems, Equations of hyperbolic type, Riemann's method, The equations of wave motions, Marcel Riesz' method, Potential theory in the plane, Subharmonic functions and the problems of Dirichlet, Equations of elliptic type in the space, The equation of heat.

This text-book will be useful for lecturers and students of mathematics or theoretical physics.

J. Terjéki (Szeged)

R. E. Edwards and G. I. Gaudry, Littlewood-Paley and Multiplier Theory (Ergebnisse der Mathematik und ihrer Grenzgebiete, 90), IX+212 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1977.

The classical Littlewood—Paley theorem asserts that to each p in $(1, \infty)$ there corresponds a pair (A_p, B_p) of positive constants such that $A_p\|f\|_p \leqq \|(\sum_{j\in\mathbf{Z}} |S_j f|^2)^{1/2}\|_p \leqq B_p\|f\|_p$ for every f in L^p, where $S_j f$ is the jth dyadic partial sum of the Fourier series $\sum \hat{f}(n)e^{inx}$ of f, defined by (1) $\sum_{2^{j-1}\leqq n<2^j} \hat{f}(n)e^{inx}$, or $\hat{f}(0)$, or $\sum_{-2^{|j|}<n\leqq -2^{|j|-1}} \hat{f}(n)e^{inx}$ according as $j>0$, $j=0$, or $j<0$.

In other words, we can say that the L^p norm of a function f can be computed, up to equivalence, by breaking up the Fourier series of f into its dyadic partial sums, putting them together in an l^2 fashion, and calculating the L^p norm of the resulting function.

The Littlewood—Paley theorem implies, among others, the following analogue of the Riesz—Fischer theorem: Suppose $p\in(1, \infty)$. A series $\sum_{n\in\mathbf{Z}} c_n e^{inx}$ is the Fourier series of a function, say f, in L^p iff $\|(\sum_{j\in\mathbf{Z}} |s_j|^2)^{1/2}\|_p < \infty$, where s_j denotes the trigonometric polynomial obtained from formula (1) by replacing $\hat{f}(n)$ by c_n. Moreover, the series $\sum_{j\in\mathbf{Z}} s_j$ converges unconditionally in L^p to f.

These results make the Littlewood—Paley theorem one of the fundamental results in L^p harmonic analysis.

The treatment proceeds along two main lines, the first relating to singular integrals on locally compact groups (Chs. 2 and 3), and the second to martingales (Ch. 5). Both (classical and modern) versions of the Littlewood—Paley theorem are dealt with for the classical groups $\mathbf{R}^n$, $\mathbf{Z}^n$, $\mathbf{T}^n$ (Chs. 7 and 8) and for certain classes of discontinuous groups (Ch. 4); $\mathbf{R}$ denoting the set of real numbers, $\mathbf{Z}$ the set of integers, and $\mathbf{T}$ the circle group.

The Littlewood—Paley theorem is then applied to Fourier multiplier theory, for instance to obtain the famous theorems of M. Riesz, Marcinkiewicz and Stečkin (Ch. 6); and to lacunary sets (Ch. 9).

For the reader's convenience there are four appendices containing a number of auxiliary topics at the end of the book. Historical Notes, References, Terminology, Index of Notation, and Index of Authors and Subjects complete the book.

The presentation is self-contained and unified. The book is intended primarily for use by graduate students and mathematicians who wish to begin studies in these areas, poorly served by existing books. This well-written book fills in the gap in the literature and satisfies all needs of a beginner as well as of a "worker in the field".

F. Móricz (Szeged)

Carl Faith, Algebra. II, Ring Theory (Grundlehren der mathematischen Wissenschaften, Band 191), XVIII+302 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1976.

This book is the second volume of the work on rings, modules and categories, Volume I (Parts I—IV) of which was published as Band 190 of the same series (and reviewed in these *Acta*, **38** (1976), 209). The present volume II is devoted entirely to ring theory. To a large extent this volume, except Chapters 18 and 26, illustrates the power of homological methods in ring theory. In contrast to what was announced in Volume I, Part VI on commutative rings, hereditary rings, separable algebras, and the Brauer group is not included in Volume II, thus Part V (Chapters 17—26) comprises all of Volume II.

This is the largest part of the book, too rich in content to list here all the topics covered in it. Summarizing in a few sentences, Chapter 17 is on modules of finite Jordan—Hölder length, while Chapter 18 deals with the Jacobson radical of a ring. In Chapter 19 quasiinjective modules are studied and, among others, the Chevalley—Jacobson density theorem is proved. Chapter 20 is devoted to the direct decompositions of rings and modules. Azuyama diagrams are discussed in Chapter 21 while the aim of Chapter 22 is to study the projective covers of modules and perfect rings. In Chapters 23 and 24 Morita's duality theory and some applications are presented; in particular, quasi-Frobenius rings are also discussed. Chapter 25 is on serial and Σ-cyclic rings and, finally, Chapter 26 is concerned with semiprimitive and semiprime rings, the main result being

Amitsur's theorem on the semiprimitivity of group algebras over transcendental fields of characteristic zero.

Each chapter ends with a list of related results aiming to help those wishing to specialize in that topic. The book is concluded with a rich bibliography up to 1975.

Á. Szendrei (Szeged)

Robert Fortet, Elements of Probability Theory, XIX+524 pages, Gordon and Breach Science Publishers, London—New York—Paris, 1977.

This book is the translation of the French original "Eléments de la Théorie des Probabilités, Vol. I", published by the Centre National de la Recherche Scientifique in 1960. A few errors of the French original have been corrected, but otherwise the text remained unaltered. The book is introductory, written in the best tradition of French scholarship. Each newly introduced concept is carefully motivated from various aspects. The author is not shy to get into philosophical problems, and discussions of questions from physics, mechanics, genetics, etc., to help the beginner get a real feeling of the subject. He recommends his book to non-mathematical research workers such as physicists, engineers, biologists and operation research workers "to provide users with an exposition of the fundamentals of probability theory, at a level of mathematical sophistications which would not repel the non-specialist reader". Chapter headings: I. Combinatorial analysis and its application to classical and quantum statistics and to the chromosome theory of heredity; II. The concept of probability, Measures or mass distributions, Hilbert spaces, Random elements and probability laws; III. Distribution functions; IV. Random variables, axiom of conditional probability; V. n-dimensional random vectors and variables; VI. Addition of independent random variables; Stochastic convergence, laws of large numbers, ergodic theorems; Convergence to a normal law, convergence to a Poisson law; Generalizations.

Gordon and Breach is to be praised for having made available this valuable book in English.

Sándor Csörgő (Szeged)

Э. Фрид—И. Пастор—И. Рейман—П. Ревес—И. Ружа, Малая математическая энциклопедия, 693 стр. Изд. АН Венгрии, Будапешт, 1976.

Эта книга является энциклопедией в менее привычном смысле слова; именно, она является обзором высшей математики, напоминающим превосходную книгу Куранта и Роббинса «Что такое математика?» В ней представлены важнейшие разделы математики: алгебра, геометрия, математический анализ, теория множеств, теория вероятностей, математическая статистика и математическая логика. Такая классификация, разумеется, отражает и личный интерес авторов, что, в свою очередь, отражает в некоторой мере и главные напривления математических исследований в Венгрии.

Книга может быть использована для первого ознакомления с различными математическими понятиями (например, группы занимают в ней 7 страниц, числовые ряды — 13, начертательная геометрия — 9, а исчисление предикатов — 11). Ее можно использовать также в качестве справочника, поскольку в ней можно найти много простых, важных теорем (без доказательства), как, например, теорему Кэли, критерий Коши, теорему Полке и теорему Гёделя о полноте. Книга в основном написана живо, местами увлекательно. Следует, однако, предупреждать читателя, что переводчики часто упожребляют нестандартную терминологию, особенно в разделах теории чисел и теории множеств.

B. Csákány (Szeged)

Morris W. Hirsch, Differential Topology (Graduate Texts in Mathematics, 33), 221 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1976.

The aim of this book is to give an introduction to the problems and results of differential topology, that is of the topology of differentiable manifolds. Although topological questions on manifolds occur in differential geometry and global analysis as well, there is always some extra structure present, e.g. Riemannian metric or a differential equation on the manifold. In differential topology the manifold itself is studied, the extra structures are used only as tools. Some typical questions: Can a given manifold be embedded in another one? If two manifolds are homeomorphic, are they necessarily diffeomorphic? Which manifolds are boundaries of compact manifolds? Do the topological invariants of a manifold have any special properties? Does every manifold admit a non-trivial action of some cyclic group? This book presents some answers to these questions.

The first three chapters are fundamental for the understanding of the book. Definitions are introduced and the basic properties of manifolds, the approximation theorems for the maps of manifolds and the unifying idea in differential topology: the transversality are treated. In Chapter 4 the elementary theory of vector bundles is developed, including the classification theorem: isomorphism classes of vector bundles over the manifold M correspond naturally to homotopy classes of maps from M into a certain Grassmann manifold. Chapter 5 is devoted to the study of the theory of degrees of maps. In this way some results of classical algebraic topology are derived. In Chapter 6 an introduction to the Morse theory is presented. Chapter 7 contains the elementary part of one of the most elegant theories in differential topology: René Thom's theory of cobordisms. (Two manifolds are cobordant if together they form the boundary of a compact manifold.) In chapter 8 the isotopy of embeddings of manifolds is investigated. Chapter 9 deals with the classification of surfaces.

Each chapter contains many interesting exercises and historical remarks.

The book is a rich, up-to-date account of differential topology. It is also very well-written. I can warmly recommend it to everyone interested in the theory of manifolds.

P. T. Nagy (Szeged)

Wu Yi Hsiang, Cohomology Theory of Topological Transformation Groups (Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 85), X+164 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1975.

The pioneering results of P. A. Smith for prime periodic maps on homology spheres, and of L. E. J. Brouwer on periodic transformations suggest a general direction of studying topological transformation groups in the framework of algebraic topology. This book is an excellent summary of these different generalizations. Chapters I and II contain general material on compact Lie groups, G-spaces and on structural and classification theory of compact Lie groups and their representations.

Let G be a compact Lie group and let X be a given G-space. Then the equivariant cohomology $H_G^*(H)$ of the G-space X is the ordinary cohomology of the total space X_G of the universal bundle $X \to X_G \to B_G$, with the given G-space as the typical fibre. In Chapter III some fundamental properties and theorems (such as the localization theorem of Borel-Atiyah-Segal) of this equivariant cohomology theory of A. Borel are formulated and proved.

In Chapter IV the relationship between the geometric structure of a given G-space X and the algebraic structure of its equivariant cohomology $H_G^*(H)$ is investigated. The reader obtains in this Chapter an answer for the following problems: How much of the cohomology structure of the fixed point set, $H_G^*(F)$, is determined by the equivariant cohomology $H_G^*(X)$? Is it possible to give a criterion for the existence of fixed points purely in terms of the equivariant cohomology $H_G^*(H)$? Suppose $F(G, X)=\emptyset$. How to determine the set of maximal isotropy subgroups from the algebraic structure of $H_G^*(X)$?

The structural splitting theorem for linear tori actions can be generalized to various structural splitting theorems of the equivariant cohomology, and combining the structural splitting theorems with the maximal tori theorem, a geometric weight system for topological transformation groups can be defined.

Such a program is carried out in Chapters IV, VI and for the special cases of acyclic manifolds and cohomology spheres in Chapter V. In Chapter VII the cohomology method is applied to study transformation groups on compact homogeneous spaces.

This book comprises a very large material. To read it certain knowledge on differential manifolds, Lie groups and cohomology theory is necessary.

Z. Szabó (Szeged)

John G. Kemeny—J. Laurie Snell—Anthony W. Knapp, Denumerable Markov Chains (Graduate Texts in Mathematics, 40), XII+484 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1976.

After the second edition of "Finite Markov Chains" by the first two authors in the series "Undergraduate Texts in Mathematics", Springer-Verlag has provided us with the second edition of another success-book, which treats discrete parameter Markov chains having countable state space and stationary transition probabilities, with special emphasis on the context of potential theory. The original was published by Van Nostrand, Princeton, N. J., 1966, in the University Series in Higher Mathematics, and has been reviewed by J. L. Doob in detail (MR 34 (1967)#6858). Doob wrote that "the potential-theoretic point of view should have a strong influence on future research", and his prediction has already been proved to be true. The book has served as a source of inspiration in the past ten years, occurring often in the reference list of research papers in the field of probabilistic potential and boundary theory. An erroneus theorem is corrected, but aside from this change, the text of the first eleven chapters of the first edition is left intact. This new edition contains a new twelfth chapter on Markov random fields, written by David Griffeath (pp. 425—428). In addition to this, it also contains a new section of Additional Notes (pp. 465—470), covering some of the developments of the past ten years, which is accompanied by a section of Additional References, listing 58 items.

Sándor Csörgő (Szeged)

A. A. Kirillov, Elements of the Theory of Representations (Grundlehren der mathematischen Wissenschaften, 220), XI+315 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1976.

This is a translation, by Edwin Hewitt, of the original Russian text. The translation is faithful except for some corrections supplied by Professor Kirillov himself. The bibliography, for an obvious reason, has been considerably modified from the original.

The material of the book grew out of courses given and seminars directed by the author at Moscow State University. The first part of the book (§§ 1—6) is not directly related to representations, it contains the facts needed from other parts of mathematics. Those topics that are not included in elementary university courses are treated here in detail. A reader familiar with this material may skip the first part of the book. The second part (§§ 7—15) is devoted to the principal concepts and methods of the theory of representations. In the third part (§§ 16—19) the general ideas of the second one are illustrated by concrete examples.

The book includes a large number of exercises playing an essential rôle in the text proper. "A majority of the proofs are given in the form of a cycle of mutually connected problems". Most problems are supplied with remarks to help the reader solve them.

Little attention has been paid to finite dimensional representations of semisimple Lie groups and Lie algebras as there exist good expositions of this subject (in both the Russian and English literature, many books being mutually translated). For the same reason the applications of the theory of group representations in the theory of special functions as well as in mathematical physics have been completely ignored in this book. However, a large space is devoted to the method of orbits, which has not yet been included in any textbook. The author hopes that some of the readers of this book will contribute to the development of the rapidly growing and important theory of orbits.

The paragraph headings are as follows. § 1. Sets, categories, topology, § 2. Groups and homogeneous spaces, § 3. Ring and modules, § 4. Elements of functional analysis, § 5. Analysis on manifolds, § 6. Lie groups and Lie algebras, § 7. Representations of groups, § 8. Decomposition of representation, § 9. Invariant integration, § 10. Group algebras, § 11. Characters, § 12. Fourier transforms and duality, § 13. Induced representations, § 14. Projective representations, § 15. The method of orbits, § 16. Finite groups, § 17. Compact groups, § 18. Lie groups and Lie algebras, § 19. Examples of wild Lie groups.

"A short historical sketch and a guide to the literature", and a "Bibliography" complete the book.

This masterly written and translated book may be mainly recommended to those wanting to begin the study of the vast field of representations.

József Szűcs (Szeged)

S. Lefschetz, Applications of Algebraic Topology (Applied Mathematical Sciences, 16), viii+189 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1975.

Solomon Lefschetz's book published posthumously consists of two independent monographs.

In Part I the author first gives a short résumé of the algebraic topology up to dimension 2 (Chapters I—V). Except for the theorem of Jordan—Schoenflies all results are presented here together with proofs. In Chapter VI Kirchoff's Laws are formulated in terms of the co-theory and a system of differential equations is deduced from them. In Chapters VII and VIII the elements of the theory of 2-dimensional complexes and surfaces are presented. They are applied in Chapter IX to the problem of planar graphs and dual networks. Maclane's and Kuratowski's characterization theorems are proved.

Part II is devoted to the demonstration of the connection between the Picard—Lefschetz theory and the theory of Feynman integrals. After a short algebraic and topological résumé with almost no proofs (Chapter I) the author treats a special phase of Picard's program: he investigates the behaviour of the abelian integral of a rational function on a complex irreducible algebraic surface near an isolated singularity (Chapter II). Chapter III deals with the extension of this theory to higher varieties.

Feynman's problem, which is treated in Chapters IV and V can be outlined as follows:

Set $x=\{x_1, \dots, x_n\}$, $y=\{y_1, \dots, y_n\}$, where x_k are real or complex coordinates and y_k are real or complex parameters. Let $Q_h(x, y)$ denote real quadratic polynomials. The Feynman problem consists of the study of the analytical character of

$$I(y)=\int_{\Gamma}\frac{dx_1\dots dx_n}{\prod\limits_h Q_h(x,y)},$$

as function of y, where Γ is the whole admissible part of the x-space. In Part II the author explains only the crucial points of proofs for a reader well versed in classical function theory.

A. Krámli (Budapest)

Edwin E. Moise, Geometric topology in dimensions 2 and 3 (Graduate Texts in Mathematics), VII+262 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1977.

The manuscript of this book was used in 1975—76 at Texas University and earlier at the University of Wisconsin in seminars conducted by R. H. Bing. It is intended to be a textbook; the text is divided into 37 sections and contains all of the most important classical and new results concerning this branch of topology. The traditional material of plane topology has been reformulated in such a way that, bringing three-dimensional ideas in sharper focus, it serves as an introduction to the methods to be used in dimension 3. The proofs of the triangulation theorem and the *Hauptvermutung* are largely new. So is the proof of Schoenflies' theorem.

At the end of each section there are sets of problems, which are composed in an unusual way. Most of the problems state true theorems, extending or elucidating the preceding section of the text. But in a large number of them false propositions are also stated as if they were true. Here it is the reader's job to discover that they are false and find counterexamples.

The book is highly recommended to anyone interested in topology and mature enough to understand abstract mathematical thinking.

L. Gehér (Szeged)

R. Narasimhan, Analysis on Real and Complex Manifolds (Advanced Studies in Pure Mathematics), X+246 pages, Masson & Cie, Paris, and North-Holland Publishing Company, Amsterdam, 1968.

This book contains the basic material for the study of differential equations on manifolds. It has three chapters.

In Chapter 1 some theorems on differentiable functions in $\mathbf{R}^n$ are proved such as the implicit function theorem, Sard's theorem and Whitneys' approximation theorem. Chapter 2 is an excellent introduction to the study of real and complex manifolds. This chapter contains, among others, the theorem of Frobenius, the lemmata of Poincaré and Grothendieck, the imbedding theorem of Whitney and Thom's transversality theorem. In chapter 3 properties of linear elliptic differential operators are formulated. This chapter deals with Peetre's and Hörmander's characterizations of linear differential operators, the inequalities of Gårding and of Friedrichs on elliptic operators, and finally with the approximation theorem of Malgrange-Lax. The Runge theorem on open Riemann surfaces is also proved.

The book is written in a very elegant style. It is an excellent graduate textbook.

Z. Szabó (Szeged)

M. S. Raghunathan, Discrete Subgroups of Lie Groups (Ergebnisse der Mathematik und ihre Grenzgebiete, 68), 226 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1972.

The theory of discrete subgroups of Lie groups, which originated among others from geometrical crystallography, has become a separate discipline as a consequence of the work of A. D. Malcev, A. Selberg, A Weil, A. Borel, L. Auslander and others in the last 20 years.

This book is a fundamental monography on this subject. Its aim is to present a detailed account of recent work on the theory of discrete subgroups of a Lie group from the geometric point of view.

Chapters I—V contain a fairly complete study of lattices in nilpotent, solvable and semisimple Lie groups, where a "lattice" in a locally compact group G means a discrete subgroup H in G such that the homogeneous space G/H carries a finite G-invariant measure. Chapter VI presents some general theorems on finitely generated subgroups of a Lie group. In Chapter VII results on the cohomology of solv-manifolds and compact symmetric spaces are treated. Chapter VIII plays a central role. Among other results it is proved here that, at least up to a point, the study of lattices in general Lie groups can be split into the study of those in solvable and semisimple groups separately. Chapters IX—XIV are devoted to further interesting results on discrete subgroups (results of Kazdan—Margolis, arithmetic groups, existence of arithmetic lattices, etc.).

It is assumed that the reader has considerable familiarity with Lie groups and algebraic groups. Most of the results used frequently in the book are summarized in the "Preliminaries"; this chapter may be useful as a reference as well.

P. T. Nagy (Szeged)

Alfréd Rényi, Selected Papers, I—III, edited by Pál Turán, co-editors: P. Bártfai, I. Csiszár, J. Fritz, G. Halász, Gy. Katona, P. Révész, D. Szász, E. Szemerédi, I. Vincze; technical editor: G. Székely, 628+646+667 pages, Akadémiai Kiadó, Budapest, 1976.

Alfréd Rényi (1921—1970) was one of the most outstanding mathematicians of the new Hungarian generation. He inventively commanded a wide area of interest. He published about 350 papers and books.

The Selected Papers includes 156 articles. The papers published in English, German, or French are kept in their original form, and those published in Hungarian, Russian or Chinese are here translated into English. The papers follow in chronological order. In the first volume 52 works, written between 1948 and 1956, in the second 48 works, written between 1956 and 1961, and in the third 56 works, written between 1962 and 1970, are reproduced. Each volume contains a general introduction by the editor, a biography of Alfréd Rényi and a list of his scientific works. Besides, the first volume contains a photograph of Alfréd Rényi.

The papers are selected very carefully. Thanks to this the present selection offers a good survey of Alfréd Rényi's main fields of interest and shows in what areas Alfréd Rényi contributed most significantly to the development of mathematics. The most important papers in the first volume are those dealing with the generalization and application of the Linnik large sample method, those concerning the Poisson process and the generalization of Kolmogorov's inequality as well as those devoted to the foundations of probability theory and to the applications of probability theory in chemistry and biology. In the second and third volumes the articles about the applications of probability theory in graph theory, number theory, group theory, and information theory are the most valuable.

The selection is greatly enhanced by the remarks of the editors after each paper. These professional remarks inform the reader about the international influence of the results and problems included in each paper: who developed and in what direction the problems in question and in what sense the problems were solved. On the basis of the papers followed by the remarks the reader obtains a picture of a widely ranging mathematical cooperation in the centre of which stood Alfréd Rényi. His untimely death was a great loss to mathematics.

Károly Tandori (Szeged)

T. G. Room—P. B. Kirkpatrick, Miniquaternion geometry. An introduction to the study of projective planes, Cambridge Tracts in Math. 60, viii+176 pages, Cambridge University Press, Cambridge, 1971.

This book is of a rather unusual nature. Its primary aim is the study of four concrete mathematical objects: the four projective planes (known at present) of order 9. It is the authors' aim that the study of these planes should also serve as an introduction to the study of projective planes in general (9 being the smallest order for which non-desarguesian projective planes exist).

The first chapter describes the relevant algebraic objects: the finite field of 9 elements: $GF(9)$ and the other near-field of order 9: Q, which is called by the authors the miniquaternion system because of its common features with the skew-field of ordinary quaternions. The automorphism group of Q is also determined, and solution of equations in Q is discussed.

Chapter 2 gives a rapid introduction to projective planes (including Bruck's theorem on the possible orders of subplanes) and collineations. The standard procedure of constructing a plane $\Pi(K)$ from a field K is then desribed and projectivities, correlations and conics of $\Pi(K)$ are discussed. In Chapter 3 these investigations are carried through in much greater detail in the case of the plane $\Pi(GF(9))$, after a brief study of $\Pi(GF(3))$.

Chapters 4 and 5 are devoted to the study of the 3 non-desarguesian planes of order 9, discovered by Veblen and Wedderburn in 1907. In Chapter 4 the translation plane Ω is defined with the aid of Q and an appropriate coordinatization. The collineation group of Ω is completely determined. Ω has subplanes of order 2 and 3; both types of subplanes are studied in detail. Then the Rodriguez oval in Ω is introduced. Finally it is proved that the dual plane of Ω: Ω^D is a plane of order 9 which is not isomorphic to Ω.

In Chapter 5 the plane Ψ is defined using Q in a different coordinatization. (Ψ is the smallest Hughes plane.) The collineation group of Ψ is determined. It is proved that Ψ is self-dual and polarities of Ψ are studied. The subplanes of Ψ are investigated and another definition of Ψ (essentially the original definition of Veblen and Wedderburn) is also given.

Throughout the book great emphasis is laid on concreteness. Points, lines etc. are given names and there are lots of tables enumerating all the objects of a certain type in the plane under study. The numerous exercises also aim usually at checking certain concrete statements rather than prove theorems.

This book seems to be particularly useful to people who want to see (or want to show their students) what certain concepts, constructions, theorems concerning projective planes, usually presented in an abstract setting, actually mean. This type of book is particularly welcome since in the modern literature there is a tendency toward the most general possible formulation without giving examples illuminating the motivations behind the various investigations. At the same time the book is a really good starting point for a further study of projective planes.

József Pelikán (Budapest)

G. B. Seligman, Modular Lie Algebras (Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 40), IX+165 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1967.

The theory of classical Lie algebras over a base field of characteristic zero was developed by Lie, Killing, Elie Cartan, and Weyl. The first papers studying the structure of Lie algebras over arbitrary fields were those of Jacobson (Rational methods in the theory of Lie algebras) and Landherr (Über einfache Liesche Ringe). On this stage of generalization it was already clear that it is not a too straight-forward problem to work out new methods to establish the analogues of characteristic-zero theorems.

A modular Lie algebra (such are the algebras to which the title of this book refers) is a Lie algebra over a field of positive characteristic. The study of these structures is now more than forty years old, and this book is the first general treatment on this active field.

In Chapter I (Fundamentals) generalities, such as restricted Lie-algebras, Iwasawa theorem, Cartan subalgebras, are formulated and proved. Chapter II (Classical semi-simple Lie Algebras) contains the Cartan decompositions of algebras with non-degenerate trace form, and the classification of the classical algebras. In Chapter III (Automorphisms of the Classical Algebras) the automorphism groups of the classical algebras are determined, and Chapter IV (Forms of the Classical Lie Algebras) is motivated by the problem of determining all Lie algebras with non-singular Killing form over an arbitrary field of characteristic $\neq 2,3$. Chapter V (Comparison of the Modular and Non-modular Cases) deals with a number of analogues of fundamental classical theorems. Chapter VI (Related Topics) is an indication of some ways in which Lie algebras, especially those of prime characteristic, have arisen in other areas of mathematics.

The book is clearly written and could serve as an excellent textbook for a graduate course in Lie algebras.

Z. Szabó (Szeged)

I. M. Singer—S. A. Thorpe, Lecture Notes on Elementary Topology and Geometry (Undergraduate Texts in Mathematics), VIII+232 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1967.

This book presenst an introduction in modern topology and modern global differential geometry. The text consists of 8 chapters. The first two chapters give a short glimpse into point set topology. Chapter 3 treats homotopy, fundamental group, and covering spaces. In Chapter 4 the concept of simplicial complexes is introduced, geometry and barycentric subdivisions of simplicial complexes, the simplicial approximation theorem and fundamental group of simplicial complexes are treated. After some preliminaries in Chapter 5 concerning the theory of differentiable manifolds and differential forms, Chapter 6 deals with simplicial homology and the de Rham theorem. The two final chapters are devoted to studying intrinsic Riemannian geometry of surfaces and imbedded manifolds in R^n.

The book is highly recommanded to anybody interested in modern differential topology.

L. Gehér (Szeged)

Ernst Specker—Volker Strassen, Komplexität von Entscheidungsproblemen: ein Seminar (Lecture Notes in Computer Science, 43), 217 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1976.

The volume contains eleven lectures of a seminar on the complexity of logical and combinatorial decision problems held at the University of Zürich, in 1973—74. Besides some lectures giving algorithms to some nontrivial decision problems of rather general art, an overview is given of the main methods of finding lower bounds on the complexity of decision problems and a systematic exposition of the underlying concepts.

The Introduction gives the definition of the classes of sets decidable (P) or verifiable (NP) on a Turning machine in polynomial time. One is further told about the problems concerning various relations between these classes (e.g. $P=NP$?), and the NP-complete-problems discovered by Cook and Karp.

The lectures are the following:

I. Time-bounded Turning machines and polynomial reduction. (W. Baur)

II. Polynomially bounded nondeterministic Turning machines and the completeness of the satisfiability problem of propositional logic. (A. Häussler)

III. Problems equivalent to the satisfiability problem of propositional logic. (P. Schuster)

IV. Further combinatorial problems equivalent to the satisfiability problem. (J. von zur Gathen and M. Sieveking)

(II—IV expose the Cook-Karp theory of NP-completeness: the part of Computer Science which contains the simplest and hardest open mathematical problems and is best known for its far-going consequences in applications.)

V. A polynomial algorithm for finding systems of independent representatives. (E. Specker) (The existence of this algorithm can be considered as a warning that even if first an exponential search intrudes itself upon us, sometimes a more detailed analysis leads to a polynomial algorithm.)

VI. Polynomial transformations and the Axiom of Choice. (M. Fürer) (The transformations (reductions) dealt with in Lectures III and IV have an immediate analogy with the transformations used in axiomatic set theory to prove the equivalence of various weakened forms of the Axiom of Choice. The lecture works out this analogy and constructs transformations suitable for both equivalences.)

VII. The spectral problem and complexity theory. (C.-A. Christen) (The spectrum of a logical formula is the set of cardinalities of its finite models. The generalized spectrum (called here projective class) in the set of structures with relations $(R_1, \ldots, R_k)$ (as defined by R. Fagin, whose work does not seem to be known by the author) of a formula Φ in the first order language with relation symbols $(R_1, \ldots, R_k, S_1, \ldots, S_n)$ is the set of structures which are restrictions to $(R_1, \ldots, R_k)$ of models of Φ. Spectra and generalized spectra were realized to be in a one-to-one correspondence with the sets recognizable nondeterministically in exponential resp. polynomial time. In this way old unsolved problems of spectral theory correspond to problems of complexity theory.)

VIII. Lower bounds on the complexity of logical decision problems. (J. Heintz) (Fischer and Rabin, to whose work this lecture is devoted, showed that any first order theory which has groups as models and allows for an element of infinite order, has exponential complexity of decision. Examples are the theory of real numbers and the Presburger arithmetic.)

IX. A decision method for the theory of real-closed fields due to Collins and Moenck—Soloway. (H. R. Wüthrich)

X. Simulation of Turning machines, by logical networks. (M. Fürer) (The theorem of Fischer—Pippenger is proved, which yields an efficient representation of Boolean mappings by logical networks, provided they are rapidly computable on many-tape Turning machines.)

XI. Lengths and formulas. (E. Specker and G. Wick) (Which "inner" properties of Boolean functions imply that the minimum of lengths of formulas representing them, is large? The first method, that of Neciporuk, estimates this by the help of the number of subfunctions; the second one, that of Hodes—Specker, shows that any function representable by a short formula contains some particularly simple subfunction.)

The papers can be read almost independently of each other. Although everything used is defined, knowledge of the elementary Turning machine and recursion theory is presupposed. The book will be useful for those interested in Computer Science and for combinatorists or algebraists with a taste for logic.

P. Gács (Budapest)

Universale Algebren und Theorie der Radikale (Studien zur Algebra und ihre Anwendungen, Band 1), Herausgegeben von Hans-J. Hoehnke, 85 Seiten, Akademie-Verlag, Berlin, 1976.

This is essentially the proceedings of the International Winter School for Universal Algebra and Radical Theory, held at Reinhardsbrunn, GDR, from January 26 to February 9, 1974. It contains quite detailed abstracts of 17 lectures, given by the following authors: V. A. Andrunakievič (with Yu. M. Ryabuhin, K. K. Kračilov, and E. I. Tebyrce), L. Budach, N. Jacobson, K. Keimel (with H. Werner), J. Lambek (with B. A. Rattray), P. Němec, L. Bican, T. Kepka, J. Rosický, B. M. Schein, D. Simson, L. A. Skornjakov, Bo Stenström, R. Strecker, A. V. Tiščenko, R. Wiegandt and B. Davey. The book, having a definite categorical flavor, acquaintes the reader with several modern directions and results in radical theory.

B. Csákány (Szeged)

John Wermer, Banach Algebras and Several Complex Variables, Second Edition (Graduate Texts in Mathematics, 35), IX+161 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1976.

The relationships between the theory of functions of one or more complex variables and that of commutative Banach algebras have been studied extensively during the past twenty years. Function theoretic methods have been applied to solve Banach algebra problems, for example, the question of the existence of idempotents in a Banach algebra. On the other hand, the concepts of the theory of Banach algebras such as the maximal ideal space, the Silov boundary, Gleason parts have fertilized function theory. About one third of the book is devoted to the most important applications of function theory in several complex variables to Banach algebras. No knowledge of function theory in several complex variables is assumed on the part of the reader. The rest of the book studies uniform approximation on compact subsets of the space of n complex variables.

The exposition is elementary and self-contained. The emphasis is put on the easy understanding of the main ideas and not on generality and completeness.

The connections between the theory of functions of one complex variable and Banach algebras are only "touched on", as this topic is well treated in other monographs.

This second edition contains the following sections: 1. Preliminaries and notations, 2. Classical approximation theorems, 3. Operational calculus in one variable, 4. Differential forms, 5. The $\bar{\partial}$-operator, 6. The equation $\bar{\partial}u=f$, 7. The Oka-Weil theorem, 8. Operational calculus in several variables, 9. The Silov boundary, 10. Maximality and Radó's theorem, 11. Analytic structure, 12. Algebra of analytic functions, 13. Appromixation on curves in C^n, 14. Uniform approximation on disks in C^n, 15. The first cohomology group of a maximal ideal space, 16. The $\bar{\partial}$-operator in smoothly bounded domains, 17. Manifolds without complex tangents, 18. Submanifolds of high dimension, 19. Generators, 20. The fibers over a plane domain, 21. Examples of hulls, 22. Solutions to some exercises. Sections 18—21 are new, Section 11 has been revised. Exercises of varying degrees of difficulty are offered, the starred exercises are solved in Section 22.

This excellent book may be recommended mainly to specialists or to those wanting to become specialists in the subject matter treated in the book.

József Szűcs (Szeged)

O. Zariski—P. Samuel, Commutative algebra, Vols. I, II, (Graduate Texts in Mathematics. 28, 29), Springer-Verlag, New York—Heidelberg—Berlin, 1975.

This is a new and essentially unchanged edition of a great classic in commutative algebra. The original was published with Van Nostrand, Princeton, N. J. in 1958 (vol. I) and 1960 (vol. II).

Since that time several excellent textbooks on commutative algebra have been written. To mention just some of them: N. BOURBAKI, *Algèbre commutative* (Hermann, Paris, 1961—1965); M. F. ATIYAH—I. G. MACDONALD, *Introduction to commutative algebra* (Addison—Wesley, 1969); I. KAPLANSKY, *Commutative rings* (Allyn and Bacon, 1970). Important new developments, like the work of Grothendieck have also taken place in commutative algebra. None the less the book of Zariski and Samuel still constitutes an excellent and thorough introduction to those classical parts of the theory which can be handled without the use of homological methods.

Let us describe the contents of the book. The first volume consists of 5 chapters while the second one contains 3 chapters and 7 appendices.

Chapter 1 deals with fundamental concepts: groups, rings, fields, unique factorization, and euclidean domains, polynomial rings, vector specas. Quotient rings are also introduced. The material in this chapter is much the same as in most textbooks of algebra (except that non-commutative structures are not studied at all).

Chapter 2 deals with the theory of fields: algebraic extensions (separable and inseparable), normal extensions and splitting fields, the elements of Galois theory, finite fields, norms, traces and the discriminant. Next come transcendental extensions, with a discussion of the transcendence degree. Algebraically closed fields are considered and algebraic function fields and derivations are discussed.

Chapter 3 contains classical material on ideals and modules. Prime, primary and maximal ideals are considered and the chain conditions introduced. A discussion of direct sums follows. Tensor products of rings and free joins of integral domains are defined and studied.

Chapter 4 discusses noetherian rings. After the Hilbert basis theorem comes a thorough presentation of the Lasker—Noether decomposition theory. Quotient rings are then studied, especially the relations between the ideals of a ring and its quotient ring. Prime ideals in noetherian and in particular principal ideal rings are discussed. There is an appendix on primary representation in noetherian modules.

Chapter 5 starts with a discussion of integral dependence and integral closure. Dedekind domains are then thoroughly discussed as well as finite algebraic extensions of quotient fields of Dedekind domains. Some sections deal with ramification theory after which some applications to quadratic and cyclotomic fields are given.

Chapter 6 discusses valuation theory. Places are introduced, the valuation ring, residue field and dimension of a place are defined. Next comes a discussion of specializations and the existence of places. The behaviour of places under field extensions is considered. Valuations are then introduced and their connection with places analyzed. The rank of a valuation is considered together with the behaviour of valuations under field extensions. Ramification theory of general valuations is presented as a generalisation of the ramification theory of Chap. 5 (which turns out to have dealt with the case of a discrete, rank 1 valuation). After a discussion of prime divisors in function fields, the abstract Riemann surface of a field is introduced with a discussion of the topological aspects. Finally normal and derived normal models are considered.

Throughout the whole chapter there is a strong evidence of the algebro-geometric motivation of the authors.

Chapter 7 deals with polynomial and power-series rings. More generally graded rings are introduced together with a study of homogeneous ideals. Algebraic varieties in affine and projective spaces are considered, the Nullstellensatz is proved. Then comes the dimension theory in finite integral domains, especially polynomial rings and then the dimension theory in power-series rings. The chapter closes with a study of characteristic functions (with a proof of the Hilbert—Serre theorem) and chains of syzygies.

The subject matter of Chapter 8 is local algebra. After an introduction to topological rings, modules and completions, Zariski rings (the term was introduced by Samuel in 1953) are considered. Hensel's lemma is proved. After a section on dimension theory in semi-local rings comes a discussion of the theory of multiplicities. Regular local rings are then discussed and a structure theorem of I. S. Cohen on certain complete local rings is proved. The final topic is analytical irreducibility and analytical normality of normal varieties.

The appendices deal with some special but interesting questions and form a valuable part of the book.

The present edition contains one alteration worth mentioning compared with the original edition: the new, stronger formulation (and modified proof) of Theorem 29 on pp. 303—305 of volume I.

J. Pelikán (Budapest)

LIVRES REÇUS PAR LA RÉDACTION

B. Andrásfai, Introductory graph theory, 268 pages, Budapest, Akadémiai Kiadó, 1977.

E. Bergmann—H. Noll, Mathematische Logik mit Informatik Anwendungen (Heidelberger Taschenbücher, Bd. 187), XV+324 S., Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 24,80.

P. M. Cohn, Skew field constructions (London Mathematical Society Lecture Note Series, 27), XII+253 pages, Cambridge University Press, 1977. — £ 7.50.

Colloquio Internazionale sulle Teorie Combinatorie con la collaborazione della American Mathematical Society, Roma, 3—15 Settembre, 1973. Tomo I—II (Atti dei Convegni Lincei, 17), 518+526 pages, Roma, Accademia Nazionale dei Lincei, 1976. — L. 25.000+34.000.

Combinatorial Surveys: Proceedings of the Sixth British Combinatorial Conference. Edited by P. J. Cameron, VII+226 pages, London—New York—San Francisco, Academic Press, 1977. — £ 7.00.

R. Courant—K. O. Friedrichs, Supersonic flow and shock wawes (Applied Mathematical Sciences, 21), XVI+464 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 45,—.

R. D. Driver, Ordinary and delay differential equations (Applied Mathematical Sciences, 20), IX+501 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 33,60.

H. M. Edwards, Fermat's last theorem. A genetic introduction to algebraic number theory (Graduate Texts in Mathematics, Vol. 50), XV+410 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 45,—.

W. Fleming, Functions of several variables (Undergraduate Texts in Mathematics), 2nd edition. Originally published by Addison-Wesley, XI+411 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 41,—.

W. Gähler, Grundstrukturen der Analysis, Bd. 1 (Mathematische Lehrbücher und Monographien, II. Abteilung: Mathematische Monographien, Bd. 41), VIII+412 Seiten, Berlin, Akademie-Verlag, 1977. — 48,— M.

L. Gårding, Encounter with mathematics, IX+270 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 22,30.

H. Goering, Asymptotische Methoden zur Lösung von Differentialgleichungen (Wissenschaftliche Taschenbücher, Reihe Mathematik-Physik, Bd. 144), 154 Seiten, Berlin, Akademie-Verlag, 1977. — 8,— M.

C. Gruber—A. Hintermann—D. Merlini, Group analysis of classical lattice systems (Lecture Notes in Physics, Vol. 60), XIV+326 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 31,—.

K. W. Gruenberg—A. J. Weir, Linear geometry (Graduate Texts in Mathematics, Vol. 49), 2nd edition. Originally published by Van Nostrand Publ. Co., X+198 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 29,10.

K. M. S. Humak, Statistische Methoden der Modellbildung, Bd. I. Statistische Inferenz für lineare Parameter (Mathematische Lehrbücher und Monographien. II. Abteilung: Mathematische Monographien, Bd. 43), XVI+516 Seiten, Berlin, Akademie-Verlag, 1977. — 68,— M.

Interactive Systems. Proceedings 6th Informatik Symposium, IBM. Germany, Bad Homburg v.d.H. September 1976. Edited by A. Blaser and C. Hackl (Lecture Notes in Computer Science, Vol. 49), VI+380 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 31,—.

VII. Internationale Konferenz über nichtlineare Schwingungen, 8—13. September 1975 in Berlin, Bd. I, 1—2: Analytische Theorie. Qualitative Theorie und Stabilitatstheorie. Dynamische Systeme und Bifurkationstheorie (Abhandlungen der Akademie der Wissenschaften der DDR, Abteilung Mathematik, Naturwissenschaften, Technik, 3—4), 498+410 Seiten, Berlin, Akademie-Verlag, 1977. — 64,—+58,— M.

K. Kending, Elementary algebraic geometry (Graduate Texts in Mathematics, Vol. 44), VIII+309 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 42,70.

E. E. Kummer, Collected papers, Vol. 1: Contributions to number theory. Edited by A. Weil, VIII+957 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1975. — DM 98,—.

W. Ledermann, Introduction to group characters, VIII+174 pages, Cambridge—London—New York, Cambridge University Press, 1977. — £ 3.30.

J. Lindenstrauss—L. Tzafriri, Classical Banach spaces I. Sequence spaces (Ergebnisse der Mathematik und ihrer Grenzgebiete — A Series of Modern Surveys in Mathematics, 92), XIII+188 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 52.—.

M. Loève, Probability theory I (Graduate Texts in Mathematics, Vol. 45), 4th edition. Originally published by Van Nostrand Publ. Co., XVII+425 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 45,—.

R. C. Lyndon—P. E. Schupp, Combinatorial group theory (Ergebnisse der Mathematik und ihrer Grenzgebiete — A Series of Modern Surveys in Mathematics, 89), XIV+339 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 72,—.

E. E. Moise, Geometric topology in dimensions 2 and 3 (Graduate Texts in Mathematics, Vol. 47), X+262 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 45,—.

S. A. Morris, Pontryagin duality and the structure of locally compact Abelian groups (London Mathematical Society Lecture Notes Series, 29), VIII+128 pages, Cambridge University Press, 1977.— £ 4.95.

J. C. C. Nitsche, Vorlesungen über Minimalflächen (Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Bd. 199), XIII+775 Seiten, Berlin—Heidelberg—New York, Springer-Verlag, 1975. — DM 196,—.

Production Theory and its Applications. Proceedings of a workshop. Edited by H. Albach and G. Bergendahl (Lecture Notes in Economics and Mathematical Systems. Operations Research, Vol. 139), VIII+186 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 18,80.

M. H. Protter—C. B. Morrey, A first course in real analysis (Undergraduate Texts in Mathematics), XII+507 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 42,70.

P. Quittner, Problems, programs, processing, results. Software techniques for Sci-Tech programs, 381 pages, Budapest, Akadémiai Kiadó, 1977.

N. Rouche—P. Habets—M. Laloy, Stability theory by Liapunov's direct method (Applied Mathematical Sciences, 22), XII+396 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977.— DM 33,60.

R. K. Sachs—H. Wu, General relativity for mathematicians (Graduate Texts in Mathematics, Vol. 48), XII+291 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 45,—.

B.-W. Schulze—G. Wildenhain, Methoden der Potentialtheorie für elliptische Differentialgleichungen beliebiger Ordnung (Mathematische Lehrbücher und Monographien, II. Abteilung: Mathematische Monographien, Bd. 46), XV+408 Seiten, Berlin, Akademie-Verlag, 1977. — M. 66,—

The State of the Art in Numerical Analysis. Proceedings of the Conference on the State of the Art in Numerical Analysis held at the University of York, April 12th—15th, 1976. Edited by D. Jacobs, XIX+978 pages, New York—London—San Francisco, Academic Press, 1977. — £ 20,00.

Theoretical Computer Science. 3rd GI Conference, Darmstadt March 28—30, 1977. Edited by H. Tzschach, H. Waldschmidt, H. K. G. Walter on behalf on the GI (Lecture Notes in Computer Science, Vol. 48), VII+418 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1977. — DM 37,—.

TOMUS XXXIX — 1977 — 39. KÖTET

BIBLIOGRAPHIE

ANNOUNCEMENT

The International Congress of Mathematicians will be held in Helsinki, Finland, during August 15—23, 1978. Further details will be issued in the autumn of 1977. Correspondence concerning the Congress should be addressed to:

International Congress of Mathematicians, ICM 78
Department of Mathematics
University of Helsinki
Hallituskatu 15
SF-00100 Helsinki 10
Finland

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 0001-6969

INDEX: 26024

77-3534 — Szegedi Nyomda — F. v.: Dobó József igazgató

Felelős szerkesztő és kiadó: Szőkefalvi-Nagy Béla
A kézirat a nyomdába érkezett: 1977. augusztus 16.
Megjelenés: 1977. december 30

Példányszám: 1200. Terjedelem: 17,70 (A/5) ív
Készült monószedéssel, íves magasnyomással,
az MSZ 5601-24 és az MSZ 5602-55 szabvány szerint