ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 41

FASC. 1—2

SZEGED, 1979

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

41. KÖTET
FASC. 1—2

SZEGED, 1979

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 41

SZEGED, 1979

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

41. KÖTET

SZEGED, 1979

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

Acta Sci. Math., 41 (1979), 3—14

The value distribution of entire functions of order at most one

I. N. BAKER and L. S. O. LIVERPOOL

§ 1. Introduction and results

Recently S. KIMURA [6] proved

Theorem A. *Let f be an entire function of order less than one and w_n a sequence such that $|w_n|\to\infty$ as $n\to\infty$. Suppose that all the roots of the equations $f(z)=w_n$ $(n=1, 2, \ldots)$ lie in a half-plane (say $\mathrm{Re}\, z\geqq 0$). Then f is a polynomial of degree at most 2.*

We begin by improving Theorem A a little to

Theorem 1. *If f is an entire function whose growth is at most order one and minimal type, and w_n is a sequence such that $|w_n|\to\infty$ while all roots of $f(z)=w_n$ $(n=1, 2, \ldots)$ lie in a half-plane, then f is a polynomial of degree at most 2.*

In this form the theorem is sharp. For any $d>0$ the function e^{dz} has type d and is bounded in $\mathrm{Re}\, z\leqq 0$ so that any sequence w_n such that $1<|w_n|\to\infty$ may be taken to satisfy the hypothesis in $\mathrm{Re}\, z\geqq 0$.

Theorem 1 has an application in the theory of iteration of entire functions (see e.g. FATOU [5] for proofs of the following results). The iterates f^n of an entire function f are defined by $f^1=f$, $f^{n+1}=f^n\circ f=f\circ f^n$ $(n=1, 2, \ldots)$. If f is non-linear the set $\mathfrak{C}(f)$ of points in whose neighbourhood $\{f^n\}$ is a normal family, is a proper open subset of the plane. The complement $\mathfrak{F}(f)$ of $\mathfrak{C}(f)$ is a non-empty, unbounded, perfect set. $\mathfrak{F}(f)$ has the *invariance property:*

If $w\in\mathfrak{F}(f)$ and $f(z)=w$, then $z\in\mathfrak{F}(f)$ and $f(w)\in\mathfrak{F}(f)$.

In iteration theory the fixed points of f are important. A fixed point z of f of order k is a solution of $f^k(z)=z$. It is proved in [5] that every point of $\mathfrak{F}(f)$ is a limit point of fixed points of f.

Received January 7, 1978.

It may happen that a component of $\mathfrak{C}(f)$ contains a half-plane. Thus for $d>0$ the function

$$g(z) = d^{-1}(e^{dz}-1) \tag{1}$$

maps $H=\{z\colon \operatorname{Re} z<0\}$ into itself so that $\{g^n\}$ is normal in H.

Suppose that conversely g is a transcendental entire function and that $\mathfrak{C}(g)$ contains a half-plane, which we may take to be $\operatorname{Re} z<0$. Then $\mathfrak{F}(g)$ lies in $\operatorname{Re} z\geqq 0$ and if we take a sequence $w_n\in\mathfrak{F}(g)$ such that $|w_n|\to\infty$, all solutions of $f(z)=w_n$ lie in $\mathfrak{F}(g)$ by the invariance property, and hence in $\operatorname{Re} z\geqq 0$. Thus from Theorem 1 we have

Theorem 2. *If g is a transcendental entire function such that the domain of normality $\mathfrak{C}(g)$ of $\{g^n\}$ contains a half-plane, then the growth of g must be at least of order* 1, *positive type.*

Example (1) shows that this is sharp with respect to growth. Related problems have been discussed under more restrictive conditions by P. Bhattacharyya [4].

If $0\in\mathfrak{F}(g)$ then every solution z of $g(z)=0$ belongs to $\mathfrak{F}(g)$. The following Theorem 3a is thus a strengthening of theorem 2.

We introduce the notation

$$A(\theta,\delta) = \{z\colon |\arg z-\theta| < \delta\}. \tag{2}$$

Theorem 3a. *Suppose* (i) *g is a transcendental entire function whose growth is at most of order* 1, *minimal type,* (ii) *all the zeros of g lie in* $\operatorname{Re} z\geqq 0$.

Then for any $\delta>0$ the set $\mathfrak{F}(g)\cap A(\pi,\delta)$ is unbounded.

Because of the importance of fixed points it is interesting that we can also prove

Theorem 3b. *If in* 3a (ii) *is replaced by the hypothesis that the first order fixed points lie in* $\operatorname{Re} z\geqq 0$, *the conclusion remains true.*

The example (1), for which all first order fixed points lie in $\operatorname{Re} z\geqq 0$, shows that 3b ceases to hold if the assumption of minimal type is dropped.

In the circumstances of Theorems 3a or 3b it follows that $A(\pi,\delta)$ must contain fixed points of some order of g. Can one be more explicit about the order of such fixed points? Let us take 3b and make the stronger hypothesis in (ii) that all the first order fixed points are real and positive. Our methods and results differ slightly according to the order of g. For order less than $\frac{1}{2}$ we have

Theorem 4a. *Suppose* (i) *g is transcendental entire of at most order $\frac{1}{2}$, minimal type, and*

(ii) *all but finitely many first order fixed points of g are real and positive.*

Then for any $\delta>0$, $A(\pi,\delta)$ contains infinitely many fixed points of order k for each $k\geqq 2$.

Indeed the fixed points of higher order, whose existence is shown in the theorem can be taken to be non-real. This is somewhat analogous to the result of the first author in [2] that if f is transcendental entire of order less than $\frac{1}{2}$ and l is a straight line, then not all solutions of $f^2(z)-z=0$ lie in l. Neither result includes the other but both show that second order fixed points tend to be scattered in their angular distribution.

If the order of g exceeds $\frac{1}{2}$ we have not been able to prove the existence of fixed points of order 2 in $A(\pi, \delta)$. However we can prove

Theorem 4b. *If in Theorem* 4a (i) *is replaced by the assumption that the order of g is strictly positive, but at most order* 1 *minimal type, then for any* θ, δ *subject to* $\frac{\pi}{2}<\theta<\frac{3\pi}{2}$, $\delta>0$, *we have that* $A(\theta, \delta)$ *contains infinitely many fixed points of order k for each* $k \geqq 3$.

Thus in particular if g is at most of order 1 minimal type and all first order fixed points are real and positive, f has fixed points of every order greater than 2 in $A(\pi, \delta)$, however small $\delta>0$ is taken.

The arguments used in this discussion can also be applied to show that functions of certain classes are not expressible as iterates of entire functions. An example is furnished by

Theorem 5. *Suppose the transcendental function F is such that*
(i) $\limsup_{r\to\infty} \{\log\log\log M(F, r)\}/\log r<1$,
(ii) *all first order fixed points of F lie in* $\operatorname{Re} z \geqq 0$, *and*
(iii) *F is bounded in* $A(\pi, \delta)$ *for some* $\delta>0$.
Then F is not expressible as $f^k, k \geqq 2$, *for any entire f.*

In (ii) we may replace fixed points by zeros without affecting the validity of the theorem. The function e^{e^z} has all its fixed points in $\operatorname{Re} z \geqq 0$ and shows that we cannot allow equality in (i).

§ 2. Proof of Theorem 1

We may assume $f(0) \neq 0$ (otherwise consider $f(z-\delta)$ for a suitable positive constant δ).

We shall use the following results about functions of minimal type whose zeros lie in a half-plane. They may be found e.g. in the proof of theorem 1 of [8], where the additional hypothesis $f(-r)=O(r^k)$ of that theorem is not used until after these facts have been derived.

Lemma 1. *Let f be a transcendental entire function of at most order one and minimal (i.e. zero) exponential type. Suppose $f(0)\neq 0$ and that all zeros a_n of f lie in the right half plane* $\operatorname{Re} z\geqq 0$. *Then there are constants A and c such that*

$$f(z) = Ae^{cz} \prod_{n=1}^{\infty}\left(1-\frac{z}{a_n}\right)e^{\frac{z}{a_n}}, \tag{3}$$

where $a_n=r_n e^{i\theta_n}$ *is such that*

$$\lambda = \operatorname{Re}\sum_{n=1}^{\infty} a_n^{-1} = \sum_{n=1}^{\infty} (\cos\theta_n)/r_n \tag{4}$$

is convergent and

$$\lambda+\operatorname{Re} c = 0. \tag{5}$$

Further, for any fixed k

$$|f(-r)|/r^k \to \infty \quad as \quad r\to\infty. \tag{6}$$

Proof of Theorem 1. We may suppose $w_1=0$ (for otherwise consider $f(z)-w_1$) and suppose first that f is transcendental entire of at most order one, minimal type and that all solutions of $f(z)=w_n$ lie in H: $\operatorname{Re} z\geqq 0$. In particular the zeros $a_n=r_n e^{i\theta_n}$ lie in H, so by Lemma 1

$$\frac{f'(z)}{f(z)} = c+\sum_{n=1}^{\infty}\left(\frac{1}{z-a_n}+\frac{1}{a_n}\right).$$

Using (4) and (5) this yields

$$\operatorname{Re}\frac{f'(z)}{f(z)} = \sum_{n=1}^{\infty}\operatorname{Re}\frac{1}{z-a_n}. \tag{7}$$

If $\operatorname{Re} z<0$ and $\operatorname{Re} a\geqq 0$ we have $\operatorname{Re}\dfrac{1}{z-a}<0$, while if $z=\varrho e^{i\varphi}$ then for fixed φ

$$|z|\operatorname{Re}\frac{1}{z-a}\to\cos\varphi \quad \text{as } \varrho\to\infty.$$

Thus by (7), if δ is a fixed number such that $0<\delta<\dfrac{\pi}{2}$,

$$|z|\operatorname{Re}\frac{f'(z)}{f(z)}\to -\infty \quad \text{as} \quad z\to\infty \quad \text{in} \quad A(\pi,\delta).$$

Take a fixed constant $K>2\pi/\delta$. Then there is a constant r_0 such that

$$|z|\left|\frac{f'}{f}\right| > K \quad \text{for} \quad z\in A(\pi,\delta), \quad |z|>r_0. \tag{8}$$

Next choose a member of the given sequence w_n so that $|f(z)|<|w_n|$ for $|z|\leqq r_0$. By (6) there is a largest r_n such that $|f(-r_n)|=|w_n|$. There is a component G of $\{z\colon |f(z)|>|w_n|\}$ which contains $\{z\colon z=-r<-r_n\}$ and this component is bounded by a level curve $\Gamma: |f(z)|=|w_n|$ which passes through $z=-r_n$. Γ cannot close in $\mathrm{Re}\, z<0$ for there are no zeros of f in this region.

If Γ meets neither of the lines $\arg z=\pi\pm\delta$, then G lies entirely in the angle $A(\pi, \delta)$. Let $r\theta(r)$ be the length of that segment γ_r of $|z|=r$ which lies in G and contains $z=-r$. By the arguments used in the proof of the Denjoy—Carleman—Ahlfors theorem in [9, pp. 310—311] it follows that for all sufficiently large $r(>r_1$ say) the maximum modulus function $M(f, r)$ of f satisfies

$$\log\log M(f, r) > \log\log \operatorname*{Max}_{\gamma_r} |f(z)| > \pi \int_{r_1}^{r} \frac{dt}{t\theta(t)} + C,$$

for a suitable constant C. Since $\theta(r)<2\delta$ this implies that f has order at least $\pi/2\delta>1$, which is impossible.

Thus there is a level curve $\Gamma: |f(z)|=|w_n|$, which starts at $z=-r_n$ and runs to either $\arg z=\pi+\delta$ or $\pi-\delta$. Moreover Γ lies in $|z|\geqq r_0$ so that the inequality (8) holds on Γ. But $w=f(z)$ maps Γ onto $|w|=|w_n|$ and as z traverses Γ, w traverses $|w|=|w_n|$ without change of direction. Further, we have

$$\frac{dw}{w} = \frac{dz}{z} \cdot \frac{zf'(z)}{f(z)},$$

whence, if $w=|w_n|e^{i\varphi}$ and $z=re^{i\theta}\in\Gamma$ we have

$$i\,d\varphi = \left(\frac{dr}{r} + i\,d\theta\right)\frac{zf'(z)}{f(z)}, \tag{9}$$

so that by (8)

$$|d\varphi| \geqq |d\theta| \left|\frac{zf'(z)}{f(z)}\right| > K|d\theta|.$$

The image of Γ is therefore an arc of $|w|=|w_n|$ whose angular measure is at least $K\delta>2\pi$. Thus Γ, and in particular $A(\pi, \delta)$ must contain a root z of $f(z)=w_n$, against the hypothesis of the theorem.

We conclude that f cannot therefore be transcendental. If f is a polynomial its degree can clearly not exceed two.

§ 3. Proof of Theorem 3

Suppose g is a transcendental entire function of growth at most order 1, minimal type and is such that

$$A(\pi, \delta) \cap \mathfrak{F}(g) \tag{10}$$

is bounded for some $\delta>0$. Without loss of generality we may assume that the set in (10) is empty — it is only necessary to shift the origin and consider the iteration of $g(z+a)-a$ for sufficiently large negative a.

Whether the zeros of $g(z)$ or the fixed points (i.e. the zeros of $g(z)-z$) are in $\operatorname{Re} z \geqq 0$ it follows from Lemma 1 that for any k

$$g(-r)/r^k \to \infty \quad \text{as} \quad r \to \infty. \tag{11}$$

Since $A = A(\pi, \delta)$ does not meet $\mathfrak{F}$, A belongs to an unbounded component G of the set $\mathfrak{C}(g)$ of normality of g^n. Indeed by [3] G is simply-connected. The boundary ∂G belongs to $\mathfrak{F}$ and is a continuum in the complex sphere. By the invariance property of $\mathfrak{F}$, $g(z)$ omits all the values of ∂G for $z \in A$.

If $M = \pi/(2\delta)$ the transformation

$$u = (1+t)/(1-t), \quad z = -u^{\frac{1}{M}} \tag{T}$$

maps $|t|<1$ onto A, so that the function

$$w = h(t) = g\left\{-\left(\frac{1+t}{1-t}\right)^{\frac{1}{M}}\right\}$$

is regular in $|t|<1$ and omits the values $w \in \partial G$.

By a result of J. E. Littlewood [7]

$$M(h, \varrho) = O\{(1-\varrho)^{-2}\} \quad \text{as} \quad \varrho \to 1^-.$$

If $z = re^{i\theta} \in A$, and $|\theta - \pi| < \delta/2$, then in (T)

$$\begin{aligned} t &= (1 - e^{Mi(\pi-\theta)} r^{-M})/(1 + e^{Mi(\pi-\theta)} r^{-M}) \\ &\sim 1 - 2e^{Mi(\pi-\theta)} r^{-M} \quad \text{as} \quad r \to \infty. \end{aligned}$$

Since $|M(\pi-\theta)| < \pi/4$ we have $1-|t| > r^{-M}$ for large r. Thus as $z = re^{i\theta} \to \infty$ in $|\theta - \pi| < \frac{1}{2}\delta$ we have

$$|g(z)| = |h(t)| < M(h, 1-r^{-M}) = O(r^{2M}).$$

But this conflicts with (11). The result follows.

§ 4. Preliminaries to the proof of Theorems 4a and 4b

Throughout this section assume that g is an entire function such that

(i) g is transcendental and of at most order one, minimal type,

(ii) all but finitely many fixed points of first order of g are real and positive.

Then we have

$$g(z) - z = p(z)e^{cz} \prod_{n=1}^{\infty} \left(1 - \frac{z}{a_n}\right) e^{\frac{z}{a_n}},$$

where p is a polynomial of degree say $d \geqq 0$, and $a_n > 0$. Applying lemma 1 to $\{g(z)-z\}/p(z)$ we see that $\sum a_n^{-1}$ converges and in fact

$$g(z)-z = p(z)\exp(i\gamma z)\prod_{n=1}^{\infty}\left(1-\frac{z}{a_n}\right),$$

where γ is real. If $\gamma \neq 0$ then

$$\operatorname{Max}|g(\pm iy)| > \exp|\gamma y|$$

so $\gamma = 0$ since g has minimal type. Thus

$$g(z) = z + h(z), \quad h(z) = p(z)Q(z) = p(z)\prod_{n=1}^{\infty}\left(1-\frac{z}{a_n}\right). \tag{12}$$

Lemma 2. *If g satisfies* (i), (ii) *then there is some $r_0 > 0$ such that $|g(-r)|$ is increasing for $r > r_0$, so that $w = g(-r)$, $r > r_0$ describes a simple curve Γ. Γ approaches infinity in a limiting direction* $\arg w = \alpha$.

For, let δ satisfy $0 < \delta < \frac{\pi}{2}$. From (12) it follows that as $z \to \infty$ in $A(\pi, \delta)$ we have $|h(z)/z| \to \infty$ and

$$\left|\frac{zh'}{h}\right| = \left|\frac{zp'}{p} + \frac{zQ'}{Q}\right| = \left|d + o(1) + \frac{zQ'}{Q}\right| \to \infty,$$

(c.f. (7) and (8) in theorem 1). Thus $|h'(z)| \to \infty$ and

$$\frac{zg'}{g} = \frac{zh'}{h}\frac{(1+1/h')}{(1+z/h)} \to \infty \quad \text{as} \quad z \to \infty \quad \text{in} \quad A(\pi, \delta). \tag{13}$$

In particular

$$g'(-r)/g(-r) = \frac{-1}{r}\left\{d + o(1) + \sum_{n=1}^{\infty}\frac{r}{r+a_n}\right\} \quad \{1+o(1)\} \tag{14}$$

as $r \to \infty$, and if $g(-r) = \mathrm{R}e^{i\varphi}$ we have

$$\frac{dR}{R} + i\,d\varphi = \frac{g'(-r)}{g(-r)}(-dr). \tag{15}$$

By (14) the argument of (15) approaches zero as $r \to \infty$, so that $\frac{dR}{R} > 0$ for large r.

Clearly $|h(-r)| \to \infty$ faster than any power of r and $\arg h(-r)$ tends to a constant value, namely the argument of the leading coefficient of $p(z)$. Hence $\arg g(-r)$ approaches the same limit. The lemma is proved.

Lemma 3. *If g satisfies* (i), (ii) *then, given any real θ_0, δ, σ such that* $0<\delta<\frac{\pi}{2}$, $0<\sigma\leqq\pi$, *there exist a constant R_1 and two branches ψ and χ of $z=g^{-1}(w)$ regular in*

$$S = A(\theta_0,\sigma)\cap\{|w|>R_1\},$$

such that the values of ψ,χ satisfy $\pi-\delta<\arg\psi<\pi$ and $\pi<\arg\chi<\pi+\delta$, respectively. For any $k>0$ we have

(16) $$\operatorname{Max}\{|\psi(w)|,|\chi(w)|\} = O\left(|w|^{\frac{1}{k}}\right) \quad as \quad w\to\infty \quad in \quad S.$$

Proof. As w traverses Γ from $w_0=g(-r_0)$ to ∞ the branch of $z=g^{-1}(w)$ such that $r_0=g^{-1}(w_0)$ has a regular continuation and the values of z are all real and negative $(<-r_0)$.

For $r_1>r_0$ put $R=|g(-r_1)|$ and consider the level-curve $\lambda=|g(z)|=R$ which passes through $z=-r_1$. Along λ we have as in (9)

$$i\,d\varphi = (zg'/g)\left\{i\,d\theta+\frac{dr}{r}\right\},$$

where $z=re^{i\theta}\in\lambda$, $g(z)=\mathrm{R}e^{i\varphi}$.

By (13) for z of sufficiently large modulus in $A(\pi,\delta)$ we have for any given $K>4\pi/\delta$ that $|zg'(z)/g(z)|>K$. Thus if R and hence r are sufficiently large we have $|d\varphi|>K|d\theta|$, $|d\varphi|>K|dr|/r$. As z leaves $-r_1$ on λ and travels in a given direction to $re^{i\theta}$ the corresponding φ changes monotonely so that

$$K|\theta-\pi| = \left|\int K d\theta\right| \leqq \int K|d\theta| \leqq \int|d\varphi| = \left|\int d\varphi\right| = \Delta\varphi,$$

and similarly $K|\log(r/r_1)|\leqq\Delta\varphi$. As $w=g(z)$ traverses $|w|=R$, increasing from $\arg g(-r)$ by 4π, z traverses λ in one direction with θ changing by at most $4\pi/K<\delta$, while r satisfies

(17) $$r_1\exp(-4\pi/K) < r < r_1\exp(4\pi/K).$$

Thus if r_1 is large enough the value of z remains in $A(\pi,\delta)$ and by (13) $g'(z)\neq 0$ on λ so the value of z gives a regular continuation of $g^{-1}(w)$ from $g(-r_1)$ in Γ round $|w|=R$ through an angle of 4π. The values of z lie in $A(\pi,\delta)$ but do not meet the negative real axis except at $z=-r_1$, since $g(-r)$ is increasing. Since Γ can be taken to lie in any sector $|\arg w-\alpha|<\varepsilon$, $\varepsilon>0$, it follows that we can derive from these values of $g^{-1}(w)$ a branch ψ which satisfies the statements of lemma 3, including either $\pi-\delta<\arg\psi<\pi$ or $\pi<\arg\psi<\pi+\delta$.

If in the above construction we begin by proceeding along λ in the opposite direction from that chosen originally we construct the other branch χ of g^{-1}.

For $re^{i\theta}=\psi(\mathrm{Re}^{i\varphi})$ we have by

$$|r| = |\psi(\mathrm{Re}^{i\varphi})| < r_1 \exp(4\pi/K)$$

and from $R=|g(-r_1)|>r_1^{2k}$ for large r_1 the estimate (16) follows.

We shall also need

Lemma 4 (Pólya [10]). *Let e, f, h be entire functions which satisfy $e=f\circ h$, $h(0)=0$. Then there is a positive constant c independent of e, f, h such that*

$$M(e, r) > M\left[g, cM\left(h, \frac{r}{2}\right)\right]. \tag{18}$$

The condition $h(0)=0$ can be dropped if (18) *is to hold only for all sufficiently large r.*

§ 5. Proofs of Theorems 4a and 4b

Theorem 4a. *Suppose g satisfies the hypotheses of the theorem. The first of these implies that the minimum modulus of g is large ($>R_n$) on a sequence of circles $|z|=R_n\to\infty$. The R_n may be chosen so that there is at least one zero of g in each $R_n<|z|<R_{n+1}$. Since $|g(-r)|/r\to\infty$ as $r\to\infty$ each of the simply-connected slit annuli*

$$A_n = \{z\colon R_n < |z| < R_{n+1},\ |\arg z| < \pi\},\quad n = 1, 2, \ldots$$

contains a zero of g and has the property that

$$|g(z)| > |z| \text{ on the boundary } \partial A_n. \tag{19}$$

Denote by φ a branch of $z=g^{-1}(w)$ which is regular in $A(0, \pi)$ for sufficiently large w, with values in $\pi>\arg z>\pi-\delta$, δ being the fixed number, $0<\delta<\frac{\pi}{2}$ chosen in § 4. Such a φ exists by lemma 3.

For any fixed $l=2, 3, \ldots,$ the $(l-1)$-th iterate $\varphi^{l-1}(w)$ is defined in $A(0, \pi)$ for sufficiently large w, with values in $\pi>\arg z>\pi-\delta$. For sufficiently large n then φ^{l-1} maps A_n univalently onto a simply-connected domain D_n in $\pi>\arg z>$ $>\pi-\delta$. For $z\in\partial D_n$ we have $g^{l-1}(z)\in\partial A_n$. Now since $|g(z)|>|z|$ for large $|z|$, $z\in A(\pi, \delta)$, it follows from $z\in\partial D_n$ that $|g^{l-1}(z)|>|z|$ and from $g^{l-1}(z)\in\partial A_n$ and (19) that at

$$|g^l(z)| = \left|g(g^{l-1}(z))\right| > |g^{l-1}(z)| > |z|,$$

at least for large n.

By Rouché's theorem $g^l(z)-z$ and $g^l(z)$ have equal numbers of zeros in D_n and $0\in g(A_n)=g^l(D_n)$. Thus the region $\pi>\arg z>\pi-\delta$ and a fortiori $A(\pi, \delta)$ contains an infinity of solutions of $g^l(z)-z=0$.

Theorem 4b. *Suppose g has order ϱ, $0<\varrho\leqq 1$, and is at most of order one, minimal type, while all but finitely many first order fixed points are positive. Suppose also that $\frac{\pi}{2}<\theta<\frac{3\pi}{2}$ and that σ, $0<\sigma<\pi/2$ is so small that $\frac{\pi}{2}<\theta\pm\sigma\leqq\frac{3\pi}{2}$. Let ψ and χ be the two branches of g^{-1} whose existence is asserted in Lemma 3, in the case $\theta_0=\theta$. Then $\psi=\chi$ has no solution in $A(\theta,\sigma)\cap\{|w|>R_1\}$.*

Suppose g has only finitely many fixed points of order k in $A(\theta,\sigma)$. Then

$$F=(g^{k-1}-\psi)/(\chi-\psi)$$

is regular and different from 0, 1, ∞ for large z in $A(\theta,\sigma)$. By applying Schottky's theorem to F in $A(\theta,\sigma)$ (or in a slightly smaller sector within $A(\theta,\sigma)$ and with origin shifted so that $F\neq 0, 1, \infty$ in this sector) we find

$$F(z)=O\{\exp(C|z|^{\pi/\sigma})\} \tag{20}$$

for some constant C as $z\to\infty$ in $A(\theta,\sigma')$, $\sigma'<\sigma$. From (16) the same estimate follows for $|g^{k-1}(z)|$ with perhaps a different C.

Now there exists δ_1 such that $0<\delta_1<\frac{\pi}{2}$ and $A(\theta,\sigma')\subset A(\pi,\delta_1)$. Thus $|g(re^{i\theta})|\to\infty$ as $r\to\infty$ and $|zg'/g|>K>2\pi/\sigma'$ for large $|z|$, $z\in A(\theta,\sigma')$. As in the proof of theorem 1 there is for large R a level curve $\Gamma(R)$: $|g(z)|=R$, which passes through $z=re^{i\theta}$, say. Such a curve cannot close in $A(\theta,\delta)$ for arbitrarily large R, since $|g(z)|\to\infty$ in $A(\theta,\delta)$ and $A(\theta,\delta)$ contains only finitely many zeros of g. As in theorem 1 Γ must run to the boundary of $A(\theta,\sigma')$ in at least one direction. If γ is an arc of Γ which goes from $re^{i\theta}$ to $\partial A(\theta,\sigma')$, then from $|zg'/g|>K$ it follows that the image of γ under $w=g(z)$ is the whole of $|w|=R$.

For large R we have that if t is the point on $|t|=R$ where $|g^{k-2}(t)|=$ $=M(g^{k-2},R)$ then for some $z\in\gamma$, $g(z)=t$

$$M(g^{k-2},R)=|g^{k-1}(z)|. \tag{21}$$

Now in $A(\theta,\sigma')\subset A(\pi,\delta_1)$, $|g(z)|/|z|^N\to\infty$ as $|z|\to\infty$, for any N. Take $N>2\pi/(\varrho\sigma)$, where ϱ is the order of g. Then for large R we have from (21)

$$\max_{\substack{|z|=r\\ z\in A(\theta,\sigma')}}|g^{k-1}(z)|>M(g^{k-2},r^N). \tag{22}$$

Since $k-2\geqq 1$ the right hand side is (for large r) at least

$$M(g,r^N)>\exp(r^{N\varrho})>\exp(r^{2\pi/\sigma})$$

for some arbitrarily large r. Thus we have a contradiction between (22) and the estimate for g^{k-1} from (20). Hence g must in fact have an infinity of fixed points in $A(\theta,\sigma)$.

§ 6. Proof of Theorem 5

Suppose F satisfies the hypotheses of Theorem 5 and that there exist an entire function f and an integer $k \geqq 2$ such that $F = f^k$. Since F is bounded on the path γ which consists of the negative axis running to $-\infty$, it follows that one of $\gamma, f(\gamma), \ldots, f^{k-1}(\gamma)$ is an unbounded path on which f is bounded. From this it follows that the lower order of f is positive.

From lemma 4 and the fact that the lower order of f is positive we easily obtain a contradiction to hypothesis (i) of the theorem, provided $k \geqq 3$.

It remains to prove the theorem for $k=2$. From hypothesis (i), $F=f^2$ and the fact that the lower order of f is positive it follows from Lemma 4 (as is proved in [1, Satz 12]) that the order of f is less than one.

Now $f(z)=z+g(z)$ where the zeros of g are fixed points of f and hence of F. Thus the zeros of g lie in $\operatorname{Re} z \geqq 0$ and the order of z is less than 1. By lemma 1 we have

$$\frac{|f(-r)|}{r^2} \text{ and } \frac{|g(-r)|}{r^2} \to \infty \quad \text{as} \quad r \to \infty, \tag{23}$$

while

$$|z|\left|\frac{g'}{g}\right| > K > 2\pi/\delta \quad \text{in} \quad |z| > r_0, \quad |\arg z - \pi| < \delta. \tag{24}$$

For a large $R\ (>M(g, r_0))$ there is a level curve Γ: $|g(z)|=R$ passing through $z=-r$ such that $|g(-r)|=R>r^2$. Just as in the proof of theorem 1 it follows that Γ must run to at least one of $\arg z=\pi+\delta$ or $\pi-\delta$, say the former, and that the image under $w=g(z)$ of this arc must cover $|w|=R$ with angular measure at least $K\delta>2\pi$. Let γ denote the arc of Γ between $-r$ and a point z' chosen so that the image $g(\gamma)$ covers exactly the angular length $K\delta$ of $|w|=R$. As in the proof of Lemma 3 (17) it follows that for all $z_1=r_1e^{i\theta_1}\in\gamma$ we have $|\log(r_1/r)|<\delta$.

The arc γ is mapped by $f(z)=z+g(z)$ onto a (not necessarily closed) curve in such a way that the image of z_1 is $z_1+\mathrm{Re}^{i\varphi_1}$ where $|z_1|<re^{\delta}$, $R>r^2$, and φ_1 increases by $K\delta>2\pi$ as z_1 describes γ. Thus $f(\gamma)$ certainly cuts the negative real axis, say in a point $w'=f(z'')$, $z''\in\gamma$. Then

$$|F(z'')| = |f(f(z''))| = |f(w')| > |w'|^2 > (R-re^{\delta})^2 > \frac{1}{4}r^4$$

if R and hence r are sufficiently large. Thus $A(\pi, \delta)$ contains points z'' of arbitrarily large modulus for which

$$|F(z'')| > \frac{1}{4}e^{-4\delta}|z''|^4,$$

which contradicts (iii). This completes the proof.

References

[1] I. N. BAKER, Zusammensetzungen ganzer Funktionen, *Math. Z.*, **69** (1958), 121—163.

[2] ——— The distribution of fixed points of entire functions, *Proc. London Math. Soc.*, (3) **16** (1966), 493—506.

[3] ——— The domains of normality of an entire function, *Ann. Acad. Sci. Fennicae*, Ser. A I, **1** (1975), 277—283.

[4] P. BHATTACHARYYA, On the domain of normality of an attractive fixed point, *Trans. Amer. Math. Soc.*, **153** (1971), 89—98.

[5] P. FATOU, Sur l'itération des fonctions transcendantes entières, *Acta Math.*, **47** (1926), 337—370.

[6] S. KIMURA, On the value distribution of entire functions of order less than one, *Kōdai Math. Sem. Rep.*, **28** (1976), 28—32.

[7] J. E. LITTLEWOOD, On inequalities in the theory of functions, *Proc. London. Math. Soc.*, **23** (1925), 481—519.

[8] L. S. O. LIVERPOOL, On entire functions with infinite domains of normality, *Aequationes Math.*, **10** (1974), 189—200.

[9] R. NEVANLINNA, *Eindeutige analytische Funktionen*, 2 Aufl., Springer (Berlin, 1953).

[10] G. PÓLYA, On an integral function of an integral function, *J. London Math. Soc.*, **1** (1926), 12—15.

DEPARTMENT OF MATHEMATICS
IMPERIAL COLLEGE
LONDON SW7 2BZ

DEPARTMENT OF MATHEMATICS
FOURAH BAY COLLEGE
UNIVERSITY OF SIERRA LEONE,
FREETOWN, SIERRA LEONE

Acta Sci. Math., 41 (1979), 15—27

C_0-Fredholm operators. I

HARI BERCOVICI

In this note we introduce the notions of C_0-Fredholm and C_0-semi-Fredholm operators, which are generalisations of the Fredholm and semi-Fredholm operators, and we study some properties of these operators. The study of index problems in connection with operators that intertwine contractions of class C_0 was suggested by [10], Theorem 2 and Conjecture.

In § 1 of this note we introduce some notions and we define and study the determinant function of an arbitrary operator of class C_0. In § 2 the notions of C_0-fredholmness, C_0-semi-fredholmness, and index are defined. Here we find (Corollary 2.8) a generalisation of [10], Theorem 2 under weaker assumptions. We also show that the index defined for C_0-semi-Fredholm operators is multiplicative. At the end of § 2 we prove a perturbation theorem. In § 3 we show that there exist C_0-Fredholm operators with given index (Proposition 3.1). We also prove that the conjecture from [10] is generally false (Proposition 3.2) but is verified in the bicommutant of a C_0 contraction of arbitrary multiplicity (Proposition 3.4). At the end of § 3 we show that the set of C_0-Fredholm operators is not generally open.

§ 1. Preliminaries. The determinant function

For any (linear and bounded) operator T acting on the Hilbert space $\mathfrak{H}$ we denote by Lat (T) the set of invariant subspaces of T and by $\mathrm{Lat}_{1/2}(T)$ the set of all semi-invariant subspaces of T (that is, subspaces of the form $\mathfrak{M}\ominus\mathfrak{N}$, where $\mathfrak{M}, \mathfrak{N}\in \mathrm{Lat}(T)$ and $\mathfrak{M}\supset\mathfrak{N}$). It is known (see [4], Lemma 0) that a subspace $\mathfrak{M}$ of $\mathfrak{H}$ is semi-invariant for T if and only if

$$T_{\mathfrak{M}} = P_{\mathfrak{M}}T|\mathfrak{M} \tag{1.1}$$

Received April 18, 1978.

is a "power-compression", that is, if

$$T_{\mathfrak{M}}^{n} = P_{\mathfrak{M}} T^{n} | \mathfrak{M}, \quad n = 1, 2, \ldots. \tag{1.2}$$

If T is a completely non-unitary contraction this is equivalent to

$$u(T_{\mathfrak{M}}) = P_{\mathfrak{M}} u(T) | \mathfrak{M}, \quad u \in H^{\infty}. \tag{1.3}$$

It is obvious that $\mathrm{Lat}_{1/2}(T) = \mathrm{Lat}_{1/2}(T^*)$ (we have $\mathfrak{M} \ominus \mathfrak{N} = \mathfrak{N}^{\perp} \ominus \mathfrak{M}^{\perp}$). Let us recall that the multiplicity μ_T of the operator T is the minimum cardinality of a subset $\mathfrak{A}$ of $\mathfrak{H}$ such that $\bigvee_{n \geq 0} T^n \mathfrak{A} = \mathfrak{H}$. For each $\mathfrak{M} \in \mathrm{Lat}_{1/2}(T)$ let us put $\mu_T(\mathfrak{M}) = \mu_{T_{\mathfrak{M}}}$. If T is an operator of class C_0, we have by [7] that $\mu_T = \mu_{T^*}$. In this case we shall have

$$\mu_T(\mathfrak{M}) = \mu_{T^*}(\mathfrak{M}), \quad M \in \mathrm{Lat}_{1/2}(T). \tag{1.4}$$

For any two operators T, T' acting on $\mathfrak{H}, \mathfrak{H}'$, respectively, we denote by $\mathscr{I}(T', T)$ the set of those operators $X: \mathfrak{H} \to \mathfrak{H}'$ which satisfy the relation

$$T'X = XT. \tag{1.5}$$

Obviously, $(\mathscr{I}(T, T'))^* = \mathscr{I}(T'^*, T^*)$.

We are now going to define the determinant function of a C_0 operator acting on a separable Hilbert space.

Definition 1.1. Let T be a C_0 operator acting on a separable space and let $S(M)$, $M = \{m_j\}_{j=1}^{\infty}$ be the Jordan model of T [2]. We define the determinant function d_T as the limit of any convergent subsequence of $\{m_1 m_2 \ldots m_j\}$ $(j = 1, 2, \ldots)$.

The function d_T is uniquely determined up to a constant factor of modulus one because $|d_T| = \prod_{j=1}^{\infty} |m_j|$. If $d_T \neq 0$ then d_T is an inner function.

The C_0 operators of finite multiplicity have nonvanishing determinant function. Indeed, if $S(m_1, m_2, \ldots, m_n)$ is the Jordan model [6] of T, we have $d_T = m_1 m_2 \ldots m_n$. For any C_0 operator T the relation $d_{T^*} = d_T^{\sim}$ holds (where $f^{\sim}(z) = \overline{f(\bar{z})}$).

With this definition of the determinant function, it is obvious that d_T is invariant with respect to quasi-affine transforms. It is also obvious that $d_T = 1$ if and only if T acts on the trivial space $\{0\}$. We shall use the general notation

$$d_T(\mathfrak{M}) = d_{T_{\mathfrak{M}}} \tag{1.6}$$

for any C_0 operator T and any $\mathfrak{M} \in \mathrm{Lat}_{1/2}(T)$.

Lemma 1.2. *A contraction T of class C_0 on a separable Hilbert space is a weak contraction if and only if $d_T \neq 0$. If T is a weak contraction of class C_0, d_T coincides with the determinant of the characteristic function of T.*

Proof. If $d_T \neq 0$ it follows that the Jordan model $S(M)$ of T is a weak contraction (cf. [3], Lemma 8.4). Thus, by Proposition 4.3.a of [3], it follows that T is a weak contraction. Conversely, if T is a weak contraction, by Lemma 8.4 and Theorem 8.5 of [3] we have $d_T \neq 0$. The coincidence of d_T with the determinant of the characteristic function of T follows from [3], Theorem 8.7.

Theorem 1.3. *For any C_0 operator T acting on a separable space and any $\mathfrak{H}' \in \mathrm{Lat}(T)$ we have $d_T = d_T(\mathfrak{H}')d_T(\mathfrak{H}'')$, where $\mathfrak{H}'' = \mathfrak{H}'^{\perp}$.*

Proof. If $d_T \neq 0$, T is a weak contraction and the Theorem follows from [3], Proposition 8.2. If $d_T = 0$ we must show that either $d_T(\mathfrak{H}') = 0$ or $d_T(\mathfrak{H}'') = 0$. Equivalently, we have to show that T is a weak contraction whenever $T_{\mathfrak{H}'}$ and $T_{\mathfrak{H}''}$ are weak contractions. So, let us assume that $T_{\mathfrak{H}'}$ and $T_{\mathfrak{H}''}$ are weak contractions. Let $S(M)$, $S(M')$, $S(M'')$ be the Jordan models of T, T', T'', respectively. We consider firstly the case $\mu_T(\mathfrak{H}') < \infty$. For every natural number k we can find a subspace $\mathfrak{H}_k \in \mathrm{Lat}(T)$ such that $T|\mathfrak{H}_k$ is quasisimilar to $S(m_1, m_2, \dots, m_k)$. The subspace $\mathfrak{H}'_k = \mathfrak{H}' \vee \mathfrak{H}_k \in \mathrm{Lat}(T)$ and $T|\mathfrak{H}'_k$ is also of finite multiplicity. From [3], Proposition 8.2 we infer

$$d_T(\mathfrak{H}'_k) = d_T(\mathfrak{H}')d_T(\mathfrak{H}''_k), \quad \mathfrak{H}''_k = \mathfrak{H}'_k \ominus \mathfrak{H}' = \mathfrak{H}'_k \cap \mathfrak{H}''. \tag{1.7}$$

Again by [3], Proposition 8.2, $m_1 m_2 \dots m_k$ divides $d_T(\mathfrak{H}'_k)$ and $d_T(\mathfrak{H}''_k)$ divides $d_T(\mathfrak{H}'')$. Thus (1.7) implies that $m_1 m_2 \dots m_k$ divides $d_T(\mathfrak{H}')d_T(\mathfrak{H}'')$. In particular $d_T \neq 0$ and by [3], Proposition 8.2, we have $d_T = d_T(\mathfrak{H}')d_T(\mathfrak{H}'')$ in this case.

Let us remark now that from the preceding argument it follows that the equality $d_T = d_T(\mathfrak{H}')d_T(\mathfrak{H}'')$ also holds under the assumption $\mu_T(\mathfrak{H}'') < \infty$. Indeed, we have only to replace T by T^* and to use the relation $d_{T^*} = d_T^{\sim}$.

We are now considering the general case. Let $\mathfrak{H}_k$, $\mathfrak{H}'_k$, $\mathfrak{H}''_k$ have the same meaning as before. It is clear that $\mu_T(\mathfrak{H}''_k) < \infty$ and by the preceding remark it follows that $T_{\mathfrak{H}'_k}$ is a weak contraction and (1.7) holds. Arguing as in the case $\mu_T(\mathfrak{H}') < \infty$ we obtain $d_T \neq 0$, that is T is a weak contraction. This finishes the proof.

Let T, T' be two operators and $X \in \mathscr{I}(T', T)$. For every $\mathfrak{M} \in \mathrm{Lat}(T)$, $(X\mathfrak{M})^- \in \mathrm{Lat}(T')$. We shall prove now a lemma which is not particularly concerned with operators of class C_0.

Lemma 1.4. *Let T, T' be two operators and let $X \in \mathscr{I}(T', T)$. The mapping $\mathfrak{K} \mapsto (X\mathfrak{K})^-$ is onto* $\mathrm{Lat}(T')$ *if and only if $\mathfrak{K}' \mapsto (X^*\mathfrak{K}')^-$ is one-to-one on* $\mathrm{Lat}(T'^*)$

Proof. Let us assume that $\mathfrak{K}' \mapsto (X^*\mathfrak{K}')^-$ is one-to-one on $\mathrm{Lat}(T'^*)$ and let us take $\mathfrak{K}' \in \mathrm{Lat}(T')$. If we put $\mathfrak{K} = X^{-1}(\mathfrak{K}')$ and $\mathfrak{K}'_1 = (X\mathfrak{K})^-$, we have $(X^*(\mathfrak{K}'^{\perp}_1))^- = (\mathrm{ran}\, X^* P_{\mathfrak{K}'^{\perp}_1})^- = (\ker P_{\mathfrak{K}'^{\perp}_1} X)^{\perp} = (X^{-1}(\mathfrak{K}'_1))^{\perp} = (X^{-1}(\mathfrak{K}'))^{\perp}$ and by the same computation $(X^*(\mathfrak{K}'^{\perp}))^- = (X^{-1}(\mathfrak{K}'))^{\perp}$. By the assumption we have $\mathfrak{K}'^{\perp}_1 = \mathfrak{K}'^{\perp}$, $\mathfrak{K}'_1 = \mathfrak{K}'$ so that $\mathfrak{K}' = (X\mathfrak{K})^-$.

Conversely, let us assume that $\mathfrak{K}\mapsto(X\mathfrak{K})^-$ is onto $\operatorname{Lat}(T')$ and let us take $\mathfrak{K}'\in\operatorname{Lat}(T'^*)$. Then $\mathfrak{K}'^{\perp}=(X\mathfrak{K})^-$ where $\mathfrak{K}=X^{-1}(\mathfrak{K}'^{\perp})$. We have $\mathfrak{K}'=(X\mathfrak{K})^{\perp}=$ $=(\operatorname{ran} XP_{\mathfrak{K}})^{\perp}=\ker P_{\mathfrak{K}}X^*=X^{*-1}(\mathfrak{K}^{\perp})=X^{*-1}((X^{-1}(\mathfrak{K}'^{\perp}))^{\perp})=X^{*-1}(\ker P_{\mathfrak{K}'}X)^{\perp}=$ $=X^{*-1}(\operatorname{ran} X^*P_{\mathfrak{K}'})^-=X^{*-1}((X^*\mathfrak{K}')^-)$ which shows that $\mathfrak{K}'$ is determined in this case by $(X^*\mathfrak{K}')^-$. The lemma follows.

Remark 1.5. Because the Jordan model of a C_0 operator acting on a non-separable Hilbert space contains uncountably many direct summands of the form $S(m)$ (cf. [1]) it is natural to extend the definition of the determinant function by taking $d_T=0$ for T acting on a non-separable space. With this extension Lemma 1.2 and Theorem 1.3 remain valid with the condition of separability dropped. For Lemma 1.4 it is enough to remark that a completely non-unitary weak contraction acts on a necessarily separable space and for the Theorem 1.3 we have to remark that T acts on a separable space if and only if $\mathfrak{H}'$ and $\mathfrak{H}''$ are separable spaces.

§ 2. C_0-Fredholm operators

Definition 2.1. Let T, T' be two operators and let $X\in\mathscr{I}(T', T)$. X is called a *(T', T)-lattice-isomorphism* if the mapping $\mathfrak{M}\mapsto(X\mathfrak{M})^-$ is an isomorphism of $\operatorname{Lat}(T)$ onto $\operatorname{Lat}(T')$.

For $T=0$ and $T'=0$ a (T', T)-lattice-isomorphism is simply an invertible operator. It is clear that a lattice-isomorphism is always a quasi-affinity but the converse is not true as shown by the example $T=0$, $T'=0$. By Lemma 1.4, X is a (T', T)-lattice-isomorphism if and only if X^* is a (T^*, T'^*)-lattice-isomorphism. We shall say simply lattice-isomorphism instead of (T', T)-lattice-isomorphism whenever it will be clear which are T and T'.

Definition 2.2. Let T and T' be two operators of class C_0 and $X\in\mathscr{I}(T', T)$. X is called a *(T', T)-semi-Fredholm* operator if $X|(\ker X)^{\perp}$ is a $(T'|(\operatorname{ran} X)^-, T_{(\ker X)^{\perp}})$-lattice-isomorphism and either $d_T(\ker X)\neq 0$ or $d_{T'}(\ker X^*)\neq 0$. A (T', T)-semi-Fredholm operator X is *(T', T)-Fredholm* if both $d_T(\ker X)$ and $d_{T'}(\ker X^*)$ are different from zero. The *index* of the (T', T)-Fredholm operator X is the meromorphic function

$$j(X)=j_{(T,T')}(X)=d_T(\ker X)/d_{T'}(\ker X^*). \tag{2.1}$$

If X is (T', T)-semi-Fredholm and not (T', T)-Fredholm we define

$$j(X)=0 \quad\text{if}\quad d_T(\ker X)=0;\quad j(X)=\infty \quad\text{if}\quad d_{T'}(\ker X^*)=0. \tag{2.2}$$

We shall say simply C_0-semi-Fredholm, C_0-Fredholm instead of (T', T)-semi-Fredholm, (T', T)-Fredholm, respectively, whenever it will be clear which are the C_0 operators T and T'. We shall denote by $\mathrm{sF}(T', T)$ (respectively $\mathrm{F}(T', T)$) the set of all (T', T)-semi-Fredholm (respectively (T', T)-Fredholm) operators. If $T=T'$ we shall write $\mathrm{sF}(T)$, $\mathrm{F}(T)$ instead of $\mathrm{sF}(T, T)$, $\mathrm{F}(T, T)$, respectively.

We can easily see how the preceding definition is related to the usual definition of Fredholm operators. Let us note that the operator $T=0$ acting on the Hilbert space $\mathfrak{H}$ is a C_0 operator; it is a weak contraction if and only if $n=\dim \mathfrak{H}<\infty$ and in this case $d_T(z)=z^n$ ($|z|<1$). If $T=T'=0$ and $X\in\mathscr{I}(T', T)=\mathscr{L}(\mathfrak{H})$ then $X|(\ker X)^\perp$ is a lattice-isomorphism if and only if X has closed range. From these remarks it follows that an operator $X\in\mathscr{I}(0, 0)$ is C_0-Fredholm if and only if it is Fredholm in the usual sense, and $j(X)(z)=z^{i(X)}$, where $i(X)=\dim\ker X-\dim\ker X^*$ is the (usual) index of the Fredholm operator X.

Proposition 2.3. *Let T, T', T'' be C_0-operators acting on $\mathfrak{H}, \mathfrak{H}', \mathfrak{H}''$, respectively, and let $A\in\mathscr{I}(T, T')$, $B\in\mathscr{I}(T, T'')$ be such that $A\mathfrak{H}'\subset(B\mathfrak{H}'')^-$. If $d_{T'}\neq 0$, we have:*

$$(A^{-1}(B\mathfrak{H}''))^- = \mathfrak{H}'; \tag{2.3}$$

$$(A\mathfrak{H}'\cap B\mathfrak{H}'')^- \supset A\mathfrak{H}'. \tag{2.4}$$

Proof. It is enough to prove (2.3) because (2.4) is a simple consequence of (2.3).

We may suppose that B is a quasi-affinity and A is one-to-one. Indeed, we have only to replace A, B respectively by $A|(\ker A)^\perp$ and $B|(\ker B)^\perp$, and $\mathfrak{H}$ by $(B\mathfrak{H}'')^-$. It follows that $d_{T''}=d_T$ and T' is quasisimilar to the restriction of T to some invariant subspace. By Theorem 1.3 we have $d_{T'}\neq 0$ and therefore

$$d_{T'\oplus T''} = d_{T'}d_{T''} = d_{T'}d_T \neq 0. \tag{2.4}$$

The operator $X\colon \mathfrak{H}'\oplus\mathfrak{H}''\to\mathfrak{H}$ defined by $X(h'\oplus h'')=Ah'-Bh''$ has dense range and satisfies $TX=X(T'\oplus T'')$.
Thus $(T'\oplus T'')_{(\ker X)^\perp}$ is a quasi-affine transform of T, in particular

$$d_{T'\oplus T''}((\ker X)^\perp) = d_T. \tag{2.5}$$

From (2.4) and (2.5) we infer

$$d_{T'\oplus T''}(\ker X) = d_{T'}. \tag{2.6}$$

The operator $Y\colon \ker X\to\mathfrak{H}'$ defined by $Y(h'\oplus h'')=h'$ is one-to-one. Indeed, $Y(h'\oplus h'')=0$ and $h'\oplus h''\in\ker X$ imply $h'=0$ and $Bh''=Ah'=0$; it follows that $h''=0$ because B is one-to-one. Moreover, we have $Y\in\mathscr{I}(T', (T'\oplus T'')|(\ker X))$. It is easy to verify that $\operatorname{ran} Y=A^{-1}(B\mathfrak{H}'')$. By the invariance of the determinant function we have

$$d_{T'}((A^{-1}(B\mathfrak{H}''))^-) = d_{T'\oplus T''}(\ker X) = d_{T'}. \tag{2.7}$$

From Theorem 1.3 and relation (2.7) it follows that

$$(2.8)\qquad d_{T'} = d_{T'}((A^{-1}(B\mathfrak{H}''))^-)d_{T'}((A^{-1}(B\mathfrak{H}''))^\perp) = d_{T'}d_{T'}((A^{-1}(B\mathfrak{H}''))^\perp)$$

and therefore

$$d_{T'}((A^{-1}(B\mathfrak{H}''))^\perp) = 1,\quad (A^{-1}(B\mathfrak{H}''))^\perp = \{0\} \text{ and (2.3) follows.}$$

The Proposition is proved.

Corollary 2.4. *Let T, T' be two C_0 operators such that $d_T \neq 0$ and let $A \in \mathscr{I}(T', T)$ be a quasi-affinity. Then A is a lattice-isomorphism.*

Proof. The correspondence $\mathfrak{K} \mapsto (A\mathfrak{K})^-$ is onto $\mathrm{Lat}\,(T')$ by Proposition 2.3. Corollary follows by Lemma 1.4 since A^* is also a quasi-affinity.

Lemma 2.5. *Let T, T' be C_0 operators and $A \in \mathscr{I}(T', T)$. We always have $d_{T'}d_T(\ker A) = d_T d_{T'}(\ker A^*)$.*

Proof. From Theorem 1.3 and the invariance of the determinant function with respect to quasi-affine transforms we infer $d_{T'} = d_{T'}(\ker A^*)d_{T'}((\mathrm{ran}\,A)^-) = d_{T'}(\ker A^*)d_T((\ker A)^\perp)$ and $d_T = d_T(\ker A)d_T((\ker A)^\perp)$. The Lemma obviously follows from these relations.

Corollary 2.6. *Let T, T' be weak contractions of class C_0. Then* $\mathrm{F}(T', T) = \mathscr{I}(T', T)$ *and* $j(A) = d_T/d_{T'}$, *for* $A \in \mathscr{I}(T', T)$.

Proof. For each $A \in \mathscr{I}(T', T)$, $A|(\ker A)^\perp$ is a lattice-isomorphism by Corollary 2.4. Also we have $d_T(\ker A) \neq 0$ and $d_{T'}(\ker A^*) \neq 0$ by Theorem 1.3. The value of $j(A)$ follows then from Lemma 2.5.

Remark 2.7. From the preceding proof it easily follows that $\mathrm{sF}(T', T) = \mathscr{I}(T', T)$ and $\mathrm{F}(T', T) = \emptyset$ if exactly one of the contractions T and T' is weak.

The following Corollary is a generalisation of [10], Theorem 2.

Corollary 2.8. *Let T and T' be weak contractions of class C_0 such that $d_T = d_{T'}$. Then each injection $A \in \mathscr{I}(T', T)$ is a lattice-isomorphism (in particular a quasi-affinity).*

Proof. Let $A \in \mathscr{I}(T', T)$ be an injection. By Corollary 2.6 $A \in \mathrm{F}(T', T)$ and $j(A) = d_T/d_{T'} = 1$; it follows that $d_{T'}(\ker A^*) = d_T(\ker A) = 1$, thus $\ker A^* = \{0\}$ and A is a quasi-affinity. The conclusion follows by Corollary 2.4.

Corollary 2.9. *Let T be a weak contraction of class C_0 and let $A \in \{T\}'$ be an injection. Then the restriction of A to each hyper-invariant subspace of T is a quasi-affinity.*

Proof. Obviously follows from the preceding Corollary.

Lemma 2.10. *For any two C_0 operators T and T' we have* $\mathrm{sF}(T, T')^* = \mathrm{sF}(T'^*, T^*)$, $\mathrm{F}(T, T')^* = \mathrm{F}(T'^*, T^*)$, *and*

$$(2.9) \qquad j(A^*) = \big(j(A)^{\sim}\big)^{-1}, \quad A \in \mathrm{sF}(T', T) \quad (\textit{here } 0^{-1} = \infty \textit{ and } \infty^{-1} = 0).$$

Proof. If $A \in \mathscr{I}(T', T)$, we have $\big(A|(\ker A)^{\perp}\big)^* = A^*|(\ker A^*)^{\perp}$, $d_{T'^*}(\ker A^*) = d_{T'}(\ker A^*)^{\sim}$ and $d_{T^*}(\ker A) = d_T(\ker A)^{\sim}$. The Lemma follows.

Theorem 2.11. *Let T, T', T'' be operators of class C_0, $A \in \mathrm{sF}(T', T)$, $B \in \mathrm{sF}(T'', T')$. If the product $j(B)j(A)$ makes sense we have $BA \in \mathrm{sF}(T'', T)$ and $j(BA) = j(B)j(A)$.*

Proof. We shall show firstly that $BA|(\ker BA)^{\perp}$ is a lattice-isomorphism. To do this we will show that the range of BA is dense in each cyclic subspace of T'', contained in $(\mathrm{ran}\, BA)^-$. The whole statement will follow from Lemma 1.4 and Lemma 2.10 and the same argument applied to $(BA)^* = A^*B^*$.

Let us remark that from the C_0-semi-fredholmness of B it follows that

$$B^{-1}\big((\mathrm{ran}\, BA)^-\big) \subset \big((\mathrm{ran}\, A)^- + \ker B\big)^-.$$

Therefore, for each $f \in (\mathrm{ran}\, BA)^-$ and $\varepsilon > 0$ we can find $g \in \big((\mathrm{ran}\, A)^- + \ker B\big)^-$ such that

$$(2.10) \qquad Bg \in \mathfrak{H}_f = \bigvee_{n \geq 0} T''^n f \quad \text{and} \quad \|Bg - f\| < \varepsilon.$$

Now, let us denote by $\mathfrak{R}$ the subspace $\big((\mathrm{ran}\, A)^- + \ker B\big)^- \ominus (\mathrm{ran}\, A)^-$ and by P the orthogonal projection of $\big((\mathrm{ran}\, A)^- + \ker B\big)^-$ onto $\mathfrak{R}$. We claim that

$$(2.11) \qquad d_{T'}(\mathfrak{R}) \neq 0.$$

Indeed, if $j(A) \neq \infty$, we have $d_{T'}(\ker A^*) \neq 0$ and $\mathfrak{R} \subset \ker A^*$. If $j(A) = \infty$ it follows from the hypothesis that $j(B) \neq 0$ and therefore $d_{T'}(\ker B) \neq 0$. But

$$(2.12) \qquad \big((\mathrm{ran}\,(P|\ker B)\big)^- = \mathfrak{R}$$

and

$$(2.13) \qquad T'_{\mathfrak{R}} = PT'|\big((\mathrm{ran}\, A)^- + \ker B\big)^-.$$

From Theorem 1.3 and the invariance of the determinant function with respect to quasi-affine transforms we infer that $d_{T'}(\mathfrak{R})$ divides $d_{T'}(\ker B)$; thus (2.11) is proved.

From the relations (2.11—13) it follows, via Proposition 2.3, that $\{k \in \mathfrak{H}_g;\ Pk \in P(\ker B)\}$ is dense in $\mathfrak{H}_g$, that is $\mathfrak{H}_g \cap \big((\mathrm{ran}\, A)^- + \ker B\big)$ is dense in $\mathfrak{H}_g$. Thus there exist $u \in (\mathrm{ran}\, A)^-$ and $v \in \ker B$ such that

$$(2.14) \qquad u + v \in \mathfrak{H}_g, \quad \|u + v - g\| < \varepsilon.$$

Now, by the C_0-semi-fredholmness of A, there exists $k\in\mathfrak{H}$ such that

$$(2.15)\qquad Ak\in\mathfrak{H}_u,\quad \|Ak-u\|<\varepsilon.$$

We have $Bu=B(u+v)\in B\mathfrak{H}_g\subset\mathfrak{H}_f$ and it follows that $B\mathfrak{H}_u\subset\mathfrak{H}_f$. Therefore $BAk\in B\mathfrak{H}_u\subset\mathfrak{H}_f$. From (2.10), (2.14) and (2.15) we infer $\|BAk-f\|\leqq\|BAk-Bu\|+$ $+\|B(u+v)-Bg\|+\|Bg-f\|<(2\|B\|+1)\varepsilon$. Because ε is arbitrarily small, the first part of the proof is done.

We obviously have

$$(2.16)\qquad \ker BA = A^{-1}(\ker B),\quad \ker(BA)^* = B^{*-1}(\ker A^*).$$

Let us consider the triangularisation $T|\ker BA=\begin{bmatrix}T_1 & X\\ 0 & T_2\end{bmatrix}$ determined by the decomposition $\ker BA=\ker A\oplus(\ker BA\ominus\ker A)$. By the C_0-semi-fredholmness of A, T_2 is a quasi-affine transform of $T'|\mathfrak{H}_1$, where

$$(2.17)\qquad \mathfrak{H}_1 = (\operatorname{ran} A)^-\cap\ker B.$$

If $d_{T'}(\ker B)\neq 0$ and $d_T(\ker A)\neq 0$ it follows that

$$(2.18)\qquad d_T(\ker BA) = d_T(\ker A)\,d_{T'}(\mathfrak{H}_1)\neq 0,$$

thus $BA\in\mathrm{sF}(T'',T)$. Analogously, if $d_{T''}(\ker B^*)\neq 0$ and $d_{T'}(\ker A^*)\neq 0$ it follows that $BA\in\mathrm{sF}(T'',T)$. From the hypothesis it follows that at least one of the situations considered must occur. Thus we always have $BA\in\mathrm{sF}(T'',T)$.

It is obvious that $d_{T''}(\ker(BA)^*)=0$ whenever $d_{T''}(\ker B^*)=0$ since $\ker(BA)^*\supset\ker B^*$. Thus the relation $j(BA)=\infty=j(B)j(A)$ is proved in this case. Let us suppose now that $j(B)=0$. Then, by Theorem 1.3 we have

$$0 = d_{T'}(\ker B) = d_{T'}(\mathfrak{H}_1)\,d_{T'}(\ker B\ominus\mathfrak{H}_1).$$

The projection onto $\ker A^*$ is one-to-one on $\ker B\ominus\mathfrak{H}_1$, thus $T'_{\ker B\ominus\mathfrak{H}_1}$ is a quasi-affine transform of some restriction of $T'_{\ker A^*}$. It follows that $d_{T'}(\ker B\ominus\mathfrak{H}_1)\neq 0$ and the preceding relation implies $d_{T'}(\mathfrak{H}_1)=0$. By (2.18), the relation $j(BA)=$ $=j(B)j(A)$ $(=0)$ is proved in this case also. If $j(A)\in\{0,\infty\}$ we have $j(BA)=$ $=\big(j((BA)^*)^\sim\big)^{-1}=\big(j(A^*)^\sim j(B^*)^\sim\big)^{-1}=j(B)j(A)$ by Lemma 2.10.

It remains now to prove the relation $j(BA)=j(B)j(A)$ for $A\in\mathrm{F}(T',T)$ and $B\in\mathrm{F}(T'',T')$. From the second relation (2.14) we infer, as before,

$$(2.18)^*\qquad d_{T''}(\ker(BA)^*) = d_{T''}(\ker B^*)\,d_{T'}(\mathfrak{H}_1^*)$$

where

$$(2.17)^*\qquad \mathfrak{H}_1^* = (\operatorname{ran} B^*)^-\cap\ker A^* = (\ker B)^\perp\cap(\operatorname{ran} A)^\perp.$$

Let us denote by Q the orthogonal projection of $\mathfrak{H}'$ onto $(\operatorname{ran} A)^\perp=\ker A^*$. If we consider the decompositions

$$(2.19)\qquad \ker B = \mathfrak{H}_1\oplus\mathfrak{H}_2,\quad \ker A^* = \mathfrak{H}_1^*\oplus\mathfrak{H}_2^*,$$

we claim that $Q|\mathfrak{H}_2$ is a quasi-affinity from $\mathfrak{H}_2$ into $\mathfrak{H}_2^*$. Indeed, if $h\in\mathfrak{H}_2$ and $g\in\mathfrak{H}_1^*$, we have $(g, Qh)=(g, h)=0$ as $g\in(\ker B)^\perp$, thus $Q\mathfrak{H}_2\subset\mathfrak{H}_2^*$. Because $\mathfrak{H}_1=\ker B\cap\cap(\operatorname{ran} A)^-=\ker(Q|\ker B)$, Q is one-to-one on $\mathfrak{H}_2$. We have only to show that $\ker A^*\ominus(Q\mathfrak{H}_2)^-=\mathfrak{H}_1^*$. If $h\in\ker A^*\ominus(Q\mathfrak{H}_2)^-$ and $g\in\ker B$ we have $(h, g)=(h, Qg)=0$ because $(Q\mathfrak{H}_2)^-=(Q(\ker B))^-$ (as $Q|\mathfrak{H}_1=0$); the inclusion $\ker A^*\ominus(Q\mathfrak{H}_2)^-\subset\mathfrak{H}_1^*$ follows and the assertion concerning $Q|\mathfrak{H}_2$ is proved.

Now, because $\mathfrak{H}_1=\ker(Q|\ker B)$, we have the intertwining relation $T'_{\mathfrak{H}_2^*}(Q|\mathfrak{H}_2)=$ $=(Q|\mathfrak{H}_2)T'_{\mathfrak{H}_2}$; in particular

$$d_{T'}(\mathfrak{H}_2)=d_{T'}(\mathfrak{H}_2^*). \tag{2.20}$$

By (2.18—20) and Theorem 1.3 we have

$$\begin{aligned} j(BA) &= d_T(\ker BA)/d_{T''}(\ker(BA)^*) = \\ &= (d_T(\ker A)/d_{T''}(\ker B^*))(d_{T'}(\mathfrak{H}_1)/d_{T'}(\mathfrak{H}_1^*)) = \\ &= (d_T(\ker A)/d_{T''}(\ker B^*))(d_{T'}(\mathfrak{H}_1)d_{T'}(\mathfrak{H}_2)/d_{T'}(\mathfrak{H}_1^*)d_{T'}(\mathfrak{H}_2^*)) = \\ &= (d_T(\ker A)/d_{T'}(\ker A^*))(d_{T'}(\ker B)/d_{T''}(\ker B^*)) = j(B)j(A). \end{aligned}$$

Theorem 2.11 is proved.

Theorem 2.12. *Let T be an operator of class C_0 acting on $\mathfrak{H}$ and let $X\in\{T\}'$ be such that $d_T((X\mathfrak{H})^-)\neq 0$. Then $I+X\in\mathrm{F}(T)$ and $j(I+X)=1$.*

Proof. We firstly show that the mapping $\operatorname{Lat}(T)\ni\mathfrak{M}\mapsto((I+X)\mathfrak{M})^-$ is onto $\operatorname{Lat}(T|((I+X)\mathfrak{H})^-)$. To do this let us take $\mathfrak{N}\in\operatorname{Lat}(T)$, $\mathfrak{N}\subset((I+X)\mathfrak{H})^-$ and let P denote the orthogonal projection of $\mathfrak{H}$ onto $(\ker X)^\perp$. Because $P\mathfrak{N}\subset(P(I+X)\mathfrak{H})^-$, $T_{(\ker X)^\perp}P=PT$ and $d_T((\ker X)^\perp)\neq 0$, it follows by Proposition 2.3 that $\mathfrak{N}'=\{h\in\mathfrak{N};\, Ph\in P(I+X)\mathfrak{H}\}$ is dense in $\mathfrak{N}$. Now we can show that $\mathfrak{N}'\subset(I+X)\mathfrak{H}$; indeed $\mathfrak{N}'\subset(I+X)\mathfrak{H}+\ker X$ and $\ker X\subset(I+X)\mathfrak{H}$ $(h=(I+X)h$ for $h\in\ker X)$. Therefore we have $N=((I+X)\mathfrak{M})^-$, where $\mathfrak{M}=(I+X)^{-1}\mathfrak{N}$.

From the preceding argument applied to $I+X^*$ and from Lemma 1.4 it follows that $(I+X)|(\ker(I+X))^\perp$ is a lattice-isomorphism. Because $\ker(I+X)\subset X\mathfrak{H}$ $(h=-Xh$ whenever $(I+X)h=0)$ and $\ker(I+X)^*\subset X^*\mathfrak{H}$, by Theorem 1.3 it follows that $I+X\in\mathrm{F}(T)$.

It remains only to compute $j(I+X)$. To do this let us consider the decomposition $\mathfrak{H}=\mathfrak{U}\oplus\mathfrak{V}$, $\mathfrak{U}=(X\mathfrak{H})^-$. With respect to this decomposition we have $I=\begin{bmatrix} I_{\mathfrak{U}} & 0 \\ 0 & I_{\mathfrak{V}} \end{bmatrix}$, $X=\begin{bmatrix} X' & X'' \\ 0 & 0 \end{bmatrix}$, where $X'\in\{T|\mathfrak{U}\}'$. Since by the hypothesis $T|\mathfrak{U}$ is a weak contraction, we infer by Corollary 2.6

$$d_T(\ker(I+X'))=d_T(\ker(I+X')^*). \tag{2.21}$$

Now, we can easily verify that $\ker(I+X)=\ker(I+X')$. The inclusion $\ker(I+X')\subset$ $\subset\ker(I+X)$ is obvious. If $h\in\ker(I+X)$ we have $h=-Xh\in\mathfrak{U}$ so that $h=-X'y$

$(X'=X|\mathfrak{U})$ and $h\in\ker(I+X')$. In particular

(2.22) $$d_T(\ker(I+X)) = d_T(\ker(I+X')).$$

It is easy to see, using the matrix representation of X, that $u\oplus v\in\ker(I+X)^*$ if and only if

(2.23) $$u\in\ker(I+X')^* \quad\text{and}\quad v=-X''^*u.$$

If we denote by Q the orthogonal projection of $\mathfrak{H}$ onto $\mathfrak{U}$, it follows from (2.23) that $Q|\ker(I+X)^*$ is an invertible operator from $\ker(I+X)^*$ onto $\ker(I+X')^*$, the inverse being given by $\ker(I+X')^*\ni u\mapsto u\oplus(-X''^*u)$. Because we have also $T_{\mathfrak{U}}^*Q=QT^*$ it follows that $T_{\mathfrak{U}}^*|\ker(I+X')^*$ and $T^*|\ker(I+X)^*$ are similar, in particular

(2.24) $$d_T(\ker(I+X)^*) = d_T(\ker(I+X')^*).$$

From (2.21), (2.22), and (2.24) it obviously follows that $j(I+X)=1$. The Theorem is proved.

§ 3. Some examples

Proposition 3.1. *For any two inner functions m and n there exist a C_0 operator T and $X\in\mathrm{F}(T)$ such that $j(X)=m/n$.*

Proof. The operator $T=(S(m)\otimes I)\oplus(S(n)\otimes I)$, where I denotes the identity operator on l^2, is of class C_0. If we denote by U_+ the unilateral shift on l^2, obviously

$$X=(I_{\mathfrak{H}(m)}\otimes U_+^*)\oplus(I_{\mathfrak{H}(n)}\otimes U_+)\in\{T\}'.$$

Moreover, X has closed range so that $X|(\ker X)^\perp$ is invertible. Because $T|\ker X$ is unitarily equivalent to $S(m)$ and $T_{\ker X^*}$ is unitarily equivalent to $S(n)$, it follows that X is C_0-Fredholm and $j(X)=m/n$.

The following proposition infirms the Conjecture from [10]. Proposition 3.4 shows however that this Conjecture is true under the assumption $X\in\{T\}''$ and with the condition $\mu_T<\infty$ dropped.

Proposition 3.2. *Let K and K_* be C_0 operators of finite multiplicities such that $d_K=d_{K_*}$. Then there exist a C_0 operator T of finite multiplicity and an $X\in\{T\}'$ such that $T|\ker X$ and $T_{\ker X^*}$ are quasisimilar to K and K_*, respectively.*

Proof. Let $S=S(m_1, m_2, \ldots, m_n)$ and $S_*=S(m_1', m_2', \ldots, m_n')$ be the Jordan models of K, K_*, respectively (it may happen that some of the m_j or m_j' be equal to 1). By the hypothesis we have

(3.1) $$m_1m_2\ldots m_n = m_1'm_2'\ldots m_n'.$$

Let us consider the operator

$$T = S(\varphi_1, \varphi_2, \dots, \varphi_n), \text{ where} \tag{3.2}$$

$$\varphi_1 = m_1 m_2 \dots m_n, \quad \varphi_2 = m_2' m_2 \dots m_n, \quad \varphi_3 = m_2' m_3' m_3 \dots m_n, \quad \dots, \tag{3.3}$$

$$\varphi_n = m_2' m_3' \dots m_n' m_n.$$

(T is generally not a Jordan operator). The matrix over H^∞ given by

$$A = \begin{bmatrix} 0 & 0 & \dots & 0 & m_1' \\ m_2' & 0 & \dots & 0 & 0 \\ 0 & m_3' & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & m_n' & 0 \end{bmatrix} = [a_{ij}]_{1 \le i, j \le n} \tag{3.4}$$

satisfies the conditions

$$a_{ij} \varphi_j \in \varphi_i H^2 \tag{3.5}$$

and therefore (cf. [2], relations (6.5—7)) the operator X defined by

$$X = [X_{ij}]_{1 \le i, j \le n}, \quad X_{ij} h = P_{\mathfrak{H}(\varphi_i)} a_{ij} h \quad (h \in \mathfrak{H}(\varphi_j)) \tag{3.6}$$

commutes with T. Now it is easy to see that

$$T|\ker X = \bigoplus_{i=1}^{n} T|(\ker X \cap \mathfrak{H}(\varphi_i)), \quad T_{\ker X^*} = \bigoplus_{i=1}^{n} T_{(\ker X^* \cap \mathfrak{H}(\varphi_i))}. \tag{3.7}$$

Using [8], p. 315, we see that $T|(\ker X \cap \mathfrak{H}(\varphi_i))$ is unitarily equivalent to $S(m_i)$ and $T_{(\ker X^* \cap \mathfrak{H}(\varphi_i))}$ is unitarily equivalent to $S(m_i')$ so that $T|\ker X$ is unitarily equivalent to S and $T_{\ker X^*}$ is unitarily equivalent to S_*. Proposition 3.2 follows.

Lemma 3.3. *If T and T' are two quasisimilar operators of class C_0 and $\varphi \in H^\infty$ then $T|\ker \varphi(T)$ and $T'|\ker \varphi(T')$ are quasisimilar.*

Proof. Let X, Y be two quasi-affinities such that $T'X = XT$ and $TY = YT'$. Then we have also $\varphi(T')X = X\varphi(T)$ and $\varphi(T)Y = Y\varphi(T')$ which shows that

$$X \ker \varphi(T) \subset \ker \varphi(T'), \quad Y \ker \varphi(T') \subset \ker \varphi(T). \tag{3.8}$$

From (3.8) it follows that $T|\ker \varphi(T)$ can be injected into $T'|\ker \varphi(T')$ and $T'|\ker \varphi(T')$ can be injected into $T|\ker \varphi(T)$. The Lemma follows by [10], Theorem 1.

Proposition 3.4. *Let T be an operator of class C_0 and $X \in \{T\}''$. Then $T|\ker X$ and $T_{\ker X^*}$ are quasisimilar. In particular we have*

$$\mathrm{sF}(T) \cap \{T\}'' = \mathrm{F}(T) \cap \{T\}'' \quad \textit{and} \quad j(X) = 1 \quad \textit{for} \quad X \in \mathrm{F}(T) \cap \{T\}''.$$

Proof. From [2] and [1] it follows that $X=(u/v)(T)$, where $u, v\in H^\infty$ and $v\wedge m_T=1$. It is easy to see that $\ker X=\ker u(T)$ and $\ker X^*=\ker u^\sim(T^*)$. By Lemma 3.3 it suffices to prove our Proposition for T a Jordan operator and $X=u(T)$. Now, a Jordan operator is a direct sum of operators of the form $S(m)$ and it is easy to see that $S(m)|\ker u(S(m))$ and $(S(m)^*|\ker(u(S(m)))^*)^*$ are both unitarily equivalent to $S(m\wedge u)$. Thus for T a Jordan operator $T|\ker u(T)$ and $T_{\ker(u(T))^*}$ are unitarily equivalent. Thus Proposition follows.

Proposition 3.5. *Let T be an operator of class C_0 and let $X\in\{T\}''$ be an injection. Then X is a lattice-isomorphism.*

Proof. Let $\mathfrak{M}\in\operatorname{Lat}(T)$; by [9] we have $X\mathfrak{M}\subset\mathfrak{M}$. Moreover we have $X|\mathfrak{M}\in\operatorname{Alg\,Lat}(T|\mathfrak{M})$ and obviously $X|\mathfrak{M}\in\{T|\mathfrak{M}\}'$. Again by [9] we infer $X|\mathfrak{M}\in\{T|\mathfrak{M}\}''$. From Proposition 3.4 applied to the injection $X|\mathfrak{M}$ we infer $\ker(X|\mathfrak{M})^*=\{0\}$ so that

$$(X\mathfrak{M})^- = \mathfrak{M}. \tag{3.9}$$

This shows that the mapping $\mathfrak{M}\mapsto(X\mathfrak{M})^-$ is the identity on $\operatorname{Lat}(T)$. The Proposition is proved.

Proposition 3.6. *There exist an operator T of class C_0 and operators X_n, $X\in\{T\}''$ such that $\lim_{n\to\infty}\|X_n-X\|=0$, $X\in\mathrm{F}(T)$ but $X_n\notin\mathrm{F}(T)$, $n=1, 2, \ldots$. Thus the set $\mathrm{F}(T)$ is not generally an open subset of $\{T\}'$.*

Proof. We shall construct Blaschke products m, b and b_n ($n=1, 2, \ldots$) such that

$$b\wedge m = 1, \quad b_n\wedge m \neq 1; \tag{3.10}$$

$$\lim_{n\to\infty}\|b_n-b\|_\infty = 0. \tag{3.11}$$

Then the required example is given by

$$T = S(m)\otimes I, \tag{3.12}$$

where I denotes the identity operator on an infinite dimensional Hilbert space, and

$$X = b(T), \quad X_n = b_n(T) \quad (n = 1, 2, \ldots). \tag{3.13}$$

It is clear that $T|\ker X_n$ is unitarily equivalent to $S(m\wedge b_n)\otimes I$ which is not a weak contraction and therefore $X_n\notin\mathrm{F}(T)$ (by Proposition 3.4, $X_n\notin\mathrm{sF}(T)$). Because $b\wedge m=1$, $b(T)$ is a lattice-isomorphism by Proposition 3.5, in particular $X\in\mathrm{F}(T)$. The convergence $X_n\to X$ follows from (3.11).

It remains only to construct the functions m, b and b_n ($n=1, 2, \ldots$). Let us put

$$b = \prod_{k=1}^{\infty} B^k, \quad b_n = \prod_{k=1}^{\infty} B_n^k \; (n = 1, 2, \ldots), \quad m = \prod_{k=1}^{\infty} B_k^k \tag{3.14}$$

where B^k (respectively B_n^k) is the Blaschke factor with the zero k^{-2} (respectively $k^{-2}\exp(it_n^k)$, $t_n^k>0$). Because $|b-b_n| \leqq \sum_{k=1}^{\infty} |B^k - B_n^k|$, one can verify that (3.11) holds whenever $\lim_{n\to\infty} \sum_{k=1}^{\infty} k^4 t_n^k = 0$. Conditions (3.10) are also verified and $b_n \wedge m = B_n^n$.

References

[1] H. Bercovici, On the Jordan model of C_0 operators, *Studia Math.*, **60** (1977), 267—284.

[2] 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.

[3] H. Bercovici, D. Voiculescu, Tensor operations on characteristic functions of C_0 contractions, *Acta Sci. Math.*, **39** (1977), 205—233.

[4] D. Sarason, On spectral sets having connected complement, *Acta Sci. Math.*, **26** (1966), 289—299.

[5] B. Sz.-Nagy, C. Foiaş, *Harmonic Analysis of Operators on Hilbert Space*, North-Holland/Akadémiai Kiadó (Amsterdam/Budapest, 1970).

[6] 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.

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

[8] B. Sz.-Nagy, C. Foiaş, Jordan model for contractions of class $C_{\cdot 0}$, *Acta Sci. Math.*, **36** (1974), 305—322.

[9] B. Sz.-Nagy, C. Foiaş, Commutants and bicommutants of operators of class C_0, *Acta Sci. Math.*, **38** (1976), 311—315.

[10] B. Sz.-Nagy, C. Foiaş, On injections, intertwining operators of class C_0, *Acta Sci. Math.*, **40** (1978), 163—167.

BD. BĂLCESCU 32—34
R—70 124 BUCUREŞTI 1
ROMANIA

Acta Sci. Math., 41 (1979), 29—31

Compléments à l'étude des opérateurs de classe C_0. IV

H. BERCOVICI, C. FOIAŞ, L. KÉRCHY, B. SZ.-NAGY

Dans la Note précédente [1] un rôle fondamental est joué par les deux propositions suivantes.

Proposition 1. *Pour tout opérateur T de classe C_0 dans l'espace de Hilbert $\mathfrak{H}$, les vecteurs $f \in \mathfrak{H}$ pour lesquels $m_{T_f} = m_T$, sont denses dans $\mathfrak{H}$.*

Proposition 2. *Pour tout $f \in \mathfrak{H}$ tel que $m_{T_f} = m_T$, il existe un sous-espace $\mathfrak{M}$ de $\mathfrak{H}$, invariant pour T, et une quasi-affinité $X: \mathfrak{H}(m) \to \mathfrak{H}_f$ $(m = m_T)$ tels que*

$$XS(m) = TX, \quad \mathfrak{H}_f \vee \mathfrak{M} = \mathfrak{H}, \quad X\mathfrak{H}(m) \cap \mathfrak{M} = \{0\}.$$

Rappelons que T_f désigne la restriction de T au sous-espace invariant $\mathfrak{H}_f = \bigvee_{n=0}^{\infty} T^n f$; $S(m)$ est l'opérateur défini sur l'espace fonctionnel $\mathfrak{H}(m) = H^2 \ominus mH^2$ par $S(m)u = P_{\mathfrak{H}(m)}(\lambda \cdot u)$ où $u = u(\lambda) \in \mathfrak{H}(m)$ $(|\lambda| < 1)$, et, pour tout opérateur V de classe C_0, $m_V = m_V(\lambda)$ est la fonction minimum de V.

Or, la démonstration qu'on a indiquée dans [1] pour la proposition 2 était trop sommaire, et même insuffisante.[1]) Nous allons remédier ce point et cela même en établissant le résultat plus fort suivant.

Proposition 2*. *Pour tout $f \in \mathfrak{H}$ tel que $m_{T_f} = m_T$, il existe un sous-espace $\mathfrak{M}$ de $\mathfrak{H}$, invariant pour T, tel que*

$$\mathfrak{H}_f \vee \mathfrak{M} = \mathfrak{H}, \quad \mathfrak{H}_f \cap \mathfrak{M} = \{0\}.$$

(Pour en déduire la proposition 2 il n'y a qu'à rappeler que, d'après la proposition 1 de [2], T_f est quasi-similaire à $S(m)$, avec $m = m_{T_f}$ $(= m_T)$.)

Reçu le 9 novembre 1978.

[1]) Cela a été remarqué par l'un des auteurs (L. K.) de la présente Note; il a d'abord proposé une démonstration portant sinon pour tous les f tels que $m_{T_f} = m_T$, mais du moins pour les éléments f d'un ensemble dense dans $\mathfrak{H}$ (ce qui suffit pour en conclûre aux théorèmes de [1]). Ensuite, on est parvenu à la démonstration suivante.

On commence par le suivant

Lemme. *Soient T et T' des opérateurs de classe C_0 dans les espaces $\mathfrak{H}$ et $\mathfrak{H}'$, et supposons que T et T' sont quasi-similaires et sans multiplicité* (c'est-à-dire $\mu_T = \mu_{T'} = 1$). *Alors, tout opérateur injectif $A\colon \mathfrak{H} \to \mathfrak{H}'$ tel que $AT = T'A$, est aussi quasi-surjectif* (c'est-à-dire que $\overline{A\mathfrak{H}} = \mathfrak{H}'$).

Remarque. Le lemme résulte aussi du théorème de [3]. Mais dans le cas particulier ($\mu_T = \mu_{T'} = 1$) qui nous occupe on a la démonstration suivante simple:

L'injection $A\colon \mathfrak{H} \to \mathfrak{H}'$ induit une quasi-affinité $B\colon \mathfrak{H} \to \mathfrak{L}$ où $\mathfrak{L} = \overline{A\mathfrak{H}}$ et on a évidemment $BT = (T'|\mathfrak{L})B$. Il s'ensuit que $T'|\mathfrak{L}$ et T, donc aussi $T'|\mathfrak{L}$ et T', ont la même fonction minimum. Puisque T' est sans multiplicité, on a alors par le théorème 2(iv) de [2] que $\mathfrak{L} = \mathfrak{H}'$, donc $\overline{A\mathfrak{H}} = \mathfrak{H}'$.

Le lemme établi, passons à la démonstration de la proposition 2*.

Puisque T_f a le vector cyclique f, son adjoint $(T_f)^*$ a aussi un vecteur cyclique, soit g; *cf.* le théorème 2 de [2]. Posons

$$\mathfrak{H}_{*g} = \bigvee_{n=0}^{\infty} T^{*n} g, \quad T_{*g} = (T^*|\mathfrak{H}_{*g})^*, \quad P = P_{\mathfrak{H}_{*g}}, \quad \mathfrak{M} = \mathfrak{H} \ominus \mathfrak{H}_{*g};$$

$\mathfrak{H}_{*g}$ étant invariant pour T, $\mathfrak{M}$ est invariant pour T.

De la définition il dérive aussitôt que $T_{*g}P = PTP$, donc $T_{*g}Px = PTx$ pour $x \in \mathfrak{H}_{*g}$. Or, la dernière équation est vérifiée pour $x \in \mathfrak{M}$ aussi, car on a $Py = 0$ pour tout $y \in \mathfrak{M}$. On a donc $T_{*g}P = PT$ et par conséquent

$$(1) \qquad T_{*g}^n X = X T_f^n \quad (n = 0, 1, \ldots) \quad \text{où} \quad X = P|\mathfrak{H}_f = P_{\mathfrak{H}_{*g}}|\mathfrak{H}_f,$$

et en passant aux adjoints,

$$(2) \qquad X^* T_{*g}^{*n} = T_f^{*n} X^* \quad (n = 0, 1, \ldots), \quad X^* = P_{\mathfrak{H}_f}|\mathfrak{H}_{*g}.$$

Puisque g est contenu dans $\mathfrak{H}_f \cap \mathfrak{H}_{*g}$, on a $X^* g = g$, et par (2)

$$(3) \qquad X^* T_{*g}^{*n} g = T_f^{*n} g \quad (n = 0, 1, \ldots).$$

Or, g étant cyclique pour T_f^*, (3) entraîne que $X^* \mathfrak{H}_{*g}$ est dense dans $\mathfrak{H}_f$. Il s'ensuit que X est *injectif*.

On en déduit que $\mathfrak{H}_f \cap \mathfrak{M} = \{0\}$. En effet, pour $x \in \mathfrak{H}_f \cap \mathfrak{M}$ on a $Xx = Px$ (parce que $x \in \mathfrak{H}_f$) $= 0$ (parce que $x \in \mathfrak{M}$), d'où $x = 0$ (parce que X est injectif).

Notons que de (1) il dérive $u(T_{*g})X = Xu(T_f)$ pour toute fonction $u \in H^\infty$, d'où, toujours par l'injectivité de X, il s'ensuit que $m_{T_f} | m_{T_{*g}}$. D'autre part, on a

$$m_{T_{*g}} = m_{(T^*|\mathfrak{H}_{*g})^*} = \tilde{m}_{T^*|\mathfrak{H}_{*g}} | \tilde{m}_{T^*} = m_T.$$

Puisque par hypothèse $m_{T_f}=m_T$, on conclut que $m_{T_f}=m_{T_{*g}}$. Les opérateurs T_f et T_{*g}, étant cycliques et ayant la même fonction minimum, sont quasi-similaires. Vu que X est injectif, et que $T_{*g}X=XT_f$, on a en vertu du lemme que X est aussi *quasi-surjectif*, donc $\overline{P\mathfrak{H}_f}=\mathfrak{H}_{*g}$.

Puisque $\overline{P\mathfrak{H}_f}$ est évidemment compris dans $\mathfrak{H}_f \vee \mathfrak{M}$ cela entraîne que

$$\mathfrak{H}=\mathfrak{H}_{*g}\oplus\mathfrak{M}=P\mathfrak{H}_f\vee\mathfrak{M}=\mathfrak{H}_f\vee\mathfrak{M}.$$

Cela achève la démonstration.

Ouvrages cités

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

[3] B. Sz.-Nagy—C. Foiaş, On injections, intertwining operators of class C_0, *Acta Sci. Math.*, **40** (1978), 163—167.

(H. B.)
BD. BĂLCESCU 32—34
70 124 BUCUREŞTI
ROMANIA

(L. K. ET B. SZ.-N.)
BOLYAI INTÉZET
ARADI VÉRTANUK TERE 1.
SZEGED, HONGRIE 6720

Acta Sci. Math., 41 (1979), 33—37

Essential spectrum for a Banach space operator

RICHARD BOULDIN

§ 1. Introduction

Essential spectrum has been much studied with papers [4], [5], [6], [10], [14] taking the point of view of describing the Weyl spectrum or showing when different notions of "essential spectrum" coincide. A principal result of the significant paper [8] says that if the Weyl spectrum of T coincides with the Fredholm spectrum and T is essentially normal then T is the sum of a normal operator and a compact operator. The papers [1], [2], [3], [7] develop theories such as triangular representations for nonnormal operators by using the fine structure of index theory. The purpose of this note is to show that points identified by the fine structure of index theory are either very bad or very nice. Points in the semi-Fredholm domain which satisfy a "modest" hypothesis are very nice.

Let X be a fixed Banach space. Throughout this note "operator" will mean a linear map of X into X which is defined on a vector space dense in X and has closed graph. We adopt the notation of [15], which is our basic source for the theory of closed operators on a Banach space.

For the operator T let $\operatorname{nul}(T-\lambda)$ be the dimension of the kernel of $T-\lambda$, denoted $N(T-\lambda)$, and let $\operatorname{def}(T-\lambda)$ be the codimension of the range of $T-\lambda$, denoted $R(T-\lambda)$. The operator $T-\lambda$ is semi-Fredholm provided $R(T-\lambda)$ is closed and either $\operatorname{nul}(T-\lambda)$ or $\operatorname{def}(T-\lambda)$ is finite; for such λ the index of $T-\lambda$, denoted $\operatorname{ind}(T-\lambda)$, is $\operatorname{nul}(T-\lambda)-\operatorname{def}(T-\lambda)$. The operator $T-\lambda$ is Fredholm provided $R(T-\lambda)$ is closed and both $\operatorname{nul}(T-\lambda)$ and $\operatorname{def}(T-\lambda)$ are finite.

Lemma 1. (Index Theorem) *If the operator $T-\mu$ is semi-Fredholm then there is a neighborhood of μ, say G, such that the following are true:*

(i) $\lambda\in G$ *implies* $T-\lambda$ *is semi-Fredholm with* $\operatorname{nul}(T-\lambda)\leqq\operatorname{nul}(T-\mu)$, $\operatorname{def}(T-\lambda)\leqq\operatorname{def}(T-\mu)$ *and* $\operatorname{ind}(T-\lambda)=\operatorname{ind}(T-\mu)$;

(ii) $\operatorname{nul}(T-\lambda)$ *and* $\operatorname{def}(T-\lambda)$ *are constant on* $G\setminus\{\mu\}$ *(that is* $\{z\colon z\in G, z\neq\mu\}$*);*

Received April 4, 1978.

(iii) *provided* $\operatorname{nul}(T-\mu)<\infty$, $\operatorname{nul}(T-\lambda)$ *is constant on G if and only if* $N(T-\mu)\subset \subset \bigcap\{R((T-\mu)^k)\colon k=1, 2, \ldots\}$;

(iv) *provided* $\operatorname{def}(T-\mu)<\infty$, $\operatorname{def}(T-\lambda)$ *is constant on G if and only if* $N(T'-\mu)\subset \bigcap\{R((T'-\mu)^k)\colon k=1, 2, \ldots\}$.

Proof. Parts (i) and (ii) are well known; part (iii) is Problem 5.32 of [12, p. 242] and (iv) results from applying (iii) to $(T'-\lambda)$.

The next lemma summarizes many useful facts. The spectrum of the operator T is denoted $\sigma(T)$. The dimension of the subspace X_0 in the lemma is called the algebraic multiplicity of λ.

Lemma 2. *Let T be an operator and let λ be an isolated point of $\sigma(T)$. Then there is a direct sum decomposition of X, say $X_0\oplus X_1$, such that X_0 and X_1 are invariant under $T-\lambda$. The restriction of $T-\lambda$ to X_0, denoted $(T-\lambda)|X_0$, is quasinilpotent and $(T-\lambda)|X_1$ is invertible. If $T-\lambda$ is semi-Fredholm then the dimension of X_0, denoted* $\dim X_0$, *is finite.*

Proof. There are many sources for the information about the decomposition corresponding to $\{\lambda\}$ and its complement (for example, see [12, pp. 178—181]). Since $T-\lambda$ is semi-Fredholm, it follows that $(T-\lambda)|X_0$ is semi-Fredholm. Since $R\left((T-\lambda)|X_0\right)$ is closed, $\operatorname{nul}'(T-\lambda)|X_0=\operatorname{nul}(T-\lambda)|X_0$ and $\operatorname{def}'(T-\lambda)|X_0= =\operatorname{def}(T-\lambda)|X_0$ by [12, Theorem 5.10, p. 233]. By [12, Theorem 5.30, p. 240] we know that $\dim X_0=\infty$ implies $\operatorname{nul}'(T-\lambda)|X_0=\infty$. This proves that $\dim X_0<\infty$.

§ 2. Main result

The set of points μ such that $T-\mu$ is a Fredholm operator is denoted $\Phi(T)$ and the set of λ for which $\operatorname{ind} T-\lambda$ is zero is denoted $\Phi_0(T)$. Provided there are nonnegative integers k such that $N(T^k)$ equals $N(T^{k+1})$, T is said to have finite ascent and the smallest such k is the ascent of T. Provided there are nonnegative integers m such that $R(T^m)$ equals $R(T^{m+1})$, T is said to have finite descent and the smallest such m is the descent of T.

To say that $N(T-\lambda)$ is not an asymptotic eigenspace for the operator T means that whenever there is a sequence of distinct eigenvalues say $\{\lambda_n\}$, converging to λ then $|\lambda_n-\lambda|=o\left(d(\lambda_n, \lambda)\right)$ where

$$d(\lambda_n, \lambda) = \sup\left\{\operatorname{dist}\left(x, N(T-\lambda)\right)\colon \|x\| = 1,\ x\in N(T-\lambda_n)\right\}.$$

This concept was introduced in [6].

Theorem 3. *If $T-\lambda$ is a semi-Fredholm operator with $\lambda\in\sigma(T)$ and one of the conditions* (1), (2), (3) *below holds then λ is an isolated eigenvalue with finite* *algebraic multiplicity. Furthermore, if μ is any isolated eigenvalue with finite algebraic multiplicity then μ belongs to $\Phi_0(T)$ and satisfies* (2) *and* (3).

(1) *λ is an isolated point of $\sigma(T)$.*

(2) *$N(T-\lambda)$ and $N(T'-\lambda')$ are not asymptotic eigenspaces for T and T', respectively.*

(3) *$T-\lambda$ has finite ascent and finite descent.*

Proof. First it is noted that (1) suffices for the conclusion about λ. Since $T-\lambda$ is semi-Fredholm, Lemma 2 implies that the spectral subspace X_0 corresponding to λ is finite dimensional. Consequently the quasinilpotent $(T-\lambda)|X_0$ is nilpotent, and $N((T-\lambda)|X_0)$ is non-trivial. Thus, λ is an isolated eigenvalue with finite algebraic multiplicity, and it suffices to show (1) is implied by each of the conditions (2) and (3).

Assume (2) holds and for the sake of a contradiction assume $\sigma(T)$ contains $\{\lambda_n\}$ which converges to λ with $\lambda_n \neq \lambda$. Lemma 1 shows that it may be assumed that each $T-\lambda_n$ is semi-Fredholm. Either nul $T-\lambda_n$ or def $T-\lambda_n$ is positive, and first we consider the case nul $T-\lambda_n > 0$. It will be shown that since $N(T-\lambda)$ is not an asymptotic eigenspace, $R(T-\lambda)$ is not closed, a contradiction. Since $\{|\lambda_n-\lambda|/d(\lambda_n, \lambda)\}$ converges to zero there is a sequence of unit vectors $\{x_n\}$ such that $x_n \in N(T-\lambda_n)$ and

$$\operatorname{dist}(x_n, N(T-\lambda)) > d(\lambda_n, \lambda) - |\lambda_n - \lambda|.$$

It follows that

$$\|(T-\lambda)x_n\|/\operatorname{dist}(x_n, N(T-\lambda)) = |\lambda_n-\lambda|/\operatorname{dist}(x_n, N(T-\lambda)) < |\lambda_n-\lambda|/(d(\lambda_n, \lambda) - |\lambda_n-\lambda|)$$

and clearly the last fraction converges to zero. Thus, $R(T-\lambda)$ is not closed (see Theorem 5.3, p. 72, [15]) and this contradiction proves that λ is an isolated point of $\sigma(T)$. If def $(T-\lambda_n)$ where positive then one would use that $N(T'-\lambda)$ is not an asymptotic subspace to show $R(T'-\lambda)$ to be not closed.

If (3) holds then (1) follows immediately from [13, Theorem 2.1, p. 200].

It only remains to establish the properties of the isolated eigenvalue μ. If Y_0 is the algebraic eigenspace associated with μ and Y_1 is the complementary subspace in X then $(T-\mu)|Y_1$ is one-to-one and onto. Since dim Y_0 is finite, it is straightforward to see that $(T-\mu)|Y_0$ is Fredholm with index zero, and conditions (2) and (3) must hold.

If λ belongs to $\sigma(T) \cap \Phi_0(T)$ then clearly λ is an eigenvalue for the operator T. Thus, the hypothesis of the next corollary would be stronger if one of the conditions (1), (2), (3), (4) was required for each eigenvalue λ. Hence, the hypothesis of the corollary is weaker than the hypotheses for similar results in [4], [5], [11].

Corollary 4. *Let T be an operator on X. If every λ in $\sigma(T) \cap \Phi_0(T)$ satisfies one of the conditions of Theorem* 3 *or* (4) *below then each λ is an isolated eigenvalue with finite algebraic multiplicity.*

(4) *$N(T-\lambda)$ and $N(T'-\lambda)$ are not subspaces of* $\bigcap_{k=1}^{\infty}\{R((T-\lambda)^k)\}$ *and* $\bigcap_{k=1}^{\infty}\{R((T'-\lambda)^k)\}$, *respectively.*

Proof. Let (4) hold and take $\varepsilon>0$ such that $0<|\lambda-\mu|<\varepsilon$ implies that all of the conclusions of Lemma 1 hold. Lemma 1 and condition (4) imply nul $(T-\mu)<$ $<$nul $(T-\lambda)$. If nul $(T-\mu)$ were positive then one of the conditions (1), (2), (3), (4) would apply and the resulting conclusion would contradict that nul $(T-\mu)$ is a positive constant for $0<|\lambda-\mu|<\varepsilon$; hence, nul $(T-\lambda)=0$ for such μ. Similarly, def $(T-\lambda)$ is zero for $0<|\lambda-\mu|<\varepsilon$ and so $T-\mu$ is invertible, which proves that λ is an isolated point of $\sigma(T)$.

The conditions (1), (2), (3) of Theorem 3 can be weakened provided the hypothesis for $T-\lambda$ is strengthened.

Corollary 5. *Let T be an operator on X. Every λ in $\sigma(T)\cap\Phi_0(T)$ which satisfies one of the conditions* (1′), (2′), (3′) *below is an isolated eigenvalue with finite algebraic multiplicity.*

(1′) *λ is an isolated point of $\sigma(T)$.*

(2′) *$N(T-\lambda)$ is not an asymptotic eigenspace for T.*

(3′) *$T-\lambda$ has finite ascent.*

Proof. If λ is an isolated point then Theorem 3 proves the desired conclusion.

The argument given in the second paragraph of the proof of Theorem 3 shows that (2′) above suffices.

That (3′) suffices follows from [14, Theorem 1.1].

In the final corollary the previous results are applied to get a simple alternative proof for a recent result on Riesz operators. An operator T is a *Riesz operator* provided the following hold for every nonzero λ:

(i) $T-\lambda$ has finite ascent and finite descent;

(ii) $N((T-\lambda)^k)$ is finite dimensional for $k=1,2,\ldots$;

(iii) $R((T-\lambda)^k)$ is closed with finite codimension for $k=1,2,\ldots$;

(iv) nonzero points of $\sigma(T)$ are eigenvalues and the only possible accumulation point of $\sigma(T)$ is zero.

Note that the sum of any quasinilpotent operator and a compact operator is a Riesz operator. For bounded T the next result was proved by Caradus [9, p. 42].

Corollary 6. *Let T be an operator with nonempty resolvent set. If $\Phi(T)$ contains $\{z\colon z\neq 0\}$ then T is a Riesz operator.*

Proof. The index, being locally constant, is continuous and integer valued; thus, it is constant on connected components, and $\Phi_0(T)$ contains $\{z\colon z\neq 0\}$. If $\sigma(T)\cap\{z\colon z\neq 0\}$ contains accumulation points of $\sigma(T)$ then the intersection of

$\{z: z\neq 0\}$ with the boundary of $\sigma(T)$ contains λ, an accumulation point of $\sigma(T)$. Since $\operatorname{nul}(T-z)$ is constant on $N=\{z: 0<|\lambda-z|<\varepsilon\}$ for some $\varepsilon>0$ and N intersects the resolvent set of T, it must be that $\operatorname{nul}(T-z)=0$ for $z\in N$. Since $\operatorname{ind}(T-z)=0$ for $z\in N$, λ is an isolated point and the only possible accumulation point of $\sigma(T)$ is zero. Now Corollary 5 and Theorem 3 complete the proof.

Because of the astonishing lack of examples of (unbounded) operators in the literature, we mention the following. If C is the complex plane endowed with Lebesgue measure and M_z is multiplication by the independent variable defined on $\{f(z)\in L^2(C): zf(z)\in L^2(C)\}$ then M_z is an operator with no λ such that $M_z-\lambda$ is semi-Fredholm. So the resolvent set of an operator might be empty.

Bibliography

[1] C. Apostol, The correction by compact perturbation of the singular behavior of operators, *preprint.*

[2] C. Apostol, Matrix models for operators, *Duke Math. J.*, **42** (1975), 779—785.

[3] C. Apostol and K. Clancey, Generalized inverses and spectral theory, *Trans. Amer. Math. Soc.*, **215** (1976), 293—300.

[4] S. K. Berberian, An extension of Weyl's theorem to a class of not necessarily normal operators, *Michigan Math. J.*, **16** (1969), 273—279.

[5] R. Bouldin, Essential spectrum for a Hilbert space operator, *Trans. Amer. Math. Soc.*, **163** (1972), 437—445.

[6] R. Bouldin, The Weyl essential spectrum, *Proc. Amer. Math. Soc.*, **28** (1971), 531—536.

[7] R. Bouldin, The triangular representation of C. Apostol, *Proc. Amer. Math. Soc.*, **57** (1976), 256—260.

[8] L. Brown, R. G. Douglas and P. A. Fillmore, Unitary equivalence modulo the compact operators and extensions of C* algebras, *Proceedings of a Conference in Operator Theory*, Springer-Verlag, (New York, 1973), 58—128.

[9] S. R. Caradus, W. E. Pfaffenberger and B. Yood, *Calkin Algebras and Algebras of Operators on Banach Spaces*, Marcel Dekker (New York, 1974).

[10] K. Gustafson, Necessary and sufficient conditions for Weyl's theorem, *Michigan Math. J.*, **19** (1972), 71—81.

[11] V. Istrăţescu, Weyl's theorem for a class of operators, *Rev. Roumaine Math. Pures. Appl.*, **13** (1968), 1103—1105.

[12] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag (New York, 1966).

[13] D. C. Lay, Spectral analysis using ascent, descent, nullity and defect, *Math. Ann.*, **184** (1970), 197—214.

[14] M. Schechter, On the essential spectrum of an arbitrary operator I, *J. Math. Anal. Appl.*, **13** (1966), 205—215.

[15] M. Schechter, *Principals of Functional Analysis*, Academic Press (New York, 1971).

DEPARTMENT OF MATHEMATICS
THE UNIVERSITY OF GEORGIA
ATHENS, GEORGIA 30 602
USA

Acta Sci. Math., 41 (1979), 39—45

On the lattice of congruence varieties of locally equational classes

G. CZÉDLI

1. Introduction

For a class $\mathscr{K}$ of algebras, let $\mathbf{Con}(\mathscr{K})$ denote the lattice variety generated by the class of congruence lattices of all members of $\mathscr{K}$. A lattice variety $\mathscr{U}$ will be called an *l-congruence variety* if $\mathscr{U}=\mathbf{Con}(\mathscr{K})$ for some locally equational class $\mathscr{K}$ of algebras. In particular, every congruence variety is an l-congruence variety. Our aim is to show that l-congruence varieties form a complete lattice, which is a join-subsemilattice of the lattice of all lattice varieties (while meet is not preserved). We also show that the minimal modular congruence varieties described by FREESE [1] and the minimal modular l-congruence varieties are the same.

The notion of locally equational class has been introduced by HU [2]. For the definition, let F be a subset of an algebra A of type τ and let t_1, t_2 be n-ary τ-terms. The identity $t_1=t_2$ is said to be valid in F if for all $(a_1, a_2, \ldots, a_n)\in F^n$ we have $t_1(a_1, a_2, \ldots, a_n)=t_2(a_1, a_2, \ldots, a_n)$. Suppose $\mathscr{K}$ is a class of algebras of type τ and denote by $\mathbf{L}(\mathscr{K})$ the class of all algebras A of type τ having the following property:

for each finite subset G of A there is a finite family $\{B_i: i\in I\}$ in $\mathscr{K}$ and there is for each $i\in I$ a finite subset $F_i\subseteq B_i$ such that every identity valid in F_i for all $i\in I$ is also valid in G.

Now, $\mathbf{L}$ is a closure operator on classes of similar algebras. $\mathbf{L}(\mathscr{K})$ is called the *locally equational class* (or, briefly, *local variety*) generated by $\mathscr{K}$, and $\mathscr{K}$ is said to be a local variety if $\mathbf{L}(\mathscr{K})=\mathscr{K}$. We often write $\mathbf{L}(A)$ instead of $\mathbf{L}(\{A\})$.

Denote by $\mathbf{H}, \mathbf{S}, \mathbf{P}_f, \mathbf{D}$ the operators of forming homomorphic images, subalgebras, direct products of finite families and directed unions, respectively, and let us recall

Received February 15, 1978.

Theorem 1.1. (Hu [2]) (a) *Every variety is a local variety. The converse does not hold, e.g. all torsion groups form a local variety.*

(b) *For a class* $\mathcal{K}$ *of similar algebras* $\mathbf{L}(\mathcal{K})=\mathbf{DHSP}_f(\mathcal{K})$; *consequently,*

(c) $\mathcal{K}$ *is locally equational if and only if it is closed under* $\mathbf{D}, \mathbf{H}, \mathbf{S}, \mathbf{P}_f$.

Our main tool is the following

Theorem 1.2. (Pixley [11]) *There is an algorithm which, for each lattice identity* λ *and pair of integers* $n, k \geqq 2$, *determines a strong Mal'cev condition (i.e., a finite set of equations of polynomial symbols of unspecified type)* $U_{n,k}=U_{n,k}(\lambda)$ *such that for an arbitrary algebra* A *of type* τ *the following three conditions are equivalent:*

(i) λ *is satisfied throughout* $\mathbf{Con}(\mathbf{L}(A))$;

(ii) *for each finite subset* F *of* A *and integer* $n \geqq 2$ *there is an integer* $k=k(n, F, \lambda)$ *and a* τ*-realization* $U^{\tau}_{n,k}$ *of* $U_{n,k}$ *such that* $U^{\tau}_{n,k}$ *is valid in* F;

(iii) *for each finite subset* F *of* A *and integer* $n \geqq 2$ *there is a* $k_0=k_0(n, F, \lambda)$ *such that for each* $k \geqq k_0$ *there is a* τ*-realization* $U^{\tau}_{n,k}$ *of* $U_{n,k}$ *which is valid in* F.

We have supplemented Pixley's theorem with condition (iii) which is implicit in the proof in [11] of the theorem. We shall make essential use of

Proposition 1.3. *In the above theorem each polynomial of* $U^{\tau}_{n,k}$ *is idempotent in* F.

This follows easily from the construction of $U_{n,k}$ described in [11].

2. Lattice of *l*-congruence varieties

A lattice variety $\mathcal{U}$ is called a *congruence variety* (Jónsson [8]) if $\mathcal{U}=\mathbf{Con}(\mathcal{K})$ for some variety $\mathcal{K}$, and $\mathcal{U}$ will be called an *l-congruence variety* if $\mathcal{U}=\mathbf{Con}(\mathcal{V})$ for some local variety $\mathcal{V}$. Let $\mathfrak{C}$ and $\mathfrak{C}^*$ denote the "sets" consisting of all *l*-congruence varieties and all *l*-congruence varieties of the form $\mathbf{Con}(\mathbf{L}(A))$, respectively. Let $\mathfrak{C}$ and $\mathfrak{C}^*$ be partially ordered by inclusion. Our main result is

Theorem 2.1. $\mathfrak{C}$ *is a complete lattice. The (infinitary) join of arbitrary l-congruence varieties in* $\mathfrak{C}$ *and their join taken in the lattice of all lattice varieties coincide.*

Although there exists a local variety which cannot be generated by a single algebra (Hu [2]), we have

Theorem 2.2. *For any local variety* $\mathcal{V}$ *there is an algebra* A *(*not necessarily *of the same type as* $\mathcal{V}$*) such that* $\mathbf{Con}(\mathcal{V})=\mathbf{Con}(\mathbf{L}(A))$. *Thus* $\mathfrak{C}=\mathfrak{C}^*$.

Proof of Theorems 2.1 and 2.2. First we show the following statement:

(1) For any algebra A of type τ there exists an algebra B such that $\mathbf{Con}(\mathbf{L}(A))=$ $=\mathbf{Con}(\mathbf{L}(B))$ and B has a one-element subalgebra.

Let $b_0\in A$, $\Phi=\{\lambda:\lambda$ is a lattice identity satisfied throughout $\mathbf{Con}(\mathbf{L}(A))\}$ and $H=\{F:F$ is a finite subset of A containing $b_0\}$. By Thm. 1.2 choose a $k=k(n, F, \lambda)$ and a τ-realization $U^\tau_{n,k}(F,\lambda)$ of $U_{n,k}(\lambda)$ for all $\lambda\in\Phi$, $F\in H$ and $n\geqq 2$ such that $U^\tau_{n,k}(F,\lambda)$ is valid in F. Denote by $P(n, F, \lambda)$ the set of τ-polynomials occuring in $U^\tau_{n,k}(F,\lambda)$ and define an algebra B as follows: B has the same carrier as A and the set of its operations is $\bigcup\{P(n, F, \lambda): n\geqq 2, F\in H, \lambda\in\Phi\}$ (i.e. B is a reduct of A). Since $U^\tau_{n,k}$ is also valid in $F\setminus\{b_0\}$, $\mathbf{Con}(\mathbf{L}(A))=\mathbf{Con}(\mathbf{L}(B))$ follows from Thm. 1.2. By Prop. 1.3, $\{b_0\}$ is a subalgebra of B, which completes the proof of (1).

Now we prove that

(2) For an arbitrary set Γ of indices and for any algebras $A_\gamma(\gamma\in\Gamma)$ there is an algebra A' such that $\bigvee_{\gamma\in\Gamma}\mathbf{Con}(\mathbf{L}(A_\gamma))=\mathbf{Con}(\mathbf{L}(A'))$ in the lattice of all lattice varieties.

We can assume $\Gamma\neq\emptyset$ (otherwise the statement is trivial) and

— $\{a_\gamma\}$ is a one-element subalgebra of A_γ for each $\gamma\in\Gamma$,

— all the algebras $A_\gamma(\gamma\in\Gamma)$ are of the same similarity type τ (otherwise the set of operations of A_γ can be supplemented with projections since for polynomially equivalent algebras B_1 and B_2 over the same carrier, $\mathbf{Con}(\mathbf{L}(B_1))=\mathbf{Con}(\mathbf{L}(B_2))$ by Thm. 1.2), and

— for each $\gamma\in\Gamma$, every τ-polynomial is equal to some τ-operation over A_γ.

Denote by τ_i the set of i-ary operation symbols in τ and regard $\tau_i'=\tau_i^\Gamma$ as a set of i-ary operation symbols $(i=0, 1, 2, \ldots)$. Now, $\tau=\bigcup_{i=0}^{\infty}\tau_i$ and set $\tau'=\bigcup_{i=0}^{\infty}\tau_i'$. For each $\gamma\in\Gamma$, A_γ can be regarded as an algebra A_γ' of type τ' if we define, for $q\in\tau'$, the operation q by $q=q(\gamma)$ ($q(\gamma)\in\tau$, A_γ and A_γ' have the same carrier). Evidently, $\mathbf{Con}(\mathbf{L}(A_\gamma'))=\mathbf{Con}(\mathbf{L}(A_\gamma))$ by Thm. 1.2. Let A' be a weak direct product of the algebras A_γ' defined by

$A'=\{f\in\prod_{\gamma\in\Gamma}A_\gamma':$ for all but finitely many $\gamma\in\Gamma$, $f(\gamma)=a_\gamma\}$.

By Thm. 1.1 $\mathbf{L}(A_\gamma')\subseteqq\mathbf{L}(A')$, therefore

$$\bigvee_{\gamma\in\Gamma}\mathbf{Con}(\mathbf{L}(A))=\bigvee_{\gamma\in\Gamma}\mathbf{Con}(\mathbf{L}(A_\gamma'))\subseteqq\bigvee_{\gamma\in\Gamma}\mathbf{Con}(\mathbf{L}(A'))=\mathbf{Con}(\mathbf{L}(A')).$$

In order to prove the converse inclusion by means of Thm. 1.2, suppose a lattice identity λ is satisfied throughout each $\mathbf{Con}(\mathbf{L}(A_\gamma))$. Fix an arbitrary finite subset F of A' and $n\geqq 2$. For each $\gamma\in\Gamma$ set $F_\gamma=\{f(\gamma): f\in F\}\subseteqq A_\gamma'$ and choose a non-empty finite $\Delta\subseteqq\Gamma$ such that $\gamma\in\Gamma\setminus\Delta$ implies $F_\gamma=\{a_\gamma\}$. Since λ holds in each $\mathbf{Con}(\mathbf{L}(A_\gamma))$, by Thm 1.2 for each $\gamma\in\Gamma$ there exist $k_\gamma\geqq 2$ and for all $k\geqq k_\gamma$ a τ-realization $U^\tau_{n,k}(\gamma)$ of $U_{n,k}$ such that $U^\tau_{n,k}(\gamma)$ is valid in F_γ. We can suppose $k_\gamma=2$

if $\gamma\in\Gamma\setminus\Delta$, because F_γ is a subalgebra consisting of a single element. Set $k=\max\{k_\gamma: \gamma\in\Gamma\}$. Then for each $\gamma\in\Gamma$ there exists a realization $U^\tau_{n,k}(\gamma)$ of $U_{n,k}$ which is valid in F_γ. Let $U^\tau_{n,k}(\gamma)$ consist of τ-operations $q_{1,\gamma}, q_{2,\gamma}, \ldots, q_{s,\gamma}$. For $i=1, 2, \ldots, s$ define $q_i\in\tau'$ by $q_i(\gamma)=q_{i,\gamma}$ over $A_\gamma(\gamma\in\Gamma)$. Then the operations $q_1, q_2, \ldots, q_s$ yield a τ'-realization of $U_{n,k}$ which is valid in F. This completes the proof of (2).

Now, let $\mathscr{V}$ be an arbitrary local variety and let Φ consist of all lattice identities which are not satisfied throughout $\mathbf{Con}(\mathscr{V})$. For each $\lambda\in\Phi$ we can choose $A_\lambda\in\mathscr{V}$ such that λ is not satisfied in the congruence lattice of A_λ. Since $\mathbf{L}(A_\lambda)\subseteq\mathscr{V}$ and λ is not satisfied throughout $\mathbf{Con}(\mathbf{L}(A_\lambda))$, it can be easily seen that $\mathbf{Con}(\mathscr{V})= =\bigvee_{\lambda\in\Phi}\mathbf{Con}(\mathbf{L}(A_\lambda))$. Hence Thm. 2.2 follows from (2). Since any complete join-semilattice having a 0-element is a complete lattice, Thm. 2.1 follows from (2) and Thm. 2.2. Q.E.D.

3. Minimal modular l-congruence varieties

Let P be the set of all prime numbers and set $P_0=P\cup\{0\}$. For $p\in P_0$ denote by Q_p the prime field of characteristic p and by $\mathscr{V}_p$ the variety of all vector spaces over Q_p. The following theorem was announced by FREESE [1]:

Theorem 3.1. *For any modular but not distributive congruence variety $\mathscr{U}$ there is a $p\in P_0$ such that $\mathbf{Con}(\mathscr{V}_p)\subseteq\mathscr{U}$. Consequently, congruence varieties do not form a sublattice in the lattice of all lattice varieties.*

Christian Herrmann has also proved the above theorem. We shall slightly modify his (unpublished) proof to obtain the following

Theorem 3.2. *For any modular but not distributive l-congruence variety $\mathscr{U}$ there is a $p\in P_0$ such that $\mathbf{Con}(\mathscr{V}_p)\subseteq\mathscr{U}$. Consequently, l-congruence varieties do not form a sublattice in the lattice of all lattice varieties.*

The proof is based on the following theorem (which is presented here in a weakened form):

Theorem 3.3. (HUHN [4]) *For an arbitrary modular lattice M and $n\geqq 3$ the following two conditions are equivalent:*

(i) *M is not n-distributive, i.e., the n-distributivity law*

$$x\wedge\bigvee_{i=0}^{n} y_i=\bigvee_{j=0}^{n}\Big(x\wedge\bigvee_{\substack{i=0\\ i\neq j}}^{n} y_i\Big)$$

(cf. HUHN [3] *and* [5]) *is not satisfied in M.*

(ii) *The lattice variety generated by M contains $L_{n+1}(Q_p)$ for some $p \in P_0$ where $L_{n+1}(Q_p)$ denotes the congruence lattice of the $(n+1)$-dimensional vector space over Q_p.*

For a pair of non-negative integers m, k let us define the divisibility condition $D(m, k)$ by the formula $(\exists x)\ (m \cdot x = k \cdot 1)$ where $m \cdot x$ and $k \cdot 1$ mean $x + x + \ldots + x$ (m times) and $1 + 1 + 1 + \ldots + 1$ (k times), respectively. We need the following

Proposition 3.4. *For any lattice identity λ there exist non-negative integers n_0, m, k such that for each $p \in P_0$ the following three conditions are equivalent:*
(i) *λ is satisfied throughout* $\mathbf{Con}(\mathcal{V}_p)$,
(ii) *there exists $n \geqq n_0$ such that λ is satisfied in $L_n(Q_p)$,*
(iii) *the divisibility condition $D(m, k)$ holds in Q_p.*

Proof. The equivalence of (i) and (iii) is a special case of [6, Thm. 3]. As for (ii)$\to$(i), we can argue as follows: Let us construct the identity $\hat{\lambda}$ from λ by replacing the operation symbols $\wedge$ and $\vee$ by $\cap$ and $\circ$ (composition of relations), respectively. By congruence permutability, (i) holds iff $\hat{\lambda}$ is satisfied by arbitrary congruences of any algebra in $\mathcal{V}_p$. Now, WILLE's theorem [12] (see also PIXLEY [11, Thm. 2.2]) involves implicitly that if $\hat{\lambda}$ is satisfied by certain congruences of the free $\mathcal{V}_p$-algebra of rank n_0, for some n_0 depending on $\hat{\lambda}$, then $\hat{\lambda}$ is satisfied by arbitrary congruences of any algebra in $\mathcal{V}_p$. Finally, the congruence lattice of the free $\mathcal{V}_p$-algebra of rank n_0 is a sublattice of $L_n(Q_p)$ whence $\hat{\lambda}$ is satisfied by arbitrary congruences of the free $\mathcal{V}_p$-algebra of rank n_0. Q.E.D.

It follows from a more general result of NATION [10, Thm. 2] that any n-distributive congruence variety is distributive $(n \geqq 1)$. Now we need the following generalization of this fact:

Proposition 3.5. *Let $n \geqq 1$ and $\mathcal{U}$ be an arbitrary l-congruence variety. If $\mathcal{U}$ is n-distributive, then $\mathcal{U}$ is distributive.*

Proof. Certain arguments using Mal'cev conditions for congruence varieties can easily be reformulated for l-congruence varieties. PIXLEY [11] has pointed out that JÓNSSON's criterion for congruence distributivity [7] remains valid for l-congruence varieties. Similarly, MEDERLY's criterion for n-distributivity [9, Theorem 2.1] also remains valid. Thus the have:

Proposition 3.6. *For an arbitrary algebra of type τ and $n \geqq 1$ the following two conditions are equivalent:*
(i) $\mathbf{Con}(\mathbf{L}(A))$ *is n-distributive,*
(ii) *For each finite $F \subseteq A$ there exist $k \geqq 2$ and $(n+2)$-ary τ-polynomials*

$t_0, t_1, \ldots, t_k$ *on* A *such that the identities*

$$t_0(x_0, x_1, \ldots, x_{n+1}) = x_0, \quad t_k(x_0, x_1, \ldots, x_{n+1}) = x_{n+1},$$

$$t_i(x_0, x_1, \ldots, x_n, x_0) = x_0 \quad (i = 0, 1, \ldots, k),$$

$$t_i(\underbrace{x, x, \ldots, x}_{j+1}, y, y, \ldots, y) = t_{i+1}(\underbrace{x, x, \ldots, x}_{j+1}, y, y, \ldots, y)$$

$(0 \leqq i < k,\ 0 \leqq j \leqq n$ *and* $i \equiv j \pmod{n+1})$ *are valid in* F.

Now, suppose $\mathbf{Con}(\mathbf{L}(A))$ is n-distributive for some $n \geqq 1$. Fix a finite $F \subseteq A$. Then, by Prop. 3.6, there are $k \geqq 2$ and τ-polynomials $t_0, t_1, \ldots, t_k$ satisfying the required identities in F. Define $j(-1)=0$ and for $i=0, 1, \ldots, k$, $j(i) \equiv i \pmod{n+1}$, $0 \leqq j(i) \leqq n$. Define ternary τ-polynomials $q_0, q_1, \ldots, q_{2k+2}$ as follows: $q_0(x, y, z) = x$ and for $i=0, 1, \ldots, k$

$$q_{2i+1}(x, y, z) = t_i(\underbrace{x, x, \ldots, x}_{j(i-1)+1}, y, y, \ldots, y, z)$$

and

$$q_{2i+2}(x, y, z) = t_i(\underbrace{x, x, \ldots, x}_{j(i)+1}, y, y, \ldots, y, z).$$

It is easy to check that the polynomials $q_0, q_1, \ldots, q_{2k+2}$ satisfy the equations of Prop. 3.6 (ii) in F for $(1, 2k+2)$ instead of (n, k). Hence, by Prop. 3.6, 1-distributivity — which is the usual distributivity — holds throughout $\mathbf{Con}(\mathbf{L}(A))$. Thus Thm. 2.2 completes the proof.

Proof of Theorem 3.2. Let $\mathcal{U}$ be an l-congruence variety as in the theorem. By Prop. 3.5, $\mathcal{U}$ is not distributive for $n=1, 2, 3, \ldots$. Hence, by Thm. 3.3, for each $n>2$ we can choose $p_n \in P_0$ such that $L_{n+1}(Q_{p_n}) \in \mathcal{U}$. Set $S = \{p_n : n>2\}$. If the set $\{n : n>2$ and $p_n = p_t\}$ is infinite for some t, then $\{L_{n+1}(Q_{p_n}) : p_n = p_t\}$ generates $\mathbf{Con}(\mathcal{V}_{p_t})$ by Prop. 3.4 (i, ii). Hence $\mathbf{Con}(\mathcal{V}_{p_t}) \subseteq \mathcal{U}$. Suppose $\{n : n>2$ and $p_n = p_t\}$ is finite for all $t>2$. Then it suffices to show that $\mathbf{Con}(\mathcal{V}_0)$ is a subvariety of the variety generated by $\{L_{n+1}(Q_{p_n}) : n>2\}$. Suppose λ holds in $L_{n+1}(Q_{p_n})$ for each $n>2$. For a sufficiently large t, λ holds throughout $\mathbf{Con}(\mathcal{V}_{p_n})$ for any $n \geqq t$ by Prop 3.4 (i, ii). Hence there exists an infinite $S' \subseteq S \setminus \{0\}$ such that λ holds in $\mathbf{Con}(\mathcal{V}_p)$ for each $p \in S'$. Then, by Prop. 3.4, the divisibility condition $D(m, k)$ associated with λ holds in Q_p for each $p \in S'$. Therefore, $D(m, k)$ holds in Q_0 (otherwise $m=0$ and $k \neq 0$, so each $p \in S'$ divides k). Hence, by Prop. 3.4, λ holds throughout $\mathbf{Con}(\mathcal{V}_0)$. Q.E.D.

Remark. If $\mathscr{K}$ is a class of similar algebras closed under S and $\mathbf{P}_f$ then $\mathbf{Con}(\mathscr{K})$ is an l-congruence variety, namely $\mathbf{Con}(\mathscr{K})=\mathbf{Con}(\mathbf{L}(\mathscr{K}))$.

The author would like to express his thanks to A. P. Huhn for the idea of introducing l-congruence varieties.

References

[1] R. FREESE, Minimal modular congruence varieties, *Notices Amer. Math. Soc.*, **23** (1976), No. 1, A—4.

[2] T. K. HU, Locally equational classes of universal algebras, *Chinese J. Math.* **1** (1973), 143—165.

[3] A. P. HUHN, Schwach distributive Verbände. I, *Acta Sci. Math.*, **33** (1972), 297—305.

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

[5] A. P. HUHN, n-distributivity and some questions of the equational theory of lattices, in *Contributions to universal algebra,* Colloq. Math. Soc. J. Bolyai, vol. 17, North-Holland, (1977), 167—178.

[6] G. HUTCHINSON, G. CZÉDLI, A test for identities satisfied in lattices of submodules, *Algebra Universalis,* **8** (1978), 269—309.

[7] B. JÓNSSON, Algebras whose congruence lattices are distributive, *Math. Scand.,* **21** (1967), 110—121.

[8] B. JÓNSSON, Varieties of algebras and their congruence varieties, in *Proc. of the International Congress of Math., Vancouver* (1974).

[9] P. MEDERLY, Three Mal'cev type theorems and their application, *Mat. Časopis Sloven. Akad. Vied.,* **25** (1975), 83—95.

[10] J. B. NATION, Varieties whose congruences satisfy certain lattice identities, *Algebra Universalis,* **4** (1974), 78—88.

[11] A. F. PIXLEY, Local Mal'cev conditions, *Canad. Math. Bull.,* **15** (1972), 559—568.

[12] R. WILLE, *Kongruenzklassengeometrien,* Lecture Notes in Math., vol. 113, Springer-Verlag (1970).

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1.
6720 SZEGED, HUNGARY

Acta Sci. Math., 41 (1979), 47—53

О свойстве перемешивания в смысле А. Реньи для числа положительных сумм

А. А. ДЖАМИРЗАЕВ

1. Пусть на вероятностном пространстве $\{\Omega, \mathscr{F}, P\}$ заданы

$$\xi_1, \xi_2, \ldots, \xi_n, \ldots \tag{1}$$

— последовательность независимых случайных величин (сл. вел.) с $M\xi_i=0$, $D\xi_i=1$ $(i=1, 2, \ldots)$ и $\{\nu_n\}$-последовательность положительных целочисленных сл. вел. Положим $S_n=\xi_1+\xi_2+\ldots+\xi_n$ $(n=1, 2, \ldots)$ и $F_k(x)=P\{S_k<x\}$. Через N_k^n обозначим число положительных сумм S_i из последовательности $S_{k+1}, S_{k+2}, \ldots, S_n$, где $k=0, 1, \ldots, n-1$. Также положим $N_n=N_0^n$.

Известно [3], что если к последовательности сл. вел. (1) применима центральная предельная теорема, тогда

$$\lim_{n\to\infty} P\left\{\frac{N_n}{n}<x\right\} = R(x), \tag{2}$$

где

$$R(x) = \begin{cases} 0 & \text{при} \quad x \leqq 0, \\ \dfrac{2}{\pi}\arcsin\sqrt{x} & \text{при} \quad 0 < x \leqq 1, \\ 1 & \text{при} \quad x > 1. \end{cases}$$

В данной статье доказывается, что последовательность сл. вел. $\left\{\dfrac{N_n}{n}\right\}$ обладает свойством перемешивания в смысле А. Реньи. Применяя этот факт, доказываем закон арксинуса для сумм независимых сл. вел. до случайного индекса.

Прежде чем формулировать результаты, приведем следующее определение из работы А. Реньи [4]. Будем говорить, что последовательность сл. вел. $\{\eta_n\}$, заданная на $\{\Omega, \mathscr{F}, P\}$, обладает свойством перемешивания с пре-

---

Поступило 3. VI. 1977.

дельной функцией распределения (ф. р.) $F(x)$, если для любого события $A\in\mathscr{F}$, где $P(A)>0$,

$$\lim_{n\to\infty} P\{\eta_n < x|A\} = F(x) \tag{3}$$

в каждой точке x, являющейся точкой непрерывности ф. р. $F(x)$.

Теорема 1. *Если к последовательности* (1) *применима центральная предельная теорема, то последовательность сл. вел.* $\left\{\frac{N_n}{n}\right\}$ *обладает свойством перемешивания с предельной ф. р.* $R(x)$.

Теорема 2. *Пусть к* (1) *применима центральная предельная теорема, существует последовательность положительных чисел* $\{k_n\}$ *такая, что* $k_n\to\infty$ *при* $n\to\infty$ *и*

$$\frac{\nu_n}{k_n} \xrightarrow{P} \nu_0, \tag{4}$$

где ν_0 *положительная сл. вел. Тогда последовательность сл. вел.* $\left\{\frac{N_{\nu_n}}{\nu_n}\right\}$ *обладает свойством перемешивания с предельной ф. р.* $R(x)$.

Теорема 3. *Если выполнены условия теоремы 2, то при* $n\to\infty$

$$P\left\{\frac{N_{\nu_n}}{k_n} < x\right\} \to \Psi(x) = \int_0^\infty R\left(\frac{x}{y}\right) dA(y), \tag{5}$$

где $A(x)=P\{\nu_0<x\}$.

Отметим, что (5) доказано в работе [1] при условии независимости ν_n от последовательности сл. вел. $\{\xi_n\}$.

2. Для доказательства теоремы 1 нам понадобится следующее вспомогательное предложение.

Лемма. *Пусть* $\{\zeta_n\}$ *и* $\{\eta_n\}$ *— две последовательности сл. вел. такие, что* $\zeta_n \xrightarrow{P} 0$ *при* $n\to\infty$ *и* $\{\eta_n\}$ *обладает свойством перемешивания с предельной ф. р.* $F_0(x)$. *Тогда* $\{\zeta_n+\eta_n\}$ *обладает свойством перемешивания с предельной ф. р.* $F_0(x)$.

Доказательство леммы является очевидной модификацией доказательства оригинальной леммы Крамера.

3. Доказательство теоремы 1. Воспользуемся одной теоремой А. Реньи (теорема 2 из [4]), которая утверждает что, если для любого x и при каждом k, $k=1, 2, \ldots$

$$\lim_{n\to\infty} P\left\{\frac{N_n}{n} < x,\ \frac{N_k}{k} < x\right\} = R(x)\cdot P\left\{\frac{N_k}{k} < x\right\}, \tag{6}$$

тогда $\left\{\frac{N_n}{n}\right\}$ обладает свойством перемешивания с предельной ф. р. $R(x)$.

Очевидно, что $N_n = N_k + N_k^n$ и при $n \to \infty$ $\frac{N_k}{n} \xrightarrow{P} 0$. Поэтому

$$\lim_{n\to\infty} P\left\{\frac{N_k^n}{n} < x\right\} = R(x).$$

Теперь, проследив доказательство леммы, нетрудно видеть, что из соотношения

$$\lim_{n\to\infty} P\left\{\frac{N_k^n}{n} < x,\ \frac{N_k}{k} < x\right\} = R(x) P\left\{\frac{N_k}{k} < x\right\} \tag{7}$$

следует (6). Следовательно, нам достаточно доказать, что при каждом k и для любого x, $0 < x \leqq 1$ имеет место (7).

Известно (см. [2], глава V) что

$$P\left\{\frac{N_k^n}{n} < x,\ \frac{N_k}{k} < x\right\} = \int_{-\infty}^{\infty} P\left\{\frac{N_k^n}{n} < x,\ \frac{N_k}{k} < x \,\middle|\, S_k = y\right\} dF_k(y), \tag{8}$$

где $P\{A|S_k = y\}$ — значение $P(A|S_k)$ при $S_k = y$ и $P(A|S_k)$ — условная вероятность события A относительно сл. вел. S_k.

Предварительно докажем, что для любого y, $|y| \leqq T$,

$$\lim_{n\to\infty} P\left\{\frac{N_k^n(y)}{n} < x\right\} = R(x), \tag{9}$$

где $N_k^n(y)$ — число положительных сумм из последовательности

$$y + \xi_{k+1},\ y + \xi_{k+1} + \xi_{k+2},\ \ldots,\ y + \xi_{k+1} + \ldots + \xi_n$$

и $T = T(\varepsilon)$ выбрано так, что для любого заданного $\varepsilon > 0$

$$\int_{|y| > T} dF_k(y) \leqq \varepsilon. \tag{10}$$

Для удобства записи индексов, введем сл. вел. $\eta_i = \xi_{k+i}$, $i = 1, 2, \ldots$ и обозначим через M_k^n число положительных сумм из последовательности

$$\eta_k,\ \eta_k + \eta_{k+1},\ \ldots,\ \eta_k + \eta_{k+1} + \ldots + \eta_n,$$

при $k \leqq n$; при $k > n$ положим $M_k^n = 0$. Ясно, что в силу (2), имеет место соотношение

$$\lim_{n\to\infty} P\left\{\frac{M_1^n}{n} < x\right\} = R(x). \tag{11}$$

Пусть $\bar{N}_n(y)$ — число положительных сумм из последовательности

$$y + \eta_1,\ y + \eta_1 + \eta_2,\ \ldots,\ y + \eta_1 + \ldots + \eta_n.$$

Тогда $N_k^n(y) = \bar{N}_{n-k}(y)$, так как в терминах $\{\eta_i\}$ $N_k^n(y)$ — число положительных сумм из последовательности $y+\eta_1, y+\eta_1+\eta_2, \ldots, y+\eta_1+\ldots+\eta_{n-k}$. Легко проверить, что при фиксированном k

$$\frac{\bar{N}_n(y)-N_k^n(y)}{n}=\frac{\bar{N}_n(y)-\bar{N}_{n-k}(y)}{n} \xrightarrow{P} 0, \quad n\to\infty.$$

Поэтому, чтобы показать (9), достаточно доказать, что при $n\to\infty$

$$P\left\{\frac{\bar{N}_n(y)}{n}<x\right\}\to R(x). \tag{12}$$

Отметим сначала, что имеют место следующие неравенства:

а) если $y>0$, то $M_1^n \leqq \bar{N}_n(y)$;

б) если $y\leqq 0$ то $M_1^n \geqq \bar{N}_n(y)$.

Теперь докажем (12) в отдельности для $y>0$ и $y\leqq 0$.

1) Пусть $y>0$. Введем величину $\mu=\mu(y)$ следующим образом:

$$P\{\mu=m\}=P\{y+\eta_1>0, \ldots, y+\eta_1+\ldots+\eta_{m-1}>0,\ y+\eta_1+\ldots+\eta_m\leqq 0\},$$

где $m=1, 2, \ldots$. Воспользовавшись первой теоремой работы [5], легко показать, что для любого y, $0<y\leqq T$, при $m\to\infty$ $P\{\mu=m\}\to 0$ и, следовательно, μ — собственная сл. вел., т. е. $P\{\mu=\infty\}=0$ для любого $y, 0<y\leqq T$. Поэтому, для произвольной последовательности возрастающих к бесконечности чисел m_n имеем, что

$$\frac{\mu}{m_n} \xrightarrow{P} 0 \quad \text{при} \quad n\to\infty. \tag{13}$$

Теперь заметим, что в силу определения μ, $y+\eta_1>0$, $y+\eta_1+\eta_2>0, \ldots, y+\eta_1+$ $+\ldots+\eta_{\mu-1}>0$ и $y+\eta_1+\ldots+\eta_\mu\leqq 0$, откуда при $\mu\leqq n$ имеем, что $\bar{N}_n(y)=$ $=(\mu-1)+\bar{N}_{\mu+1}^n(y)$, где $\bar{N}_k^n(y)$ — число положительных сумм в последовательности

$$y+\eta_1+\ldots+\eta_k,\ y+\eta_1+\ldots+\eta_k+\eta_{k+1}, \ldots, y+\eta_1+\ldots+\eta_n,$$

при $k\leqq n$ и $\bar{N}_k^n(y)=0$ при $k>n$.

При $\mu\leqq n$ будем сравнивать $M_{\mu+1}^n$ и $\bar{N}_{\mu+1}^n(y)$, т. е., соответственно число положительных членов последовательностей $\eta_{\mu+1}, \eta_{\mu+1}+\eta_{\mu+2}, \ldots, \eta_{\mu+1}+\ldots$ $\ldots+\eta_n$ и $y+\eta_1+\ldots+\eta_\mu+\eta_{\mu+1}, \ldots, y+\eta_1+\eta_2+\ldots+\eta_\mu+\eta_{\mu+1}+\ldots+\eta_n$. Тогда, в силу того, что $y+\eta_1+\ldots+\eta_\mu\leqq 0$, имеем

$$\bar{N}_{\mu+1}^n(y)\leqq M_{\mu+1}^n.$$

Итак, если $\mu\leqq n$, то для любого y, $0<y\leqq T$,

$$\bar{N}_n(y)\leqq(\mu-1)+M_{\mu+1}^n. \tag{14}$$

Если же $\mu > n$, то (14) очевидно, так как всегда $\overline{N}_n(y) \leqq n$. Теперь из а) и (14) имеем, что

$$M_1^n \leqq \overline{N}_n(y) \leqq (\mu-1)+M_{\mu+1}^n, \tag{15}$$

откуда для любого x

$$P\left\{\frac{\mu-1}{n}+\frac{M_{\mu+1}^n}{n}<x\right\} \leqq P\left\{\frac{\overline{N}_n(y)}{n}<x\right\} \leqq P\left\{\frac{M_1^n}{n}<x\right\}. \tag{16}$$

В силу определения μ события $\{\mu=m\}$ и $\left\{\frac{M_{m+1}^n}{n}<x\right\}$ независимые. Поэтому

$$P\left\{\frac{M_{\mu+1}^n}{n}<x\right\} = \sum_{m=1}^{\infty} P\left\{\frac{M_{m+1}^n}{n}<x\right\} P\{\mu=m\}.$$

Так как сл. вел. μ независит от n, то для любого заданного $\delta>0$ и для всех y, $0<y\leqq T$, можно выбрать целое $T_1=T_1(\delta)$ так, чтобы $P\{\mu>T_1\}\leqq\delta$. Тогда

$$P\left\{\frac{M_{\mu+1}^n}{n}<x\right\} = \sum_{m=1}^{T_1} P\left\{\frac{M_{m+1}^n}{n}<x\right\} P\{\mu=m\}+P_n, \tag{17}$$

где $P_n<\delta$. Теперь для любого фиксированного T_1 нетрудно проверить, что при $m=1, 2, \ldots, T_1$

$$\lim_{n\to\infty} P\left\{\frac{M_{m+1}^n}{n}<x\right\} = R(x). \tag{18}$$

В силу произвольности $\delta>0$ из (17) и (18) следует, что для любого y, $0<y\leqq T$,

$$\lim_{n\to\infty} P\left\{\frac{M_{\mu+1}^n}{n}<x\right\} = R(x). \tag{19}$$

Принимая во внимание (13), из (19) имеем

$$\lim_{n\to\infty} P\left\{\frac{\mu-1}{n}+\frac{M_{\mu+1}^n}{n}<x\right\} = R(x). \tag{20}$$

Из (11), (16) и (20) получаем, что для любого y, $0<y\leqq T$, имеет место (12) и, следовательно, (9).

2) Пусть $y\leqq 0$. В этом случае величину $\mu=\mu(y)$ введем следующим образом. Для $m=1, 2, \ldots$ положим

$$P\{\mu=m\} = P\{y+\eta_1\leqq 0, \ldots, y+\eta_1+\ldots+\eta_{m-1}\leqq 0,\ y+\eta_1+\ldots+\eta_m>0\}.$$

Опять нетрудно проверить, что μ — собственная сл. вел. для каждого y, $-T\leqq$ $\leqq y\leqq 0$. В этом случае вместо неравенства (15) будем имееть, неравенство

$$M_{\mu+1}^n \leqq \overline{N}_n(y) \leqq M_1^n,$$

причем анологично случаю $y>0$, доказывается, что

$$P\left\{\frac{M^n_{\mu+1}}{n}<x\right\}\to R(x),\quad n\to\infty.$$

Далее, тем же способом, что и при $y>0$ получим доказательство (12) и, следовательно, (9) для случая $y\leqq 0$.

Теперь снова вернемся к соотношению (8). Известно [2] (см. глава V, § 3), что подинтегральное выражение (которое мы обозначим через $P_n(x, y)$) в (8) можно написать в следующем виде

$$P_n(x, y)=\lim_{h\to+0} P\left\{\frac{N^n_k}{n}<x,\ \frac{N_k}{k}<x|y\leqq S_k<y+h\right\}. \tag{21}$$

Легко видеть, что при условии $\{y\leqq S_k<y+h\}$, $h>0$,

$$N^n_k(y)\leqq N^n_k\leqq N^n_k(y+h).$$

Поэтому

$$\begin{aligned}P\left\{\frac{N^n_k(y+h)}{n}<x,\ \frac{N_k}{k}<x|y\leqq S_k<y+h\right\}\leqq \bar P_n\leqq\\ \leqq P\left\{\frac{N^n_k(y)}{n}<x,\ \frac{N_k}{k}<x|y\leqq S_k<y+h\right\},\end{aligned} \tag{22}$$

где

$$\bar P_n=P\left\{\frac{N^n_k}{n}<x,\ \frac{N_k}{k}<x|y\leqq S_k<y+h\right\}.$$

Используя независимость $N^n_k(y)$ от $\xi_1, \xi_2, \ldots, \xi_k$, (22) перепишем в следующем виде:

$$\begin{aligned}P\left\{\frac{N^n_k(y+h)}{n}<x\right\}\cdot P\left\{\frac{N_k}{k}<x|y\leqq S_k<y+h\right\}\leqq \bar P_n\leqq\\ \leqq P\left\{\frac{N^n_k(y)}{n}<x\right\}\cdot P\left\{\frac{N_k}{k}<x|y\leqq S_k<y+h\right\}.\end{aligned} \tag{23}$$

В неравенстве (23) переходим к пределу сначала по h, потом по n и интегрируем по y от $-T$ до T. Тогда, при помощи (21), (9) и теоремы Лебега, а также учитывая монотонность $N^n_k(y)$ по y, имеем

$$\lim_{n\to\infty}\int_{-T}^{T} P_n(x, y)\,dF_k(y)=R(x)\int_{-T}^{T} P\left\{\frac{N_k}{k}<x|S_k=y\right\}dF_k(y).$$

Отсюда, принимая во внимание (8) и (10) получаем, что

$$\lim_{n\to\infty} P\left\{\frac{N_k^n}{n} < x,\ \frac{N_k}{k} < x\right\} = R(x)P\left\{\frac{N_k}{k} < x\right\} + \Delta,$$

где $|\Delta| < 2\varepsilon$. В силу произвольности $\varepsilon > 0$ получаем (7). Теорема 1 доказана.

Доказательство теоремы 2 следует из теоремы 1 и одного результата Ш. Чёргё ([7], теорема 1).

Доказательство теоремы 3 следует из теоремы 2 и одного результата Й. Модьороди ([6], теорема 1) в силу замечания 2 и следствия 1 Ш. Чёргё в [7].

Автор выражает глубокую благодарность проф. Й. Модьороди за ценные советы и внимание оказанное при выполнении настоящей работы.

Литература

[1] М. Гофман, Э. А. Даниелян, К закону арксинуса для сумм независимых случайных величин до случайного индекса, Сб. »*Предельные теоремы для случайных процессов*«, Ташкент, Изд-во »Фан« УзССР (1977).

[2] А. Н. Колмогоров, *Основные понятия теории вероятностей*, Изд-во »Наука« (Москва, 1974).

[3] P. Erdős and M. Kac, On the number of positive sums of independent random variables, *Bull. Amer. Math. Soc.*, **53** (1947), 1011—1020.

[4] A. Rényi, On mixing sequences of sets, *Acta Math. Acad. Sci. Hung.* **9** (1958), 215—228.

[5] P. Erdős and M. Kac, On certain limit theorems of the theory of probability, *Bull. Amer. Math. Soc.*, **52** (1946), 292—302.

[6] J. Mogyoródi, A remark on the stable sequences of random variables and limit distribution theorem for the sum of independent random variables, *Acta Math. Acad. Sci. Hung.*, **17** (1966), 401—409.

[7] S. Csörgő, On limit distributions of sequences of random variables with random indices, *Acta Math. Acad. Sci. Hung.*, **25** (1974), 227—232.

ТАШКЕНТСКИЙ ГОС. УНИВЕРСИТЕТ
МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ
ТАШКЕНТ, ГСП

Acta Sci. Math., 41 (1979), 55—62

Extensions of Lomonosov's invariant subspace theorem

C. K. FONG, E. A. NORDGREN, M. RADJABALIPOUR,
H. RADJAVI and P. ROSENTHAL

1. Introduction

The famous invariant subspace theorem of V. LOMONOSOV [9] includes the assertion that each algebra of operators on a Banach space which commutes with a nonzero compact operator has a nontrivial invariant subspace. That is, if K is a compact operator other than 0, and if $AK=KA$ for all A in some algebra $\mathcal{A}$, then $\mathcal{A}$ has an invariant subspace. In [10] it was shown that this could be generalized, in the case where K is injective and $\mathcal{A}$ is uniformly closed, to the same conclusion under the assumption that $\mathcal{A}K\subset K\mathcal{A}$ (in the sense that $A\in\mathcal{A}$ implies that $AK=KA_1$ for some $A_1\in\mathcal{A}$). In [12] it was shown that the hypothesis that K be injective is not needed.

In the present note we prove that $\mathcal{A}$ uniformly closed and $\mathcal{A}K_1\subset K_2\mathcal{A}$, for K_1 and K_2 compact and nonzero, implies $\mathcal{A}$ has an invariant subspace (Theorem 3) and the commutant of $\mathcal{A}$ has an invariant subspace (Theorem 4). In fact, we obtain results slightly more general than this. The proofs presented are considerably simpler than those in [10] and [12].

Our work is merely a perturbation of LOMONOSOV's [9]; it relies on the following lemma.

Lomonosov's Lemma. ([9], [13, p. 156], [11]) *If $\mathcal{A}$ is an algebra of bounded operators on a Banach space which has no nontrivial invariant subspace, and if K is any nonzero compact operator, then there is a vector $x\neq 0$ and an A in $\mathcal{A}$ such that $AKx=x$.*

Received July 20, 1978.

2. Preliminary results: An operator equation and operator ranges

We need to consider maps which may be nonlinear, but which are bounded in a certain sense.

Definition. A function S taking a Banach space $\mathfrak{X}$ into a Banach space $\mathfrak{Y}$ is a *bounded map* if there is a constant M such that $\|Sx\| \leqq M\|x\|$ for all $x \in \mathfrak{X}$; a *bounded operator* is a bounded map which is linear.

Note that a nonlinear bounded map need not be continuous.

The next lemma is implicit in [10]. We are grateful to Ivan Kupka for providing a suggestion which led to the simpler proof given below.

Lemma 1. *Suppose that S is a bounded map taking $\mathfrak{X}$ into itself, K is a bounded linear operator on $\mathfrak{Y}$ with spectral radius $r(K)$, and T is a bounded linear operator taking $\mathfrak{X}$ into $\mathfrak{Y}$. If $T=KTS$, if $\varepsilon>0$, and if $\|Sx\| \leqq (r(K)+\varepsilon)^{-1}\|x\|$ for all $x \in \mathfrak{X}$ then $T=0$.*

Proof. Fix $x \in \mathfrak{X}$. For each positive integer n, $Tx=K^nTS^nx$ (just keep applying K and S on the left and the right, respectively). Thus, for all n,

$$\|Tx\| \leqq \|K^n\|\|T\|r((K)+\varepsilon)^{-n}\|x\|.$$

Given any $\delta>0$, $\|K^n\|^{1/n}<r(K)+\delta$ for n sufficiently large. For sufficiently large n, then,

$$\|Tx\| \leqq (r(K)+\delta)^n\|T\|(r(K)+\varepsilon)^{-n}\|x\|.$$

If $\delta<\varepsilon$, then $\left\{\left(\dfrac{r(K)+\delta}{r(K)+\varepsilon}\right)^n\right\} \to 0$ as $n\to\infty$, so $Tx=0$.

Recall that a *Riesz operator* is an operator with spectral properties like those of a compact operator; i.e., a Riesz operator is a noninvertible operator whose nonzero spectrum consists of eigenvalues of finite multiplicity with no limit points other than 0.

Definition. The operator K is *decomposable at* 0 if for each $\varepsilon>0$ there is an invariant subspace $\mathfrak{M}\neq\{0\}$ of K which has an invariant complement and is such that the spectral radius of the restriction of K to $\mathfrak{M}$ is less than ε.

Theorem 1. *If $T=KTS$, where S is a bounded map on $\mathfrak{X}$, K is a bounded operator on $\mathfrak{Y}$ and T is a bounded operator taking $\mathfrak{X}$ into $\mathfrak{Y}$, then*

(i) *K quasinilpotent implies $T=0$;*

(ii) *K a Riesz operator implies T has finite rank;*

(iii) *K decomposable at 0 implies the range of T is not dense.*

Proof of (i): For ε sufficiently small and positive, $\|Sx\| \leqq \varepsilon^{-1}\|x\|$, so the result follows immediately from Lemma 1.

Proof of (ii): Choose ε sufficiently small so that $\|Sx\| \leqq (2\varepsilon)^{-1}\|x\|$ for all x. Then the Riesz functional calculus yields an idempotent P which commutes with K such that the spectral radius of PK is less than ε. From $T=KTS$ it follows that $PT=PKTS=(PK)(PT)S$, so Lemma 1 implies that $PT=0$. Hence $T=(1-P)T$, and the range of T is contained in the range of the finite-rank operator $1-P$.

Proof of (iii): Begin as in (ii) above; get P by the assumption of decomposability at 0. Then $T=(1-P)T$, and the range of T is contained in the range of $1-P$ and thus is not dense.

HALMOS and DOUGLAS showed (see [4]) that if A and B are operators on Hilbert space, and if the range of A is contained in the range of B, then $A=BS$ for some operator S. This result is false, in general, on Banach spaces (cf. [5]), unless B is injective. We note that the result is true in general if we do not require S to be linear.

Lemma 2. *Let A be a bounded operator taking $\mathfrak{X}$ into $\mathfrak{Y}$ and B a bounded operator taking $\mathfrak{Z}$ into $\mathfrak{Y}$. If the range of A is contained in the range of B, then there is a bounded mapping S from $\mathfrak{X}$ into $\mathfrak{Z}$ such that $A=BS$.*

Proof. Let $\ker B=\{z\in\mathfrak{Z}\colon Bz=0\}$. Define $\hat{B}\colon(\mathfrak{Z}/\ker B)\to\mathfrak{Y}$ by

$$\hat{B}(z+\ker B) = Bz;$$

then $\hat{B}$ is an injective bounded operator. Now $\hat{B}^{-1}A\colon\mathfrak{Y}\to\mathfrak{Z}/\ker B$ is trivially seen to be a closed operator, so the closed graph theorem implies that $\hat{B}^{-1}A=\hat{S}$ for some bounded operator $\hat{S}\colon\mathfrak{X}\to\mathfrak{Z}/\ker B$. Then $A=\hat{B}\hat{S}$. Define the map $S\colon\mathfrak{X}\to\mathfrak{Z}$ by letting, for each $x\in\mathfrak{X}$, Sx be any element in $\hat{S}x$ of norm at most $\|\hat{S}x\|+\|x\|$; the definition of the norm on a quotient space implies that such an Sx exists. Then $\|Sx\|\leqq(\|\hat{S}\|+1)\|x\|$. Also $A=BS$, for if $x\in\mathfrak{X}$, then $Ax=\hat{B}\hat{S}x=Bz$ for any $z\in\hat{S}x$. Since Sx is such a z, $Ax=BSx$, and the lemma is proven.

Definition. A linear manifold $\mathfrak{M}$ in a Banach space $\mathfrak{X}$ is an *operator range* if there is a Banach space $\mathfrak{Y}$ and a bounded operator $T\colon\mathfrak{Y}\to\mathfrak{X}$ such that $T(\mathfrak{Y})=\mathfrak{M}$.

A comprehensive treatment of operator ranges in Hilbert space is given in [6]. GRABINER [7] contains some results about operator ranges in Banach spaces, including part (i) of the next theorem (with a proof different from ours).

Theorem 2. *If $\mathfrak{M}$ is an operator range in $\mathfrak{Y}$, and if K is a bounded operator on $\mathfrak{Y}$ such that $\mathfrak{M}\subset K\mathfrak{M}$, then*

(i) ([7]) *K quasinilpotent implies $\mathfrak{M}=\{0\}$;*

(ii) *K a Riesz operator implies $\mathfrak{M}$ is finite-dimensional;*

(iii) *K decomposable at 0 implies $\mathfrak{M}$ is not dense.*

Proof. Suppose that $T: \mathfrak{X} \to \mathfrak{Y}$ and $T(\mathfrak{X}) \doteq \mathfrak{M}$. Then the range of T is contained in the range of KT, so Lemma 2 implies that $T=KTS$ for some bounded map S. Now parts (i), (ii) and (iii) of this theorem follow from the corresponding parts of Theorem 1.

3. Invariant subspaces for certain operator algebras

If $\mathscr{A}$ is an algebra of operators contained in the commutant of a compact operator K, then the closure of $\mathscr{A}$ in any of the standard operator topologies is also contained in the commutant of K. Thus no closure assumption on such an $\mathscr{A}$ will be helpful in obtaining invariant subspaces. In the case where $\mathscr{A}$ merely intertwines a compact operator some closure assumption is essential (cf. remark (iii), p. 118 of [10]). For certain applications discussed below, however, we need to include cases where $\mathscr{A}$ is not closed even in the norm topology. It turns out to be sufficient that $\mathscr{A}$ be an operator range, in the sense that there is a bounded linear operator taking some Banach space into the space of operators such that the range of T is $\mathscr{A}$. (If $\mathscr{A}$ is uniformly closed it is an operator range; it is the range of the injection of $\mathscr{A}$ into the space of operators.)

Theorem 3. *If $\mathscr{A}$ is an algebra of operators and $\mathscr{A}$ is an operator range, and if there exist a nonzero compact operator K_1 and an operator K_2 which is decomposable at 0 such that $\mathscr{A}K_1 \subset K_2\mathscr{A}$, then $\mathscr{A}$ has a nontrivial invariant subspace.*

Proof. If $\mathscr{A}$ had no invariant subspaces, then Lomonosov's Lemma would imply that $A_0K_1x=x$ for some $A_0 \in \mathscr{A}$ and some $x \neq 0$. Now $\mathscr{A}=S\mathfrak{Y}$ for some Banach space $\mathfrak{Y}$. Define $Ty=(Sy)(x)$ for each $y \in \mathfrak{Y}$. Then the range of $\mathfrak{T}$ is $\mathscr{A}x= =\{Ax: A \in \mathscr{A}\}$, so $\mathscr{A}x$ is an operator range. If $\mathscr{A}x=\{0\}$ then the one-dimensional space spanned by x is invariant under $\mathscr{A}$. If $\mathscr{A}x \neq \{0\}$ then $\mathscr{A}x$ is an operator range invariant under $\mathscr{A}$. For $A \in \mathscr{A}$,

$$Ax = AA_0K_1x = K_2A_2x \quad \text{for some} \quad A_2 \in \mathscr{A}.$$

Hence $\mathscr{A}x \subset K_2\mathscr{A}x$. Thus part (iii) of Theorem 2 implies $\mathscr{A}x$ is not dense, so its closure is a proper invariant subspace for $\mathscr{A}$.

Remark. If K_2 is compact then the linear manifold $\mathscr{A}x$ occurring in the proof of Theorem 3 is finite-dimensional. This does not prove, however, the obviously false assertion that the hypotheses of Theorem 3 and the additional requirement that K_2 be compact yield a finite-dimensional invariant subspace for $\mathscr{A}$. We get the finite-dimensional subspace $\mathscr{A}x$ via Lomonosov's Lemma, on the assumption that we have no invariant subspaces at all.

On the other hand, if $\mathscr{A}$ is any algebra of operators with a finite-dimensional invariant subspace $\mathfrak{M}$, then $\mathfrak{M}$ could arise from Theorem 3. For let $\mathscr{A}_0$ be the set of all operators leaving $\mathfrak{M}$ invariant and let P denote an idempotent with range $\mathfrak{M}$. Then $\mathscr{A}_0 P \subset P\mathscr{A}_0$, so Theorem 3 applies to $\mathscr{A}_0$ (with $K_1 = K_2 = P$). An invariant subspace for $\mathscr{A}_0$ is also invariant under its subalgebra $\mathscr{A}$. In particular, the answer to question 1 of [12] is "no"; $\mathscr{A}_0$ is a counter-example.

Theorem 4. *If $\mathscr{A}$ is an algebra of operators which is an operator range, if there exist compact operators K_1 and K_2 different from 0 such that $\mathscr{A}K_1 \subset K_2\mathscr{A}$, and if $\mathscr{A}$ contains an operator which is not a multiple of the identity, then the commutant of $\mathscr{A}$ has a nontrivial invariant subspace.*

Proof. If the commutant of $\mathscr{A}$ had no invariant subspace then Lomonosov's Lemma would imply that there exists a B commuting with $\mathscr{A}$ and an $x \neq 0$ such that $BK_1x = x$. For A in $\mathscr{A}$, then,

$$Ax = ABK_1x = BAK_1x = (BK_2)A_1x$$

for some $A_1 \in \mathscr{A}$. Thus the linear manifold $\mathscr{A}x$ satisfies $\mathscr{A}x \subset (BK_2)(\mathscr{A}x)$. Part (ii) of Theorem 2 above implies that $\mathscr{A}x$ is finite-dimensional, (since BK_2 is compact). Choose an A_0 in $\mathscr{A}$ which is not a multiple of the identity. Since A_0 has the finite-dimensional invariant subspace $\mathscr{A}x$, A_0 has a nontrivial eigenspace (if $\mathscr{A}x = \{0\}$, then A_0 has nullspace). Since an eigenspace of A_0 is invariant under all operators commuting with A_0, the commutant of $\mathscr{A}$ has a nontrivial invariant subspace.

Corollary 1. *If A is an operator for which there exist a bounded open set D containing $\sigma(A)$, an analytic function φ taking D into D and a nonzero compact operator K such that $AK = K\varphi(A)$, then A has a nontrivial hyperinvariant subspace (unless A is a multiple of the identity).*

Proof. Let $H^\infty(D)$ denote the Banach algebra of all bounded analytic functions on D, with supremum norm, and let

$$\mathscr{A} = \{f(A)\colon f \in H^\infty(D)\}.$$

Choose a fixed Cauchy domain S contained in D and containing $\sigma(A)$. Then for $f \in H^\infty(D)$

$$\|f(A)\| = \frac{1}{2\pi}\Big\| \int_{\partial S} f(z)(z-A)^{-1}\,dz \Big\| \leqq$$

$$\leqq \frac{1}{2\pi}\cdot(\text{length of } \partial S)\cdot\|f\|_\infty \cdot \sup_{z\in\partial S}\|(z-A)^{-1}\|.$$

Hence there is a constant M such that $\|f(A)\| \leqq M\|f\|_\infty$ for $f \in H^\infty(D)$, and it follows that $\mathscr{A}$ is the range of the operator $f \to f(A)$ (that $\mathscr{A}$ is an algebra follows from the fact that this map is an algebraic homomorphism).

Also, if $f \in H^\infty(D)$ then $f(A)K = Kf(\varphi(A))$. One way to verify this is to note that, regarded as operators on $\mathfrak{X} \oplus \mathfrak{X}$, $\begin{pmatrix} 0 & K \\ 0 & 0 \end{pmatrix}$ commutes with $\begin{pmatrix} A & 0 \\ 0 & \varphi(A) \end{pmatrix}$, and hence with $f\left(\begin{pmatrix} A & 0 \\ 0 & \varphi(A) \end{pmatrix}\right) = \begin{pmatrix} f(A) & 0 \\ 0 & f(\varphi(A)) \end{pmatrix}$. Now $f(\varphi(A)) = (f\circ\varphi)(A)$ is again in $\mathscr{A}$, so Theorem 4 applies.

Corollary 2. *If A is power bounded (i.e., there exists a constant M such that $\|A^n\| \leqq M$ for all positive integers n), and if there exist an integer k and a nonzero compact operator K such that $AK = KA^k$, then A has a nontrivial hyperinvariant subspace (or is a multiple of the identity).*

Proof. Let $\mathscr{A} = \left\{\sum_{n=0}^{\infty} a_n A^n \colon \sum_{n=0}^{\infty} |a_n| < \infty\right\}$. The fact that A is power bounded implies that the map of $\{a_n\}$ into $\sum_{n=0}^{\infty} a_n A^n$ is a continuous map of l^1 into the bounded operators, so $\mathscr{A}$ is an operator range. Note that $\mathscr{A}$ is an algebra, since l^1 is an algebra under convolution. Now $AK = KA^k$ yields $A^n K = KA^{nk}$ for all n, so $\left(\sum_{n=0}^{\infty} a_n A^n\right) K = K\left(\sum_{n=0}^{\infty} a_n A^{nk}\right)$ and Theorem 4 applies.

Note. Corollary 2 follows from Corollary 1 only under the additional assumption that $\sigma(A) \subset \{z\colon |z| < 1\}$, in which case the function $\varphi(z) = z^k$ will serve.

Examples. The hypotheses of Corollary 2 hold under various circumstances.

(i) Let $\{e_n\}_{n=0}^{\infty}$ be an orthonormal basis for a Hilbert space H and let $\{k_n\}$ be a sequence converging to 0. If λ is a complex number of modulus 1 and A is defined by $Ae_n = \lambda^{2^n} e_n$, then $AK = KA^2$ where K is the compact weighted shift defined by $Ke_n = k_n e_{n+1}$. Then the unitary operator A satisfies the hypotheses of Corollary 2.

(ii) Let K_0 be a compact operator and B and C be power bounded operators such that $BK_0 = K_0 C^2$. If A is the operator $B \oplus C$ and K is the operator on $\mathfrak{X} \oplus \mathfrak{X}$ defined by $K(x_1 \oplus x_2) = K_0 x_2 \oplus 0$, then $AK = KA^2$, and A satisfies the hypotheses of Corollary 2.

A natural question is whether Theorems 3 and 4 hold if the intertwining takes place on the other side; i.e., if $K_1 \mathscr{A} \subset \mathscr{A} K_2$. Upon reading a preliminary version of this manuscript L. G. Brown discovered the following two theorems. We are grateful to him for permission to include them here. These results were also obtained independently by S. Grabiner [14].

Theorem 5. *If $\mathscr{A}$ is an algebra of operators and $\mathscr{A}$ is an operator range, and if there exist a nonzero compact operator K_1 and an operator K_2 that is decomposable at 0 such that $K_1 \mathscr{A} \subset \mathscr{A} K_2$, then $\mathscr{A}$ has a nontrivial invariant subspace.*

Proof. If we suppose $\mathscr{A}$ has no invariant subspace, then, as in the proof of Theorem 3, Lomonosov's Lemma produces an A_0 in $\mathscr{A}$ with $1 \in \sigma(K_1 A_0)$. Hence $1 \in \sigma(A_0 K_1)$, and taking Banach space adjoints yields $1 \in \sigma(A_0^* K_1^*)$. Note that $\mathscr{A}^*=\{A^*\colon A\in\mathscr{A}\}$ is also an operator range, K_1^* is compact, K_2^* is decomposible at 0 and $\mathscr{A}^*K_1^*\subset K_2^*\mathscr{A}^*$. It follows as in the proof of Theorem 3, that there is a nonzero vector x^* in $\mathfrak{X}^*$ such that $\mathscr{A}^*x^*$ is not dense in $\mathfrak{X}^*$. In fact an examination of the proofs of Theorems 1 and 2 reveals that there is a nontrivial projection P on $\mathfrak{X}$ such that $\mathscr{A}^*x^*$ is included in the range of $1-P^*$. Since that range is weak* closed as well as nontrivial, there is a nonzero vector x in $\mathfrak{X}$ that annihilates $\mathscr{A}^*x^*$. Hence either $\mathscr{A}x=\{0\}$, in which case x spans a one dimensional invariant subspace of $\mathscr{A}$, or else the closure of $\mathscr{A}x$ is a proper invariant subspace of $\mathscr{A}$. The contradiction of the original supposition establishes the result.

Theorem 6. *If $\mathscr{A}$ is an algebra of operators wich is an operator range, if there exist compact operators K_1 and K_2 different from 0 such that $K_1\mathscr{A}\subset\mathscr{A}K_2$, and if $\mathscr{A}$ contains an operator that is not a multiple of the identity, then the commutant of $\mathscr{A}$ has a nontrivial invariant subspace.*

Proof. Suppose the commutant of $\mathscr{A}$ has no invariant subspace. Then Lomonosov's Lemma implies the existence of a B commuting with $\mathscr{A}$ such that $1\in\sigma(K_1B)$, and hence $1\in\sigma(B^*K_1^*)$. As in the proof of Theorem 4, there exists a nonzero vector x^* in $\mathfrak{X}^*$ such that $\mathscr{A}^*x^*$ is finite dimensional.

Choose an A_0 in $\mathscr{A}$ that is not a multiple of the identity. Either $\mathscr{A}^*x^*=\{0\}$, in wich case A_0^* has a nontrivial null space, or else $\mathscr{A}^*x^*$ is a finite dimensional invariant subspace of A_0^*. In either event A_0^* has an eigenvector. If λ is the corresponding eigenvalue, then it follovs that the closure of the range of $A_0-\lambda$ is a nontrivial subspace of $\mathfrak{X}$ wich is invariant under the commutant of $\mathscr{A}$.

It might be worth noting that the compactness assumption on K_1 in Theorem 3 can be replaced by the hypothesis that K_1 has nonzero eigenvalues.

Theorem 7. *If $\mathscr{A}$ is an algebra of operators which is an operator range, if $\mathscr{A}K_1\subset K_2\mathscr{A}$ where K_2 is decomposable at 0 and K_1 has a nonzero eigenvalue, then $\mathscr{A}$ has a nontrivial invariant subspace.*

Proof. If $K_1x_0=\lambda x_0$ with $x_0\neq 0$ and $\lambda\neq 0$, then, for any $A\in\mathscr{A}$,

$$Ax_0=\lambda^{-1}AK_1x_0=\lambda^{-1}K_1A_1x_0 \quad \text{for some} \quad A_1\in\mathscr{A}.$$

Thus $\mathscr{A}x_0$ is contained in $K_1(\mathscr{A}x_0)$, so part (iii) of Theorem 3 implies $\mathscr{A}x_0$ is not dense.

Remarks. It is shown in [15] that there is an operator that does not satisfy the hypothesis of Lomonosov's invariant subspace theorem. In light of Theorem 4 above we can ask: if B is an operator on a Hilbert space must B commute with some uniformly closed algebra $\mathscr{A}$ (containing operators other than scalars) which intertwines two nonzero compact operators?

In [10] the following question was raised. If $\mathscr{A}$ is a uniformly closed algebra of operators such that $\mathscr{A}K \subset KB(\mathfrak{X})$ must $\mathscr{A}$ have a nontrivial invariant subspace? If $\mathscr{A}$ is not required to be closed but is merely required to be an operator range then the answer is no, as is seen by letting $\mathscr{A} = KB(\mathfrak{X})$ for an injective compact operator K with dense range.

Some other variants of Lomonosov's Theorem can be found in [3], [8] and [11]. We are grateful to L. Fialkow for providing us with a copy of [1], where it is shown that $AK = \lambda KA$ for K compact and λ a complex number implies A has a hyperinvariant subspace. In the case where $|\lambda| \leqq 1$ this follows from Corollary 1 above; when $|\lambda| > 1$ it follows from the analogous corollary to Theorem 6.

References

[1] Scott Brown, Connections between an operator and a compact operator that yield hyperinvariant subspaces, *preprint*.

[2] I. Colojoară and C. Foiaş, *Theory of Generalized Spectral Operators*, Gordon Breach (New York, 1968).

[3] J. Daughtry, An invariant subspace theorem, *Proc. Amer. Math. Soc.*, **49** (1975), 267—168.

[4] R. G. Douglas, On majoration, factorization and range inclusion of operators on Hilbert space, *Proc. Amer. Math. Soc.*, **17** (1966), 413—415.

[5] M. Embry, Factorization of operators on Banach space, *Proc. Amer. Math. Soc.*, **38** (1973), 587—590.

[6] P. A. Fillmore and J. P. Williams, On operator ranges, *Advances in Math.*, **7** (1971), 254—281.

[7] S. Grabiner, Ranges of products of operators, *Can. J. Math.*, **26** (1974), 1430—1441.

[8] H. W. Kim, C. Pearcy and A. L. Shields, Rank one commutators and hyperinvariant subspaces, *Mich. Math. J.*, **22** (1975), 193—194.

[9] V. Lomonosov, Invariant subspaces for operators commuting with compact operators, *Funct. Anal. Appl.*, **7** (1973), 55—56.

[10] E. A. Nordgren, M. Radjabalipour, H. Radjavi and P. Rosenthal, Algebras intertwining compact operators, *Acta Sci. Math.*, **39** (1977), 115—119.

[11] C. Pearcy and A. L. Shields, A survey of the Lomonosov technique in the theory of invariant subspaces, pp. 220—229 of *Topics in Operator Theory*, Amer. Math. Soc. Surveys No. 13 (Providence, R. I., 1974).

[12] M. Radjabalipour and H. Radjavi, Compact-operator ranges and transitive algebras, *J. Lond. Math. Soc.*, to appear.

[13] H. Radjavi and P. Rosenthal, *Invariant Subspaces*, Springer-Verlag (Berlin—Heidelberg—New York, 1973).

[14] S. Grabiner, Operator ranges and invariant subspaces, *to appear*.

[15] D. W. Hadwin, E. A. Nordgren, Heydar Radjavi, and Peter Rosenthal, An operator not satisfying Lomonosov's hypothesis, *to appear*.

(NORDGREN)
DEPT. MATH. AND COMP. SCI.
KINGSBURY HALL,
UNIVERSITY OF NEW HAMPSHIRE
DURHAM, NH 03824
USA

Acta Sci. Math., 41 (1979), 63—76

Kernel systems of directed graphs

ANDRÁS FRANK

0. In graph theory there is a number of min—max theorems of quite similar type such that one is not a direct consequence of the other. For instance, a theorem of J. Edmonds states that in a directed graph there exist k edge disjoint spanning arborescences rooted at a fixed vertex r (see the exact definitions and formulation below) if and only if the indegree of every subset of vertices, not containing r, is at least k. A version of Menger's theorem resembles Edmonds' one: in a directed graph there exist k edge disjoint paths from r to another fixed vertex s if and only if the indegree of every subset of vertices, containing s but not r, is at least k.

It is a natural question whether there exists a common generalization of these theorems of similar type. The purpose of this paper is to present a tool, by means of which such a unification can be obtained on the one hand, and new min—max theorems can be concluded on the other hand. This tool is the notion of a kernel system, which is, roughly, a family of subsets of vertices of a directed graph which is closed under intersection.

Perhaps the most interesting consequences of min—max theorems concerning kernel systems are the following:

a) A conjecture of J. Edmonds and R. Giles concerning directed cuts is solved for graphs possessing an arborescence.

b) A min—max formula is given for the maximum number of edges which can be covered by K spanning arborescences rooted at a fixed vertex.

Some further corollaries of our results will be published in another paper [7] where, among others, a min—max formula is given for the maximum number of edges of a digraph which can be covered by k branchings.

At this point we refer to a recent, fundamental article of Edmonds and Giles [2] concerning min—max relations for submodular functions.

Some of our notions are similar to those of Edmonds and Giles and in the proof of Theorem 3 a relevant idea of their work will be used. However our results

Received October 18, 1977.

seem to be independent of the main theorem of [2]. The exact relation will be explained in the last section.

Let $G=(V, E)$ be a finite directed graph with vertex set V and edge set E. Multiple edges are allowed, loops are excluded. Let r be a distinguished vertex, called the *root* of G. An arborescence rooted at r (or briefly r-arborescence) is a directed spanning tree such that every vertex can be reached by a directed path from r (see [1]). An *$r-s$-path* is a directed path from r to the vertex s.

We say that a directed edge e *enters* a subset X of vertices if the head of e is in X but the tail is not. We say that a subset E' of edges *enters* a subset X of V if at least one element of E' enters X. The *indegree* $\varrho(X)$ and the *outdegree* $\delta(X)$ of a subset X of V is the number of edges entering X or $V \setminus X$, respectively. It is well known that the function $\varrho(X)$ is submodular, i.e. $\varrho(X)+\varrho(Y) \geqq \varrho(X \cup Y)+$ $+\varrho(X \cap Y)$ for every pair X, Y of subsets of vertices.

For an arbitrary set X, $X' \sqsubset X$ means that X' is a family of not necessarily distinct elements of X. $|X|$ denotes the cardinality of X. We shall use the notation $V \setminus r$ instead of $V \setminus \{r\}$. Two subsets X and Y of $V \setminus r$ are called *crossing* if $X \cap Y \neq \emptyset$, $X \setminus Y \neq \emptyset$, $Y \setminus X \neq \emptyset$. Otherwise X and Y are *non-crossing*. A family of subsets of $V \setminus r$ is called *laminar* if its members are pairwise non-crossing. (These notions occur slightly more generally in previous papers [2, 9].) A *directed cut* of G is a nonempty set of edges entering a vertex set X provided $\delta(V \setminus X)=0$.

1. Definition. A family $\mathcal{M}$ of distinct subsets of vertices of $V \setminus r$ is called a *kernel system* with respect to G if

1) $\varrho(M)>0$ for every $M \in \mathcal{M}$;

2) if M, $N \in \mathcal{M}$ and $M \cap N \neq \emptyset$ then $M \cap N$, $M \cup N \in \mathcal{M}$. The members of $\mathcal{M}$ are called *kernels*.

Examples. 1. $\mathcal{M}_1=\{M: M \subseteqq V \setminus r\}$. The second axiom is trivially satisfied, the first one holds if G has an r-arborescence.

2. Let s be another fixed vertex of G and $\mathcal{M}_2=\{M: M \subseteqq V \setminus r, s \in M\}$. The first axiom holds if there exists an $r-s$-path.

3. $\mathcal{M}_3=\{M: M \subseteqq V \setminus r, \delta(M)=0\}$. If G is connected (in the undirected sense) then the first axiom is fulfilled. The proof of the second one, as an easy exercise, is left to the reader.

4. If $\mathcal{M}$ is an arbitrary kernel system with respect to G then the kernels of minimum indegree form another kernel system

$$\mathcal{M}' = \left\{M: M \in \mathcal{M}, \varrho(M) = \min_{X \in \mathcal{M}} \varrho(X)\right\}.$$

The proof of the second axiom is as follows: Let $k = \min_{X \in \mathcal{M}} \varrho(X)$ and $M, N \in \mathcal{M}'$.

Then

$$k+k = \varrho(M)+\varrho(N) \geqq \varrho(M \cup N)+\varrho(M \cap N) \geqq k+k$$

whence $\varrho(M \cup N)=\varrho(M \cap N)=k$, therefore $M \cup N, M \cap N \in \mathcal{M}'$.

5. Let $\mathcal{M}$ be a kernel system and F be a subset of edges, then

$$\mathcal{M}_F = \{M\colon M \in \mathcal{M},\ F \text{ does not enter } M\}$$

is again a kernel system. The axioms trivially hold.

2. Let k be a positive integer.

Definition. A subset E' of edges is called *k-entering* with respect to the kernel system $\mathcal{M}$, if in the subgraph formed by E', the indegree of every kernel is at least k.

Theorem 1. *A subset E' of edges is k-entering if and only if E' can be partitioned into k 1-entering subsets.*

Proof. The necessity is trivial. For the sufficiency it can be assumed that $E'=E$. We are going to prove that E can be partitioned into a 1-entering subset E_1 and a $(k-1)$-entering subset E_2. This assertion proves our theorem.

The subset E_1 will be constructed sequentially and once an edge has been inserted into E_1 it is never changed. In an intermediate stage of the algorithm a kernel M is called *dangerous* with respect to the current E_1 if

$$\varrho_{G-E_1}(M) = k-1.$$

Starting from the empty set E_1, in every step we consider a maximal kernel M such that E_1 does not enter M. Insert an edge e into E_1 which enters M but does not enter any dangerous kernel, and then we say that e was inserted into E_1 because of M. The process stops when E_1 is 1-entering.

To verify this algorithm we have to justify that the required edge e always exists.

Claim 1. *If $f \in E_1$ then the head of f is not in M.*

Proof. Suppose the contrary then the tail of f is also in M, by the algorithm. Let E_f denote the set of edges which were inserted into E_1 before f, and suppose that f was inserted into E_1 because of M_f. Now $M_f \cap M \neq \emptyset$ therefore $M_f \cap M$ is a kernel. E_f does not enter $M_f \cap M$ and $M_f \cup M \neq M_f$ which contradict the maximality of M_f. □

Claim 2. *If M_D is dangerous with respect to E_1 then $M_D \subsetneqq M$.*

Proof. Since M_D is dangerous, there exists an edge $e_1 \in E_1$ entering M_D. The head of this edge is in M_D but not in M by Claim 1. □

Claim 3. *If M and N are dangerous kernels and $M\cap N$ is nonempty, then $M\cap N$ is dangerous as well.*

Proof. $k-1+k-1=\varrho_{G-E_1}(M)+\varrho_{G-E_1}(N)\geqq\varrho_{G-E_1}(M\cup N)+\varrho_{G-E_1}(M\cap N)\geqq \geqq k-1+k-1$ whence $\varrho_{G-E_1}(M\cap N)=k-1$. □

If every dangerous kernel is disjoint from M then an arbitrary edge entering M can be inserted into E_1 and we are done since the new set $E\setminus E_1$ remains $(k-1)$-entering. Otherwise let M_D be a dangerous kernel such that $M_D\cap M\neq 0$ and $M_D\setminus M$ is as small as possible.

By Claim 2, $M_D\setminus M\neq\emptyset$. There exists an edge e with tail in $M_D\setminus M$ and head in $M_D\cap M$ since otherwise

$$k-1 = \varrho_{G-E_1}(M_D) \geqq \varrho_{G-E_1}(M_D\cap M) \geqq k-1$$

whence $M_D\cap M$ is a dangerous kernel, contradicting Claim 2.

We assert that the edge e enters no dangerous set. If e entered a dangerous set M_e then $M'=M_e\cap M_D$ would also be dangerous by Claim 3. The existence of such an M' is in contradiction with the minimum property of M_D. □

Corollary 1. (J. Edmonds [4]) *A digraph G has k edge-disjoint r-arborescences if and only if the indegree of every subset of $V\setminus r$ is at least k.*

Proof. Apply Theorem 1 to the first example. The corollary follows from the simple fact that a 1-entering edge set surely contains an r-arborescence. □

Corollary 2. (Directed edge version of Menger's theorem [1]) *In a digraph there exist k edge disjoint $r-s$-paths if and only if the indegree of every subset of $V\setminus r$ containing s is at least k.*

Proof. Apply Theorem 1 for the second example. The corollary follows from the simple fact that a 1-entering edge set surely contains an $r-s$-path. □

The next consequence settles in the affirmative a conjecture of J. Edmonds and R. Giles [2] in a special case.

Conjecture. An edge set E' is a k-covering of directed cuts of a directed graph if and only if E' can be partitioned into k 1-coverings of directed cuts. (An edge set E' is called a *k-covering of directed cuts* if every directed cut contains at least k edges of E').

Corollary 3. *The conjecture of Edmonds—Giles is true for graphs possessing an arborescence.*

Proof. Applying Theorem 1 to the third example we obtain that a k-covering (that is a k-entering edge set) of those directed cuts which are directed away from

r can be partitioned into k 1-coverings. However when the graph has an r-arborescence then all of the directed cuts are of this type.

Remark. The proof of Theorem 1 can be considered as a generalization of LovÁsz' proof in [8] of the afore mentioned theorem of Edmonds. It is, in fact, a polynomial bounded algorithm provided that some simple operations can be carried out in polynomial time on the kernels. These operations are as follows:

a) Find a maximal kernel M such that E' does not enter M for an arbitrary edge set E'.

b) Decide whether E'' is k-entering for arbitrary edge set E''.

The above three corollaries are of this type. In Corollary 1 we obtain LovÁsz' algorithm. In Corollary 2 our proof does not mean a new algorithm for Menger's theorem since the only way at hand to check b) is to use the classical augmenting path method.

In Corollary 3 operation a) is simple because the required maximal kernel M consists of those vertices which cannot be reached by a directed path from r in the graph arising from G after contracting the edges of E'. Operation b) can be carried out as follows: Let G^+ denote the graph which arises from G after inserting $k-1$ reversed copies of all the edges of E''. It can easily be checked that E'' is k-entering if and only if there exist k edge disjoint $r-s$-paths in G^+ for every vertex $s \in V \setminus r$. This latter problem is polynomially solvable.

3. Let $\mathbf{c}$ be a nonnegative integer function defined on the edge set E of G. $\mathbf{c}(e)$ is called the *weight* of e.

Definition. A family $\mathcal{M}'$ of not necessarily distinct kernels of $\mathcal{M}$ (i.e. $\mathcal{M}' \sqsubset \mathcal{M}$) is called *$\mathbf{c}$-edge-independent* if each edge e enters at most $\mathbf{c}(e)$ members of $\mathcal{M}'$.

Theorem 2.

$$\text{(1)} \qquad \max |\mathcal{M}'| = \min \sum_{e \in E'} \mathbf{c}(e)$$

where the maximum is taken over all the $\mathbf{c}$-edge-independent subfamilies $\mathcal{M}'$ of $\mathcal{M}$ while the minimum is taken over all the 1-entering edge sets E'.
(2) *The maximum can be realized by a laminar $\mathcal{M}'$ too.*

Proof. max$\leqq$min. A simple enumeration shows that $|\mathcal{M}'| \leqq \sum_{e \in E'} \mathbf{c}(e)$ for any c-edge-independent $\mathcal{M}'$ and for any 1-intering E'.

max$=$min. We are going to construct a c-edge-independent family $\mathcal{M}'$ and a 1-entering edge set E' such that $|\mathcal{M}'| = \sum_{e \in E'} \mathbf{c}(e)$.

The algorithm consists of two parts constructing $\mathcal{M}'$ and E', respectively. It has the interesting feature that both of its parts are of the greedy type, i.e. both

$\mathcal{M}'$ and E' will be produced sequentially and once a kernel or edge has been inserted into $\mathcal{M}'$ or E', respectively, it is never changed.

First part: Construction of $\mathcal{M}'$.

First let $\mathcal{M}'$ be empty. In the general step we decide whether there exists a kernel M which can be inserted into the current $\mathcal{M}'$ without destroying its c-edge-independence. If the answer is "no" then the construction of $\mathcal{M}'$ terminates.

Otherwise let M be a minimal kernel which can be inserted into $\mathcal{M}'$ and let us insert into $\mathcal{M}'$ as many copies of M as possible without destroying the c-edge-independence.

The family $\mathcal{M}'$ produced by the first part is obviously c-edge-independent.

In order to describe the second part we need some notations. Let the different kernels of $\mathcal{M}'$ be $M_1, M_2, \dots, M_k$ (i.e. the first part terminated at the $(k+1)$-th step), and suppose that these kernels have been inserted into $\mathcal{M}'$ in this order. We call an edge e *saturated* with respect to $\mathcal{M}'$ (or briefly saturated) if it enters exactly $\mathbf{c}(e)$ members of $\mathcal{M}'$. Let E_i $(i=1, 2, \dots, k)$ denote the set of those saturated edges which have been saturated in the i^{th} step of the first part. It is easy to see that

(3a) $E_i \neq \emptyset$ for $i=1, 2, \dots, k$;

(3b) $E_i \cap E_j = \emptyset$ for $1 \leqq i < j \leqq k$;

(3c) If $e \in E_i$ then e enters M_i;

(3d) If $e \in E_i$, $i<j$ then e does not enter M_j.

Taking into consideration the construction of $\mathcal{M}'$, the following claim can be checked easily.

Claim 1. *If $M_i \in \mathcal{M}'$, $M \subset M_i$, and $M \in \mathcal{M}$ then there exists a saturated edge e which enters M but not M_i, and then e is in E_h where $h<i$.* □

In order to verify (2) we show that $\mathcal{M}'$ is laminar. For, otherwise, let M_i and M_j be two crossing members of $\mathcal{M}'$ $(i<j)$. Applying Claim 1 with the choice M^i and $M = M_i \cap M_j$ we obtain that there exists an edge e in E_h (for some $h<i$) which enters M but not M_i. Then e enters M_j, a contradiction to (3d).

Second part: Construction of E'.

First let E' be empty. In the general step we decide whether E' is 1-entering. If the answer is "yes" then the second part terminates.

Otherwise, let M be a maximal kernel such that the current E' does not enter M. Let i be the minimum index for which E_i enters M. Let us insert an edge e of E_i which enters M into E'. (We say that e has been inserted because of M.)

The set E' produced by the second part is obviously 1-entering.

To verify (1) and the algorithm we have to show that there exists a unique edge of E' entering M_i for each member M_i of $\mathcal{M}'$. This implies $|\mathcal{M}'| = \sum_{e \in E'} \mathbf{c}(e)$, taking into consideration the fact that the edges of E' are saturated.

Claim 2. *If an edge e has been inserted into E' because of N, and e enters a member M_i of $\mathcal{M}'$, then $N \supseteq M_i$.*

Proof. Since e enters M_i, using (3d) we obtain that e is in E_j for some $j \geqq i$. If $N \not\supseteq M_i$ then with the choice M_i and $M = N \cap M_i$ Claim 1 implies that there exists an edge e' in E_h (for some $h < i$) which enters $M_i \cap N$ but not M_i. Then e' enters N which is in contradiction with the minimality of j, since $h < j$. □

Now suppose, indirectly, that two edges e_1, e_2 of E' enter a kernel M_i of $\mathcal{M}'$. Suppose that e_1 and e_2 have been inserted into E' because of N_1 and N_2, respectively, and e_2 was inserted later than e_1. By Claim 2, $N_1, N_2 \supseteq M_i$ and e_1 does not enter N_2. Hence $N_1 \cup N_2 \neq N_1$ which contradicts the maximality of N_1. □

Remark. The proof of Theorem 2 can be considered as a generalization of that of FULKERSON [5] given for maximum packing of rooted r-cuts. Our algorithm is polynomial bounded provided that the following simple operations can be carried out in polynomial time.

a) Find a minimal kernel M such that E' does not enter M for an arbitrarily given edge set E'.

b) Decide whether E'' is 1-entering for a given edge set E'', and if it does then find a maximal kernel M such that E'' does not enter M.

All the following corollaries and problems are of such type.

Apply Theorem 2 to the first example:

Corollary 4. (EDMONDS [3], FULKERSON [5]) *In an edge-weighted digraph the minimum weight of an r-arborescence is equal to the maximum number of* **c***-edge-independent vertex sets of $V \setminus r$.* □

(A family of **c**-edge-independent vertex sets corresponds to a packing of r-directed cuts in [5]).

Apply Theorem 2 for the second example:

Corollary 5. (FORD—FULKERSON [6]) *In an edge-weighted digraph the minimum weight of an $r-s$-path is equal to the maximum number of* **c***-edge-independent vertex sets containing s but not r.* □

The following corollaries seem to be new.

Problem 1. Suppose that the maximum number of edge disjoint r-arborescences of a (weakly) connected digraph $G=(V, E)$ is k ($k \geqq 0$). We want to increase this maximum by using new edges. Let the set E_1 of possible new edges be such that $G^+ = (V, E \cup E_1)$ has $k+1$ arborescences. Assign to each edge e of E_1 a nonnegative integer weight $\mathbf{c}(e)$. What is the minimum sum of weights of the required new edges?

Solution. Let us define a kernel system $\mathcal{M}$ with respect to $G_1=(V, E_1)$ as follows:

$$\mathcal{M} = \{M: \varrho_G(M) = k,\ M \subseteqq V \setminus r\}.$$

(Observe that the kernel system $\mathcal{M}$ with respect to G_1 is defined by means of G.) Due to the above theorem of Edmonds (Corollary 1) we have to assure that the indegree of all the subsets of $V \setminus r$ is at least $k+1$, that is, we have to find a minimum weight 1-entering subset of kernel system $\mathcal{M}$. Applying Theorem 2 for this $\mathcal{M}$ we get:

Corollary 6. *The minimum value of the weight sum of those edges of E_1 whose insertion into G increases the maximum number of edge disjoint r-arborescences by one, is equal to the maximum number of not necessarily distinct subsets of $V \setminus r$ such that* (i) *the indegree of the set in G is minimum $(=k)$ and* (ii) *an arbitrary edge e of E_1 enters at most $\mathbf{c}(e)$ subsets of them.* □

Remark. A possible generalization arises naturally. Let $G=(V, E)$ be strongly k-edge-connected and E_1 be a set of new edges. Find a minimum subset E_2 of E_1 such that $G^+=(V, E \cup E_2)$ is strongly $(k+1)$-edge-connected. However it is easy to check that the Hamilton circuit problem is contained in this one in the case $k=0$. Therefore this problem is NP-hard and this direction is hopeless.

Now a simple application of Corollary 6 will be presented.

Problem 2. Let us suppose that $G=(V, E)$ has an r-arborescence. Let $F=(E, A)$ be the hypergraph of all r-arborescence of G. Here the vertex set E of F is the edge set of G and the edge set of F is the family of r-arborescences of G. Determine the rank-function $\mathbf{r}$ of F. We recall the definition of the rank-function $\mathbf{r}$ of an arbitrary hypergraph:

$$\mathbf{r}(E') = \max_{a \in A} |a \cap E'| \quad (E' \subseteqq E) \tag{4}$$

(i.e. $\mathbf{r}(E')$ shows at most how many edges of E' can occur in an r-arborescence). Since every arborescence consists of $|V|-1$ edges, our problem is equivalent to the following:

Let us complete E' by a minimum number edges of $E \setminus E'$ so that the completed E' contains an r-arborescence. Applying Corollary 6 for the case when the original graph is $G'=(V, E')$, $E_1=E \setminus E'$, $\mathbf{c}\equiv 1$ and $k=0$, we obtain

Corollary 7. $\mathbf{r}(E') = \min_{V_1, V_2, \ldots, V_t} (|V|-1-t)$ *where the minimum is taken over all those laminar families of subsets $V_1, V_2, \ldots, V_t$ of $V \setminus r$ for which E' does not enter any V_i and an arbitrary edge of $E \setminus E'$ enters at most one V_i.*

Hence one can easily obtain

Corollary 8. *A subset E' of edges of G is a subset of an r-arborescence if and only if $|V|-1 \geqq |E'|+t$ for an arbitrary 1-edge-independent laminar family of subsets $V_1, V_2, \ldots, V_t$ of $V \setminus r$ such that E' enters no V_i.*

Remarks 1. One can immediately prove a slightly sharper version of this corollary when in the necessary and sufficient condition the cardinalities of all but one V_i are one.

2. Some further special cases of the above corollaries are interesting for their own sake. Let us apply Corollary 6 in the case if $k=0$ and E_1 consists of the reversed copies of all edges of E. We obtain a theorem of Lucchesi—Younger type (but not the Lucchesi—Younger theorem itself), which simply follows from the theorem of Edmonds—Giles [2], too (although our proof provides a polynomial algorithm as well). The reader may find it interesting to specialize for the case $k \geqq 1$, $E_1 = E$ and $\mathbf{c} \equiv 1$. In this way a min—max theorem can be obtained for the minimum number of edges of G whose duplication increases the maximum number of edge disjoint r-arborescences.

4. In this section a generalization of Theorem 2 will be given. Unlike the proof of Theorem 2, this does not provide a polynomial algorithm. This is the reason why Theorem 2 was discussed in the previous paragraph.

Let $\mathcal{M}$ be a kernel system with respect to $G=(V, E)$ and let $\mathbf{f}$ be a nonnegative integer function defined on the kernels.

Definition. The function $\mathbf{f}$ is called *weakly supermodular* on $\mathcal{M}$ if $M, N \in \mathcal{M}$, $\mathbf{f}(M)>0$, $\mathbf{f}(N)>0$, $M \cap N \neq \emptyset$ imply that

$$\mathbf{f}(M)+\mathbf{f}(N) \leqq \mathbf{f}(M \cup N)+\mathbf{f}(M \cap N). \tag{5}$$

If already $M, N \in \mathcal{M}$ and $M \cap N \neq \emptyset$ imply this inequality then $\mathbf{f}$ is called *supermodular*.

Definition. A family E' of not-necessarily distinct edges of E (i.e. $E' \sqsubset E$) is called $\mathbf{f}$-*entering*, if in the subgraph $G'=(V, E')$ the indegree of every kernel M is at least $\mathbf{f}(M)$.

Let $\mathbf{c}$ be a nonnegative integer function defined on the edges of G.

Theorem 3. *Let $\mathbf{f}$ be a weakly supermodular function on $\mathcal{M}$. Then*

$$\max_{\mathcal{M}' \subset \mathcal{M}} \sum_{M \in \mathcal{M}'} \mathbf{f}(M) = \min_{E' \sqsubset E} \sum_{e \in E'} \mathbf{c}(e) \tag{6}$$

where $\mathcal{M}'$ is $\mathbf{c}$-edge-independent, E' is $\mathbf{f}$-entering.

(7) *The maximum can be realized by a laminar $\mathcal{M}'$.*

Proof. First we will prove (7) which will be used in the proof of (6), too. We note that this technique is due to N. Robertson for $\mathbf{f}\equiv 1$ and to Edmonds and Giles for an arbitrary supermodular function $\mathbf{f}$. It can be assumed that the optimum $\mathcal{M}'$ consists of kernels with positive weights only. If M, N are crossing members of $\mathcal{M}'$ then replace them by $M\cup N$ and $M\cap N$ i.e. $\mathcal{M}''=\mathcal{M}'\setminus\{M, N\}\cup \cup\{M\cup N, M\cap N\}$. It is easy to check that $\mathcal{M}''$ is c-edge-independent again and, since $\mathbf{f}$ is weakly supermodular,

$$\sum_{M\in\mathcal{M}''}\mathbf{f}(M)\geqq\sum_{M\in\mathcal{M}'}\mathbf{f}(M).$$

Hence $\mathcal{M}''$ is another optimum c-edge-independent family. Apply this method as long as there exist crossing members in the optimum family. The process terminates since $\sum_{M\in\mathcal{M}'}|M|^2$ increases at each step.

We need two simple claims.

Claim 1. *Let e be an edge of G and let $\mathbf{f}$ be a weakly supermodular function on $\mathcal{M}$. Let*

$$\mathbf{f}_e(M)=\begin{cases}\mathbf{f}(M), & \textit{if } e \textit{ does not enter } M\\ 0, & \textit{if } \mathbf{f}(M)=0\\ \mathbf{f}(M)-1, & \textit{otherwise,}\end{cases}$$

then $\mathbf{f}_e$ is weakly supermodular.

The proof of the claim is trivial.

We note that the analogous property for supermodular functions is not necessarily true.

Claim 2. *Let $\mathbf{c}_1(e)=k\cdot\mathbf{c}(e)$ for a natural number k. If $\mathcal{M}''\subset\mathcal{M}$ is a laminar $\mathbf{c}_1$-edge-independent family, then it can be partitioned into k $\mathbf{c}$-edge-independent families.*

Proof. The members of $\mathcal{M}''$ will be colored one by one with colors $0, 1, \dots k-1$. In the general step let M be a maximal non-colored member of $\mathcal{M}''$. If there exist no previously colored member M' of $\mathcal{M}''$ containing M then let M be colored by 0. Otherwise let M' be a previously colored kernel with $M'\supseteqq M$, which received its color last. If the color of M' is i then we color M by $i+1 \bmod k$.

It is an easy exercise to verify that each subfamily of kernels with the same color is c-edge-independent. □

For the proof of (6) a simple enumeration shows that max≦min. Let $v_\mathbf{f}$ denote the left-hand side in (6). We use induction on $v_\mathbf{f}$. If $v_\mathbf{f}=0$ then the statement is trivial.

Let M be an arbitrary kernel such that $\mathbf{f}(M)>0$ and not all the edges entering M are of zero weight. There are two cases.

(a) There is an edge e with positive weight, entering M such that all the optimum (of weight $v_{\mathbf{f}}$) $\mathbf{c}$-edge-independent families saturate e (i.e. e enters just $\mathbf{c}(e)$ kernels of the family with positive weight).
In this case $v_{\mathbf{f}_e}=v_{\mathbf{f}}-\mathbf{c}(e)$. By the induction hypothesis there exists an $E'_e \subset E$ for which $v_{\mathbf{f}_e}=\sum_{e'\in E'_e}\mathbf{c}(e')$ and E'_e is $\mathbf{f}_e$-entering. Let $E'=E'_e\cup\{e\}$. Since $v_{\mathbf{f}}=\sum_{e'\in E'}\mathbf{c}(e')$ and E' is $\mathbf{f}$-entering we are finished with the proof.

(b) For each edge e_i with positive weight and entering M there exists an optimum $\mathbf{c}$-edge-independent family $\mathcal{M}_i$ which does not saturate e_i. Let $\mathcal{M}''=$ $=\mathcal{M}_1\cup\mathcal{M}_2\cup\ldots\cup\mathcal{M}_k\cup\{M\}$. Then $\mathcal{M}''$ is $\mathbf{c}_1$-edge-independent where $\mathbf{c}_1=k\cdot\mathbf{c}$ and

$$\sum_{N\in\mathcal{M}''}\mathbf{f}(N)=k\cdot v_{\mathbf{f}}+\mathbf{f}(M).$$

By the proof of (7) there exists a laminar family $\mathcal{M}'''$ such that

$$\sum_{N\in\mathcal{M}'''}\mathbf{f}(N)\geqq\sum_{N\in\mathcal{M}''}\mathbf{f}(N).$$

Now by Claim 2, $\mathcal{M}'''$ can be partitioned into k $\mathbf{c}$-edge-independent subfamilies. However, the weight of one of these subfamilies is greater than $v_{\mathbf{f}}$ which is impossible. Hence case (b) cannot occur. □

Theorem 3 reduces to Theorem 2 in the case $\mathbf{f}\equiv 1$, therefore the corollaries of Theorem 2 can be generalized. However, we emphasize only one consequence of Theorem 3.

Problem 3. Let $G=(V,E)$ be a digraph in which the maximum number of edge-disjoint r-arborescences is k $(k>0)$. We want to increase this maximum to K $(K>k)$ by multiplying edges. What is the minimum number of the required new edges?

Solution. Due to the theorem of Edmonds (Corollary 1) we have to assure just that in the extended graph the indegree of every subset of $V\setminus r$ is at least K.

Let $\mathcal{M}$ be the kernel system defined in the first example. Let the function $\mathbf{f}$ be defined as follows:

$$\mathbf{f}(M)=\max\{K-\varrho(M),0\} \tag{8}$$

that shows the number of edges still required to reach K as the indegree of M. In this way our question is translated into the problem of a minimum $\mathbf{f}$-entering edge set.

Claim. *The above defined* $\mathbf{f}$ *is weakly supermodular.*

Proof. Trivial. □

We note that $\mathbf{f}$ is not supermodular in general.

Applying Theorem 3 for this $\mathbf{f}$ in the case $\mathbf{c}\equiv 1$ we obtain a min—max formula for the minimum number of new edges. Instead of the exact formulation of this theorem we mention another problem which is equivalent to this one but is more illustrative.

Problem 4. What is the maximum number of edges which can be covered by K r-arborescences?

Solution. If there exist K edge disjoint r-arborescences then this number is obviously $K\cdot(|V|-1)$. Otherwise let $a_1, a_2, \dots, a_K$ be r-arborescences whose union is as large as possible. Suppose that this union consists of m edges. Let us multiply every edge of G by the number of r-arborescences from $a_1, a_2, \dots, a_K$ containing it. Of course this graph has already K r-arborescences. This means that $s=K\cdot(|V|-1)-m$ new copies of original edges assure the existence of K edge disjoint r-arborescences. Conversely, if the insertion of s new copies of edges yields the existence of K edge disjoint r-arborescences, then $m=K\cdot(|V|-1)-s$ edges can be covered by K r-arborescences in G. In this way Problem 4 is equivalent to Problem 3. Hence, as a consequence of Theorem 3, we obtain

Corollary 9. *The maximum number of edges which can be covered by K r-arborescences is equal to the minimum value of*

$$K(|V|-1)-\sum_{i=1}^{t}\mathbf{f}(V_i)$$

where the minimum is taken over all the 1*-edge-independent laminar families of subsets $V_1, V_2, \dots, V_t$ of $V\setminus r$ where t is arbitrary and function $\mathbf{f}$ is defined in* (8).

There is an interesting special case of this corollary.

Corollary 10. *The edges of G can be covered by K r-arborescences if and only if for an arbitrary laminar* 1*-edge-independent family of subsets $V_1, V_2, \dots V_t$ of $V\setminus r$, the number e_t of edges entering no V_i satisfies*

$$e_t \leq K(|V|-1-t). \tag{9}$$

Remark. K. Vidyasankar [11] has proved a similar but simpler necessary and sufficient condition for the problem in Corollary 10. He requires (9) only in the case if the cardinality of all but one of the V_i's is one, with the two side-conditions that the indegree of each vertex is at most K and every edge is in an r-arborescence. The necessity of these two latter conditions is trivial (and obviously our conditions imply them).

Now we formulate Corollary 9 in another way. Suppose again that G has an r-arborescence. Let E' be a subset of edges of G and let $\mathbf{r}(E')$ denote the maximum number of edges E' can have in common with an r-arborescence, i.e. $\mathbf{r}$ is the rank-function of the hypergraph of r-arborescences. We recall that function $\mathbf{r}$ was determined by a min—max formula in Corollary 7.

Corollary 9a. *The maximum number of edges which can be covered by K r-arborescences is equal to the*

$$\min_{E''\subseteq E}(K\cdot\mathbf{r}(E'')+|E\setminus E''|).$$

Proof. max$\leqq$min is true for any hypergraph. For the equality we show that

$$\min_{E''\subseteq E'}(K\cdot\mathbf{r}(E'')+|E\setminus E''|)\leqq K(|V|-1)-\sum_{i=1}^{t}\mathbf{f}(V_i) \tag{10}$$

where $V_1, V_2, \ldots, V_t$ form a 1-edge-independent family. It can be assumed that $\mathbf{f}(V_i)>0$ whence $\mathbf{f}(V_i)=K-\varrho(V_i)$. Let E'' be the set of edges which do not enter any V_i. We have $\sum_{i=1}^{t}\varrho(V_i)=|E\setminus E''|$. Obviously, an arbitrary r-arborescence contains at least t edges entering one of the V_i's. Thus $\mathbf{r}(E'')\leqq|V|-1-t$. Hence (10) follows, as required. □

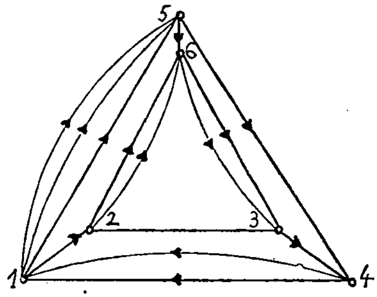
A similar version of Corollary 10 easily follows.

Corollary 10a. *The edges of G can be covered by K r-arborescences if and only if $K\cdot\mathbf{r}(E')\geqq|E'|$ for every $E'\subseteq E$.*

The reader can easily observe the similarity between Corollary 10a and a Theorem of C. St. J. A. Nash-Williams [10] on the covering of a matroid by K bases.

5. In this last section we discuss the relationship between our results and those of J. Edmonds and R. Giles. Roughly speaking the main difference is that we consider entering edges only while they deal with entering and outcoming edges together.

Edmonds and Giles have defined the notion of crossing family. Our theorems concern a special type of crossing family (when the members of the family do not contain a fixed vertex), but they cannot, however, be generalized for arbitrary crossing family. The remark after Corollary 6 justifies this statement for Theorem 2. The example in the Figure shows that Theorem 1 also fails for general crossing families.



Let $\mathcal{M}=\{M\colon \varrho(M)=2\}=\{(1, 2, 3, 4, 6), (2, 3, 6), (2), (1, 2, 4, 5, 6), (4)\}$. The edges cannot be colored with two colors so that both of the color classes enter every kernel.

References

[1] C. Berge, *Graphs and Hypergraphs*, North-Holland (Amsterdam, 1973).

[2] J. Edmonds—R. Giles, A min-max relation for submodular functions on graphs, in *Annals of Discrete Math.*, **1** (1977), 185—204.

[3] J. Edmonds, Optimum branchings, *J. Res. Nat. Bur. Standards Sect. B.*, **71** (1967), 233—240.

[4] J. Edmonds, Edge disjoint branchings, in *Combinatorial Algorithms*, Academic Press (New York, 1973), 91—96.

[5] D. R. Fulkerson, Packing rooted directed cuts in a weighted directed graph, *Math. Programming*, **6** (1974), 168—194.

[6] L. R. Ford—D. R. Fulkerson, *Flows in Networks*, University Press (Princeton, 1962).

[7] A. Frank, Covering branchings, *Acta Sci. Math.*, **41** (1979), 77—81.

[8] L. Lovász, On two minimax theorems in graph, *J. Combinatorial Theory Ser. B.*, **2** (1976), 96—103.

[9] C. Lucchesi—D. H. Younger, A minimax theorem for directed graphs, *to appear*.

[10] C. St. J. A. Nash-Williams, An application of matroids to graph theory, in *Theory of Graphs*, Journées Internationales d'Études, Rome (Dunod, Paris, 1966), 263—265.

[11] K. Vidyasankar, Covering the edge set of a directed graph with trees, *Discrete Math.*, to appear.

BOLYAI INSTITUTE
ARADI VÉRTANUK TERE 1.
6720 SZEGED, HUNGARY

and

RESEARCH INSTITUTE FOR TELECOMMUNICATION
GÁBOR Á. U. 65.
1026 BUDAPEST, HUNGARY

Acta Sci. Math., 41 (1979), 77—81

Covering branchings

ANDRÁS FRANK

In a previous paper [4] we proved, among others, a min-max theorem concerning cuts of a directed graph. Now this theorem will be applied in order to get some new min-max theorems about branchings and arborescences. For example, a good characterization is given for the problem of the existence of k branchings covering all of the edges of a directed graph. This theorem can be considered as a directed counterpart of a theorem of Nash-Williams about covering forests.

Another corollary is a directed analogue of Tutte's theorem about edge disjoint spanning trees. A directed graph has k edge disjoint spanning arborescences (possibly rooted at different vertices) if and only if, for every family of t disjoint subsets of vertices, the sum of their indegrees is at least $k(t-1)$. This theorem differs from Edmonds' one concerning the existence of k edge disjoint spanning arborescences rooted at a fixed vertex. However we shall use Edmonds' result in the proof.

Let $G=(V,E)$ be a finite directed graph with vertex set V and edge set E. Multiple edges are allowed, loops are excluded. Let r be a distinguished vertex of G. We use the notation $U=V\setminus\{r\}$.

An *arborescence* a is a directed tree such that every edge is directed toward a different vertex. It is well known that an arborescence has a unique vertex (of indegree 0) from which every other vertex can be reached by a directed path. This vertex is called the *root* of a. A spanning arborescence of G rooted at r is called an *r-arborescence*.

A *branching* b is a directed forest, the components of which are arborescences.

We say that a directed edge e *enters* a set X of vertices if the head of e is in X but its tail is not. We say that a subset E' of edges *enters* X if at least one element of E' enters X.

The *indegree* $\varrho_G(X)$ of a subset X of V is the number of edges entering X. The following inequality is straightforward: $\varrho_G(X)+\varrho_G(Y)\geqq\varrho_G(X\cup Y)+\varrho_G(X\cap Y)$.

Received December 30, 1977.

For an arbitrary set X, $X' \sqsubset X$ means that X' is a family of not necessarily distinct elements of X.

A family $\mathcal{F}$ of subsets of U is called *laminar* if at least one of $X \setminus Y$, $Y \setminus X$, $X \cap Y$ is empty for any two members of $\mathcal{F}$.

Let $\mathbf{f}$ be a non-negative integer valued function defined on the subsets of U. $\mathbf{f}$ is called *weakly supermodular* if $X, Y \subseteq U$, $\mathbf{f}(X)$, $\mathbf{f}(Y) > 0$ and $X \cap Y \neq \emptyset$ imply $\mathbf{f}(X) + \mathbf{f}(Y) \leqq \mathbf{f}(X \cup Y) + \mathbf{f}(X \cap Y)$. If $X, Y \subseteq U$ and $X \cap Y \neq \emptyset$ already imply it then $\mathbf{f}$ is called *supermodular*.

A family E' of not necessarily distinct edges of G (i.e. $E' \sqsubset E$) is called $\mathbf{f}$-*entering* if in the graph $G' = (V, E')$ the indegree of every subset X is at least $\mathbf{f}(X)$.

Let $\mathbf{c}$ be a non-negative integer valued function on E. A family $\mathcal{F}$ of not necessarily distinct subsets of U is called $\mathbf{c}$-*edge-independent* if each edge e of G enters at most $\mathbf{c}(e)$ members of $\mathcal{F}$.

The following theorem was proved in a slightly other form in [4].

Theorem 1. *If* $\mathbf{f}$ *is weakly supermodular and* $\varrho(Y) = 0$ *implies* $\mathbf{f}(Y) = 0$ *then*

$$\max_{\mathcal{F}} \sum_{X \in \mathcal{F}} \mathbf{f}(X) = \min_{E' \sqsubset E} \sum_{e \in E'} \mathbf{c}(e)$$

where $\mathcal{F}$ *is* $\mathbf{c}$-*edge-independent* ($\mathcal{F} \sqsubset 2^U$) *and* $E' \sqsubset E$ *is* $\mathbf{f}$-*entering. The maximum can be realized by a laminar* $\mathcal{F}$.

Let k be a natural number and $F \subseteq E$.

Problem 1. What is the maximum number M of edges of F which can be covered by k r-arborescences of G?

The case $F = E$ was discussed in [4]. We formulate this problem in another form.

Problem 1a. What is the minimum number m of not necessarily distinct edges of G which, together with F, contain k edge disjoint r-arborescences?

The two problems are equivalent because $M \geqq k(|V|-1) - m$ and $m \leqq k(|V|-1) - M$, hence

(1) $$m + M = k(|V|-1).$$

By a theorem of J. Edmonds [3, 5] a digraph has k edge disjoint r-arborescences if and only if the indegree of every subset of $V \setminus \{r\}$ is at least k. Therefore $m = \min_{E' \sqsubset E} |E'|$ where E' is $\mathbf{f}$-entering and the function $\mathbf{f}$ is defined as follows:

$$\mathbf{f}(X) = \max\left(0, k - \varrho_H(X)\right) \quad \text{for} \quad X \subseteq U$$

where $\varrho_H(X)$ is the indegree of X in the subgraph $H = (V, F)$. Obviously $\mathbf{f}$ is weakly supermodular. (Observe that F is used only to define $\mathbf{f}$). Applying Theorem 1 to G and to this function $\mathbf{f}$, with the choice $\mathbf{c}(e) = 1$ $(e \in E)$, we get $m = \max_{\mathcal{F}} \sum_{X \in \mathcal{F}} \mathbf{f}(X)$ where $\mathcal{F}$ is 1-edge-independent. This, together with (1), proves

Theorem 2. *If* $H=(V, F)$ *is a subgraph of* $G=(V, E)$ *then the maximum number of edges of* H *which can be covered by* k *r-arborescences of* G *is equal to*

$$\min\left[k(|V|-1-t)+\sum_{i=1}^{t}\varrho_H(V_i)\right]$$

where the minimum is taken over all 1-edge-independent laminar families $\mathcal{F}=\{V_1, V_2, \ldots, V_t\}$ $(V_i\subseteqq U)$.

Problem 2. Let $H=(U, F)$ be a directed graph (there is no distinguished vertex). What is the maximum number M of edges which can be covered by k branchings?

Complete H by a new vertex r and by $|U|$ new edges which are joined from r to all other vertices of U, i.e. $V=U\cup\{r\}$ and $E=F\cup\{(\overrightarrow{r, x}): x\in U\}$. It is easy to check that the maximum number of edges of H which can be covered by k r-arborescences of $G=(V, E)$ is M. Apply Theorem 2 and observe that in this case a laminar family of subsets of U consists of pairwise disjoint subsets. Thus we have

Theorem 3. *The maximum number of edges of* $H=(U, F)$ *which can be covered by* k *branchings is equal to*

$$\min\left[k(|U|-t)+\sum_{i=1}^{t}\varrho_H(V_i)\right]$$

where the minimum is taken over all families of disjoint subsets V_i $(i=1, 2, \ldots, t)$ *of* U.

A simple application of this theorem provides an analogue of Tutte's disjoint spanning trees theorem [8].

Theorem 4. $H=(U, F)$ *has* k *edge disjoint spanning arborescences (possibly rooted at different vertices) if and only if*

$$\sum_{i=1}^{t}\varrho_H(V_i)\geqq k(t-1) \tag{2}$$

for every family of disjoint subsets V_i $(i=1, 2, \ldots, t)$ *of* U.

Proof. H has k edge disjoint spanning arborescences if and only if at least $k(|U|-1)$ edges of H can be covered by k branchings, i.e., by Theorem 3, $k(|U|-t)+\sum_{i=1}^{t}\varrho_H(V_i)\geqq k(|U|-1)$, which is equivalent to (2). $\square$

Another consequence of Theorem 3 is

Theorem 5. *The edges of H can be covered by k branchings if and only if*

(3) $$k(|U|-t) \geqq e_t$$

for every family of disjoint subsets $V_1, V_2, \ldots, V_t$ of U, where e_t denotes the number of edges not entering any V_i.

Proof. By Theorem 3 we have to assure that $k(|U|-t)+\sum_{i=1}^{t} \varrho_H(V_i) \geqq |F|$. But this is equivalent to (3), because $e_t+\sum_{i=1}^{t} \varrho_H(V_i)=|F|$. □

Theorem 5a. *The edges of H can be covered by k branchings if and only if*
(4a) *the indegree of every vertex is at most k, and*
(4b) *the edges of H (in the undirected sense) can be covered by k forests.*

Proof. The necessity of the conditions is obvious. For the sufficiency we verify that (4a) and (4b) imply (3). Let $V_1, V_2, \ldots, V_t$ be disjoint subsets of U. Let $V_0=U \setminus \bigcup_{i=1}^{t} V_i$ (V_0 may be empty) and let $e(X)$ denote the number of edges with tails and heads both in X. Then

$$e_t = \sum_{x \in V_0} \varrho_H(x)+\sum_{i=1}^{t} e(V_i) \leqq k|V_0|+\sum_{i=1}^{t} k(|V_i|-1) = k(|U|-t). \quad \square$$

Remark. The last theorem can be considered as a new "linking" theorem. Let $\mathcal{M}_1$ denote the circuit matroid (on F) of H considering H as an undirected graph. Let $\mathcal{M}_2$ denote the matroid on F in which a subset is defined to be independent if it contains no two edges directed toward the same vertex. Now Theorem 5a states that if F can be covered by k independent sets of $\mathcal{M}_1$ and can be covered by k independent sets of $\mathcal{M}_2$ then F can be covered by k sets which are independent in both $\mathcal{M}_1$ and $\mathcal{M}_2$.

Another special case of this statement, when $\mathcal{M}_1$ and $\mathcal{M}_2$ are transversal matroids, was proved by Brualdi [2]. However, this statement is not true in general: Let $\mathcal{M}_1$ be the circuit matroid of K_4 (the complete graph on 4 vertices) and $\mathcal{M}_2$ be defined such that a subset in independent if it contains no disjoint edges of K_4.

Now we prove a Vizing type theorem which is due to Mosesyan [6] for $\gamma=1$.

Theorem 6. *If in $H=(U, F)$ the indegree of every vertex is at most K and H does not contain $\gamma+1$ edges with the same heads and tails then F can be covered by $k=K+\gamma$ branchings.*

Proof. (4a) holds obviously. To prove (4b) we have to verify that $e(X) \leqq \leqq k(|X|-1)$ for $X \subseteqq U$. This condition is equivalent to (4b) by a well-known theorem of Nash-Williams [7]. If $|X|\gamma \leqq k$ then $e(X) \leqq |X|(|X|-1)\gamma \leqq \leqq k(|X|-1)$. If in turn $|X|\gamma \geqq k$ then $e(X) \leqq |X| \cdot K = |X|(k-\gamma) \leqq k(|X|-1)$. □

Finally, a theorem is stated which is also a consequence of Theorem 1. The proof is left to the reader.

Theorem 7. *The edges of $H=(U, F)$ can be covered by k spanning arborescences if and only if $k(|U|-1-t+d) \geqq e_t$ for every 1-edge-independent laminar family $\mathcal{F}=\{V_1, \ldots, V_t\}$, where e_t is the number of edges not entering any V_i and d denotes the maximum number of V_i's containing any vertex.*

References

[1] C. Berge, *Graphs and Hypergraphs*, North-Holland, (Amsterdam, 1973).

[2] R. A. Brualdi, Common Partial Transversals and Integral Matrices, *Amer. Math. Soc.*, **155** (1971), 475—492.

[3] J. Edmonds, Edge disjoint branchings, in *Combinatorial Algorithms*, Academic Press (New York, 1973), 91—96.

[4] A. Frank, Kernel systems of directed graphs, *Acta. Sci Math.*, **41** (1979), 63—76.

[5] L. Lovász, On two minimax theorems in graph theory, *J. Combinatorial Theory*, Ser. B., **2** (1976), 96—103.

[6] K. M. Mosesyan, oral communication.

[7] C. St. J. A. Nash-Williams, Decompositions of finite graphs into forests, *J. London Math. Soc.*, **39** (1964), 12.

[8] W. T. Tutte, On the problem of decomposing a graph into n connected factors, *J. London Math. Soc.*, **36** (1961), 221—230.

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1.
6720 SZEGED, HUNGARY

and

RESEARCH INSTITUTE FOR TELECOMMUNICATION
GÁBOR ÁRON U. 65.
1026 BUDAPEST, HUNGARY

Acta Sci. Math., 41 (1979), 83—100

The dual discriminator function in universal algebra

E. FRIED and A. F. PIXLEY

1. Introduction and summary of results

For any set S the *(ternary) discriminator* t of S is the function from S^3 to S defined by $t(x, y, z)=x$ if $x \neq y$ and $=z$ if $x=y$.

The discriminator function has proved useful in the study of varieties generated by quasi-primal algebras — which includes the variety of Boolean algebras — and related areas of universal algebra. (See [14] and, for example, [15], [18], [19], [21].) Indeed, in the two element Boolean algebra $(\{0, 1\}, \vee, \wedge, ')$,

$$(x \wedge y') \vee (y' \wedge z) \vee (x \wedge z) \tag{1.1}$$

is a polynomial representing the discriminator of the set $\{0, 1\}$.

In the present paper we introduce the study of a closely related function, the *dual discriminator*, the function d from S^3 to S defined by $d(x, y, z)=x$ if $x=y$, and $=z$ if $x \neq y$.

We may think of the dual discriminator as playing a role which generalizes the "median" polynomial on the two element lattice in the same way that t generalizes (1.1); indeed, for the lattice $(\{0, 1\}, \vee, \wedge)$, the median,

$$(x \wedge y) \vee (y \wedge z) \vee (x \wedge z), \tag{1.2}$$

is a polynomial representing the dual discriminator of $\{0, 1\}$. More generally, the algebras in which the dual discriminator is a polynomial stand, roughly speaking, in the same relation to the two element lattice as quasi-primal algebras stand to the two element Boolean algebra; the purpose of the present paper is to give some grounds for this analogy. The two element lattice is, however, the only lattice in which the dual discriminator function is a polynomial. On the other hand, weakly associative lattices with the unique bound property ([5]) provide important examples of this extension of the theory of distributive lattices. Within this extended theory the special results of [7] are also of particular interest.

Received September 16, 1977.

As is suggested by the examples of Boolean algebras and distributive lattices, the discriminator is strictly "stronger" than the dual discriminator. In fact, from the definition we obviously have

$$d(x, y, z) = t(x, t(x, y, z), z), \tag{1.3}$$

but there is no way of expressing t in terms of d (as one sees by considering the two element lattice). There are, however, interesting relations connecting the two. Among these are the following dual functional equations:

$$t(x, y, d(x, y, z)) = x, \quad d(x, y, t(x, y, z)) = x. \tag{1.4}$$

Also, if f is the *4-ary discriminator* defined by

$$f(x, y, u, v) = u \quad \text{if} \quad x = y, \quad \text{and} \quad = v \quad \text{if} \quad x \neq y, \tag{1.5}$$

then

$$t(x, y, z) = f(x, y, z, x), \quad \text{and} \quad d(x, y, z) = f(x, y, x, z), \tag{1.6}$$

dually. (As is well known, f is equivalent to t since $f(x, y, u, v) = t(t(x, y, u), t(x, y, v), v)$.)

For terminology in the paper we shall generally follow GRÄTZER [9]. In particular, for a given type of algebras a *polynomial symbol* $p(x, y, \ldots)$ is simply a term in the first order theory of that type. If $\mathbf{A} = (A, F)$ is an algebra of this type, the *polynomial* $p^{\mathbf{A}}(x, y, \ldots)$ of $\mathbf{A}$ is the mapping induced on A by $p(x, y, \ldots)$. An *algebraic function* is a mapping of A obtained by inserting fixed elements of A in some of the argument places of a polynomial.

Summary of results. In Section 2 we shall discuss some simple relations between "discriminator" and "dual discriminator" varieties. We also show (Theorem 2.3) that a finite algebra $\mathbf{A}$ of more than two elements is functionally complete if and only if the dual discriminator is an algebraic function of $\mathbf{A}$. Finite algebras in which the dual discriminator is a polynomial are characterized (Theorem 2.4) in a way which generalizes a characterization of quasi-primal algebras. In Sections 3 and 4 we obtain an equational characterization (Theorem 3.2) of dual discriminator varieties. This result parallels an earlier result of MCKENZIE [11] for discriminator varieties. We also show (Theorem 3.11) that dual discriminator varieties have equationally definable principal congruences in the sense of [8], and examine the duality between "principal" and "co-principal" congruences in discriminator varieties. It is further shown (Theorem 4.2) that the discriminator behaves, in certain respects, like a generalized complementation operation. In Section 5 we examine weakly associative lattices and, in particular, obtain an explicit finite equational base for the variety generated by all weakly associative lattices having the unique bound property (Theorem 5.8).

2. Discriminator and dual discriminator varieties; functional completeness

A *discriminator variety* is a variety V having a ternary polynomial symbol $p(x, y, z)$ such that for each subdirectly irreducible (SI) algebra $\mathbf{A} \in V$, $p^{\mathbf{A}}(x, y, z)$ is the discriminator of A. Finite SI members of a discriminator variety are usually called *quasi-primal* algebras ([14]). Dually, let us say that a variety V is a *dual discriminator variety* if V has a ternary polynomial $q(x, y, z)$ such that $q^{\mathbf{A}}(x, y, z)$ is the dual discriminator of $\mathbf{A}$ for each SI algebra $\mathbf{A} \in V$.

Note that if V is a discriminator (respectively, dual discriminator) variety in which the polynomial symbols p and p' (respectively q and q') each induce the discriminator (respectively, dual discriminator) on each SI member of V, then $p=p'$ (respectively $q=q'$) is an equation of V. Briefly, the discriminator and dual discriminator are unique. Also note that by (1.3) each discriminator variety is a dual discriminator variety.

In addition to the variety of Boolean algebras, discriminator varieties include, as a few examples, all varieties of arithmetical rings (i.e.: varieties generated by finite sets of finite fields), varieties generated by simple relation algebras, and simple cylindric algebras. (See [21] for other examples.) Beyond the variety of distributive lattices the simplest dual discriminator variety is the variety W_3 generated by the "triangle" algebra $\mathbf{W}_3=(\{0, 1, a_1\}, \vee, \wedge)$ where $\vee, \wedge$ are the l.u.b. and g.l.b. respectively for the following reflexive and antisymmetric relation: $0 \leqq 1 \leqq a_1 \leqq 0$. (See [6].) In this case the polynomial

$$[(z \wedge (x \wedge y)) \vee (x \vee y)] \wedge [z \vee (x \wedge y)]$$

will be shown, in Section 5, to be the dual discriminator of the set $\{0, 1, a_1\}$. More generally, varieties generated by weakly associative lattices having the unique bound property ([5]) will also be shown to be dual discriminator varieties. Interesting special cases include the varieties $W_n, 2 \leqq n < \omega$, generated by the algebras $\mathbf{W}_n=(\{0, 1, a_1, \ldots, a_{n-2}\}, \vee, \wedge)$ where $\vee, \wedge$ are l.u.b. and g.l.b. for the reflexive and antisymmetric relation $0 \leqq 1 \leqq a_i \leqq 0$, $i=1, \ldots, n-2$. These were introduced in [6].

The following are some simple comparative properties of the discriminator and dual discriminator.

2.1 Lemma. *If V is a discriminator variety or a dual discriminator variety, each nontrivial SI member of V is simple and has only simple nontrivial subalgebras.*

Proof. For the dual discriminator, if $\theta > \omega$ is a congruence of any subalgebra of $\mathbf{A} \in V$, let $(x, y) \in \theta$, $x \neq y$. Then for any z in the subalgebra,

$$x = q^{\mathbf{A}}(x, x, z) \; \theta \; q^{\mathbf{A}}(x, y, z) = z$$

For the discriminator,

$$x = p^{\mathbf{A}}(x, y, z)\ \theta\ p^{\mathbf{A}}(x, x, z) = z.$$

2.2 Lemma. i) *A discriminator variety V is arithmetical (i.e.: congruence permutable and congruence distributive); equivalently,*

$$p(x, x, z) = z, \quad p(x, y, x) = x, \quad p(x, z, z) = x$$

are equations of V.

ii) *A dual discriminator variety V is congruence distributive, in fact 2-distributive; equivalently,*

$$q(x, x, z) = x, \quad q(x, y, x) = x, \quad q(x, z, z) = z$$

are equations of V.

iii) *A dual discriminator variety is a discriminator variety if and only if it is congruence permutable.*

Proof. i) and ii) are well known; see, e.g., [14] and [10]. For iii), if V is congruence permutable and $m(x, y, z)$ is any Mal'cev polynomial symbol for V, then for any SI member $\mathbf{A} \in V$, $m^{\mathbf{A}}(x, q^{\mathbf{A}}(x, y, z), z)$ is clearly the discriminator of A. If V is a discriminator variety, then $p(x, y, z)$ is a Mal'cev polynomial symbol for V, by i). (Note that for *any* Mal'cev polynomial symbol $m(x, y, z)$ for V, and in particular for $m=p$, $m^{\mathbf{A}}(x, p^{\mathbf{A}}(x, y, z), z)$ is the dual discriminator of any SI $\mathbf{A} \in V$. In general the discriminator and its dual are interdefinable through any Mal'cev function.)

Recall that a finite algebra $\mathbf{A}$ is *functionally complete* ([13]) if each function $f\colon A^n \to A$, $0 \leqq n < \omega$, is an algebraic function of $\mathbf{A}$. A well known criterion (due to Werner [20]) for functional completeness is that the discriminator of A be an algebraic function of $\mathbf{A}$. (Hence a quasi-primal algebra is obviously functionally complete.) From the remarks above it is clear that a finite SI algebra in a dual discriminator variety is quasi-primal (respectively, functionally complete), if and only if there is a polynomial $m(x, y, z)$ (respectively, algebraic function) of $\mathbf{A}$, satisfying $m(x, x, y) = y$ and $m(x, y, y) = x$. According to the following theorem, if $|A| > 2$ and the dual discriminator is an algebraic function of $\mathbf{A}$, then such an algebraic function $m(x, y, z)$ always exists.

2.3 Theorem. *Let $\mathbf{A} = (A, F)$ be a finite algebra of order greater than 2. The dual discriminator of A is an algebraic function of $\mathbf{A}$ if and only if $\mathbf{A}$ is functionally complete.*

Proof. The "if" direction is trivial. To prove the "only if" direction we establish two claims:

Claim 1. If **A** is a finite algebra having a majority polynomial (i.e.: a ternary polynomial satisfying the equations of Lemma 2.2, ii)) and if $\mathbf{A}\times\mathbf{A}$ has only two subalgebras (the diagonal $\Delta=\{(a, a): a\in A\}$ and $\mathbf{A}\times\mathbf{A}$), then **A** is primal.

Proof of Claim 1. Since **A** has a majority polynomial it follows from [3, Corollary 5.1] that the polynomials of **A** are exactly the functions $f\colon A^n\to A$ such that each subalgebra of $\mathbf{A}\times\mathbf{A}$ is closed under f; i.e.: if **S** is a subalgebra of $\mathbf{A}\times\mathbf{A}$ and $(x_i, y_i)\in S$, $i=1, \dots, n$, then

$$f((x_1, y_1), \dots, (x_n, y_n)) = (f(x_1, \dots, x_n), f(y_1, \dots, y_n))\in S.$$

But since the subalgebras Δ and $\mathbf{A}\times\mathbf{A}$ are closed under any $f\colon A^n\to A$ and since these are the only subalgebras of $\mathbf{A}\times\mathbf{A}$, the claim follows.

Claim 2. Suppose $\mathbf{A}=(A, F)$ is finite with $n>2$ elements and the dual discriminator of A is an algebraic function of **A**. Let **S** be a subalgebra of $\mathbf{A}\times\mathbf{A}$ such that $\Delta\subseteqq S$. Then $S=\Delta$ or $S=A\times A$.

Proof of Claim 2. Let the distinct elements of A be $a_1, \dots, a_n$, $n>2$. Since the dual discriminator of A is an algebraic function q of **A** and since $\Delta\subseteqq S$, q extends (coordinate-wise) to S. Suppose $(a_i, a_j)\in S$ for some $i\neq j$, i.e.: suppose S contains some off-diagonal element. Then for all r,

$$q((a_i, a_i), (a_i, a_j), (a_r, a_r)) = (q(a_i, a_i, a_r), q(a_i, a_j, a_r)) = (a_i, a_r)\in S.$$

Hence for all $s, r, r\neq i$,

$$q((a_r, a_r), (a_i, a_r), (a_s, a_s)) = (q(a_r, a_i, a_s), q(a_r, a_r, a_s)) = (a_s, a_r)\in S.$$

Finally, choose m different from both i and j, which is possible since $n>2$. Then for all s,

$$q((a_s, a_j), (a_s, a_m), (a_i, a_i)) = (q(a_s, a_s, a_i), q(a_j, a_m, a_i)) = (a_s, a_i)\in S.$$

Hence $S=A\times A$ so Claim 2 is proved.

To complete the proof of Theorem 2.3 let **A** satisfy the hypotheses. Let $\mathbf{A}^+$ be the algebra obtained from **A** by adjoining as new nullary operations all elements of A. Then the dual discriminator is a majority polynomial of $\mathbf{A}^+$ and for each subalgebra **S** of $\mathbf{A}^+\times\mathbf{A}^+$, $\Delta\subseteqq S$. By Claim 2, $\mathbf{A}^+\times\mathbf{A}^+$ has only Δ and $\mathbf{A}^+\times\mathbf{A}^+$. as subalgebras. Hence by Claim 1, $\mathbf{A}^+$ is primal. Thus **A** is functionally complete.

Note. The two element lattice, for which the median is the dual discriminator, is not functionally complete — since algebraic functions of lattices are isotone. Hence the condition that **A** be of order greater than 2 is essential.

Quasi-primal algebras were originally defined ([14]) as finite algebras **A** having the following property:

Let $f\colon A^n \to A$ be any function. If each subalgebra $\mathbf{S}$ of $\mathbf{A}\times\mathbf{A}$ with subuniverse of the form $S=\{(x, x\alpha): x\in \operatorname{dom}(\alpha)\}$, α an internal isomorphism of $\mathbf{A}$, is closed under the coordinate-wise extension $f\times f$ of f, then f is a polynomial of $\mathbf{A}$. It is therefore natural to ask for a similar characterization of finite algebras having the dual discriminator as a polynomial. To do this we introduce the following definition:

For an algebra $\mathbf{A}$, a subalgebra $\mathbf{S}$ of $\mathbf{A}\times\mathbf{A}$ is *projectively rectangular* (or, briefly *p-rectangular*) if S has the following two properties:

i) $(x, y_1),\ (x, y_2),\ (u, v)\in S$ and $y_1\neq y_2$ imply $(x, v)\in S$,

ii) $(x_1, y),\ (x_2, y),\ (u, v)\in S$ and $x_1\neq x_2$ imply $(u, y)\in S$.

2.4. Theorem. *For a finite algebra* $\mathbf{A}$ *the following are equivalent:*

a) *The dual discriminator of* A *is a polynomial of* $\mathbf{A}$.

b) *If* $f\colon A^n\to A$ *is any function such that each p-rectangular subalgebra of* $\mathbf{A}\times\mathbf{A}$ *is closed under the coordinate-wise extension* $f\times f$ *of* f, *then* f *is a polynomial of* $\mathbf{A}$.

Proof. a)$\Rightarrow$b). The dual discriminator $q^{\mathbf{A}}$ is a majority polynomial of $\mathbf{A}$. Hence, by [3, Corollary 5.1], the polynomials of $\mathbf{A}$ are just the functions $f\colon A^n\to A$ such that all subalgebras $\mathbf{S}$ of $\mathbf{A}\times\mathbf{A}$ are closed under f. Hence we need only show that each subalgebra $\mathbf{S}$ of $\mathbf{A}\times\mathbf{A}$ is p-rectangular. But if $(x, y_1),\ (x, y_2),\ (u, v)\in S$ and $y_1\neq y_2$, then $q^{\mathbf{A}}((x, y_1), (x, y_2), (u, v))=(x, v)\in S$, and if $(x_1, y),\ (x_2, y),\ (u, v)\in S$ and $x_1\neq x_2$, then $q^{\mathbf{A}}((x_1, y), (x_2, y), (u, v))=(u, y)\in S$.

b)$\Rightarrow$a). Let $\mathbf{S}$ be a p-rectangular subalgebra of $\mathbf{A}\times\mathbf{A}$ and let $(x_1, y_1),\ (x_2, y_2),\ (x_3, y_3)\in S$. Then

$$\begin{aligned} d((x_1, y_1), (x_2, y_2), (x_3, y_3)) &= (d(x_1, x_2, x_3),\ d(y_1, y_2, y_3)) = \\ &= (x_1, y_1) \quad \text{if} \quad x_1 = x_2 \quad \text{and} \quad y_1 = y_2 \\ &= (x_1, y_3) \quad \text{if} \quad x_1 = x_2 \quad \text{and} \quad y_1 \neq y_2 \\ &= (x_3, y_1) \quad \text{if} \quad x_1 \neq x_2 \quad \text{and} \quad y_1 = y_3 \\ &= (x_3, y_3) \quad \text{if} \quad x_1 \neq x_2 \quad \text{and} \quad y_1 \neq y_3, \end{aligned}$$

hence $\mathbf{S}$ is closed under the dual disciminator d. Thus d is a polynomial of $\mathbf{A}$.

Note. The conditions defining a p-rectangular subalgebra are equivalent to the following (p_1, p_2 are, respectively, the first and second projections):

i)′ If S contains 2 points of $\{x\}\times Sp_2$ then S contains $\{x\}\times Sp_2$,

ii)′ If S contains 2 points of $Sp_1\times\{y\}$ then S contains $Sp_1\times\{y\}$.

To compare quasi-primal algebras with algebras having the dual discriminator as a polynomial, let $\mathbf{A}$ be quasi-primal and suppose $\mathbf{S}$ is a subalgebra of $\mathbf{A}\times\mathbf{A}$ such that S contains two points, $(x, y_1), (x, y_2)$ of $\{x\}\times Sp_2$. Then for any $(u_1, v_1), (u_2, v_2)\in S$,

$f((x, y_1), (x, y_2), (u_1, v_1), (u_2, v_2)) = (u_1, v_2) \in S$ if f is the 4-ary discriminator (1.5). Hence $S = Sp_1 \times Sp_2$. On the other hand, Theorem 2.3 shows that for sets with at least three elements, the constant functions together with either the discriminator or the dual discriminator, generate all functions.

3. Characterization and properties of dual discriminator varieties

The following theorem, which gives an equational characterization of discriminator varieties, appears in McKenzie [11].

3.1 Theorem. *For a variety V and ternary polynomial symbol $p(x, y, z)$, the following are equivalent:*

1) *V is a discriminator variety with $p(x, y, z)$ the discriminator on each SI member of V.*

2) *The following are equations of V:*

a) $p(x, z, z) = x, \quad p(x, y, x) = x, \quad p(x, x, z) = z,$

b) $p(x, p(x, y, z), y) = y,$

c) *for each operation symbol f of V,*

$$p(x, y, f(z_1, \dots, z_k)) = p(x, y, f(p(x, y, z_1), \dots, p(x, y, z_k)))$$

(where f is k-ary).

The proof depends essentially on observing that for any $\mathbf{A} \in V$ and $a, b \in A$, the principal congruence $\theta(a, b)$ is given by

$$\theta(a, b) = \{(x, y) \in A \times A : p^{\mathbf{A}}(a, b, x) = p^{\mathbf{A}}(a, b, y)\}.$$

For dual discriminator varieties we have the following corresponding result:

3.2 Theorem. *For a variety V and ternary polynomial symbol $q(x, y, z)$, the following are equivalent:*

1) *V is a dual discriminator variety with $q(x, y, z)$ the dual discriminator on each SI member of V.*

2) *The following are equations of V:*

a) $q(x, z, z) = z, \quad q(x, y, x) = x, \quad q(x, x, z) = x,$

b) $q(x, y, q(x, y, z)) = q(x, y, z),$

c) $q(z, q(x, y, z), q(x, y, w)) = q(x, y, z),$

d) *for each operation symbol f of V,*

$$q(x, y, f(z_1, \dots, z_k)) = q(x, y, f(q(x, y, z_1), \dots, q(x, y, z_k)))$$

(where f is k-ary).

If V is an idempotent variety the equations d) *may be replaced by*

d′) $q(x, y, f(z_1, \ldots, z_k)) = f(q(x, y, z_1), \ldots, q(x, y, z_k))$ [1].

Proof. If 1) holds then it is easy to check that a)—d) are equations of each SI member of V and hence of V. If V is idempotent then 1) clearly implies d′).

Conversely, suppose a)—d) are equations of V. For each $\mathbf{A} \in V$ and $a, b \in A$, define the *co-principal* congruence $\gamma(a, b)$ by

$$\gamma(a, b) = \{(x, y) \in A \times A\colon\ q^{\mathbf{A}}(a, b, x) = q^{\mathbf{A}}(a, b, y)\}.$$

By d) (or d′)) if V is idempotent) $\gamma(a, b)$ is easily seen to be a congruence of $\mathbf{A}$. Next observe that, by b),

$$(3.3) \qquad (q^{\mathbf{A}}(x, y, z), z) \in \gamma(x, y) \quad \text{for all} \quad x, y, z \in A.$$

Now $q^{\mathbf{A}}(x, x, z) = x$ by a). Hence to complete the proof it will suffice to show that if $\mathbf{A}$ is SI and $x, y \in A$, then

$$(3.4) \qquad x \neq y \quad \text{implies} \quad \gamma(x, y) = \omega,$$

for, by (3.3), this will mean $q^{\mathbf{A}}(x, y, z) = z$ if $x \neq y$.

As a preliminary we first establish the following implication:

$$(3.5) \quad x \neq y \quad \text{implies} \quad \gamma(q^{\mathbf{A}}(x, y, z), z) \neq \omega, \quad \text{for any} \quad \mathbf{A} \in V \quad \text{and} \quad x, y, z \in A.$$

To prove (3.5) we observe that, by c), we have

$$q^{\mathbf{A}}(z, q^{\mathbf{A}}(x, y, z), q^{\mathbf{A}}(x, y, w)) = q^{\mathbf{A}}(z, q^{\mathbf{A}}(x, y, z), q^{\mathbf{A}}(x, y, z))$$

and hence $(q^{\mathbf{A}}(x, y, z), q^{\mathbf{A}}(x, y, w)) \in \gamma(z, q^{\mathbf{A}}(x, y, z))$ for all $x, y, z, w \in A$. If $\gamma(z, q^{\mathbf{A}}(x, y, z)) = \omega$ for some x, y, z, then $q^{\mathbf{A}}(x, y, z) = q^{\mathbf{A}}(x, y, w)$ for all $w \in A$. In particular, using a), we would then have

$$x = q^{\mathbf{A}}(x, y, x) = q^{\mathbf{A}}(x, y, z) = q^{\mathbf{A}}(x, y, y) = y.$$

This establishes (3.5).

Now let $\mathbf{A}$ be SI in V. Choose $a, b \in A$, $a \neq b$, such that $(a, b) \in \theta$ for all congruences $\theta \neq \omega$. Let us suppose we have a pair $x, y \in A$ contradicting (3.4), i.e.: such that $x \neq y$ and $\gamma(x, y) \neq \omega$. Then $(a, b) \in \gamma(x, y)$ so $q^{\mathbf{A}}(x, y, a) = q^{\mathbf{A}}(x, y, b)$. Denote the common value of these expressions by $c \in A$. By (3.5) we have

$$\gamma(a, q^{\mathbf{A}}(x, y, a)) = \gamma(a, c) \neq \omega \quad \text{and} \quad \gamma(b, q^{\mathbf{A}}(x, y, b)) = \gamma(b, c) \neq \omega$$

Hence $(a, b) \in \gamma(a, c) \cap \gamma(b, c)$ so that, by a),

$$a = q^{\mathbf{A}}(a, c, a) = q^{\mathbf{A}}(a, c, b) \quad \text{and} \quad b = q^{\mathbf{A}}(b, c, b) = q^{\mathbf{A}}(b, c, a).$$

[1]) In Theorem 3.1, if V is idempotent, the equations c) may be analogously simplified.

Since $a \neq b$, we have $c \neq a$ or $c \neq b$. If $c \neq a$ then, taking $x=a$, $y=c$, $z=b$ in (3.5), we obtain $\gamma(q^{\mathbf{A}}(a, c, b), b)=\gamma(a, b) \neq \omega$. Hence $(a, b) \in \gamma(a, b)$ which implies $a=q^{\mathbf{A}}(a, b, a)=q^{\mathbf{A}}(a, b, b)=b$, a contradiction. If $c \neq b$ then, taking $x=b$, $y=c$, $z=a$ in (3.5), we obtain $\gamma(q^{\mathbf{A}}(b, c, a), a)=\gamma(b, a) \neq \omega$ which again leads to the contradiction $a=b$. Hence (3.4) is established, completing the proof.

Notice that on SI members of a dual discriminator variety we have:

$$\theta(a, b) = \omega \quad \text{if} \quad a = b, \quad \text{and} \quad = \iota \quad \text{if} \quad a \neq b,$$

$$\gamma(a, b) = \iota \quad \text{if} \quad a = b, \quad \text{and} \quad = \omega \quad \text{if} \quad a \neq b.$$

We compare these congruences more closely. First observe that in a dual discriminator variety V, $q(x, y, u)$ and $q(x, y, v)$ are *principal intersection polynomials* in the sense of BAKER [1], i.e.: the polynomial symbols $D_1(x, y, u, v)=q(x, y, u)$ and $D_2(x, y, u, v)=q(x, y, v)$ have the property that on any SI member $\mathbf{A}$ of V,

$$D_1^{\mathbf{A}}(x, y, u, v) = D_2^{\mathbf{A}}(x, y, u, v) \quad \text{iff} \quad x = y \quad \text{or} \quad u = v. \tag{3.6}$$

From [1] it then follows that for any $\mathbf{A} \in V$, the meet of principal congruences $\theta(a, b)$ and $\theta(c, d)$ is principal and is given by

$$\theta(a, b) \wedge \theta(c, d) = \theta(q^{\mathbf{A}}(a, b, c),\ q^{\mathbf{A}}(a, b, d)). \tag{3.7}$$

Using this observation we have

3.8 Theorem. *Let $\mathbf{A}$ be any algebra in a dual discriminator variety. For any $a, b \in A$, the principal and co-principal congruences $\theta(a, b)$ and $\gamma(a, b)$ are complements and, in particular,*

$$\gamma(a, b) \circ \theta(a, b) \circ \gamma(a, b) = \iota \quad (\circ \text{ denotes relation product}).$$

Proof. For all $x, y \in A$, using equations a), b) of Theorem 3.2, we have

$$x\ \gamma(a, b)\ q^{\mathbf{A}}(a, b, x)\ \theta(a, b)\ q^{\mathbf{A}}(a, a, x) = a = q^{\mathbf{A}}(a, a, y)\ \theta(a, b)\ q^{\mathbf{A}}(a, b, y)\ \gamma(a, b)\ y$$

Hence $\gamma(a, b) \circ \theta(a, b) \circ \gamma(a, b)=\iota$. For the meet, if $(x, y) \in \theta(a, b) \wedge \gamma(a, b)$ then $(x, y) \in \theta(a, b) \wedge \theta(x, y)=\theta(q^{\mathbf{A}}(a, b, x), q^{\mathbf{A}}(a, b, y))$ by (3.7). But $q^{\mathbf{A}}(a, b, x)= =q^{\mathbf{A}}(a, b, y)$ since $(x, y) \in \gamma(a, b)$. Hence $x=y$, whence $\theta(a, b) \wedge \gamma(a, b)=\omega$.

3.9 Corollary. *In a dual discriminator variety the join of co-principal congruences is co-principal and is given by*

$$\gamma(a, b) \vee \gamma(c, d) = \gamma(q^{\mathbf{A}}(a, b, c),\ q^{\mathbf{A}}(a, b, d)). \tag{3.10}$$

Proof. Apply Theorem 3.8 and congruence distributivity, taking complements of both sides of (3.7).

Theorem 3.8 has several important consequences. Recall from [12] that a variety V has *definable principal congruences* if there is a formula $\beta(u, v, x, y)$ in the firs-order language of V such that for all $\mathbf{A}\in V$ and $a, b, c, d\in A$,

$$(c, d)\in\theta(a, b) \quad \text{iff} \quad \mathbf{A}\models\beta(a, b, c, d).$$

A stronger concept, *equationally definable principal congruences,* was introduced in [8].

3.11 Theorem. *If V is a dual discriminator variety, then V has equationally definable principal congruences. In particular, for $\mathbf{A}\in V$, $a, b, c, d\in A$,*

$$(3.12) \qquad (c, d)\in\theta(a, b) \quad \text{iff} \quad A\models(\forall u)[q(c, d, u) = q(c, d, q(a, b, u))].$$

Proof. First notice that

$$(c, d)\in\theta(a, b) \quad \text{iff} \quad \theta(c, d)\leqq\theta(a, b) \quad \text{iff} \quad \gamma(a, b)\leqq\gamma(c, d)$$

by Theorem 3.8. But $\gamma(a, b)\leqq\gamma(c, d)$ is equivalent to the condition

$$(3.13) \quad (\forall u_1, u_2\in A)[q^{\mathbf{A}}(a, b, u_1) = q^{\mathbf{A}}(a, b, u_2) \Rightarrow q^{\mathbf{A}}(c, d, u_1) = q^{\mathbf{A}}(c, d, u_2)].$$

Clearly the right side of (3.12) implies (3.13) and, taking $u_2=q^{\mathbf{A}}(a, b, u_1)$, the right side of (3.12) follows from (3.13) and equation b) of Theorem 3.2.

3.14 Corollary. *If V is a dual discriminator variety and $\mathbf{A}\in V$ is a subdirect product of algebras $\mathbf{A}_i, i\in I$, then for $a, b, c, d\in A$,*

$$(3.15) \qquad (c, d)\in\theta(a, b) \quad \text{iff} \quad (\forall i\in I)[(c_i, d_i)\in\theta(a_i, b_i)].$$

Corollary 3.14 is immediate from the universal form of the formula appearing in (3.12); it asserts that V has *factor determined principal congruences* in the sense of [8].

From [8, Theorem 4.5] we also obtain:

3.16 Corollary. *A dual discriminator variety has the congruence extension property.*

McKenzie [12] has shown that if a variety V of finite type has only finitely many SI members, all finite, and has definable principal congruences, then V has a finite equational base. On the other hand Baker's finite basis theorem [2] asserts that a congruence distributive variety of finite type generated by a finite algebra always has a finite equational base. Hence if $\mathbf{A}$ is a finite algebra in a dual discriminator variety (of finite type), then $\mathbf{A}$ has a finite equational base either as a result of Baker's theorem and Lemma 2.2, ii), or more briefly, from McKenzie's result and Theorem 3.11. Despite the fact that we have these two proofs it is still instructive to establish the result directly, illustrating McKenzie's method. We do this as follows:

If $d(x_0, x_1, x_2)$ is the dual discriminator of a set, define inductively

$$\begin{aligned} d_1(x_0, x_1, x_2) &= d(x_0, x_1, x_2), \\ d_2(x_0, x_1, x_2, x_3) &= d(x_0, d_1(x_0, x_1, x_2), x_3), \\ &\vdots \\ d_n(x_0, \ldots, x_{n+1}) &= d(x_0, d_{n-1}(x_0, \ldots, x_n), x_{n+1}), \end{aligned}$$

and observe that

$$\begin{aligned} d_n(x_0, \ldots, x_{n+1}) &= x_0 \quad \text{if} \quad x_0 \text{ equals any of } x_1, \ldots, x_n, \\ &= x_{n+1} \quad \text{otherwise.} \end{aligned}$$

It follows that on any set S the sentence

$$(\forall x_0, \ldots, x_n) \bigvee_{0 \leqq i < j \leqq n} (x_i = x_j), \quad (\text{meaning } |S| \leqq n),$$

is true in S if and only if the following equation N_n holds in S:

$$\begin{aligned} &d_n(x_0, \ldots, x_n, d_{n-1}(x_1, \ldots, x_n, d_{n-2}(x_2, \ldots, d_2(x_{n-2}, x_{n-1}, x_n, x_n) \ldots) = \\ &= d_n(x_0, \ldots, x_n, d_{n-1}(x_1, \ldots, x_n, d_{n-2}(x_2, \ldots, d_2(x_{n-2}, x_{n-1}, x_n, x_{n-1}) \ldots). \end{aligned}$$

For example, N_3 is

$$d_3(x_0, x_1, x_2, x_3, d_2(x_1, x_2, x_3, x_3)) = d_3(x_0, x_1, x_2, x_3, d_2(x_1, x_2, x_3, x_2)).$$

Now let V be a dual discriminator variety of finite type and let $\mathbf{A}$ be a finite algebra in V. By congruence distributivity and either [13, Theorem 2.5] or [10], the variety generated by $\mathbf{A}$ is

$$V(\mathbf{A}) = IP_S HS(\mathbf{A})$$

and thus we can effectively determine from the finitely many isomorphism types of $HS(\mathbf{A})$, all of the isomorphism types of the SI members of $V(\mathbf{A})$. Let these be $K_0 = \{\mathbf{A}_1, \ldots, \mathbf{A}_k\}$ and let $n = \max\{|A_i| : i = 1, \ldots, k\}$. Then the equations a), b), c), d) of Theorem 3.2, together with N_n (with q_i replacing $d_i, i = 1, \ldots, n$) are an equational base for the variety generated by

$K_1 = \{\mathbf{B} : \mathbf{B}$ is SI of the given type, $|B| \leqq n$, and $q^{\mathbf{B}}$ is the dual discriminator of $\mathbf{B}\}$.

To this extent the basis is canonical. Next we can obviously effectively list the members of K_1 and, by congruence distributivity, for each $\mathbf{B}_i$ in K_1, either $\mathbf{B}_i$ is isomorphic with some algebra of K_0 or there is an equation e_i which is an identity of each member of K_0 but not of $\mathbf{B}_i$. Let $e_{i(1)}, \ldots, e_{i(m)}$, $m \geqq 0$, be such "exclusion" equations, which can clearly be effectively determined. Then the equations a), b), c), d), $N_n, e_{i(1)}, \ldots, e_{i(m)}$ are a finite equational base for $\mathbf{A}$.

In Section 5 we shall apply Theorem 3.2 even more directly to obtain explicit equational bases for the varieties a) generated by all weakly associative lattices having the unique bound property, and b) the variety generated by the triangle algebra $\mathbf{W}_3$.

4. Dual discriminator varieties and the discriminator

Recall that a congruence permutable distributive lattice is necessarily relatively complemented. (Of course the converse is always true whether the lattice is distributive or not.) The following theorem generalizes this fact to dual discriminator varieties. (Cf. Lemma 2.2.)

4.1 Lemma. *Let* **A** *be any algebra in a dual discriminator variety and let* $f\colon A^3 \to A$ *be any function which is compatible with all of the congruences of* **A** *and which satisfies the following equations for all* $x, y, z \in A$:

$$q^{\mathbf{A}}(x, y, f(x, y, z)) = x, \quad f(x, x, z) = z.$$

Then f *induces the discriminator on each SI homomorphic image of* **A**.

Proof. $q^{\mathbf{A}}(x, y, f(x, y, z)) = x = q^{\mathbf{A}}(x, y, x)$ so that $(f(x, y, z), x) \in \gamma(x, y)$ in any SI homomorphic image. But if $x \neq y$, $\gamma(x, y) = \omega$, so that $f(x, y, z) = x$.

4.2 Theorem. *If* **A** *is any algebra in a dual discriminator variety the following are equivalent:*

a) **A** *is congruence permutable.*

b) *There is a ternary function* $f\colon A^3 \to A$ *which is compatible with all congruences of* **A** *and which induces the discriminator on each SI homomorphic image of* **A**.

Example. If **A** is a congruence permutable distributive lattice, then, by Theorem 4.2, such a function f exists and induces the discriminator on the two element lattice. Hence on $\{0, 1\}$

$$f(1, y, 0) = f(0, y, 1) = y' \quad \text{(complement)}.$$

From this it follows that for y in the interval $[x, z]$ of **A**, $f(x, y, z)$ is the relative complement of $y = q^{\mathbf{A}}(x, y, z) = (x \wedge y) \vee (y \wedge z) \vee (x \wedge z)$. (Cf. (1.1) and (1.2).)

Proof of Theorem 4.2. Suppose **A** is congruence permutable and $x, y, z \in A$. Then (as noted in the proof of Theorem 3.8),

$$x = q^{\mathbf{A}}(x, x, z)\ \theta(x, y)\ q^{\mathbf{A}}(x, y, z)\ \gamma(x, y)\ z.$$

Hence, by permutability, there is a $c \in A$ such that

$$x\ \gamma(x, y)\ c\ \theta(x, y)\ z,$$

which means $[x]\gamma(x, y) \cap [z]\theta(x, y) \neq \emptyset$. Thus, by the Axiom of Choice, there is a function $f\colon A^3 \to A$ such that for each $x, y, z \in A$,

$$x\ \gamma(x, y)\ f(x, y, z)\ \theta(x, y)\ z.$$

Hence,

$$\text{i) } q^{\mathbf{A}}(x, y, f(x, y, z)) = q^{\mathbf{A}}(x, y, x) = x, \quad \text{ii) } f(x, y, z)\ \theta(x, y)\ z.$$

To show that f is compatible with the congruences of $\mathbf{A}$ first let φ be a completely meet irreducible congruence (i.e.: such that $\mathbf{A}/\varphi$ is SI). Let (x, x_1), (y, y_1), $(z, z_1) \in \varphi$. If $(x, y) \in \varphi$ then $(x_1, y_1) \in \varphi$, so $\theta(x, y) \leqq \varphi$ and $\theta(x_1, y_1) \leqq \varphi$ and, by ii), $f(x, y, z)\ \varphi\ z\ \varphi\ z_1\ \varphi\ f(x_1, y_1, z_1)$. If $(x, y) \notin \varphi$ then $(x_1, y_1) \notin \varphi$ so that

$$f(x, y, z)\ \varphi\ q^{\mathbf{A}}(x, y, f(x, y, z)) = x\ \varphi\ x_1 = q^{\mathbf{A}}(x_1, y_1, f(x_1, y_1, z_1))\ \varphi\ f(x_1, y_1, z_1)$$

by i). Hence f is compatible with φ. Since any congruence is the meet of completely meet irreducible congruences, it easily follows that f is compatible with all congruences of A. Hence f meets the conditions of Lemma 4.1 so that it induces the discriminator on each SI homomorphic image of A.

Conversely, if f satisfying b) exists and θ_1, θ_2 are congruences of $\mathbf{A}$ with $x\ \theta_1\ y\ \theta_2\ z$ then, by the compatibility of f,

$$x = f(x, z, z)\ \theta_2\ f(x, y, z)\ \theta_1\ f(x, x, z) = z$$

so that $\mathbf{A}$ is congruence permutable.

Finally we observe that in any discriminator variety V (which by Lemma 2.2 is necessarily a dual discriminator variety) we have, in addition to formulas (3.6), (3.7), and (3.10), their duals. Indeed, we may call the polynomial symbols $p(x, y, u)$ and $p(x, y, v)$ *principal join polynomials,* since for SI algebras $\mathbf{A} \in V$,

(3.6)′ $\qquad p^{\mathbf{A}}(x, y, u) = p^{\mathbf{A}}(y, x, v)$ iff $x = y$ and $u = v$.

(Note the reversal of x and y.) Using (3.6)′ we can deduce

(3.7)′ $\qquad \theta(a, b) \vee \theta(c, d) = \theta(a, b) \circ \theta(c, d) = \theta(p^{\mathbf{A}}(a, b, c),\ p^{\mathbf{A}}(b, a, d))$.

($\vee = \circ$ since V is congruence permutable by Lemma 2.2.)

To prove (3.7)′ we observe that by the remark following Theorem 3.1,

$(x, z) \in \theta(a, b) \circ \theta(c, d)$ iff $(\exists y)[p^{\mathbf{A}}(a, b, x) = p^{\mathbf{A}}(a, b, y)$ and $p^{\mathbf{A}}(c, d, y) = p^{\mathbf{A}}(c, d, z)]$.

Hence on each SI factor $\mathbf{A}_i$ of $\mathbf{A}$,

$$p^{\mathbf{A}_i}(a_i, b_i, x_i) = p^{\mathbf{A}_i}(a_i, b_i, y_i) \quad \text{and} \quad p^{\mathbf{A}_i}(c_i, d_i, y_i) = p^{\mathbf{A}_i}(c_i, d_i, z_i),$$

from which we directly infer

(4.3) $\quad p^{\mathbf{A}_i}(p^{\mathbf{A}_i}(a_i, b_i, c_i),\ p^{\mathbf{A}_i}(b_i, a_i, d_i),\ x_i) = p^{\mathbf{A}_i}(p^{\mathbf{A}_i}(a_i, b_i, c_i),\ p^{\mathbf{A}_i}(b_i, a_i, d_i),\ z_i)$,

using (3.6)′. From (4.3) we have $(x, z) \in \theta(p^{\mathbf{A}}(a, b, c), p^{\mathbf{A}}(b, a, d))$ by Corollary 3.14.

Conversely, if $(x, z) \in \theta(p^{\mathbf{A}}(a, b, c), p^{\mathbf{A}}(b, a, d))$ then (4.3) holds on each SI factor $\mathbf{A}_i$ of $\mathbf{A}$. Thus if $a_i = b_i$ and $c_i = d_i$, then $x_i = z_i$ while $\theta(a_i, b_i) = \iota$ if $a_i \neq b_i$ and likewise for $c_i \neq d_i$. From this it follows that

$$x_i\ \theta(a_i, b_i)\ p^{\mathbf{A}_i}(p^{\mathbf{A}_i}(a_i, b_i, x_i),\ p^{\mathbf{A}_i}(a_i, b_i, z_i),\ z_i)\ \theta(c_i, d_i)\ z_i$$

since the middle term is x_i if $a_i=b_i$ and z_i if $a_i\neq b_i$. Hence, by Corollary 3.14,

$$x\,\theta(a,b)\,p^{\mathbf{A}}(p^{\mathbf{A}}(a,b,x),\,p^{\mathbf{A}}(a,b,z),\,z)\,\theta(c,d)\,z,$$

establishing (3.7)′. Complementing both sides of (3.7)′ we obtain

$$\text{(3.10)}' \qquad \gamma(a,b)\wedge\gamma(c,d)=\gamma(p^{\mathbf{A}}(a,b,c),\,p^{\mathbf{A}}(b,a,d)).$$

Formulas (3.6)′ and (3.7)′ together with some additional properties of principal congruences in discriminator varieties can also be found in [4].

5. Weakly associative lattices

Recall from [7] that an algebra $\mathbf{A}=(A,\vee,\wedge)$ is a *weakly associative lattice* (WAL) if the operations $\vee$ and $\wedge$ are binary and satisfy the following identities in A:

$$\text{(5.1)}\qquad \begin{array}{lll} x\vee x=x, & x\wedge x=x, & \text{(idempotence)}\\ x\vee y=y\vee x, & x\wedge y=y\wedge x, & \text{(commutativity)}\\ x\wedge(x\vee y)=x, & x\vee(x\wedge y)=x, & \text{(absorption)}\\ \multicolumn{2}{c}{\left.\begin{array}{l}((x\wedge z)\vee(y\wedge z))\vee z=z\\ ((x\vee z)\wedge(y\vee z))\wedge z=z\end{array}\right\}} & \text{(weak associativity)}\end{array}$$

WALs have also been called *trellises* by SKALA [17]. *Tournaments* ([7]) and the algebras $\mathbf{W}_n$ of Section 2 are special cases of WALs. A WAL has the *unique bound property* (UBP) ([5]) if for distinct $a,b\in A$, $a\leqq c$ and $b\leqq c$ imply $c=a\vee b$ and, dually, $d\leqq a$, $d\leqq b$ imply $d=a\wedge b$. For brevity we shall call a WAL with the UBP simply a UBP. In [6] it was shown that for a WAL $\mathbf{A}$ the following are equivalent:

a) $\mathbf{A}$ is a UBP,

b) $\mathbf{A}$ is SI and satisfies the congruence extension property.

c) Each subalgebra of $\mathbf{A}$ is simple.

The following theorem adds a new equivalence to this list. Combined with Theorem 2.3 it also provides a new proof that a finite UBP of more than two elements is functionally complete. (See [5] for the original proof.)

5.2 Theorem. *A WAL $\mathbf{A}$ is a UBP if and only if the dual discriminator is a polynomial of $\mathbf{A}$. In particular the WAL polynomial symbol $q_u(x,y,z)$, explicitly constructed in the proof below, has the property: For any UBP $\mathbf{A}$, $q_u^{\mathbf{A}}(x,y,z)$ is the dual discriminator of A.*

Proof. If the dual discriminator is a polynomial q of $\mathbf{A}$ and if $\mathbf{A}$ is not a UBP, then for some $a,b\in A$, $a\neq b$, $a\leqq c$, $b\leqq c$, and $a\vee b<c$; while $q(c,a\vee b,a)=a$. But for the three element chain with elements $\{a,a\vee b,c\}$ the mapping α onto $\mathbf{2}$

given by: $a\alpha=0$, $(a\vee b)\alpha=c\alpha=1$, is a homomorphism. But since $q(1,1,0)=1$, this is a contradiction. Hence, **A** must have unique upper bounds, and dually, unique lower bounds. (Essentially the same proof is used in [5, Theorem 4].)

Conversely, define the polynomial symbols h, h', g by

$$h(x,y,z) = (z\wedge y)\vee(((z\wedge x)\vee y)\wedge x),$$
$$h'(x,y,z) = (z\vee x)\wedge(((z\vee y)\wedge x)\vee y),$$
$$g(x,y,z) = h(h(x,y,z),\, h'(x,y,z),\, z).$$

Let **A** be any UBP. Then h and h' are easily seen to induce majority functions on A and hence so does g.

Consider a pair $a, b\in A$ such that $a<b$. Then

$$a < b \leqq (c\wedge a)\vee b \quad \text{for all} \quad c\in A. \tag{5.3}$$

Since $c\wedge a$ is a lower bound for the elements a and $(c\wedge a)\vee b$, which by (5.3) are distinct, we have

$$((c\wedge a)\vee b)\wedge a = c\wedge a,$$

because **A** is a UBP. Therefore $h^{\mathbf{A}}(a,b,c)=(c\wedge b)\vee(c\wedge a)$ for $a<b$. Thus, since c is an upper bound for both $c\wedge b$ and $c\wedge a$,

$$h^{\mathbf{A}}(a,b,c) = c \quad \text{unless} \quad c\wedge b = c\wedge a < c.$$

In the latter case both $c\wedge a$ and a are lower bounds for the distinct elements a and b. Hence $c\wedge a=a$ (which means $a\leqq c$). Also $c\wedge b=c\wedge a$ implies $b\neq c$. Hence we have

$$\left.\begin{aligned} h^{\mathbf{A}}(a,b,c) &= a \quad \text{if} \quad a=b \quad \text{or} \quad b\neq c>a \\ &= c \quad \text{otherwise} \end{aligned}\right\} \quad \text{for} \quad a\leqq b \quad \text{and arbitrary } c. \tag{5.4}$$

Dually, from the definition of h', we obtain

$$\left.\begin{aligned} h'^{\mathbf{A}}(a,b,c) &= b \quad \text{if} \quad a=b \quad \text{or} \quad a\neq c<b \\ &= c \quad \text{otherwise} \end{aligned}\right\} \quad \text{for} \quad a\leqq b \quad \text{and arbitrary } c. \tag{5.4$'$}$$

Now consider $g^{\mathbf{A}}(a,b,c)$ for $a\leqq b$. By (5.4) and (5.4)$'$ and the fact that $h^{\mathbf{A}}$ is a majority function, we have four cases: $g^{\mathbf{A}}(a,b,c)$ equals one of $h^{\mathbf{A}}(a,b,c)$, $h^{\mathbf{A}}(a,c,c)=c$, $h^{\mathbf{A}}(c,b,c)=c$, or $h^{\mathbf{A}}(c,c,c)=c$, i.e.:

$$g^{\mathbf{A}}(a,b,c) = h^{\mathbf{A}}(a,b,c) \quad \text{or} \quad c.$$

The case $g^{\mathbf{A}}(a,b,c)=h^{\mathbf{A}}(a,b,c)$ occurs when $a=b$ (yielding $g^{\mathbf{A}}(a,b,c)=b$) or when $b\neq c>a$ and $a\neq c<b$. Since **A** is a UBP these two inequalities together with $a<b$ yield the contradiction $a=b$. Hence we have

$$g^{\mathbf{A}}(a,b,c) = b \quad \text{if} \quad a=b, \quad \text{and} \quad = c \quad \text{if} \quad a<b. \tag{5.5}$$

Now let $u(x, y, z)$ be any WAL ternary majority polynomial symbol. The simplest is apparently

$$u(x, y, z) = [(x \wedge z) \vee (y \wedge z)] \vee (x \wedge y).$$

Put

$$f(x, y, z, w) = u\big(g(x, y, z),\ g(y, w, z), z\big).$$

Since $u^{\mathbf{A}}$ is a majority function, if $a \leqq b \leqq d$ and c is arbitrary, from (5.5) we obtain:

$$\begin{aligned} f^{\mathbf{A}}(a, b, c, d) &= u^{\mathbf{A}}(b, b, c) = b \quad \text{if} \quad a = b = d, \\ &= u^{\mathbf{A}}(b, c, c) = c \quad \text{if} \quad a = b < d, \\ &= u^{\mathbf{A}}(c, b, c) = c \quad \text{if} \quad a < b = d, \\ &= u^{\mathbf{A}}(c, c, c) = c \quad \text{if} \quad a < b < d. \end{aligned}$$

Thus for $a \leqq b \leqq d$ and c arbitrary we have

(5.6) $\qquad f^{\mathbf{A}}(a, b, c, d) = b \quad \text{if} \quad a = b = d, \quad \text{and} \quad = c \quad \text{otherwise.}$

Finally, put

$$q_u(x, y, z) = f(x \wedge y, x, z, x \vee y).$$

By (5.6) we have

$$q_u^{\mathbf{A}}(a, b, c) = a \quad \text{if} \quad a \wedge b = a = a \vee b, \quad \text{and} \quad = c \quad \text{otherwise.}$$

Since $a \wedge b = a = a \vee b$ is equivalent to $a = b$, $A_u^{\mathbf{A}}$ is the dual discriminator of A.

Next we shall prove that the explicit polynomial symbol given in Section 2 is the dual discriminator for the triangle algebra $\mathbf{W}_3$. In fact we prove more.

5.7 Theorem. *The polynomial symbol*

$$q_t(x, y, z) = \big[(z \wedge (x \wedge y)) \vee (x \vee y)\big] \wedge [z \vee (x \wedge y)]$$

induces the dual discriminator on the triangle algebra $\mathbf{W}_3$ *and on no other* WAL *of more than two elements.*

Proof. It is routine to check that $q_t^{\mathbf{A}}$ is a majority polynomial on any WAL $\mathbf{A}$. Hence on $\mathbf{W}_3$ $q_t^{\mathbf{W}_3}$ agrees with the dual discriminator if any two of its arguments are equal. Otherwise, using symmetry, we may suppose $z = a_1$ and either $x = 0$, $y = 1$, or $x = 1$, $y = 0$. In both cases $x \wedge y = 0$ and $x \vee y = 1$. Thus

$$q_t^{\mathbf{W}_3}(x, y, z) = [(a_1 \wedge 0) \vee 1] \wedge [a_1 \vee 0] = a_1 = z$$

so that q_t induces the dual discriminator on $\mathbf{W}_3$.

To complete the proof observe that if $\mathbf{A}$ is a WAL which is not $\mathbf{W}_3$ and has more than two elements, then it must contain an incomparable pair b and c. Put $a = b \wedge c$. Then $a \wedge b = a$ and $a \vee b = b$, so

$$q_t^{\mathbf{A}}(a, b, c) = [(c \wedge a) \vee b] \wedge [c \vee a] = b \wedge c = a \neq c.$$

But $a \neq b$ since b and c are incomparable. Hence $q_t^{\mathbf{A}}$ fails to be the dual discriminator on A.

In contrast to the unique property of the polynomial symbol q_t expressed by Theorem 5.7, any polynomial symbol which induces the dual discriminator on $\mathbf{W}_5$ also induces the dual discriminator on every $\mathbf{W}_n$, $n>5$, This is so since each 3-generated subalgebra of any such $\mathbf{W}_n$ is evidently isomorphic to a subalgebra of $\mathbf{W}_5$. (See Problem 2, Section 6.)

Theorems 5.6 and 5.7 together with Theorem 3.2 yield the following result immediately. (Since it is routine to check that q_u and q_t induce majority polynomials in any WAL, equations a) of Theorem 3.2 are omitted from ii) below. Also we may use d′) of Theorem 3.2 since WALs are idempotent.)

5.8. Theorem. *Let B denote the set of identities whose members are the following:*

i) *the identities* (5.1) *defining WALs,*

ii) *the identities (from Theorem* 3.2*):*

$$q(x, y, q(x, y, z)) = q(x, y, z), \quad q(z, q(x, y, z), q(x, y, w) = q(x, y, z)),$$

$$q(x, y, z \vee w) = q(x, y, z) \vee q(x, y, w), \quad q(x, y, z \wedge w) = q(x, y, z) \wedge q(x, y, w).$$

Then the set B_u, obtained from B by taking for q the polynomial symbol q_u defined in Theorem 5.3, *is an equational base for the variety U generated by all* UBPs. *The set B_t, obtained from B by taking for q the polynomial symbol q_t of Theorem* 5.7, *is an equational base for the variety T generated by the triangle algebra* $\mathbf{W}_3$.

Since the identities B of Theorem 5.8 contain only four variable symbols, we have the following corollary.

5.9 Corollary. *Let* $\mathbf{A}$ *be a WAL. If each subalgebra of* $\mathbf{A}$ *which is generated by four or fewer elements is contained in the variety U (respectively T), then* $\mathbf{A}$ *is contained in U (respectively T).*

In [7] a weaker version of Corollary 5.9 was established, namely for the variety T only and with "five" instead of "four". Hence Corollary 5.9 solves the problem raised in [7] (following Corollary 1 of Theorem 2).

6. Problems

1. Find a simple property P of varieties such that: A variety V is a dual discriminator variety if and only if V has a) a majority polynomial, b) the congruence extension property, and c) property P.

2. Does Theorem 5.7 have an analog for $\mathbf{W}_4$, i.e.: is there a WAL polynomial symbol $q(x, y, z)$ which induces the dual discriminator on the UBP $\mathbf{W}_4$ and on no other WAL of more than three elements?

References

[1] K. A. BAKER, Primitive satisfaction and equational problems for lattices and other algebras, *Trans. Amer. Math. Soc.*, **190** (1974), 125—150.

[2] ———, Finite equational bases for finite algebras in a congruence distributive equational class, *Advances in Math.*, **24** (1977), 207—243.

[3] ——— and A. F. PIXLEY, Polynomial interpolation and the Chinese remainder theorem for algebraic systems, *Math. Z.*, **143** (1976), 165—174.

[4] S. BULMAN-FLEMING and H. WERNER, *Equational compactness in quasi-primal varieties*, T. H. Darmstadt, Preprint Nr. 215, August 1975.

[5] E. FRIED, Subdirectly irreducible weakly associative lattices with the congruence extension property, *Ann. Univ. Sci. Budapest, Sect. Math.*, **17** (1974), 59—68.

[6] ———, Weakly associative lattices with the congruence extension property, *Algebra Universalis*, **4**, (1974), 151—162.

[7] ——— and G. GRÄTZER, A non-associative extension of the class of distributive lattices, *Pac. J. Math.*, **49** (1973), 59—78.

[8] ——— and — and R. W. QUACKENBUSH, *Uniform congruence schemes*, Preprint, Winnipeg, 1976.

[9] G. GRÄTZER, *Universal Algebra*, Van Nostrand (Princeton, N. J.; 1968).

[10] B. JÓNSSON, Algebras whose congruence lattices are distributive, *Math. Scand.* **21** (1967), 110—121.

[11] R. MCKENZIE, On spectra and the negative solution of the decision problem for identities having a finite non-trivial model, *J. Symb. Logic*, **40** (1975), 186—195.

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

[13] A. F. PIXLEY, Functionally complete algebras generating distributive and permutable classes, *Math. Z.*, **114** (1970), 361—372.

[14] ———, The ternary discriminator function in universal algebra, *Math. Ann.*, **191** (1971), 167—180.

[15] R. W. QUACKENBUSH, Structure theory for equational classes generated by quasi-primal algebras, *Trans. Amer. Math. Soc.*, **187** (1974), 127—145.

[16] ———, Semi-simple equational classes with distributive congruence lattices, *Ann. Univ. Sci. Budapest, Sect. Math.*, **17** (1974), 15—19.

[17] H. SKALA, Trellis theory, *Algebra Universalis*, **1** (1971), 218—233.

[18] W. TAYLOR, The fine spectrum of a variety, *Algebra Universalis*, **5** (1975), 263—303.

[19] ———, Pure compactifications in quasi-primal varieties, *Canad. J. Math.*, **28** (1976), 50—62.

[20] H. WERNER, Eine Charakterisierung funktional vollständiger Algebren, *Arch. Math.*, *(Basel)*, **21** (1970), 381—385.

[21] ———, *Algebraic representation and model theoretic properties of algebras with the ternary discriminator*, T. H. Darmstadt, Preprint Nr. 237, 1975.

E. FRIED
EÖTVÖS LORÁND UNIVERSITY
MÚZEUM KÖRÚT 6—8.
1088 BUDAPEST, HUNGARY

A. F. PIXLEY
DEPARTMENT OF MATHEMATICS
HARVEY MUDD COLLEGE
CLAREMONT, CALIFORNIA 91711, U.S.A.

Acta Sci. Math., 41 (1979), 101—118

Quasisimilar operators with different spectra

DOMINGO A. HERRERO

1. Introduction. Let $\mathscr{L}(\mathfrak{X})$ be the Banach algebra of all (bounded linear) operators acting on the complex Banach space $\mathfrak{X}$. $T\in\mathscr{L}(\mathfrak{X})$ and $A\in\mathscr{L}(\mathfrak{Y})$ are called *quasisimilar* (q.s.) provided there exist quasi-invertible continuous linear maps $X\colon \mathfrak{Y}\to\mathfrak{X}$ and $Y\colon \mathfrak{X}\to\mathfrak{Y}$ such that $TX=XA$ and $YT=AY$ (X is quasi-invertible if $\operatorname{Ker} X=\{0\}$ and $\operatorname{Ran} X$ is dense in $\mathfrak{X}$; [48]).

As in [38], the four (weakly closed identity containing) subalgebras naturally associated with $T\in\mathscr{L}(\mathfrak{X})$ will be denoted by $\mathscr{A}(T)$, $\mathscr{A}^a(T)$, $\mathscr{A}'(T)$ and $\mathscr{A}''(T)$ (the algebra generated by the polynomials in T, the algebra generated by the rational functions of T with poles outside the spectrum $\sigma(T)$ of T, the commutant and the double commutant of T, resp.). Then $\mathscr{A}(T)\subset\mathscr{A}^a(T)\subset\mathscr{A}''(T)\subset\mathscr{A}'(T)$ and the corresponding invariant subspace lattices satisfy the reverse inclusions: $\operatorname{Lat} T\overset{\mathrm{def}}{=}\operatorname{Lat}\mathscr{A}(T)\supset\operatorname{Lat}\mathscr{A}^a(T)\supset\operatorname{Lat}\mathscr{A}''(T)\supset\operatorname{Lat}\mathscr{A}'(T)$. (These are called the lattices of invariant, analytically invariant, bi-invariant and hyperinvariant subspaces, resp. As usual, *subspace* will denote a *closed linear manifold* of $\mathfrak{X}$.)

Quasisimilarity was first studied by B. Sz.-Nagy and C. Foiaş ([48]; see also [17]) in connection with the invariant subspace problem in Hilbert spaces; namely, if A is q.s. to T, and T has a non-trivial hyperinvariant subspace, then so does A ([17; 39; 41]). A and T need not have the same spectrum ([48]); however, $\sigma(A)\cap\sigma(T)$ cannot be empty ([39]). Furthermore, every component of $\sigma(A)$ $(\sigma(T))$ intersects $\sigma(T)$ $(\sigma(A)$, resp.; [32]).

Several results scattered through the literature assert that, under suitable restrictions on T or A or both, $\sigma(A)$ actually contains $\sigma(T)$ or coincides with it ([9; 11; 39]) and there also exist examples of q.s. operators with different spectra ([39; 48]; see also Section 2, below).

This article is primarily concerned with the following questions:

(1) Under what conditions on T does "A is q.s. to T" imply "A is similar to T"?

Received May 6, 1977.

This work has been partially supported by CONICIT (Venezuela), Grant No. 51. 26. S1—0894.

(2) Under what conditions on T does "A is q.s. to T" imply $\sigma(A)=\sigma(T)$?

(3) When can we assert that $\sigma(A)$ is strictly larger (or strictly smaller) than $\sigma(T)$ for some A q.s. to T?

It is completely apparent that if T satisfies (1), then it also satisfies (2). On the other hand, two q.s. nilpotent operators with infinite dimensional range acting on a separable Hilbert space need not be similar ([3]; see also [18; 36]), so that a T satisfying (2) need not satisfy (1).

In [2], C. Apostol proved that A is q.s. to a normal operator if and only if Lat A contains a countable *basic system of subspaces* $\{\mathfrak{K}_n\}_1^m$ $(1\leqq m\leqq\infty)$ such that $A|\mathfrak{K}_n$ (A restricted to $\mathfrak{K}_n$) is similar to a normal operator for every n. (A countable family $\{\mathfrak{X}_n\}_1^m$ of subspaces of the Banach space $\mathfrak{X}$ is called basic if the subspaces $\mathfrak{X}_n$ and $\mathfrak{X}_n'=\bigvee_{k\neq n}\mathfrak{X}_k$ are complementary for every n and $\bigcap_1^m \mathfrak{X}_n'=\{0\}$; [2]). In Section 2 it will be shown that, under suitable (very general) conditions, an operator T having a *denumerable* basic system of invariant subspaces is q.s. to operators A and B such that either $\sigma(A)$ is strictly smaller than $\sigma(T)$, or $\sigma(B)$ is strictly larger than $\sigma(T)$, or both. To the best of the author's knowledge, this is the only known way to produce q.s. operators with different spectra. Recently, L. A. Fialkow showed that two q.s. non-invertible injective bilateral weighted shifts need not be similar; however, they necessarily have the same spectrum and this spectrum can be a disc of positive radius. Since Fialkow's operators do not admit any non-trivial pair of complementary invariant subspaces (see [22]), they add some extra support to the following

Conjecture 1. *Assume that* Lat T *does not contain any denumerable basic system of subspaces. Then* $\sigma(A)=\sigma(T)$ *for every* A *q.s. to* T.

The *strict multiplicity* $\bar\mu(\mathscr{A})$ of a subalgebra $\mathscr{A}$ of $\mathscr{L}(\mathfrak{X})$ is defined as the infimum of card(Γ), taken over all the subsets Γ of $\mathfrak{X}$ such that $\mathfrak{X}=$ $=\left\{\sum_1^n A_jx_j\colon A_j\in\mathscr{A},\ x_j\in\Gamma,\ n=1,2,\dots\right\}$. If Γ can be taken equal to the singleton $\{x_0\}$, then $\mathscr{A}$ is called a *strictly cyclic algebra* and x_0 is called a *strictly cyclic vector* for $\mathscr{A}$. According to [28, Theorem 8], if $\bar\mu[\mathscr{A}''(T)]<\infty$, then T satisfies (1). The main part of this paper is devoted to exploit this result and the constructions in [6] in order to show the existence and/or the density of operators satisfying certain properties related with quasisimilarity and an approximation problem, acting on a *complex separable infinite dimensional Hilbert space* $\mathfrak{K}$ (throughout this paper $\mathfrak{K}$ will always denote a space of this type).

Recall that $T\in\mathscr{L}(\mathfrak{K})$ is *biquasitriangular* (*BQT*) if ind$(\lambda-T)=0$, whenever $\lambda-T$ is a semi-Fredholm operator ([4]). C. Foiaş, C. Pearcy and D. Voiculescu [19] proved that for every $T\in\mathscr{L}(\mathfrak{K})$ and $\varepsilon>0$, there exists $T_\varepsilon\in\mathscr{L}(\mathfrak{K})$ such that

$\|T-T_\varepsilon\|<\varepsilon$, $T-T_\varepsilon\in\mathscr{K}$ (the ideal of compact operators), $T_\varepsilon=$norm-lim $U_nTU_n^*$ for a suitable sequence $\{U_n\}$ of unitary operators, Lat T_ε contains a denumerable family of pairwise orthogonal subspaces and T_ε is q.s. to a BQT operator ($T_\varepsilon\in(BQT)_{qs}$, in the notation of [19]). This strong result suggested to the authors of that article the following question

$$Is\ (BQT)_{qs}=\mathscr{L}(\mathfrak{H})?$$

The answer is no. Indeed, the following sets are (norm-)dense in $\mathscr{L}(\mathfrak{X})$:
(A)$=\{T$: T is q.s. to some $A\in(BQT)$ with $\sigma(A)=\sigma(T)\}$ [19];
(B)$=\{T$: T is q.s. to some $A\in(BQT)$ with $\sigma(A)\supset\sigma(T)$, $\sigma(A)\neq\sigma(T)\}$;
(C)$=\{T$: T is q.s. to some $A\in(BQT)$ with $\sigma(A)\subset\sigma(T)$, $\sigma(A)\neq\sigma(T)\}$;
(D)$=\{T$: T is similar to $A\oplus B$, $\bar\mu[\mathscr{A}''(A)]=\bar\mu[\mathscr{A}''(B^*)]=1$, $\sigma(A)\cap\sigma(B)=\emptyset$, λ_A-A and λ_B-B^* are semi-Fredholm operators of index $-\infty$ for suitably chosen points $\lambda_A,\lambda_B\in\mathbf{C}\}$.

Clearly, for every such T and every L q.s. to T, L is actually similar to T and it has the same spectrum as T. Therefore, (D)$\subset\{T$: T satisfies (1)$\}\setminus(BQT)_{qs}$.

$(E)_{mn}=\{T$: T, A and B are as in (D), except that $\bar\mu[\mathscr{A}''(A)]=m$ and $\bar\mu[\mathscr{A}''(B^*)]=n\}$ (for every m,n such that $m,n=1,2,\dots$ or c, the power of the continuum);

(F)$=\{T$: $\sigma(T)=\sigma(L)$ for every L q.s. to T, but $\mathscr{S}(T)\neq\mathscr{S}_{qs}(T)\}$, where $\mathscr{S}(T)$ ($\mathscr{S}_{qs}(T)$, resp.)$=\{A\in\mathscr{L}(\mathfrak{H})$: $A=WTW^{-1}$ for some invertible $W\in\mathscr{L}(\mathfrak{H})$ (A is q.s. to T, resp.)$\}$.

Recall that $\mathscr{A}\subset\mathscr{L}(\mathfrak{X})$ is a *reflexive algebra* if $\mathscr{A}=\operatorname{Alg}\operatorname{Lat}\mathscr{A}$, where $\operatorname{Alg}\Sigma=\{A\in\mathscr{L}(\mathfrak{X})$: Lat $A\supset\Sigma\}$ ($\Sigma=$any family of subspaces of $\mathfrak{X}$). $T\in\mathscr{L}(\mathfrak{X})$ is called reflexive if $\mathscr{A}(T)$ is. The following results are "in the air": The sets
(G)$=\{T$: T is reflexive$\}$;
(H)$=\{T$: $\mathscr{A}^a(T)$ is reflexive$\}$;
(I) $=\{T$: $\mathscr{A}''(T)$ is reflexive$\}$;
(J) $=\{T$: $\mathscr{A}'(T)$ is reflexive$\}$,
as well as their complements in $\mathscr{L}(\mathfrak{H})$, are dense in $\mathscr{L}(\mathfrak{H})$.

There are at least two different extensions of the notion of similarity related with approximation problems: A and T are *asymptotically similar* if their similarity orbits have the same closure (i.e., $\mathscr{S}(A)^-=\mathscr{S}(T)^-$; [7; 33]). They are *approximately similar* if $A=$norm-lim $W_nTW_n^{-1}$ for a sequence $\{W_n\}$ of invertible operators with $\sup\|W_n\|\,\|W_n^{-1}\|<\infty$ ([24]). Since asymptotic similarity (and, a fortiori, approximate similarity) preserves the spectrum and every part of it (see [33]), it will not be difficult to conclude from the results and examples of this article and the results of [7; 8; 33; 34; 35] that, in general, $\mathscr{S}(T)$ is a proper subset of $\mathscr{S}_{ap}(T)\cap\cap\mathscr{S}_{qs}(T)$ ($\mathscr{S}_{ap}(T)=\{A$: A is approximately similar to $T\}\subset\mathscr{S}(T)^-$) and the equality $\mathscr{S}(T)=\mathscr{S}(T)^-$, $T\in\mathscr{L}(\mathfrak{H})$, implies that T is similar to a normal operator with a finite spectrum and therefore $\mathscr{S}(T)=\mathscr{S}_{qs}(T)=\mathscr{S}_{ap}(T)=\mathscr{S}(T)^-$ (this is false for arbitrary Banach spaces; see [7; 35]); however, the equality $\mathscr{S}(T)=\mathscr{S}_{qs}(T)$ does not imply $\mathscr{S}(T)=\mathscr{S}(T)^-$ (even for Hilbert spaces; [28; 35]. Since approximate similarity preserves every Schatten p-ideal and asymptotic similarity does not preserve them, it is immediate that these two notions are different; see [33; 46] for details).

In [24], D. W. Hadwin defined the *approximate double commutant* of $T\in\mathscr{L}(\mathfrak{K})$ by $\operatorname{appr}(T)''=\{L\in\mathscr{L}(\mathfrak{K})\colon \|LA_n-A_nL\|\to 0\ (n\to\infty)$ whenever $\{A_n\}$ is a *bounded* sequence such that $\|TA_n-A_nT\|\to 0\ (n\to\infty)\}$. He proved that $\operatorname{appr}(T)''\subset\mathscr{A}''(T)\cap C^*(T)$ (where $C^*(T)$ denotes the C^*-algebra generated by T) and conjectured ([24, Conjecture 2.5]) that $\operatorname{appr}(T)''=\mathscr{A}''(T)$ if and only if T is algebraic. This conjecture is false. Indeed, $(K)=\{T\colon \operatorname{appr}(T)''=\mathscr{A}''(T)\}$, as well as its complement, is dense in $\mathscr{L}(\mathfrak{K})$.

The interested reader will have no trouble to prove the density in $\mathscr{L}(\mathfrak{K})$ of new different classes of operators somehow related with $(A)-(K)$.

The author is deeply indebted to R. G. Douglas, L. A. Fialkow, D. W. Hadwin and C. Pearcy for sending him their unpublished papers (the reader will find very useful information in Fialkow's papers [14; 15; 16], which have several points in common with the present article). The author also wishes to thank J. Barría, M. Cotlar, A. Etcheberry, B. Margolis and M. B. Pecuch for many helpful suggestions.

2. **Operators quasisimilar to orthogonal direct sums.** Given a family $\{\mathscr{X}_n\}$ of Banach spaces, let $\mathscr{Y}=\bigoplus_1^\infty \mathscr{X}_n$ denote the *hilbertian sum* of the $\mathscr{X}_n$'s (i.e., $\mathscr{Y}$ is the closure of the algebraic direct sum with respect to the norm $\|\{x_n\}\|=\left(\sum_1^\infty \|x_n\|_n^2\right)^{1/2}$.

Lemma 1. *Let $\mathscr{Y}$ be the hilbertian sum of the family $\{\mathscr{X}_n\}$ of Banach spaces and let $\{T_n\}$ $(T_n\in\mathscr{L}(\mathscr{X}_n))$ be a uniformly bounded family of operators. Let $T=\bigoplus_1^\infty T_n$ be the operator defined in the usual fashion on $\mathscr{Y}$ and assume that $\|(\lambda-T_n)^{-1}\|\leq\Phi[d_n(\lambda)-\varepsilon_n]$ for $d_n(\lambda)>\varepsilon_n$, where $\Phi(t)$ is a non-increasing function of t $(0<t<\infty)$ independent of n, $\{\varepsilon_n\}$ is a sequence of non-negative reals converging to 0 and $d_n(\lambda)=\operatorname{dist}[\lambda,\sigma(T_n)]$. Then $\sigma(T)=\sigma^-$, where $\sigma=\bigcup_1^\infty\sigma(T_n)$.*

Proof. Clearly, $\lambda-T$ is invertible in $\mathscr{L}(\mathscr{Y})$ if and only if $\lambda-T_n$ is invertible in $\mathscr{L}(\mathscr{X}_n)$ for every n and $\|(\lambda-T_n)^{-1}\|\leq C$ for a constant C depending only on λ.

From our hypothesis about the growth of $\|(\lambda-T_n)^{-1}\|$, we can easily see that, given $\varepsilon>0$, $\|(\lambda-T_n)^{-1}\|\leq\Phi[\operatorname{dist}\{\lambda,\sigma\}-\varepsilon]$ for every λ such that $\operatorname{dist}[\lambda,\sigma]>\varepsilon$ and for all $n>n_0(\varepsilon)$, whence the result follows. □

Example 1. Clearly, the function Φ must satisfy $\Phi(t)\geqq 1/t$, but the condition of Lemma 1 cannot be replaced by $\|(\lambda-T_n)^{-1}\|=O[1/d_n(t)]$. Indeed, if $H=\sum_1^{\infty}(1/n)P_n$, where $\{P_n\}$ is a sequence of pairwise orthogonal projections of infinite rank in the Hilbert space $\mathfrak{K}$ whose partial sums strongly converge to the identity I, then H is an hermitian operator unitarily equivalent to $H^{(\infty)}$ (the orthogonal direct sum of denumerable many copies of H), $\sigma(H)=\{0\}\bigcup_1^{\infty}\{1/n\}=E(H)$ ($E(\cdot)$ denotes the *essential spectrum*) and $\|W(\lambda-H)^{-1}W^{-1}\|\leqq\|W\|\,\|W^{-1}\|/d(\lambda)$ for every invertible $W\in\mathcal{L}(\mathfrak{K})$ and for every $\lambda\notin\sigma(H)$.

$\mathcal{S}(H)^-$ contains a BQT operator A such that $\sigma(A)=E(A)=\sigma(H)\cup K$, where K is an arbitrary compact connected set containing the origin ([34]), i.e., there exists a sequence $A_n=W_nHW_n^{-1}$ converging to A in the norm. It readily follows that $B=\bigoplus_1^{\infty}A_n$ is q.s. to H and it can be shown as in [13] that $\sigma(B)=E(B)=\sigma(A)$.

Example 2. If $\lim_{t\to 0}t^r\Phi(t)=\infty$ for every $r>0$, then there exists a *universal quasinilpotent operator* Q in $\mathcal{L}(\mathfrak{K})$ (i.e., $\mathcal{S}(Q)^-$ contains every nilpotent) such that $\|(\lambda-Q)^{-1}\|\leqq\max\{\Phi(|t|),(1+\varepsilon)|\lambda|^{-1}\}$ (for an arbitrary prescribed $\varepsilon>0$), $Q\cong Q^{(\infty)}$, Q is the orthogonal direct sum of denumerable many nilpotent operators acting on finite dimensional Hilbert spaces and is q.s. to a compact quasinilpotent operator (see [3; 8; 31]).

Proceeding as in Example 1 it is not difficult to construct a BQT operator $B=\bigoplus_1^{\infty}B_n$ q.s. to Q such that $\sigma(B)=E(B)$ is an arbitrary connected compact set containing the origin.

Theorem 1. *Assume that $T\in\mathcal{L}(\mathcal{X})$ admits a denumerable basic system of invariant subspaces $\{\mathcal{X}_n\}$ and let $T_n=T|\mathcal{X}_n$ for $n=1,2,\dots$; let $\mathcal{Y}$ be the hilbertian sum of the $\mathcal{X}_n$'s and let $B\in\mathcal{L}(\mathcal{Y})$ be defined by $B=\bigoplus_1^{\infty}T_n$. Then B is q.s. to T, $\sigma^-\subset\subset\sigma(B)\subset\sigma(T)$, every component of $\sigma(B)$ or $\sigma(T)$ intersects σ^-, $\sigma_p(T)=\sigma_p(B)==\bigcup_1^{\infty}\sigma_p(T_n)\subset\sigma$* ($\sigma_p(\cdot)$ denotes the point spectrum) *and $\sigma_p(T^*)=\sigma_p(B^*)==\bigcup_1^{\infty}\sigma_p(T_n^*)\subset\sigma$. Assume, moreover, that $\mathcal{X}_n$ is actually* (isomorphic with) *a Hilbert space for every n; then there exist operators L_n similar to T_n, $n=1,2,\dots$, such that $A=\bigoplus_1^{\infty}L_n$ is q.s. to T and $\sigma(A)=\sigma^-$.*

Note. In the case when $\mathcal{X}$ is a Hilbert space and T^* is defined via inner product, $\sigma(T^*)=\sigma(T)^*$, where $K^*=\{\bar{\lambda}\colon\lambda\in K\}$ is the symmetric of the set $K\subset\mathbf{C}$ with respect to the real axis. In this case the corresponding inclusion should be read $\sigma_p(T^*)\subset\sigma^*$. It is convenient to remark that $\mathscr{X}_n$ can be isomorphic to a Hilbert space for every n even if $\mathscr{X}$ is not; namely, it T is a diagonal operator with respect to a Schauder basis of $\mathscr{X}$ and the $\mathscr{X}_n$'s are the one-dimensional subspaces spanned by the elements of that basis.

Proof. That B and T (and A when $\mathscr{X}_n$ is a Hilbert space for every n) are actually q.s. follows by standard arguments (see, e.g., [2; 39]). It is clear that $\lambda\in\sigma(B)$ if and only if either $\lambda\in\sigma(T_n)$ for some n or the family $\{(\lambda-T_n)^{-1}\}$ is not uniformly bounded. Now, if $\|(\lambda-T_{n(k)})x_{n(k)}\|\to 0$ $(k\to\infty)$ for a suitable subsequence $\{n(k)\}_1^\infty$ of natural numbers and for suitably chosen unitary vectors $x_{n(k)}\in\mathscr{X}_{n(k)}$, then $\lim_{k\to\infty}\|(\lambda-T)x_{n(k)}\|=0$ and therefore $\lambda\in\sigma(T)$. Hence, $\sigma^-\subset\sigma(B)$ and $\sigma(B)\setminus\sigma\subset\sigma(T)$.

Now assume that $\mathscr{X}_n$ is a Hilbert space for every n. According to [30], for each $n=1, 2, \ldots$, there exists an operator $L_n\in\mathscr{L}(\mathscr{X}_n)$ similar to T_n such that $\|(\lambda-L_n)^{-1}\|\leqq 1/[d_n(\lambda)-1/n]$ for all λ such that $d_n(\lambda)>1/n$. Define $A\in\mathscr{L}(\mathscr{Y})$ q.s. to T and B by $A=\bigoplus_1^\infty L_n$. By Lemma 1, $\sigma(A)=\sigma^-$.

The remaining spectral inclusions follow from [13; 25; 32]. □

By using [12, Theorem 1.4], we obtain

Corollary 1. *Let T be as in Theorem* 1. *If $\sigma(T_n)\cap\sigma(T_m)=\emptyset$ for a pair of indices n, m then T has a nontrivial hyperinvariant subspace.*

Example 3. (The main example) Combining the arguments of the previous examples and the results of [2; 13; 29; 31; 34; 35; 39] it is possible to show that if T is a Hilbert space operator such that Lat T contains a denumerable basic system of subspaces $\{\mathfrak{R}_n\}$ such that $T_n=T|\mathfrak{R}_n$ either satisfies $A_n\oplus(\lambda+Q_n)\in\mathscr{S}(T_n)^-$ for some A_n and some nilpotent Q_n with $Q_n^n\neq 0$ or a universal quasinilpotent, or $\sigma(T_n)$ contains more than n points, then given an arbitrary compact set $K\subset\mathbf{C}$ such that every $\lambda\in K\setminus\sigma^-$ belongs to a component of K that intersects $\sigma_\infty=\bigcap_{m=1}^\infty[\bigcup_{n=m}^\infty\sigma(T_n)]^-$, then there exist A and B q.s. to T such that $\sigma(A)=K\cup\sigma^-$ and $\sigma(B)=K\cup\sigma(T)$. The details of the construction are left to the reader.

Remarks. a) Let $\mathscr{B}$ be a Banach algebra with identity. It is well known that the mapping $a\to\sigma(a)$ from $\mathscr{B}$ into the family of nonempty compact subsets of $\mathbf{C}$ is upper semi-continuous with respect to the Hausdorff metric, but it is not continuous, in general ([5; 25; 29; 40; 42; 44]). In certain special cases (e.g., $a=\lim a_n$ for a *commutative* sequence $\{a_n\}$, or $\sigma(a)=a$ totally disconnected set, etc.) it is possible to prove that $a\to\sigma(a)$ is actually a continuous mapping. By a minor modification of the proof of Lemma 1, we can obtain the following sufficient condition: "If $a=\lim a_n$ for a sequence $\{a_n\}$ satisfying the conditions of Lemma 1, then $\sigma(a)=\lim \sigma(a_n)$ (in the Hausdorff metric)". Examples 1 and 2 show that this condition cannot be too relaxed.

b) In Lemma 1 and Theorem 1: The results remain true if the hilbertian sum is replaced by $\|\{x_n\}\|=(\sum_{n=1}^{\infty}\|x_n\|^p)^{1/p}$ for some p, $1\leqq p<\infty$, etc.

c) If T is decomposable, then $\sigma(T)\subset\sigma(A)$ for every A q.s. to T ([10]). Furthermore, if $\mathfrak{M}\in\operatorname{Lat} T$ and A is q.s. to T, then $\sigma(T|\mathfrak{M})\cap\sigma(A)\neq\emptyset$ ([14]); thus, if for every $\lambda\in\sigma(T)$ and every $\varepsilon>0$ there exists an $\mathfrak{M}_{\lambda,\varepsilon}\in\operatorname{Lat} T$ such that $\sigma(T|\mathfrak{M}_{\lambda,\varepsilon})\subset\subset\Delta(\lambda,\varepsilon)=\{z\colon|\lambda-z|<\varepsilon\}$, then it readily follows that $\sigma(T)\subset\sigma(A)$ for every A q.s. to T. The hyponormal operators have the same property ([9]). The spectral inclusion could be strict, e.g., for the operators of Examples 1,2,3. However, by combining the results of [23] and the examples of [28] it is possible to show that for every infinite dimensional separable Banach space $\mathscr{X}$, there exist operators $A, Q\in\mathscr{L}(\mathscr{X})$ such that A and Q are nuclear operators, Q is quasinilpotent, $\sigma(A)$ is the union of $\{0\}$ and a sequence of points converging "very fast" to 0, $\mathscr{A}(A)$ and $\mathscr{A}(Q)$ are strictly cyclic algebras, $\mathscr{A}(A)$ ($\mathscr{A}(Q)$, resp.) is semisimple (a radical algebra, resp.; see definitions in [42]), every L q.s. to A (to Q, resp.) is actually similar to it and it has the same spectrum as A (as Q, resp.). Moreover, for every finite m, $\operatorname{Lat} A$ contains a basic system $\{\mathscr{X}_n\}_1^m$ of invariant subspaces, which are maximal spectral subspaces for the decomposable operator A ([10]); however, $\operatorname{Lat} A$ does not contain any denumerable basic system of subspaces (see [28]).

d) Every subspace in a basic system of invariant subspaces of $T\in\mathscr{L}(\mathscr{X})$ is actually *bi-invariant*. Many examples regarding operators T such that $\sigma(A)\neq\sigma(T)$ for some A q.s. to T deal with operators having a denumerable basic system of *hyperinvariant* subspaces. This is not always the case: indeed, a straightforward computation shows that for the q.s. operators A and T involved in the example of Hoover [39], every pair of non-trivial hyperinvariant subspaces of T (or A) has a non-trivial intersection.

e) There is little hope to improve [12, Theorem 1.4] or Corollary 1. Indeed, if U denotes the bilateral shift "multiplication by $e^{i\theta}$" in L^2 (Unit circle, Lebesgue measure) and $u(e^{i\theta})=\operatorname{sign}\theta$ $(-\pi<\theta<+\pi)$, then H^2 and uH^2 are invariant (but not bi-invariant!) subspaces of U such that $H^2\cap uH^2=\{0\}$, $L^2=H^2\vee uH^2$, but (by Apostol's result; [2]) U cannot be q.s. to $(U|H^2)\oplus(U|uH^2)$.

f) [15, Theorem 2.1] admits the following mild generalization, which follows from Theorem 1 and the same proof as in [15]: If $T\in\mathscr{L}(\mathfrak{H})$ and $\operatorname{Lat} T$ contains a basic system of subspaces $\{\mathfrak{H}_n\}$ such that $T_n=T|\mathfrak{H}_n$ is a spectral operator for every n, then T is q.s. to a spectral operator.

3. **The subsets** (A), (B) **and** (C) **are dense in** $\mathscr{L}(\mathfrak{H})$. From this point on, we shall only consider Hilbert space operators. The density of (A) follows from [19].

Lemma 2. *Given* $T\in\mathscr{L}(\mathfrak{K})$ *and* $\varepsilon>0$, *there exists* $T_\varepsilon\in\mathscr{L}(\mathfrak{K})$ *such that* $\|T-T_\varepsilon\|<\varepsilon$ *and* T_ε *is similar to* $(\lambda+Q)\oplus C$, *where* $\sigma(\lambda+Q)$ *lies in the unbounded component of* $\mathbf{C}\setminus\sigma(C)$, $E(T)=E(C)$ *and* Q *is an arbitrary operator such that* $\sigma(Q)\subset\Delta(0,\varepsilon/5)$.

Proof. Proceeding as in [45], we can find an $L\in\mathscr{L}(\mathfrak{K})$ such that $\|T-L\|<3\varepsilon/4$ and

$$L=\begin{pmatrix}\lambda I & B\\ 0 & C\end{pmatrix}$$

with respect to an orthogonal direct sum decomposition $\mathfrak{K}=\mathfrak{K}_\lambda\oplus\mathfrak{K}_\lambda^\perp$ of $\mathfrak{K}$ into two infinite dimensional subspaces, where $\operatorname{dist}[\lambda,\sigma(T)]=\operatorname{dist}[\lambda,\sigma(C)]=\varepsilon/2$ and λ lies in the unbounded component of $\mathbf{C}\setminus\sigma(C)$.

By the corollary of Rota [43] (see also [30]), we can find a Q' similar to Q such that $\|Q'\|<\varepsilon/4$. Then

$$T_\varepsilon=\begin{pmatrix}\lambda+Q' & B\\ 0 & C\end{pmatrix}$$

is similar to $(\lambda+Q)\oplus C$, by Rosenblum's corollary ([41, Corollary 0.15]) and $\|T-T_\varepsilon\|\leqq\|T-L\|+\|Q'\|<\varepsilon$. $\square$

As in Hoover [39], we can find two q.s. operators Q_1 and Q_2 such that Q_1 is quasinilpotent and $\sigma(Q_2)=\Delta(0,\varepsilon/6)^-$, and $\|Q_j\|<\varepsilon/4$, $j=1,2$. By using the results of [19], C can be replaced by an operator $C_\varepsilon\in(BQT)_{qs}$ with the same spectrum as C such that $\|C-C_\varepsilon\|<\varepsilon$. Then the operator $T_{\varepsilon j}$ given by

$$T_{\varepsilon j}=\begin{pmatrix}\lambda+Q_j & B\\ 0 & C_\varepsilon\end{pmatrix}$$

satisfies $\|T-T_{\varepsilon j}\|<\varepsilon$, $j=1,2$, and it is immediate from our construction that $T_{\varepsilon 1}$ and $T_{\varepsilon 2}$ are q.s. operators of the class $(BQT)_{qs}$.

Since $\sigma(T_{\varepsilon 1})$ is a proper subset of $\sigma(T_{\varepsilon 2})$, it follows at once that (B) and (C) are dense in $\mathscr{L}(\mathfrak{K})$.

Given $T\in\mathscr{L}(\mathfrak{K})$, let T_ε be constructed as in Lemma 2 with $Q=V=$the Volterra operator, and let W be an invertible operator such that $T_\varepsilon=W[(\lambda+V)\oplus\oplus C]W^{-1}$. Then $\mathscr{A}(T_\varepsilon)=W[\mathscr{A}(V)\oplus\mathscr{A}(C)]W^{-1}$, $\operatorname{Alg\,Lat}\mathscr{A}(T_\varepsilon)=W[\operatorname{Alg\,Lat}\mathscr{A}(V)\oplus\oplus\operatorname{Alg\,Lat}\mathscr{A}(C)]W^{-1}$ and similarly for the other three algebras naturally associated with T_ε (all these facts can be easily checked by using the results of [41]); moreover, $\operatorname{appr}(T_\varepsilon)''=W[\operatorname{appr}(V)''\oplus\operatorname{appr}(C)'']W^{-1}$ ([24]).

Since $\mathscr{A}(V)=\mathscr{A}^a(V)=\mathscr{A}''(V)=\mathscr{A}'(V)\neq\operatorname{Alg\,Lat}V$ ([41]) and $\mathscr{A}''(V)\neq\neq\operatorname{appr}(V)''$ (see [41] or [20, Proposition 6]), it follows that none of the four algebras associated with T_ε is reflexive and $\operatorname{appr}(T_\varepsilon)''\neq\mathscr{A}''(T_\varepsilon)$. Thus, we have

Corollary 2. *The complements of the sets* (G), (H), (I), (J) *and* (K) *are dense in* $\mathscr{L}(\mathfrak{K})$.

4. (D) **is dense in** $\mathcal{L}(\mathfrak{K})$. The main ingredient is the construction of a large family of operators with a strictly cyclic double commutant.

Let Ω be a nonempty bounded connected open subset of the plane such that $\partial\Omega$ (the boundary of Ω) consists of finitely many pairwise disjoint regular analytic Jordan curves (We shall say that "Ω is an open set with analytic boundary" or "$\partial\Omega$ is analytic" as a shorthand notation) and let $\Lambda=\{\lambda_1, \lambda_2, \dots, \lambda_m\}$ be a finite subset of $\mathbf{C}\setminus\Omega^-$ having exactly one point in every component of this last set. Let $\varepsilon>0$ be small enough so that $\Lambda\cap(\Omega^-+\varepsilon t)=\emptyset$ (where $K+\lambda=\{z+\lambda: z\in K\}$, $K\subset\mathbf{C}$) for every $t\in[0, 1]$ and define $\Gamma=\{(z, t)\in\mathbf{C}\times(0, 1): z-\varepsilon t\in\partial\Omega\}$. It is apparent that there exists an analytic diffeomorphism φ: $\{(z, t)\in\mathbf{C}\times(-1, 2): z-\varepsilon t\in\partial\Omega\}\to\Omega_1$, where Ω_1 is the union of m open annulus with pairwise disjoint closures in the plane; then $\varphi|\Gamma: \Gamma\to\Omega_0\stackrel{\text{def}}{=}\varphi(\Gamma)$ is an analytic diffeomorphism such that if dm_0 denotes the planar Lebesgue measure on Ω_0 and dm_Γ is the area measure on Γ induced by Lebesgue measure in $\mathbf{R}^3$, then there exists $\delta, 0<\delta<1$, such that $\delta m_0[\varphi(B)]\leqq m_\Gamma(B)\leqq(1/\delta)m_0[\varphi(B)]$ for every Borel set $B\subset\Gamma$; moreover, φ can be chosen to be a conformal mapping.

The Sobolev space $\mathbf{W}^{2,2}(\Omega_0)$ of all distributions u on Ω_0 whose distributional partial derivatives of order $m, 0\leqq m\leqq 2$, belong to $L^2(\Omega_0, dm_0)$ can be identified with a Banach algebra (under an equivalent norm) of continuous functions on Ω_0^- (see [1, Chapter V]) and it is clear that φ induces an isomorphism between this space and $\mathbf{W}_\infty=\mathbf{W}^{2,2}(\Gamma)$ (defined in the obvious way on the analytic differentiable manifold Γ). Furthermore, by using this isomorphism, it is easily seen that there exists a constant C such that, given $f, g\in\mathbf{W}_\infty$, the pointwise product $(fg)(z, t)=f(z, t)\cdot g(z, t)$ defines an element of $\mathbf{W}_\infty$ and $\|fg\|\leqq C\|f\|\,\|g\|$ (where $\|\cdot\|$ denotes the norm in $\mathbf{W}_\infty$); hence, $\mathbf{W}_\infty$ is a semisimple Banach algebra with identity $e(z, t)\equiv 1$, under an equivalent norm. The Gelfand spectrum $\mathcal{M}(\mathbf{W}_\infty)$ can be naturally identified (via point evaluations; see [1; 21]) with Γ^-.

Let $\mathbf{A}_\infty=\mathbf{A}^{2,2}(\Gamma)$ be the closure in $\mathbf{W}_\infty$ of the functions of the form

$$f(z, t) = \sum_{k=0}^{n} t^k f_k(z), \quad n = 1, 2, \dots, \tag{\#}$$

where the f_k's are rational functions with poles in a subset of Λ (these are the "analytic elements" of $\mathbf{W}_\infty$). By using the maximum modulus principle and Runge's theorem (see, e.g., [21]), it is easily seen that every $f\in\mathbf{A}_\infty$ can be continuously extended to a unique function defined on $\Xi=\{(z, t)\in\mathbf{C}\times[0, 1]: z-\varepsilon t\in\Omega^-\}$, analytic with respect to $z\in\Omega+\varepsilon t$ for every $t\in[0, 1]$ and, on the other hand, every function $f(z, t)$ satisfying these conditions such that $f|\Gamma\in\mathbf{W}_\infty$, is an element of $\mathbf{A}_\infty$.

$\mathbf{A}_\infty$ is a subspace of $\mathbf{W}_\infty$ invariant under $T=M_z\in\mathcal{L}(\mathbf{W}_\infty)$ defined by $Tf(z, t)=zf(z, t)$ (here and in what follows, M_g denotes the operator "multiplication by g"). Moreover, $\mathbf{A}_\infty$ is a Banach algebra with identity e, and $\mathcal{M}(\mathbf{A}_\infty)$ can be naturally identified with Ξ.

Let $L=T|\mathbf{A}_\infty$ and let Pr: $\mathbf{C}\times\mathbf{R}\to\mathbf{C}$ be the projection onto $\mathbf{C}$ $(Pr(z,t)=z)$; then

Lemma 3. *With the above notation:*

(i) $\sigma(T)=E_l(T)=E_r(T)=E_l(L)=Pr(\Gamma^-)$, *where* $E_l(\cdot)$ $(E_r(\cdot)$, resp.) *denotes the left (right,* resp.) *essential spectrum.*

(ii) $\sigma(L)=E_r(L)=Pr(\Xi)$.

(iii) $\operatorname{Ker}(\lambda-L)=\{0\}$ *and* $\dim\operatorname{Ker}(\lambda-L)^*=\infty$ *(so that* $\operatorname{ind}(\lambda-L)=-\infty$*) for every* $\lambda\in\sigma(L)\setminus E_l(L)$.

(iv) $\mathcal{A}''(L)=\mathcal{A}'(L)=\{M_g\colon g\in\mathbf{A}_\infty\}$, *i.e., the double commutant of* L *is the maximal abelian subalgebra of* $\mathcal{L}(\mathbf{A}_\infty)$ *consisting of all multiplications by elements of* $\mathbf{A}_\infty$ *and this is a strictly cyclic algebra with strictly cyclic vector* e.

Proof. (i), (ii) and (iii) follow from the previous observations. The proof is left to the reader.

(iv) By using several well known results about strictly cyclic algebras ([26; 27; 37]), it suffices to show that, if $A\in\mathcal{A}'(L)$, then $A=M_g$, where $g=Ae$. The remaining of the proof is an "ad hoc" modification of an argument used in [28].

Given $\eta,\tau\in[0,1]$, $\eta\neq\tau$, choose δ, $0<\delta<|\eta-\tau|/8$, and let $h_\eta(z,t)\in\mathbf{A}_\infty$ be the restriction to Γ^- of the function defined by

$$h_\eta(z,t)=\begin{cases}0 & \text{outside } (\eta-3\delta,\eta+3\delta),\\ (t-\eta+3\delta)^2/2\delta^2 & \text{in } [\eta-3\delta,\eta-2\delta],\\ 1-(t-\eta+\delta)^2/2\delta^2 & \text{in } [\eta-2\delta,\eta-\delta],\\ 1 & \text{in } [\eta-\delta,\eta+\delta],\\ 1-(t-\eta-\delta)^2/2\delta^2 & \text{in } [\eta+\delta,\eta+2\delta],\\ (t-\eta-3\delta)^2/2\delta^2 & \text{in } [\eta+2\delta,\eta+3\delta].\end{cases}$$

Define $h_\tau(z,t)=h_\eta(z,t-\eta+\tau)$ and let $\psi\colon[0,1]\to[\tau-4\delta,\tau+4\delta]\cap[0,1]$ be an arbitrary C^∞ bijection such that $\psi(t)=t$ in $[\tau-3\delta,\tau+3\delta]\cap[0,1]$ and $\min\{\psi'(t)\colon t\in[0,1]\}>0$.

Define $\mathbf{W}_{\tau,\delta}=\mathbf{W}^{2,2}(\{(z,t)\in\Gamma\colon|t-\tau|<4\delta\})$ exactly in the same way as $\mathbf{W}_\infty$ and let $T_{\tau,\delta}$ be the "multiplication by z" in this new space. Let $\mathbf{A}_{\tau,\delta}$ be the subalgebra of the "analytic elements" of $\mathbf{W}_{\tau,\delta}$ (defined in the obvious way) and $L_{\tau,\delta}=T_{\tau,\delta}|\mathbf{A}_{\tau,\delta}$.

The properties of ψ make it clear that $S\colon\mathbf{A}_{\tau,\delta}\to\mathbf{A}_\infty$ defined by $Sf(z,t)=f(z+\varepsilon[\psi(t)-t],\psi(t))$ is a (not necessarily isometric) isomorphism of Hilbert spaces.

Our choice of δ makes it possible to find a disc $\Delta=\Delta(\lambda(\eta,\tau),\varepsilon\delta/2)$ contained in $\Omega+\varepsilon\eta$ such that $\Delta^-\cap\sigma(L_{\tau,\delta})=\emptyset$.

Finally, let $R\colon \mathrm{A}_\infty \to H^2(\Delta)$ be the "restriction in the η-fiber" mapping defined by $Rf(z)=f(z,\eta)|_{z\in\Delta}$ and let L_Δ be the "multiplication by z" in $H^2(\Delta)$.

Clearly, if M_η and M_τ are the multiplications by h_η and h_τ, respectively, then $M_\eta A M_\tau \in \mathscr{A}'(L)$, so that $L(M_\eta A M_\tau)-(M_\eta A M_\tau)L=0$ whence we obtain $0=RLM_\eta AM_\tau S-RM_\eta AM_\tau LS=L_\Delta(RM_\eta AM_\tau S)-(RM_\eta AM_\tau S)L_{\tau,\delta}$ (Beware! $LS\neq \neq SL_{\tau,\delta}$; however, it is not difficult to check that $\psi(t)=t$ in $[\tau-3\delta,\tau+3\delta]\cap[0,1]$ yields $M_\tau LS=M_\tau SL_{\tau,\delta}$).

Since $\sigma(L_\Delta)=\Delta^-$ is disjoint from $\sigma(L_{\tau,\delta})$ by construction, it follows from Rosenblum's corollary ([41, Corollary 0.13]) that $RM_\eta AM_\tau S=0$; moreover, since S is an isomorphism, $RM_\eta AM_\tau=0$. Since Ω is connected, the vanishing of $f(z,\eta)$ on Δ implies that $f(z,\eta)\equiv 0$, whence we conclude that the value of $Af(z,\tau)$ only depends on the values of $f(z,t)$ for t in a neighborhood of τ.

We shall need a little more: A straightforward computation shows that $\|(t-\tau)^k h_\tau(z,t)\|\to 0$ as $\delta\to 0$, uniformly with respect to k $(k=1,2,\ldots)$. Let f be any function of the form (#) and let $F(z,t)=f(z,\tau)$; then

$$f(z,t)=F(z,t)+\sum_{k=1}^{n}(t-\tau)^k f_k(z),$$

where the f_k's are rational functions of z with poles in a subset of Λ. Since A commutes with M_z, it is clear that $AM_F=M_FA$ and $AM_{f_k}=M_{f_k}A$ for $k=1,2,\ldots,n$, and therefore $AF(z,t)=AM_Fe(z,t)=[M_F(Ae)](z,t)=F(z,t)g(z,t)=g(z,t)f(z,\tau)$, which is equal to $g(z,\tau)f(z,\tau)$ for $t=\tau$. Hence,

$$Af(z,\tau)=AF(z,\tau)+\sum_{k=1}^{n}f_k(z)\lim_{\delta\to 0}A[(t-\tau)^k h_\tau](z,\tau)=$$

$$=g(z,\tau)f(z,\tau),\quad \text{for every}\quad \tau\in[0,1].$$

Therefore, $Af(z,t)=g(z,t)f(z,t)$ on Γ^- for every f of the form (#). By continuity, we conclude that $A=M_g$. □

By a formal repetition of the proof of [28, Theorem 8] and the above result, we can easily obtain

Lemma 4. *Let Ω, ε and A be as in Lemma 3 and let n be a positive integer. Define $\mathrm{W}_n=\bigoplus_{k=1}^{n}\mathrm{W}^{2,1}(\partial\Omega+k\varepsilon/n, dm_k)$, where dm_k is the "arc length measure" on $\partial\Omega+k\varepsilon/n$ and $\mathrm{W}^{2,1}(\partial\Omega+k\varepsilon/n, dm_k)$ is the Sobolev space of all distributions u on $\partial\Omega+k\varepsilon/n$ with distributional derivative (with respect to "arc length") in $L^2(\partial\Omega+k\varepsilon/n, dm_k)$, with the norm*

$$\|f\|=\Big\{\int_{\partial\Omega+k\varepsilon/n}[|f(z)|^2+|df/dm_k(z)|^2]\,dm_k\Big\}^{1/2},$$

and let $\mathbf{A}_n$ *be the subspace of "analytic elements" of* $\mathbf{W}_n$ (i.e., $\mathbf{A}_n = \mathbf{W}_n$*-closure* $\{(f_1, f_2, \dots, f_n)\colon f_k$ is rational with poles in a subset of $\Lambda\}$).

Then $\mathbf{W}_n$ *and* $\mathbf{A}_n$ *are semisimple Banach algebras of continuous functions with identity (under an equivalent norm),* $\mathcal{M}(\mathbf{W}_n)$ ($\mathcal{M}(\mathbf{A}_n)$, resp.) *can be naturally identified with* $\bigcup_{k=1}^{n} (\partial\Omega + k\varepsilon/n) \times \{k/n\}$ ($\bigcup_{k=1}^{n} (\Omega^- + k\varepsilon/n) \times \{k/n\}$, resp.)$\subset \mathbf{C} \times [0, 1]$.

Furthermore, if $T_n = M_z$ *in* $\mathbf{W}_n$, *then* $\mathbf{A}_n$ *is invariant under* T_n, *and* T_n *and its restriction* $L_n = T_n | \mathbf{A}_n$ *satisfy*

(i) $\sigma(T_n) = E_l(T_n) = E_r(T_n) = E_l(L_n) = E_r(L_n) = Pr[\mathcal{M}(W_n)]$.

(ii) $\sigma(L_n) = Pr[\mathcal{M}(\mathbf{A}_n)]$.

(iii) $\operatorname{Ker}(\lambda - L_n) = \{0\}$ *and* $\dim \operatorname{Ker}(\lambda - L_n)^* = n$ (*so that* $\operatorname{ind}(\lambda - L_n) = -n$)

for every $\lambda \in \bigcap_{k=1}^{n} (\Omega + k\varepsilon/n) \subset \sigma(L_n) \setminus E(L_n)$.

(iv) $\mathcal{A}'(L_n) = \mathcal{A}''(L_n) = \{M_g\colon g \in \mathbf{A}_n\}$ *is a maximal abelian strictly cyclic subalgebra of* $\mathcal{L}(\mathbf{A}_n)$.

The proof is left to the reader.

Given Ω with analytic boundary, $\varepsilon > 0$ and Λ as indicated, and an index n, $-\infty \leqq n < 0$, we shall denote by $T(\Omega, \varepsilon, n)$ and $L(\Omega, \varepsilon, n)$ the operators defined by Lemma 3 (for $n = -\infty$) or by Lemma 4 (for $-\infty < n < 0$). If $0 < n \leqq +\infty$, we shall use the adjoint operators $T(\Omega^*, \varepsilon, -n)^*$ and $L(\Omega^*, \varepsilon, -n)^*$.

Now we are in a position to prove the main result of this paper.

Theorem 3. *The subset (D) of those operators T similar to* $A \oplus B$, *where*

(i) $\sigma(A) \cap \sigma(B) = \emptyset$;

(ii) $\mathcal{A}''(A)$ *and* $\mathcal{A}''(B^*)$ *are strictly cyclic algebras;*

(iii) $\lambda_A - A$ *and* $\lambda_B - B^*$ *are semi-Fredholm operators of index* $-\infty$, *for suitably chosen* λ_A *and* λ_B;

(iv) $\mathcal{S}(A \oplus B) = \mathcal{S}_{qs}(A \oplus B)$ *and this set does not intersect* $(BQT)_{qs}$; *is dense in* $\mathcal{L}(\mathfrak{R})$.

Proof. The result follows by modifying the proofs in [6].

By [6, Proposition 1.4] (Indeed, by a minor modification of it), given $T \in \mathcal{L}(\mathfrak{R})$ and $\varepsilon > 0$, there exists an operator T_1 such that $\|T - T_1\| < \varepsilon$ and

$$T_1 = \begin{pmatrix} N_1 & 0 & * & * & * \\ 0 & N_2 & * & * & * \\ 0 & 0 & S_1 & * & * \\ 0 & 0 & 0 & N_3 & 0 \\ 0 & 0 & 0 & 0 & N_4 \end{pmatrix}$$

(with respect to a suitable orthogonal direct sum decomposition of $\mathfrak{H}$ into five subspaces), where

a) N_j is normal and $\sigma(N_j)=E(N_j)$ for $j=1,2,3,4$;

b) $\sigma(N_2)\cup\sigma(N_3)$ is the closure of a nonempty open subset Ω_0 with analytic boundary;

c) $\sigma(N_1)\cap\sigma(N_4)=\emptyset$, $\sigma(N_1)$ and $\sigma(N_4)$ are disjoint unions of pairwise disjoint regular analytic Jordan curves, $\sigma(N_1)\subset\partial\sigma(N_2)\cap\partial\Omega_0$ and $\sigma(N_4)\subset\partial\sigma(N_3)\cap\partial\Omega_0$; $\sigma(N_1)$ ($\sigma(N_4)$, resp.) is contained in the open set $\{\lambda: (\lambda-T)$ is semi-Fredholm of negative (positive, resp.) index$\}$;

d) S_1 is similar to a direct sum $F\oplus S_2$, where F is a normal operator with simple eigenvalues (i.e., cyclic) acting on a finite dimensional subspace, such that $\sigma(F)\cap[\sigma(N_2)\cup\sigma(N_3)\cup\sigma(S_2)]=\emptyset$, and $\partial\sigma(S_2)\subset\Omega_0^-$;

e) The Weyl spectrum $w(T)$ of T satisfies the inclusions $w(T)\overset{\text{def}}{=}$ $=\sigma(T)\setminus\{\lambda: (\lambda-T)$ is a Fredholm operator of index $0\}\subset\sigma(N_2)\cup\sigma(N_3)\cup\sigma(S_2)\subset$ $\subset w(T)_\varepsilon$, where $K_\varepsilon=\{\lambda: \operatorname{dist}(\lambda, K)\leqq\varepsilon\}$ $(K\subset\mathbf{C})$;

f) $\operatorname{min.ind}(\lambda-S_2)\overset{\text{def}}{=}\min\{\dim\operatorname{Ker}(\lambda-S_2), \dim\operatorname{Ker}(\lambda-S_2)^*\}=0$ for every λ such that $(\lambda-S_2)$ is semi-Fredholm.

Clearly, $\sigma(T_1)$ is the disjoint union of its clopen subsets $\sigma(F)$ and $\sigma(T_1)\setminus\sigma(F)$ so that, by ROSENBLUM [41, Corollary 0.15), T_1 is similar to $F\oplus T_2$, where

$$T_2=\begin{pmatrix} N_1 & * & * & * & * \\ 0 & N_2 & * & * & * \\ 0 & 0 & S_2 & * & * \\ 0 & 0 & 0 & N_3 & * \\ 0 & 0 & 0 & 0 & N_4 \end{pmatrix}.$$

According to c), $N_1=\bigoplus_{k=1}^{m} N_{1k}$ $\left(N_4=\bigoplus_{j=1}^{p} N_{4j}\right)$, where $\sigma(N_{1k})=E(N_{1k})$ $(\sigma(N_{4j})=$ $=E(N_{4j})$, resp.) is the boundary of a unique component Ω_k (Ω_j, resp.) of the semi-Fredholm domain of T_2, where $\operatorname{ind}(\lambda-T_2)=n_k<0$ $(=n_j>0$, resp.) for all $\lambda\in\Omega_k$, $k=1,2,\dots,m$ ($\lambda\in\Omega_j$, $j=1,2,\dots,p$, resp.).

Let $\Lambda M=\{\lambda_1,\lambda_2,\dots,\lambda_q,\mu_1,\mu_2,\dots,\mu_q\}$ be a finite set having exactly two points, λ_h and μ_h, in each of the q components of Ω_0. Replacing, if necessary, ε by an ε', $0<\varepsilon'<\varepsilon$, we can assume that the three sets $\sigma(F)$, $(\Lambda M)_\varepsilon$ and $(\Omega_0)_\varepsilon$ are pairwise disjoint.

Let $T(\Omega_k,\varepsilon,n_k)$ $(k=1,2,\dots,m)$ and $T(\Omega_j^*,\varepsilon,-n_j)^*$ $(j=1,2,\dots,p)$ be the operators constructed as above indicated. Since $d_H\{\sigma(N_{1k}),\sigma[T(\Omega_k,\varepsilon,n_k)]\}\leqq\varepsilon$ (d_H denotes the Hausdorff distance), it follows from [35] that there exists T_k' similar to $T(\Omega_k,\varepsilon,n_k)$ such that $\|T_k'-N_{1k}\|<2\varepsilon$, $k=1,2,\dots,m$. Analogously, there exists T_j'' similar to $T(\Omega_j^*,\varepsilon,-n_j)^*$ such that $\|T_j''-N_{4j}\|<2\varepsilon$, $j=1,2,\dots,p$; thus, if

$M_1=\bigoplus_{k=1}^{m} T_k'$ and $M_4=\bigoplus_{j=1}^{p} T_j''$, and T_3 is the operator obtained from T_2 by replacing N_1 by M_1 and N_4 by M_4, then $\|T_2-T_3\|<2\varepsilon$.

It is clear that M_1 has an invariant subspace $\mathfrak{M}_1$ such that $L_1=M_1|\mathfrak{M}_1$ is similar to $\bigoplus_1^m L(\Omega_k, \varepsilon, n_k)$ and that M_4 has an invariant subspace $\mathfrak{M}_4$ such that the *compression* L_4 of M_4 to $\mathfrak{M}_4^\perp$ is similar to $\bigoplus_1^p L(\Omega_j^*, \varepsilon, -n_j)^*$. Since the spectra of the components of these direct sums are pairwise disjoint, it follows as in the proof of Corollary 2 that $\mathscr{A}''(L_1)$ and $\mathscr{A}''(L_4^*)$ are strictly cyclic operator algebras.

Proceeding exactly as in the proof of [6, Proposition 2.1], we find out that

$$T_3=\begin{pmatrix} L_1 & * & * \\ 0 & S_3 & * \\ 0 & 0 & L_4 \end{pmatrix}$$

(with respect to a suitable orthogonal direct sum decomposition), where $S_3\in BQT$ and $\sigma(S_3)=E(S_3)=\Omega_0^-$.

Let $C_1=\bigoplus_{h=1}^{q} (\lambda_h+L[\Delta(\lambda_h, \varepsilon/2), \varepsilon/2, -\infty])$ and $C_4=\bigoplus_{h=1}^{q} (\mu_h+L[\Delta(\mu_h, \varepsilon/2), \varepsilon/2, -\infty]^*)$. By using the results of [34; 35] and Rosenblum [41, Corollary 0.15], we can find an operator

$$S_4=\begin{pmatrix} C_1' & * \\ 0 & C_4' \end{pmatrix},$$

with C_i' similar to C_i, $i=1, 4$, such that $\|S_3-S_4\|<\varepsilon$, so that if T_4 is the operator obtained from T_3 by replacing S_3 by S_4, then a formal repetition of previous arguments shows that $\|T_3-T_4\|<\varepsilon$ and T_4 is similar to $L_1\oplus C_1'\oplus C_4'\oplus L_4$ which, in turn, is similar to $A_0\oplus B$, where $A_0=\{\bigoplus_1^m L(\Omega_k, \varepsilon, n_k)\}\oplus C_1$ and

$$B=C_4\oplus\{\bigoplus_1^p L(\Omega_j^*, \varepsilon, -n_j)^*\}.$$

Thus, if $A=F\oplus A_0$, it readily follows that there exists an operator T_5 similar to $A\oplus B$ such that $\|T-T_5\|<4\varepsilon$.

Since A and B clearly satisfy (i)—(iv), we are done. □

Corollary 3. *$(E)_{mn}$, (F), (G), (H), (I) and (J) are dense in $\mathscr{L}(\mathfrak{K})$.*

Proof. The proof will be just sketched. Repeat exactly the same proof as above replacing ΛM by $\Lambda MN\Pi=\{\lambda_1, \ldots; \lambda_q, \mu_1, \ldots, \mu_q, \nu_1 \ldots, \nu_q; \pi_1, \ldots, \pi_4\}$ with the same characteristics as ΛM and four points, $\lambda_h, \mu_h, \nu_h, \pi_h$, in each component of Ω_0.

$(E)_{mn}$: Replace A and B by $A\oplus\{\bigoplus_{h=1}^{q} \nu_h I_m\}$ and $B\oplus\{\bigoplus_{h=1}^{q} \pi_h I_n\}$, resp., where I_m (I_n) is the identity on a Hilbert space of algebraic dimension m (n, resp.), and use the results of [27].

(F) Replace A by $A\oplus\{\bigoplus_{h=1}^{q} (\nu_h+Q)\}$, where Q is any nilpotent of infinite rank. The result follows as in Theorem 3 by using the results of [3].

(I) and (J): These two cases follow at once from Theorem 3, the fact that $\mathbf{A}_\infty$ and $\mathbf{A}_n$ are *semisimple* Banach algebras and [47]; it is easily seen that $\mathscr{A}''(A)=\mathscr{A}'(A)$ and $\mathscr{A}''(B)=\mathscr{A}'(B)$ are reflexive.

(G) and (H): These two cases follow at once from the above observations abou $\mathscr{A}''(A)$ and $\mathscr{A}''(B)$ and the results of [37; 38]. □

Remark. An alternative proof for the cases $(G)-(J)$ can be obtained by using the Apostol—Morrel dense class $C_0(\mathfrak{R})$ (see definition and properties in [6]) and the results of [41].

5. (K) **is dense in** $\mathscr{L}(\mathfrak{R})$. The proof is a "trivialization" of that of the case (D).

Lemma 5. *Let Ω be an open set with analytic boundary, let $\Gamma_0=\partial\Omega\times(0,1)$ and $\Xi_0=\Omega^-\times[0,1]$. $\mathbf{W}_{0\infty}=\mathbf{W}^{2,2}(\Gamma_0)$ (defined as in Section 4) has the same properties as $\mathbf{W}_\infty$ and the subalgebra $\mathbf{A}_{0\infty}$ of "analytic elements" of $\mathbf{W}_{0\infty}$ ($\mathbf{A}_{0\infty}=\{f\in\mathbf{W}_{0\infty};\ f(z,t)$ is analytic with respect to $z\in\Omega$ for every $t\in[0,1]\}$) has the same properties as $\mathbf{A}_\infty$.*

If $T_0=M_z$ in $\mathbf{W}_{0\infty}$ and $L_0=T_0|\mathbf{A}_{0\infty}$, then:

(i) $\sigma(T_0)=E_l(T_0)=E_r(T_0)=E_l(L_0)=\partial\Omega$.

(ii) $\sigma(L_0)=E_r(L_0)=\Omega^-$.

(iii) $\operatorname{Ker}(\lambda-L_0)=\{0\}$ *and* $\dim\operatorname{Ker}(\lambda-L_0)^*=\infty$ *(so that* $\operatorname{ind}(\lambda-L_0)=-\infty$*) for every $\lambda\in\Omega$.*

(iv) $\mathscr{A}'(L_0)\supset\{M_g\colon g\in\mathbf{A}_{0\infty}\}$, *so that* $\bar\mu[\mathscr{A}'(L_0)]=1$.

(v) $\mathscr{A}^a(L_0)=\mathscr{A}''(L_0)=\{M_g\colon g\in\mathbf{A}_{0\infty}, g(z,t)$ is constant with respect to t for every (fixed) $z\in\Omega^-\}$ = *norm-closure of the rational functions of L_0 with poles outside Ω^-;*

(vi) $\operatorname{appr}(L_0)''=\mathscr{A}''(L_0)$.

Proof. The statements relative to $\mathbf{W}_{0\infty}$ and $\mathbf{A}_{0\infty}$ (in particular, $\mathscr{M}(\mathbf{W}_{0\infty})\approx\Gamma_0^-$ and $\mathscr{M}(\mathbf{A}_{0\infty})\approx\Xi_0$) can be proved exactly as in the previous section. Now (i), (ii) and (iii) are clear and (iv) is obvious.

(v) Let $A\in\mathscr{A}''(L_0)$. Since A commutes with the maximal abelian algebra of all multiplications by the elements of $\mathbf{A}_{0\infty}$, A must be a multiplication too: $A=M_g$, where $g=Ae\in\mathbf{A}_{0\infty}$.

For every $\tau\in[0,1]$, define C_τ by $C_\tau f(z,t)=f[z,1/2+(t-\tau)/2]$. By using, e.g., [1], it is not difficult to see that C_τ is bounded and commutes with A; therefore, $g(z, \tau)=Ae(z, \tau)=AC_\tau e(z, \tau)=C_\tau Ae(z, \tau)=C_\tau g(z, \tau)=g(z, 1/2)$, i.e., g depends only on z.

By the definition and properties of $\mathbf{A}_{0\infty}$, it follows that $g(z, t)$ is the norm-limit of a sequence of rational functions with poles outside Ω^-. Since $\mathscr{A}'(L_0)$ is strictly cyclic, this implies that $A=M_g$ is a norm-limit of rational functions of L_0 with poles outside Ω^- (see [37]). This proves (v).

(vi) It is obvious that for every $C\in\mathscr{L}(\mathfrak{K})$, $\operatorname{appr}(C)''$ is inverse-closed, so that $\operatorname{appr}(C)''$ always contains the norm-closure of the rational functions of C with poles outside $\sigma(C)$. Now (vi) follows from (v). □

Lemma 6. *Let Ω be an open set with analytic boundary, let n be a natural number, let $\mathbf{W}_{0n}$ be the direct sum of n copies of $\mathbf{W}^{2,1}(\partial\Omega, dm)$ and let $\mathbf{A}_{0n}$ be the subspace of "analytic elements" of $\mathbf{W}_{0n}$. Then $\mathbf{W}_{0n}$ and $\mathbf{A}_{0n}$ are Banach algebras with identity (under an equivalent norm), $\mathscr{M}(\mathbf{W}_{0n})\approx\partial\Omega\times\{1/n, 2/n, \dots, 1\}$ and $\mathscr{M}(\mathbf{A}_{0n})\approx\approx\Omega^-\times\{1/n, 2/n, \dots, 1\}$.*

If $T_{0n}=M_z$ in $\mathbf{W}_{0n}$, then $\mathbf{A}_{0n}$ is invariant under T_{0n} and its restriction $L_{0n}=T_{0n}|\mathbf{A}_{0n}$ satisfy

(i) $\sigma(T_{0n})=E_l(T_{0n})=E_r(T_{0n})=E_l(L_{0n})=E_r(L_{0n})=\partial\Omega$.

(ii) $\sigma(L_{0n})=\Omega^-$.

(iii) $\operatorname{Ker}(\lambda-L_{0n})=\{0\}$ *and* $\dim\operatorname{Ker}(\lambda-L_{0n})^*=n$ *(so that* $\operatorname{ind}(\lambda-L_{0n})=-n$*) for every $\lambda\in\Omega$.*

(iv) $\mathscr{A}'(L_{0n})\cong\mathbf{A}_{0n}^{(n\times n)}$ *is the algebra of all $n\times n$ operator matrices with entries in $\{M_g\colon g\in\mathbf{A}_{0n}\}$, so that $\bar{\mu}[\mathscr{A}'(L_{0n})]=1$.*

(v) $\mathscr{A}^a(L_{0n})=\mathscr{A}''(L_{0n})\cong\{M_g\colon g\in\mathbf{A}_{0n}\}$ *=norm-closure of the rational functions of L_{0n} with poles outside Ω^-.*

(vi) $\operatorname{appr}(L_{0n})''=\mathscr{A}''(L_{0n})$.

The proof (that can be easily "modelled" on that of Lemma 5) is left to the reader.

Now it is clear that if $T\doteq F\oplus\{\bigoplus_{1}^{m} L(\Omega_k, n_k)\}\oplus\{\bigoplus_{j=1}^{p} L(\Omega_j^*, -n_j)^*\}$, where F is an operator acting on a finite dimensional space, $L(\Omega, n)$ is the operator defined by Lemma 5 (for $n=-\infty$) and by Lemma 6 (for $-\infty<n<0$) and $\{\sigma(F), \{\Omega_k^-\}_{k=1}^m, \{\Omega_j^-\}_{j=1}^p\}$ $(0\leqq m, p<\infty)$ is a family of pairwise disjoint compact sets, then $\operatorname{appr}(T)''=\mathscr{A}''(T)=$norm-closure of the rational functions of T with poles outside $\sigma(T)$.

A formal repetition of the proof of Theorem 3 shows that the operators in $\mathscr{L}(\mathfrak{K})$ that are similar to some T as above form a dense subset, whence we obtain

Corollary 4. *(K) is dense in $\mathscr{L}(\mathfrak{K})$.*

References

[1] R. A. ADAMS, *Sobolev spaces*, Academic Press (New York—San Francisco—London, 1975).

[2] C. APOSTOL, Operators quasi-similar to a normal operator, *Proc. Amer. Math. Soc.*, **53** (1975), 104—106.

[3] C. APOSTOL, R. G. DOUGLAS and C. FOIAŞ, Quasi-similar models for nilpotent operators, *Trans. Amer. Math. Soc.*, **224** (1976), 407—415.

[4] C. APOSTOL, C. FOIAŞ and D. VOICULESCU, Some results on non-quasitriangular operators. IV, *Rev. Roum. Math. Pures Appl.*, **18** (1973), 487—514.

[5] C. APOSTOL, C. FOIAŞ and D. VOICULESCU, On the norm-closure of nilpotents. II, *Rev. Roum. Math. Pures Appl.*, **19** (1974), 549—577.

[6] C. APOSTOL and B. MORREL, On approximation of operators by simple models, *Indiana Univ. Math. J.*, **26** (1976), 427—442.

[7] J. BARRÍA and D. A. HERRERO, Closure of similarity orbits of nilpotent operators. I: Finite rank operators, *J. Operator Theory*, **1** (1979), to appear.

[8] J. BARRÍA and D. A. HERRERO, Closure of similarity orbits of nilpotent operators. II *(to appear)*.

[9] S. CLARY, Equality of spectra of quasi-similar hyponormal operators, *Proc. Amer. Math. Soc.* **53** (1975), 88—90.

[10] I. COLOJOARĂ and C. FOIAŞ, *Theory of generalized spectral operators*, Gordon and Breach (New York, 1968).

[11] R. G. DOUGLAS, On the operator equation $S^*XS=X$ and related topics, *Acta Sci. Math.* **30** (1969), 19—32.

[12] R. G. DOUGLAS and C. PEARCY, Hyperinvariant subspaces and transitive algebras, *Mich. Math. J.*, **19** (1972), 1—12.

[13] L. A. FIALKOW, A note on direct sums of quasinilpotent operators, *Proc. Amer. Soc.*, **48** (1975), 125—131.

[14] L. A. FIALKOW, A note on quasisimilarity of operators, *Acta Sci. Math.*, **39** (1977), 67—85.

[15] L. A. FIALKOW, A note on quasisimilarity. II, *Pacific J. Math.*, **70** (1977), 151—162.

[16] L. A. FIALKOW, A note on the operator $X\to AX-XB$, *Transactions Amer. Math. Soc.*, **243** (1978), 147—168.

[17] P. A. FILLMORE, *Notes in operator theory*, D. Van Nostrand, Reinhold Mathematical Studies No. 30 (New York, 1970).

[18] C. FOIAŞ and C. PEARCY, A model for quasinilpotent operators, *Mich. Math. J.*, **21** (1974), 399—404.

[19] C. FOIAŞ, C. PEARCY and D. VOICULESCU, Biquasitriangular operators and quasisimilarity, *Linear spaces and approximation* (Proc. Conf. Oberwolfach, 1977), Birkhäuser (Basel, 1978), 47—52.

[20] C. FOIAŞ, C. PEARCY and J. P. WILLIAMS, Some remarks on the Volterra operator, *Proc. Amer. Math. Soc.*, **31** (1972), 177—184.

[21] T. W. GAMELIN, *Uniform algebras*, Prentice Hall Inc., (Englewood Cliffs, New Jersey, 1969).

[22] R. GELLAR, Two sublattices of weighted shifts invariant subspaces, *Indiana Univ. Math. J.*, **23** (1973), 1—10.

[23] S. GRABINER, Derivations and automorphisms of Banach algebras of power series, *Memoirs Amer. Math. Soc.* **146** (Providence, Rhode Island, 1974).

[24] D. W. HADWIN, An asymptotic double commutant theorem for C*-algebras, *Transactions Amer. Math. Soc.*, **244** (1978), 273—297.

[25] P. R. HALMOS, *A Hilbert space problem book*, D. Van Nostrand (Princeton, New Jersey, 1967).

[26] D. A. HERRERO, Álgebras de operadores transitivas que contienen una subálgebra de multiplicidad estricta finita, *Rev. Un. Mat. Argentina*, **26** (1972), 77—84.

[27] D. A. HERRERO, Operator algebras of finite strict multiplicity, *Indiana Univ. Math. J.*, **22** (1972), 13—24.
[28] D. A. HERRERO, Operator algebras of finite strict multiplicity. II, *Indiana Univ. Math. J.*, **27** (1978), 9—18.
[29] D. A. HERRERO, Normal limits of nilpotent operators, *Indiana Univ. Math. J.*, **23** (1974), 1097—1108.
[30] D. A. HERRERO, A Rota universal model for operators with multiply connected spectrum, *Rev. Roum. Math. Pures Appl.*, **21** (1976), 15—23.
[31] D. A. HERRERO, Universal quasinilpotent operators, *Acta Sci. Math.*, **38** (1976), 291—300.
[32] D. A. HERRERO, On the spectra of the restrictions of an operator, *Trans. Amer. Math. Soc.*, **233** (1977), 45—58.
[33] D. A. HERRERO, Clausura de las órbitas de similaridad de operadores en espacios de Hilbert, *Rev. Un. Mat. Argentina*, **27** (1976), 244—260.
[34] D. A. HERRERO, Closure of similarity orbits of Hilbert space operators. II: Normal operators, *J. London Math. Soc.*, (2) **13** (1976), 299—316.
[35] D. A. HERRERO, Closure of similarity orbits of Hilbert space operators. III, *Math. Ann.* **232** (1978), 195—204.
[36] D. A. HERRERO, Quasisimilarity does not preserve the hyperlattice, *Proc. Amer. Math. Soc.*, **65** (1977), 80—84.
[37] D. A. HERRERO and A. LAMBERT, On strictly cyclic algebras, $\mathscr{P}$-algebras and reflexive operators, *Trans. Amer. Math. Soc.*, **185** (1975), 229—235.
[38] D. A. HERRERO and N. SALINAS, Analytically invariant and bi-invariant subspaces, *Trans. Amer., Math. Soc.*, **173** (1972), 117—136.
[39] T. B. HOOVER, Quasi-similarity of operators, *Illinois J. Math.*, **16** (1972), 678—686.
[40] T. KATO, *Perturbation theory for linear operators*, Springer-Verlag (New York, 1966).
[41] H. RADJAVI and P. ROSENTHAL, *Invariant subspaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, B. 77, Springer-Verlag (New York—Heidelberg—Berlin, 1973).
[42] C. RICKART, *General theory of Banach algebras*, D. Van Nostrand, (Princeton, New Jersey, 1960).
[43] G.-C. ROTA, On models for linear operators, *Comm. Pure Appl. Math.*, **13** (1960), 469—472.
[44] N. SALINAS, Operators with essentially disconnected spectrum, *Acta Sci. Math.*, **33** (1972), 193—206.
[45] N. SALINAS, A characterization of the Browder spectrum, *Proc. Amer. Math. Soc.*, **38** (1973), 369—373.
[46] R. SCHATTEN, *Norm ideals of completely continuous operators*, 2nd. ed., Ergebnisse der Mathematik und ihrer Grenzgebiete, B. 27, Springer-Verlag (Berlin—Heidelberg—New York, 1970).
[47] V. S. ŠULMAN, Operator algebras with strictly cyclic vectors, *Mat. Zametki*, **16** (2) (1974), 253—257. (Russian)
[48] B. SZ.-NAGY and C. FOIAŞ, *Analyse harmonique de opérateurs de l'espace de Hilbert*, Masson et Cie.—Akadémiai Kiadó (Paris—Budapest, 1967).
[49] D. VOICULESCU, A non-commutative Weyl—von Neumann theorem, *Rev. Roum. Math. Pures et Appl.*, **21** (1976), 97—113.

DEPARTMENTO DE MATEMATICAS, APARTADO 1827
INSTITUTO VENEZOLANO DE INVESTIGACIONES CIENTIFICAS
CARACAS 101, VENEZUELA

Acta Sci. Math., 41 (1979), 119—125

Affine algebras in congruence modular varieties

CHRISTIAN HERRMANN

Algebras which are polynomially equivalent to a module have been characterized by CSÁKÁNY [3], [4] in terms of the associated system of congruence classes. Recently, SMITH [10] and GUMM [7] characterized such algebras within congruence permutable classes following the lines of "Remak's Principle", cf. [2, p. 167]. In this note their results will be extended to congruence modular classes.

Definition. A (general) algebra A is called *abelian* if in the congruence lattice of $A\times A$ there exists a common complement of the kernels of the two projections.

Theorem. *Every abelian algebra in a congruence modular variety is polynomially equivalent to a module over a suitable ring. The abelian algebras form a subvariety.*

Here the polynomial equivalence of two algebras with the same base set means that the sets of their algebraic functions coincide.

Corollary A. *Let $\mathscr{A}$ and $\mathscr{B}$ be subvarieties of a congruence modular variety, $\mathscr{A}$ abelian and $\mathscr{B}$ congruence distributive. Then every algebra in the join of $\mathscr{A}$ and $\mathscr{B}$ is a direct product of an algebra in $\mathscr{A}$ and an algebra in $\mathscr{B}$.*

Now, the finite base theorems of BAKER [1] and MCKENZIE [9] join into one.

Corollary B. *There exists a finite equational base for every congruence modular variety which is generated by finitely many finite algebras each of which is either abelian or generates a congruence distributive subvariety.*

The idea of the proof can be easily stated: For an abelian group A the difference is a homomorphism of A^2 onto A which has the diagonal $D=\{(x, x)\mid x\in A\}$ as its kernel. Thus, the group structure can be recovered from the natural homomorphism $A^2\to A^2/D$ via the identification $x\mapsto (x, 0)+D$ of A and A^2/D. In general,

Received August 1, 1977.

it is still true that an abelian algebra has a congruence $\varkappa$ which has D as a class — cf. HAGEMANN and HERRMANN [8] — and we may define the difference by the natural homomorphism $A^2 \to A^2/\varkappa$. Assume for a moment that 0 is an idempotent element of A. Then $x \mapsto [(x, 0)]\varkappa$ is an embedding of A into $A^2/\varkappa$ but not necessarily onto. Therefore, a limit construction is used to embed A into an algebra B which is closed under the group operations. Using DAY's [5] terms for congruence modularity one sees that A is a subgroup, too.

1. The centring congruence. The proofs rely on results of HAGEMANN and HERRMANN [8]. Thus, a general assumption to be made is that the algebras are *strictly modular* which means that every "diagonal" subdirect product $B \subseteqq A^n$ — with n finite, $(x, \ldots, x) \in B$ for all x in A — is congruence modular. We write xy for pairs, $xyzu$ for quadruples, $[a]\alpha$ for the congruence class of a modulo α. Let η_0, η_1 denote the kernels of the two projections of A^2 onto A.

Proposition 1. *A strictly modular algebra A is abelian if and only if there is a congruence $\varkappa$ on A^2 such that*

(C) $\eta_0 \cap \varkappa = \eta_1 \cap \varkappa = 0$

(RR) *$xx\varkappa uu$ for all x and u in A.*

If A is abelian then $\varkappa = \zeta(A)$ is uniquely determined and it holds

(RS) *$xy\varkappa uv$ implies $yx\varkappa vu$*

(RT) *$xy\varkappa uv$ and $yz\varkappa vw$ imply $xz\varkappa uw$*

(SW) *$xy\varkappa uv$ if and only if $xu\varkappa yv$.*

Proof. Everything but (SW) is shown in [8], Thm. 1.4 and Prop. 1.6. Now, define λ by $xy\lambda uv$ if and only if $xu\varkappa yv$. Due to (RR), (RS) and (RT) λ is a congruence on A^2. Since $\varkappa$ is reflexive it satisfies (RR). Finally, assume $xy\lambda xv$, i.e. $xx\varkappa yv$. By (RR) we have $yy\varkappa xx$, hence $yy\varkappa yv$ and $y = v$ by (C). This proves $\eta_0 \cap \lambda = 0$ and, by symmetry, $\eta_1 \cap \lambda = 0$. By the uniqueness of $\varkappa$ it follows $\varkappa = \lambda$ which means (SW).

Lemma 2. *Let A be strictly modular and abelian, $\varkappa = \zeta(A)$. Then $A^2/\varkappa$ is strictly modular and abelian, too, and with $\lambda = \zeta(A^2/\varkappa)$ it holds for all a, b, c, e in A*

$$([ae]\varkappa, [be]\varkappa)\lambda([ab]\varkappa, [ee]\varkappa) \tag{1}$$

$$([ae]\varkappa, [bc]\varkappa)\lambda([ce]\varkappa, [ba]\varkappa). \tag{2}$$

Proof. Consider A^4 and let $\Theta_0, \Theta_1, \Theta_2, \Theta_3$ be the kernels of the projections. For each $i < j$ there is a "copy" $\varkappa_{ij}$ of $\varkappa$ on A^4 given by

$$x_0x_1x_2x_3\varkappa_{ij}y_0y_1y_2y_3 \quad \text{if and only if} \quad x_ix_j\varkappa y_iy_j.$$

Because of $\varkappa_{01} \supseteqq \Theta_0 \cap \Theta_1$ and $\varkappa_{23} \supseteqq \Theta_2 \cap \Theta_3$ both permute and have join 1. Therefore, the map φ with $\varphi(x_0x_1x_2x_3) = ([x_0x_1]\varkappa, [x_2x_3]\varkappa)$ is a homomorphism of

A^4 onto C^2 where $C=A^2/\varkappa$. Its kernel is $\varepsilon=\varkappa_{01}\cap\varkappa_{23}$. We claim that the image of $\mu=\varepsilon+\varkappa_{12}\cap\varkappa_{03}$ is the congruence $\zeta(C)$ on C^2. We have to show

$$\varkappa_{01}\cap\mu=\varkappa_{23}\cap\mu=\varepsilon \quad\text{and}\quad xxxx\mu uuuu \quad\text{for all } x \text{ and } u \text{ in } A.$$

The second is obvious. By modularity we get $\varkappa_{01}\cap\mu=\varepsilon+\varkappa_{01}\cap\varkappa_{12}\cap\varkappa_{03}$. Now, consider $x_0x_1x_2x_3\varkappa_{01}\cap\varkappa_{12}\cap\varkappa_{03}y_0y_1y_2y_3$. By (RS) we have $x_2x_1\varkappa y_2y_1$ and $x_1x_0\varkappa y_1y_0$, hence $x_2x_0\varkappa y_2y_0$ by (RT). With $x_0x_3\varkappa y_0y_3$ and a second application of (RT) it follows $x_2x_3\varkappa y_2y_3$. This shows $x_0x_1x_2x_3\varkappa_{23}y_0y_1y_2y_3$, i.e. $\varkappa_{01}\cap\varkappa_{12}\cap\cap\varkappa_{03}\subseteq\varkappa_{23}$ and $\varkappa_{01}\cap\mu=\varepsilon$. $\varkappa_{23}\cap\mu=\varepsilon$ follows by symmetry.

By Proposition 1 the image of μ has properties (RT) and (SW), i.e.

(3) $\qquad xyuv\mu abcd$ and $uvst\mu cdef$ imply $xyst\mu abef$, and

(4) $\qquad xyuv\mu abcd$ if and only if $xyab\mu uvcd$.

On the other hand, all the arguments about μ remain valid if we interchange $\varkappa_{01}$ and $\varkappa_{23}$ with $\varkappa_{12}$ and $\varkappa_{03}$. In particular, property (SW) reads then

(5) $\qquad xyuv\mu abcd$ if and only if $byuc\mu axvd$.

Moreover, recall that $\varkappa$ is reflexive and satisfies (RR). Thus, since $\mu\supseteq\varkappa_{01}\cap\varkappa_{23}$ and $\mu\supseteq\varkappa_{12}\cap\varkappa_{03}$, we have

(6) $xxuv\mu aauv$, (7) $xyuu\mu xycc$, (8) $xyux\mu ayua$, (9) $xyyv\mu xbbv$.

Now, we are ready to prove (1): $aaba\mu babb$ holds by (8) and $baaa\mu bbba$ by (9) whence $aaaa\mu baba$ by (3). $eeaa\mu aaaa$ holds by (6) and it follows $eeaa\mu baba$ by the transitivity of μ. An application of (5) yields $aeab\mu beaa$. Since $beaa\mu beee$ by (7) one concludes $aeab\mu beee$ by the transitivity of μ. Thus, $aebe\mu abee$ by (4). To prove (2) substitute in $aaaa\mu baba$ b by c to get $aaaa\mu caca$. By (6) it holds $eeaa\mu aaaa$ and by (7) $eebb\mu eeaa$ whence $eebb\mu caca$ by the transitivity of μ. Thus, $aebc\mu ceba$ by (5).

2. Embedding into a "linear" algebra. Call an algebra A *linear* — with respect to an abelian group structure $(A, +, -, 0)$ on A — if 0 is an idempotent element of A and if "$-$" (and "$+$") are homomorphisms of A^2 into A. Linear algebras are just reducts of modules: If A is linear let R be the set of all unary functions on A which are induced by terms in the language of A with 0 added as a constant. With pointwise addition and with composition R becomes a unitary ring. Its operation on A makes A a faithful unitary R-module A_R. Given any fundamental operation f of A one has

$$f(x_1\ldots x_n)=f(x_10\ldots 0)+\ldots+f(0\ldots 0x_n),$$

i.e. f is desciibed by a term in the language of A_R — cf. SMITH [10]. For a class

$\mathscr{C}$ of algebras let $\mathbf{D}\mathscr{C}$, $\mathbf{H}\mathscr{C}$, $\mathbf{S}\mathscr{C}$, $\mathbf{P}_s^f\mathscr{C}$ denote the class of all direct unions, homomorphic images, subalgebras, and finite subdirect products of algebras in $\mathscr{C}$ resp.

Lemma 3. *Let A be a strictly modular abelian algebra having an idempotent element* 0. *Then A can be embedded into an algebra B in $\mathbf{DHP}_s^f A$ which is linear with respect to an abelian group* $(B, +, -, 0)$.

Proof. In view of Lemma 2 we may define a series of strictly modular abelian algebras:

$$A_0 = A, \quad A_{n+1} = A_n^2/\zeta(A_n).$$

Let π_n be the canonical homomorphism of A_n^2 onto A_{n+1}. Clearly, for every n, $0_{n+1} = [xx]\zeta(A_n)$ is an idempotent element of A_{n+1}. Thus, with $0_0=0$ and $\varepsilon_n x = [x0]\zeta(A_n)$ one gets due to (*C*) for every n an embedding $\varepsilon_n\colon A_n \to A_{n+1}$ such that $\varepsilon_n 0_n = 0_{n+1}$. Let A_∞ be the direct union over the system (A_n, ε_n) and idenify A_n with its image in A_∞. Applying Lemma 2(1) to A_n we see that for each n $\varepsilon_{n+1}\circ\pi_n = \pi_{n+1}\circ(\varepsilon_n\times\varepsilon_n)$. Therefore, $a-b=\pi_n(a,b)$ if a and b are in A_n, defines a map of A_∞^2 into A_∞. By definition it is compatible with the fundamental operations of A_∞ and it holds $a-0=a$, $a-a=0$. Moreover, by Lemma 2(2) it follows $a-(b-c)=c-(b-a)$. Thus, with $a+b=a-(0-b)$ one gets an abelian group structure on A_∞ which makes it linear.

3. Using the Day terms. For all of the following suppose that we work within a fixed congruence modular variety $\mathscr{V}$. Then, due to Day [5] there are a number n and 4-variable terms $m_0, \dots, m_n$ in the language of $\mathscr{V}$ such that the following identities hold in $\mathscr{V}$:

(m1) $m_0(xyzu) = x$ and $m_n(xyzu) = y$,

(m2) $m_i(xxzz) = x$ for all $i = 0, \dots, n$,

(m3) $m_i(xyzz) = m_{i+1}(xyzz)$ for i even,

(m4) $m_i(xyxy) = m_{i+1}(xyxy)$ for i odd.

We define by induction $p_0(xzu)=x$,

$$p_{i+1}(xzu) = \begin{cases} m_{i+1}(p_i(xzu), p_i(xzu), u, z) & \text{for } i \text{ even}, \\ m_{i+1}(p_i(xzu), p_i(xzu), z, u) & \text{for } i \text{ odd}. \end{cases}$$

Obviously, in $\mathscr{V}$ it holds $p_i(xzz)=x$ for all i. Put $p(xzu)=p_{n-1}(xzu)$. Then $p(xzz)=x$ holds in $\mathscr{V}$.

Call an algebra A *affine* if it is polynomially equivalent to a linear algebra A^∇ or, in other words, if there is an abelian group structure $(A, +, -, 0)$ on A

such that for every fundamental operation f of A there is an f^∇ linear with respect to $(A, +, -, 0)$ such that

$$f(x_1 \dots x_n) = f^\nabla(x_1 \dots x_n) + f(0 \dots 0).$$

Lemma 4. *In an affine algebra* $A \in \mathcal{V}$ *it holds* $p(xzu) = x - z + u$.

Proof. Since A is polynomially equivalent to an R-module A_R for each $i = 0, \dots, n$ there are $\alpha_i, \beta_i, \gamma_i, \delta_i$ in R and c_i in A such that

$$m_i(xyzu) = \alpha_i x + \beta_i y + \gamma_i z + \delta_i u + c_i$$

holds in A. $(m2)$ yields $0 = m_i(0000) = c_i$, $x = m_i(xx00) = (\alpha_i + \beta_i)x$, and $0 = m_i(00zz) = (\gamma_i + \delta_i)z$. Since A_R is faithful it follows $\alpha_i + \beta_i = 1$ and $\gamma_i + \delta_i = 0$. In particular, we get

$$m_i(xxvw) = x - \delta_i v + \delta_i w \quad \text{for} \quad i = 0, \dots, n.$$

By induction one concludes

$$p_k(xzu) = x - \sum_{i=1}^{k} (-1)^i \delta_i z + \sum_{i=1}^{k} (-1)^i \delta_i u. \tag{10}$$

On the other hand, $(m1)$ yields $0 = m_0(0y00) = \beta_0 y$ and $0 = m_0(000u) = \delta_0 u$, as well as $0 = m_n(000u) = \delta_n u$ and $y = m_n(0y00) = \beta_n y$ whence $\beta_0 = \delta_0 = \delta_n = 0$ and $\beta_n = 1$. Finally, $(m3)$ and $(m4)$ imply $\beta_i y = m_i(0y00) = m_{i+1}(0y00) = \beta_{i+1} y$ for i odd and $(\beta_i + \delta_i)y = m_i(0y0y) = m_{i+1}(0y0y) = (\beta_{i+1} + \delta_{i+1})y$ for i even. Thus, it holds $\beta_{i+1} = \beta_i$ for i odd and $\beta_{i+1} = \beta_i + \delta_i - \delta_{i+1}$ for i even. By induction one gets $\beta_k = \beta_{k+1} = \sum_{i=1}^{k} (-1)^i \delta_i$ for k odd. In particular, with $m = n - 1$ if n even and $m = n$ if n odd we have $1 = \beta_n = \sum_{i=1}^{m} (-1)^i \delta_i$. Then with (10) it follows $p(xzu) = x - z + u$.

Corollary 5. *If α is a congruence of $A \in \mathcal{V}$ such that A/α is affine then α permutes with every congruence of A.*

Proof. Let β be a congruence of A and suppose $x\alpha y\beta z$. Then $p(xyz)\beta x$ since $p(xyy) = x$ holds in $\mathcal{V}$ and $p(xyz)\alpha z$ by Lemma 4. Thus, $z\alpha p(xyz)\beta x$.

4. Proof of the Theorem. First, suppose that the abelian algebra $A \in \mathcal{V}$ has an idempotent element 0. Construct the linear algebra $A_\infty \supseteq A$ according to Lemma 3. By Lemma 4 there is a term $p(xyz)$ in the language of $\mathcal{V}$ such that $p(xyy) = x = p(yyx)$ holds in A_∞. In particular, all subalgebras of A_∞ are congruence permutable and each of the embeddings ε_n is onto: $\eta_1 \circ \varkappa = 1$ implies that for every xy there is uv such that $00\eta_1 uv\varkappa xy$ which means $u0\varkappa xy$. Thus, in fact $A_\infty = A$ and A is linear itself. Since $x - y + z = p(xyz)$ is represented by a term in the language of A we get every term of A_R after joining 0 as a constant. In general, choose an arbitrary element 0 of A and consider the map ε: $A \to A^2/\varkappa$ with $x = [x0]\varkappa$. $A^2/\varkappa$

has the idempotent element $[xx]\varkappa$ hence it is linear by the above. ε is still one-to-one by (C) and in view of Lemma 2 (1) it satisfies

$$f(x_1, \dots, x_n) = \varepsilon f(x_1, \dots, x_n) - \varepsilon f(0, \dots, 0). \tag{11}$$

for every fundamental operation f of A. Hence, it holds

$$p(\varepsilon x, \varepsilon y, \varepsilon z) = \varepsilon p(x, y, z) - \varepsilon p(0, 0, 0) = \varepsilon p(x, y, z) - \varepsilon 0 = \varepsilon p(x, y, z), \tag{12}$$

since p is a term and $\varepsilon 0 = [00]\varkappa$ is the neutral element of the linear algebra $A^2/\varkappa$. Therefore, $\varepsilon(A)$ is closed under the operation $p(xyz) = x - y + z$ and an abelian group with zero $0 = \varepsilon 0$, $x + z = p(x0z)$, and $x - y = p(xy0)$. If we transfer the group operations via ε^{-1} to A then (11) states that A is affine. Moreover, by (12) we have $p(xyz) = x - y + z$ on A. Indeed, A and A^2 are congruence permutable and ε is an onto map, too. Moreover, the full module structure of A_R can be recovered from A after adding the constant 0.

That the abelian algebras in a congruence modular variety form a subvariety is obvious by Proposition 1. As a defining set of identities one can use $p(xyy) = p(yyx) = x$ and the identities expressing the compatibility of p and the fundamental operations of $\mathscr{V}$; cf. GUMM [7].

5. Proof of Corollary A. First, observe that $\mathscr{A}$ and $\mathscr{B}$ have only the trivial algebra in common. Every algebra in the join of $\mathscr{A}$ and $\mathscr{B}$ is a homomorphic image C/Θ of a subdirect product $C \subseteq A \times B$ with $A \in \mathscr{A}$ and $B \in \mathscr{B}$. Let α and β denote the kernels of the projections of C onto A and B, respectively. Since $C/\alpha + \beta$ is in both $\mathscr{A}$ and $\mathscr{B}$ it must hold $\alpha + \beta = 1$. Then, by Corollary 5, C is the direct product of A and B.

Since B generates a congruence distributive variety, β is a neutral element of the congruence lattice of C (see [8, Thm. 4.1]) which implies $\Theta = \Theta + \alpha \cap \beta = (\Theta + \alpha) \cap \cap (\Theta + \beta)$. Thus, C/Θ is itself a subdirect product of an algebra in $\mathscr{A}$ and one in $\mathscr{B}$ and, by the above argument, even a direct product.

6. Proof of Corollary B. Let $\mathscr{C}$ be congruence modular and generated by finite algebras $A_1, \dots, A_n, B_1, \dots, B_m$ where each A_i is abelian and each B_i generates a congruence distributive subvariety. Let $\mathscr{A}$ and $\mathscr{B}$ be the subvarieties generated by the $A_1, \dots, A_n$ and the $B_1, \dots, B_m$, respectively. Then $\mathscr{B} = \mathbf{DHP}_s^f\mathbf{S}\{B_1, \dots, B_m\}$ is congruence distributive due to [8, Cor. 4.3] and has a finite equational base due to BAKER [1]. The variety $\mathscr{A}$ is polynomially equivalent (via finitely many constants) to the variety of all modules over a fixed ring R: take the free algebra on countably many generators in $\mathscr{A}$ and apply the Theorem. Since $\mathscr{A}$ is locally finite, R has to be finite. Thus, $\mathscr{A}$ has a finite equational base, too.

By Corollary A $\mathscr{A}$ and $\mathscr{B}$ are independent in the sense of GRÄTZER, LAKSER, and PŁONKA [6, Thm. 2]. In particular, one can define predicates for the congruences which yield the direct product decomposition. Therefore, $\mathscr{C}=\mathscr{A}\vee\mathscr{B}$ is finitely axiomatizable, i.e. it has a finite equational base.

The author wishes to express his warmest thanks to the J. Bolyai Society and to B. Csákány and A. P. Huhn for inviting him to Szeged, the most appropriate place where to write this paper.

Added in March 78. Since several reformulations of our Theorem have been discovered meanwhile it seems necessary to add the following

Scholion. *For a strictly modular algebra A the following are equivalent:*
(1) *A is abelian.*
(2) *For the commutator introduced in* [8] *it holds* $[1_A, 1_A]=0_A$.
(3) *The diagonal D is a congruence class of $A\times A$.*

Implications (1)⇒(2), (2)⇒(1), and (2)⇔(3) are instances of Thm. 1.4, Observation 1.2, and Cor. 2.4 in [8] respectively. Moreover, using Cor. 1.2 it is easily seen that for projective quotients α/β and γ/δ $[\gamma,\gamma]\subseteqq\delta$ implies $[\alpha,\alpha]\subseteqq\beta$. Thus, by Thm. 1.4 A is abelian if there is B and $\alpha\in\operatorname{con}(B)$ such that $B/\beta\cong A$ and $1_B/B$ is projective to a quotient of a sublattice of $\operatorname{con}(B)$ which is isomorphic to the 5-element lattice M_3.

References

[1] K. Baker, Finite equational bases for finite algebras in a congruence distributive equational class, *Advances in Math.*, **24** (1977), 207—243.
[2] G. Birkhoff, *Lattice Theory*, 3rd ed., Amer. Math. Soc. (Providence, 1967).
[3] B. Csákány, Abelian properties of primitive classes of universal algebras, *Acta Sci. Math.*, **25** (1964), 202—208. (Russian).
[4] B. Csákány, Varieties of affine modules, *Acta Sci. Math.*, **37** (1975), 3—10.
[5] A. Day, A characterization of modularity for congruence lattices of algebras, *Canad. Math. Bull.*, **12** (1969), 167—173.
[6] G. Grätzer—H. Lakser—J. Płonka, Joins and direct products of equational classes, *Canad. Math. Bull.*, **12** (1969), 741—744.
[7] H. P. Gumm, Algebras in congruence permutable varieties: geometrical properties of affine algebras, *Algebra Universalis*, **9** (1979), 8—34.
[8] J. Hagemann—C. Herrmann, A concrete ideal multiplication for algebraic systems and its relation to congruence distributivity, *Arch. Math. (Basel)*, to appear.
[9] R. McKenzie, Para primal varieties: A study of finite axiomatizability and definable principal congruences in locally finite varieties, *preprint*.
[10] J. D. H. Smith, *Mal'cev varieties*, Springer Lecture Notes 554 (Berlin, 1976).

DEPARTMENT OF MATHEMATICS
TECHNICAL UNIVERSITY OF DARMSTADT
HOCHSCHULSTRASSE 1
61 DARMSTADT, GFR

Acta Sci. Math., 41 (1979), 127—132

A characterisation of binary geometries of types $K(3)$ and $K(4)$

L. KÁSZONYI

In connection with halfplanar geometries I introduced the property $K(r)$ of binary geometries (see [2]). The aim of this paper is to give a new characterisation of such geometries for $r=3$ and 4.

First I give some definitions.

Definition 1. Let us consider the binary geometry G, which is embedded in the binary projective geometry Γ (i.e. in the geometry over $GF(2)$, $r(G)=r(\Gamma)$, $G\subseteqq\Gamma$). A subspace (point, line, hyperplane) m of Γ is called an *outer subspace* (point, line, hyperplane) of G, if m is not spanned by G-points.

We note that by the homogeneity of Γ this definition depends only on G.

Definition 2. A binary geometry G of rank n has the *property* $K(r)$ if every subgeometry of G of rank $n-r$ is contained in an outer hyperplane ($n\geqq r$, $r\geqq 2$).

Definition 3. A set $H=\{h_1, h_2, \ldots, h_m\}$ ($m\geqq 3$) of hyperplanes of the binary projective geometry Γ is a *hypercircuit* if

$$r\left(\bigcap_{\substack{i=1\\ i\neq j}}^{m} h_i\right)=r\left(\bigcap_{i=1}^{m} h_i\right)=r(\Gamma)-m+1$$

holds for every $j\in\{1, 2, \ldots, m\}$; m is called the *length* of H.

We shall frequently use the following

Theorem 1. (Two Colour Theorem) *$b\cap G\neq\emptyset$ holds for all hyperplanes b of a binary projective geometry $\Gamma(\supset G)$ if and only if G contains an odd circuit.*

The geometrical dual of Theorem 1 may be formulated as follows:

Theorem 2. *A set $\mathscr{L}$ of hyperplanes of a binary projective geometry Γ covers all Γ-points if and only if $\mathscr{L}$ contains an odd hypercircuit.*

Received June 20, 1977, in revised form March 3, 1978.

Our theorem is the following:

Theorem 3. *A binary geometry G of rank n $(\geqq r+1)$ is of type $K(r)$ $(r=3,4)$ if and only if for an arbitrary subspace a of G of rank $n-r-1$, the set $\mathcal{M}(a)$ of outer hyperplanes containing a contains an odd hypercircuit.*

First we prove

Lemma 1. *A binary geometry of rank 4 is of type $K(3)$ if and only if the set of its outer hyperplanes contains an odd hypercircuit.*

Proof. Sufficiency is clear by Theorem 2. We have to prove that if the set of outer planes covers the points of G then it covers the outer points of G as well (see Theorem 2).

Let us assume indirectly that the set $\mathcal{M}_3$ of outer planes of G covers the points in G but there is an outer point γ_1 which is not covered by $\mathcal{M}_3$. Consider the outer plane b_1 and let f be a line of b_1 which covers $b_1 \cap G$. The existence of such a line is trivial. Denote the planes incident in Γ to f and distinct from b_1, by d_1 and d_2. The set of planes $\{b_1, d_1, d_2\}$ covers all Γ-points. $\gamma_1 \notin b_1$, thus $\gamma_1 \in d_i \setminus f$ holds for an $i \in \{1, 2\}$ by our indirect hypothesis. Let for example $\gamma_1 \in d_1 \setminus f$ (see Fig. 1).

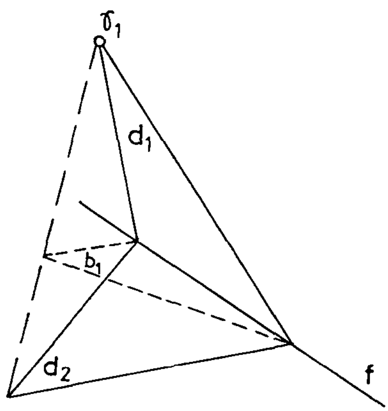


Fig. 1

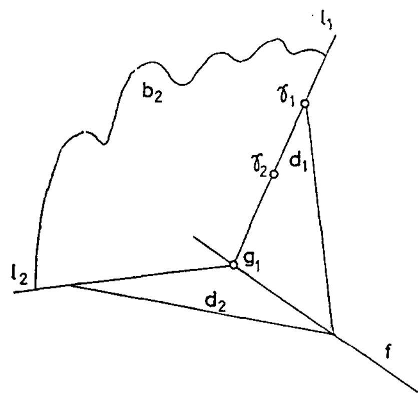


Fig. 2

We prove that for the set $U=d_1 \setminus f \setminus \{\gamma_1\}$, $U \subset G$ holds. Let us assume indirectly that U has a point γ_2 not in G. Consider the line $l_1=\sigma(\gamma_1, \gamma_2)$ *) and set $g_1=f \cap l_1$. Choose a point g of d_2 not on f and let $l_2=\sigma(g_1, g)$ (see Fig. 2). It is easy to see that the plane $b_2=\sigma(l_1 \cup l_2)$ is an outer plane containing γ_1, a contradiction.

*) $\sigma(\ldots)$ denotes the subspace of Γ spanned by the set given in the parentheses.

We show that U is an oval of d_1. Let $u_1, u_2 \in U$ be arbitrary. $\sigma(u_1, u_2) \cap f \neq \emptyset$ and $f \cap U = \emptyset$, thus every line of U consits of two points. But $|U|=3$, therefore U spans d_1. It is easy to see that γ_1 is a nucleus of U (i.e. the common point of tangentials to U).

Let $u_1 \in U$ be arbitrary, consider the outer plane b_3 containing u_1. The line $f_1 = b_3 \cap d_1$ cannot be tangential to U, thus f_1 intersects U in two points, say at u_1 and u_2 (see Fig. 3). This means that f_1 covers all the points of b_3, and the points of $f_2 = d_2 \cap b_3$ and not on f are outer points. Therefore the plane $\sigma(f_2 \cup \{\gamma_1\})$ is an outer plane containing γ_1, a contradiction.

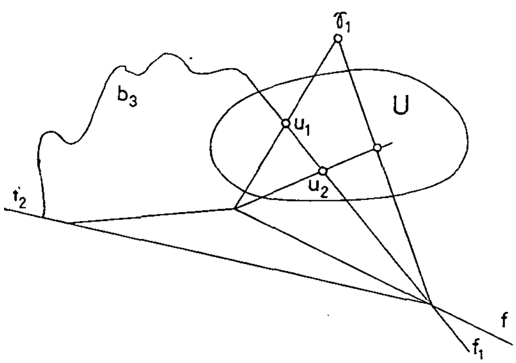


Fig. 3

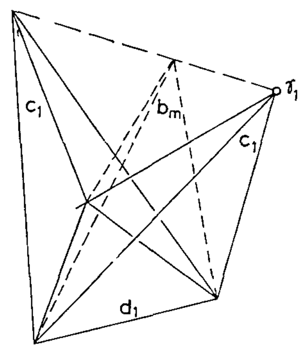


Fig. 4

Lemma 2. *A geometry G of rank 5 is of type $K(4)$ iff the set $\mathcal{M}_4$ of its outer hyperplanes contains an odd hypercircuit.*

Proof. We have to prove that if the elements of $\mathcal{M}_4$ cover G, then they cover the outer points as well. Let us assume indirectly that G is covered by the elements of $\mathcal{M}_4$ but G has an outer point γ_1 which is not covered by them. Let b_m be an element of $\mathcal{M}_4$ for which $|b_m \cap G|$ is maximal.

Let d_1 be a plane of b_m which covers the points of $b_m \cap G$. Let us denote by c_1 and c_2 the hyperplanes containing d_1 and distinct from b_m. The set $\{b_m, c_1, c_2\}$ covers all the Γ-points. Let $\gamma_1 \in c_1$ (see Fig. 4). Denote the set $(G \cap c_1) \setminus d_1$ by V. Let us project V from γ_1 to d_1. We prove that the projection V' meets all lines of d_1. Let us assume indirectly that d_1 has a line l for which $l \cap V' = \emptyset$ holds. Let g_2 be an arbitrary point of c_2 not on d_1. It is easy to see that $\sigma(\{g_2\} \cup l \cup \{\gamma_1\})$ is an outer hyperplane of G containg γ_1, a contradiction (see Fig. 5). Therefore V' contains a full line f_1. Let us consider the plane $d_2 = \sigma(f_1 \cup \{\gamma_1\})$. Making use of the fact that V' contains f_1, we can see that $U = V \cap d_2$ is an oval on d_2, the nucleous of which is γ_1.

Let $u \in U$ be arbitrary, denote the outer hyperplane containg the point u by b_u. We prove the validity of the following three statements:

(i) $d_2 \cap b_u$ is a line and a secant to U (i.e. $|U \cap (d_2 \cap b_u)| = 2$),

(ii) $b_u \cap c_1 \cap G \subseteqq d_2$,

(iii) $|b_u \cap G| \geqq 3$.

Part (i). $\gamma_1 \notin b_u$, thus $d_2 \not\subset b_u$, therefore $d_2 \cap b_u$ is a line. The lines of d_2 containing u are either secants or tangentials to U. The lines which are tangential to U contain γ_1, thus $d_2 \cap b_u$ is a secant.

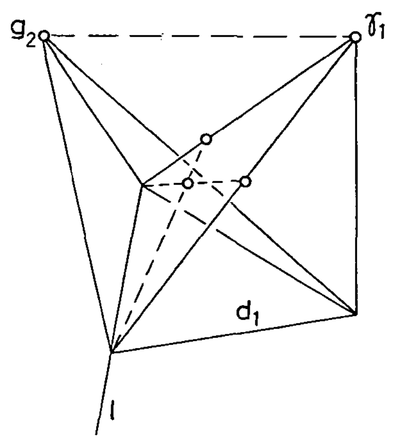


Fig. 5

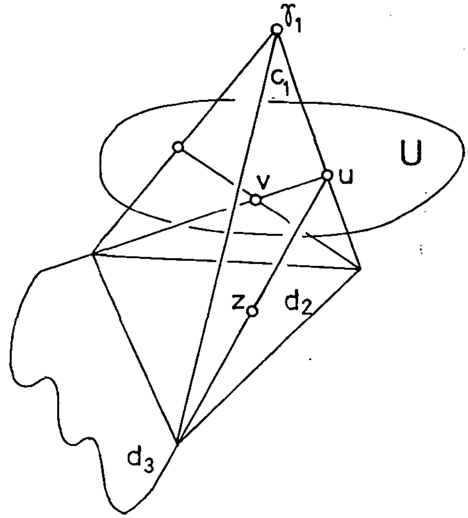


Fig. 6

Part (ii). Let $U \cap b_u = \{u, v\}$; assume indirectly that $b_u \cap G \cap c_1$ contains a point z not in U (see Fig. 6). The plane $d_3 = \sigma(u, v, z)$ covers $G \cap b_u$, because $b_u \in \mathcal{M}_4$. Therefore $d_3 \cap d_1$ covers $c_2 \cap b_u$. Using this fact we can see that $\sigma(c_2 \cap b_u \cup \{\gamma_1\})$ is an outer hyperplane, a contradiction (see Fig. 7).

Part (iii). $|b_u \cap U| = 2$, thus we have to prove $c_2 \cap b_u \cap G \neq \emptyset$. Assuming $c_2 \cap b_u \cap G = \emptyset$ we can see that $\sigma(c_2 \cap b_u \cup \{\gamma_1\})$ is an outer hyperplane.

Let $U = \{u_1, u_2, u_3\}$, and assume that $U \cap b_{u_2} = \{u_1, u_2\}$ and $U \cap b_{u_3} = \{u_1, u_3\}$. Set $\sigma(u_1, u_2) = s_1$, $\sigma(u_1, u_3) = s_2$, $c_1 \cap b_{u_2} = d_4$, $c_1 \cap b_{u_3} = d_5$, $d_1 \cap d_4 = s_4$, $d_1 \cap d_5 = s_5$, $g_{45} = s_4 \cap s_5$ (see Fig. 8). We prove that the plane $\sigma(s_4 \cup \{\gamma_1\})$ contains a point g_6 not on $\sigma(u_3, \gamma_1)$. Let us assume indirectly that

$$\{\sigma(s_4 \cup \{\gamma_1\}) \cap G\} \setminus \{u_3, \gamma_1\} = \emptyset.$$

Using (ii) we can see that $\sigma(\{c_2 \cap b_{u_2}\} \cup \{\gamma_1\})$ is an outer hyperplane. It is easy to see as well that g_6 is the point of the line $\sigma(\gamma_1, g_{45})$ distinct from γ_1 and g_{45}. Consider b_{g_6}. The line $b_{g_6}\cap d_2$ cannot be a secant to U by (ii), and cannot be tangential to U by $\gamma_1\notin b_{g_6}$, therefore $b_{g_6}\cap d_2=f_1$.

We prove that $|f_1\cap G|\leqq 1$. Let us assume indirectly that $|f_1\cap G|\geqq 2$. Let $g_7, g_8\in f_1\cap G$. Then the plane $\sigma(g_6, g_7, g_8)$ covers $b_{g_6}\cap G$. Using this fact we can see that the hyperplane $\sigma(\{c_2\cap b_{g_6}\}\cup\{\gamma_1\})$ is an outer hyperplane, a contradiction. Therefore $|\{f_1\cup s_4\cup s_5\}\cap G|\leqq 1$, thus $|b_m\cap G|=|d_1\cap G|\leqq 2$ holds, contradicting (iii) and the maximality property of b_m.

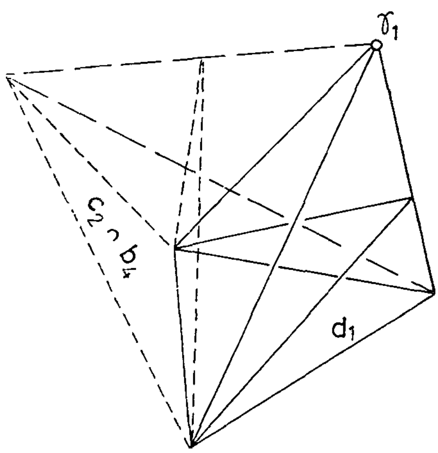


Fig. 7

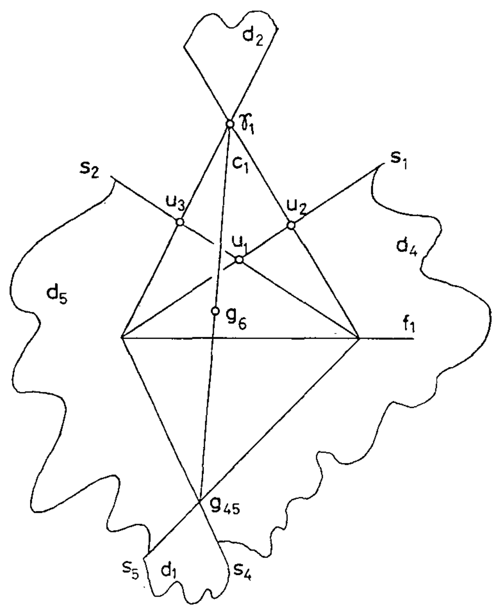


Fig. 8

Our Theorem 3 is now an easy consequence of Lemmas 1 and 2 and of the "Scum Theorem" (see [1]).

The following assertion may be proved by induction on $n=r(G)$ for $r=2, 3, 4$: if the outer hyperplanes of the binary geometry G cover all of its subspaces with rank $n-r$ then they cover all those of Γ as well.

To prove it for general r it would suffice to settle the case $n=r+1$:

Conjecture. If the outer hyperplanes of a binary geometry G cover all G-points, then they cover all Γ-points as well.

The conjecture is proved for $r(G)=4$ and $r(G)=5$ in Lemmas 1 and 2. The case $r(G)=3$ is trivial. For $r(G)>5$ the proof (if it exists) seems to be hard.

References

[1] H. Crapo—G. C. Rota, *On the Foundations of Combinatorial Theory: Combinatorial Geometries*, M.I.T. Press (Cambridge, Mass., 1970).

[2] L. Kászonyi, On the structure of binary geometries having the $K(3)$-property, *Ann. Univ. Sci. Budapest, Sect. Math.*, to appear.

DEPARTMENT OF PROBABILITY THEORY
EÖTVÖS LORÁND UNIVERSITY
MÚZEUM KÖRÚT 6—8.
1088 BUDAPEST, HUNGARY

Acta Sci. Math., 41 (1979), 133—136

Multiplicative periodicity in rings

REYADH R. KHAZAL

A well known result of JACOBSON [4] establishes that a ring R is commutative if for every $a \in R$ there is an integer $n > 1$ (depending on a) such that $a = a^n$. This has been generalized by HERSTEIN [1]. On the other hand, ISKANDER [3] characterizes via polynomial identities varieties of rings in which every element generates a finite subring, while KRUSE [5] and L'VOV [6, 7] characterize via polynomial identities varieties generated by finite rings.

In the present paper we consider rings in which every element generates a finite multiplicative semigroup. It turns out that such rings are precisely the rings in which a power of every element generates a finite subring. A semigroup is called *periodic* if every element is of finite order. We call a ring R *periodic* if for every $a \in R$ there are a positive integer r and a polynomial $h(t)$ with integral coefficients such that $a^r + a^{r+1}h(a) = 0$. The term "periodic" has been used in literature for the case $r = 1$, (cf. OSBORN [8]). We will use the term periodic to mean also the case $r > 1$. The main result is:

Theorem 1. *The following statements about a ring R are equivalent:*
(i) *R is periodic;*
(ii) *if $a \in R$ then a power of a generates a finite subring;*
(iii) *the multiplicative semigroup of R is periodic.*

It is clear that (ii) implies (iii) and (iii) implies (i). Before we show that (i) implies (ii) we give some preliminaries.

Theorem 2. (HERSTEIN [2]) *If R is a ring with centre C such that for every $a \in R$ there exists a polynomial $p_a(t)$ such that $a^2 p_a(a) - a \in C$, then R is commutative.*

Proposition 3. *If R is a periodic division ring then R is a field. Also R is an algebraic extension of $\mathbf{Z}_p$ (the integers modulo p) for some prime p.*

Proof. Let $a \in R$. As R is periodic, there are $r > 0$ and a polynomial $h(t)$ such that $a^r + a^{r+1}h(a) = 0$. Thus $a^{r-1}(a + a^2 h(a)) = 0$. But R is a division ring, hence $a + a^2 h(a) = 0$. Thus, by Herstein's Theorem 2, R is commutative. Since

Received November 4, 1977.

$\mathbf{Z}$, the ring of integers, is not periodic $(2+2^2h(2)=0$ is impossible), the prime field of R is $\mathbf{Z}_p$ for some prime p and so R is an algebraic extension of $\mathbf{Z}_p$.

Proposition 4. *Let R be a primitive ring. If R is periodic then R is isomorphic to a dense ring of algebraic linear transformations of a vector space V over a field F that is an algebraic extension of some prime field $\mathbf{Z}_p$.*

Proof. By Jacobson's Density Theorem R is isomorphic to a dense ring of linear transformations of a vector space V over a division ring D. However D is a homomorphic image of a subring of R. Hence D is periodic and thus D is a fieldt which is also an algebraic extension of $\mathbf{Z}_p$. In this case periodicity implies tha, the linear transformations involved are algebraic over $\mathbf{Z}_p$.

Proposition 5. *Let R be a periodic ring. Then*

(i) *$J(R)$ (the Jacobson radical of R) is nil;*

(ii) *$R/J(R)$ is isomorphic to a subdirect sum of dense rings of algebraic linear transformations of vector spaces over fields each of which is an algebraic extension of $\mathbf{Z}_p$ for some prime p.*

Proof. Statement (ii) follows from Proposition 4 and Jacobson's Structure Theorem [2, 4]. Let $a\in J(R)$. Then $a^r+a^{r+1}h(a)=0$ for some positive integer r and $h(t)\in\mathbf{Z}[t]$. Hence, $a^r=-a^{r+1}h(a)=a^{r+1}g(a)=a^{r+2}g(a)^2=a^{2r}g(a)^r$, and $(ag(a))^r$ is an idempotent. Hence $a^rg(a)^r=0$, as the only idempotent in $J(R)$ is 0. Hence $a^r=a^ra^rg(a)^r=0$ and $J(R)$ is nil.

The converse of Proposition 5 is not true. The ring of integers $\mathbf{Z}$ is a subdirect sum of $\mathbf{Z}_p$ for all primes p and $\mathbf{Z}$ is not periodic.

Proposition 6. *The following conditions on a ring R are equivalent:*

(i) *R is periodic;*

(ii) *every subring of R generated by one element is an extension of a nilpotent ring by a finite direct sum of finite fields;*

(iii) *every subring of R generated by one element is an extension of a nil ring by a finite ring;*

(iv) *for every $a\in R$ there are integers $s,t>1$ such that $(a-a^s)^t=0$.*

Proof. It is obvious that (ii) implies (iii) and (iv) implies (i). Let A be the subring of R generated by $a\in A$. Then every ideal of A is finitely generated as A is commutative and is generated by one element. If R is periodic then $J(A)$ is nil (by Proposition 5) and hence nilpotent. $A/J(A)$ is isomorphic to a subdirect sum of periodic primitive rings generated by one element. Thus $A/J(A)$ is isomorphic to a subdirect sum of finite fields $F(i)$. $F(i)$ is generated by one element a_i. Also $a_i^r+a_i^{r+1}h(a_i)=0$. But $a_i^{r-1}=0$ is impossible in $F(i)$, so $a_i+a_i^2h(a_i)=0$. Hence $e_i=-a_ih(a_i)$ is idempotent $\neq 0$ and it is the identity element of $F(i)$. Thus $\bar{a}=a+J(A)$ satisfies $\bar{a}+\bar{a}^2h(\bar{a})=0$ in $A/J(A)$ and $e=-\bar{a}h(\bar{a})$ is the identity element of $A/J(A)$. Thus $A/J(A)$ is isomorphic to a finite direct sum of finite fields. This establishes that (i) implies (ii).

Let N be a nil ideal in A such that $A/N=F$ is finite. Hence F is periodic and is generated by one element. By (ii), $J(F)$ is nilpotent and $F/J(F)\cong F(1)\oplus\ldots\oplus F(k)$, where $F(i)$ is a finite field of characteristic p_i, $1\leqq i\leqq k$. Thus there is $s>1$ such that $F/J(F)$ satisfies $x-x^s=0$. Thus $\bar{a}=a+N$ satisfies $\bar{a}-\bar{a}^s\in J(F)$. As $J(F)$ is nilpotent, there is a positive integer r such that $(\bar{a}-\bar{a}^s)^r=0$, i.e. $(a-a^s)^r\in N$. Thus for some $t>0$, $(a-a^s)^{rt}=0$. This establishes that (iii) implies (iv) and concludes the proof of Proposition 6.

Now, we conclude the proof of Theorem 1. By Statement (ii) of Proposition 6, if $a\in R$ then $J(A)$ is nilpotent and $A/J(A)\cong F(1)\oplus\ldots\oplus F(k)$ where $F(i)$ is a finite field of characteristic p_i, $1\leqq i\leqq k$. Thus $ma\in J(A)$, where $m=\text{l.c.m.}(p_1,\ldots,p_k)$. Hence $(ma)^r=0$ for some $r>0$. Thus for every $a\in R$ some power a^r is torsion in the additive group of R. By (iv) of Proposition 6, $(b-b^s)^t=0$, $b=a^r$. b^{st} is a polynomial of degree less than st in b, and $nb=0$ for some $n>0$. In the subring B of R generated by b, every element has an expression in the form $\sum\{s_ib^i\colon 1\leqq i\leqq st, 0\leqq s_i<n\}$. Hence B is finite, it has at most n^{st-1} elements. Thus Statement (i) of Theorem 1 implies Statement (ii). This concludes the proof of Theorem 1.

If R is a periodic ring and $a\in R$, we define: Index $(a)=\inf\{r\colon r>0,\ a^r+a^{r+1}h(a)=0,\ h(t)\in\mathbf{Z}[t]\}$, Index $(R)=\sup\{\text{Index}(a)\colon a\in R\}$, $N(R)=\sup\{n\colon n>0$, for some $a\in R$, a is nilpotent, $a^n=0$ and $a^{n-1}\neq 0\}$. Degree $(a)=\inf\{\deg h(a)\colon a^r+a^{r+1}h(a)=0,\ r>0,\ h(t)\in\mathbf{Z}[t]\}$. Degree $(R)=\sup\{\text{Degree}(a)\colon a\in R\}$.

It turns out that

Proposition 7. *If R is a periodic ring then* $N(R)=\text{Index}(R)$.

Proof. Clearly, $N(R)\leqq\text{Index}(R)$. If $a\in R$ then by Proposition 6 (iv), $(a-a^s)^r=0$. One can assume that $r\leqq N(R)$. But $\text{Index}(a)\leqq r\leqq N(R)$. Hence $\text{Index}(R)\leqq N(R)$.

We conclude this paper by establishing some properties of periodic rings of bounded Index or Degree.

Proposition 8. *Let F be a periodic field. Then* Degree $(F)=d$ *iff* $F\cong GF(p,d+1)$ *(where $GF(p,t)$ is the Galois field of p^t elements).*

Proof. If F is periodic and Degree $(F)=d$, then F is an algebraic extension of $\mathbf{Z}_p$ for some prime p; furthermore, for any $a\in F$, there is $h(t)\in\mathbf{Z}[t]$ such that $a+a^2h(a)=0$ and $\deg h(t)\leqq d$, on the other hand, there is $b\in F$ such that Degree $(b)=d$.

Now $[\mathbf{Z}_p(b)\colon\mathbf{Z}_p]=d+1=$ the degree of the minimal polynomial of b over $\mathbf{Z}_p$. In fact $F=\mathbf{Z}_p(b)$. It is obvious that F contains $\mathbf{Z}_p(b)$. Let $a\in F$. If $a\notin\mathbf{Z}_p(b)$ then $(\mathbf{Z}_p(b))(a)\neq\mathbf{Z}_p(b)$. Now a being algebraic over $\mathbf{Z}_p$, $H=(\mathbf{Z}_p(b))(a)$ is a finite subfield of F and $[H:\mathbf{Z}_p]=n>d+1$. The field H is generated by one element c whose minimal polynomial over $\mathbf{Z}_p$ is of degree n. Thus Degree $(c)=n-1>d$, which is impossible. Therefore $F=\mathbf{Z}_p(b)$. Conversely, since F is a finite field of p^{d+1} elements, F is periodic. Now any $0\neq a\in F$ is algebraic over $\mathbf{Z}_p$ and $[\mathbf{Z}_p(a):\mathbf{Z}_p]=$ $=k\leqq d+1$. Thus the minimal polynomial of a is of degree at most $d+1$ and so Degree $(a)\leqq d$. Also F is generated by an element b such that Degree $(b)=d$.

Thus from Propositions 5 and 8 it follows that a periodic ring R whose Degree is d is such that $J(R)$ is nil and $R/J(R)$ is isomorphic to a subdirect sum of dense rings of algebraic linear transformations of vector spaces over $GF(p, k)$ with $k\leqq d+1$ for some primes p.

Proposition 9. *R is a periodic primitive ring and* Index $(R)=n$ *iff R is isomorphic to F_n (the ring of $n\times n$ matrices over F) for some algebraic extension F of $\mathbf{Z}_p$ for some prime p.*

Proof. Let F be an algebraic extension of $\mathbf{Z}_p$. If $A\in F_n$ then the matrix A has n^2 entries and involves only a finite number of elements of F. Thus $A\in G_n$ where G is a finite subfield of F, i.e. A belongs to a finite subring of F_n. By Theorem 1, F_n is periodic. It is well known that F_n is primitive. Since the minimal polynomial of $A\in F_n$ is of degree at most n, $N(F_n)\leqq n$. Also $A=[a_{ij}]$, $a_{ij}=1$ if $i<j$ and $a_{ij}=0$ if $i\geqq j$, satisfies $A^n=0$ and $A^{n-1}\neq 0$. Thus $N(F_n)=n$, and by Proposition 7, Index $(F_n)=N(F_n)=n$. Conversely, let R be a periodic primitive ring and Index $(R)=n$. Then $R\cong F_m$ or F_s is a homomorphic image of a subring of R, for every positive integer s, where F is an algebraic extension of $\mathbf{Z}_p$ for some p. Now, Index (R) does not increase by taking subrings or homomorphic images and so $s=$ Index $(F_s)\leqq$ Index $(R)=n$. Thus $R\cong F_n$.

References

[1] I. N. Herstein, The structure of a certain class of rings, *Amer. J. Math.*, **75** (1953), 864—871.

[2] I. N. Herstein, *Non-Commutative Rings*, Carus Monograph Series, Math. Association of America (Menasha, Wis., 1968).

[3] A. A. Iskander, Locally finite ring varieties, *Proc. Amer. Math. Soc.*, **50** (1975), 28—32.

[4] N. Jacobson, *Structure of Rings*, Amer. Math. Soc. (Providence, 1964).

[5] R. L. Kruse, Identities satisfied by a finite ring, *J. Algebra*, **26** (1973), 298—318.

[6] I. V. L'vov, Varieties of associative rings. I, *Algebra i Logika*, **12** (1973), 269—297; English translation: pp. 150—167.

[7] I. V. L'vov, Varieties of associative rings. II, *Algebra i Logika*, **12** (1973), 667—688; English translation: pp. 381—393.

[8] J. M. Osborn, Varieties of algebras, *Advances in Math.*, **8** (1972), 163—369.

DEPARTMENT OF MATHEMATICS
OUWAIT UNIVERSITY
P0Box 5486
KUWAIT

Acta Sci. Math., 41 (1979), 137—150

Sublattices of a distributive lattice

VÁCLAV KOUBEK

At the Mini-Conference on Lattice Theory in Szeged, 1974, M. SEKANINA has formulated the following problem: Is it true that if a lattice B contains an arbitrarily large finite number of pairwise disjoint sublattices, isomorphic to a lattice A, then B also contains an infinite number of such sublattices? The aim of the present paper is to construct two countable distributive lattices A and B which are counter-examples, i.e. such that for any $m=1, 2, 3, \ldots$, B contains m disjoint copies of A, but it does not contain infinitely many such copies. An independent solution of Sekanina's problem was found by I. KOREC in a paper to appear (personal communication).

An analogous problem can be formulated for other structures than lattices and various concepts of subobject, e.g. summand. In the second part a general formulation of this problem is exhibited.

1. We recall that a graph (X, R) (i.e. $R \subset X \times X$) is bipartite if it is symmetric and there exists a subset M of X such that if $(x, y) \in R$ then $x \in M$ iff $y \notin M$.

Definition. A graph (X, R) is *strongly reduced* if for any distinct points $x, y \in X$ there exists at most one point z with $(z, x), (z, y) \in R$.

Convention. Denote by $\mathbf{N}$ the set of all natural numbers, by $\mathbf{Z}$ the set of all integers.

Construction 1.1. We shall construct countable, connected, strongly reduced, bipartite graphs (X_i, R_i) with $i \in \mathbf{N}$, $i > 1$ such that

a) for every $x \in X_i$, card $\{z: (x, z) \in R_i\} \in \{2, 3\}$;

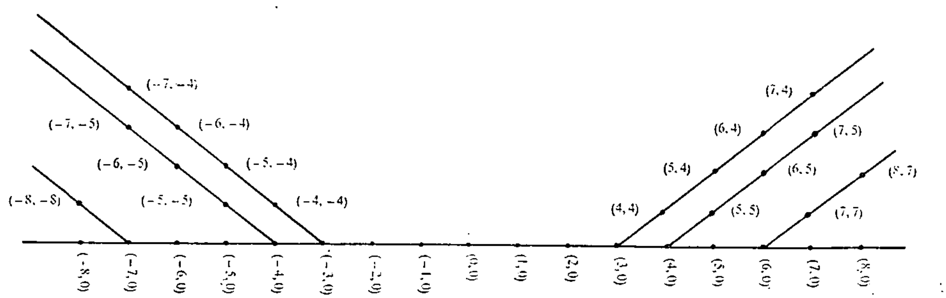
b) if $f: (X_i, R_i) \to (X_j, R_j)$ is a one-to-one compatible mapping then $i=j$ and f is the identity.

Received October 3, 1977.

Put

$$X_i = \{(x, y): (x, y \in \mathbf{Z}),\ (y \neq 0 \Rightarrow y \in \{i, -i\} \cup \\ \cup \{i+2k+1:\ k \in \mathbf{N}\} \cup \{-i-3k-1:\ k \in \mathbf{N}\})\ (\operatorname{sgn} x = \operatorname{sgn} y)\},$$

$$R_i = \{((x, 0), (x+1, 0)):\ x \in \mathbf{Z}\} \cup \{((x, 0), (x-1, 0)):\ x \in \mathbf{Z}\} \cup \\ \cup \{((y+s, y), (y+t, y)):\ (y \in \{i+2k+1:\ k \in \mathbf{N}\} \cup \\ \cup \{-i-3k-1:\ k \in \mathbf{N}\} \cup \{i, -i\}), (|s-t| = 1), (sy, ty \geqq 0)\} \cup \\ \cup \{((y - \frac{y}{|y|}, 0), (y, y)), ((y, y), (y - \frac{y}{|y|}, 0)):\ y \in \{i+2k+1:\ k \in \mathbf{N}\} \cup \\ \cup \{-i-3k-1:\ k \in \mathbf{N}\} \cup \{i, -i\}\}.$$



$(X_4,\ R_4)$

Fig. 1

It is clear that (X_i, R_i) is a countable, symmetric, strongly reduced graph. Set $M_i = \{(x, y) \in X_i: x \text{ is even}\}$, then $R_i \subset ((X_i - M_i) \times M_i) \cup (M_i \times (X_i - M_i))$ and therefore (X_i, R_i) is a bipartite graph. Further, for every $x \in X_i$,

$$\operatorname{card} \{z:\ (x, z) \in R_i\} \in \{2, 3\}.$$

We shall prove Property b). If $f:\ (X_i, R_i) \to (X_j, R_j)$ is a one-to-one compatible mapping then for $x \in \{i-1, 1-i\} \cup \{i+2k: k \in \mathbf{N}\} \cup \{-i-3k: k \in \mathbf{N}\}$, $f(x, 0) \in \{(j-1, 0), (1-j, 0)\} \cup \{(j+2k, 0): k \in \mathbf{N}\} \cup \{(-j-3k, 0): k \in \mathbf{N}\}$. Hence $f(\{(i-1, 0), (1-i, 0)\}) \in \in \{(j-1, 0), (1-j, 0)\}$ and therefore $i=j$. Further, $f(x, 0) \in \{(y, 0): y \in \mathbf{Z}\}$. If $f(i-1, 0) = (1-i, 0)$ then $f(i+2k, 0) = (-i-2k, 0)$ but the latter is impossible, thus $f(i-1, 0) = (i-1, 0)$ and so is $f(x, 0) = (x, 0)$ for every $x \in \mathbf{Z}$. Hence f is the identity.

Let us introduce the notation $\mathfrak{X}_i = (X_i, R_i, M_i)$, $i \in \mathbf{N}$, $i > 1$.

Construction 1.2. Let $\mathfrak{X}=(X, R, M)$ where (X, R) is a bipartite graph and $M\subset X$ such that if $(x, y)\in R$ then $x\in M$ iff $y\notin M$. Set

$A_1^{\mathfrak{X}} = \{Z\subset X: (\exists (x, y)\in R)(x\in M \text{ and } Z = (M-\{x\})\cup\{y\})\}$;

$A_2^{\mathfrak{X}} = \{Z\subset X: (\exists x\in M)(\exists K\subset \{y: (y, x)\in R\})(K \text{ is finite and } Z = (M-\{x\})\cup K)\}$;

$A_3^{\mathfrak{X}} = \{Z\subset X: (\exists x\in X-M)(\exists K\subset \{y: (y, x)\in R\})(K \text{ is finite and } Z=(M-K)\cup\{x\})\}$;

$A_4^{\mathfrak{X}} = \{Z\subset X: (\exists K\subset M)(K \text{ is finite and } Z = M-K)\}$;

$A_5^{\mathfrak{X}} = \{Z\subset X: (\exists K\subset X-M)(K \text{ is finite and } Z = M\cup K)\}$.

Put $A^{\mathfrak{X}}=\bigcup_{i=1}^{5} A_i^{\mathfrak{X}}$, $B^{\mathfrak{X}}=A^{\mathfrak{X}}\cup\{\emptyset, X\}$. For $Z, V\in B^{\mathfrak{X}}$ define $Z\vee V=Z\cup V$, $Z\wedge V= =Z\cap V$, then it is easy to verify that $(A^{\mathfrak{X}}, \cup, \cap)$ and $(B^{\mathfrak{X}}, \cup, \cap)$ are lattices (and hence they are distributive). Moreover, $A^{\mathfrak{X}}$ and $B^{\mathfrak{X}}$ are countable iff X is countable.

Let $\mathfrak{X}=(X, R, M)$, $\mathfrak{Y}=(Y, S, N)$ where (X, R), (Y, S) are bipartite graphs and for $(x, y)\in R$ (or $(x, y)\in S$), $x\in M$ iff $y\notin M$ (or $x\in N$ iff $y\notin N$, respectively). If $f: X\to Y$ such that $f(M)\subset N$ and $f: (X, R)\to(Y, S)$ is a one-to-one compatible mapping then $\varphi: B^{\mathfrak{X}}\to B^{\mathfrak{Y}}$ (or $\varphi/A^{\mathfrak{X}}: A^{\mathfrak{X}}\to A^{\mathfrak{Y}}$) is a one-to-one lattice homomorphism, where $\varphi(Z)=(f(Z)\cup N)-f(M-Z)$ if $Z\neq\emptyset, X$, $\varphi(\emptyset)=\emptyset$, $\varphi(X)=Y$. We shall write $\Psi\mathfrak{X}=(A^{\mathfrak{X}}, \cup, \cap)$, $\Psi f=\varphi/A^{\mathfrak{X}}$, $\Phi\mathfrak{X}=(B^{\mathfrak{X}}, \cup, \cap)$, $\Phi f=\varphi$.

Note 1.3. Denote by **Gr** the category whose objects are triples (X, R, M) where (X, R) is a bipartite graph and $M\subset X$ such that if $(x, y)\in R$ then $x\in M$ iff $y\notin M$ and whose morphisms $f: (X, R, M)\to(Y, N, S)$ are one-to-one mappings $f: (X, R)\to(Y, S)$ with $f(M)\subset N$. Denote by **DLat** the category of distributive lattices and one-to-one lattice homomorphisms. Then Φ, Ψ are faithful functors from **Gr** to **DLat**.

Definition. Let $\mathfrak{A}$ be a lattice. An element x of $\mathfrak{A}$ is called *meet-infinite* (or *join-infinite*) if there exists an infinite subset B of $\mathfrak{A}$ such that for any distinct points $a, b\in B$, $a\wedge b=x$ (or $a\vee b=x$, respectively).

Lemma 1.4. *Let $\mathfrak{X}=(X, R, M)$ be an object of* **Gr** *such that M and $X-M$ are infinite and for every $x\in X$ the set $\{y: (x, y)\in R\}$ is finite. Then for $Z\in A^{\mathfrak{X}}$ we have*

a) *Z is a meet-infinite element iff $Z\supset M$;*

b) *Z is a join-infinite element iff $Z\subset M$.*

Proof. If $V\supset M$ then it is clear that V is meet-infinite $(V=(V\cup\{x\})\cap(V\cup\{y\})$ for every $x\neq y$, $x, y\in X-V)$. Let V be meet-infinite. Let $\mathscr{B}\subset A^{\mathfrak{X}}$ be an infinite set with $W_1\cap W_2=V$ for every $W_1\neq W_2$, $W_1, W_2\in\mathscr{B}$. If $M-V\neq\emptyset$ then $M-W\neq \neq M-V$ holds only for finitely many $W\in\mathscr{B}$, and so $\mathscr{B}$ is finite because the set $\{y: (x, y)\in R\}$ is finite for every $x\in X$, a contradiction; thus $M-V=\emptyset$ and hence $V\supset M$. The proof of case b) is analogous.

Lemma 1.5. *Let* $f\colon \mathfrak{A}\to\mathfrak{B}$ *be a one-to-one lattice homomorphism. If* $a\in\mathfrak{A}$ *is a meet-infinite (join-infinite) element then* $f(a)$ *is meet-infinite (join-infinite), too.*

Proof. The proof is easy and is therefore omitted.

Lemma 1.6. *Let* $\mathfrak{X}=(X, R, M)$ *be an object of* **Gr** *such that for every* $x\in X$ *the set* $\{y\colon (x,y)\in R\}$ *is finite. Let* $Z, V\in B^{\mathfrak{X}}$ *be such that there exists an infinite set* $\mathscr{B}\subset B^{\mathfrak{X}}$ *with the following properties:* 1) *for every* $W_1, W_2\in\mathscr{B}$, $W_1\cap W_2=V$ *(or* $W_1\cup W_2=V$*);* 2) $Z\supset W$ *(or* $Z\subset W$*) for every* $W\in\mathscr{B}$*. Then* $Z=X$ *(or* $Z=\emptyset$*, respectively).*

Proof. Clearly, X is finite iff $B^{\mathfrak{X}}$ is finite. If the set $\{y\colon (x,y)\in R\}$ is finite for every $x\in X$ then X is finite iff M and $X-M$ are finite. By Lemma 1.4 we get that $V\in A_5^{\mathfrak{X}}$ and therefore either $Z=X$ or $Z\in A_5^{\mathfrak{X}}$. If $Z\in A_5^{\mathfrak{X}}$, we have that $Z-V$ is finite and therefore $\mathscr{B}$ is not infinite, a contradiction.

Proposition 1.7. *Let* $\mathfrak{X}=(X, R, M)$, $\mathfrak{Y}=(Y, S, N)$ *be objects of* **Gr** *such that*

a) (X, R), (Y, S) *are strongly reduced;*

b) *for every* $x\in X$ *the set* $\{y\colon (y,x)\in R\}$ *is finite and has at least two points;*

c) $M, X-M, N, Y-N$ *are infinite.*

If $f\colon \Psi\mathfrak{X}\to\Psi\mathfrak{Y}$ *(or* $f\colon \Phi\mathfrak{X}\to\Phi\mathfrak{Y}$*) is a one-to-one lattice homomorphism then there exists a morphism* $g\colon (X,R,M)\to(Y,S,N)$ *of* **Gr** *with* $\Psi g=f$ *(or* $\Phi g=f$*, respectively).*

Proof. By Lemmas 1.4 and 1.5, $f(A_5^{\mathfrak{X}})\subset A_5^{\mathfrak{Y}}$, $f(A_4^{\mathfrak{X}})\subset A_4^{\mathfrak{Y}}$. Now we shall prove $f(A_1^{\mathfrak{X}})\subset A_1^{\mathfrak{Y}}$. Since for every $Z\in A_1^{\mathfrak{X}}$, $Z-M$ and $M-Z$ are nonempty, we get that $f(Z)\in A_1^{\mathfrak{Y}}\cup A_2^{\mathfrak{Y}}\cup A_3^{\mathfrak{Y}}$, hence $f(A_1^{\mathfrak{X}})\subset A_1^{\mathfrak{Y}}\cup A_2^{\mathfrak{Y}}\cup A_3^{\mathfrak{Y}}$. Assume that there exists $Z\in A_1^{\mathfrak{X}}$ with $f(Z)\in A_2^{\mathfrak{Y}}$. Then there exists $V_1\in A_1^{\mathfrak{X}}$ with $V_1\cup Z\in A_5^{\mathfrak{X}}$ and $V_1\cap Z\notin A_4^{\mathfrak{X}}$. Then $f(V_1)\cup f(Z)\in A_5^{\mathfrak{Y}}$ and $f(V_1)\cap f(Z)\notin A_4^{\mathfrak{X}}$. Therefore $(f(V_1)-N)\cap\cap(f(Z)-N)\neq\emptyset$ but $(N-f(V_1))\cap(N-f(Z))=\emptyset$. We shall prove $f(V_1)-N==f(Z)-N$, hence we get a contradiction because (Y, S) is strongly reduced. Choose $V_2\in A_1^{\mathfrak{X}}$ with $V_2\cup Z, V_2\cup V_1\in A_5^{\mathfrak{X}}$, $V_2\cap Z$, $V_2\cap V_1\in A_4^{\mathfrak{X}}$. Then $V_2\cup Z==V_2\cup V_1=Z\cup V_2\cup V_1$ (we use that $V_1\cap Z\notin A_4^{\mathfrak{X}}$ and therefore $V_1-M=Z-M$). Then $f(V_2)\cup f(Z)=f(V_2)\cup f(V_1)=f(Z)\cup f(V_2)\cup f(V_1)$, hence $(f(V_2)\cup f(Z))--N=(f(V_2)\cup f(V_1))-N$. Since $f(V_1)\cap f(V_2)$, $f(Z)\cap f(V_2)\in A_4^{\mathfrak{Y}}$, we have $(f(V_1)-N)\cap(f(V_2)-N)=\emptyset$ and $(f(Z)-N)\cap(f(V_2)-N)=\emptyset$. Thus $f(Z)-N==f(V_1)-N$. We obtain that $f(A_1^{\mathfrak{X}})\subset A_1^{\mathfrak{Y}}$ because it can be proved analogously that $f(A_1^{\mathfrak{X}})\cap A_3^{\mathfrak{Y}}=\emptyset$. Hence $f(A_2^{\mathfrak{X}})\subset A_2^{\mathfrak{Y}}$, $f(A_3^{\mathfrak{X}})\subset A_3^{\mathfrak{Y}}$. Define $g\colon X\to Y$ as follows:

for $x\in M$, $g(x)=y$ where $f(M-\{x\})=N-\{y\}$,

for $x\notin M$, $g(x)=y$ where $f(M\cup\{x\})=N\cup\{y\}$.

(Since $f(A_1^{\mathfrak{X}})\subset A_1^{\mathfrak{Y}}$, we get that for every $v\in M$, $f(M-\{v\})=N-\{w\}$ where $w\in N$ and for every $v\in X-M$, $f(M\cup\{v\})=N\cup\{w\}$ where $w\in Y-N$.) It is clear that $g(M)\subset N$ and g is one-to-one. If $(x, y)\in R$ with $x\in M$ then $Z=(M-\{x\})\cup\{y\}\in A_1^{\mathfrak{X}}$ and therefore $f(Z)\in A_1^{\mathfrak{Y}}$. Since $Z\supset M-\{x\}$, we get that $f(Z)\supset N-\{g(x)\}$ and since $Z\subset M\cup\{y\}$, we get that $f(Z)\subset N\cup\{g(y)\}$. Hence $f(Z)=(N-\{g(x)\})\cup$ $\cup\{g(y)\}$ and so $(g(x), g(y))\in S$. It is clear that $\Psi g=f$. If $f\colon \Phi\mathfrak{X}\to\Phi\mathfrak{Y}$ then by Lemmas 1.4 and 1.5 $f(A_5^{\mathfrak{X}})\subset A_5^{\mathfrak{Y}}$, $f(A_4^{\mathfrak{X}})\subset A_4^{\mathfrak{Y}}$. Therefore by Lemma 1.6 $f(\emptyset)=\emptyset$, $f(X)=Y$ and the rest follows from the foregoing part of the proof.

Corollary 1.8. *Put $\mathfrak{A}_i=\Psi\mathfrak{X}_i$, $\mathfrak{B}_i=\Phi\mathfrak{X}_i$ (for $\mathfrak{X}_i$, see Construction 1.1). If $f\colon \mathfrak{A}_i\to\mathfrak{A}_j$ (or $f\colon \mathfrak{B}_i\to\mathfrak{B}_j$) is a one-to-one lattice homomorphism then $i=j$ and f is the identity.*

Construction 1.9. Let T be a set. Put

$$Y=\{Z\colon (Z\subset\exp T), (Z\neq\emptyset), (Z \text{ is finite}), (V\in Z\Rightarrow (V\neq\emptyset \text{ and } V \text{ or } T-V \text{ is finite})), (\forall V_1, V_2\in Z)(V_1-V_2\neq\emptyset)\}.$$

Define a partial ordering $\leqq$ on Y as follows: $Z_1\leqq Z_2$ iff for every $V\in Z_1$ there exists $W\in Z_2$ with $V\supset W$. Clearly, $\leqq$ is a reflexive and transitive relation. Since for every $Z\in Y$, $V_1, V_2\in Z$ implies $V_1-V_2\neq\emptyset$, we get that $Z_1\leqq Z_2\leqq Z_1$ iff $Z_1=Z_2$; thus $\leqq$ is a partial ordering.

Now if we put

$$Z_1\vee Z_2=\{V\in Z_1\cup Z_2\colon (W\in Z_1\cup Z_2\Rightarrow W-V\neq\emptyset \text{ or } W=V)\};$$

$$Z_1\wedge Z_2=\{V\colon (\exists V_1\in Z_1)(\exists V_2\in Z_2)(V=V_1\cup V_2), (\forall W_1\in Z_1)(\forall W_2\in Z_2)((W_1\cup W_2)\subset V\Rightarrow W_1\cup W_2=V)\},$$

we get that $(Y, \leqq)$ is a partial ordering induced by a lattice $(Y, \vee, \wedge)$ and it is easy to verify that $(Y, \wedge, \vee)$ is a distributive lattice. Put $\mathscr{D}(T)=(Y, \vee, \wedge)$. We shall identify $t\in T$ with $\{\{t\}\}\in Y$, i.e. $T\subset Y$. It is clear that the sublattice of $\mathscr{D}(T)$ generated by T is a free distributive lattice over T. Furthermore, no element Z of $\mathscr{D}(T)$ is join-infinite and $Z\in Y$ is meet-infinite iff there exists an infinite set $V\subset T$ with $V\in Z$.

Let U be a set and let $\{U_{i,j}\colon i, j\in\mathbf{N}\}$ be a cover of U. Define

$$\bar{Y}=\{Z\subset\exp U\colon (Z \text{ is finite}), (V\in Z\Rightarrow (V\neq\emptyset), (V \text{ is finite or } (\exists i, j\in\mathbf{N})(U_{i,j}-V \text{ is finite}))), (\forall V_1, V_2\in Z)(V_1-V_2\neq\emptyset)\}.$$

Define a partial ordering $\leqq$ on $\bar{Y}$ as follows: $Z_1\leqq Z_2$ iff for every $V\in Z_1$ there exists $W \in Z_2$ with $V \supset W$. Clearly, $\leqq$ is a partial ordering and if we put

$$Z_1 \vee Z_2 = \{V \in Z_1 \cup Z_2 \colon (\forall W \in Z_1 \cup Z_2)(W \subset V \Rightarrow W = V)\};$$

$$Z_1 \wedge Z_2 = \{V \colon (\exists V_1 \in Z_1)(\exists V_2 \in Z_2)(V = V_1 \cup V_2), (\forall W_1 \in Z_1), (\forall W_2 \in Z_2)((W_1 \cup W_2) \subset V \Rightarrow W_1 \cup W_2 = V)\}$$

then $(\bar{Y}, \vee, \wedge)$ is a distributive lattice induced by the ordering $\leqq$. Put

$$\bar{\bar{Y}} = \{Z \in \bar{Y} \colon V \in Z \Rightarrow (V \text{ is infinite}, (\exists i, j, m, n \in \mathbf{N}) ((i, j) \neq (m, n), V - U_{i,j} \neq \emptyset, V - U_{m,n} \neq \emptyset))\},$$

then $\bar{\bar{Y}}$ is an ideal in $\bar{Y}$. Let $\sim$ be the congruence relation generated by $\bar{\bar{Y}}$. Then $Z_1 \sim Z_2$ iff $V \in \bar{\bar{Y}}$ whenever $V \in (Z_1 - Z_2) \cup (Z_2 - Z_1)$. Hence if we put

$$\tilde{Y} = \{Z \in \bar{Y} \colon V \in Z \Rightarrow (V \text{ is finite or } (\exists i, j \in \mathbf{N})(U_{i,j} - V \text{ is finite}, V \subset U_{i,j}))\},$$

we get that $(\tilde{Y}, \leqq)$ induces operations sup and inf as follows: $\sup\{Z_1, Z_2\} = Z_1 \vee Z_2$, $\inf\{Z_1, Z_2\} = Z_1 \wedge Z_2$ if $Z_1 \wedge Z_2 \in \tilde{Y}$, $= \emptyset$ otherwise. Clearly, $(\tilde{Y}, \sup, \inf)$ is a lattice. Since $(\tilde{Y}, \sup, \inf)$ is isomorphic to $(\bar{Y}, \vee, \wedge)/\sim$, we get that it is distributive. We shall identify $u \in U$ with $\{\{u\}\} \in \tilde{Y}$, i.e. $U \subset \tilde{Y}$. Notice that the sublattice of $(\tilde{Y}, \sup, \inf)$ generated by U is a free distributive lattice over U. Introduce the notation $\mathscr{C}(U, U_{i,j} \colon i, j \in \mathbf{N}) = (\tilde{Y}, \sup, \inf)$ (further on we shall write only $\vee$, $\wedge$ instead of sup, inf).

Lemma 1.10. *For every cover $\{U_{i,j} \colon i, j \in \mathbf{N}\}$ of U no element of $\mathscr{C}(U, U_{i,j} \colon i, j \in \mathbf{N}\}$ is join-infinite. An element Z of $\mathscr{C}(U, U_{i,j} \colon i, j \in \mathbf{N})$ is meet-infinite iff*

a) *either $Z \neq \emptyset$ and there exists $V \in Z$ such that V is infinite,*

b) *or $Z = \emptyset$ and there exist infinitely many $i, j \in \mathbf{N}$ such that $U_{i,j}$ is infinite.*

Proof. Let $Z \in \tilde{Y}$, we prove that it is not join-infinite. Let $\mathscr{T}$ be a subset of $\tilde{Y}$ such that $Z_1 \vee Z_2 = Z$ for any distinct $Z_1, Z_2 \in \mathscr{T}$. Then $Z_1 \cup Z_2 \supset Z$ and for every $V \in (Z_1 \cup Z_2) - Z$ there exists $W \in Z$ with $V \supset W$. Hence, if $V \in Z - Z_i$ for $Z_i \in \mathscr{T}$ then $V \in Z_j$ for every $Z_j \in \mathscr{T} - \{Z_i\}$ and if $Z_i \supset Z$ where $Z_i \in \mathscr{T}$ then $Z_i = Z$. Therefore we get that $\mathscr{T}$ is finite and Z is not join-infinite.

Let $Z \in \tilde{Y}$, $Z \neq \emptyset$ be such that every $V \in Z$ is finite. We shall prove that Z is not meet-infinite. Let $\mathscr{T} \subset \tilde{Y}$ be such that $Z_1 \wedge Z_2 = Z$ for any distinct $Z_1, Z_2 \in \mathscr{T}$. Hence if $V \in Z$, $V_1 \in Z_1$ with $V \supset V_1$ then for every $W_2 \in Z_2$, $V \not\supset V_1 \cup W_2$ and there exists $V_2 \in Z_2$ with $V = V_1 \cup V_2$. On the other hand, for every $V \in Z$ there exists $V_1 \in Z_1$ with $V \supset V_1$. Now, for every $V \in Z$ and every $Z_i \in \mathscr{T}$ we choose $W_{V,i} \in Z_i$ with $W_{V,i} \subset V$. Then for $i \neq j$, $W_{V,i} \cup W_{V,j} = V$. Therefore for every $V \in Z$

the set $\{W_{V,i}: Z_i \in \mathscr{T}\}$ is finite and if $W_{V,i} \neq V$ then $W_{V,i} \neq W_{V,j}$ for every $Z_j \neq Z_i$, $Z_j \in \mathscr{T}$. Hence the set $\{Z_i \in \mathscr{T}: (\exists V \in Z)\ (W_{V,i} \neq V)\}$ is finite. Let $\mathscr{T}'$ be a subset of $\mathscr{T}$ with $Z_i \in \mathscr{T}'$ iff $Z_i \supset Z$. It suffices to prove that $\mathscr{T}'$ is finite. For any distinct $Z_1, Z_2 \in \mathscr{T}'$ and every $V_1 \in Z_1 - Z$, $V_2 \in Z_2 - Z$, there exists $V \in Z$ with $V_1 \cup V_2 \supset V$. For every $Z_i \in \mathscr{T}' - \{Z\}$, we choose $V_i \in Z_i - Z$ and put $W_i = V_i \cap \bigcap \bigcup_{V \in Z} V$. Now if $Z_i \neq Z_j$ then $W_i \cup W_j \subset V$ for some $V \in Z$. Since $\bigcup_{V \in Z} V$ is a finite set, we get that there exists only a finite set $\mathscr{T}'' \subset \mathscr{T}'$ such that if $Z_i \in \mathscr{T}''$ then $V - W_i \neq \emptyset$ for every $V \in Z$. Hence $\mathscr{T}'$ is finite because if $W_i \supset V$ for some $V \in Z$ then $W_i = V_i = V$, a contradiction (notice that $V \in Z_i$). Thus $\mathscr{T}$ is finite and Z is not meet infinite.

If there exists an infinite set $V \in Z$ then put $\mathscr{T} = \big\{\{W: W \in Z - \{V\}\} \cup \cup \{V - \{x\}\}: x \in V\big\}$. Clearly, if $Z_1, Z_2 \in \mathscr{T}$, $Z_1 \neq Z_2$ then $Z_1 \wedge Z_2 = Z$ and $\mathscr{T}$ is infinite since V is infinite.

If $Z = \emptyset$ and $\mathbf{M} = \{(i, j): U_{i,j}$ is infinite$\}$ is infinite, then put $\mathscr{T} = \big\{\{U_{i,j}\}: (i, j) \in \mathbf{M}\big\}$. Then $\mathscr{T}$ is infinite and for distinct $Z_1, Z_2 \in \mathscr{T}$, $Z_1 \wedge Z_2 = \emptyset = Z$.

Let $\mathscr{T}$ be an infinite subset of $\mathscr{C}(U, U_{i,j}: i, j \in \mathbf{N})$ such that for distinct $Z_1, Z_2 \in \mathscr{T}$, $Z_1 \wedge Z_2 \neq \emptyset$. Then for every $Z_i \in \mathscr{T} - \{\emptyset\}$ there exists an infinite set $V_i \in Z_i$ and if $Z_i \neq Z_j$ then $V_i \cup V_j$ is not a subset of any $U_{m,n}$, $m, n \in \mathbf{N}$, but every V_i is a subset of some U_{m_i,n_i}. Hence if $Z_i \neq Z_j$ then $(m_i, n_i) \neq (m_j, n_j)$ and U_{m_i,n_i} is infinite.

Construction 1.11. Choose countably infinite sets T and U and a covering $\{U_{i,j}: i, j \in \mathbf{N}\}$ of U such that $U_{i,j}$ is infinite, and if $i+j = m+n$ then $U_{i,j} \cap U_{m,n} = \emptyset$, otherwise the intersection is a singleton. Choose a mapping $\varepsilon: U \to T$ such that $\varepsilon | U_{i,j}: U_{i,j} \to T$ is a bijection for every $(i, j) \in \mathbf{N} \times \mathbf{N}$ and choose a bijection $\mu: \mathbf{N} \to T$. Set $\mathbf{K} = \left\{(p, q): (p \in \mathbf{N}), \left[q \in \varepsilon^{-1}\left(\mu\left(\left[\frac{p}{2}\right]\right)\right)\right]\right\}$.

Let $\mathfrak{M}$ be the sublattice of $\mathscr{D}(T) \times \prod_{i \in \mathbf{N}} \mathfrak{B}_i$ (for $\mathfrak{B}_i$ see Corollary 1.8) generated by the set

$$S = \left\{(t, \{a_i\}_{i \in \mathbf{N}}): (t \in T), \left((i \text{ is odd}), \left(\mu\left(\left[\frac{i}{2}\right]\right) \neq t\right) \Rightarrow a_i = \emptyset\right),\right.$$

$$\left.\left((i \text{ is even}), \left(\mu\left(\left[\frac{i}{2}\right]\right) \neq t\right) \Rightarrow a_i = X_{\left[\frac{i}{2}\right]}\right), \left(\mu\left(\left[\frac{i}{2}\right]\right) = t \Rightarrow a_i \in \mathfrak{B}_j\right)\right\} \cup$$

$$\cup \{(\{V\}, \{a_i\}_{i \in \mathbf{N}}): (T - V \text{ is finite}), (\forall i \in \mathbf{N})(a_i = \emptyset)\}.$$

It is clear that S is a countable set and therefore $\mathfrak{M}$ is a countable distributive lattice. For $t \in T$ set $\alpha(t) = (t, \{a_i\}_{i \in \mathbf{N}})$ where $\alpha(t) \in S$ and $a_i = M_i$ if $\mu\left(\left[\frac{i}{2}\right]\right) = t$.

Let $\mathfrak{N}$ be the sublattice of $\mathscr{C}(U, U_{i,j}: i, j\in\mathbf{N})\times\prod_{i\in\mathbf{N}}\mathfrak{B}_i^{s_i}$ with $s_i = \operatorname{card}\varepsilon^{-1}\left(\mu\left(\left[\frac{i}{2}\right]\right)\right)$ generated by the set

$$Q=\Big\{\big(u, \{a_{p,q}\}_{(p,q)\in\mathbf{K}}\big): (u\in U), \big((p \text{ is odd}), (q\neq u)\Rightarrow$$
$$\Rightarrow a_{p,q}=\emptyset\big), \big((p \text{ is even}), (q\neq u)\Rightarrow a_{p,q}=X_{\left[\frac{p}{2}\right]}\big),$$
$$\big(q=u\Rightarrow a_{p,q}\in\mathfrak{B}_{\left[\frac{p}{2}\right]}\big)\Big\}\cup\Big\{\big(\{V\}, \{a_{p,q}\}_{(p,q)\in\mathbf{K}}\big):$$
$$(\exists i, j\in\mathbf{N})\big((U_{i,j}-V \text{ is finite}), (V\subset U_{i,j})\big),$$
$$\big(\forall(p,q)\in\mathbf{K}(a_{p,q}=\emptyset)\big)\Big\}.$$

Since Q is a countable set, $\mathfrak{N}$ is a distributive lattice. For $u\in U$, put $\beta(u)=(u, \{a_{p,q}\}_{(p,q)\in\mathbf{K}})\in Q$ where $a_{p,q}=M_{\left[\frac{p}{2}\right]}$ if $q=u$.

Lemma 1.12. *Let $(t, \{a_i\}_{i\in\mathbf{N}})$ be a point where $t\in\mathscr{D}(T)$ and $a_i\in\mathfrak{B}_{[i/2]}$ for every $i\in\mathbf{N}$. Then $(t, \{a_i\}_{i\in\mathbf{N}})$ is a point of $\mathfrak{M}$ iff the following conditions hold:*

a) *if i is odd and $a_i\neq\emptyset$ then t is greater than or equal to $\mu\left(\left[\frac{i}{2}\right]\right)$;*

b) *if i is even and $a_i\neq X_{[i/2]}$ then either t is less than or equal to $\mu\left(\left[\frac{i}{2}\right]\right)$ or for every $i\in\mathbf{N}$, $a_i=\emptyset$ and every set $V\in t$ is infinite.*

Let $(u, \{a_{p,q}\}_{(p,q)\in\mathbf{K}})$ be a point where $u\in\mathscr{C}(U, U_{i,j}: i, j\in\mathbf{N})$ and $a_{p,q}\in\mathfrak{B}_{[p/2]}$ for every $(p,q)\in\mathbf{K}$. Then $(u, \{a_{p,q}\}_{(p,q)\in\mathbf{K}})$ is a point of $\mathfrak{N}$ iff the following conditions hold:

a) *if p is odd and $a_{p,q}\neq\emptyset$ then u is greater than or equal to q;*

b) *if p is even and $a_{p,q}\neq X_{[p/2]}$ then either u is less than or equal to q or for every $(p,q)\in\mathbf{K}$, $a_{p,q}=\emptyset$ and either $u=\emptyset$ or every set $V\in u$ is infinite.*

Proof. Easy.

Notice, if $(u, \{a_{p,q}\}_{(p,q)\in\mathbf{K}})$ is a point of $\mathfrak{N}$ then there exist only finitely many $(p,q)\in\mathbf{K}$ with $a_{p,q}=\emptyset$, $X_{[p/2]}$. Hence we get

Corollary 1.13. *An element $(u, \{a_{p,q}\}_{(p,q)\in\mathbf{K}})$ of $\mathfrak{N}$ is meet-infinite iff either there exists an infinite set $V\subset U$ with $V\in u$ or $u=\emptyset$, or there exists $(p,q)\in\mathbf{K}$ with $X_{[p/2]}\neq a_{p,q}\supset M_{[p/2]}$. An element $(u, \{a_{p,q}\}_{(p,q)\in\mathbf{K}})$ of $\mathfrak{N}$ is join-infinite iff there exists $(p,q)\in\mathbf{K}$ with $M_{[p/2]}\supset a_{p,q}\neq\emptyset$.*

Proof. The statement follows from Lemmas 1.4, 1.10, 1.12 and the fact that if $Z_1\cap Z_2=\emptyset$ then either $Z_1=\emptyset$ or $Z_2=\emptyset$ and if $Z_1\cup Z_2=X_i$ then either $Z_1=X_i$ or $Z_2=X_i$ in each $\mathfrak{B}_i$.

Proposition 1.14. *Let $\varphi\colon\mathfrak{M}\to\mathfrak{N}$ be a one-to-one lattice homomorphism. Then for every $t\in T$, $\varphi(\alpha(t))=\beta(u)$ where $\varepsilon(u)=t$.*

Proof. Set $\varphi(\alpha(t))=(u^t, \{a^t_{p,q}\}_{(p,q)\in \mathbf{K}})$. Since $\alpha(t)$ is join-infinite, we get according to Corollary 1.13 and Lemma 1.12 that there exists $\bar{u}^t\in U$ such that either $u^t\leqq\bar{u}^t$ or $u^t\geqq\bar{u}^t$.

a) First we prove that $u^t=\bar{u}^t$. Assume the contrary, i.e. for some $t_0\in T$, $u^{t_0}\neq\bar{u}^{t_0}$. We know that there exists a finite set $W\subset U$ with $W\in u^{t_0}$. Now if $u^{t_0}<\bar{u}^{t_0}$ (in the case $u^{t_0}>\bar{u}^{t_0}$, the proof is analogous) then put $L_t=\{u\in U: u>u^t\}$ for $t\in T$. Clearly, L_t is a finite set for every $t\in T$. Now there exists a finite subset $T'\subset T$ with $\bigcap_{t\in T} L_t=\bigcap_{t\in T'} L_t$. Then $\bigvee\{\alpha(t): t\in T'\}$ is join-infinite and therefore $\bigcap_{t\in T} L_t\neq\emptyset$ (see Corollary 1.13 and Lemma 1.12). For $t\in T-T'$ put

$$E_t=\left\{(t, \{a_i\}_{i\in\mathbf{N}})\in\mathfrak{M}: \left(\left(\mu\left(\left[\frac{i}{2}\right]\right)=t\right), (i \text{ is even})\Rightarrow a_t=M_{[i/2]}\right),\right.$$

$$\left.(i \text{ is odd}\,(\exists x\in M_{[i/2]})(a_i=M_{[i/2]}-\{x\}))\right\}.$$

Hence, if e_1, e_2 are distinct points of E_t then $e_1\vee e_2=\alpha(t)$ and $e_1\vee c\neq e_2\vee c$, $e_1\wedge c=e_2\wedge c$ where $c=\bigvee\{\alpha(t): t\in T'\}$. For $w\in\mathfrak{M}$, let $\varphi(w)=\{(v^w, b^w_{p,q})\}_{(p,q)\in\mathbf{K}}$. Since no element of $\mathscr{C}(U, U_{i,j}: i, j\in\mathbf{N})$ is join-infinite, the set $\bar{E}_t=\{e\in E_t: v^e=u^t\}$ is infinite for every $t\in T-T'$ (because for infinitely many $e_1, e_2\in E_t$, $v^{e_1}=v^{e_2}$ and then necessarily $v^{e_1}=u^t$). Hence for $e\in\bar{E}_t$, $v^{e\vee c}=v^{\alpha(t)\vee c}$. Now, for distinct t_1, $t_2\in T-T'$ and for $e_1\in\bar{E}_{t_1}$, $e_2\in\bar{E}_{t_2}$, we have $e_1\wedge e_2=\alpha(t_1)\wedge\alpha(t_2)$. Thus, for every distinct points $t_1, t_2\in T-T'$,

(a) $e_1\vee c\neq e_2\vee c$ for any $e_1, e_2\in\bar{E}_{t_1}$ and $v^{e_1\vee c}=v^{e_2\vee c}$,

(b) $e\wedge\bar{e}=\alpha(t_1)\wedge\alpha(t_2)$ for every $e\in\bar{E}_{t_1}$, $\bar{e}\in\bar{E}_{t_2}$.

Since for every $t\in T-T'$ there exists only a finite subset $\mathbf{K}_t\subset\mathbf{K}$ such that $(p,q)\in\mathbf{K}_t$ whenever $b_{p,q}\neq\emptyset$, $X_{[p/2]}$ and $(v^{e\vee c}, \{b_{p,q}\}_{(p,q)\in\mathbf{K}})\in\mathfrak{N}$ where $e\in E_t$, therefore there exists $(p_t, q_t)\in\mathbf{K}$ and an infinite set $\tilde{E}_t\subset\bar{E}_t$ such that $b^{e_1\vee c}_{p_t,q_t}\neq b^{e_2\vee c}_{p_t,q_t}$ whenever e_1, e_2 are distinct points of $\tilde{E}_t$. Since $\alpha(t)\wedge\alpha(t')\leqq e\vee c$ for every $t, t'\in T-T'$ and $e\in E_t$, we get that $b^{\alpha(t)\wedge\alpha(t')}_{p_t,q_t}=\emptyset$ (see Lemma 1.6) and since in $\mathfrak{B}_i$ $Z_1\cap Z_2=\emptyset$ implies either $Z_1=\emptyset$ or $Z_2=\emptyset$, we have that for every $t'\in T-T'$, $t\neq t'$ and every $e\in\bar{E}_{t'}$, $b^e_{p_t,q_t}=\emptyset$. Since $q_{\bar{t}}\in\bigcap_{t\in T} L_t$ for every $\bar{t}\in T-T'$ and since $\bigcap_{t\in T} L_t$ is finite, we get a contradiction. Hence $u^{t_0}=\bar{u}^{t_0}$.

b) Now, we prove that $a_{\bar{p},u}=M_{[\bar{p}/2]}$. Assume the contrary, i.e. $a_{\bar{p},u}=Z\neq M_{[\bar{p}/2]}$. If $\bar{p}$ is odd then for $t', t''\in T$, $t'\neq t''\neq t$, $t'\neq t$, the element $e=(\alpha(t)\vee\alpha(t'))\wedge\wedge(\alpha(t)\vee\alpha(t''))$ is both meet- and join-infinite. On the other hand, if $\varphi(e)==(v^e, \{b^e_{p,q}\}_{(p,q)\in\mathbf{K}})$ then $b^e_{p,q}=\emptyset$ or $X_{[p/2]}$ if $(p,q)\neq(\bar{p}, u)$ and $b^e_{\bar{p},u}=Z$, which contradicts Lemmas 1.4 and 1.5. If $\bar{p}$ is even, the proof is analogous. Thus $a_{\bar{p},u}=M_{[\bar{p}/2]}$.

c) Now we prove that $\varepsilon(u^t)=t$. Let i_0 be an odd natural number with $\mu\left(\left[\frac{i_0}{2}\right]\right)=t$, let p_0 be an odd natural number with $\mu\left(\left[\frac{p_0}{2}\right]\right)=\varepsilon(u^t)$. It suffices to prove that $i_0=p_0$. Define $\psi\colon \mathfrak{B}_{i_0}\to\mathfrak{B}_{p_0}$ as follows: $\psi(Z)=b^{e_Z}_{p,u^t}$ where $e_Z=(t, \{a_i\}_{i\in\mathrm{N}})$ and if $\mu\left(\left[\frac{i}{2}\right]\right)=t$ and i is odd then $a_i=Z$, while if i is even then $a_i=M_{[i/2]}$ (recall that $\varphi(e_Z)=(v^{e_Z}, \{b^{e_Z}_{p,q}\}_{(p,q)\in\mathrm{K}})$). It is clear that ψ is a lattice homomorphism (it is a composition of the embedding of $\mathfrak{B}_{i_0}$ into $\mathfrak{M}$, of φ and of the projection from $\mathfrak{N}$ to $\mathfrak{B}_{p_0}$). We shall prove that ψ is one-to-one. By Lemma 1.6 it suffices to prove that $\psi|\mathfrak{A}_{i_0}$ is one-to-one. First we shall prove that for every $Z\in\mathfrak{A}_{i_0}$, $v^{e_Z}=u^t$. Hence, it follows immediately that ψ is one-to-one and by Corollary 1.8, $i_0=p_0$. Put

$$E_1=\left\{(t, \{a_i\}_{i\in\mathrm{N}})\colon \left(\left(\mu\left(\left[\frac{i}{2}\right]\right)=t\right), (i \text{ is even})\Rightarrow a_i=M_{[i/2]}\right),\right.$$
$$\left.(i \text{ is odd}\Rightarrow(\exists x\in M_{[i/2]})(a_i=M_{[i/2]}-\{x\}))\right\},$$
$$E_2=\left\{(t, \{a_i\}_{i\in\mathrm{N}})\colon \left(\left(\mu\left(\left[\frac{i}{2}\right]\right)=t\right), (i \text{ is even})\Rightarrow a_i=M_{[i/2]}\right),\right.$$
$$\left.(i \text{ is odd}\Rightarrow(\exists x\in X_{[i/2]}-M_{[i/2]})(a_i=M_{[i/2]}\cup\{x\}))\right\}.$$

Clearly, if we verify that for $e\in E_1\cup E_2$, $v^e=u^t$, then for every $Z\in\mathfrak{A}_{i_0}$, $v^{e_Z}=u^t$. Since for any distinct $e_1, e_2\in E_1$ ($e_1, e_2\in E_2$), $e_1\vee e_2=\alpha(t)$ ($e_1\wedge e_2=\alpha(t)$, resp.) we get that there exists at most one $e_1\in E_1$ (or $e_2\in E_2$) with $v^{e_1}\neq u^t$ (or $v^{e_2}\neq u^t$) because for $u\in U$, if $u_1\vee u_2=u$ (or $u_1\wedge u_2=u$) then either $u_1=u$ or $u_2=u$. Then necessarily $v^{e_1}\leqq u^t\leqq v^{e_2}$. Choose a homomorphism $\sigma\colon\mathscr{D}(T)\to\mathfrak{M}$ such that $\sigma(t)=\alpha(t)$ for every $t\in T$ (clearly, such a homomorphism exists). Now we can choose $t'\in\mathscr{D}(T)$ such that $t'>t$ and $v^{\sigma(t')}$ and v^{e_2} are incomparable. Then $\sigma(t')\wedge e_2=\alpha(t)$ (observe that if $\sigma(t')=(t, \{a_i\}_{i\in\mathrm{N}})$ then for an odd i with $\mu\left(\left[\frac{i}{2}\right]\right)=t$, $a_i=M_{[i/2]}$), but $\varphi(\sigma(t'))\wedge\varphi(e_2)\neq\varphi(\alpha(t))$, a contradiction. Thus for every $e\in E_2$, $v^e=u^t$. Analogously, we prove that $v^{e_1}=u^t$. The proof is concluded.

Theorem 1.15. *For every natural number i, there exist pairwise disjoint sublattices $\mathfrak{N}_0, \mathfrak{N}_1, \dots, \mathfrak{N}_{i-1}$ of the lattice $\mathfrak{N}$ which are isomorphic to $\mathfrak{M}$, but there are not infinitely many pairwise disjoint sublattices $\mathfrak{N}_0, \mathfrak{N}_1, \dots$ of $\mathfrak{N}$ which are isomorphic to $\mathfrak{M}$.*

Proof. Let $\{\varphi_k\}_{k\in\mathrm{N}}$ be a sequence of one-to-one lattice homomorphisms from $\mathfrak{M}$ to $\mathfrak{N}$. Then for arbitrary $k\in\mathrm{N}$ and $t\in T$, $\varphi_k(\alpha(t))=\beta(u)$ where $\varepsilon(u)=t$. Further, for every finite set $T' \subset T$ there exists a point $\gamma(T') \in \mathfrak{M}$ such that $\gamma(T') \leqq \alpha(t)$ iff $t \in T - T'$. On the other hand, if $U' \subset U$ is an infinite set and $U' - U_{i,j} \neq \emptyset$ for any $i, j \in \mathbf{N}$, then $e \leqq \beta(u)$ for every $u \in U'$ iff $e = (\emptyset, \{a_{p,q}\}_{(p,q) \in \mathbf{K}})$ where $a_{p,q} = \emptyset$ for every $(p, q) \in \mathbf{K}$. Therefore there exist i_k, $j_k \in \mathbf{N}$ with $\varphi_k(\alpha(t)) = (u^t, \{a_{p,q}\}_{(p,q) \in \mathbf{K}})$ where $\{u^t\} = \varepsilon^{-1}(t) \cap U_{i_k, j_k}$ for every $t \in T$. Therefore there exist $k_1 \neq k_2$ with $i_{k_1} + j_{k_1} \neq i_{k_2} + j_{k_2}$. Put $\{u\} = U_{i_{k1}, j_{k1}} \cap U_{i_{k2}, j_{k2}}$ $\varepsilon(u) = t$. Then $\varphi_{k_1}(\alpha(t)) = \varphi_{k_2}(\alpha(t))$ and $\{\varphi_k(\mathfrak{M})\}_{k \in \mathbf{N}}$ are not pairwise disjoint.

Let k be a natural number. For every $j \leqq k$ define $\psi_j \colon \mathscr{D}(T) \to \mathscr{C}(U, U_{i,j} \colon i, j \in \mathbf{N})$ as follows: $\psi_j(Z) = \{\varepsilon^{-1}(V) \cap U_{(k-j),j} \colon V \in Z\}$. Clearly, the ψ_j's are one-to-one homomorphisms and $\{\psi_j(\mathscr{D}(T))\}_{j \leqq k}$ are pairwise disjoint. Define $\varphi_j \colon \mathfrak{M} \to \mathfrak{N}$, $\varphi_j(t, \{a_i\}_{i \in \mathbf{N}}) = (\psi_j(t), \{b_{p,q}\}_{(p,q) \in \mathbf{K}})$ where $b_{i, \psi_j(t)} = a_i$. Then $\{\varphi_j \colon j \leqq k\}$ is a family of pairwise disjoint one-to-one lattice homomorphisms. The proof is concluded.

2. Let us formulate the above problem in a general category with a class $\mathfrak{M}$ of its morphisms.

Definition. Let $\mathscr{K}$ be a category with a cosingleton $\emptyset$. Let $f, g \colon A \to B$ be morphisms of $\mathscr{K}$. We shall say that f, g are *disjoint* if

$$\begin{tikzcd} \emptyset \arrow[r] \arrow[d] & A \arrow[d, "f"] \\ A \arrow[r, "g"] & B \end{tikzcd}$$

is a pull back.

Definition. Let $\mathscr{K}$ be a category, let $\mathfrak{M}$ be a class of its morphisms. A pair (A, B) of objects is said to have the property $(S_{\mathfrak{M}})$ if for every $n = 1, 2, \dots$ there exist n pairwise disjoint $\mathfrak{M}$-morphisms from A to B, but there do not exist infinitely many such morphisms. We say that $\mathscr{K}$ fulfils Sekanina's axiom with respect to $\mathfrak{M}$ if no pair of objects has the property $(S_{\mathfrak{M}})$.

Now we can formulate the foregoing result as follows: The pair $(\mathfrak{M}, \mathfrak{N})$ of countable distributive lattices has the property $(S_{\mathscr{M}})$ with $\mathscr{M}$ the class of all monomorphisms.

Now we establish some other results:

Theorem 2.1. *The category of sets, the category of vector spaces and the category of unary algebras with one operation fulfil Sekanina's axiom with respect to $\mathfrak{M}$ for every class $\mathfrak{M}$ containing all monomorphisms.*

Proof. Easy.

Theorem 2.2. *The category of complete, completely distributive Boolean algebras fulfils Sekanina's axiom with respect to the class of all monomorphisms.*

Proof. The statement follows immediately from the well-known fact that every complete, completely distributive Boolean algebra is the algebra of all subsets of some set.

Theorem 2.3. *The category of graphs or unary algebras with α operations (α is a cardinal, $\alpha>0$) fulfils Sekanina's axiom with respect to the class of all summands.*

Proof. The statement follows from the fact that $f\colon A\to B$ is sumand iff A is isomorphic to the sum of some components of B.

Now we recall that a monomorphism f in a category $\mathcal{K}$ is an *extremal monomorphism* if any epimorphism e is an isomorphism whenever $f=g\circ e$ for some morphism g of $\mathcal{K}$. In the category of graphs or topological spaces extremal monomorphisms are embeddings to full subgraphs or subspaces.

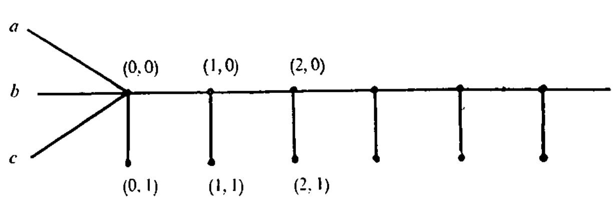
I. Korec showed that there exists a pair (A, B) of countable graphs or countable unary algebras with two operations which have the property $(S_{\mathfrak{M}})$ where $\mathfrak{M}$ is the class of all extremal monomorphisms.

Theorem 2.4. *There exists a pair (A, B) of connected, countable, bipartite graphs with the property $(S_{\mathfrak{M}})$ where $\mathfrak{M}$ is an arbitrary class of monomorphisms containing all extremal monomorphisms.*

Theorem 2.5. *There exists a pair (A, B) of continua with the property $(S_{\mathfrak{M}})$ where A is a subcontinuum of the plane, B is a subcontinuum of the cube and $\mathfrak{M}$ is an arbitrary class of monomorphisms containing all extremal monomorphisms.*

Proof of Theorems 2.4 and 2.5. Put $X=\{a, b, c\}\cup(\mathbf{N}\times\{0, 1\})$,

$$R = \{((0,0),a), ((0,0),b), ((0,0),c), (a,(0,0)), (b,(0,0)), (c,(0,0))\}\cup$$
$$\cup\{((i,0),(i,1)), ((i,1),(i,0))\colon i\in\mathbf{N}\}\cup$$
$$\cup\{((i,0),(i+1,0)), ((i+1,0),(i,0))\colon i\in\mathbf{N}\}.$$



(X, R)

Fig. 2

Clearly, (X, R) is a connected, countable bipartite graph. Choose a bijection φ from $\mathbf{L}=\{(x, y, z, v):(x, y, z, v\in\mathbf{N}), (x+y\neq z+v)\}$ to $\mathbf{N}$. Put $(Y, S)=(X, R)\times\times(\mathbf{N}\times\mathbf{N}, \Delta)/\sim$ where $(\mathbf{N}\times\mathbf{N}, \Delta)$ is the smallest reflexive relation on $\mathbf{N}\times\mathbf{N}$ and $\sim$ is the smallest equivalence relation on $X\times\mathbf{N}\times\mathbf{N}$ with

$$(k, 1, x, y) \sim (k, 1, z, v) \quad \text{whenever} \quad \varphi(x, y, z, v) = k.$$

Clearly, (Y, S) is a connected, countable graph. To verify that it is bipartite, it suffices to put $M=\{(k, i, x, y)\in Y\colon k+i \text{ is even}\}/\sim$. Let k be a natural number, $i\leqq k$. Define $f_i^k\colon (X, R)\to(Y, S)$ as follows: $f_i^k(x)$ is the $\sim$-class containing $(x, k, k-i)$. Clearly, $f_i^k, i=0, 1, \dots, k$, are pairwise disjoint extremal monomorphisms. Let $\{f_i\}$ be a sequence of one-to-one morphisms from (X, R) to (Y, S). Since $\operatorname{card}\{y\colon(y, (0, 0))\in R\}=4$, we get that for every i there exists $(p_i, q_i)\in\mathbf{N}\times\mathbf{N}$ such that $f_i(0, 0)$ is the $\sim$-class containing $(0, 0, p_i, q_i)$. Hence we easily get that $f_i(j, 0)$ is the $\sim$-class containing $(j, 0, p_i, q_i)$ and $f_i(j, 1)$ is the $\sim$-class containing $(j, 1, p_i, q_i)$. Further, there exist i_0, i_1 with $p_{i_0}+q_{i_0}\neq p_{i_1}+q_{i_1}$. Let $k==\varphi(p_{i_0}, q_{i_0}, p_{i_1}, q_{i_1})$. Then $f_{i_0}(k, 1)=f_{i_1}(k, 1)$ and therefore f_{i_0} and f_{i_1} are not disjoint. If we set $A=(X, R)$, $B=(Y, S)$, then the proof of Theorem 2.4 is concluded.

Let K be a circle with the usual topology. Choose two distinct points $a, b\in K$. Let $S=\{\{x, y\}\colon (x, y)\in R\}$ be equipped with the discrete topology where $R\subset X\times X$ is the relation defined above. Let P_1 be the one-point compactification of $K\times S/\sim$ with $\sim$ standing for the smallest equivalence relation such that:

$$(a, \{x, y\}) \sim (a, \{x, z\}) \quad \text{for every} \quad \{x, y\}, \{x, z\}\in S \quad \text{with} \quad x\in\{(i, j)\colon i+j \text{ is even}\};$$

$$(b, \{x, y\}) \sim (b, \{x, z\}) \quad \text{for every} \quad \{x, y\}, \{x, z\}\in S \quad \text{with} \quad x\in\{(i, j)\colon i+j \text{ is odd}\}.$$

Clearly, P_1 is a subcontinuum of the plane. We shall assume that $\mathbf{N}$ has the discrete topology. Let P_2 be the one-point compactification of $P_1\times\mathbf{N}\times\mathbf{N}/\approx$ where $\approx$ is the smallest equivalence relation such that if $\varphi(x, y, z, v)=k$ then

$$\big((a, \{(k, 0), (k, 1)\}], x, y\big) \approx \big([a, \{(k, 0), (k, 1)\}], z, v\big) \quad \text{if } k \text{ is odd},$$

$$\big([b, \{(k, 0), (k, 1)\}], x, y\big) \approx \big([b, \{(k, 0), (k, 1)\}], z, v\big) \quad \text{if } k \text{ is even},$$

where $[x]$ denotes the $\sim$-class containing x. Clearly, P_2 is a subcontinuum of the cube. The proof that (P_1, P_2) has the property $(S_{\mathfrak{M}})$ is analogous to that of the similar statement for (X, R) and (Y, S). It suffices to realize that if $f\colon K\to K$ is one-to-one then f is a homeomorphism.

Theorem 2.6. *There exists a pair (A, B) of 0-dimensional compact Hausdorff spaces on sets of power $\aleph_1$, which has the property $(S_{\mathfrak{M}})$ where $\mathfrak{M}$ is the class of all summands.*

Proof. Define topological spaces S_n by induction as follows: S_1 is the one-point compactification of a countable discrete set; S_n is the one-point compactification of $S_{n-1}\times\mathrm{N}$ where N has the discrete topology. Put R_n to be the one-point compactification of $\aleph_1$ copies of S_n. Let T_1 be the one-point compactification of the disjoint union of $R_1, R_2, \ldots$. Clearly, T_1 is a 0-dimensional compact Hausdorff space on a set of power $\aleph_1$.

Let U be a countable set and let $\{U_{i,j}: i, j\in\mathrm{N}\}$ be a cover of U such that every $U_{i,j}$ is infinite and $U_{i,j}\cap U_{m,n}=\emptyset$ if $i+j=m+n$, $U_{i,j}\cap U_{m,n}$ is infinite if $i+j\neq m+n$. Choose a mapping $\psi\colon U\to\{1,2,3,\ldots\}$ such that $\psi|U_{i,j}$ is a bijection from $U_{i,j}$ onto $\{1,2,3,\ldots\}$ for every $i, j\in\mathrm{N}$.

Let T_2 be the one-point compactification of $T_1\times\mathrm{N}\times\mathrm{N}/\approx$ ($\mathrm{N}\times\mathrm{N}$ has the discrete topology) where $\approx$ is the smallest equivalence relation such that $(x, i, j)\approx(x, m, n)$ if $x\in R_p$ and $p\in\psi(U_{i,j}\cap U_{m,n})$. Clearly, T_2 is a 0-dimensional compact Hausdorff space on a set of power $\aleph_1$.

Let k be a natural number, $i=0, 1, \ldots, k$. Define $f_i^k\colon T_1\to T_2$, by $f_i^k(x)$ being the $\approx$-class containing $(x, k, k-i)$. It is easy to verify that f_i^k is a summand and that f_i^k and f_j^k are disjoint whenever $i\neq j$.

Let $\{f_i\colon T_1\to T_2\}$ be a sequence of summands. Then for every i, there exist $j_i, k_i\in\mathrm{N}$ and $i_1, i_2, \ldots, i_n\in\mathrm{N}$ such that $f_i(T_1-\bigcup_{m=1}^{n} R_{i_m})\subset T_1\times\{(j_i, k_i)\}/\approx$, therefore if $j_{i_0}+k_{i_0}\neq j_{i_1}+k_{i_1}$, we get that $\mathrm{Im} f_{i_0}\cap\mathrm{Im} f_{i_1}\neq\emptyset$ and thus f_{i_0} and f_{i_1} are not disjoint. On the other hand, there exist i_0, i_1 such that either $j_{i_0}+k_{i_0}\neq j_{i_1}+k_{i_1}$ or $(j_{i_0}, k_{i_0})=(j_{i_1}, k_{i_1})$. Hence if we set $A=T_1, B=T_2$, the proof of the theorem is complete.

FACULTY OF MATHEMATICS AND PHYSICS, CHARLES UNIVERSITY
MALOSTRANSKÉ NÁMĚSTÍ 25
118 00 PRAHA 1, CZECHOSLOVAKIA

Acta Sci. Math., 41 (1979), 151—159

Mean ergodic semigroups on W^*-algebras

BURKHARD KÜMMERER and RAINER NAGEL

In 1966 I. KOVÁCS and J. SZŰCS [5] proved the following: If G is a group of *-automorphisms on a W^*-algebra $\mathscr{A}$ having a faithful family of normal G-invariant states, then the weak*-closed convex hull of Gx, $x \in \mathscr{A}$, contains a unique G-invariant element. As its forerunner, the von Neumann mean ergodic theorem, this result has found many applications in mathematical physics. For that reason it is interesting to ask whether the theorem may be generalized to semigroups of bounded operators on $\mathscr{A}$.

On the basis of an abstract mean ergodic theory (see [8], [14] or section 1 below) we will prove in section 2 that the desired result holds for semigroups of operators on $\mathscr{A}$ satisfying a certain contraction property. This answers a question raised in [9]. Our technique can be applied with particular success to semigroups of completely positive contractions on $\mathscr{A}$ (theorem 2.4). In section 3 we investigate the relation between mean ergodicity and compactness of such semigroups.

Some of the results have been announced in [6].

1. The abstract mean ergodic theory

In this section we recall the basic theory of mean ergodic operator semigroups. Let E denote a Banach space, E^* its (topological) dual and $\mathscr{L}(E)$ the space of all bounded linear operators on E. We call a semigroup $S \subset \mathscr{L}(E)$ *(norm)mean ergodic* if the strong (or weak) operator closure $\overline{\text{co}}\ S$ in $\mathscr{L}(E)$ of the convex hull of S contains a projection P satisfying

$$TP = PT = P \quad \text{for all} \quad T \in S.$$

Endow E^* with the weak* topology $\sigma(E^*, E)$ and denote the weak* closed convex hull of the orbit of $\varphi \in E^*$ under the adjoint semigroup $S^*: =\{T^* \in \mathscr{L}(E^*) : T \in S\}$

Received October 3, 1977.

by $\overline{\text{co}}\, S^*\varphi$. The following property of a mean ergodic semigroup is then immediate.

(1.1) Proposition. *Let* $S \subset \mathscr{L}(E)$ *be mean ergodic with projection* P. *Then* P *(resp. its adjoint* P^**) is a projection onto the* S*-fixed space* F *in* E *(resp.* S^**-fixed space* F^* *in* E^**) and we have* $Px \in \overline{\text{co}}\, Sx$, $P^*\varphi \in \overline{\text{co}}\, S^*\varphi$ *for* $x \in E$, $\varphi \in E^*$. *Moreover, the dual of* PE *is isomorphic to* P^*E^*.

For bounded semigroups we have the following useful characterization of mean ergodicity.

(1.2) Theorem [8]. *If the semigroup* $S \subset \mathscr{L}(E)$ *is bounded, the following conditions are equivalent:*

(a) S *is mean ergodic.*

(b) S^* *is weak* mean ergodic, i.e. the closed convex hull of* S^* *with respect to the weak* operator topology* $\sigma(\mathscr{L}(E^*), E^* \otimes E)$ *contains a weak* continuous projection* P^* *satisfying*

$$P^*T^* = T^*P^* = P^* \quad \textit{for all} \quad T \in S.$$

(c) *The* S*-fixed space* F *separates the* S^**-fixed space* F^* *and* $F^* \cap \overline{\text{co}}\, S^*\varphi \neq \emptyset$ *for all* $\varphi \in E^*$.

Remarks: 1. The proof of (c)$\Rightarrow$(a) follows as in [8], 1.7.

2. $\sigma(\mathscr{L}(E^*), E^* \otimes E)$ is the topology of pointwise convergence on $(E^*, \sigma(E^*, E))$.

3. Weak* mean ergodicity of S^* should be distinguished from norm mean ergodicity of S^* on the Banach space E^*.

The main objective of the theory is to show that certain semigroups on certain Banach spaces are mean ergodic. The oldest result in this direction is the von Neumann mean ergodic theorem which states that the semigroup $S := \{T^n : n \in \mathbb{N}\}$ generated by a unitary operator T on a Hilbert space is mean ergodic. This has been generalized considerably with emphasis either on the geometry of the underlying Banach space or on particular properties of the semigroup. We quote two typical results (see [8], 1.4 and 1.9 or [14], III. 7.11 and 7.9).

(1.3) Examples. 1. Every contraction semigroup on a Hilbert space is mean ergodic (Alaoglu-Birkhoff).

2. A group G in $\mathscr{L}(E)$ which is compact in the strong (or weak) operator topology is mean ergodic.

2. A non-commutative mean ergodic theorem

In what follows $\mathscr{A}$ shall always be a W^*-algebra with dual $\mathscr{A}^*$ and predual $\mathscr{A}_*$ (see [3] or [12]). On $\mathscr{A}$ we consider a bounded semigroup S of weak* continuous linear operators whose preadjoints $S_* := \{T_* \in \mathscr{L}(\mathscr{A}_*) : T \in S\}$ then exist. For such a pair $(\mathscr{A}, S)$, called here a *dynamical system*, we state our main result.

(2.1) Theorem. *Let $(\mathscr{A}, S)$ be a dynamical system. If there exists a faithful family Φ of normal states on $\mathscr{A}$ satisfying*

$$(*) \qquad \varphi\big((Tx)^*(Tx)\big) \leqq \varphi(x^*x) \quad \text{for all} \quad \varphi \in \Phi,\ T \in S,\ x \in \mathscr{A},$$

then S is weak mean ergodic.*

The proof of the theorem will be based on example 1.3.1 and on the coincidence of certain topologies. For this purpose we denote by $\mathscr{T}_A$, A contained in $\mathscr{A}_*^+$, the set of all normal positive linear forms on $\mathscr{A}$, the topology on $\mathscr{A}$ generated by the seminorms

$$x \to \varphi(x^*x)^{1/2}, \quad \varphi \in A.$$

In particular, we write $\mathscr{T}_A = \mathscr{T}_\varphi$ if $A = \{\varphi\}$ and we have $\mathscr{T}_A = s(\mathscr{A}, \mathscr{A}_*)$, the strong topology, if $A = \mathscr{A}_*^+$ (see [12], 1.8.6). Take now $\varphi \in \mathscr{A}_*^+$ and denote its support by p_φ and the orthogonal complement $1 - p_\varphi$ by $p_\varphi^\perp$. Then we have $\mathscr{A} = K_\varphi \oplus L_\varphi$ setting

$$K_\varphi := \mathscr{A} p_\varphi \quad \text{and} \quad L_\varphi := \mathscr{A} p_\varphi^\perp = \{x \in \mathscr{A} : \varphi(x^*x) = o\}.$$

If $\mathscr{A}$ is a weakly closed *-subalgebra of $\mathscr{L}(H)$, H Hilbert space, it follows from [3], chap. I, § 4, prop. 4 and [12], 1.15.2 that the topologies $\mathscr{T}_\varphi$, $s(\mathscr{A}, \mathscr{A}_*)$ and the strong operator topology coincide on norm bounded sets of K_φ.

Assume now that A is a faithful family of normal states on $\mathscr{A}$. It follows from the considerations above that $\mathscr{T}_A$ coincides on norm bounded subsets of $\mathscr{A}$ with the topology of pointwise convergence on $\bigcup_{\varphi \in A} p_\varphi H$. By [3], chap. I, § 4, sect. 6 this set is total in H. Therefore we can apply [13], III.4.5. and have the following lemma.

(2.2) Lemma. *Let $\mathscr{A}$ be a W^*-algebra in $\mathscr{L}(H)$, H Hilbert space, with unit ball $\mathscr{A}_1$. If $\varphi \in \mathscr{A}_*^+$ (resp. if $A \subset \mathscr{A}_*^+$ faithful) then $\mathscr{T}_\varphi$ (resp. $\mathscr{T}_A$) coincides on $\mathscr{A}_1 \cap K_\varphi$ (resp. $\mathscr{A}_1$) with the strong operator topology.*

Proof of the Theorem. *Step 1.* For $\varphi \in \Phi$ it follows from $(*)$ that L_φ is S-invariant. Therefore S induces an operator semigroup S_φ on K_φ by

$$T_\varphi(xp_\varphi) := T(xp_\varphi)p_\varphi = (Tx)p_\varphi \quad \text{for all} \quad T \in S.$$

The completion of K_φ with respect to $\mathscr{T}_\varphi$ is a Hilbert space H_φ on which S_φ induces a contraction semigroup $\tilde{S}_\varphi$ (again by $(*)$). This contraction semigroup is mean ergodic by (1.3.1) with corresponding projection $P_\varphi \in \mathscr{L}(H_\varphi)$ and has the property that

$$x_\varphi^0 := P_\varphi x, \quad x \in H_\varphi,$$

is contained in the $\mathscr{T}_\varphi$-closed convex hull of $\tilde{S}_\varphi x$. Now take $x \in K_\varphi$. Then $\mathrm{co}\, \tilde{S}_\varphi x$ is a norm bounded subset of $K_\varphi \subset \mathscr{A}$ whose strong operator closure is a complete subset of K_φ ([3], chap. I, 3, sect. 1). From (2.2) it follows that $x_\varphi^0 \in K_\varphi$. Since the bounded convex subsets of $\mathscr{A}$ have the same closure for the strong operator topology, the weak operator topology and the weak* topology ([3], chap. I, § 3, th. 1), we have shown that for every $x \in K_\varphi$ there is a unique S_φ-fixed point x_φ^0 contained in the weak* closed convex hull $\overline{\mathrm{co}}\, S_\varphi x$ of $S_\varphi x$.

Step 2. For $x \in \mathscr{A}$, $\varphi \in \Phi$ we define the non-empty, weak* compact set

$$Q_\varphi(x) := \{y \in \mathscr{A} \colon x_\varphi^0 = y p_\varphi \text{ and } \|y\| \leqq r\|x\|\}$$

where $r := \sup\{\|T\| \colon T \in S\}$. For $\varphi_1, \varphi_2 \in \Phi$ and $p_{\varphi_1} \leqq p_{\varphi_2}$ it follows from the construction above that $x_{\varphi_2}^0 p_{\varphi_1} = x_{\varphi_1}^0$ and therefore $Q_{\varphi_1}(x) \supset Q_{\varphi_2}(x)$. Since we may assume that Φ is convex, we obtain that $\{Q_\varphi(x) \colon \varphi \in \Phi\}$ is filtered downwards. Because of compactness there exists an element $x^0 \in \bigcap_{\varphi \in \Phi} Q_\varphi(x)$. From $(Tx^0 - x^0)p_\varphi = T_\varphi(x^0 p_\varphi) - x^0 p_\varphi = 0$ for all $\varphi \in \Phi$ it follows that x^0 is an S-fixed point. Moreover, x^0 is contained in the $\mathscr{T}_\varphi$-closed, $\varphi \in \Phi$, hence in the $\mathscr{T}_\Phi$-closed convex hull of Sx. Again we conclude by (2.2) that this closure coincides with $\overline{\mathrm{co}}\, Sx$ for the weak* topology on $\mathscr{A}$.

Step 3. Take $0 \neq x \in \mathscr{A}$ such that $Tx = x$ for all $T \in S$. Since Φ is faithful there exists $\varphi \in \Phi$ such that

$$0 \neq x p_\varphi =: x_\varphi = T_\varphi x_\varphi \in K_\varphi \quad \text{for all } T_\varphi \in S_\varphi.$$

Since $\tilde{S}_\varphi$ is mean ergodic, we can find a continuous linear form $\tilde{\psi}$ on the Hilbert space H_φ such that $\tilde{\psi}(x_\varphi) \neq 0$ and $\tilde{\psi}(T_\varphi y_\varphi) = \tilde{\psi}(y_\varphi)$ for all $T \in S$, $y \in \mathscr{A}$ and $y_\varphi := y p_\varphi \in K_\varphi$. The formula $\psi(y) := \tilde{\psi}(y_\varphi)$, $y \in \mathscr{A}$, defines an $s(\mathscr{A}, \mathscr{A}_*)$-continuous, hence a weak* continuous linear form on $\mathscr{A}$ (see [12], 1.8.10 and 1.8.12). Obviously ψ is S_*-invariant and does not vanish on x. Therefore the fixed space of S_* separates the fixed space of S and S is weak* mean ergodic by (1.2.c).

Remarks. 1. If S satisfies the above assumptions, then the preadjoint semigroup S_* is (norm)mean ergodic on $\mathscr{A}_*$ by (1.2).

2. The same result holds if Φ satisfies

$$(*) \qquad \varphi\big((Tx)(Tx)^*\big) \leqq \varphi(xx^*) \quad \text{for all } \varphi \in \Phi, T \in S, x \in \mathscr{A}.$$

3. If $(\mathscr{A}_1, S_1)$ and $(\mathscr{A}_2, S_2)$ are dynamical systems having faithful families Φ_1 and Φ_2 of normal states which satisfy $(*)$ and if one defines the semigroup

$S_1 \otimes S_2$ on the W^*-tensor product $\mathscr{A}_1 \tilde{\otimes} \mathscr{A}_2$ in the usual way, then $\Phi_1 \otimes \Phi_2$ is a faithful family of normal states on $\mathscr{A}_1 \tilde{\otimes} \mathscr{A}_2$ satisfying $(*)$.

(2.3) Corollary. *Let* $(\mathscr{A}, S)$ *be a dynamical system with* Φ *as in* (2.1). *For* $\varphi \in \Phi$, $x \in \mathscr{A}$ *(and with the notation as in the proof of* (2.1)*) the following are equivalent:*

(a) x_φ *is* $\tilde{S}_\varphi$*-invariant in* H_φ.

(b) $y \mapsto \varphi(x^* y)$, $y \in \mathscr{A}$, *defines an* S_**-invariant, weak* continuous linear form on* $\mathscr{A}$.

Proof. (a)$\Rightarrow$(b): The contraction semigroup $\tilde{S}_\varphi$ on H_φ and its adjoint semigroup have the same fixed space containing $x_\varphi = xp_\varphi$. Therefore we obtain

$$\begin{aligned}\varphi(x^* y) &= \varphi(p_\varphi x^* y p_\varphi) = \langle y_\varphi, x_\varphi \rangle_\varphi = \langle T_\varphi y_\varphi, x_\varphi \rangle_\varphi = \\ &= \langle (Ty) p_\varphi, x_\varphi \rangle_\varphi = \varphi\big(p_\varphi x^* (Ty) p_\varphi\big) = \varphi(x^* Ty)\end{aligned}$$

for all $T \in S$.

(b)$\Rightarrow$(a): Assume $T_\varphi(xp_\varphi) \neq xp_\varphi$ for some $T_\varphi \in \tilde{S}_\varphi$ and therefore $T_\varphi^*(xp_\varphi) \neq xp_\varphi$. Since $\mathscr{A} p_\varphi$ is dense in H_φ, there exists $y_\varphi = yp_\varphi$ such that

$$\begin{aligned}\varphi(x^* y) &= \varphi(p_\varphi x^* y p_\varphi) = \langle y_\varphi, x_\varphi \rangle_\varphi \neq \langle y_\varphi, T_\varphi^* x_\varphi \rangle_\varphi = \\ &= \langle T_\varphi y_\varphi, x_\varphi \rangle_\varphi = \varphi(x^* Ty).\end{aligned}$$

The most important application of the above theorem will be to semigroups of completely positive operators on $\mathscr{A}$ (see [16] for the definition). In particular we will obtain useful information on the corresponding mean ergodic projection and on the fixed space.

(2.4) Theorem. *Let* $(\mathscr{A}, S)$ *be a dynamical system where* S *consists of completely positive contractions. If there exists a faithful family* Φ *of* S_**-subinvariant normal states on* $\mathscr{A}$*, then* S *is weak* mean ergodic. Moreover, the corresponding projection* P *is a conditional expectation onto the weak* closed fixed point subalgebra* $P\mathscr{A}$ *of* $\mathscr{A}$.

Proof. Completely positive contractions $T \in S$ satisfy a Schwarz inequality

$$(Tx)^*(Tx) \leqq T(x^* x) \quad \text{for all} \quad x \in \mathscr{A} \quad \text{(see [16])}.$$

For $\varphi \in \Phi$ we get

$$\varphi\big((Tx)^*(Tx)\big) \leqq \varphi\big(T(x^* x)\big) \leqq \varphi(x^* x) \quad \text{for all} \quad x \in \mathscr{A},$$

hence $(*)$ is satisfied and S is weak* mean ergodic by (2.1). Obviously, the projection P is positive and therefore $P(y^*) = (Py)^*$ for all $y \in \mathscr{A}$. Consequently, it suffices to show that $P(x_0 y)=x_0 Py$ for all $x_0\in P\mathscr{A}$, $y\in\mathscr{A}$. Take $\varphi\in\Phi$, $x_0\in P\mathscr{A}$ and denote the weak* continuous S_*-invariant linear form $x\mapsto\varphi(x_0^* x)$ by $L_{x_0}\varphi$ (use (2.3)). But $L_{x_0}\varphi$ is a linear combination of positive elements in $P_*\mathscr{A}_*$. Again by (2.3) we obtain $L_{z_0}\varphi\in P_*\mathscr{A}_*$ for $z_0:=y_0x_0$ and $y_0\in P\mathscr{A}$. Preserving the notation of (2.3) we remark that the mean ergodic projection P_φ is self adjoint on H_φ. Therefore we obtain by an analogous computation that $\langle x_\varphi, P_\varphi z_{0\varphi}\rangle=\langle x_\varphi, z_{0\varphi}\rangle$ for all $x\in\mathscr{A}$. Since Φ is faithful we conclude $z_0=y_0x_0\in P\mathscr{A}$. Consequently $x_0Py\in P\mathscr{A}$, and for $\varphi\in P_*\mathscr{A}_*$ we have $\varphi(P(x_0y))=P_*\varphi(x_0y)=\varphi(x_0y)=L_{x_0^*}\varphi(y)=$ $=P_*(L_{x_0^*}\varphi)(y)=L_{x_0^*}\varphi(Py)=\varphi(x_0Py)$. Since $P_*\mathscr{A}_*$ separates $P\mathscr{A}$, the assertion follows.

The known mean ergodic theorems for W^*-algebras follow from the fact that *-homomorphisms on arbitrary W^*-algebras and positive operators on commutative W^*-algebras are completely positive.

(2.5) Corollary. (Kovács—Szűcs [5], [1]) *Let S be a semigroup of normal *-homomorphisms on a W^*-algebra $\mathscr{A}$. If there exists a faithful family Φ of S_*-invariant normal states on $\mathscr{A}$, then S is weak* mean ergodic and the corresponding projection is a conditional expectation.*

(2.6) Corollary (Nagel [8]): *Let S be a semigroup of weak* continuous positive contractions on a commutative W^*-algebra $\mathscr{A}$. If there exists a faithful normal S_*-subinvariant state on $\mathscr{A}$, then S is weak* mean ergodic.*

(2.7) Corollary. *If $(\mathscr{A}, S)$ satisfies all assumptions of* (2.4), *the following are equivalent:*

(a) $T1=1$ *for all* $T\in S$.

(b) *The mean ergodic projection P is strictly positive.*

(c) *Φ consists of S-invariant states.*

Proof. (a)⇒(c): $T1=1$ and $\varphi\in\Phi$ implies $Tp_\varphi\leqq p_\varphi$ and $T_\varphi p_\varphi=p_\varphi$. From (2.3) it follows that $x\mapsto\varphi(p_\varphi x)=\varphi(x)$ is T_*-invariant. (c)⇒(b) is trivial and (b)⇒(a) follows since $Py=0$ for $y:=1-T1\geqq 0$.

3. Compactness and mean ergodic properties

Compactness in some form underlies many ergodic theorems. In the case of automorphism groups on W^*-algebras this has been studied by several authors (e.g. [15], [4], [7], [10]). We will generalize these results to bounded semigroups.

As in the previous section, let $\mathscr{A}$ denote a W^*-algebra with predual $\mathscr{A}_*$. On $\mathscr{L}(\mathscr{A})$ (resp. $\mathscr{L}(\mathscr{A}_*)$) we consider the weak* operator topology $\sigma(\mathscr{L}(\mathscr{A}), \mathscr{A}\otimes\mathscr{A}_*)$

(resp. the weak operator topology $\sigma(\mathscr{L}(\mathscr{A}_*), \mathscr{A}_*\otimes\mathscr{A})$) and recall that the unit ball in $\mathscr{L}(\mathscr{A})$ is weak* operator compact.

(3.1) Proposition. *For a dynamical system $(\mathscr{A}, S)$ the following are equivalent:*
(a) S_* *(and* co S_**) are relatively compact.*
(b) *The closure of S (and of* co S*) contains only weak* continuous operators.*
(c) $S_*(W)$ *is relatively weakly compact for any weakly compact set $W\subset\mathscr{A}_*$.*
(d) S *is equicontinuous for the Mackey topology $\tau(\mathscr{A}, \mathscr{A}_*)$.*

Proof. The implications (a)$\Leftrightarrow$(b)$\Leftarrow$(c)$\Leftrightarrow$(d) hold in any (dual) Banach space and follow from topological vector space theory (use [13], IV.3.2, IV.11.4).

(a)$\Rightarrow$(c): By Eberlein's theorem ([13], IV.11.2) it suffices to show that $\{T_*\psi_i\colon T\in S, i\in\mathbf{N}\}$ is relatively weakly compact for any weakly convergent sequence $\{\psi_i\}$ in $\mathscr{A}_*$. To that purpose we choose a sequence $\{p_n\}$ of mutually orthogonal projections in $\mathscr{A}$ and show that $\lim\limits_{n\to\infty}\psi_i(Tp_n)=0$ uniformly for $i\in\mathbf{N}$, $T\in S$ (use the compactness criterion from [2]). Take $\varepsilon>0$ and denote the limit of $\{\psi_i\}$ by ψ_0 and the r-ball in $\mathscr{A}$ for $r:=\sup\{\|T\|\colon T\in S\}$ by $\mathscr{A}_r$. Then define

$$Q_i := \{x\in\mathscr{A}_r\colon |(\psi_j-\psi_0)(x)| \leqq \varepsilon/4 \text{ for all } j\geqq i\}.$$

Each Q_i is weak* closed and $\mathscr{A}_r=\bigcup\limits_{i=1}^{\infty} Q_i$. Since $\mathscr{A}_r$ is weak* compact, there exists $i_0\in\mathbf{N}$ such that Q_{i_0} contains a weak* open set in $\mathscr{A}_r$, i.e. there exists $x_0\in\mathscr{A}_r$ and $\varphi_1, \dots, \varphi_m\in\mathscr{A}_*$ such that $\bigcap\limits_{i=1}^{m}\{x\in\mathscr{A}_r\colon |\varphi_i(x)-\varphi_i(x_0)|<1\}$ is contained in Q_{i_0}. By (a) the set $\{T_*\varphi_i\colon T\in S, 1\leqq i\leqq m\}$ is relatively weakly compact, hence there exists $n_1\in\mathbf{N}$ such that $|\varphi_i(Tp_n)|<1$ for all $T\in S$, $1\leqq i\leqq m$ and $n\geqq n_1$ (use [2] again). Then $(Tp_n+x_0)\in Q_{i_0}$ and, since $x_0\in Q_{i_0}$,

$$(1) \qquad |(\psi_j-\psi_0)(Tp_n)| \leqq \varepsilon/2 \quad\text{for}\quad T\in S,\ j\geqq i_0 \quad\text{and}\quad n\geqq n_1.$$

We apply now the hypothesis (a) and the compactness criterion to the set $\{T_*(\psi_j-\psi_0)\colon T\in S,\ 1\leqq j\leqq i_0-1\}$ and find $n_2\in\mathbf{N}$ such that

$$(2) \qquad |(\psi_j-\psi_0)(Tp_n)| \leqq \varepsilon/2 \quad\text{for}\quad T\in S,\quad 1\leqq j\leqq i_0-1,\quad n\geqq n_2.$$

Finally, there exists $n_3\in\mathbf{N}$ such that

$$(3) \qquad |\psi_0(Tp_n)| \leqq \varepsilon/2 \quad\text{for}\quad T\in S \quad\text{and}\quad n\geqq n_3.$$

Define $n_0:=\max\{n_1, n_2, n_3\}$ and conclude from (1), (2) and (3) that

$$|\psi_j(Tp_n)| \leqq \varepsilon \quad\text{for}\quad T\in S,\quad j\in\mathbf{N} \quad\text{and}\quad n\geqq n_0.$$

Remarks 1. SAITÔ [10] proved the equivalence of (a) and (c) for mean ergodic groups of *-automorphisms.

2. Instead of the weak* operator topology we could use equally well the topology of pointwise convergence on $\mathscr{A}$ where $\mathscr{A} \subset \mathscr{L}(H)$ is endowed with the weak operator topology.

Our next result shows once more the usefulness of condition $(*)$ of section 2. Not only does it imply mean ergodicity but also compactness.

(3.2) Proposition. *Let $(\mathscr{A}, S)$ be a dynamical system. If there exists a faithful family Φ of normal states on $\mathscr{A}$ satisfying*

$$(*) \qquad \varphi((Tx)^*(Tx)) \leqq \varphi(x^*x) \quad \textit{for all} \quad T \in S, \quad x \in \mathscr{A}, \quad \varphi \in \Phi,$$

then S satisfies (a)–(d) *of* (3.1).

Proof. Take $\psi \in \mathscr{A}_*$ and assume that $S_*\psi$ is not relatively weakly compact. By [2] there exists $\varepsilon > 0$, a sequence $\{T_n\} \subset S$ and a sequence $\{p_n\}$ of mutually orthogonal projections in $\mathscr{A}$ such that $|\psi(T_n p_n)| \geqq \varepsilon$ for all $n \in \mathbf{N}$. But for $\varphi \in \Phi$ we have $\lim_{n\to\infty} \varphi(p_n) = 0$ and $\lim_{n\to\infty} \varphi((T_n p_n)^*(T_n p_n)) = 0$ (apply $(*)$), i.e. $T_n p_n$ converges to 0 for the topology $\mathscr{T}_\Phi$. Since this topology is stronger than the weak* topology on norm bounded sets of $\mathscr{A}$ (use (2.2)), we obtain the contradiction that $\psi(T_n p_n)$ converges to 0.

Remark. Examples show that for arbitrary dynamical systems weak compactness does not imply weak* mean ergodicity [10] nor does weak* mean ergodicity imply compactness.

(3.3) Corollary. *Let $(\mathscr{A}, S)$ be a dynamical system where S is a group of $*$-automorphisms on $\mathscr{A}$. The following properties are equivalent:*

(a) *There exists a faithful family of S_*-invariant normal states.*

(b) *S is weak* mean ergodic.*

(c) *S_* is relatively weak operator compact in $\mathscr{L}(\mathscr{A}_*)$.*

Proof. The implications (a)$\Rightarrow$(b) and (a)$\Rightarrow$(c) follow from (2.1) and (3.2), while a proof of (c)$\Rightarrow$(a) can be found in [15]. Assume now that S is weak* mean ergodic with projection P. To prove (a) it suffices to show that P is strictly positive: Assume $P(x^*x) = 0$ for some non-zero $x \in \mathscr{A}$. Since P is normal we can find a maximal projection $0 \neq p \in \mathscr{A}$ such that $P(p) = 0$. This and the assumption $T1 = 1$ imply $Tp \leqq p$ for all $T \in S$. But S is a group and therefore $Tp = p$ for all $T \in S$, which is a contradiction.

References

[1] S. M. ABDALLA—J. SZŰCS, On an ergodic type theorem for von Neumann algebras, *Acta Sci. Math.*, **36** (1974), 167—172.

[2] C. A. AKEMANN, The dual space of an operator algebra, *Trans. Amer. Math. Soc.*, **126** (1967), 286—302.

[3] J. DIXMIER, *Les algèbres d'opérateurs dans l'espace Hilbertien*, Gauthier-Villars (Paris, 1969).

[4] W. L. GREEN, Compact groups of automorphisms of von Neumann algebras, *Math. Scand.*, **37** (1975), 284—296.

[5] I. KOVÁCS—J. SZŰCS, Ergodic type theorems in von Neumann algebras, *Acta Sci. Math.*, **27** (1966), 233—246.

[6] B. KÜMMERER, Mean ergodic semigroups of contractions in W^*-algebras, in *Group theoretical methods in Physics*, Lecture Notes in Physics, **79**, 479—481, Springer (Berlin—Heidelberg—New York, 1978).

[7] A. T. LAU, W^*-algebras and invariant functionals, *Studia Math.*, **56** (1976), 253—261.

[8] R. NAGEL, Mittelergodische Halbgruppen linearer Operatoren, *Ann. Inst. Fourier*, Grenoble, **23** (1973), 75—87.

[9] R. NAGEL, *Mean ergodic semigroups and invariant ideals in ordered Banach spaces*. Lecture Notes in Physics, **29**, Springer (Berlin—Heidelberg—New York, 1974), pp. 309—315.

[10] K. SAITÔ, Automorphism groups of von Neumann algebras and ergodic type theorems, *Acta Sci. Math.*, **36** (1974), 119—130.

[11] K. SAITÔ, Groups of *-automorphisms and invariant maps of von Neumann algebras, *Pacific J. Math.*, **57** (1975), 553—557.

[12] S. SAKAI, *C^*-algebras and W^*-algebras*, Springer (Berlin—Heidelberg—New York, 1971).

[13] H. H. SCHAEFER, *Topological vector spaces*, Springer (Berlin—Heidelberg—New York, 1971).

[14] H. H. SCHAEFER, *Banach lattices and positive operators*, Springer (Berlin—Heidelberg—New York, 1974).

[15] E. STØRMER, Invariant states of von Neumann algebras, *Math. Scand.*, **30** (1972), 253—256.

[16] E. STØRMER, *Positive linear maps of C^*-algebras*, Lecture Notes in Physics **29**, 85—106, Springer (Berlin—Heidelberg—New York, 1974).

MATHEMATISCHES INSTITUT
UNIVERSITÄT TÜBINGEN,
AUF DER MORGENSTELLE 10
7400 TÜBINGEN, BRD.

Acta Sci. Math., 41 (1979), 161—172

Purely contractive analytic functions and characteristic functions of non-contractions

BRIAN W. McENNIS

1. Introduction

If T is a bounded operator on a Hilbert space $\mathfrak{H}$, then the *characteristic function* of T is the operator valued analytic function

$$\Theta_T(\lambda) = [-TJ_T + \lambda J_{T^*} Q_{T^*}(I-\lambda T^*)^{-1} J_T Q_T] \,|\, \mathfrak{D}_T,$$

where $J_T = \mathrm{sgn}\,(I-T^*T)$, $J_{T*} = \mathrm{sgn}\,(I-TT^*)$, $Q_T = |I-T^*T|^{1/2}$, $Q_{T*} = |I-TT^*|^{1/2}$, and $\mathfrak{D}_T = J_T\mathfrak{H}$. $\Theta_T(\lambda)$ is defined for those complex numbers λ for which $I-\lambda T^*$ is boundedly invertible, and takes values which are continuous operators from $\mathfrak{D}_T$ to the space $\mathfrak{D}_{T*} = J_{T*}\mathfrak{H}$.

The characteristic function of a contraction T appears in the work of Sz.-Nagy and Foiaş [13] as the Fourier representation of a projection in the space of the unitary dilation of T. From this representation is obtained a functional model for T in terms of Θ_T. If $\Theta(\lambda)\colon \mathfrak{D} \to \mathfrak{D}_*$ is an operator valued analytic function, then a contraction T can be constructed (using the same type of functional model) such that $\Theta = \Theta_T$, if and only if Θ is purely contractive, i.e., if $\|\Theta(\lambda)a\| < \|a\|$ whenever $|\lambda| < 1$, $a \in \mathfrak{D}$, $a \neq 0$.

In this paper we also consider operator valued analytic functions $\Theta(\lambda)\colon \mathfrak{D} \to \mathfrak{D}_*$, but $\mathfrak{D}$ and $\mathfrak{D}_*$ are *Krein spaces* rather than Hilbert spaces (see Sec. 2 below), and thus the inner product is not assumed to be positive definite. We show that $\Theta \doteq \Theta_T$ for some bounded operator T if and only if Θ is purely contractive and

Received October 11, 1977, in revised form June 24, 1978.

The results in this paper first appeared in the author's doctoral thesis [12] at the University of Toronto, under the supervision of Ch. Davis. Additional work on the paper, done at the University of Georgia, Athens, Georgia and the University of Missouri, Rolla, Missouri, was partially supported by an NSF grant.

fundamentally reducible, with respect to the indefinite inner products of $\mathfrak{D}$ and $\mathfrak{D}_*$ (see Sec. 3).

There have been previous papers (Brodskiĭ [3], Brodskiĭ, Gohberg, and Kreĭn [4], [5], Clark [6], and Ball [1]) giving necessary and sufficient conditions for an operator valued analytic function to be a characteristic function, but the conditions lack the simplicity of those presented here. In each paper, it is required that a certain function have an analytic extension to the unit disk (cf. Sec. 7); this requirement is eliminated here. In [6] and [1] it is also shown that the positive definiteness of certain functions is necessary and sufficient for Θ to be a characteristic function. We use this kind of result from [1] to prove our theorem, by showing that if Θ is a purely contractive fundamentally reducible analytic function, taking values which are operators between Krein spaces, then the associated kernel matrix function is positive definite (Theorem 3).

2. Krein spaces

An *indefinite inner product space* is a complex vector space $\mathfrak{K}$ on which is defined an inner product $[.,.]$ that is not assumed to be positive, i.e., it is possible for $[x, x]$ to be negative for some $x \in \mathfrak{K}$. We call $\mathfrak{K}$ a *Krein space* if there is an operator J on $\mathfrak{K}$ such that $J^2=I$, $J=J^*$, (i.e., $[Jx, y]=[x, Jy]$), and the *J-inner product*

$$(x, y) = [Jx, y] \quad (x, y \in \mathfrak{K}) \tag{2.1}$$

makes $\mathfrak{K}$ a Hilbert space. Such an operator J is called a *fundamental symmetry*. (See [2, Chapter V].)

The spaces $\mathfrak{D}_T = J_T\mathfrak{H}$ and $D_{T*} = J_{T*}\mathfrak{H}$, considered in Sec. 1, are Krein spaces with the indefinite inner products

$$[x, y] = (J_T x, y) \quad (x, y \in \mathfrak{D}_T)$$

and

$$[x, y] = (J_{T*} x, y) \quad (x, y \in \mathfrak{D}_{T*}).$$

(Here $(.,.)$ denotes the inner product on the Hilbert space $\mathfrak{H}$.) Clearly, J_T and J_{T*} are fundamental symmetries on $\mathfrak{D}_T$ and $\mathfrak{D}_{T*}$, respectively. We will always consider $\mathfrak{D}_T$ and $\mathfrak{D}_{T*}$ as Krein spaces, rather than as subspaces of the Hilbert space $\mathfrak{H}$.

In Krein spaces, the emphasis is always on the indefinite inner product, with the *J-norm* $\|x\|_J = [Jx, x]^{1/2}$ serving mainly to define the topology (the *strong* topology). Accordingly, if A is a continuous operator between Krein spaces $\mathfrak{K}$ and $\mathfrak{K}'$, we use A^* to denote the adjoint of A with respect to the indefinite inner products. If J and J' are fundamental symmetries on $\mathfrak{K}$ and $\mathfrak{K}'$, respectively, then the adjoint of A with respect to the J- and J'-inner products (2.1) will be denoted by $A^{(*)}$ and is given by $A^{(*)}=JA^*J'$ (see [2, Sec. VI. 2]).

Different fundamental symmetries J on a Krein space $\mathfrak{K}$ define different J-norms, but the strong topologies obtained coincide (see [12, Sec. 1.4] and [2]). Thus we can talk about *the* strong topology on a Krein space.

We will be needing the following:

Lemma 1. *If J is a symmetry on a Krein space $\mathfrak{K}$ (i.e., $J^2=I$, $J=J^*$) such that the J-inner product $(x, y)=[Jx, y]$ is positive, then J is a fundamental symmetry, i.e., the J-inner product makes $\mathfrak{K}$ a Hilbert space.*

Proof. See [2, Corollary V.1.2].

3. Purely contractive analytic functions. The main theorem

An *operator valued analytic function* is a function θ which is defined and analytic in D, the open unit disk in the complex plane, and which takes values that are continuous operators from a Krein space $\mathfrak{D}$ to a Krein space $\mathfrak{D}_*$. Θ is said to be *purely contractive* if, for each $\lambda\in D$,

$$\text{(3.1)}\qquad [\Theta(\lambda)a, \Theta(\lambda)a] < [a, a] \quad (a\in\mathfrak{D},\ a\neq 0)$$

and

$$\text{(3.2)}\qquad [\Theta(\lambda)^* a_*, \Theta(\lambda)^* a_*] < [a_*, a_*] \quad (a_*\in\mathfrak{D}_*,\ a_*\neq 0).$$

Remarks. In Hilbert space, (3.2) is implied by (3.1), but this is not true in general. (See, for example [9, Sec. 3]. Note that the results in [9] concerning expansive operators apply to contractive operators, with the inner product $-[.\,,.]$.) Also, in Hilbert space, it is not necessary to assume that the inequalities (3.1) and (3.2) are strict inequalities, except at $\lambda=0$, since the maximum modulus principle can then be used to get strict inequality for all $\lambda\in D$ (cf. [13, Sec. V.2.2]).

An operator A on a Krein space $\mathfrak{K}$ is said to be *fundamentally reducible* if there is a fundamental symmetry on $\mathfrak{K}$ commuting with A [2, Sec. VIII.1]. Suppose Θ is an operator valued analytic function, and let $\Theta_0=\Theta(0)$. We call the function Θ *fundamentally reducible* if the operators $\Theta_0^*\Theta_0$ (on $\mathfrak{D}$) and $\Theta_0\Theta_0^*$ (on $\mathfrak{D}_*$) are fundamentally reducible.

If $\mathfrak{D}$ and $\mathfrak{D}'$ are two Krein spaces, then an operator $\tau\colon\mathfrak{D}\to\mathfrak{D}'$ is said to be *unitary* if it is continuous and invertible, and if $[\tau x, \tau x]=[x, x]$ for all $x\in\mathfrak{D}$. Two operator valued analytic functions $\Theta(\lambda)\colon\mathfrak{D}\to\mathfrak{D}_*$ and $\Theta'(\lambda)\colon\mathfrak{D}'\to\mathfrak{D}'_*$ are said to *coincide* if there are unitary operators $\tau\colon\mathfrak{D}\to\mathfrak{D}'$ and $\tau_*\colon\mathfrak{D}_*\to\mathfrak{D}'_*$ such that $\Theta'(\lambda)=\tau_*\Theta(\lambda)\tau^{-1}$ for all $\lambda\in D$.

We can now state the main result of this paper.

Theorem 1. *Suppose Θ is an operator valued analytic function. For Θ to coincide with the characteristic function of a bounded operator on a Hilbert space, it is necessary and sufficient that Θ be purely contractive and fundamentally reducible.*

The condition that Θ be fundamentally reducible can not be omitted from Theorem 1, as P. Jonas of Berlin has constructed an example of a (constant) purely contractive analytic function which does not coincide with the characteristic function of any bounded operator on a Hilbert space. The author is indebted to J. Bognár and B. Sz.-Nagy for pointing out the existence of this example.

4. Proof of necessity in Theorem 1

We assume here that Θ is an operator valued analytic function coinciding with Θ_T, for some operator T. Thus $\Theta_T(\lambda)$ is defined and analytic in the open unit disk D, and takes values which are continuous operators between the Krein spaces $\mathfrak{D}_T$ and $\mathfrak{D}_{T*}$. To prove that Θ is purely contractive and fundamentally reducible, it clearly suffices to show that Θ_T is purely contractive and fundamentally reducible.

Recalling the definitions of the indefinite inner products on $\mathfrak{D}_T$ and $\mathfrak{D}_{T*}$, we obtain

$$\Theta_T(\lambda)^* = [-T^*J_{T^*}+\bar{\lambda}Q_T(I-\bar{\lambda}T)^{-1}Q_{T^*}]|\mathfrak{D}_{T^*}, \tag{4.1}$$

and it follows, similarly to [13], relation (VI.1.4) (cf. [11]), that for $\lambda\in D$

$$I-\Theta_T(\lambda)^*\Theta_T(\lambda) = (1-|\lambda|^2)Q_T(I-\bar{\lambda}T)^{-1}(I-\lambda T^*)^{-1}J_TQ_T$$

and

$$I-\Theta_T(\lambda)\Theta_T(\lambda)^* = (1-|\lambda|^2)J_{T^*}Q_{T^*}(I-\lambda T^*)^{-1}(I-\bar{\lambda}T)^{-1}Q_{T^*}.$$

Consequently, for $\lambda\in D$,

$$[(I-\Theta_T(\lambda)^*\Theta_T(\lambda))a, a] = (1-|\lambda|^2)\|(I-\lambda T^*)^{-1}J_TQ_Ta\|^2 > 0,$$

where $a\in\mathfrak{D}_T$, $a\neq 0$; and

$$[(I-\Theta_T(\lambda)\Theta_T(\lambda)^*)a_*, a_*] = (1-|\lambda|^2)\|(I-\bar{\lambda}T)^{-1}Q_{T^*}a_*\|^2 > 0,$$

where $a_*\in\mathfrak{D}_{T*}$, $a_*\neq 0$. Therefore, Θ_T is purely contractive.

Since $\Theta_T(0)^*\Theta_T(0)=T^*T$, it follows that $\Theta_T(0)^*\Theta_T(0)$ commutes with the fundamental symmetry $J_T=\operatorname{sgn}(I-T^*T)$ on $\mathfrak{D}_T$. Likewise, $\Theta_T(0)\Theta_T(0)^*=TT^*$ commutes with the fundamental symmetry $J_{T*}=\operatorname{sgn}(I-TT^*)$ on $\mathfrak{D}_{T*}$. Consequently, Θ_T is fundamentally reducible.

The proof of sufficiency in Theorem 1 occupies the next four sections.

5. Fundamental symmetries on $\mathfrak{D}$ and $\mathfrak{D}_*$

Suppose Θ is a purely contractive, fundamentally reducible analytic function. Let J_0 be a fundamental symmetry on $\mathfrak{D}$ commuting with $\Theta_0^* \Theta_0$. Then the operator $I-\Theta_0^*\Theta_0$ is not only self-adjoint but also J_0-self-adjoint (i.e., self-adjoint with respect to the J_0-inner product). Thus we can define

$$J = \operatorname{sgn}(I-\Theta_0^* \Theta_0) \quad \text{and} \quad Q = |I-\Theta_0^* \Theta_0|^{1/2},$$

where these are computed using the J_0-self-adjoint functional calculus on the Hilbert space $\mathfrak{D}$ with the J_0-inner product. J and Q commute with J_0 and hence are self-adjoint (as well as J_0-self-adjoint).

Since Θ is assumed to be purely contractive, $I-\Theta_0^*\Theta_0$ is injective. Therefore we conclude that $J^2=I$ and Q has range dense in $\mathfrak{D}$. Also, we have

$$[JQh, Qh] = [(I-\Theta_0^* \Theta_0)h, h] \geqq 0 \quad (h \in \mathfrak{D}),$$

and consequently $[Jh, h] \geqq 0$ for all $h \in \mathfrak{D}$. It follows from Lemma 1 that J is a fundamental symmetry on $\mathfrak{D}$.

Similarly, if J_{0*} is a fundamental symmetry on $\mathfrak{D}_*$ commuting with $\Theta_0\Theta_0^*$, and if we define

$$J_* = \operatorname{sgn}(I-\Theta_0\Theta_0^*),$$

using the J_{0*}-self-adjoint functional calculus on $\mathfrak{D}_*$, then J_* is a fundamental symmetry on $\mathfrak{D}_*$.

The operators J and J_* do not depend on the choice of fundamental symmetries J_0 and J_{0*}. Indeed, suppose J_1 is another fundamental symmetry on $\mathfrak{D}$ commuting with $\Theta_0^*\Theta_0$, and let $J'=\operatorname{sgn}(I-\Theta_0^*\Theta_0)$ and $Q'=|I-\Theta_0^*\Theta_0|^{1/2}$ be computed with respect to the J_1-inner product. Then J' and Q' are also J_0-self-adjoint, and $JQ^2=J'Q'^2$. Since both J and J' are J_0-unitary, and both Q^2 and Q'^2 are J_0-positive, the uniqueness of the J_0-polar decomposition of JQ^2 implies $J=J'$. The argument for J_* is similar.

Proposition 1. *Θ_0^* is the same as $\Theta_0^{(*)}$, the adjoint of Θ_0 with respect to the J- and J_*-inner products on $\mathfrak{D}$ and $\mathfrak{D}_*$.*

Proof. $\Theta_0^{(*)}=J\Theta_0^*J_*=J(J\Theta_0^*)=\Theta_0^*$ (cf. [7, Sec. 2]).

Although Clark [6] and Ball [1] do not approach the subject of fundamental symmetries on $\mathfrak{D}$ and $\mathfrak{D}_*$ in the same manner as we have done here, the end result is the same. Their approach is to begin with $\mathfrak{D}$ and $\mathfrak{D}_*$ as Hilbert spaces and subsequently impose on them the Krein space structures derived from the symmetries $J=\operatorname{sgn}(I-\Theta_0^*\Theta_0)$ and $J_*=\operatorname{sgn}(I-\Theta_0\Theta_0^*)$. Thus it is implicit in the definitions of the inner products in [6] and [1] that Θ is to be taken to be fundamentally reducible.

BRODSKIĬ [3], and BRODSKIĬ, GOHBERG, and KREĬN [4], [5] consider $\mathfrak{D}$ and $\mathfrak{D}_*$ as Hilbert spaces, but do not, however, assume $J=\operatorname{sgn}(I-\Theta_0^*\Theta_0)$. Instead, they deal with a more general situation in which the object studied is not a single operator T but an *uzel*, a collection of operators and Hilbert spaces. A $\mathscr{U}$-uzel is a collection of spaces $\mathfrak{H}$, $\mathfrak{G}$ and operators T, R, J, where $T\colon \mathfrak{H}\to\mathfrak{H}$, $R\colon \mathfrak{G}\to\mathfrak{H}$, J is a symmetry on $\mathfrak{G}$, and $I-T^*T=RJR^*$. The particular case of interest to us is when $\mathfrak{G}=\mathfrak{D}_T$, $R=Q_T$, and $J=J_T$.

6. The theorem of Ball

We wish to apply [1, Theorem 2], but some differences in notation need to be cleared up first. Let us define $\bar{\Theta}(\lambda)=\Theta(\bar{\lambda})^*$. In [1], the characteristic function B_T is $B_T=\bar{\Theta}_T$ (cf. (4.1)), and so the condition given in [1, Theorem 2] for B to be a characteristic function must be written with $B=\bar{\Theta}$. Also, operators in [1] are assumed to act between Hilbert spaces, whereas here we consider them as acting between Krein spaces. The concept of adjoint must therefore be interpreted appropriately. Proposition 1 shows that the definitions of J and J_* given here are the same as those in [1] (with $B=\bar{\Theta}$).

It should be noted that in [1, Theorem 2] it is being asserted that $\tau B=B_T\tau_*$ for some operator T, where τ and τ_* are unitary operators between Hilbert spaces. Since $J=\operatorname{sgn}(I-B(0)B(0)^*)$ and $J_T=\operatorname{sgn}(I-B_T(0)B_T(0)^*)$, it follows that $\tau J=J_T\tau$. Similarly, $\tau_*J_*=J_{T*}\tau_*$, and we deduce that τ and τ_* are also unitary operators between Krein spaces. Thus we have coincidence of B and B_T in the sense of Sec. 3.

We can now state Ball's theorem, using our notation:

Theorem 2. ([1, Theorem 2]) *Let $\Theta(\lambda)\colon \mathfrak{D}\to\mathfrak{D}_*$ be an operator valued analytic function. Then θ coincides with the characteristic function of some bounded operator on a Hilbert space if and only if*

(i) *Θ is fundamentally reducible,*

(ii) *$I-\Theta_0^*\Theta_0$ and $I-\Theta_0\Theta_0^*$ are injective, and*

(iii) *the operator matrix*

$$(6.1)\qquad k(\mu,\lambda)=\begin{pmatrix}(1-\lambda\bar{\mu})^{-1}(I-\Theta(\mu)^*\Theta(\lambda)) & (\lambda-\bar{\mu})^{-1}(\Theta(\bar{\lambda})^*-\Theta(\mu)^*)\\ (\lambda-\bar{\mu})^{-1}(\Theta(\lambda)-\Theta(\bar{\mu})) & (1-\lambda\bar{\mu})^{-1}(I-\Theta(\bar{\mu})\Theta(\bar{\lambda})^*)\end{pmatrix}$$

is positive definite on some neighborhood $\mathscr{U}$ of zero, i.e.,

$$(6.2)\qquad \sum_{i=1}^{n}\sum_{j=1}^{n}[k(\mu_j,\mu_i)(f_i\oplus g_i),(f_j\oplus g_j)]\geqq 0$$

for all $n\geqq 1$ and for all $f_i\in\mathfrak{D}$, $g_i\in\mathfrak{D}_*$, and $\mu_i\in\mathscr{U}$ $(i=1,2,\ldots,n)$.

Note. $k(\mu, \lambda)$ is considered as an operator on the Krein space $\mathfrak{D}\oplus\mathfrak{D}_*$. The matrix in [1] is obtained by setting $B=\bar{\Theta}$ and considering $(I\oplus J_*)k(\bar{z}, \bar{w})(J\oplus I)$. The neighborhood D considered in [1] is denoted by $\mathcal{U}$ in the above theorem. (In this paper, D denotes the open unit disk.)

7. The functions Φ and Ω

In view of Theorem 2, it remains to show that (6.2) is valid whenever Θ is purely contractive and fundamentally reducible.

If we consider, for the moment, $\mathfrak{D}$ and $\mathfrak{D}_*$ as Hilbert spaces, with the J- and J_*-inner products, then Proposition 1 shows that Θ_0 has a polar decomposition

$$\Theta_0 = U(\Theta_0^*\Theta_0)^{1/2} = (\Theta_0\Theta_0^*)^{1/2}U. \tag{7.1}$$

(This decomposition can also be done in Krein space. See [10].) Following the argument used in [1, Sec. 1.3 and Sec. 2.1], we can assume that the codimension of $(\Theta_0^*\Theta_0)^{1/2}\mathfrak{D}$, in the Hilbert space $\mathfrak{D}$, equals the codimension of $(\Theta_0\Theta_0^*)^{1/2}\mathfrak{D}_*$, in the Hilbert space $\mathfrak{D}_*$. Then U can be chosen to be a unitary operator between the Hilbert spaces $\mathfrak{D}$ and $\mathfrak{D}_*$. From (7.1) we have $U(\Theta_0^*\Theta_0)=(\Theta_0\Theta_0^*)U$, and hence $UJ=J_*U$. Thus U is also a unitary operator between the Krein spaces $\mathfrak{D}$ and $\mathfrak{D}_*$.

If the vector $f\in\mathfrak{D}$ satisfies $(I+U^*\Theta(\lambda))f=0$, for some $\lambda\in D$, then it follows that

$$[f,f] = [U^*\Theta(\lambda)f, U^*\Theta(\lambda)f] = [\Theta(\lambda)f, \Theta(\lambda)f].$$

Hence, since Θ is purely contractive, $f=0$ and we conclude that $I+U^*\Theta(\lambda)$ is injective for all $\lambda\in D$. Similarly, $I+\Theta(\lambda)^*U$ is injective and thus $I+U^*\Theta(\lambda)$ has range dense in $\mathfrak{D}$, for all $\lambda\in D$. Hence, for each $\lambda\in D$, we can define

$$\Phi(\lambda) = \big(I-U^*\Theta(\lambda)\big)\big(I+U^*\Theta(\lambda)\big)^{-1}J, \tag{7.2}$$

an operator with domain dense in $\mathfrak{D}$.

Note that the operator $I+U^*\Theta_0=I+(\Theta_0^*\Theta_0)^{1/2}$ is boundedly invertible, and hence $(I+U^*\Theta(\lambda))^{-1}$ is analytic for all λ belonging to some neighborhood $\mathcal{U}$ of zero, with $\mathcal{U}\subset D$ (cf. [8, Sec. VII.1.1]). Consequently, $\Phi(\lambda)$ is analytic for $\lambda\in\mathcal{U}$. Clearly, we can assume that $\mathcal{U}$ is closed under complex conjugation.

Remark. In [3], [4], [5], [6], and [1] assumptions are made which amount to assuming that $\Phi(\lambda)$ $(\lambda\in\mathcal{U})$ extends to an analytic function on D. However, it is not necessary to make this assumption here.

We obtain from (7.2) that, for each $\lambda\in D$,

$$\Phi(\lambda)J = 2\big(I+U^*\Theta(\lambda)\big)^{-1}-I, \tag{7.3}$$

and thus $\Phi(\lambda)J$ is closed. Consequently, for each $\lambda \in D$, $\Phi(\lambda)$ is closed. Also, (7.3) implies that

$$\Phi(\lambda)^{(*)}J = 2(I+\Theta(\lambda)^*U)^{-1} - I$$

(recall the notation introduced in Sec. 2), and hence

$$\Phi(\lambda)^{(*)} = (I-\Theta(\lambda)^*U)(I+\Theta(\lambda)^*U)^{-1}J. \tag{7.4}$$

For $\lambda \in D$, an arbitrary vector in the domain of $\Phi(\lambda)$ is of the form $f = J(I+U^*\Theta(\lambda))g$, where $g \in \mathfrak{D}$. If $(.\,,.)$ denotes the J-inner product on $\mathfrak{D}$, then we can readily deduce that

$$\operatorname{Re}(\Phi(\lambda)f, f) = [g, g] - [\Theta(\lambda)g, \Theta(\lambda)g] \geqq 0. \tag{7.5}$$

It follows that

$$\|(I+\Phi(\lambda))f\|_J^2 \geqq \|f\|_J^2 + \|\Phi(\lambda)f\|_J^2 \geqq \|(I-\Phi(\lambda))f\|_J^2, \tag{7.6}$$

and hence $I+\Phi(\lambda)$ is injective with closed range.

By a similar argument, we deduce from (7.4) that $I+\Phi(\lambda)^{(*)}$ is injective, and thus $I+\Phi(\lambda)$ has dense range. Therefore, the operator $I+\Phi(\lambda)$ is bijective and closed, and consequently (by the closed graph theorem) boundedly invertible (for all $\lambda \in D$).

We can now make the definition

$$\Omega(\lambda) = (I-\Phi(\lambda))(I+\Phi(\lambda))^{-1} \quad (\lambda \in D).$$

By the preceding paragraph, $\Omega(\lambda)$ has domain equal to $\mathfrak{D}$ and, by (7.6), $\Omega(\lambda)$ is a contraction on the Hilbert space $\mathfrak{D}$ with the J-inner product. Since Θ is purely contractive, equality holds in (7.5) only if $f=0$, and hence the same is true in (7.6). It then readily follows that Ω is purely contractive.

$\Phi(\lambda)$ is known to be analytic only in a neighborhood of zero, and thus the analyticity of $\Omega(\lambda)$ in D is not immediately obvious. We can write

$$I+\Phi(\lambda) = \Psi(\lambda)(I+U^*\Theta(\lambda))^{-1}J, \tag{7.7}$$

where $\Psi(\lambda) = J(I+U^*\Theta(\lambda)) + (I-U^*\Theta(\lambda))$. Since $I+\Phi(\lambda)$ is boundedly invertible, (7.7) implies that $\Psi(\lambda)$ is boundedly invertible. Since $\Psi(\lambda)$ is analytic in D, so is $\Psi(\lambda)^{-1}$ (see [8, Sec. VII.1.1]), and therefore the function

$$(I+\Phi(\lambda))^{-1} = J(I+U^*\Theta(\lambda))\Psi(\lambda)^{-1}$$

is analytic in D. Finally, we note that

$$\Omega(\lambda) = 2(I+\Phi(\lambda))^{-1} - I \quad (\lambda \in D),$$

and thus $\Omega(\lambda)$ is analytic in D.

We have now shown that Ω is a purely contractive analytic function on the Hilbert space $\mathfrak{D}$ (with the J-inner product) and thus (by [13, Theorem VI.3.1]) Ω coincides with the characteristic function of some contraction S, i.e., $\tau_* \Omega(\lambda) =$

$=\Theta_S(\lambda)\tau$ for some unitary operators $\tau:\mathfrak{D}\to\mathfrak{D}_S$ and $\tau_*:\mathfrak{D}\to\mathfrak{D}_{S*}$ ($\mathfrak{D}$ considered as a Hilbert space). It then follows that, for each $n\geqq 1$, and for $a_i, b_i\in\mathfrak{D}$, $\mu_i\in D$ $(i=1, 2, \ldots, n)$ we have (using the J-inner product)

(7.8)

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\{((1-\mu_i\bar\mu_j)^{-1}(I-\Omega(\mu_j)^{(*)}\Omega(\mu_i))a_i, a_j)+((\mu_i-\bar\mu_j)^{-1}(\Omega(\mu_i)-\Omega(\bar\mu_j))a_i, b_j)+$$

$$+((\mu_i-\bar\mu_j)^{-1}(\Omega(\bar\mu_i)^{(*)}-\Omega(\mu_j)^{(*)})b_i, a_j)+((1-\mu_i\bar\mu_j)^{-1}(I-\Omega(\bar\mu_j)\Omega(\bar\mu_i)^{(*)})b_i, b_j)\}=$$

$$=\left\|\sum_{i=1}^{n}\{(I-\mu_iS^*)^{-1}Q_S\tau a_i+(I-\mu_iS)^{-1}Q_{S*}\tau_*b_i\}\right\|^2\geqq 0$$

(cf. [1, Sec. 1.5]).

8. Positive definiteness of $k(\mu, \lambda)$

We now prove a series of results that will enable us to write (6.2) in the form (7.8) and thus establish the positive definiteness of $k(\mu, \lambda)$.

Proposition 2. *For $\lambda\in D$, $I+\Phi(\lambda)J$ is boundedly invertible and $U^*\Theta(\lambda)=(I-\Phi(\lambda)J)(I+\Phi(\lambda)J)^{-1}$.*

Proof. The equation $I+\Phi(\lambda)J=2(I+U^*\Theta(\lambda))^{-1}$, obtained from (7.3), shows that $I+\Phi(\lambda)J$ is boundedly invertible and that

$$U^*\Theta(\lambda)(I+\Phi(\lambda)J)=[(I+U^*\Theta(\lambda))-(I-U^*\Theta(\lambda))](I+U^*\Theta(\lambda))^{-1}=I-\Phi(\lambda)J.$$

Lemma 2. *Suppose that A_1 and A_2 are bounded operators with $I+A_i$ boundedly invertible $(i=1, 2)$. Then if $B_i=(I-A_i)(I+A_i)^{-1}$ $(i=1, 2)$, we have*

$$I-B_1B_2=2(I+A_1)^{-1}(A_1+A_2)(I+A_2)^{-1}$$

and

$$B_2-B_1=2(I+A_1)^{-1}(A_1-A_2)(I+A_2)^{-1}.$$

Proof. We simply perform the calculations:

$$I-B_1B_2=(I+A_1)^{-1}[(I+A_1)(I+A_2)-(I-A_1)(I-A_2)](I+A_2)^{-1}=$$
$$=2(I+A_1)^{-1}(A_1+A_2)(I+A_2)^{-1};$$

$$B_2-B_1=(I+A_1)^{-1}[(I+A_1)(I-A_2)-(I-A_1)(I+A_2)](I+A_2)^{-1}=$$
$$=2(I+A_1)^{-1}(A_1-A_2)(I+A_2)^{-1}.$$

In the following, $\mathscr{U}$ is the neighborhood of zero where $\Phi(\lambda)$ is analytic (Sec. 7).

Lemma 3. *For $\lambda, \mu \in \mathscr{U}$ we have*

(i) $I-\Theta(\mu)^*\Theta(\lambda)=$

$$= (I+J\Phi(\mu)^*)^{-1}(I+\Phi(\mu)^{(*)})[I-\Omega(\mu)^{(*)}\Omega(\lambda)](I+\Phi(\lambda))J(I+\Phi(\lambda)J)^{-1},$$

(ii) $\Theta(\lambda)-\Theta(\bar{\mu})=$

$$= U(I+\Phi(\bar{\mu})J)^{-1}(I+\Phi(\bar{\mu}))[\Omega(\lambda)-\Omega(\bar{\mu})](I+\Phi(\lambda))J(I+\Phi(\lambda)J)^{-1},$$

(iii) $\Theta(\bar{\lambda})^*-\Theta(\mu)^*=$

$$= (I+J\Phi(\mu)^*)^{-1}(I+\Phi(\mu)^{(*)})[\Omega(\bar{\lambda})^{(*)}-\Omega(\mu)^{(*)}](I+\Phi(\bar{\lambda})^{(*)})J(I+\Phi(\bar{\lambda})^*)^{-1}U^*,$$

(iv) $I-\Theta(\bar{\mu})\Theta(\bar{\lambda})^*=$

$$= U(I+\Phi(\bar{\mu})J)^{-1}(I+\Phi(\bar{\mu}))[I-\Omega(\bar{\mu})\Omega(\bar{\lambda})^{(*)}](I+\Phi(\bar{\lambda})^{(*)})J(I+J\Phi(\bar{\lambda})^*)^{-1}U^*.$$

Proof. Since $\mathscr{U}$ is assumed to be self-conjugate, both $\Phi(\lambda)$ and $\Phi(\bar{\lambda})$ are bounded for $\lambda\in\mathscr{U}$. It follows from Sec. 7 and Proposition 2 that all operators appearing above are bounded.

We know by Proposition 2 that

$$U^*\Theta(\lambda) = (I-\Phi(\lambda)J)(I+\Phi(\lambda)J)^{-1},$$

and we obtain from this (and the adjoint relation), by means of Lemma 2, the equations

$$I-\Theta(\mu)^*\Theta(\lambda) = 2(I+J\Phi(\mu)^*)^{-1}(J\Phi(\mu)^*+\Phi(\lambda)J)(I+\Phi(\lambda)J)^{-1} =$$

$$= 2(I+J\Phi(\mu)^*)^{-1}(\Phi(\mu)^{(*)}+\Phi(\lambda))J(I+\Phi(\lambda)J)^{-1}.$$

We also have $\Omega(\lambda)=(I-\Phi(\lambda))(I+\Phi(\lambda))^{-1}$ and hence, using Lemma 2 again, we obtain

$$I-\Omega(\mu)^{(*)}\Omega(\lambda) = 2(I+\Phi(\mu)^{(*)})^{-1}(\Phi(\mu)^{(*)}+\Phi(\lambda))(I+\Phi(\lambda))^{-1}.$$

Combining these two results gives (i).

Equations (ii), (iii), and (iv) are proved similarly.

Theorem 3. *$k(\mu, \lambda)$ is positive definite on a neighborhood $\mathscr{U}$ of zero.*

Proof. Let $\mathcal{U}$ be the neighborhood of zero on which $\Phi(\lambda)$ is analytic. Then for each $n \geqq 1$, and for $f_i \in \mathfrak{D}$, $g_i \in \mathfrak{D}_*$, and $\mu_i \in \mathcal{U}$ $(i=1, 2, \dots, n)$ we have (from (6.1))

$$\sum_{i=1}^{n} \sum_{j=1}^{n} [k(\mu_j, \mu_i)(f_i \oplus g_i), (f_j \oplus g_j)] = \tag{8.1}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \{[(1-\mu_i\bar{\mu}_j)^{-1}(I-\Theta(\mu_j)^*\Theta(\mu_i))f_i, f_j] + [(\mu_i-\bar{\mu}_j)^{-1}(\Theta(\mu_i)-\Theta(\bar{\mu}_j))f_i, g_j] +$$

$$+[(\mu_i-\bar{\mu}_j)^{-1}(\Theta(\bar{\mu}_i)^*-\Theta(\mu_j)^*)g_i, f_j] + [(1-\mu_i\bar{\mu}_j)^{-1}(I-\Theta(\bar{\mu}_j)\Theta(\bar{\mu}_i)^*)g_i, g_j]\}.$$

By Lemma 3,

$$[(I-\Theta(\mu_j)^*\Theta(\mu_i))f_i, f_j] = \tag{8.2}$$

$$= [(I+J\Phi(\mu_j)^*)^{-1}(I+\Phi(\mu_j)^{(*)})(I-\Omega(\mu_j)^{(*)}\Omega(\mu_i))(I+\Phi(\mu_i))J(I+\Phi(\mu_i)J)^{-1}f_i, f_j] =$$

$$= [(I+\Phi(\mu_j)^{(*)})(I-\Omega(\mu_j)^{(*)}\Omega(\mu_i))(I+\Phi(\mu_i))J(I+\Phi(\mu_i)J)^{-1}f_i, (I+\Phi(\mu_j)J)^{-1}f_j] =$$

$$= (J(I+\Phi(\mu_j)^{(*)})(I-\Omega(\mu_j)^{(*)}\Omega(\mu_i))(I+\Phi(\mu_i))J(I+\Phi(\mu_i)J)^{-1}f_i, (I+\Phi(\mu_j)J)^{-1}f_j) =$$

$$= ((I-\Omega(\mu_j)^{(*)}\Omega(\mu_i))a_i, a_j),$$

where $a_i = (I+\Phi(\mu_i))J(I+\Phi(\mu_i)J)^{-1}f_i$ $(i=1, 2, \dots, n)$.

We also make the definition

$$b_i = (I+\Phi(\bar{\mu}_i)^{(*)})J(I+J\Phi(\bar{\mu}_i)^*)^{-1}U^*g_i \quad (i = 1, 2, \dots, n).$$

Then, by applying Lemma 3 to the terms of (8.1) in the same manner as (8.2) above, we deduce that (8.1) is the same as (7.8). Therefore, $k(\mu, \lambda)$ is positive definite on $\mathcal{U}$.

Theorem 3, in conjunction with Theorem 2, completes the proof of Theorem 1.

9. Conclusion

Theorem 1 establishes for certain non-contractions a result which generalizes a result of Sz.-Nagy and Foiaş for contractions. The proof, however, does not make use of a Sz.-Nagy and Foiaş type construction of the operator model, but instead relies on the much less geometric model obtained by Ball. When Θ is bounded, i.e., $\sup_{\lambda \in D} \|\Theta(\lambda)\| < \infty$, a model can be obtained that closely resembles the Sz.-Nagy and Foiaş model (see [12]), and this will be the subject of a future paper.

References

[1] J. A. BALL, Models for non-contractions, *J. Math. Anal. Appl.*, **52** (1975), 235—254.

[2] J. BOGNÁR, *Indefinite inner product spaces*, Springer-Verlag (New York, 1974).

[3] V. M. BRODSKIĬ, On operator nodes and their characteristic functions, *Soviet Math. Dokl.*, **12** (1971), 696—700. (Translated from *Dokl. Akad. Nauk SSSR*, **198** (1971), 16—19.)

[4] V. M. BRODSKIĬ, I. C. GOHBERG and M. G. KREĬN, Determination and fundamental properties of the characteristic function of a $\mathcal{Y}$-node, *Functional Anal. Appl.*, **4** (1970), 78—80. (Translated from *Funk. Analiz i ego Pril.*, **4**:1 (1970), 88—90.)

[5] V. M. BRODSKIĬ, I. C. GOHBERG and M. G. KREĬN, On characteristic functions of an invertible operator (In Russian), *Acta Sci. Math.*, **32** (1971), 141—164.

[6] D. N. CLARK, On models for non-contractions, *Acta Sci. Math.*, **36** (1974), 5—16.

[7] CH. DAVIS, *J*-unitary dilation of a general operator, *Acta Sci. Math.*, **31** (1970), 75—86.

[8] T. KATO, *Perturbation theory for linear operators*, Springer-Verlag (New York, 1966).

[9] M. G. KREĬN and JU. L. ŠMUL'JAN, Plus-operators in a space with indefinite metric, *Amer. Math. Soc. Transl.*, (2) **85** (1969), 93—113. (Translated from *Mat. Issled.*, **1** (1966), No. 1, 131—161.)

[10] M. G. KREĬN and JU. L. ŠMUL'JAN, *J*-polar representations of plus-operators, *Amer. Math. Soc. Transl.*, (2) **85** (1969), 115—143. (Translated from *Mat. Issled.*, **1** (1966), No. 1, 172—210.)

[11] O. V. KUŽEL', The characteristic operator-function of an arbitrary bounded operator, *Amer. Math. Soc. Transl.*, (2) **90** (1970), 225—228. (Translated from *Dopovidi Akad. Nauk. Ukraïn. RSR*, Ser. A 1968, 233—235.)

[12] B. W. MCENNIS, *Characteristic functions and the geometry of dilation spaces*, Ph. D. thesis, University of Toronto (1977).

[13] B. SZ.-NAGY and C. FOIAŞ, *Harmonic analysis of operators on Hilbert space*, North Holland (Amsterdam—London, 1970).

DEPARTMENT OF MATHEMATICS
THO OHIO STATE UNIVERSITY
MARION CAMPUS
1465 MT. VERNON AVENUE
MARIO, OHIO 43302
USA

Acta Sci. Math., 41 (1979), 173—185

Über die Veränderung der Länge der Vektoren in Weyl–Otsukischen Räumen

ARTHUR MOÓR

§ 1. Einleitung

T. OTSUKI entwickelte in seiner Arbeit [4] eine Übertragungstheorie in Punkträumen, die in lokaler Schreibweise dadurch charakterisiert werden kann, daß die *Übertragungsparameter* $'\Gamma_j{}^i{}_k$ *bzw.* $''\Gamma_j{}^i{}_k$ für die ko- bzw. kontravarianten Indizes der Tensoren im allgemeinen voneinander verschieden sind, zwar sind sie miteinander durch eine Relation verbunden (vgl. [4], (3.13)). Außerdem sind die durch diese Übertragungsparameter gebildeten affinen kovarianten Ableitungen mit einem „a priori" angegebenen und die geometrische Struktur des Raumes bestimmenden *Tensor* P^i_j kontrahiert. In [3] bestimmten wir solche Übertragungsparameter für diese Otsukischen Räume, die aus einem metrischen und symmetrischen *Grundtensor* $g_{ij}(x)$ [1] abgeleitet waren, und die für g_{ij} rekurrente kovariante Ableitung bestimmten. Somit vereinigten wir die Weylschen und die Otsukischen Übertragungstheorien (vgl. [4] und [5]) — wir wollen im folgenden diese Räume *Weyl—Otsukische Räume* nennen — und untersuchten in erster Reihe die Eigenschaften der *Eigentensoren* bezüglich des invarianten Differentials.

In den Weylschen Räumen verändert sich bekanntlich die Länge der Vektoren bei einer Parallelverschiebung proportional mit der ursprünglichen Länge (vgl. z. B. [2], § 4, wo aber im Falle der Punkträume durchwegs $\gamma^*_k \equiv 0$ gesetzt werden muß). Wir wollen im folgenden die Veränderung der Länge der Vektoren, ferner die Veränderung gewisser Invarianten der Tensoren zweiter Stufe bei Parallelverschiebung in den Weyl—Otsukischen Räumen untersuchen und die Veränderung, genauer den Differentialquotienten dieser Invarianten nach einem Parameter „t" bestimmen, falls die Parallelverschiebung längs einer vorgegebenen Kurve $C: x^i = x^i(t)$ durch-

Eingegangen am 3. März 1977.

[1]) $x = (x^1, x^2, \ldots, x^n)$ bedeutet jetzt und im folgenden einen Punkt im n-dimensionalen Punktraum.

geführt wird; vgl. Formeln: (3.10), (4.4) und (5.4), die unsere wichtigste Resultate ausdrücken.

Besondere Wichtigkeit haben in diesen Räumen die sog. Eigentensoren. Unsere Sätze 2, 4 und 7 beziehen sich eben auf die Veränderung der charakteristischen Invarianten der Eigenvektoren und Eigentensoren des Raumes bei Parallelverschiebungen längs gewisser Kurven.

§ 2. Fundamentalformeln der Weyl—Otsukischen Räume

Die Grundgrößen eines Weyl—Otsukischen Raumes sind der in (i, k) symmetrische metrische Grundtensor $g_{ik}(x)$, der kovariante Vektor $\gamma_k(x)$ und der gemischte Tensor $P^i_j(x)$, von dem wir im folgenden durchwegs annehmen wollen, daß er der Relation

$$P^r_i g_{rj} = P^r_j g_{ir} \tag{2.1}$$

genügt, d. h. P_{ij} ist in (i, j) symmetrisch. Der inverse Tensor von P^i_j soll durch die Formeln

(2.2a) $\quad Q^i_r P^r_j = \delta^i_j, \qquad$ (2.2b) $\quad Q^r_j P^i_r = \delta^i_j$

festgelegt sein, wo (2.2b) nach der Tensoralgebra — bekanntlich — eine Folgerung von (2.2a) ist. Aus (2.1) folgt leicht, daß neben P_{ij} auch Q_{ij} symmetrisch ist.

Die Übertragungsparameter $''\Gamma_j{}^i{}_k$ bzw. $'\Gamma_j{}^i{}_k$ — die bei der Bildung der kovarianten Ableitung der Tensoren für ko- bzw. kontravariante Indizes verwendet werden — sind durch die Relationen (vgl. (3.13) von [4]):

$$\partial_k P^i_j + ''\Gamma_r{}^i{}_k P^r_j - '\Gamma_j{}^r{}_k P^i_r = 0, \quad \partial_k \equiv \frac{\partial}{\partial x^k} \tag{2.3}$$

miteinander verbunden.

Die kovarianten Ableitungen bzw. das invariante Differential für einen gemischten Tensor $V_j{}^i{}_k$ sind die folgenden:

$$V_j{}^i{}_{k|m} = \partial_m V_j{}^i{}_k + '\Gamma_r{}^i{}_m V_j{}^r{}_k - ''\Gamma_j{}^r{}_m V_r{}^i{}_k - ''\Gamma_k{}^r{}_m V_j{}^i{}_r, \tag{2.4}$$

$$\nabla_m V_j{}^i{}_k = P^i_r V_s{}^r{}_{t|m} P^s_j P^t_k, \tag{2.5}$$

$$D V_j{}^i{}_k = \nabla_m V_j{}^i{}_k \, dx^m, \tag{2.6}$$

wo die Struktur des Raumes im wesentlichen durch (2.5) und (2.6) festgelegt ist (vgl. [4] § 2 und § 3, insbesondere (2.14), (2.15) und (3.6)—(3.8)). Aus (2.4)—(2.6) sieht man schon, wie diese Operationen auf beliebige Tensoren erweitert werden können. Für ein Skalarfeld $S(x)$ gilt selbstverständlich

$$S_{|m} \equiv \nabla_m S = \partial_m S, \quad DS = dS.$$

Die Übertragungsparameter $''\Gamma_j{}^i{}_k$ sind durch die Gleichungen

$$\nabla_k g_{ij} = \gamma_k(x) g_{ij}(x) \tag{2.7}$$

festgelegt, die für die in (j, k) symmetrischen $''\Gamma_j{}^i{}_k$ ein Gleichungssystem bilden, die leicht gelöst werden kann (vgl. [3], Formel (2.3) und die nachfolgenden Zeilen).

Neben (2.6) benötigen wir im folgenden das invariante Differential

$$\bar{D}V_j{}^i{}_k = V_j{}^i{}_{k|m}\,dx^m \equiv dV_j{}^i{}_k + ('\Gamma_r{}^i{}_m V_j{}^r{}_k - ''\Gamma_j{}^r{}_m V_r{}^i{}_k - ''\Gamma_k{}^r{}_m V_j{}^i{}_r)\,dx^m, \tag{2.8}$$

womit nach (2.5) und (2.6) das invariante Differential D in der Form:

$$DV_j{}^i{}_k = P_r^i(\bar{D}V_s{}^r{}_t) P_j^s P_k^t \tag{2.9}$$

geschrieben werden kann. Für ein Skalarfeld $S(x)$ ist selbstverständlich $DS = \bar{D}S = dS$. Längs einer Kurve $C: x^i = x^i(t)$ verwenden wir im weiteren statt des invarianten Differentials immer den längs C gebildeten invarianten Differentialquotienten D/dt bzw. $\bar{D}/dt$.

§ 3. Veränderung der Länge der Vektoren bei Parallelverschiebung

Die quadrierte Länge V^2 eines Vektors $\vec{V}$ mit den lokalen Komponenten V^i ist durch die Formel:

$$V^2 \stackrel{\text{def}}{=} g_{ij}(x) V^i V^j \tag{3.1}$$

festgelegt. Längs einer Kurve $C: x^i = x^i(t)$ ist nun der invariante Differentialquotient mit dem gewöhnlichen identisch. Aus (3.1) folgt somit längs C:

$$\frac{DV^2}{dt} = \frac{\bar{D}V^2}{dt} = \frac{dV^2}{dt} = \frac{dg_{ij}}{dt} V^i V^j + 2g_{ij} \frac{dV^i}{dt} V^j. \tag{3.2}$$

Beachten wir nun die Formel (2.8) der Operation: $\bar{D}$, so folgt unmittelbar die Formel:

$$\frac{\bar{D}\delta_r^i}{dt} = ('\Gamma_r{}^i{}_k - ''\Gamma_r{}^i{}_k) \frac{dx^k}{dt}, \tag{3.3}$$

womit (3.2) in folgende Form verwandelt werden kann:

$$\frac{DV^2}{dt} = \frac{\bar{D}g_{ij}}{dt} V^i V^j + 2g_{ij} \frac{\bar{D}V^i}{dt} V^j - 2g_{ij} \frac{\bar{D}\delta_r^i}{dt} V^r V^j. \tag{3.4}$$

Bemerkung. Die Übereinstimmung der Formeln (3.2) und (3.4) könnte auch unmittelbar bestätigt werden, wenn in (3.4) für $\bar{D}g_{ij}$ bzw. $\bar{D}V^i$ die Übertragungsparameter $''\Gamma_j{}^i{}_k$ bzw. $'\Gamma_i{}^j{}_k$ verwendet werden, und auch noch (3.3) beachtet wird. —

Die Operation $\bar{D}$ kann auf Grund von (2.9) auch mit der Operation D ausgedrückt werden, wenn beachtet wird, daß für P^i_j auf Grund von (2.2a) die Existenz eines inversen Tensors: Q^i_j postuliert wurde. Nach Kontraktionen wird:

(3.5) $$\bar{D}g_{ij} = Q^r_i Q^s_j D g_{rs},$$

(3.6) $$\bar{D}\delta^i_r = Q^i_m Q^s_r D\delta^m_s,$$ (3.7) $$\bar{D}V^i = Q^i_r DV^r,$$

wo wir — Einfachheit halber — immer nur $\bar{D}$ bzw. D statt $\frac{\bar{D}}{dt}$ bzw. $\frac{D}{dt}$ geschrieben haben.

Beachten wir nun diese Formeln, so wird aus (3.4):

(3.8) $$\frac{DV^2}{dt} = \frac{Dg_{rs}}{dt} Q^r_i Q^s_j V^i V^j + 2g_{ij}\left(Q^i_r \frac{DV^r}{dt} V^j - Q^i_m Q^s_r \frac{D\delta^m_s}{dt} V^r V^j\right).$$

Auf Grund der Rekurrenz des metrischen Grundtensors, d. h. auf Grund von (2.7) wird nun im Hinblick auf (2.6):

(3.9) $$\frac{DV^2}{dt} = Q^r_i Q^s_j g_{rs} V^i V^j \gamma_k \frac{dx^k}{dt} + 2g_{ij}\left(Q^i_r \frac{DV^r}{dt} V^j - Q^i_m Q^s_r \frac{D\delta^m_s}{dt} V^r V^j\right).$$

Da $DV^2 = 2VDV$ ist, bestimmt diese Formel die allgemeinste Form für die Veränderung der Länge eines kontravarianten Vektors bei einer Verschiebung längs einer Kurve: C. Ist diese Verschiebung eine Parallelverschiebung, d. h. ist $DV^r = 0$, so gilt auf Grund von (3.9) der

Satz 1. *In einem Weyl—Otsukischen Raum ist die Veränderung der Länge V eines kontravarianten Vektors V^i bei einer Parallelverschiebung längs einer Kurve $C: x^i = x^i(t)$ durch*

(3.10) $$\frac{DV}{dt} = \frac{1}{2} V^{-1} Q^r_i Q^s_j g_{rs} V^i V^j \gamma_k \frac{dx^k}{dt} - V^{-1} g_{ij} Q^i_m Q^s_r \frac{D\delta^m_s}{dt} V^r V^j$$

bestimmt.

Nehmen wir nun an, daß V^i längs C ein Eigenvektor mit dem Eigenfunktion $\tau(t)$ ist, d. h. es besteht:

(3.11) $$P^i_m V^m = \tau V^i \quad (\tau(t) \neq 0).$$

Nach einer Überschiebung mit Q^i_r folgt aus (3.11) nach (2.2a):

(3.12) $$Q^r_i V^i = \tau^{-1} V^r.$$

Beachten wir noch die Symmetrie von Q_{ij} in (i, j), was aus der Bedingung (2.1) nach einer Kontraktion mit $Q^i_h Q^j_k$ unmittelbar folgt, so wird:

(3.13) $$g_{ij} Q^i_m = g_{im} Q^i_j$$

und (3.9) geht im Hinblick auf (3.12) in

$$(3.14)\qquad \frac{DV^2}{dt}=\tau^{-2}V^2\gamma_k\frac{dx^k}{dt}+2g_{ij}Q^i_r\frac{DV^r}{dt}V^j-2\tau^{-2}g_{im}\frac{D\delta^m_s}{dt}V^sV^i$$

über. Aus (3.11) folgt aber auch die Relation:

$$(3.15)\qquad P^m_k\frac{DV^k}{dt}+\frac{D\delta^m_s}{dt}V^s=\tau\left(\frac{DV^m}{dt}+\frac{d\tau}{dt}V^m\right)$$

(vgl. [4], (5.8); oder [3], (3.8)).

Mit Hilfe von (3.15) können wir $\frac{D\delta^m_s}{dt}V^s$ aus (3.14) eliminieren. Ist längs C auch $\frac{DV^r}{dt}=0$, so gilt wegen

$$\frac{DV^2}{dt}\equiv\frac{dV^2}{dt}=2V\frac{DV}{dt}$$

(nur für Skalare ist $D/dt\equiv d/dt$) der folgende

Satz 2. *Ist V^i längs C ein Eigenvektor mit der Eigenfunktion τ, gilt* (2.1), *und ist ferner $\frac{DV^i}{dt}=0$, so ist $\frac{DV}{dt}$ zur ursprünglichen Länge V proportional:*

$$(3.16)\qquad \frac{DV}{dt}=\left(\frac{1}{2}\tau^{-1}\gamma_k\frac{dx^k}{dt}-\frac{d\tau}{dt}\right)\tau^{-1}V.$$

Aus diesem Satz bzw. aus der Formel (3.16) folgt unmittelbar das

Korollar 2*. *Ist V^i längs C ein Eigenvektor mit der Eigenfunktion $\tau(t)$, ist ferner längs C*

$$\gamma_k(x(t))\frac{dx^k}{dt}=\frac{d\tau^2}{dt},$$

und besteht noch die Bedingung (2.1), *so ist bei Parallelverschiebung längs C die Länge V von V^i eine Konstante.*

§ 4. Veränderung der Fundamentalinvariante der symmetrischen Tensoren

Zu einem in (i, j) symmetrischen rein kontravarianten Tensor T^{ij} ordnen wir durch die Definitionsformel

$$(4.1)\qquad T\overset{\text{def}}{=}g_{ij}T^{ij}$$

eine Invariante: T, die wir als *Tensorlänge* des symmetrischen Tensors T^{ij} nennen wollen. Diese Invariante ist für antisymmetrische Tensoren zweiter Stufe offenbar identisch Null, und für allgemeine kontravariante Tensoren ist T nur mit ihrem symmetrischen Teil gebildet. Die Invariante T hat also wirklich nur für einen symmetrischen Tensor einen Sinn. Für die Tensorlänge der antisymmetrischen Tensoren werden wir im folgenden Paragraphen eine andere Definition angeben.

Auf Grund der Definition des invarianten Differentialquotienten ist

$$\frac{DT}{dt} \equiv \frac{dT}{dt} = \frac{dg_{ij}}{dt} T^{ij} + g_{ij} \frac{dT^{ij}}{dt}.$$

eaBchten wir nun, daß nach (2.8) die Operation $\bar{D}/dt$ für ko- bzw. kontravariante Tensoren mit den Übertragungsparameter $''\Gamma_j{}^i{}_k$ bzw. mit $'\Gamma_j{}^i{}_k$ gebildet werden muß, so bekommt man aus unserer letzten Formel in Hinsicht auf (3.3):

$$\frac{DT}{dt} = \frac{\bar{D}g_{ij}}{dt} T^{ij} + g_{ij} \frac{\bar{D}T^{ij}}{dt} - g_{ij} \left(\frac{\bar{D}\delta^i_r}{dt} T^{rj} + \frac{\bar{D}\delta^j_r}{dt} T^{ir} \right), \tag{4.2}$$

was offenbar längs der Kurve $C: x^i = x^i(t)$ gültig ist.

Drücken wir jetzt die Operation $\bar{D}/dt$ durch D/dt aus, was auf Grund der Formel (2.9) mit Hilfe des inversen Tensors Q^i_j von P^i_j leicht durchführbar ist (vgl. unsere Formeln (3.5) und (3.6)), so geht (4.2) im Hinblick auf (2.7) in

$$\frac{DT}{dt} = Q^r_i Q^s_j g_{rs} T^{ij} \gamma_k \frac{dx^k}{dt} + g_{ij} \left(Q^i_r Q^j_s \frac{DT^{rs}}{dt} - Q^i_s Q^e_r \frac{D\delta^s_e}{dt} T^{rj} - Q^j_s Q^e_r \frac{D\delta^s_e}{dt} T^{ir} \right) \tag{4.3}$$

über. Aus dieser Formel folgt der

Satz 3. *Die Veränderung der durch* (4.1) *bestimmten Tensorlänge T ist bei einer Parallelverschiebung von T^{ij}, d. h. im Falle $DT^{ij} = 0$ durch die folgende Formel angegeben:*

$$\frac{DT}{dt} = Q^r_i Q^s_j g_{rs} T^{ij} \gamma_k \frac{dx^k}{dt} - g_{ij} \left(Q^i_s Q^e_r \frac{D\delta^s_e}{dt} T^{rj} + Q^j_s Q^e_r \frac{D\delta^s_e}{dt} T^{ir} \right). \tag{4.4}$$

Nehmen wir nun an, daß der symmetrische Tensor T^{ij} längs C ein Eigentensor ist, d. h. es gilt längs C:

$$P^i_r P^j_s T^{rs} = \tau(t) T^{ij} \quad (\tau(t) \neq 0). \tag{4.5}$$

Da aus (4.5) auf Grund der Relation (2.2a), längs C

$$\tau^{-1} T^{rs} = Q^r_i Q^s_j T^{ij} \tag{4.5a}$$

folgt, bekommt man aus (4.3) in Hinsicht auf (3.13) und (4.1):

$$\frac{DT}{dt} = \tau^{-1} \gamma_k \frac{dx^k}{dt} T + g_{ij} Q^i_r Q^j_s \frac{DT^{rs}}{dt} - 2\tau^{-1} g_{is} T^{ie} \frac{D\delta^s_e}{dt}. \tag{4.6}$$

Aus der Formel (4.5) folgt nach der Operation D/dt:

$$\frac{D}{dt}(P^i_r P^j_s T^{rs}) = P^i_r P^j_s T^{rs}\frac{d\tau}{dt}+\tau\frac{DT^{ij}}{dt}.$$

Berechnen wir nun auf der linken Seite die Operation D/dt, beachten wir ferner auf der rechten Seite die Formel (4.5) selbst, so wird

$$P^i_a P^j_b\left\{T^{rs}\frac{d}{dt}(P^a_r P^b_s)+P^a_r P^b_s\frac{dT^{rs}}{dt}+({}'\Gamma_p{}^a{}_k P^p_r P^b_s+{}'\Gamma_p{}^b{}_k P^a_r P^p_s)T^{rs}\frac{dx^k}{dt}\right\}=$$

$$=\tau\left(\frac{d\tau}{dt}T^{ij}+\frac{DT^{ij}}{dt}\right).$$

Eliminieren wir jetzt aus dieser Formel (dP^a_r) bzw. (dP^b_s) mittels der Formel

$$\frac{dP^a_r}{dt}=({}'\Gamma_r{}^p{}_k P^a_p-{}''\Gamma_p{}^a{}_k P^p_r)\frac{dx^k}{dt}, \tag{4.7}$$

was aus (2.3) nach einer Kontraktion durch $\dfrac{dx^k}{dt}$ unmittelbar folgt, und beachten wir noch die aus (3.3) folgende Relation

$$\frac{D\delta^i_r}{dt}=P^i_p P^m_r({}'\Gamma_m{}^p{}_k-{}''\Gamma_m{}^p{}_k)\frac{dx^k}{dt}\equiv P^i_p P^m_r\frac{\bar{D}\delta^p_m}{dt}, \tag{4.8}$$

so wird nach entsprechenden Vertauschungen der Indizes:

$$P^i_a P^j_b\frac{DT^{ab}}{dt}+\left(P^b_s P^j_b\frac{D\delta^i_r}{dt}+P^a_r P^i_a\frac{D\delta^j_s}{dt}\right)T^{rs}=\tau\left(\frac{DT^{ij}}{dt}+\frac{d\tau}{dt}T^{ij}\right). \tag{4.9}$$

Da aber T^{rs} in (r, s) symmetrisch vorausgesetzt wurde, wird nach einer Kontraktion von (4.9) mit $Q^h_i Q^f_j$, und auf Grund von (2.2a)

$$\frac{DT^{hf}}{dt}+(P^f_r Q^h_e+P^h_r Q^f_e)\frac{D\delta^e_s}{dt}T^{rs}=\tau\left(\frac{DT^{ij}}{dt}+\frac{d\tau}{dt}T^{ij}\right)Q^h_i Q^f_j. \tag{4.10}$$

Überschiebung von beiden Seiten der Formel (4.10) mit g_{hf} gibt in Hinsicht auf (2.1), (2.2b) bzw. auf der rechten Seite nach (4.5a):

$$g_{hf}\frac{DT^{hf}}{dt}+2g_{re}T^{rs}\frac{D\delta^e_s}{dt}=\tau g_{hf}Q^h_i Q^f_j\frac{DT^{ij}}{dt}+\frac{d\tau}{dt}T. \tag{4.11}$$

Elimineren wir das Glied $\dfrac{D\delta^e_s}{dt}$ von (4.6) mittels (4.11), so erhält man dem folgenden

Satz 4. *Ist für den symmetrischen Eigentensor T^{ij} längs einer Kurve $C: DT^{ij}=0$, besteht* (2.1), *so ist die Veränderung der durch* (4.1) *angegebenen Tensorlänge T mit T selbst proportional; in expliziter Form:*

$$\frac{DT}{dt} = \left(\gamma_k \frac{dx^k}{dt} - \frac{d\tau}{dt}\right) \tau^{-1} T. \tag{4.12}$$

Analog zur Formel (3.16) und zum Korollar 2*, folgt aus (4.12) das

Korollar 4*. *Ist der symmetrische Tensor T^{ij} längs C ein Eigentensor mit der Eigenfunktion τ, gilt ferner längs C:*

$$\gamma_k(x(t)) \frac{dx^k}{dt} = \frac{d\tau}{dt},$$

***so ist** bei Parallelverschiebung längs C die Tensorlänge T eine Konstante.*

§ 5. Veränderung der Fundamentalinvariante der antisymmetrischen Tensoren

In diesem Paragraphen soll $T^{ij} = -T^{ji}$ durchwegs einen antisymmetrischen Tensor bedeuten. Da der Tensor

$$g_{ijhk} \overset{\text{def}}{=} \frac{1}{2}(g_{ih}g_{jk} - g_{ik}g_{jh})$$

in (i, j) und auch in (h, k) antisymmetrisch ist, ist

$$T_{(a)} \overset{\text{def}}{=} \sqrt{g_{ijhk}T^{ij}T^{hk}} \equiv \sqrt{g_{ir}g_{js}T^{ij}T^{rs}} \tag{5.1}$$

eine Invariante des antisymmetrischen Tensors T^{ij}, die wir als *Tensorlänge* von T^{ij} bezeichnen wollen[2]). Die Einführung des Tensors g_{ijhk} stammt von Prof. A. KAWAGUCHI (vgl. [1], § 3); für Bivektoren definiert $T_{(a)}$ — bis auf eine Konstante — in arealen Räumen eben das Maß der Bivektoren (vgl. [1], (3.22)).

Längs einer Kurve $C: x^i = x^i(t)$ ist nach (5.1):

$$\frac{DT^2_{(a)}}{dt} = \frac{d}{dt}(g_{ir}g_{js}T^{ij}T^{rs}) \equiv \frac{\bar{D}g_{ir}}{dt} g_{js}T^{ij}T^{rs} + \tag{5.2}$$

$$+ g_{ir}\frac{\bar{D}g_{js}}{dt}T^{ij}T^{rs} + g_{ir}g_{js}\left(\frac{''\bar{D}T^{ij}}{dt}T^{rs} + T^{ij}\frac{''\bar{D}T^{rs}}{dt}\right),$$

---

[2]) Der Index „a" bei $T_{(a)}$ bedeutet die Antisymmetrie von T^{ij}; es ist also nicht ein tensorieller Index.

wo $\dfrac{''\bar{D}}{dt}$ den mit $''\Gamma_{j\ k}^{\ i}$ gebildeten invarianten Differentialquotienten bedeutet. Offenbar ist für rein kovariante Tensoren $''\bar{D}=\bar{D}$.

Benützen wir nun die Antisymmetrie von T^{ij} in „i" und „j", so vereinfacht sich die Formel (5.2) auf

$$\frac{1}{2}\frac{DT^2_{(a)}}{dt}=\frac{\bar{D}g_{ir}}{dt}g_{js}T^{ij}T^{rs}+g_{ir}g_{js}\frac{''\bar{D}T^{ij}}{dt}T^{rs}.$$

Mit Hilfe der Formel (3.3) können die Übertragungsparameter $''\Gamma_{i\ k}^{\ j}$ aus $''\bar{D}T^{ij}$ eliminiert werden; da jetzt statt dieser Übertragungsparameter die $'\Gamma_{i\ k}^{\ j}$ hineinkommen, wird wegen $'\bar{D}T^{ij}\equiv\bar{D}T^{ij}$ aus der letzten Gleichung:

$$\frac{1}{2}\frac{\bar{D}T^2_{(a)}}{dt}=\frac{\bar{D}g_{ir}}{dt}g_{js}T^{ij}T^{rs}+\left(\frac{\bar{D}T^{ij}}{dt}-\frac{\bar{D}\delta^i_k}{dt}T^{kj}-\frac{\bar{D}\delta^j_k}{dt}T^{ik}\right)T_{ij}.$$

Unter einer wiederholten Beachtung der Antisymmetrie von T^{ij} und selbstverständlich auch die von T_{ij}, erhält man auf Grund von (2.9), welche Formel die Operationen D und $\bar{D}$ miteinander verknüpft, und ferner wegen (2.2a), (2.2b):

$$(5.3)\qquad T_{(a)}\frac{DT_{(a)}}{dt}=\frac{Dg_{cb}}{dt}Q^c_iQ^b_rT^i{}_sT^{rs}+\frac{DT^{cb}}{dt}Q^i_cQ^j_bT_{ij}-2T_{ip}T^{kp}Q^b_kQ^i_c\frac{D\delta^c_b}{dt},$$

da für Skalare offenbar die Operationen $\bar{D}/dt$, D/dt und d/dt übereinstimmen.

Auf Grund von (5.3) folgt im Hinblick auf (2.7) der

Satz 5. *In einem Weyl—Otsukischen Raum ist die Veränderung der Tensorlänge $T_{(a)}$ eines antisymmetrischen Tensors T^{ij} bei einer Parallelverschiebung längs einer Kurve $C:x^i=x^i(t)$ durch:*

$$(5.4)\qquad \frac{DT_{(a)}}{dt}=T^{-1}_{(a)}\left(Q^c_iQ^b_rg_{cb}T^i{}_sT^{rs}\gamma_k\frac{dx^k}{dt}-2T_{is}T^{rs}Q^b_rQ^i_c\frac{D\delta^c_b}{dt}\right)$$

bestimmt.

Nehmen wir nun an, daß der Grundtensor P^i_j die Form

$$(5.5)\qquad P^i_j=\varrho\delta^i_j\quad(\varrho=Konst.)$$

hat. In diesem Fall ist nach (2.2a): $Q^i_r=\varrho^{-1}\delta^i_r$, ferner wird nach (2.3): $'\Gamma=''\Gamma$ und somit auch $\dfrac{D\delta^i_r}{dt}=0$. Es besteht also nach (5.4) und (5.1) das

Korollar 5*. *Gilt in einem Weyl—Otsukischen Raum die Relation* (5.5), *so ist die Veränderung der Tensorlänge $T_{(a)}$ eines antisymmetrischen Tensors T^{ij} bei einer Parallelverschiebung zu $T_{(a)}$ selbst proportional. Es ist*

$$\frac{DT_{(a)}}{dt} = \sigma T_{(a)}, \quad \sigma \overset{\text{def}}{=} \varrho^{-2}\gamma_k \frac{dx^k}{dt}. \tag{5.6}$$

§ 6. Bemerkungen über dem Tensor g_{ij}

In den klassischen Weylschen Räumen, wo $P^i_j = \delta^i_j$ und folglich $\nabla_k \delta^i_j = 0$ ist, folgt bekanntlich aus (2.7), d. h. aus $\nabla_k g_{ij} = \gamma_k g_{ij}$ auch $\nabla_k g^{ij} = -\gamma_k g^{ij}$ (vgl. z. B. [2], Formel (7.1) a), wo aber der Raum ein Linienelementraum, d. h. allgemeiner, als ein Punktraum ist). Wir wollen die Formel von $\nabla_k g^{ij}$ in den Weyl—Otsukischen Räumen bestimmen.

Der Tensor g^{ih} ist bekanntlich durch die Formel

$$g_{ij} g^{jh} = \delta^h_i$$

bestimmt. Die mit Hilfe von $''\Gamma_a{}^b{}_c$ gebildete kovariante Ableitung von beiden Seiten gibt

$$(''\nabla_k g_{ij}) g^{jh} + (''\nabla_k g^{jh}) g_{ij} = 0, \tag{6.1}$$

wo $''\nabla_k$ die mit $''\Gamma$ gebildete gewöhnliche affine kovariante Ableitung bezeichnet (vgl. [4], § 3, S. 111). Offenbar ist $''\nabla_k g_{ij} = g_{ij|k}$, hingegen gilt wegen

$$\delta^h_{r|k} = {}'\Gamma_r{}^h{}_k - {}''\Gamma_r{}^h{}_k \tag{6.2}$$

für $''\nabla_k g^{ij}$ die Relation

$$''\nabla_k g^{jh} = g^{jh}{}_{|k} - \delta^j_{r|k} g^{rh} - \delta^h_{r|k} g^{jr}.$$

Aus (6.1) wird somit

$$g_{ij|k} g^{jh} + g^{jh}{}_{|k} g_{ij} = \delta^j_{r|k} g^{rh} g_{ij} + \delta^h_{i|k}. \tag{6.3}$$

Die in (6.3) vorkommende kovariante Ableitung von g_{ij} bzw. g^{jh} kann auf Grund von (2.5) mit Hilfe der kovarianten Ableitung ∇_k ausgedrückt werden. Aus (6.3) wird somit:

$$g^{jh} Q^r_i Q^s_j \nabla_k g_{rs} + g_{ij} Q^j_r Q^h_s \nabla_k g^{rh} = \delta^j_{r|k} g^{rh} g_{ij} + \delta^h_{i|k}. \tag{6.4}$$

Wir müssen jetzt aus dieser Gleichung $\nabla_k g^{jh}$ unter Beachtung von (2.7) bestimmen. Ziehen wir in der Symmetrieforderung (2.1) die Indizes herauf, so wird:

$$g^{ab} P^i_b = g^{ib} P^a_b,$$

beachten wir (2.7), so wird aus (6.4) nach einer Kontraktion auf beiden Seiten mit $g^{ac}P^i_c P^b_h$:

$$\nabla_k g^{ab} = -\gamma_k g^{ab} + (\nabla_k \delta^a_r) g^{rb} + (\nabla_k \delta^b_r) g^{ar}. \tag{6.5}$$

Aus dieser Formel folgt

Satz 6. *Die Relation*

$$g^{rb} \nabla_k \delta^a_r + g^{ra} \nabla_k \delta^b_r = 0 \tag{6.6}$$

ist notwendig und hinreichend dafür, daß neben g_{ij} *auch* g^{ij} *rekurrente kovariante Ableitung habe mit* $(-\gamma_k)$ *als Rekurrenzvektor.*

Wir gehen jetzt zur Diskussion des Falles über, in dem der Tensor g_{ij} außer (2.7), längs einer vorgegebenen Kurve $C: x^i = x^i(t)$ auch der Bedingung

$$P^a_i P^b_j g_{ab} = \tau(t) g_{ij} \quad (\tau(t) \neq 0) \tag{6.7}$$

genügt, d. h. g_{ij} ist nicht nur ein rekurrenter metrischer Fundamentaltensor, sondern längs der Kurve C auch ein zur Funktion $\tau(t)$ gehöriger Eigentensor.

Es kann leicht ein solcher Weyl—Otsukischer Raum konstruiert werden, in dem also (2.7) und (6.7) gleichzeitig erfüllt sind. Die Forderung (6.7) ist nämlich eine Bedingung für $g_{ij}(x)$ und P^a_i, sogar nur längs einer Kurve C. Befriedigt nun g_{ij} die Bedingung (6.7), so bestimmt (2.7) ein Gleichungssystem für die Übertragungsparameter $''\Gamma_a{}^b{}_c$, und auf Grund von (2.3) können auch die Übertragungsparameter $'\Gamma_a{}^b{}_c$ bestimmt werden (vgl. [3], Formel (2.3)).

Aus der Bedingung (6.7) kann nun die Formel

$$g_{mj} \frac{D\delta^m_i}{dt} + g_{mi} \frac{D\delta^m_j}{dt} = g_{ij} \frac{d\tau}{dt} \tag{6.8}$$

abgeleitet werden, wie wir das in [3] durchgeführt haben (vgl. [3], Satz 5, insbesondere Gleichung (4.4)).

Wir wollen nun mit Hilfe von (6.8), das also eine Folgerung von (6.7) ist, die Formeln (3.10), (4.4) und (5.4) unformen. Führen wir die Bezeichnungen

$$\tilde{V}^i \overset{\text{def}}{=} Q^i_r V^r, \quad \tilde{V}^2 = g_{ij} \tilde{V}^i \tilde{V}^j \tag{6.9}$$

ein, so geht (3.10) unter Beachtung der Symmetriebedingung (3.13) in

$$\frac{DV}{dt} = \frac{1}{2} V^{-1} \tilde{V}^2 \gamma_k \frac{dx^k}{dt} - V^{-1} g_{im} \tilde{V}^i \tilde{V}^s \frac{D\delta^m_s}{dt} \tag{6.10}$$

über. Nach einer Kontraktion von (6.8) mit $\tilde{V}^i \tilde{V}^j$ und dann unter Beachtung der mit (6.7) äquivalenten Relation

$$Q^r_i Q^s_j g_{rs} = \tau^{-1} g_{ij}, \tag{6.11}$$

bekommt man

$$\frac{D\delta_j^m}{dt} g_{im}\tilde{V}^i\tilde{V}^j = \frac{1}{2}\tilde{V}^2\frac{d\tau}{dt}, \quad \tilde{V}^2 = \tau^{-1}V^2,$$

wodurch (6.10) in der Form:

$$\frac{DV}{dt} = \frac{1}{2}\tau^{-1}\left(\gamma_k\frac{dx^k}{dt} - \frac{d\tau}{dt}\right)V \tag{6.12}$$

geschrieben werden kann.

Wir gehen nun zur Umformung von (4.4) über, wenn (6.7) bzw. (6.8) besteht. Die Formel (4.4) kann mit Hilfe der Bezeichnung:

$$\tilde{T}^{rs} \overset{\text{def}}{=} Q_i^r Q_j^s T^{ij},$$

ferner in Hinsicht auf die Symmetrie von T^{ij}, bzw. nach (3.13), in der Form

$$\frac{DT}{dt} = g_{rs}\tilde{T}^{rs}\gamma_k\frac{dx^k}{dt} - 2g_{is}\tilde{T}^{ie}\frac{D\delta_e^s}{dt} \tag{6.13}$$

geschrieben werden. Aus (6.8) folgt aber nach einer Kontraktion mit $\tilde{T}^{ij}$ und in Hinsicht auf (6.11):

$$2g_{mj}\tilde{T}^{ij}\frac{D\delta_i^m}{dt} = \frac{d\tau}{dt}g_{ij}\tilde{T}^{ij} = \tau^{-1}\frac{d\tau}{dt}T, \tag{6.14}$$

und wieder in Hinsicht auf (6.11) hat man noch

$$g_{rs}\tilde{T}^{rs} = g_{rs}Q_i^r Q_j^s T^{ij} = \tau^{-1}g_{ij}T^{ij} = \tau^{-1}T. \tag{6.15}$$

Substituieren wir (6.14) und (6.15) in (6.13), so wird:

$$\frac{DT}{dt} = \tau^{-1}\left(\gamma_k\frac{dx^k}{dt} - \frac{d\tau}{dt}\right)T. \tag{6.16}$$

Bei der Umformung von (5.4) benützen wir im Hinblick auf (6.11) und (5.1) die folgenden Bezeichnungen:

$$\hat{T}^{ps} \overset{\text{def}}{=} Q_i^p T^{is}, \quad \hat{T}^2 \overset{\text{def}}{=} g_{pr}\hat{T}^r{}_s\hat{T}^{ps} \equiv \tau^{-1}T_{(a)}^2.$$

Aus der Formel (5.4) folgt somit:

$$\frac{DT_{(a)}}{dt} = T_{(a)}^{-1}\left(\hat{T}^2\gamma_k\frac{dx^k}{dt} - 2\hat{T}^{bs}T_{is}Q_c^i\frac{D\delta_b^c}{dt}\right). \tag{6.17}$$

Nun ist nach (3.13):

$$T_{is}Q_c^i = g_{im}T^m{}_sQ_c^i = g_{ic}\hat{T}^i{}_s = \hat{T}_{cs},$$

und aus (6.17) wird

$$\frac{DT_{(a)}}{dt} = T_{(a)}^{-1}\left(\hat{T}^2\gamma_k\frac{dx^k}{dt} - 2\hat{T}^{bs}\hat{T}_{cs}\frac{D\delta_b^c}{dt}\right). \tag{6.18}$$

Eine Kontraktion von (6.8) mit $\hat{T}^j{}_s\hat{T}^{is}$ gibt nach entsprechenden Vertauschungen der Indizes:

$$2\hat{T}_{ms}\hat{T}^{bs}\frac{D\delta_b^m}{dt} = \frac{d\tau}{dt}\hat{T}_{is}\hat{T}^{is} = \frac{d\tau}{dt}\hat{T}^2,$$

wodurch (6.18) in

$$\frac{DT_{(a)}}{dt} = T_{(a)}^{-1}\hat{T}^2\left(\gamma_k\frac{dx^k}{dt} - \frac{d\tau}{dt}\right) \equiv \tau^{-1}\left(\gamma_k\frac{dx^k}{dt} - \frac{d\tau}{dt}\right)T_{(a)} \tag{6.19}$$

übergeht.

Wir fassen unsere Resultate im folgenden Satz zusammen:

Satz 7. *Ist in einem Weyl—Otsukischen Raum der metrische Grundtensor* g_{ij} *längs einer Kurve C ein Eigentensor, so sind die Differentialquotienten der entsprechenden Tensorlängen zu den ursprünglichen Tensorlängen proportional und die Formeln* (6.12), (6.16) *und* (6.19) *bestimmen der Reihe nach die Veränderung der Länge der Vektoren bzw. die der symmetrischen und antisymmetrischen Tensoren bei einer Parallelverschiebung längs C. In allen drei Fällen ist* $\gamma_k\frac{dx^k}{dt} = \frac{d\tau}{dt}$ *notwendig und hinreichend dafür, daß die Länge bei Parallelverschiebung konstant sei, wenn nur* g_{ij} *eine positiv-definite Metrik bestimmt.*

Schriftenverzeichnis

[1] A. Kawaguchi, On areal spaces. I, *Tensor,* N. S. **1** (1950), 14—45.

[2] A. Moór, Über eine Übertragungstheorie der metrischen Linienelementräume mit rekurrentem Maßtensor, *Tensor,* N. S. **29** (1975), 47—63.

[3] A. Moór, Otsukische Übertragung mit rekurrentem Maßtensor. *Acta Sci. Math.,* **40** (1978), 129—142.

[4] T. Otsuki, On general connections. I, *Math. J. Okayama Univ.,* **9** (1959—60), 99—164.

[5] H. Weyl, Reine Infinitesimalgeometrie, *Math. Zeitschrift,* **2** (1918), 384—411.

LEHRSTUHL FÜR MATHEMATIK
UNIV. FÜR FORST UND HOLZINDUSTRIE
H—9401 SOPRON, PF. 132

Acta Sci. Math., 41 (1979), 187—190

Remarks on finitely projected modular lattices

E. T. SCHMIDT

1. **Introduction.** Let **K** be a variety of lattices. A lattice L in **K** is called finitely **K**-projected if for any surjective $f: K \twoheadrightarrow L$ in **K** there is a finite sublattice of K whose image under f is L. These lattices are important by the investigations of subvarieties of **K**, in fact, every finite **K**-projected subdirectly irreducible lattice L is splitting in **K**, i.e. there is a largest subvariety of **K** not containing L (see DAY [1]). Let $\mathbf{B}_2$ be the variety generated by all breadth 2 modular lattices. In [2] there is given a necessary condition for a lattice $L \in B_2$ to be $\mathbf{B}_2$-projected. Our goal here is to give some further necessary conditions for a lattice to be **M**-projected, where **M** denotes the variety of all modular lattices.

2. **Preliminaries.** Let M be a finite modular lattice and let Q be the chain of bounded rationals, say $Q=[0, 1]$. $M(Q)$ is the lattice of all continous monotone maps of the compact totally ordered disconnected space X of all ultrafilters of Q into the discrete space M. The constant mappings form a sublattice of $M(Q)$ which is isomorphic to M; we identify M with this sublattice. If a/b is a prime quotient of M then the corresponding quotient a/b of $M(Q)$ is isomorphic to Q, we have a natural isomorphism ε_{ab}: $Q \to a/b$. If a/b runs over all prime quotients then all a/b generate a sublattice $M[Q]$ of $M(Q)$.

Let A and B be two modular lattices with isomorphic sublattices $C \cong C'$ where C is a filter of A and C' is an ideal of B. Then $L=A \cup B$ can be made into a modular lattice by defining $x \leqq y$ if and only if one of the following conditions is satisfied: $x \leqq y$ in A or $x \leqq y$ in B or $x \leqq c$ in A and $c' \leqq y$ in B where c, c' are corresponding elements under the isomorphism $C \cong C'$. We say that L is the lattice obtained by gluing together A and B identifying the corresponding elements under the isomorphism $C \cong C'$. This useful construction is due to Hall and Dilworth. In this case A is an ideal and B is a filter of L, $L=A \cup B$ and $C=A \cap B$. Conversely if A is an ideal and B is a filter of a lattice L such that $L=A \cup B$ then L is obviously the lattice obtained by gluing together A and B.

Received February 11, 1978; in revised form September 16, 1978.

3. The Hall-Dilworth construction.

Theorem 1. *Let A be an ideal and let B be the filter of the finite modular lattice M such that $M=A\cup B$ and $C=A\cap B$ is a chain. Let a/b and c/d be two different prime quotients of C which are projective in A and in B. Then M is not finitely* **M**-*projected.*

Proof. $A[Q]$ is an ideal and $B[Q]$ is a filter of $M[Q]$. Consequently, $M[Q]=$ $=A[Q]\cup B[Q]$. It is easy to see that $A[Q]\cap B[Q]=C[Q]$. Let $B'[Q]$ be a disjoint copy of $B[Q]$ with the isomorphism $\varphi\colon B[Q]\to B'[Q]$ $(x\to x')$. The restriction of φ to $C[Q]$ give a sublattice $C'[Q]$ of $B'[Q]$.

Let a/b and c/d two different prime quotients of C. Then we can assume that $a>b\geqq c>d$. First we define an injection $\psi\colon C[Q]\to C'[Q]$ which is different from φ. To define this ψ we distinguish two cases:

(a) We assume that there exists a $u\in C$ covering a. The quotients u/b, a/b, u/a of $C[Q]$ are all isomorphic to Q. Let further δ be an automorphism of u/b and we set $a_0=a$, $a_1=\delta a_0, \dots, a_{i+1}=\delta a_i$ and $\bar{a}_1=\delta^{-1}a_0, \dots, a_{i+1}=\delta^{-1}\bar{a}_i$. Obviously, if r is an arbitrary irrational number between 0 and 1 then there exists an automorphism δ of u/b satisfying the following two conditions (see Fig. 1.).

$$a_1 < a, \tag{1}$$

$$\varepsilon_{ab}(\inf\{a_i\}) = \varepsilon_{ua}(\sup\{\bar{a}_i\}) = r \tag{2}$$

(ε_{ab} (resp. ε_{ua})) denotes the natural isomorphism $a/b\to Q$ (resp. $u/a\to Q$)). Defining ψ_0 to be the product $\varphi\circ\delta$, ψ_0 is an isomorphism of u/b onto u'/b'. ψ_0 can be extended to an isomorphism $\psi\colon C[Q]\to C'[Q]$ as follows:

$$\psi(x)=\begin{cases}\varphi(x) & \text{if } x\notin u/b\\ \psi_0(x) & \text{if } x\in u/b.\end{cases}$$

(b) In the second case a is a maximal element of C. Then we can choose an arbitrary t such that $a'>t>b'$. t/b' is isomorphic to Q, hence there exists an isomorphism $\psi_0\colon a/b\to t/b'$. The extension of ψ_0 is defined by

$$\psi(x)=\begin{cases}\varphi(x) & \text{if } x\notin a/b\\ \psi_0(x) & \text{if } x\in a/b.\end{cases}$$

We take in both cases the lattice L obtained by gluing together $A[Q]$ and $B'[Q]$ identifying the corresponding elements of $C[Q]$ and $\psi(C[Q])$ under the isomorphism ψ (Fig. 2).

We prove that there exists a surjection $f\colon L\twoheadrightarrow M$. Let Θ be the congruence relation of Q defined as follows: $x\equiv y(\Theta)$ if and only if either $x, y>r$ or $x, y<r$. Then $A[Q]$ has a congruence relation Θ_A such that the restriction of Θ_A to a quotient

a/b — where a/b is a prime quotient of A — is the image of Θ by the isomorphism ε_{ab}: $Q \to a/b$. The corresponding factor lattice $A[Q]/\Theta_A$ is isomorphic to A. Similarly $B'[Q]$ has a congruence relation Θ_B corresponding to Θ, and the factor lattice is isomorphic to B. By the definition of δ, $x \equiv y(\Theta_A)$ $(x, y \in C[Q])$ if and only if $\delta x \equiv \delta y(\Theta_A)$. That means that the restriction of Θ_A to $\delta C[Q]$ corresponds by φ to the restriction of Θ_B to $C'[Q]$. If follows that the join $\Theta_A \cup \Theta_B$ has an extension $\bar{\Theta}$ to L such that the restriction of $\bar{\Theta}$ to $A[Q]$ is Θ_A and the restriction $B'[Q]$ is Θ_B. Hence $M/\bar{\Theta}$ is isomorphic to M.

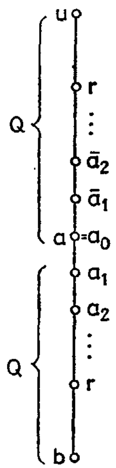


Fig. 1

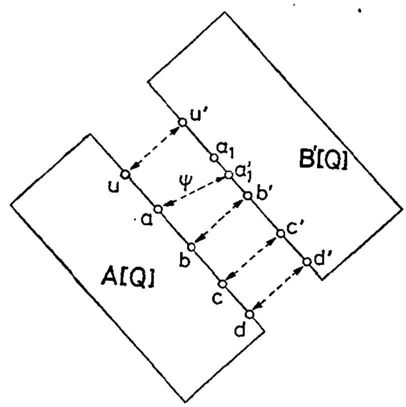


Fig. 2

Let π be the projectivity $a/b \approx c/d$ in A. ϱ denotes the projectivity $c/d \approx a/b$ in B. Thus we get the projectivity $\varrho \circ \pi : a/b \approx a/b$. It is easy to show that this projectivity has no inverse in L, i.e. by Lemma 1 of [2] we get that M is not finitely projected.

Theorem 2. *Let A be an ideal and let B be a filter of the finite modular lattice M such that $M = A \cup B$ and $C = A \cap B$ is a Boolean lattice. Let a/b and c/d be two prime quotients of C which are projective in A and in B. If M is finitely* **M**-*projected then a/b and c/d are projective in C.*

Proof. The proof is similar to the previous one. We define an injective endomorphism δ of $C[Q]$. a/b is isomorphic to Q, hence we can choose an arbitrary t such that $b < t < a$. Let u be the relative complement of b in the quotient a/o where o denotes the least element of $C[Q]$. Finally u' denotes the complement of u in $C[Q]$. Then the ideal $(t \vee u']$ of $C[Q]$ is isomorphic to $C[Q]$. We have therefore an injective endomorphism δ for which $\delta a = t$ and $\delta x = x$ for every $x \leqq u'$. We assume that $c, d \leqq u'$.

If φ denotes the isomorphism $C[Q]\to C'[Q]$ then $\psi=\varphi\circ\delta$ is an injection of $C[Q]$ into $C'[Q]$, such that the image of $C[Q]$ is an ideal of $C'[Q]$, $\psi a>a$, $\psi b=b$, $\psi c=c$, $\psi d=d$. Let L be the lattice obtained by gluing together $A[Q]$ and $B[Q]$ identifying the corresponding elements under ψ. We can finish the proof as in Theorem 1.

It is easy to generalize the previous theorems if we introduce the following notion.

Definition. Let M be a finite lattice. An injective endomorphism δ of $M[Q]$ is called a compression if the following properties are satisfied.

(i) $\delta(x)\leqq x$ for every $x\in M[Q]$ and $\delta M[Q]$ is an ideal of $M[Q]$;

(ii) there exists a $\Theta\in\mathrm{Con}(Q)$ with exactly two Θ-classes such that $\delta^{-1}(x)\equiv \equiv x\equiv\delta x(\bar{\Theta})$ for every x where $\bar{\Theta}$ denotes the extension of Θ to $M[Q]$.

Theorem 3. *Let A be an ideal and let B be a filter of the finite modular lattice M, such that $M=A\cup B$. Let further a/b and c/d be two prime quotients of $C=A\cap B$ which are projective in A and in B. If C has a compression δ such that $a>\delta a>\delta b=b$ and $\delta c=c$, $\delta d=d$ then M is not finitely* M-*projected.*

4. **Stable quotients.** Let a/b be a prime quotient of a finite modular lattice M. We define a new element t to M for which $a>t>b$. Then $M\cup\{t\}$ is a partial lattice with the sublattice M. $t\vee m$, $t\wedge m$ $(m\in M)$ are not defined. It is easy to show that there exists a lattice $\overline{M}$ freely generated by this partial lattice. We say that a/b is stable if $\overline{M}$ is finite. A. Mitschke and R. Wille have proved that every prime quotient of M_3 is stable. The prime-quotients of M_4 are not stable.

Conjecture. A finite modular lattice is finitely M-projected if and only if every prime quotient is stable.

It is easy to show — applying [2] — that a finite planar modular lattice is finitely M-projected if and only if every prime quotient is stable.

References

[1] A. Day, Splitting algebras and weak notion of projectivity, *Algebra Universalis*, **5** (1975), 153—162.

[2] A. Mitschke, E. T. Schmidt, R. Wille, On finitely projected modular lattices of breadth two, *in preparation.*

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

Acta Sci. Math., 41 (1979), 191—207

On curvature measures

L. L. STACHÓ

1. Introduction

It is well-known that Steiner's famous polynomial formula for the volume function of convex parallel sets is based on the following heuristical idea:

If A is a convex open subset of $\mathbf{R}^n$ (the Euclidean n-space) whose boundary ∂A is a C^2 submanifold of $(n-1)$-dimensions of $\mathbf{R}^n$ and $\varrho > 0$ then for its parallel set (of radius ϱ) $A_\varrho \equiv \{c \in \mathbf{R}^n : \operatorname{dist}(x, A) < \varrho\}$ we have that $\partial(A_\varrho)$ is also an $(n-1)$-dimensional C^2-submanifold of $\mathbf{R}^n$, and denoting its infinitesimal surface piece by dF one can find the following relation between the $(n-1)$-dimensional Hausdorff measures of dF and its projection on $\bar{A}$ (the closure of A): [1])

$$\operatorname{vol}_{n-1} dF = (1+\varrho\varkappa_1) \dots (1+\varrho\varkappa_{n-1}) \operatorname{vol}_{n-1} dF^0 \quad \text{with} \quad dF^0 \equiv pr_{\bar{A}}\, dF$$

where $\varkappa_1, \dots, \varkappa_{n-1}$ denote the values of the main curvatures of ∂A at the place dF^0.

Hence one easily deduces that for all bounded Borel sets $Q \subset \mathbf{R}^n$ the n-dimensional Hausdorff measure (which, by definition, coincides with Lebesgue measure on $\mathbf{R}^n$) of the figures $T(Q, \varrho) \equiv A \cap \{t \in \mathbf{R}^n : pr_{\bar{A}} t \in Q\}$ is a polynomial of degree n in the variable ϱ, of the form

$$\operatorname{vol}_n T(Q, \varrho) = \sum_{j=0}^{n} a_j(Q) \varrho^j \tag{1}$$

where for the coefficients we have

$$a_0(Q) = \operatorname{vol}_n Q \cap A, \quad a_1(Q) = \operatorname{vol}_{n-1} Q \cap \partial A,$$

and

$$a_j(Q) = \int_{Q \cap \partial A} \frac{1}{j} \sum_{\substack{I \subset \{1, \dots, n-1\} \\ \operatorname{card} I = j}} \prod_{i \in I} \varkappa_i(p)\, d(\operatorname{vol}_{n-1} p)$$

Received February 27, 1978.

[1]) For any closed subset B of $\mathbf{R}^n$ and for $x \in \mathbf{R}^n$ we define $\operatorname{pr}_B x \equiv \{b \in B : \operatorname{dist}(x, b) = \operatorname{dist}(x, B)\}$. For $G \subset \mathbf{R}^n$ we define $\operatorname{pr}_B G \equiv \bigcup_{x \in G} \operatorname{pr}_B x$.

for $j=2, \dots, n$ (card=cardinality); $\varkappa_1(p), \dots, \varkappa_{n-1}(p)$ are the main curvatures of ∂A at the point $p \in \partial A$.

This result was considerably generalized by FEDERER [1]: *If a closed set $A \subset \mathbf{R}^n$ is such that*

$$\operatorname{reach} A \equiv \sup \{\delta \geqq 0 \colon \forall x \in A_\delta, \ \operatorname{card} pr_A x = 1\} > 0 \quad (\text{with } A_0 = A),$$

then there exist (uniquely determined) *signed Borel measures $a_0, \dots, a_n$ over $\mathbf{R}^n$ such that* (1) *holds for all bounded Borel subsets Q of $\mathbf{R}^n$ and for all ϱ with* $0 < \varrho < \operatorname{reach} A$.

Our purpose in the present article is to prove a result analogous to this theorem which applies to every $A \subset \mathbf{R}^n$ and $\varrho > 0$ and which allows us to extend the concept of curvature measure to the boundary of every $A \subset \mathbf{R}^n$ in a reasonable manner.

2. Summary and alternative formulation of some of Federer's arguments

Theorem A. *Let A be a non-empty closed subset of $\mathbf{R}^n$ and f denote the function $x \mapsto \operatorname{dist}(x, A)$ on $\mathbf{R}^n \setminus A$. The function f is totally derivable exactly at those points of $\mathbf{R}^n \setminus A$ which admit a unique projection on A, and for such a point x,* $\operatorname{grad} f(x)$ *coincides with the unit vector $(x - pr_A x)/\operatorname{dist}(x, A)$. The function f satisfies a Lipschitz condition of order one with (exact) Lipschitz constant* 1, *and the set of the singular points $Z \equiv \{x \in \mathbf{R}^n \setminus A \colon \operatorname{card} pr_A x > 1\}$ has* vol_n*-measure* 0. *Removing Z from $\mathbf{R}^n \setminus A$, the remaining set $Q \equiv \mathbf{R}^n \setminus (A \cup Z) = \{x \in \mathbf{R}^n \setminus A \colon \operatorname{card} pr_A x = 1\}$ can de uniquely decomposed into a family $\mathbf{Q}$ of pairwise disjoint straight line segments so that for any member L of $\mathbf{Q}$ there exists a (unique) point p in ∂A such that* $\{p\} = pr_A L = \bar{L} \cap \partial A$.

Proof. See [2] p. 93, [3] pp. 271 and 216.

Definition. We shall call the members of the family $\mathbf{Q}$ described in Theorem A the *prenormals* of the set A. I.e. $L (\subset \mathbf{R}^n)$ is a prenormal of A if there exist a point $p \in \partial A$ and a unit vector $k (\in \mathbf{R}^n)$ such that $L = \{x \in \mathbf{R}^n \setminus A \colon pr_A x = \{p\}$ and $(x-p)/\|x-p\| = k\}$.

Definition. A mapping f will be called C^{1+}*-smooth* if it is defined on some open subset Ω of some space $\mathbf{R}^s$ with $f \in C^1(\Omega)$ (i.e. if f has a continuous gradient on Ω) and its gradient locally satisfies a Lipschitz condition (i.e. for all compact subsets K of Ω, $\operatorname{Lip}(\operatorname{grad} f|_K) < \infty$).

Since the composition of C^{1+}-mappings is also a C^{1+}-mapping, it makes sense to speak of $k (\leqq n)$-dimensional C^{1+}-submanifolds of the space $\mathbf{R}^n$. In particular,

F is an $(n-1)$-dimensional C^{1+}-submanifold of $\mathbf{R}^n$ if, for any $y \in F$, one can find an open neighborhood G of the point y so that for some C^{1+}-smooth function $f\colon G \to \mathbf{R}$ with nonvanishing gradient and a suitable constant γ we have $G \cap F = \{x\colon f(x) = \gamma\}$.

Theorem B. *If $A \subset \mathbf{R}^n$ is a closed set with $\partial A \neq \emptyset$ such that $\varrho_0 \equiv \operatorname{reach} A > 0$ then the function $f(.) \equiv \operatorname{dist}(., A)$ is C^{1+}-smooth on the domain $G \equiv \equiv \{x \in \mathbf{R}^n\colon 0 < \operatorname{dist}(x, A) < \varrho_0\}$. The figures $\partial(A_\varrho) = \{x\colon \operatorname{dist}(x, A) = \varrho\}$ $(0 < \varrho < \varrho_0)$ are $(n-1)$-dimensional C^{1+}-submanifolds of $\mathbf{R}^n$. By setting $B \equiv A_{\varrho_1}$, we have $\operatorname{reach} B \geqq \varrho_1$ and $\partial(A_\varrho) = \partial(B_{\varrho-\varrho_1})$ whenever $0 < \varrho < \varrho_1 \leqq \varrho_0$, that is, also introducing parallle sets of negative radius[2]) we have $\partial(A_\varrho) = \partial[(A_{\varrho_1})_{\varrho-\varrho_1}]$ for all $0 < \varrho < \infty$. The main curvatures $\varkappa_1(p), \dots, \varkappa_{n-1}(p)$ of the hypersurface ($(n-1)$-dimensional C^{1+}-submanifold) $M \equiv \partial(A_{\varrho_1})$ of $\mathbf{R}^n$ oriented by its normal $\operatorname{grad} f$ exist at vol_{n-1}-almost every point $p \in M$ and their elementary symmetrical polynomials, i.e. the functions $\varkappa_1(.) + \dots + \varkappa_{n-1}(.), \dots, \varkappa_1(.) \dots \varkappa_{n-1}(.)$, are vol_{n-1}-measurable. Further, we have $-1/(\varrho_0 - \varrho_1) \leqq \varkappa_i \leqq 1/\varrho_1$ $(i = 1, \dots, n-1)$. If T is any subset of $\mathbf{R}^n$ formed by the union of some prenormals of the set A such that $T \cap A\varrho_0$ is vol_n-measurable then, for $0 < \varrho < \operatorname{reach} A$,*

$$\operatorname{vol}_{n-1} T \cap \partial A_\varrho = \int_{T \cap M} [1 + (\varrho - \varrho_1)\varkappa_1] \dots [1 + (\varrho - \varrho_1)\varkappa_{n-1}]\, d \operatorname{vol}_{n-1} \tag{2}$$

and

$$\operatorname{vol}_n T \cap A_\varrho = \int_0^\varrho \int_{T \cap M} [1 + (\tau - \varrho_1)\varkappa_1] \dots [1 + (\tau - \varrho_1)\varkappa_{n-1}]\, d \operatorname{vol}_{n-1}\, d\tau. \tag{2'}$$

Proof. See sections "Sets with positive reach" and "Curvature measures" in [1].

We remark that the connection between (2) and (2′) is established by the following more general observation:

Lemma 1. *If $\emptyset \neq A \subset \mathbf{R}^n$ and T is a vol_n-measurable subset of $\mathbf{R}^n \setminus \bar{A}$ then*

$$\operatorname{vol}_n T = \int_0^\infty (\operatorname{vol}_{n-1} T \cap \partial(A_\varrho))\, d\varrho. \tag{3}$$

Proof. See e.g. [3] p. 271.

[2]) For $\delta < 0$ and $A \subset \mathbf{R}^n$, $A_\delta \equiv \{x \in \mathbf{R}^n\colon \operatorname{dist}(x, \mathbf{R}^n \setminus A) > -\delta\}$.

3. A separability argument

Definition. We shall call a subset $S\neq\emptyset$ of the product space $\mathbf{R}^n\times\mathbf{R}^n$ a *generalized oriented surface* (GOS) if for all $(y, k)\in S$ we have $\|k\|=1$ and one can find an $\varepsilon>0$ (depending on (y, k)) so that

$$\operatorname{dist}(y, y+\varrho k) = \varrho \geqq \operatorname{dist}(y', y+\varrho k) \quad \text{for any} \quad (y', k')\in S \quad \text{and} \quad 0\leqq\varrho\leqq\varepsilon.$$

If A is a non-empty proper subset of $\mathbf{R}^n$ then let d^+A denote the figure in $\mathbf{R}^n\times\mathbf{R}^n$ defined by

$$d^+A \equiv \{(y, k)\colon y\in\partial A,\ \|k\| = 1 \text{ and } \exists L \text{ prenormal of } A\ \ L\supset y+(0, \text{length}\, L)\cdot k\}.$$

It is clear from Theorem A that all the sets d^+A are GOS-s.

Lemma 2. *Suppose that A is a subset of non-empty compact boundary in* $\mathbf{R}^n$ *with* $\varrho_0\equiv\text{reach}\, A>0$. *Then*

a) *the figure d^+A is compact (with respect to the topology of* $\mathbf{R}^n\times\mathbf{R}^n$*)*

b) *the mapping* $\Phi\colon(d^+A)\times(0, \varrho_0)\to\mathbf{R}^n$, $\Phi((y, k), \varrho)\equiv y+\varrho\cdot k$ *is a homeomorphism between the sets* $(d^+A)\times(0, \varrho_0)$ *and* $A_{\varrho_0}\setminus\bar{A}$, *and* $\Phi(d^+A\times\{\varrho\})=\partial A_\varrho$ *whenever* $0<\varrho<\varrho_0$.

Proof. a) The GOF d^+A is bounded in $\mathbf{R}^n\times\mathbf{R}^n$ because it is contained in the product of the compact figures ∂A and $\partial\mathbf{B}^n=\{k\in\mathbf{R}^n\colon\|k\|=1\}$. On the other hand, it is also closed since in case of any sequence $\{(y_i, k_i)\colon i\in I\}\subset d^+A$ with $(y_i, k_i)\to(y, k)$ we necessarily have $y\in\partial A$ and $\|k\|=1$, and for $0<\varrho<\varrho_0$ the equalities $\varrho=\operatorname{dist}(y_i+\varrho\cdot k_i, y_i)=\operatorname{dist}(y_i+\varrho\cdot k_i, \partial A)=\operatorname{dist}(y_i+\varrho\cdot k_i, A)$ imply (by continuity of the function $\operatorname{dist}(.\,, A)$) $\varrho=\operatorname{dist}(y+\varrho\cdot k, y)=\operatorname{dist} y+\varrho\cdot k, A)$ i.e, $y\in pr_A(y+\varrho k)$. This shows that $\{y\}=pr_A(y+\varrho k)$ (since $\varrho<\text{reach}\, A$). Therefore, by taking $L\equiv\{y+\varrho\cdot k\colon 0<\varrho<\infty \text{ and } \{y\}=pr_A(y+\varrho k)\}$, we obtain from Theorem A that L is a prenormal of A and $L=y+(0, \text{length}\, L)k$ i.e. $(y, k)\in d^+A$.

b) By Theorem A and the definition of d^+A, the condition reach $A=\varrho_0>0$ means that the mapping Φ is one-to-one. By fixing an arbitrary pair ϱ_1, ϱ_2 such that $0<\varrho_1<\varrho_2<\varrho_0$, we see that the figure $D(\varrho_1, \varrho_2)\equiv(d^+A)\times[\varrho_1, \varrho_2]$ is a compact subset of dom Φ (since the GOS d^+A is compact). Since Φ is obviously continuous, $\Phi|D(\varrho_1, \varrho_2)$ is a homeomorphism (because the inverse of any continuous function with compact domain between Hausdorff spaces is coontinuous). But then the inverse of Φ is necessarily continuous over the open set $A_{\varrho_2}\setminus\overline{A_{\varrho_1}}$ contained in $\Phi(D(\varrho_1, \varrho_2))$. Thus the relation range $\Phi=A_{\varrho_0}\setminus\bar{A}=\bigcup_{0<\varrho_1<\varrho_2<\varrho_0}(A_{\varrho_2}\setminus\overline{A_{\varrho_1}})$ immediately implies continuity of Φ^{-1}.

Lemma 3. *Let A, ϱ_0, Φ be defined as in the previous lemma with the same assumptions. Then there exists a Borel measure μ and there are μ-measurable func-*

tions $a_0, \ldots, a_{n-1}$ *over* d^+A *such that for each* $0<\varrho<\varrho_0$ *and* vol_{n-1}*-measurable* $F\subset\partial(A_\varrho)$, *we have*

$$\mathrm{vol}_{n-1}F=\int_{d^+A} 1_F(y+\varrho\cdot k)\sum_{j=0}^{n-1} a_j(y,k)\varrho^j\,d\mu(y,k). \tag{4}$$

(Here $1_F(.)$ *stays for the characteristic function of* F.*)*

Proof. Fix (arbitrarily) a value $0<\varrho_1<\varrho_0$. Consider the mapping $\Psi(.)\equiv\equiv(.,\varrho_1)$. Observe that $\Psi\colon d^+A\leftrightarrow\partial(A_{\varrho_1})$ and that $\Phi(\Psi^{-1}(.),\varrho)\colon\partial(A_{\varrho_1})\leftrightarrow\partial(A_\varrho)$ for $0<\varrho<\varrho_0$ are homeomorphisms. Therefore the measure

$$d\mu\equiv d\,\mathrm{vol}_{n-1}\circ\psi\ ^{3)} \tag{5'}$$

is a Borel measure on d^+A. Further, if $\varkappa_1,\ldots,\varkappa_{n-1}$ denote the main curvatures of the hypersurface $M\equiv\partial(A_{\varrho_1})$ oriented by its normal directed outward from A_{ϱ_1} then the functions $a_0,\ldots,a_{n-1}$ defined implicitly by

$$\begin{gathered}[1+(\tau-\varrho_1)\cdot\varkappa_1(y+\varrho_1k)]\ldots[1+(\tau-\varrho)\cdot\varkappa_{n-1}(y+\varrho_1k)]\equiv\sum_{j=0}^{n-1}a_j(y,k)\tau^j\\ (\text{for } 0<\tau<\varrho_0,\ (y,k)\in d^+A)\end{gathered} \tag{5''}$$

are μ measurable (cf. Theorem B). Now let $T(F)$ denote the union of those prenormals of A which intersect F (the surface piece of $\partial(A_\varrho)$ occurring in (4)). Then we have $T(F)\cap A_{\varrho_0}=\Phi(\Psi^{-1}(F)(0,\varrho_0))$. This shows that for any Borel measurable F, the figure $T(F)\cap A_{\varrho_0}$ is also Borel measurable. Then performing the substitutions (5′) and (5″) in the right hand side of (4), we obtain from Theorem B (cf. also (2)) that (4) holds for any Borel subset F of $\partial(A_\varrho)$. Hence we derive (4) for any vol_{n-1} measurable F from the Borel regularity of the measures vol_{n-1} and μ, respectively.

Remark. a) It is clear that the system $\mu, a_0,\ldots,a_{n-1}$ is not uniquely determined. However, it is discovered from the proof that the measures $d\nu\equiv a_0d\mu,\ldots,d\nu_{n-1}\equiv\equiv a_{n-1}d\mu$ depend only on the GOS d^+A (in the sence that if $A^{(1)}$ and $A^{(2)}$ are sets in $\mathbf{R}^n$ of positive reach and $(\mu^{(i)},a_0^{(i)},\ldots,a_{n-1}^{(i)})$ are systems satisfying (4) for $A=A^{(i)}$ $(i=1,2)$, respectively, then for the measures $d\nu_j^{(i)}\equiv a_j^{(i)}d\mu^{(i)}$ $(i=1,2,;\ j=0,\ldots,n-1)$ we have

$$d\nu_j^{(1)}|(d^+A^{(1)})\cap(d^+A^{(2)})=d\nu_j^{(2)}|(d^+A^{(1)})\cap(d^+A^{(2)})\quad(j=0,\ldots,n-1).$$

b) For any $(y,k)\in d^+A$, the roots of the polynomial $\sum_{j=0}^{n-1}a_j(y,k)\varrho^j$ are real (cf. (5″)) and lie outside of the open interval $(0,\varrho_0)$ (cf. with the relations $-1/(\varrho_0-\varrho_1)\leqq\varkappa_1,\ldots,\varkappa_{n-1}\leqq 1/\varrho_1$ in Theorem B).

---

[3]) The measure $\mathrm{vol}_{n-1}\circ\Psi$ is defined on the family of subsets of d^+A $\mathcal{F}\equiv\{\Psi^{-1}(E)\colon E\subset\partial(A_{\varrho_1}),$ E is vol_{n-1}-measurable$\}$ by $(\mathrm{vol}_{n-1}\circ\Psi)(D)\equiv\mathrm{vol}_{n-1}(\Psi(D))$ for any $D\in\mathcal{F}$.

Corollary (with the notations and assumptions of Lemma 2). *Formula* (4) *implies that for all* vol_n*-measurable subsets* T *of* $A_{\varrho_0}\setminus\bar{A}(=\Phi((d^+A)\times(0,\varrho_0)))$ *we have*

$$\mathrm{vol}_n T = \int_{d^+A}\int_0 1_T(y+\varrho\cdot k)\sum_{j=0}^{n-1} a_j(y,k)\varrho^j\,d\varrho\,d\mu(y,k). \tag{4'}$$

Proof. Consider the family of surface pieces $F(\varrho)\equiv T\cap\partial(A_\varrho)$. For $\varrho\geqq\varrho_0$ we have $F(\varrho)=\emptyset$ and for almost every $0<\varrho<\varrho_0$, $F(\varrho)$ is a vol_{n-1}-measurable subset of $\partial(A_\varrho)$. Thus we can apply Lemma 2 for almost every $0<\varrho<\varrho_0$ whence we obtain that

$$\mathrm{vol}_{n-1} T\cap\partial(A_\varrho) = \mathrm{vol}_{n-1} F(\varrho) = \int_{d^+A} 1_{T\cap\partial(A_\varrho)}(y+\varrho k)\sum_{j=0}^{n-1} a_j(y,k)\varrho^j\,d(y,k) =$$

$$= \int_{d^+A} 1_T(y+\varrho k)\sum_{j=0}^{n-1} a_j(y,k)\varrho^j\,d\mu(y,k).$$

Hence, by Lemma 1,

$$\mathrm{vol}_n T = \int_0^{\varrho_0}\int_{d^+A} 1_T(y+\varrho k)\sum_{j=0}^{n-1} a_j(y,k)\varrho^j\,d\mu(y,k)\,d\varrho.$$

Observe that in the above formula, $y+\varrho k=\Phi((y,k),\varrho)$ stays in the argument of the function $1_T(.)$. Since Φ is a homeomorphism between $(d^+A)\times(0,\varrho_0)$ and $A_{\varrho_0}\setminus\bar{A}$ and since the measures $d\mu$, $d\,\mathrm{vol}_n$ and $d\varrho$ are Borel regular measures, respectively, this means that the product measure $d\tau\equiv d\mu\times d\varrho$ (i.e. $=d\mu\times d\,\mathrm{vol}_1$) satisfies

$$\mathrm{vol}_n T = \int_{(d^+A)\times(0,\varrho_0)} 1_T(y+\varrho k)\sum_{j=0}^{n-1} a_j(y,k)\varrho^j\,d\tau(y,k,\varrho).$$

This immediately yields (4′) by Fubini's theorem.

Notation. If $A\subset\mathbf{R}^n$ is closed with $\partial A\neq\emptyset$, then for any $(y,k)\in d^+A$ let $L^A(y,k)$ denote in the sequel the prenormal of A issued from the point $y(\in\partial A)$ in the direction of the vector k, and let $h^A(y,k)$ denote the length of the line segment $L^A(y,k)$.

Remark. It is easy to see that the value reach A is not other than the greatest lower bound of the function h^A (i.e. reach $A=\inf h^A(=\inf\{h^A(y,k)\colon(y,k)\in d^+A\})$).

Lemma 4. *Let* A *be closed and* $\partial A\neq\emptyset$.

a) *For any* $\varepsilon>0$, *the GOS* $\{(y,k)\in d^+A\colon h^A(y,k)\geqq\varepsilon\}$ *is closed (in* $\mathbf{R}^n\times\mathbf{R}^n$*)*

b) d^+A *is Borel measurable (moreover it is an* $\mathscr{F}_\sigma$*).*

c) *For almost every* $\varrho>0$, *the set* $\partial(A_\varrho)$ *is of* σ*-finite* vol_{n-1}*-measure.*

d) *For the set* $Z^*\equiv\{y+h^A(y,k)k\colon(y,k)\in d^+A$ with $h^A(y,k)<\infty\}$, *we have* $\mathrm{vol}_{n-1}Z^*\cap\partial(A_\varrho)=0$ *except for countably many values of* $\varrho>0$.

Proof. a) From Theorem A we know that

$$d^+A = \{(y, k)\colon \exists x \in \mathbf{R}^n \setminus A,\ y \in \mathrm{pr}_A x \text{ and } k = (x-y)/\|x-y\|\}. \tag{6}$$

Now if $\{(y_i, k_i)\colon i \in I\} \subset d^+A$ is a convergent sequence such that $h^A(y_i, k_i) \geqq \varepsilon$ $(i \in I)$ and $(y_i, k_i) \to (y, k)$, then for $x \equiv y + \varepsilon k$ we have $x_i \to x$ and $y_i \in pr_A x_i$ with $k_i = (x_i - y_i)/\|x_i - y_i\|$ (for all $i \in I$). Since, in general, the condition $y' \in pr_A x'$ is equivalent to $\mathrm{dist}\,(x', A) = \mathrm{dist}\,(x', y')$, we infer from the continuity of the functions $\|.\|$ and $\mathrm{dist}\,(.\,, A)$ that $\mathrm{dist}\,(x, y) = \mathrm{dist}\,(x, A) = \varepsilon$ i.e. $y \in pr_A x$ and $k = (x-y)/\|x-y\|$. This shows by (6) that $(y, k) \in d^+A$.

b) Since $d^+A = \bigcup_{m=1}^{\infty} \{(y, k)\colon h^A(y, k) \geqq 1/m\}$.

c) Applying Lemma 1 we obtain

$$\infty > \mathrm{vol}_n\,[r\mathbf{B}^n \cap (\mathbf{R}^n \setminus A)] = \int_{-\infty}^{\infty} \mathrm{vol}_{n-1}\,[r\mathbf{B}^n \cap \partial(A_\varrho)]\,d\varrho$$

for all $r > 0$ [4]). Thus for any $r > 0$, there exists a set $\Delta_r \subset (0, \infty)$ such that $\mathrm{vol}_{n-1}\,[r\mathbf{B}^n \cap \partial(A_\varrho)] < \infty$ whenever $\varrho \in (0, \infty) \setminus \Delta_r$. Thus if $\varrho \notin \bigcup_{m=1}^{\infty} \Delta_m$ then the vol_{n-1}-measure of $\partial(A_\varrho)$ $(= \bigcup_{m=1}^{\infty} [m\mathbf{B}^n \cap \partial(A_\varrho)])$ is σ-finite.

d) Fix (an arbitrary) $\delta > 0$ such that $\partial(A_\delta)$ has σ-finite vol_{n-1}-measure, and for all $\varrho > \delta$ let Λ_ϱ denote the binary relation $\Lambda_\varrho \equiv \{(x, z)\colon z \in \partial(A_\varrho),\ x \in pr_{\bar{A}_\delta} z\} (= \{(y + \delta k, y + \varrho k)\colon (y, k) \in d^+A \text{ and } h^A(y, k) \geqq \varrho\}$. Now we know (see [4] p 254) that for distinct $z_1, z_2\ (\in \mathbf{R}^n)$ there cannot be found any $x (\in \mathbf{R}^n)$ with (x, z_1), $(x, z_2) \in \Lambda_\varrho$ and that the mapping λ_ϱ defined by $\lambda_\varrho(x) = z \overset{\text{def}}{\Leftrightarrow} (x, z) \in \Lambda_\varrho$ is Lipschitzian with $\mathrm{dom}\, \lambda_\varrho = pr_{\bar{A}_\varrho} \partial(A_\varrho)$ and $\mathrm{range}\, \lambda_\varrho = \partial(A_\varrho)$ for any $\varrho > \delta$. So for each $\varrho > 0$, we have $\mathrm{vol}_{n-1}\, Z^* \cap \partial(A_\varrho) = 0$ whenever the set $\lambda_\varrho^{-1}(Z^* \cap \partial(A_\varrho)) = \{y + \delta k\colon h^A(y, k) = \varrho\}$ has vol_{n-1}-measure 0. But the sets $\{y + \delta k\colon h^A(y, k) = \varrho\}$ $(\varrho > \delta)$ are all pairwise disjoint subsets of $\partial(A_\delta)$. From a) we infer that they are Borel measurable. Therefore the σ-finiteness of $\mathrm{vol}_{n-1}\, \partial(A_\delta)$ implies that there exist at most countably many $\varrho > \delta$ such that $\mathrm{vol}_{n-1}\{y + \delta k\colon h^A(y, k) = \varrho\} > 0$. This suffices for d) since the value of $\delta > 0$ can be chosen arbitrarily small.

Theorem 1. *Let $A \subset \mathbf{R}^n$ be closed and $\partial A \neq \emptyset$. If one can find a sequence $A^1, A^2, \ldots (\subset \mathbf{R}^n)$ of sets with non empty compact boundary such that*

a) $d^+A \subset \bigcup_{i=1}^{\infty} d^+A^i$,

b) $h_i \equiv \mathrm{reach}\, A^i > 0$ *for* $i = 1, 2, \ldots,$

[4]) $\mathbf{B}^n$ is the standard notation for the open unit ball of $\mathbf{R}^n$.

c) *for all* $(y,k)\in d^+A$ *we have* $h^A(y,k)\equiv\sup\{h_i:(y,k)\in d^+A^i\}$, *then there exists a Borel measure* μ *on* d^+A *and there are* μ*-measurable functions* $a_0,\dots,a_{n-1}$ *(over* d^+A*) such that*

$$\mathrm{vol}_n T=\int\limits_{d^+A}\int\limits_0^{h^A(y,k)} 1_T(y+\varrho k)\sum_{j=0}^{n-1} a_j(y,k)\varrho^j\,d\varrho\,d\mu(y,k) \tag{7}$$

for all vol_n*-measurable* $T\subset\mathbf{R}^n\setminus A$.

Proof. Set $S_1\equiv(d^+A)\cap(d^+A^1),\dots,S_i\equiv[(d^+A)\cap(d^+A^i)]\setminus\bigcup_{j<i}S_j,\dots$ and for $i=1,2,\dots$ let $(\mu^i,a_0^i,\dots,a_{n-1}^i)$ be a fixed system satisfying (7) (putting A^i in the place of A, μ^i instead of μ etc. in Lemma 3). Now $S_1,S_2,\dots$ is a sequence of Borel-measurable GOS-s forming a partition of d^+A. We also have $S_i\subset d^+A^i$ $(i=1,2,\dots)$. So we can define the system $(\mu,a_0,\dots,a_{n-1})$ in the following way:

$$\mu(E)\equiv\sum_{i=1}^{\infty}\mu^i(E\cap S_i)\quad\text{for}\quad E\subset d^+A\quad(\Leftrightarrow d\mu|S_i\equiv d\mu^i|S_i\ \text{for}\ i=1,2,\dots) \tag{8'}$$

(in the sense that a set E is μ-measurable if and only if for all indices i, the sets $E\cap S_i$ are μ^i-measurable), and

$$a_j(y,k)\equiv a_j^i(y,k)\quad\text{for}\quad(y,k)\in S_i\quad(j=0,\dots,n-1\ \text{and}\ i=1,2,\dots). \tag{8''}$$

Consider now a simple Borel function $f\colon d^+A\to[0,\infty]$ such that $f<h^A$ and range $f=\{c_1,c_2,\dots\}$, and set $G_f\equiv\{y+\varrho k:(y,k)\in d^+A,\ 0<\varrho<f(y,k)\}$. Then it easily follows from Lemma 3 that

$$\mathrm{vol}_n T\cap G_f=\int\limits_{d^+A}\int\limits_0^{f(y,k)} 1_T(y+\varrho k)\sum_{j=0}^{n-1} a_j(y,k)\varrho^j\,d\varrho\,d\mu(y,k) \tag{9}$$

for each vol_n-measurable $T\subset\mathbf{R}^n\setminus A$.

To prove (9), take the following Borel-measurable partition $\{S_{im}:i,m=1,2,\dots\}$ of d^+A defined by

$$S_{im}\equiv\{(y,k)\in f^{-1}(\{c_m\}):\ i\ \text{is the smallest index with}\ (y,k)\in d^+A^i\ \text{and}\ h_i>c_m\}.$$

Then consider the partition $\{B_{im}:i,m=1,2,\dots\}$ of G_f defined by $B_{im}\equiv\{y+\varrho k:(y,k)\in S_{im},\ 0<\varrho<c_m\}$. Then fix an arbitrary pair of indices i,m. Applying Lemma 2b) to A^i, we see that the domain B_{im} is Borel measurable. Since for any $(y,k)\in S_{im}$ and $0<\varrho<h^A(y,k)$ we have $1_{T\cap B_{im}}(y+\varrho k)=1_T(y+\varrho k)\cdot 1_{S_{im}}(y,k)1_{(0,c_m)}(\varrho)$, using Lemma 3 (with A^i instead of A and

with $\varrho_0=h_i$), we have

$$\operatorname{vol}_n T\cap B_{im}=\int_{d^+A}\int_0^{h_i} 1_T(y+\varrho k)\cdot 1_{S_{im}}(y,k)1_{(0,c_m)}(\varrho)\sum_{j=0}^{n-1}a_j^i(y,k)\varrho^j\,d\varrho\,d\mu(y,k)=$$

$$=\int_{S_{im}}\int_0^{c_m} 1_T(y+\varrho k)\sum_{j=0}^{n-1}a_j(y,k)\varrho^j\,d\varrho\,d\mu(y,k)=$$

$$=\int_{d^+A}\int_0^{f(y,k)} 1_T(y+\varrho k)\cdot 1_{S_{im}}(y,k)1_{(0,c_m)}(\varrho)\sum_{j=0}^{n-1}a_j(y,k)\varrho^j\,d\varrho\,d\mu(y,k).$$

Summing this for $i,m=1,2,\dots$, we obtain (9).

In possession of (9) we can conclude as follows: Lemma 4a) shows that the function $h^A\colon d^+A\to(0,\infty]$ is Borel-measurable (moreover that it is upper semi-continuous). Therefore there exists a sequence $0\le f_1\le f_2\le\dots$ of simple Borel-functions such that $f_i\nearrow h^A$ (pointwise). For any such a sequence $\{f_i\}_1^\infty$, we have $\bigcup_{i=1}^\infty G_{f_i}=\{y+\varrho k\colon (y,k)\in d^+A,\ 0<\varrho<h^A(y,k)\}=\mathbf{R}^n\setminus(A\cup Z^*)$ where $Z^*\equiv\equiv\{y+h^A(y,k)\cdot k\colon h^A(y,k)<\infty\}$. So, for $i\to\infty$, it follows from (9) that

$$\operatorname{vol}_n T\setminus Z^*=\int_{d^+A}\int_0^{h^A(y,k)} 1_T(y+\varrho k)\sum_{j=0}^{n-1}a_j(y,k)\varrho^j\,d\varrho\,d\mu(y,k). \tag{7'}$$

But now the relation $Z^*=(\mathbf{R}^n\setminus A)\setminus\bigcup_{i=1}^\infty G_{f_i}$ shows that Z^* is a Borel-set. Thus we may apply Lemma 1 to Z^* (in place of T there) which implies (by Lemma 4d)) that $\operatorname{vol}_n Z^*=0$.

4. Some convexity properties of parallel sets

Our aim in this section will be to prove that there always exist sets $A^1, A^2, \dots$ satisfying the conditions of Theorem 1.

Lemma 5. *Let $x_0\in\mathbf{R}^n$ and $\varrho_0>0$. Then the function* $g(.)\equiv\operatorname{dist}(.,x_0)-\frac{1}{2\varrho_0}\|.\|^2$ *is concave on the domain* $G\equiv\{x\colon\operatorname{dist}(x,x_0)>\varrho_0\}$. (A function f is said to be *concave on a domain H* if it is concave in the usual sence when restricted to any convex subset of H.)

Proof. Evaluate the eigenvalues of the second derivative tensor[5]) of the function f at a point $x_1\in G$. It is convenient to use a Cartesian coordinate system

[5]) The second derivative tensor of a function $f\colon H(\subset\mathbf{R}^n)\to\mathbf{R}$ at a point $x\in H$ is considered here as the bilinear form $D_2f(x)\colon\mathbf{R}^n\times\mathbf{R}^n\to\mathbf{R}$, $(v_1,v_2)\mapsto\partial_{v_1}\partial_{v_2}f(x)$ where the symbol ∂_v means the directional derivation in the direction $v(\in\mathbf{R}^n)$ i.e. $\partial_v f(y)\equiv\lim_{\lambda\searrow 0}\lambda^{-1}[f(y+\lambda v)-f(y)]$.

with origin x_0 and first unit vector $e_1=\frac{x_1-x}{\|x_1-x\|}$. Then, independently of the choice of the further basic vectors $e_2, \ldots, e_n$, the function $f(.)\equiv\operatorname{dist}(., x)$ is represented by the form $\varphi(\xi_1, \ldots, \xi_n)=f(x_0+\xi_1 e_1+\ldots+\xi_n e_n)=\sqrt{\xi_1^2+\ldots+\xi_n^2}$ in this coordinate system. Since $x_1=x_0+\|x_1-x_0\| e_1$, the eigenvalues of $D_2 f(x_1)$ coincide with those of the matrix $M\equiv\left(\frac{\partial^2\varphi}{\partial\xi_i\partial\xi_j}\Big|_{(\|x_1-x_0\|, 0, \ldots, 0)}\right)_{i,j=1}^n$. But it is easy to see that M is of diagonal form with $0, \|x_1-x_0\|^{-1}, \ldots, \|x_1-x_0\|^{-1}$ in its main diagonal. On the other hand, $D_2\|.\|^2$ is represented in any Cartesian system by the matrix $I\equiv(2\cdot\delta_{ij})_{i,j=1}^n$ (δ_{ij} denotes the "Kronecker δ"). Therefore the eigenvalues of $D_2 f(x_1)$ are $-\frac{1}{\varrho_0}$ and $\|x_1-x_0\|^{-1}-\frac{1}{\varrho_0}$ (with multiplicity $n-1$), all negative numbers. This completes the proof by recalling that any function of negative definite second derivative tensor is concave on any open convex subset of its domain.

Theorem 2. *Let $A\subset\mathbf{R}^n$ be such that $\partial A\neq\emptyset$ and fix $\varrho_0>0$. Then the function $g(.)\equiv\operatorname{dist}(., A)-\frac{1}{2\varrho_0}\|.\|^2$ is concave on the domain $G\equiv\{x\in\mathbf{R}^n: \operatorname{dist}(x, A)>\varrho_0\}$.*

Proof. f is the infimum of the function family $F\equiv\left\{\operatorname{dist}(., A)-\frac{1}{2\varrho_0}\|.\|^2: x\in A\right\}$. By Lemma 5, all members of F are concave functions on G. But the infimum of any family of concave functions in concave.

Corollary. *All directional derivatives of the function $f(.)\equiv\operatorname{dist}(., A)$ exist in $\mathbf{R}^n\setminus A$. For a fixed $x_0\in\mathbf{R}^n\setminus A$, the function $t\mapsto\partial_t f(x_0)$ is continuous and superlinear (i.e. positive homogeneous and concave).*

Proof. Apply Theorem 2 with $\varrho_0\equiv\frac{1}{2}\operatorname{dist}(x_0, A)$. This shows that the function $g(.)=f(.)-\frac{1}{2\varrho_0}\|.\|^2$ is concave on some neighborhood of the point x_0. Therefore $\partial_t f(x_0)$ exists for all $t\in\mathbf{R}^n$ and satisfies $\partial_t f(x_0)=\partial_t g(x_0)+\frac{1}{\varrho_0}\langle t, x_0\rangle$. Thus $t\mapsto\partial_t f(x_0)$ is the sum of a continuous superlinear and a linear form of t (since the directional derivatives at a fixed point of any concave $\mathbf{R}^n\to\mathbf{R}$ function are continuous and superlinear.)

Theorem 3. *Let $A\subset\mathbf{R}^n$ be closed and $f(.)\equiv\operatorname{dist}(., A)$. Then for any $x_0\notin A$ and for any $t\in\mathbf{R}^n$ we have*

$$\partial_t f(x_0) = \min\left\{\left\langle t, \frac{y-x_0}{\|y-x_0\|}\right\rangle: y\in pr_A x_0\right\}.$$

Proof. Consider an arbitrary $y_0 \in pr_A x_0$. Now we have $f(x_0+\lambda t)-f(x_0)= = \operatorname{dist}(x_0+\lambda t, A)-\operatorname{dist}(x_0, A) = \operatorname{dist}(x_0+\lambda t, A)-\operatorname{dist}(x_0, y_0) \leqq \operatorname{dist}(x_0+\lambda t, y_0)- -\operatorname{dist}(x_0, y_0)$. Thus, by writing $h(.)\equiv\operatorname{dist}(., y_0)$, we obtain $\partial_t f(x_0)\leqq\partial_t h(x_0)= =\langle t, \operatorname{grad} h(x)_0\rangle=\left\langle t, \dfrac{y_0-x_0}{\|y_0-x_0\|}\right\rangle\leqq\min\left\{\left\langle t, \dfrac{y-x_0}{\|y-x_0\|}\right\rangle : y\in pr_A x_0\right\}$.

The proof of the inequality in the converse direction: Let us associate with any $x\in\mathbf{R}^n\setminus A$ a point $y(x)$ from the set $pr_A x$ and then let φ_x denote the function $\varphi_x(.)\equiv\operatorname{dist}(., y(x))$. Now we have $f=\inf_{x\in\mathbf{R}^n\setminus A}\varphi_x$ and for all $x\notin A$, $f(x)=\varphi_x(x)$. Thus, by writing $\psi(.)\equiv\varphi_{x_0+\lambda t}(.)$, we obtain

$$\frac{1}{\lambda}[f(x_0+\lambda t)-f(x_0)] \geqq \frac{1}{\lambda}[f(x_0+\lambda t)-\psi(x_0)] \geqq \frac{1}{\lambda}[\psi(x_0+\lambda t)-\psi(x_0)] \geqq \partial_t\psi(x_0)$$

for any arbitrarily fixed $t\in\mathbf{R}^n$ and $\lambda>0$. (The last inequality is a consequence of the convexity of ψ.) Hence from the relation $\operatorname{grad}\psi(x_0)=\dfrac{x_0-y(x_0+\lambda t)}{\|x_0-y(x_0+\lambda t)\|}$, we deduce that

$$\frac{1}{\lambda}[f(x_0+\lambda t)-f(x_0)] \geqq \left\langle t, \frac{x_0-y(x_0+\lambda t)}{\|x_0-y(x_0+\lambda t)\|}\right\rangle \quad \text{whenever} \quad \lambda>0. \tag{10}$$

Since for any bounded $G\subset\mathbf{R}^n\setminus A$ the set $\{y(x): x\in G\}$ is also bounded, there can be found a sequence $\lambda_i\searrow 0$ such that the sequence $\{y(x_0+\lambda_i t)\}_1^\infty$ be convergent. Fix such a sequence $\{\lambda_i\}_1^\infty$ and set $y^*\equiv\lim_i y(x_0+\lambda_i t)$. Now by (10) we have

$$\partial_t f(x_0) \geqq \left\langle t, \frac{x_0-y^*}{\|x_0-y^*\|}\right\rangle. \tag{10'}$$

On the other hand from the equivalence of the relations $\operatorname{dist}(x_0+\lambda_i t, A)= =\operatorname{dist}\big(x_0+\lambda_i t, y(x_0+\lambda_i t)\big)$ and $y(x_0+\lambda_i t)\in pr_A(x_0+\lambda_i t)$ we infer for $i\to\infty$ that $y^*\in pr_A x_0$. Thus for some $y^*\in pr_A x_0$, (10′) holds.

From now on, throughout the remaining part of this section, let A denote a fixed closed subset of $\mathbf{R}^n$, let $x_0\in\mathbf{R}^n\setminus A$ (also fixed), $r\equiv\operatorname{rad} pr_A x_0$ [6]), $\varrho\equiv\operatorname{dist}(x_0, A)$ and $f(.)\equiv\operatorname{dist}(., A)$.

Lemma 6. $\max_{t\neq 0}(\partial_t f(x_0)/\|t\|)=\sqrt{1-(r/\varrho)^2}$ *if* $r<\varrho$ *and* $\max_{t\neq 0}(\partial_t f(x_0)/\|t\|)$ 0 *if and only if* $r=\varrho$. *(Since* $pr_A x_0\subset\{y: \|y-x_0\|=\varrho\}$, *the possibility* $r>0$ *is excluded).*

[6]) For any set $H\subset\mathbf{R}^n$, $\operatorname{rad} H\equiv\inf\{\delta\geqq 0: \exists p\in\mathbf{R}^n\ H\subset p+\delta\overline{\mathbf{B}}^n\}$.

Proof. Since the function $t \mapsto \partial_t f(x_0)$ is superlinear and continuous, a simple compactness argument shows that $\max_{t \neq 0} \partial_t f(x_0)/\|t\|$ is always attained for some $t_0 \in \mathbf{R}^n$ with $\|t_0\|=1$. Now if $\partial_{t_0} f(x_0)>0$, then the set $pr_A x_0$ is contained in the spherical cap

$$K \equiv \{y \in \mathbf{R}^n \colon \|y-x\| = \varrho,\ \langle t_0, y-x_0\rangle \geqq \varrho \cdot \partial_{t_0} f(x_0)\}.$$

But then, by writing $p \equiv x_0 - (\varrho \cdot \partial_{t_0} f(x_0)) t_0$, we have $K \subset \{y \colon \|y-p\| \leqq \leqq \sqrt{\varrho^2 - (\varrho \cdot \partial_{t_0} f(x_0))^2}\}$. Thus $\partial_{t_0} f(x_0) > 0$ implies that $r \leqq \sqrt{1-(\partial_{t_0} f(x_0))^2}$ and therefore $\partial_{t_0} f(x_0) \geqq \sqrt{1-(r/\varrho)^2}$. •

On the other hand, if $r<\varrho$ then, because of the compactness of the set $pr_A x_0$, there exists a unique closed ball $B(\subset \mathbf{R}^n)$ of radius r such that $pr_A x_0 \subset B$. Consider the spherical cap $K' \equiv \{y \in B \colon \|y-x_0\| = \varrho\}$. It is not hard to prove that the closed ball $B'(\subset \mathbf{R}^n)$ of minimal radius containing the set K' is that whose center and radius coincide with those of the $(n-1)$-dimensional sphere $S' \equiv \equiv \{y \in \partial B \colon \|y-x_0\| = \varrho\}$, respectively. Since $pr_A x_0 \subset K' \subset B'$, we necessarily have $B'=B$. Let q denote the center of B and set $t_1 \equiv x_0 - q$. Since the point q is the center of S', we have $\text{angle}(t_1, y-q) = \pi/2$ for all $y \in S'$. Hence we deduce $\|t_1\|^2 = \sqrt{\|x_0-y\|^2 - \|y-q\|^2} = \sqrt{\varrho^2 - r^2}$ (with arbitrary $y \in S'$). Observe now that

$$K' = \{y \colon \|y-x_0\| = \varrho \text{ and } \text{angle}(t_1, y-q) \geqq \pi/2\} = \{y \colon \|y-x_0\| = \varrho,\ t\langle_1, y-q\rangle \leqq 0\}.$$

Therefore, by Theorem 5 we obtain

$$\partial_{t_1} f(x_0) \geqq \min\left\{\left\langle t_1, \frac{x_0-y}{\varrho}\right\rangle \colon \|x_0-y\| = \varrho,\ \langle t, y-q\rangle \leqq 0\right\} \geqq \left\langle t_1, \frac{x_0-q}{\varrho}\right\rangle = \|t_1\|^2/\varrho.$$

So $r<\varrho$ implies that $\max_{t \neq 0} \partial_t f(x_0)/\|t\| \geqq \|t_1\|/\varrho = \sqrt{1-(r/\varrho)^2}$.

Definition. We call a vector $t(\in \mathbf{R}^n)$ a *tangent vector* of a set $S(\subset \mathbf{R}^n)$ at the point $x \in S$ if $t=0$ if there is a sequence $x \neq x_1, x_2, \ldots \in S$ such that $x_i \to x$ and angle $(t, x_i - x) \to 0$ (for $i \to \infty$). $\left(\text{For } t_1, t_2 \in \mathbf{R}^n,\ \text{angle}(t_1, t_2) \equiv \arccos\left\langle \frac{t_1}{\|t_1\|}, \frac{t_2}{\|t_2\|}\right\rangle.\right)$ The set of the tangent vectors of S of x will be denoted by $\text{Tan}(x, S)$.

Lemma 7. *If $r<\varrho$ then for any $t \in \mathbf{R}^n$ we have*

a) $t \in \text{Tan}(x_0, \partial(A_\varrho))$ *if and only if* $\partial_t f(x_0) = 0$,

b) $t \in \text{Tan}(x_0, \mathbf{R}^n \setminus A_\varrho)$ *if and only if* $\partial_t f(x_0) \geqq 0$.

(I.e. $\text{Tan}(x_0, \mathbf{R}^n \setminus A_\varrho)$ *is a closed convex cone with non-empty interior and boundary and its boundary coincides with* $\text{Tan}(x_0, \partial(A_\varrho))$.)

Proof. Since $\mathbf{R}^n \setminus A_\varrho = \{x: f(x) \geqq \varrho\}$ and $f(x_0) = \varrho$, we can immediately establish that $\partial_t f(x_0) > 0$ implies $t \in \operatorname{Tan}(x_0, \mathbf{R}^n \setminus A_\varrho)$ and that in case of $t \in \operatorname{Tan}(x_0, \mathbf{R}^n \setminus A_\varrho)$ we have $\partial_t f(x_0) \geqq 0$. Therefore it suffices to prove just the statement a).

Since $\partial(A_\varrho) = \{x: f(x) = \varrho\}$, it is clear that $\partial_t f(x_0) = 0$ for all $t \in \operatorname{Tan}(x, \partial(A_\varrho))$. To prove $\partial_t f(x_0) = 0 \Rightarrow t \in \operatorname{Tan}(x_0, \partial(A_\varrho))$ we can proceed as follows. Let $C \equiv \{t: \partial_t f(x_0) = 0\}$ and $F(t) \equiv \partial_t f(x_0)$. From the continuity and superlinearity of the functional F it follows that C is a closed convex cone. Lemma 6 ensures that, for some $t_0 \in C$, we have $F(t_0) > 0$. Since there also exists a vector t_1 such that $F(t_1) < 0$ (e.g. the vector $t_1 \equiv y - x_0$ with an arbitrary $y \in pr_A x_0$), from the superlinearity and continuity of F we easily deduce that

$$F(t) > 0 \Leftrightarrow t \in \mathring{C} \text{ (the interior of } C), \quad F(t) = 0 \Leftrightarrow t \in \partial C, \text{ and } F(t) < 0 \Leftrightarrow t \notin C \, (\forall t \in \mathbf{R}^n).$$

Therefore we have to show that for any $0 \neq t \in \partial C$ and $\varepsilon > 0$ there exists a point $x \in \partial(A_\varrho)$ such that $0 < \|x - x_0\| < \varepsilon$ and $\operatorname{angle}(t, x - x_0) < \varepsilon$. But it is a directe corollary from continuity of F.

Lemma 8. *If S is any subset of $\mathbf{R}^n$, $x \in S$ and L denotes the smallest cone containing the unit vectors $k (\in \mathbf{R}^n)$ satisfying $(x, k) \in d^+ S$ then* $\operatorname{Tan}(x, S) \subset \operatorname{dual} L$ [7]) *(or which is the same* $L \subset \operatorname{dual} \operatorname{Tan}(s, S)$*).*

Proof. We must prove that in case of $(x, k) \in d^+ S$, for any $t \in \operatorname{Tan}(x, S)$ we have $\langle t, k \rangle < 0$. Proceed by contradiction. Suppose that $(x, k) \in d^+ S$ and $t \in \operatorname{Tan}(x, S)$ are such that $\langle t, k \rangle > 0$. Since the figure $\operatorname{Tan}(x, S)$ is a cone, we may assume without loss of generality that $\|t\| = 1$. Consider a sequence $x \neq x_1, x_2, \ldots \to x$ in S such that $\operatorname{angle}(t, x_i - x) \to 0$ $(i \to \infty)$ and set $h_i \equiv \|x_i - x\|$ and $t_i \equiv \frac{1}{h_i}(x_i - x)$ $(i = 1, 2, \ldots)$. Observe now that $t_i \to t$ and that for any arbitrarily fixed $\varrho' > 0$, the function $\psi(.) \equiv \operatorname{dist}(., x + \varrho' k)$ satisfies

$$\lim_i \frac{1}{h_i} [\operatorname{dist}(x_i, x + \varrho' k) - \varrho'] = \lim_i \frac{1}{h_i} [\psi(x + h_i t_i - \psi(x)] =$$

$$= \lim_{h \searrow 0} \frac{1}{h} [\psi(x + ht) - \psi(x)] = \partial_t \psi(x) = \langle t, k \rangle > 0.$$

This shows that $\operatorname{dist}(x_i, x + \varrho' k) < \varrho'$ holds for some index i. Thus we necessarily have $(y, k) \notin d^+ S$ by the arbitrariness of $\varrho' > 0$ and the definition of the GOS $d^+ S$.

[7]) For any set $H \subset \mathbf{R}^n$ we define its dual by $\operatorname{dual} H \equiv \{t \in \mathbf{R}^n: \forall u \in H \langle t, u \rangle \leqq 0\}$.

Remark. The converse inclusion $L \supset \text{dual Tan}(x, S)$ fails in general. Example: in $n=2$ dimensions for $S \equiv \{(\xi, \eta) \in \mathbf{R}^2 : \eta \leqq |\xi|^{3/2}\}$, $x \equiv (0, 0)$ and $k \equiv (0, 1)$ we have $\text{Tan}(x, S) = \{(\tau_1, \tau_2) : \tau_2 \leqq 0\} = \{t \in \mathbf{R}^2 : \langle k, t\rangle \leqq 0\}$ while $(y, k) \notin d^+ S$. However, one can conjecture that if $S \equiv \mathbf{R}^n \setminus A_\varrho$ and $x \equiv x_0$ then $L = \text{dual Tan}(x_0, \mathbf{R}^n \setminus A_\varrho)$ always holds. It will suit our requirements the following simpler special case:

Theorem 4. *Suppose $r < \varrho$. Then*

a) *the figure $D \equiv \{y : x_0 \in pr_{\mathbf{R}^n \setminus A_\varrho} y\}$ is convex and closed (this holds even for $r = \varrho$),*

b) *one can represent the set $D^0 \equiv \text{conv}(\{x_0\} \cup pr_A x_0)$ [8]) as the union of straight line segments issued from the point x_0 and of length $\sqrt{\varrho^2 - r^2}$.*

c) *If $L \equiv [0, \infty)\{k : (x_0, k) \in d^+(\mathbf{R}^n \setminus A_\varrho)\}$ then we have*

$$L = [0, \infty)(D - x_0) = [0, \infty)(D^0 - x) = \text{dual Tan}(x_0, \mathbf{R}^n \setminus A_\varrho)$$

d) *$h^{\mathbf{R}^n \setminus A_\varrho}(x_0, k) \geqq \sqrt{\varrho^2 - r^2}$ whenever $(x_0, k) \in d^+(\mathbf{R}^n \setminus A_\varrho)$.*

Proof. a) From the definition of $pr_{\mathbf{R}^n \setminus A_\varrho} y$ we infer that

$$D = \{y : \forall x \in \mathbf{R}^n \setminus A_\varrho,\ \text{dist}(y, x_0) \leqq \text{dist}(y, x)\} = \bigcap_{x \in \mathbf{R}^n \setminus A_\varrho} \{y : \|y - x_0\| \leqq \|y - x\|\}.$$

Thus D is the intersection of some family of closed half spaces (or $D = \mathbf{R}^n$ if $\{x_0\} = \mathbf{R}^n \setminus A_\varrho$).

b) For the sake of simplicity, we can assume (without loss of generality) that $x_0 = 0$.

It is well-known that, in general, the closed convex hull of any compact subset of $\mathbf{R}^n$ coincides with its algebraic convex hull. Hence

$$\text{conv}(\{x_0\} \cup pr_A x_0) =$$

$$= \Big\{\alpha \sum_1^m \lambda_i y_i : 0 \leqq \alpha \leqq 1,\ \lambda_1, \ldots, \lambda_m \geqq 0,\ \sum_1^m \lambda_i = 1 \text{ and } y_1, \ldots, y_m \in pr_A x_0\Big\}.$$

Thus we can write $D^0 = [0, 1] \cdot \text{conv}(pr_A x_0) = \bigcup \{[0, 1] \cdot c : c \in \text{conv}(pr_A x_0)\}$. Therefore it suffices to see that for any $c \in \text{conv}(pr_A x_0)$ we have $\|c\| \geqq \sqrt{\varrho^2 - r^2}$. Let t_0 be a unit vector such that $\partial_{t_0} f(x_0) = \sqrt{1 - (r/\varrho)^2}$ (its existence is established by Lemma 6).

[8]) For $H \subset \mathbf{R}^n$, conf H denotes the closed convex hull of H (i.e. the smallest closed convex subset of $\mathbf{R}^n$ containing H).

From Theorem 5 we infer that for any finite convex linear combination $c=\lambda_1 y_1+\ldots+\lambda_m y_m$ of some points of $pr_A x_0$ we have

$$\langle t_0, c\rangle = \sum_1^m \lambda_i \langle t_0, y_i\rangle = -\sum_1^m \lambda_i \langle t_0, x_0-y_i\rangle = -\varrho \left|\sum_1^m \lambda_i \left\langle t_0, \frac{x_0-y_i}{\|x_0-y_i\|}\right\rangle \geqq\right.$$

$$\geqq -\varrho \sum_1^m \lambda_i \partial_{t_0} f(x_0) = -\varrho\, \partial_{t_0} f(x_0) = -\sqrt{\varrho^2-r^2},$$

whence $\|c\|=\|t_0\|\cdot\|c\|\geqq|\langle t_0, c\rangle|=\sqrt{\varrho^2-r^2}$.

c) The relation $L=[0, \infty)(D-x_0)$ directly follows from the definitions. From Lemma 7b) and Theorem 5 we also have that $t\in \mathrm{Tan}\,(x_0, \mathbf{R}^n\setminus A_\varrho)\Leftrightarrow\partial_t f(x_0)\geqq 0\Leftrightarrow$ $\Leftrightarrow\forall y\in pr_A x_0 \langle t, x_0-y\rangle\geqq 0, \Leftrightarrow t\in \mathrm{dual}\,[(pr_A x_0)-x_0]\Leftrightarrow t\in \mathrm{dual}\,(D^0-x_0)\Leftrightarrow t\in \mathrm{dual}\,[0, \infty)\cdot$ $\cdot(D^0-x_0)$. Thus $\mathrm{Tan}\,(x_0, \mathbf{R}^n\setminus A_\varrho)=\mathrm{dual}\,[0, \infty)(D^0-x_0)$. Since both $\mathrm{Tan}\,(x_0, \mathbf{R}^n\setminus A_\varrho)$ and $[0, \infty)(D^0-x_0)$ are closed convex cones in $\mathbf{R}^n$, respectively, from Farkas's well-known theorem we infer $[0, \infty)(D^0-x_0)=\mathrm{dual}\,\mathrm{Tan}\,(x_0, \mathbf{R}^n\setminus A_\varrho)$. Then observe that from the definition of the set D it follows $x_0\in D$ and $pr_A x_0\subset D$. This implies by a) that $D^0\subset D$ and therefore $[0, \infty)(D^0-x_0)\subset[0, \infty)(D-x_0)$. At this point the proof of c) is completed by Lemma 8 which shows (for $S\equiv\mathbf{R}^n\setminus A_\varrho$ and $x\equiv x_0$) that $L\subset\mathrm{dual}\,\mathrm{Tan}\,(x, \mathbf{R}^n\setminus A_\varrho)$, since we have proved here $L=[0, \infty)(D-x_0)\supset$ $\supset[0, \infty)(D^0-x_0)=\mathrm{dual}\;\mathrm{Tan}\,(x_0, \mathbf{R}^n\setminus A_\varrho)$.

d) is immediate from b) and c).

Corollary. *If* $\varrho>0$, $A\subset\mathbf{R}^n$ *is closed and* $\mathrm{rad}\,A<\varrho$ *then* $h^{\mathbf{R}^n\setminus A_\varrho}\geqq$ $\geqq\sqrt{\varrho^2-(\mathrm{rad}\,A)^2}$.

Proof. Let $(x_0, k)\in d^+(\mathbf{R}^n\setminus A_\varrho)$. Now we have $x_0\in\partial(\mathbf{R}^n\setminus A_\varrho)=\partial(A_\varrho)$ and $r=\mathrm{rad}\,pr_A x_0\leqq\mathrm{rad}\,A<\varrho$. Thus Theorem 4d) can be applied.

5. Main Theorem

On the basis of the previous section we can construct the sets $A^1, A^2, \ldots$ required by Theorem 1.

Lemma 9. *For any closed subset A of the space $\mathbf{R}^n$ with $\partial A\neq\emptyset$ there exists a countable family $\mathcal{A}\equiv\{A^\alpha\colon \alpha\in I\}$ of subsets of $\mathbf{R}^n$ with positive reach and compact boundary such that* $\bigcup_{\alpha\in I} d^+A\supset d^+A^\alpha$ *and* $h^A(y, k)\leqq\sup\,\{\mathrm{reach}\,A^\alpha\colon (y, k)\in d^+A^\alpha\}$ *hold for any* $(y, k)\in d^+A$.

Proof. Let $\varrho_1, \varrho_2, \ldots$ be an enumeration of the positive rational numbers and for $i=1, 2, \ldots$ let the set B^i defined by $B^i\equiv\partial(A_{\varrho_i})$. Now we obtain from the definition of the function $h^A(:d^+A\to(0,\infty))$ that

$$B^i = \partial(A_{\varrho_i}) = \{y+\varrho_i k\colon (y,k)\in d^+A \text{ and } h^A(y,k)\geqq \varrho_i\} \quad (i=1,2,\ldots). \tag{11}$$

Then let each set B^i be covered by a countable family $K^{i,1}, K^{i,2}, \ldots$ of closed balls of radius $\varrho_i/(2i)$ and define the sets $A^{i,s}$ $(i,s=1,2,\ldots)$ as follows: set $G^{i,s}\equiv \equiv B^i\cap K^{i,s}$ and let $A^{i,s}\equiv\mathbf{R}^n\setminus(G^{i,s})_{\varrho_i}\,(=\{y\colon \mathrm{dist}\,(y, G^{i,s})\geqq\varrho_i\})$.

Observe that if $(y,k)\in d^+A$ is such that $h^A(y,k)\geqq\varrho_i$ and $y+\varrho_i k\in G^{i,s}$ then (for the same pair of indices i,s) we have $\mathrm{dist}\,(y+\varrho_i k, A^{i,s})=\varrho_i$ and hence $(y,k)\in d^+A^{i,s}$ $(i,s=1,2,\ldots)$. Since $\bigcup\limits_{s=1}^{\infty} G^{i,s}=B^i$, this means by (11) that

$$\{(y,k)\in d^+A\colon h^A(y,k)\geqq\varrho_i\}\subset\bigcup_{s=1}^{\infty} d^+A^{i,s} \quad (i=1,2,\ldots). \tag{12}$$

It follows from (12) that $d^+A\subset\bigcup\limits_{i,s=1}^{\infty} d^+A^{i,s}$.

Since the figure $G^{i,s}$ is contained in the ball $K^{i,s}$ whose radius equals to $\varrho_i/(2i)$, we have from the Corollary of Theorem 4 that reach $A^{i,s}=\inf h^{A^{i,s}}=\inf h^{\mathbf{R}^n\setminus(G^{i,s})_{\varrho_i}}\geqq \geqq\varrho_i\sqrt{1-1/(4i^2)}>0$ $(i,s=1,2,\ldots)$. So from (12) we also infer that

$$\sup\{\text{reach}\,A^{i,s}\colon (y,k)\in d^+A^{i,s}\}\geqq h^A(y,k)$$

for each $(y,k)\in d^+A$. Finally, the inclusions $\partial A^{i,s}=\partial[\mathbf{R}^n\setminus(G^{i,s})_{\varrho_i}]=\partial((G^{i,s})_{\varrho_i}]\subset \subset\overline{(G^{i,s})}_{\varrho_i}\subset\overline{(K^{i,s})}_{\varrho_i}$ immediately imply compactness of $\partial A^{i,s}$ $(i,s=1,2,\ldots)$. Thus the choice $A\equiv\{A^{i,s}\colon i,s=1,2,\ldots\}$ suits our requirements.

Theorem 5. *For every closed $A\subset\mathbf{R}^n$ of non-empty boundary there exists a Borel measure μ over the generalized oriented surface d^+A and there can be found μ-measurable functions $a_0(.),\ldots,a_{n-1}(.)$ such that for any Lebesgue integrable function $\varphi\colon\mathbf{R}^n\setminus A\to\mathbf{R}^n$ we have*

$$\int\limits_{\mathbf{R}^n\setminus A}\varphi\,d\mathrm{vol}_n=\int\limits_{d^+A}\int\limits_0^{h^A(y,k)}\varphi(y+\varrho k)\sum_{j=0}^{n-1}a_j(y,k)\varrho^j\,d\varrho\,d\mu(y,k)= \tag{13}$$

$$=\int\limits_D\varphi(y+\varrho k)\sum_{j=0}^{n-1}a_j(y,k)\varrho^j\,d\tau(y,k,\varrho)$$

where $D\equiv\{(y,k,\varrho)\colon(y,k)\in d^+A$ and $0<\varrho<h^A(y,k)\}$ and $d\tau$ denotes the product measure $d\mu\times d$length over (d^+A).

Proof. From Lemma 9 and Theorem 1 we immediately obtain (13) for characteristic functions of vol_n-measurable subsets of $\mathbf{R}^n\setminus A$. By taking linear combinations we can pass to simple $\mathbf{R}^n\setminus A\to\mathbf{R}$ functions and then a standard density argument establishes (13) for arbitrary Lebesgue integrable $\mathbf{R}^n\setminus A\to\mathbf{R}$ functions.

Corollary. *For μ-almost every $(y,k)\in d^+A$, the zeros of the polynomial $\sum_{j=0}^{n-1} a_j(y,k)\varrho^j$ are real and lie outside $(0, h^A(y,k))$.*

Proof. Recall the construction of the measure μ and the functions a_j in Theorem 1 (8′) and (8″). Applying the same notations (and definitions) as in Theorem 1, we can proceed as follows: From Remark a) after Lemma 3 we infer that for any fixed pair of indices i_1, i_2 one can write $a_j^{i_1}d\mu^{i_1}=a_j^{i_1}d\mu^{i_2}$ $(j=0,\dots,n-1)$ when restricted to the set $(d^+A^{i_1})\cap(d^+A^{i_2})$. This shows now that there exists a subset R^{i_1,i_2} of $(d^+A^{i_1})\cap(d^+A^{i_2})$ such that $\mu^{i_1}(R^{i_1,i_2})=\mu^{i_2}(R^{i_1,i_2})=0$ and there is a function $c_{i_1,i_2}\colon[(d^+A^{i_1})(d^+A^{i_2})]\setminus R^{i_1,i_2}\to(0,\infty)$ such that $a_j^{i_1}(y,k)=$ $=c_{i_1,i_2}(y,k)a_j^{i_2}(y,k)$ $(j=0,\dots,n-1)$ for any $(y,k)\in\mathrm{dom}\,c_{i_1,i_2}$. This is equivalent to the condition that the roots of the polynomials $\sum_{j=0}^{n-1}a_j^{i_1}(y,k)\varrho^j$ and $\sum_{j=0}^{n-1}a_j^{i_2}(y,k)\varrho^j$ are the same with the same multiplicity (for all $(y,k)\in\mathrm{dom}\,c_{i_1,i_2}$). Let then $(y,k)\in(d^+A)\setminus\bigcup_{i_1,i_2=1}^{\infty}R^{i_1,i_2}$ be arbitrarily fixed. Now Remark b) after Lemma 3 implies that the zeros of the polynomial $\sum_{j=0}^{n-1}a_j(y,k)\varrho^j$ are real and lie outside the interval $(0,\mathrm{reach}\,A^i)$ for any i, such that $(y,k)\in d^+A^i$. Therefore $p(.)$ cannot have any zero inside $\bigcup\{(0,\mathrm{reach}\,A^i)\colon(y,k)\in d^+A^i\}=(0,\sup\{\mathrm{reach}\,A^i\colon(y,k)\in d^+A^i\})\supset$ $\supset(0,h^A(y,k))$. Since by (8′) we have $\mu((d^+A)\cap\bigcup_{i_1,i_2=1}^{\infty}R^{i_1,i_2}=0$, the previous statement holds for μ-almost every $(y,k)\in d^+A$.

References

[1] H. Federer, Curvature measures, *Trans. American Math. Soc.*, **93** (1959), 418—491.
[2] G. Bouligand, *Introduction à la géométrie infinitésimale directe,* Vuibert (Paris, 1932).
[3] H. Federer, *Geometric measure theory,* Springer (New York, 1969).
[4] M. Kneser, Über den Rand von Parallelkörpern, *Math. Nachr.*, **5** (1951), 251—258.

BOLYAI INSTITUTE
UNIVERSITY SZEGED
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Acta Sci. Math., 41 (1979), 209—213

Type sets and nilpotent multiplications

A. E. STRATTON and M. C. WEBB

Introduction. The nilstufe, $\nu(G)$, of a torsion-free abelian group was defined by SZELE [7] to be the largest positive integer n such that there is an associative ring $(G, \circ)$ on G having a non-zero product of n elements. If no such integer exists then $\nu(G)$ is set equal to ∞. FEIGELSTOCK [2] defines the strong nilstufe, $N(G)$, in a similar manner but allows non-associative ring structures on G. In § 2 we define the *solvable degree* $\varrho(G)$ in an analogous way.

Several authors [1, 4, 5, 6, 7, 9, 10] have studied related problems of nilpotency in torsion-free rings. They have mainly restricted their attention to associative ring structures and have often demanded that the group G be completely decomposable. In [8] WEBB showed that if G is torsion-free with finite rank r then either $\nu(G)=\infty$ or $\nu(G)\leqq r$, and either $N(G)=\infty$ or $N(G)\leqq 2^{(r-1)}$.

In this note we obtain improved bounds on both $\nu(G)$ and $N(G)$ under certain conditions on the type set, $\mathbf{T}(G)$, of G. Here the *type set* of G means the partially ordered set of types $\mathbf{t}(g)$ of non-zero elements g in G. Our new bounds are expressed in terms of the *length* $l(G)$ of G by which we mean the length of the longest chain in $\mathbf{T}(G)$. If no longest chain exists we put $l(G)=\infty$, and observe the usual conventions about the ordering on $\mathbf{Z}\cup\{\infty\}$. We observe that if G has finite rank r then $l\leqq r$ (cf. FUCHS [3], page 112, Ex. 10).

We require the following notions. If α, β are types we say that α *absorbs* β if $\alpha\beta=\alpha$. If in particular α is self-absorbing then we say that α is *idempotent;* (many authors have used the term non-nil for this last notion, we prefer the word idempotent, for the existence of idempotent types in the type set of the additive group of a ring is closely related to the existence of idempotent elements in the ring itself).

Throughout the remainder of this note G denotes a torsion-free group of rank r and length l (both of which may be ∞). With these conventions we have:

Proposition 1.1. *If* $\mathbf{T}(G)$ *contains no absorbing elements, then*

$$\text{i) } N(G) \leqq 2^{l-1}, \quad \text{ii) } \nu(G) \leqq l, \quad \text{iii) } \varrho(G) \leqq l.$$

Received April 12, 1977.

Proposition 1.2. *If* $\mathbf{T}(G)$ *contains no idempotent elements, then*

$$\nu(G) \leqq \min\{2^l-1, r\}, \quad \varrho(G) \leqq l.$$

Proposition 1.3. *If* $\mathbf{T}(G)$ *contains no absorbing elements and* G *has length* 2, *then every ring on* G *is associative, and nilpotent of degree at most two.*

Basic ideas.

Lemma 2.1 *Suppose that* $(G, *)$ *is a non-associative ring on* G, *and that* $g_i \in G$ $(i=1, 2)$ *are such that* $g_1 * g_2 \neq 0$.

(i) *If neither* $\mathbf{t}(g_i)$ *absorbs the other then*

$$\mathbf{t}(g_1 * g_2) > \mathbf{t}(g_i) \quad (i = 1, 2).$$

(ii) *If neither* $\mathbf{t}(g_i)$ *is idempotent then either*

$$\mathbf{t}(g_1 * g_2) > \mathbf{t}(g_1) \quad \text{or} \quad \mathbf{t}(g_1 * g_2) > \mathbf{t}(g_2).$$

Proof. Clearly $\mathbf{t}(g_1 * g_2) \geqq \mathbf{t}(g_1)\mathbf{t}(g_2) \geqq \mathbf{t}(g_i)$ $(i=1, 2)$. If $\mathbf{t}(g_1 * g_2) = \mathbf{t}(g_1)$, then

$$\text{(A)} \qquad \mathbf{t}(g_1) = \mathbf{t}(g_1)\mathbf{t}(g_2) \geqq \mathbf{t}(g_2)$$

and $\mathbf{t}(g_1)$ absorbs $\mathbf{t}(g_2)$. This proves (i).

If $\mathbf{t}(g_1)$ is not idempotent then (A) implies that $\mathbf{t}(g_1) > \mathbf{t}(g_2)$ and (ii) follows.

For each positive integer k, let $V_k = \{x \in G \mid$ there is a chain $\mathbf{t}(x) > \tau_2 > \ldots > \tau_k$ of types in $\mathbf{T}(G)\}$ and let G_k be the subgroup of G generated by V_k. We clearly have a descending chain $G = G_1 \supset G_2 \supset G_3 \supset \ldots$ of subgroups of G.

Corollary 2.2. *Under the same hypothesis on* G *as in Lemma* 2.1 *we have:*

(i) *If* $\mathbf{T}(G)$ *has no absorbing elements then*

$$G * G_i \subset G_{i+1} \quad \text{and} \quad G_i * G \subset G_{i+1} \quad \text{for all positive integers } i.$$

(ii) *If* $\mathbf{T}(G)$ *has no idempotents then*

$$G_i * G_i \subset G_{i+1} \quad \text{for all positive integers } i.$$

Remark. In both cases $(G_i, *)$ is a subring of $(G, *)$ and in case (i) $(G_i, *)$ is an ideal in $(G, *)$.

Let R be a non-associative ring. For each positive integer k we may define four 'powers' of R as follows.

(i) $R^{(k)}$ is the subring of R generated by all products of k elements in R, however the products are associated.

(ii) $R^{[1]} = R, \quad R^{[k]} = R^{[k-1]}R^{[k-1]}$ for all $k > 1$.

(iii) $\vec{R}^1 = R, \quad \vec{R}^k = \vec{R}^{k-1}R$ for all $k > 1$.

(iv) $\overleftarrow{R}^1 = R, \quad \overleftarrow{R}^k = R\overleftarrow{R}^{k-1}$ for all $k > 1$.

We observe that each of these 'k-th powers' is contained in $R^{(k)}$. A simple induction shows that

(2.3) $$R^{[k]} \subset R^{(2^{k-1})} \quad \text{for all integers } k > 1.$$

Recall that R is *nilpotent* if there is an index k such that $R^{(k+1)} = 0$ and R is *solvable* if $R^{[k+1]} = 0$. If G is a group we say that the *solvable degree*, $\varrho(G)$, of G is k if $[G, *]^{[k+1]} = 0$ for all multiplications $*$ on G and there is a multiplication $\circ$ with $[G, \circ]^{[k]} \neq 0$.

The following inclusions are an easy consequence of Corollary 2.2.

Proposition 2.4.

(i) *If* $\mathbf{T}(G)$ *has no absorbing types then*

$$(\overleftarrow{G, *})^n \subseteqq G_n \quad \textit{and} \quad (\overrightarrow{G, *})^n \subseteqq G_n \quad \textit{for all positive integers } n.$$

(ii) *If* $\mathbf{T}(G)$ *has no idempotents then*

$$(G, *)^{[n]} \subseteqq G_n \quad \textit{for all positive integers } n.$$

In order to obtain information about $(G, *)^{(k)}$ we need a further notion. Denote by $F(R)$ the subring of the (associative) ring $E(R)$ of endomorphisms of the additive group of R, generated by the left and right multiplications L_a, R_a, $a \in R$ where

$$xL_a = ax; \quad xR_a = xa \quad \text{for all} \quad x \in R.$$

Lemma 2.5. *Let R be a torsion-free ring. Let n and k be positive integers satisfying $k > 2^{n-1}$. Then*

$$R^{(k)} \subseteqq R[F(R)]^n.$$

Proof. We proceed by induction on n the result being clear when $n=1$. Suppose that the result holds for $n = m \geqq 1$ and that $k > 2^m$. Let x be the product of k elements in R. Then $x = uv$ where at least one of u or v, u say, is the product of at least 2^{m-1} elements of R. Thus by hypothesis u belongs to $R[F(R)]^m$ and $uv \in R[(F(R)]^{m+1}$.

Corollary 2.6. *Let G be a torsion-free group whose type set contains no absorbing elements. Let $(G, *)$ be a (non-associative) ring on G. Let n and k be positive integers satisfying $k > 2^{n-1}$. Then*

$$(G, *)^{(k)} \subset G_{n+1}.$$

Proof. In fact we show that $G[F(G, *)]^{(n)} \subset G_n$, for all positive integers n. We may assume without loss of generality that $G[F(G, *)]^{(n)}$ is non null. In particular there exists non-zero monomials in $G[F(G, *)]^{(n)}$. Recalling that $F(G, *)$ is associative we see that such a monomial may be written in the form

$$\xi = (\dots((gX_1)X_2)\dots X_n) \neq 0$$

where $g \in G$ and, for each i, X_i denotes $*$ multiplication on the left or right by an element of G. It follows from Lemma 2.1 (i) that

$$\mathbf{t}(g) < \mathbf{t}(gX_1) < \mathbf{t}((gX_1)X_2) < \dots < \mathbf{t}(\xi)$$

is a strictly ascending chain in $\mathbf{T}(G)$ of length $n+1$ and so $\xi \subset V_{n+1}$. However, the monomials generate $G[F(G, *)]^{(n)}$ and the corollary follows.

Proof of Propositions 1.1, 1.2 and 1.3. Suppose that G has finite length l. Then, by definition, $G_{l+1}=0$. If $\mathbf{T}(G)$ has no absorbing types Proposition 2.4 gives

(a) $$(\overrightarrow{G, *})^{l+1} = (\overleftarrow{G, *})^{l+1} = 0$$

whilst Corollary 2.6 yields

(b) $$(G, *)^{(2^{l-1}+1)} = 0.$$

Since $*$ is an arbitrary multiplication on G we conclude from (a) that $\nu(G) \leqq l$ and from (b) that $N(G) \leqq 2^{l-1}$. Furthermore putting $k=l+1$ in equation 2.3 gives

$$(G, *)^{[l+1]} \subset (G, *)^{(2^l)} \subset (G, *)^{(2^{l-1}+1)} = 0$$

and we deduce that $\varrho(G) \leqq l$. This proves Proposition 1.1 (if l is infinite the result is trivial!). Substitution of $l=2$ in (b) gives $(G, *)^{(3)}=0$, and we deduce that in this case $(G, *)$ is always associative thus proving Proposition 1.3.

If all we know about $\mathbf{T}(G)$ is that it contains no idempotents then Proposition 2.4 (ii) gives $(G, *)^{[l+1]}=0$ and we have $\varrho(G) \leqq l$. If $*$ is an associative multiplication then

$$0 = (G, *)^{[l+1]} = (G, *)^{2^l}$$

whence $\nu(G) \leqq 2^l - 1$. Webb [8] has shown that $\nu(G) \leqq r$ and we have proved Proposition 1.2.

Finally we construct a group G which has

(i) finite type set, (ii) no idempotent types, (iii) $N(G)=\infty$.

Let $R_1 \subset R_2$ be subgroups of the rationals containing 1. Suppose that neither R_1 nor R_2 is a subring of Q, but that $R_1R_2=R_2$, the multiplication being the usual one on Q. Put $G=Rx \oplus Ry$, then G satisfies conditions (i) and (ii) above. We put two multiplications $*$ and $\circ$ on G as follows

$$\begin{array}{c|cc} * & x & y \\ \hline x & 0 & y \\ y & y & 0 \end{array} \qquad \begin{array}{c|cc} \circ & x & y \\ \hline x & y & 0 \\ y & 0 & 0 \end{array}$$

If n is a positive integer then

$$x*(\ldots*(x*(x*y))\ldots) = y \neq 0,$$

x appearing n times. It follows that $N(G)=\infty$, and reference to Proposition 1.2 shows that $(G, *)$ is nonassociative.

It is easily checked that

$$0 \neq (G, *)^{[2]} \subset R_2 y, \quad (G, *)^{[3]} = 0$$

so that $\varrho(G)=2=l(G)$. Lastly we see that $(G, \circ)$ is associative, indeed $(G, \circ)^{(3)}=0$, so $\nu(G)=2$.

References

[1] S. Feigelstock, Nilstufe of direct sums of rank 1 torsion-free groups, *Acta Math. Acad. Sci. Hung.*, **24** (1973), 269—272.

[2] S. Feigelstock, The nilstufe of homogeneous groups, *Acta Sci. Math.*, **36** (1974), 27—28.

[3] L. Fuchs, *Infinite Abelian Groups*, Vol. II, Academic Press (1973).

[4] B. J. Gardner, Rings on completely decomposable torsion-free abelian groups, *Commentationes Math. Univ. Carolinae*, **15** (1974), 381—392.

[5] B. J. Gardner, Some aspects of T-nilpotence, *Pacific J. Math.*, **53** (1974), 117—130.

[6] R. Ree and R. J. Wisner, A note on torsion-free nil groups, *Proc. A.M.S.*, **7** (1956), 6—8.

[7] T. Szele, Gruppentheoretische Beziehungen bei gewissen Ringkonstruktionen, *Math. Z.*, **54** (1951), 168—180.

[8] M. C. Webb, A bound for the nilstufe of a group, *Acta Sci. Math.*, **39** (1977), 185—188.

[9] W. J. Wickless, Abelian groups which admit only nilpotent multiplications, *Pacific J. Math.*, **40** (1972), 251—259.

[10] C. Vinsolhaler and W. J. Wickless, Completely decomposable groups which admit only nilpotent multiplications, *Pacific J. Math.*, **53** (1974), 273—280.

(STRATTON)
UNIVERSITY OF EXETER
DEPARTMENT OF MATHEMATICS
NORTH PARK ROAD
EXETER, EX4 4QE
ENGLAND

(WEBB)
URIVERSITY OF GHANA
P.O. BOX 62
LEGON, GHANA

Acta Sci. Math., 41 (1979), 215—219

A characterization of .3

GERNOT STROTH

The objective of this paper is the proof of the following theorem.

Theorem. *Let G be a finite simple group and H a 2-local subgroup of G. Assume that $H/\mathbf{O}(H)$ is an extension of $Z_4 * Q_8 * D_8$ by Σ_6. Assume further that $\mathbf{Z}(H/\mathbf{O}(H))$ is of order two. Then G is isomorphic to .3, the Conway simple group.*

Lemma 1. *Put $H_1 = H/\mathbf{O}(H)$. Then $H_1/\mathbf{Z}(H_1)$ splits over $\mathbf{O}_2(H_1/\mathbf{Z}(H_1))$.*

Proof. Put $H_2 = H_1/\mathbf{Z}(\mathbf{O}_2(H_1))$. Then $\mathbf{O}_2(H_2)$ is a symplectic space of dimension four. Thus $H_2/\mathbf{O}_2(H_2)$ is isomorphic to a subgroup of $Z_2 \times \Sigma_6$. In this group there are exactly two subgroups isomorphic to Σ_6. Since $\mathbf{Z}(H_1)$ is of order two we get that H_2 is uniquely determined. Thus we get in H_2 a subgroup isomorphic to Σ_6. Since $\mathbf{Z}(H_1)$ is of order two we get in $H_1/\mathbf{Z}(H_1)$ a subgroup isomorphic to Σ_6. This proves the lemma.

Lemma 2. *Let z be the involution in $\mathbf{Z}(H)$. Then $H \neq \mathbf{C}_G(z)$.*

Proof. By way of contradiction we assume $H = \mathbf{C}_G(z)$.

Assume first that z is conjugate to an involution x contained in $\mathbf{O}_2(H) - \langle z \rangle$. Then there is an element ϱ centralizing x such that $\varrho^3 \in \mathbf{O}(H)$. Thus $\langle \varrho, \mathbf{O}(H) \rangle$ is contained in $\mathbf{C}_G(x)$. Let π be an element of $\mathbf{O}(H)$. Then $\mathbf{C}_G(\pi)$ contains $\mathbf{O}_2(H)$. Let v be an element in $\varrho\mathbf{O}(H)$. Then a Sylow 2-subgroup of $\mathbf{C}_H(v)$ is isomorphic to $(Z_4 * Q_8)\langle a \rangle$ where $a^2 \in (Z_4 * Q_8)$. Thus 64 does not divide the order of $\mathbf{C}_G(v)$. Let ω be an element in $H - \mathbf{O}(H)$ such that $\omega^3 \in \mathbf{O}(H)$ and $\omega\mathbf{O}(H)$ is not conjugate to $\varrho\mathbf{O}(H)$ in $H/\mathbf{O}(H)$. Let μ be an element of $\omega\mathbf{O}(H)$. Then $\mathbf{C}_H(\mu)$ possesses a Sylow 2-subgroup S such that S is of order at least 8 and $\Phi(S)$ is equal to $\langle z \rangle$. Thus 16 does not divide the order of $\mathbf{C}_G(\mu)$. Let g be an element of G such that $x^g = z$. Then ϱ^g is contained in H. Since 16 divides the order of $\mathbf{C}_G(\varrho)$ but 64 does not divide the order of $\mathbf{C}_G(\varrho)$ we may assume $\varrho^g \in \varrho\mathbf{O}(H)$. Thus we may assume that g is contained in $\mathbf{N}_G(\langle \varrho \rangle)$. Let T be a Sylow 2-subgroup of $\mathbf{C}_H(\varrho) \cap \mathbf{C}_G(x)$.

Received March 18, 1975.

Then it is easy to see that T' is equal to $\langle z\rangle$. Thus x is not conjugate to z in $\mathbf{N}_G(\langle\varrho\rangle)$. We have proved that $\langle z\rangle$ is strongly closed in $\mathbf{O}_2(H)$ with respect to G.

Assume now that z is conjugate to an involution y in $H'-\langle z\rangle$. Then $\mathbf{C}_{\mathbf{O}_2(H)}(y)$ is isomorphic to $Z_4\times Z_2$. Thus there is an involution s in $\mathbf{O}_2(H)-\langle z\rangle$ such that y is conjugate to sy in G. Let U be a Sylow 2-subgroup of H. Then every involution a of $U-\langle z\rangle$ is conjugate to za in U. Thus s is conjugate to sy in G. But then s is conjugate to z in G, which is a contradiction. Thus we have proved that $\langle z\rangle$ is strongly closed in H' with respect to G.

Assume now that z is conjugate to an involution u of $H-H'$. Then z is a non-square in $\mathbf{C}_H(u)$. Thus $\mathbf{C}_{\mathbf{O}_2(H)}(u)$ is elementary abelian of order eight. But then there is an involution b in $\mathbf{O}_2(H)-\langle z\rangle$ such that u is conjugate to bu in G. As above we get a contradiction.

Thus we have proved that $\langle z\rangle$ is strongly closed in a Sylow 2-subgroup of G. Hence [2; Corollary 1, p. 404] yields the assertion.

Lemma 3. *Let M be a finite simple group which possesses a 2-local subgroup L such that $L/\mathbf{O}(L)$ is isomorphic to a faithful extension of E_{16} by A_6. Then M is isomorphic to $L_4(q)$, $q\equiv 5(8)$; $U_4(q)$, $q\equiv 3(8)$; M_{22}, M_{23} or M^c.*

Proof. By [6; Theorem 3], L contains a Sylow 2-subgroup of M. Now [4] yields the assertion.

Lemma 4. *Let M be a finite group which possesses an involution z such that $\mathbf{C}_M(z)/\mathbf{O}(\mathbf{C}_M(z))$ is isomorphic to one of the following groups:*

(i) $\mathrm{SL}_4(q)$, $q\equiv 5(8)$;
(ii) $\mathrm{SU}_4(q)$, $q\equiv 3(8)$.

Then $z\in\mathbf{Z}^*(M)$.

Proof. In $\mathbf{C}_M(z)$ there are only two classes of involutions. Let v be an involution of $\mathbf{C}_M(z)$ not equal to z.

Put $C=\mathbf{C}_M(z)$. Then $\mathbf{C}_C(v)$ contains a subgroup $E=S_1\times S_2$ where S_1 and S_2 are isomorphic to $SL_2(q)$. Now we get $\mathbf{Z}(S_1)=\langle v\rangle$ and $\mathbf{Z}(S_2)=\langle zv\rangle$, implying that $\mathbf{C}_C(v)/\mathbf{O}(\mathbf{C}_C(v))$ is equal to $\mathbf{Z}(C/\mathbf{O}(C))*(E\langle a\rangle)$ where a induces the diagonal automorphism on S_1 and S_2. Let R be a Sylow 2-subgroup of $\mathbf{C}_C(v)$. Then R' is isomorphic to $Z_4\times Z_4$ and $\mathbf{C}_R(R')$ is isomorphic to $Z_2\times Z_4\times Z_8$. Since $\mho^2(\mathbf{C}_R(R'))$ is equal to $\langle z\rangle$ we get that z is not conjugate to v in G. Hence [2; Corollary 1] yields the assertion.

Lemma 5. *Let M be a finite group. Assume that z is an involution in M such that $\mathbf{C}_M(z)/\mathbf{O}(\mathbf{C}_M(z))$ is isomorphic to one of the following groups:*

(i) *$SL_4(q)\langle x\rangle$, $q\equiv 5(8)$, x induces the graph-automorphism on $SL_4(q)$ and $x^2\in\mathbf{Z}(SL_4(q))$;*

(ii) $SU_4(q)\langle x\rangle$, $q\equiv 3(8)$, *x induces the field-automorphism of order* 2 *on* $SU_4(q)$ *and* $x^2\in \mathbf{Z}(SU_4(q))$.
Then $z\in \mathbf{Z}^*(M)$.

Proof. Put $C=\mathbf{C}_M(z)$. Then $\mathbf{C}_C(x)/\mathbf{O}(\mathbf{C}_C(x))$ is isomorphic to $Sp_4(q)\langle x\rangle$. Let T be a Sylow 2-subgroup of $\mathbf{C}_C(x)$. Then $\langle z\rangle=\mathbf{Z}(T)\cap T'$. Thus $\mathbf{C}_C(x)$ contains a Sylow 2-subgroup of $\mathbf{C}_G(x)$.

Assume that x is an element of order two. Then 2^9 does not divide the order of $\mathbf{C}_G(x)$. Thus x is not conjugate to an involution of $SL_4(q)$ or $SU_4(q)$. Thus [12; Lemma (5.38)] yields that M possesses a subgroup M_1 of index two. Consequently, $\mathbf{C}_{M_1}(z)/\mathbf{O}(\mathbf{C}_{M_1}(z))$ is isomorphic to $SL_4(q)$, $q\equiv 5(8)$ or $SU_4(q)$, $q\equiv 3(8)$ whence by Lemma 4 the assertion follows.

Put $\langle u\rangle=\mathbf{Z}(SL_4(q))$, resp. $\mathbf{Z}(SU_4(q))$. Then we may assume that $\langle u, x\rangle$ is isomorphic to Q_8.

We shall prove that $\langle z\rangle$ is strongly closed in C' with respect to M. Let v be an involution of $C'-\langle z\rangle$. Then $\mathbf{C}_C(v)/\mathbf{O}(\mathbf{C}_C(v))$ contains a subgroup $E=S_1\times S_2$ where S_1 and S_2 are isomorphic to $SL_2(q)$. We may assume $\mathbf{Z}(S_1)=\langle v\rangle$ and $\mathbf{Z}(S_2)=\langle zv\rangle$. Now $\mathbf{C}_C(v)$ contains a subgroup Q isomorphic to Q_8 such that Q' is equal to $\langle z\rangle$. Then $\mathbf{C}_C(v)/\mathbf{O}(\mathbf{C}_C(v))$ is equal to an extension of order 2 of $Q*E$. Assume that z is conjugate to v in M. Then there is a Sylow 2-subgroup B of $Q*E$ such that z is conjugate to v in $\mathbf{N}_M(B)$. Now B is isomorphic to $Q_8*(Q_8\times Q_8)$. Thus $\mathbf{N}_M(\mathbf{Z}(B))/\mathbf{C}_M(\mathbf{Z}(B))$ is isomorphic to Σ_3. However, since $\mathbf{C}_B(\mathbf{O}_3(\mathbf{C}_M(\mathbf{Z}(B))/B))$ is isomorphic to Q_8, we get a contradiction. Thus $\langle z\rangle$ is strongly closed in C' with respect to M.

Now we know that $\mathbf{C}_C(x)$ contains an element s such that sx is an involution and sx is centralized by s. Thus z is a square in $\mathbf{C}_M(xs)$. This implies that xs is not conjugate to an element of C'. Hence by [12; Lemma (5.38)] M possesses a subgroup M_1 of index two. Thus $\mathbf{C}_{M_1}(z)/\mathbf{O}(\mathbf{C}_{M_1}(z))$ is isomorphic to $SL_4(q)$, $q\equiv 5(8)$ or $SU_4(q)$, $q\equiv 3(8)$, which by Lemma 4 yields the assertion.

Lemma 6. *Let M be a finite group and z a* 2*-central involution in M such hatt* $\mathbf{C}_M(z)/\mathbf{O}(\mathbf{C}_M(z))$ *is isomorphic to a split extension of an elementary abelian group E of order* 32 *by* A_6 *where* A_6 *acts undecomposable on E. Then* $z\in \mathbf{Z}^*(M)$.

Proof. Assume first that z is conjugate in M to an involution u of $\mathbf{C}_M(z)-(E\mathbf{O}(\mathbf{C}_M(z)))$. Put $C=\mathbf{C}_M(z)$. Then there are only two classes of involutions in $C-\mathbf{O}_{2',2}(C)$. Thus $\mathbf{C}_C(u)/\mathbf{O}(\mathbf{C}_C(u))$ is isomorphic to a split extension of E_8 by D_8. Hence $C/\mathbf{O}(C)$ involves a subgroup A_5 such that EA_5 is equal to $\langle z\rangle\times(E_{16}A_5)$ where A_5 acts intransitively on E_{16}. Thus we may assume that there is an involution r in $\mathbf{Z}(\mathbf{C}_C(u)/\mathbf{O}(\mathbf{C}_C(u)))$ such that u is conjugate to ru and r is contained in $(\mathbf{C}_C(u)/\mathbf{O}(\mathbf{C}_C(u)))'$. Let S be a Sylow 2-subgroup of $\mathbf{C}_M(u)$ containing a Sylow 2-subgroup of $\mathbf{C}_C(u)$. Assume that z is conjugate neither to r nor to zr. Then $\mathbf{Z}(S)$

is equal to $\langle r, u\rangle$. But this is a contradiction. Thus we have proved that $\langle z\rangle$ is not strongly closed in E with respect to M if z is conjugate to an involution of $C-\mathbf{O}_{2',2}(C)$.

Assume now that $\langle z\rangle$ is not strongly closed in E with respect to M. Let T be a Sylow 2-subgroup of C. Since all involutions of E are conjugate to involutions of $\mathbf{Z}(T)$ in C we get that all involutions of E are conjugate in M. If z is not conjugate to an involution of $C-\mathbf{O}_{2',2}(C)$ in M we get that E is strongly closed in T with respect to M. Then it follows from [3] that $E\mathbf{O}(M)$ is normal in M. Thus $|M/\mathbf{O}(M):C\mathbf{O}(M)/\mathbf{O}(M)|$ is equal to 31, which is impossible.

Thus we have proved there are only two possibilities for the fusion of involutions in M. The first is that $\langle z\rangle$ is strongly closed in T with respect to M. Then [2] yields the assertion. The second is that all involutions of M are conjugate in M. Thus all 2-local subgroups of $M/\mathbf{O}(M)$ are 2-constrained, so that applying [1] we get a contradiction. Thus the lemma is proved.

Lemma 7. *Put* $\langle u\rangle=\mathbf{Z}(\mathbf{O}_2(H))$. *Then* $\mathbf{N}_G(\langle u\rangle)/\mathbf{O}(\mathbf{N}_G(\langle u\rangle))$ *is isomorphic to one of the following groups:*

(i) $H/\mathbf{O}(H)$;

(ii) $SL_4(q)\langle x\rangle$, $q\equiv 5(8)$, $x^2\in\mathbf{Z}(SL_4(q))$ *and* x *induces the graph-automorphism on* $SL_4(q)$;

(iii) $SU_4(q)\langle x\rangle$, $q\equiv 3(8)$, $x^2\in\mathbf{Z}(SU_4(q))$ *and* x *induces the field-automorphism on* $SU_4(q)$.

Proof. Put $N=\mathbf{N}_G(\langle u\rangle)$. Assume that N is not equal to H. Let M be a minimal normal subgroup of $N/(\mathbf{O}(N)\langle u\rangle)$. Then M is simple. Further, M possesses a 2-local subgroup isomorphic to a split extension of E_{16} by A_6. Then, by Lemma 3, M is isomorphic to $L_4(q)$; $q\equiv 5(8)$, $U_4(q)$; $q\equiv 3(8)$, M_{22}, M_{23} or M^c. Applying [5] we get that M is isomorphic to $L_4(q)$; $q\equiv 5(8)$ or $U_4(q)$; $q\equiv 3(8)$. Thus $N/\mathbf{O}(N)$ contains a subgroup of index 2 isomorphic to $SL_4(q)$ or $SU_4(q)$. Now the structure of $\mathrm{Aut}(SL_4(q))$ and $\mathrm{Aut}(SU_4(q))$ yields the assertion.

Lemma 8. *The group* $\mathbf{C}_G(z)/\mathbf{O}(\mathbf{C}_G(z))$ *is isomorphic to* $\widehat{Sp_6(2)}$.

Proof. Put $C=\mathbf{C}_G(z)/(\mathbf{O}(\mathbf{C}_G(z))\langle z\rangle)$. Assume first that $N=\mathbf{N}_G(\langle u\rangle)$ is not equal to H. Let F be a minimal normal subgroup of C. Assume that F is not simple. Then F is contained in $N/(\mathbf{O}(\mathbf{C}_G(z))\langle z\rangle)$. Then $\mathbf{C}_G(z)$ is equal to N, which by Lemmas 7 and 4 leads to a contradiction. Thus F is simple. Let T be a Sylow 2-subgroup of N. Since $u\langle z\rangle$ is not a square in $T/\langle z\rangle$ but all other involutions in $\mathbf{Z}(T/\langle z\rangle)$ are squares in $T/\langle z\rangle$ we get that T is a Sylow 2-subgroup of G. Thus C (possesses a Sylow 2-subgroup of type A_{12}. Since all involutions of $(N/(\mathbf{O}(\mathbf{C}_G(z))\langle z\rangle)$ are conjugate to involutions of $\mathbf{Z}(T/\langle z\rangle)$ we get that F possesses a Sylow 2-subgroup of type A_{12}. Then by [9], F is isomorphic to A_{12}, A_{13}, $PSp_6(2)$ or has the involution-fusion-pattern of $\Omega_7(3)$.

Assume now that N is equal to H. Let F be a minimal normal subgroup of C. Lemma 2 implies that F is simple and Lemma 6 yields that $N/(\mathbf{O}(\mathbf{C}_G(z))\langle z\rangle)$ is contained in F since a Sylow 2-subgroup of N is a Sylow 2-subgroup of $\mathbf{C}_G(z)$. Hence, by [9], F is isomorphic to $A_{12}, A_{13}, PSp_6(2)$ or has the involution-fusion-pattern of $\Omega_7(3)$.

Thus in both cases we have proved that a minimal normal subgroup of C is isomorphic to $A_{12}, A_{13}, PSp_6(2)$ or has the involution-fusion-pattern of $\Omega_7(3)$.

Assume first that a minimal normal subgroup of C has the involution-fusion-pattern of $\Omega_7(3)$. Applying [10] and [7] we get that $\mathbf{C}_G(z)/\mathbf{O}(\mathbf{C}_G(z))$ is an odd extension of $\mathrm{Spin}_7(q)$, $q\equiv 3, 5(8)$. Now [11; Theorem (3.4)] yields a contradiction.

Assume now that a minimal normal subgroup of C is isomorphic to A_{12} or A_{13}. Then $\mathbf{C}_G(z)/\mathbf{O}(\mathbf{C}_G(z))$ is isomorphic to $\hat{A}_{12}$ or $\hat{A}_{13}$, so that G possesses only one class of involutions. Now [8; Corollary] yields a contradiction.

Thus we have proved that a minimal normal subgroup of C is isomorphic to $PSp_6(2)$. The structure of $\mathrm{Aut}\,(PSp_6(2))$ shows now that C is isomorphic to $PSp_6(2)$. Thus the lemma is proved.

Lemma 9. *The group G is isomorphic to .3, the Conway simple group.*

Proof. By Lemma 8, a Sylow 2-subgroup of G is of type .3, which by [11] implies the assertion.

References

[1] B. Beisiegel, Über endliche einfache Gruppen mit Sylow 2-Untergruppe der Ordnung höchstens 2^{10}, *Comm. Algebra*, **5** (1977), 113—170.

[2] G. Glauberman, Central elements in core-free groups, *J. Algebra*, **4** (1966), 403—421.

[3] D. M. Goldschmidt, 2-fusion in finite groups, *Ann. of Math.*, **99** (1974), 70—117.

[4] D. Gorenstein and K. Harada, Finite groups whose 2-subgroups are generated by at most 4 elements, *Mem. Amer. Math. Soc.*, **147** (1974).

[5] R. Griess, *Schur multipliers of the known finite simple groups*, Thesis, University of Chicago (1971).

[6] K. Harada, On finite groups having self-centralizing 2-subgroups of small order, *J. Algebra*, **33** (1975), 144—160.

[7] J. B. Olsson, Odd-order extensions of some orthogonal groups, *J. Algebra*, **28** (1974), 573—596.

[8] R. Solomon, Finite groups with components of alternating type, *Notices Amer. Math. Soc.*, **21** (1974), A—103.

[9] R. Solomon, Finite groups with Sylow 2-subgroups of type A_{12}, *J. Algebra*, **24** (1973), 346—378.

[10] R. Solomon, Finite groups with Sylow 2-subgroups of type $\Omega(7, q)$, $q\equiv \pm 3 \pmod 8$, *J. Algebra*. **28** (1974), 174—181.

[11] R. Solomon, Finite groups with Sylow 2-subgroups of type .3, *J. Algebra*, **28** (1974), 182—198.

[12] J. G. Thompson, Nonsolvable groups all of whose local subgroups are solvable, *Bull. Amer. Math. Soc.*, **74** (1968), 383—437.

UNIVERSITÄT HEIDELBERG
MATHEMATISCHES INSTITUT
6900 HEIDELBERG, BRD

Partial translations of semigroups

R. P. SULLIVAN

To Jeanette Ryan with deep affection

1. Introduction. If S is semigroup, an element ϱ of $\mathscr{P}_S$ will be called a *partial right translation* of S if whenever $a\varrho$, $a\in S$, is defined then $(xa)\varrho$ is defined for all $x\in S$ and $x(a\varrho)=(xa)\varrho$. This definition was introduced in [6] to extend that of "one-to-one partial right translation" given by CLIFFORD and PRESTON ([1], vol. 1, page 32). Clearly both the domain of ϱ and the range of ϱ are left ideals of S and so the notion of partial right translation coincides with that of "left S-translation" discussed by STEINFELD [4]. However we prefer the former terminology since the concept of partial right translation generalises in a natural way that of "right translation".

Our first aim in this paper will be to describe a class of semigroups embeddable in inverse semigroups and thus provide an alternative account of some work done by SCHEIN [2]. After this we investigate the problem of determining conditions under which a semigroup S can be embedded in a semigroup T such that every partial right translation of S is induced by a right translation of T.

This paper was completed while visiting the Mathematics Institute of the Hungarian Academy of Sciences. I would like to thank Professor O. Steinfeld for his kind assistance and many interesting conversations during my stay in Budapest. My gratitude is also extended to the referee for recommending a number of changes in an earlier version of this work.

2. Replete semigroups. All terminology will be that of [1]. Thus $\Lambda(S)$ $[P(S)]$ denotes the semigroup of all left [right] translations of S and $\Lambda_0(S)$ $[P_0(S)]$ the semigroup of all inner left [right] translations of S. We let $\mathscr{L}(S)$ $[\mathscr{R}(S)]$ denote the set of all partial left [right] translations of S.

Suppose $\varrho_1, \varrho_2\in\mathscr{R}(S)$ and $\varrho_1\varrho_2\neq\square$ (that is, ran $\varrho_1\cap$dom $\varrho_2\neq\square$) and put

Received December 22, 1977, in revised form April 17, 1978.

$L=(\text{ran } \varrho_1 \cap \text{dom } \varrho_2)\varrho_1^{-1}$. Then $\text{dom } \varrho_1\varrho_2 = L$ and if $a \in L$ and $x \in S$, then $a \in \text{dom } \varrho_1$, $(xa)\varrho_1 \in \text{dom } \varrho_2$, and we have

$$(xa)\varrho_1\varrho_2 = ((xa)\varrho_1)\varrho_2 = (x(a\varrho_1))\varrho_2 = x(a\varrho_1)\varrho_2 = x(a\varrho_1\varrho_2).$$

Since therefore the only way $\mathscr{L}(S)$ $[\mathscr{R}(S)]$ can fail to be a semigroup under composition is that some product of two elements equals $\square$, we shall on occasion regard $\square$ as a partial left [right] translation of S: this accords with the standard terminology for $\mathscr{I}_X$. We note in passing that if $S=S^0$, then $0 \in L$ for every left ideal of S and so $\varrho_1\varrho_2 \neq \square$ for all $\varrho_1, \varrho_2 \in \mathscr{R}(S)$; thus, when S contains a zero, $\mathscr{L}(S)$ $[\mathscr{R}(S)]$ is a semigroup properly containing $\Lambda(S)$ $[P(S)]$ (since the identity on $\{0\}$ is a partial left (right] translation of S).

If L is a left ideal of S and $a \in S$, we call $\varrho \in \mathscr{P}_S$ such that $\text{dom } \varrho = L$ and $x\varrho = xa$ for all $x \in L$ the *partial inner right translation* (briefly, "pirt") of L *induced by* a and let $\mathscr{R}_0(S)$ denote the set of all such pirts. With the above convention $\mathscr{R}_0(S)$ is a subsemigroup of $\mathscr{R}(S)$ containing $P_0(S)$. The notions of partial inner left translation ("pilt" for short) and of $\mathscr{L}_0(S)$ are defined dually.

As usual (see [1], vol. 2), a non-empty subset A of a semigroup S is called *left unitary* if $xy \in A$ implies $y \in A$. A non-trivial example of a semigroup in which every left ideal is left unitary is provided by Exercise 6, § 1.11 [1]. The following result is analogous to those of Exercises 4 and 6 of § 1.3 [1]; the proof is straightforward and so is omitted.

Proposition 1. *If S is a semigroup then*
(i) *if each left ideal of S contains a right identity then $\mathscr{R}(S)=\mathscr{R}_0(S)$;*
(ii) *if $S=S^2$ then every $\varrho \in \mathscr{R}(S)$ with left unitary domain commutes with every $\lambda \in \mathscr{L}(S)$ with right unitary domain.*

STEINFELD [3] has used the notion of a one-to-one partial right translation to characterise completely 0-simple semigroups as those semigroups $S=S^0$ having the form $S=\bigcup\{Se_i : i \in I\}$ where $e_i^2=e_i$ and Se_i are 0-minimal left ideals such that for each $i, j \in I$, there exists a one-to-one partial right translation from Se_i onto Se_j. In [4] he omits the assumption of 0-minimality to consider a wider class of semigroups, the so-called similarly decomposable semigroups, and extends the Rees theory for completely 0-simple semigroups to certain matrix semigroups over semigroups with a zero and identity (for a survey of this area see [5]). In doing this he proves the following interesting results; we shall call two left ideals L_1, L_2 *right equivalent* if there is a one-to-one partial right translation from L_1 onto L_2.

Proposition 2. *If $S=S^0$ and e_1, e_2 are non-zero idempotents in S then the following are equivalent:*

(i) *Se_1 and Se_2 are right equivalent,*
(ii) *there exist $u, v \in S$ such that $uv = e_1$ and $vu = e_2$,*
(iii) *$e_1 S$ and $e_2 S$ are left equivalent.*

Proposition 3. *If $S = S^0$ and e_1, e_2 are non-zero idempotents in S and Se_1, Se_2 are right equivalent, then the semigroups $e_1 Se_1$ and $e_2 Se_2$ are isomorphic.*

Proposition 4. *If $S = S^0$ and Se_1, Se_2 are 0-minimal left ideals of S then either Se_1 and Se_2 are right equivalent or the only partial inner right translation of Se_1 induced by any $a \in Se_2$ is the zero translation of Se_1.*

In this section we use the notion of a one-to-one partial right translation to provide an alternative account of Schein's characterisation of those semigroups with identity that are embeddable in inverse semigroups ([3]; see also [1], § 11.4).

Let S be a semigroup satisfying:

(1) for each $a \in S$, there exists an idempotent $e \in S$ such that $ea = a$ and if $xa = ya$ then $xe = ye$.

If $S = S^0$, write $e_0 = 0$ and more generally for each non-zero $a \in S$, choose and fix an idempotent, denoted by e_a, satisfying (1) and put $H(S) = \{e_a : a \in S\}$. We call S a *replete* semigroup if S satisfies (1) and $H(S)$ can be chosen to satisfy the further properties: (2) $e_a e_b = e_b a_a$ and (3) $ae_b = e_{ab} a$.

Before proceeding we note that if S is replete then for $a, b \in S$, $e_a ab = e_{ab} ab$ implies that $e_a e_{ab} = e_{ab}$, and this latter identity will be used without further mention.

A result similar to the following appeared in [6]; its proof is motivated by that of Theorem 1.20 [1].

Theorem 1. *Every inverse semigroup is replete and every replete semigroup is embeddable in an inverse semigroup.*

Proof. If S is inverse and $a \in S$, put $e_a = aa^{-1}$ and $H(S) = \{e_a : a \in S\}$. It is easy to check (using Theorem 1.17 and Lemma 1.18 of [1]) that S is replete with respect to $H(S)$.

Suppose S is replete and for each $a \in S$, put $a\varphi = \varrho_a | Se_a$; note that if $S = S^0$ then $0\varphi = \varrho_0 | \{0\}$. We assert that φ embeds S into $\mathscr{I}_S$. For, if $x, y \in Se_a$ and $x\varrho_a = y\varrho_a$, then using (1), $x = xe_a = ye_a = y$. Hence $a\varphi \in \mathscr{I}_S$ for all $a \in S$. If $a\varphi = b\varphi$ then $Se_a = Se_b$ and so using (2) we have $e_a = e_b$. Hence

$$a = e_a a = e_a \varrho_a = e_a \varrho_b = e_a b = e_b b = b$$

and so φ is one-to-one. Now suppose $ab \neq 0$ and note that $x \in \operatorname{dom}(a\varphi b\varphi)$ if and only if $xe_a = x$ and $xae_b = xa$. Moreover these last two equations are together equivalent to $xe_{ab} = x$ which means $x \in \operatorname{dom}(ab)\varphi$. If $ab = 0$ then $ae_b b = 0 = 0b$

implies by (1) that $ae_b=0e_b=0$. Suppose $xe_a=x$ and $xae_b=xa$ for some non-zero $x\in S$. Then $xa=0=0a$ implies $x=xe_a=0e_a=0$, a contradiction. Hence $\operatorname{dom}(a\varphi b\varphi)=0$. In both cases therefore $\operatorname{dom}(a\varphi b\varphi)=\operatorname{dom}(ab)\varphi$ and so φ is a morphism.

3. Partial translations induced by total translations. It is readily observed that if $S=S^0$ and $\varrho\in\mathscr{R}(S)$ then $0\varrho=0$. In particular, if $\operatorname{dom}\varrho=\{0\}$ then ϱ can be regarded as the restriction of any $\sigma\in P(S)$ to $\{0\}$. It is therefore natural to ask whether every $\varrho\in\mathscr{R}(S)$ is induced by some $\bar\varrho\in P(S)$ in the sense that $\varrho=\bar\varrho\,|\operatorname{dom}\varrho$. This is certainly true for example when every left ideal of S equals Se for some idempotent $e\in S$. For then if $\varrho\in\mathscr{R}(S)$ has domain Se, we have for all $s\in S$

$$(se)\varrho = s(e\varrho) = s(e^2\varrho) = se(e\varrho) = (se)\varrho_a$$

where $a=e\varrho\in S$.

Example. Let S be the semigroup of all natural numbers under ordinary multiplication and let $\varrho\in\mathscr{P}_S$ have domain $2S$ and satisfy $(2n)\varrho=n$ for all $n\in S$. Then $m(2n)\varrho=mn=(m2n)\varrho$ for all $m\in S$ and so $\varrho\in\mathscr{R}(S)$. However if there exists $\bar\varrho\in P(S)$ such that $\varrho=\bar\varrho|\operatorname{dom}\varrho$, then $\bar\varrho=\varrho_a$ for some $a\in S$ (since $1\in S$) and so $n=(2n)\varrho=2na$ for all $n\in S$, a contradiction. Hence ϱ is not induced by any right translation of S.

In the light of this example, we now ask: can a semigroup S always be embedded in a semigroup T such that every $\varrho\in\mathscr{R}(S)$ is induced by some $\bar\varrho\in P(T)$? The next result introduces a sufficient condition under which an embedding can be achieved quite simply.

Theorem 2. *If $S=S^0$ is a semigroup in which every non-zero left ideal is left unitary, then every $\varrho\in\mathscr{R}(S)$ is induced by some $\bar\varrho\in P(S)$.*

Proof. For each $\varrho\in\mathscr{R}(S)$, we define $\bar\varrho\in\mathscr{I}_S$ by $x\bar\varrho=\begin{cases} x\varrho & \text{if } x\in\operatorname{dom}\varrho\\ 0 & \text{if } x\notin\operatorname{dom}\varrho.\end{cases}$
Then if $y\in\operatorname{dom}\varrho$ and $x\in S$, we have $x(y\bar\varrho)=x(y\varrho)=(xy)\varrho=(xy)\bar\varrho$ and if $y\notin\operatorname{dom}\varrho$, then $xy\notin\operatorname{dom}\varrho$, and we have $x(y\bar\varrho)=x0=0=(xy)\bar\varrho$. Finally, if $y\in S$ then $0(y\bar\varrho)=0=(0y)\bar\varrho$, and so we have shown that $\bar\varrho\in P(S)$ and by definition it induces $\varrho\in\mathscr{R}(S)$.

Clearly, if S is any semigroup in which every left ideal is left unitary (or equivalently, S is a disjoint union of minimal left ideals) then Theorem 2 can be applied to S^0, the semigroup S with a zero adjoined.

Finally we note that TAMURA and GRAHAM [7, Theorem 3] have determined the necessary and sufficient conditions on a semigroup S which ensure that any $\sigma\in P(S)$ is induced by an inner right translation of some semigroup T containing S as an ideal: we do not know if their result extends to the partial case.

References

[1] A. H. CLIFFORD and G. B. PRESTON, *The Algebraic Theory of Semigroups.* **1** (1961) and **2** (1967), Amer. Math. Soc. (Providence, Rhode Island).

[2] B. M. SCHEIN, Embedding semigroups in generalised groups, *Mat. Sb.* (N. S.), **55** (97) (1961), 379—400.

[3] O. STEINFELD, On semigroups which are unions of completely 0-simple subsemigroups, *Czechoslovak Math. J.*, **16** (191) (1966), 63—69.

[4] O. STEINFELD, On a generalization of completely 0-simple semigroups, *Acta Sci. Math.*, **28** (1967), 135—145.

[5] O. STEINFELD, On similarly decomposable semigroups, *to appear.*

[6] R. P. SULLIVAN, *A study in the theory of transformation semigroups,* Ph. D. Thesis, Monash University, 1969.

[7] T. TAMURA and N. GRAHAM, Certain embedding problems of semigroups, *Proc. Japan Acad.*, **40** (1964), 8—13.

MATHEMATICS DEPARTMENT
UNIVERSITY OF W.A.
NEDLANDS, 6009
WESTERN AUSTRALIA

Acta Sci. Math., 41 (1979), 227—251

On structural properties of functions arising from strong approximation of Fourier series

V. TOTIK

Introduction

Let $f(x)$ be an integrable and 2π-periodic function, and let

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{1}$$

be its Fourier series. Denote by $s_n(x) = s_n(f; x)$ and $\omega(f; \delta)$ the n-th partial sum of (1) and the modulus of continuity of f, respectively; $\|\cdot\|$ always stays for the supremum norm.

FREUD [1] proved that

$$\left\| \sum_{k=1}^{\infty} |s_k - f|^p \right\| < \infty \quad \text{for some} \quad p > 1 \quad \text{implies} \quad f \in \mathrm{Lip}\,\frac{1}{p}.$$

An analogous problem with $p=1$ was investigated by LEINDLER and NIKIŠIN [6], and this result was generalized by LEINDLER [4] as follows: If r is a nonnegative integer and

$$\left\| \sum_{k=1}^{\infty} k^r |s_k - f| \right\| < \infty,$$

then

$$|f^{(r)}(x+h) - f^{(r)}(x)| \leqq K \cdot h \cdot \log \frac{1}{h} \quad (x \in [0, 2\pi])$$

for all x, and this estimation is best possible.

From this result it follows that

$$\left\| \sum_{k=1}^{\infty} |s_k - f| \right\| < \infty \quad \text{does not imply} \quad f \in \mathrm{Lip}\, 1.$$

Received January 25, 1978.

LEINDLER raised the question whether the condition

$$\left\|\sum_{k=1}^{\infty}|s_k-f|^p\right\|<\infty \quad \textit{with some } p \quad (0<p<1) \quad \textit{implies} \quad f\in \mathrm{Lip}\,1.$$

The answer was given in the affirmative by OSKOLKOV [7] and SZABADOS [8]. They also proved

Theorem A. *For an arbitrary modulus of continuity Ω,*

$$\left\|\sum_{k=1}^{\infty}\Omega(|s_k-f|)\right\|<\infty \tag{2}$$

and

$$\int_0^1 \frac{dx}{\Omega(x)}<\infty \tag{3}$$

imply $f\in \mathrm{Lip}\,1$.

Under a certain restriction on Ω they also proved the necessity of condition (3). In [10] we proved the necessity of (3) without any further assumption, and more generally, we proved Theorem B (below).

In order to simplify our assertions, $\Omega(x)$ will always denote an increasing convex or concave function on $[0, \infty)$, with the properties

$$\Omega(x)>0 (x>0) \quad \lim_{x\to 0+0}\Omega(x)=\Omega(0)=0, \tag{4}$$

and we suppose that the inverse of $\Omega(x)$ (denoted by $\bar{\Omega}(x)$) exists in the interval $[0; 1]$. With these notations we proved

Theorem B. *If f satisfies* (2), *then*

$$\omega(f;\delta)=O\left(\delta\int_\delta^1\frac{\bar{\Omega}(x)}{x^2}dx\right), \tag{5}$$

but no estimate better than this can be given. Moreover, if Ω is concave, then we can replace

$$\int_\delta^1\frac{\bar{\Omega}(x)}{x^2}dx \quad by \quad \int_{\Omega(\delta)}^1\frac{dx}{\Omega(x)}.$$

The following theorem answers the analogous problem for the conjugate function.

Theorem 1. (i) *If Ω is concave, then* (2) *implies* $\tilde{f}\in \mathrm{Lip}\,1$. (ii) *Let Ω be convex. From* (2) *the continuity of $\tilde{f}$ follows if and only if*

$$\int_0^1\frac{\bar{\Omega}(x)}{x}dx<\infty. \tag{6}$$

If (6) *is fulfilled, then* (2) *implies that*

$$\omega(\tilde{f};\,\delta) = O\left(\int_0^\delta \frac{\bar{\Omega}(x)}{x}\,dx\right). \tag{7}$$

Furthermore, there exists a function f_0 *for which* (2) *is true, but*

$$\omega(\tilde{f}_0;\,\delta) \geqq c\int_0^\delta \frac{\bar{\Omega}(x)}{x}\,dx \quad (c>0). \tag{8}$$

We note that part (i) is a known result of LEINDLER [4].

Recently KROTOV and LEINDLER [2] investigated the problem to give a necessary and sufficient condition for a monotonic sequence $\{\lambda_k\}$ such that

$$\left\|\sum_{k=0}^{\infty}\lambda_k|s_k-f|^p\right\| < \infty \quad \text{with some } p \quad (0<p<\infty) \tag{9}$$

should imply $\omega(f;\delta)=O(\omega(\delta))$, where $\omega(\delta)$ is a fixed modulus of continuity. They proved

Theorem C. *Let* $\{\lambda_k\}$ *be a positive nondecreasing sequence,* $\omega(\delta)$ *be a modulus of continuity and* $0<p<\infty$. *Then* (9) *implies* $\omega(f;\delta)=O(\omega(\delta))$ *if and only if*

$$\sum_{k=1}^{n}(k\cdot\lambda_k)^{-\frac{1}{p}} = O\left(n\cdot\omega\left(\frac{1}{n}\right)\right). \tag{10}$$

As a common generalization of Theorem B and C we shall prove

Theorem 2. *Let* Ω *be a convex or concave function with properties* (4), *and let* $\{\lambda_k\}_0^\infty$, $\{\mu_k\}_0^\infty$ *be positive nondecreasing sequences. If*

$$\left\|\sum_{k=0}^{\infty}\lambda_k\Omega(\mu_k|s_k-f|)\right\| < \infty \tag{11}$$

then

$$\omega\left(f;\,\frac{1}{n}\right) = O\left(\frac{1}{n}\sum_{k=1}^{n}\frac{1}{\mu_k}\bar{\Omega}\left(\frac{1}{k\cdot\lambda_k}\right)\right). \tag{12}$$

Furthermore, there exists a function f_0 *satisfying* (11), *for which*

$$\omega\left(f_0;\,\frac{1}{n}\right) \geqq c\cdot\frac{1}{n}\sum_{k=1}^{n}\frac{1}{\mu_k}\bar{\Omega}\left(\frac{1}{k\cdot\lambda_k}\right) \quad (c>0). \tag{13}$$

Corollary 1. *Condition* (11) *implies* $f\in\operatorname{Lip}1$ *if and only if*

$$\sum_{k=1}^{\infty}\frac{1}{\mu_k}\bar{\Omega}\left(\frac{1}{k\cdot\lambda_k}\right) < \infty.$$

Corollary 2. *Let* $\gamma\geqq 0$. *Then*

$$\left\|\sum_{k=0}^{\infty} k^{\gamma}\Omega(|s_k-f|)\right\|<\infty$$

implies

$$\omega(f;\delta)=O\left(\delta\int_{\delta}^{1}\frac{\bar{\Omega}(x^{1+\gamma})}{x^2}\,dx\right).$$

It is easy to see that (12) reduces to (5) and (10) if $\lambda_k=\mu_k=1$ and $\mu_k=1$, $\Omega(x)=x^p$, respectively. Thus Theorem B and C, and hence all of the above results are consequences of Theorem 2.

We remark that for $\Omega(x)=x^p$ Leindler [5] proved some general statements of similar type.

It is a very interesting problem to find the analogue of Theorem 2 for the conjugate function.

We shall now generalize Theorem B in another direction. Let β be a nonnegative number and consider the condition

$$\left\|\sum_{k=0}^{\infty} k^{\beta}\Omega(|s_k-f|)\right\|<\infty \tag{14}$$

instead of (2). We ask for the differentiability properties of f and $\tilde{f}$. We prove

Theorem 3. *Let* Ω *be a concave function with properties* (4), *and let* $\beta\geqq 0$, $r=[\beta]$ *). (14) *implies that* $f, \tilde{f}$ *are* r *times differentiable, and if* r *is odd then*

$$f^{(r)}\in \mathrm{Lip}\,1 \tag{15}$$

$$\omega(\tilde{f}^{(r)};\delta)=O\left(\delta\int_{\delta}^{1}\frac{\bar{\Omega}(x^{1+\beta-r})}{x^2}\,dx\right), \tag{16}$$

while if r *is even then the role of* f *and* $\tilde{f}$ *in* (15) *and* (16) *must be inverted. Furthermore, there are functions* f_β *satisfying* (14) *with*

$$\omega(\tilde{f}_\beta^{(r)};\delta)\quad or\quad \omega(f_\beta^{(r)};\delta)\geqq c\delta\int_{\delta}^{1}\frac{\bar{\Omega}(x^{1+\beta-r})}{x^2}\,dx\quad (c>0), \tag{17}$$

according as r *is odd or even.*

The example $\Omega(x)=e^{-\frac{1}{x}}$, $f(x)=\sum_{n=1}^{\infty}\frac{1}{n^2}\sin nx$ shows that for certain convex Ω condition (14) — with arbitrary large β — does not guarantee the differentiability

*) $[\beta]$ denotes the integral part of β.

of f. On this account for convex Ω we shall investigate the condition

(18) $$\left\|\sum_{k=0}^{\infty} \Omega(k^{\beta}|s_k-f|)\right\| < \infty$$

rather than (14).

Before we state our result concerning (18), we need the following

Definition. If ω is a modulus of continuity for which $\sum_{k=0}^{\infty} \omega\left(\frac{1}{2^k}\right) < \infty$, or equivalently $\int_0^1 \frac{\omega(x)}{x}\,dx < \infty$, let

$$\omega^*(\delta) = \sup_{\{\varepsilon_k\}} \sum_{k=0}^{\infty} \omega\left(\frac{\varepsilon_k \delta}{2^k}\right),$$

where the supremum is taken over the sequences $\{\varepsilon_k\}$ which satisfy the conditions:

$$\varepsilon_k \geqq 0 \quad (k = 0, 1, \ldots), \quad \sum_{k=0}^{\infty} \varepsilon_k \leqq 1.$$

It is easy to verify that $\omega^*(\delta)$ is again a modulus of continuity, and that

$$\omega(\delta) \leqq \omega^*(\delta) \leqq \int_0^{\delta} \frac{\omega(x)}{x}\,dx.$$

With these notations we prove

Theorem 4.*) *Let Ω be convex with properties* (4), $\beta \geqq 0$, *and* $[\beta] = r$.

(i) *If* $\beta \neq [\beta]$ *then* (18) *implies*

(19) $$\omega(f^{(r)}; \delta) = O(\bar{\Omega}(\delta)).$$

(ii) *Let* $\beta = [\beta] > 0$. *From* (18) *it follows that*

(20) $$\omega(f^{(r-1)}; \delta) = O\left(\delta \int_{\delta}^{1} \frac{\bar{\Omega}(x)}{x}\,dx\right)$$

and this estimation cannot be improved. Thus if (18) *implies the existence of* $f^{(r)}$ *then*

(21) $$\int_0^1 \frac{\bar{\Omega}(x)}{x}\,dx < \infty.$$

In each of the above statements we can put $\tilde{f}$ *in place of* f.

*) We mention, that Krotov proved for a subclass of convex functions much more general results. His proofs are totally different from ours.

(iii) *Let us suppose that* (21) *is satisfied and* $r>0$. *Then* (18) *implies*

$$\omega(\tilde f^{(r)};\delta) = O\left(\int_0^\delta \frac{\bar\Omega(x)}{x}\,dx\right), \tag{22}$$

$$\omega(f^{(r)};\delta) = O\left(\bar\Omega^*(\delta)+\delta\int_\delta^1 \frac{\bar\Omega(x)}{x^2}\,dx\right), \tag{23}$$

if r is even, and the roles of $\tilde f$ *and f must be interverted in the odd case. Furthermore there are functions* f_r *satisfying* (18), *for which*

$$\omega(\tilde f_r^{(r)};\delta) \quad or \quad \omega(f_r^{(r)};\delta) \geqq c\int_0^\delta \frac{\bar\Omega(x)}{x}\,dx \quad (c>0) \tag{24}$$

according as r is even or odd.

Remark. *Estimation* (23) *is best possible also in the following sense: If*

$$\bar\Omega^*(\delta)+\delta\int_\delta^1 \frac{\bar\Omega(x)}{x^2}\,dx \neq O(\omega(\delta)), \tag{25}$$

where $\omega(\delta)$ *is an arbitrary modulus of continuity, then there is an f satisfying* (18), *but*

$$\omega(f^{(r)};\delta) \quad or \quad \omega(\tilde f^{(r)};\delta) \neq O(\omega(\delta)) \tag{26}$$

according as r is even or not.

We mention that from the proof of (i) the stronger estimation

$$\omega(f^{(r)};\delta) = O\left(\delta\int_\delta^1 \frac{\bar\Omega(x)}{x^{2-\beta+r}}\,dx\right)$$

also follows and with the aid of the function $f_0(x)=\sum_{n=1}^{\infty}\frac{1}{8n^{1+\beta}}\bar\Omega\left(\frac{1}{n}\right)\sin nx$ one can prove that this is the best possible if r is even, but we do not know what is the best estimation if r is odd.

I am grateful to Professor L. Leindler, who called my attention to these problems, and whose permanent interest and advises helped me very much in my work.

§ 1. Lemmas

Lemma 1 ([10], Lemma 2). *Let* $\{\varrho_n\}$ *be a decreasing sequence of positive numbers and let*

$$\varrho(x) = \sum_{n=1}^{\infty}\varrho_n\frac{1}{n}\sin nx.$$

Then

$$\varrho\left(\frac{\pi}{m}\right) \geqq \frac{1}{2}\frac{1}{m}\sum_{n=1}^{m}\varrho_n \quad (m=2,3,\ldots).$$

Lemma 2. *Let $\omega(x)$ be a modulus of continuity, $\beta \geqq 0$, and suppose that $E_n(f)=O\left(\frac{1}{n^\beta}\omega\left(\frac{1}{n}\right)\right)$. The following statements are true:*

(i) *if* $\beta>0$ *then* $E_n(\tilde{f})=O\left(\frac{1}{n^\beta}\omega\left(\frac{1}{n}\right)\right)$,

(ii) *if* $\beta=0$, *and* $\int\limits_0^1\frac{\omega(x)}{x}\,dx<\infty$ *then* $E_n(\tilde{f})=O\left(\int\limits_0^{1/n}\frac{\omega(x)}{x}\,dx\right)$,

(iii) *if* $\beta>[\beta]=r$, *then* $E_n(f^{(r)})=O\left(\frac{1}{n^{\beta-r}}\omega\left(\frac{1}{n}\right)\right)$,

(iv) *if* $\beta=[\beta]>0$, *then* $E_n(f^{(\beta-1)})=O\left(\frac{1}{n}\omega\left(\frac{1}{n}\right)\right)$,

(v) *if* $\beta=[\beta]$, *and* $\int\limits_0^1\frac{\omega(x)}{x}\,dx<\infty$ *then* $E_n(f^{(\beta)})=O\left(\int\limits_0^{1/n}\frac{\omega(x)}{x}\,dx\right)$.

These statements can be easily proved using the estimations below (see [9], pages 321 and 304):

$$E_n(\tilde{f})\leqq c\left(E_n(f)+\sum_{\nu=n+1}^{\infty}\frac{1}{\nu}E_\nu(f)\right),\quad E_n(f^{(r)})\leqq c_r\left(n^rE_n(f)+\sum_{\nu=n+1}^{\infty}\nu^{r-1}E_\nu(f)\right).$$

To prove (ii) and (v) use the inequality

$$\sum_{\nu=n}^{\infty}\frac{1}{\nu}\omega\left(\frac{1}{\nu}\right)\leqq\int\limits_0^{1/n}\frac{\omega(x)}{x}\,dx.$$

We omit the details.

Lemma 3. *If Ω is concave, and $\{\lambda_k\}_0^\infty$, $\{\mu_k\}_0^\infty$ are nondecreasing positive sequences then*

$$\sum_{k=0}^{\infty}\lambda_k\Omega(\mu_k|s_k(x)-f(x)|)\leqq K \tag{1.1}$$

implies that

$$E_{4n}=O\left(\log n\left(n^2\lambda_n^2\mu_n\Omega\left(\frac{\log n}{n\lambda_n}\right)\right)^{-1}\right). \tag{1.2}$$

Proof. Using the known Lebesgue estimation

$$|s_n(x)-f(x)|\leqq 3E_n(f)\log n$$

and the inequality

$$\frac{\Omega(ay_1)}{y_1}=a\frac{\Omega(ay_1)}{ay_1}\geqq a\frac{\Omega(ay_2)}{ay_2}=\frac{\Omega(ay_2)}{y_2}\quad(a>0;\ 0<y_1<y_2)$$

coming from the concavity of Ω, we get from (1.1)

$$K \geqq \left\| \sum_{k=n+1}^{2n} \lambda_k \Omega(\mu_k |s_k - f|) \right\| = \left\| \sum_{k=n+1}^{2n} \lambda_k |s_k - f| \frac{\Omega(\mu_k |s_k - f|)}{|s_k - f|} \right\| \geqq$$

$$\geqq \frac{\Omega(\mu_n E_n \log n)}{3E_n \log 2n} \lambda_n n \left\| \frac{\sum_{k=n+1}^{2n} |s_k - f|}{n} \right\| \geqq \frac{\Omega(\mu_n E_n \log n)}{6E_n \log n} \lambda_n n E_{2n}$$

i.e.

$$E_{4n} = O\big(E_{2n} \log n (n\lambda_n \Omega(\mu_{2n} E_{2n} \log n))^{-1}. \tag{1.3}$$

Now it follows from (1.1) that

$$\sum_{k=0}^{\infty} \lambda_k \mu_k |s_k(x) - f(x)| \leqq K',$$

and from this that $E_{2n} = O\big((n\lambda_n \mu_n)^{-1}\big)$. If we write this estimation in (1.3) we obtain (1.2)

Lemma 4. *If $\Omega(x)$ is convex, $\{\lambda_k\}$, $\{\mu_k\}$ are nondecreasing positive sequences, and*

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{8n\mu_n} \bar{\Omega}\left(\frac{1}{n\lambda_n}\right) \sin nx$$

then

$$\left\| \sum_{k=0}^{\infty} \lambda_k \Omega(\mu_k |s_k - f|) \right\| < \infty.$$

Proof. We introduce the notation

$$A_n(x) = \frac{1}{8n\mu_n} \bar{\Omega}\left(\frac{1}{n\lambda_n}\right) \sin nx.$$

Since $f(x)$ is odd, it is enough to consider the case $x > 0$. Let $\frac{\pi}{N} < x \leqq \frac{\pi}{N-1}$, where N is an integer. With these notations we have

$$\sum_{k=0}^{\infty} \lambda_k \Omega(\mu_k |s_k(x) - f(x)|) = \left(\sum_{k=0}^{N-1} + \sum_{k=N}^{\infty} \right) \lambda_k \Omega(\mu_k |s_k(x) - f(x)|) = B_1(x) + B_2(x). \tag{1.4}$$

Using the well-known estimation

$$\left| \sum_{l=p}^{\infty} a_l \sin lx \right| \leqq \frac{4}{x} a_p \quad (a_p \geqq a_{p+1} \geqq \ldots),$$

we get

(1.5)

$$B_2(x) = \sum_{k=N}^{\infty} \lambda_k \Omega\left(\mu_k \left|\sum_{n=k+1}^{\infty} A_n(x)\right|\right) \leqq \sum_{k=N}^{\infty} \lambda_k \Omega\left(\mu_k \frac{4}{x} \frac{1}{8(k+1)\mu_{k+1}} \bar{\Omega}\left(\frac{1}{(k+1)\lambda_{k+1}}\right)\right) \leqq$$

$$\leqq \sum_{k=N}^{\infty} \lambda_k \Omega\left(\frac{1}{Nx} \frac{N}{k+1} \bar{\Omega}\left(\frac{1}{(k+1)\lambda_{k+1}}\right)\right) \leqq \sum_{k=N}^{\infty} \lambda_k \Omega\left(\frac{N}{k+1} \bar{\Omega}\left(\frac{1}{(k+1)\lambda_{k+1}}\right)\right) \leqq$$

$$\leqq \sum_{k=N}^{\infty} \lambda_k \frac{N}{k+1} \Omega\left(\bar{\Omega}\left(\frac{1}{(k+1)\lambda_{k+1}}\right)\right) \leqq \sum_{k=N}^{\infty} \frac{N}{(k+1)^2} \leqq 1.$$

From the convexity of Ω it follows that

$$B_1(x) \leqq \sum_{k=0}^{N-1} \lambda_k \Omega\left(\mu_k \left|\sum_{n=k+1}^{N-1} A_n(x)\right| + \mu_k \left|\sum_{n=N}^{\infty} A_n(x)\right|\right) \leqq \tag{1.6}$$

$$\leqq \sum_{k=0}^{N-1} \frac{1}{2} \lambda_k \Omega\left(2\mu_k \left|\sum_{n=k+1}^{N-1} A_n(x)\right|\right) + \sum_{k=0}^{N-1} \frac{1}{2} \lambda_k \Omega\left(2\mu_k \left|\sum_{n=N}^{\infty} A_n(x)\right|\right) = B_{11}(x) + B_{12}(x).$$

Similarly to (1.5) we get

$$B_{12}(x) \leqq \sum_{k=0}^{N-1} \frac{1}{2} \lambda_k \Omega\left(2\mu_k \frac{4}{xN} \frac{1}{8\mu_N} \bar{\Omega}\left(\frac{1}{N\lambda_N}\right)\right) \leqq \frac{1}{2} \sum_{k=0}^{N-1} \lambda_N \Omega\left(\bar{\Omega}\left(\frac{1}{N\lambda_N}\right)\right) = \frac{1}{2}. \tag{1.7}$$

Finally, using the inequality $\sin x \leqq x$ $(x \geqq 0)$, we obtain

$$2B_{11}(x) \leqq \sum_{k=0}^{N-1} \lambda_k \Omega\left(2\mu_k \sum_{n=k+1}^{N-1} \frac{1}{8n\mu_n} \bar{\Omega}\left(\frac{1}{n\lambda_n}\right) nx\right) \leqq \sum_{k=0}^{N-2} \lambda_k \Omega\left(\frac{\sum_{n=k+1}^{N-1} \bar{\Omega}\left(\frac{1}{n\lambda_n}\right)}{N-1}\right) \leqq$$

$$\leqq \sum_{k=0}^{N-2} \lambda_k \frac{\sum_{n=k+1}^{N-1} \Omega\left(\bar{\Omega}\left(\frac{1}{n\lambda_n}\right)\right)}{N-1} \leqq \frac{1}{N-1} \sum_{k=0}^{N-2} \sum_{n=k+1}^{N-1} \frac{1}{n} = \frac{1}{N-1} \sum_{n=1}^{N-1} n \frac{1}{n} = 1;$$

and this — together with (1.4)—(1.7) — verifies our Lemma.

Lemma 5. *If $\Omega(x)$ is concave, $\{\lambda_k\}$, $\{\mu_k\}$ are positive nondecreasing sequences and*

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{\mu_{n^2}} \bar{\Omega}\left(\frac{1}{n^2 \lambda_{n^2}}\right) \sin nx,$$

then

$$\left\| \sum_{k=0}^{\infty} \lambda_k \Omega(\mu_k |s_k - f|) \right\| < \infty.$$

Proof. Let $A_n(x)=\frac{1}{\mu_{n^2}}\bar{\Omega}\left(\frac{1}{n^2\lambda_{n^2}}\right)\sin nx$ and $\frac{\pi}{N}<x\leqq\frac{\pi}{N-1}$. From the concavity of Ω we obtain

$$(1.8)\quad \sum_{k=0}^{\infty}\lambda_k\Omega(\mu_k|s_k(x)-f(x)|)=\left(\sum_{k=0}^{N-1}+\sum_{k=N}^{\infty}\right)\lambda_k\Omega(\mu_k|s_k(x)-f(x)|)=B_1(x)+B_2(x).$$

(1.9)

$$B_2(x)=\sum_{k=N}^{\infty}\lambda_k\Omega\left(\mu_k\left|\sum_{n=k+1}^{\infty}A_n(x)\right|\right)\leqq\sum_{k=N}^{\infty}\lambda_k\Omega\left(\mu_k\frac{4}{x}\frac{1}{\mu_{(k+1)^2}}\bar{\Omega}\left(\frac{1}{(k+1)^2\lambda_{(k+1)^2}}\right)\right)\leqq$$

$$\leqq\sum_{k=N}^{\infty}\lambda_k\Omega\left(\frac{4N}{\pi}\bar{\Omega}\left(\frac{1}{(k+1)^2\lambda_{(k+1)^2}}\right)\right)\leqq\sum_{k=N}^{\infty}\lambda_k\frac{4N}{\pi}\Omega\left(\bar{\Omega}\left(\frac{1}{(k+1)^2\lambda_{(k+1)^2}}\right)\right)\leqq$$

$$\leqq\sum_{k=N}^{\infty}\frac{4N}{\pi}\frac{1}{(k+1)^2}\leqq\frac{4}{\pi}.$$

$$(1.10)\quad B_1(x)\leqq\sum_{k=0}^{N-1}\lambda_k\Omega\left(\mu_k\sum_{n=k+1}^{N-1}A_n(x)\right)+\sum_{k=0}^{N-1}\lambda_k\Omega\left(\mu_k\left|\sum_{n=N}^{\infty}A_n(x)\right|\right)=B_{11}(x)+B_{12}(x).$$

Similarly to (1.9), we get

$$(1.11)\quad B_{12}(x)\leqq\sum_{k=0}^{N-1}\lambda_k\Omega\left(\mu_k\frac{4}{x}\frac{1}{\mu_{N^2}}\bar{\Omega}\left(\frac{1}{N^2\lambda_{N^2}}\right)\right)\leqq\frac{4N}{\pi}\lambda_N\sum_{k=0}^{N-1}\Omega\left(\bar{\Omega}\left(\frac{1}{N^2\lambda_{N^2}}\right)\right)\leqq\frac{4}{\pi}.$$

In order to estimate $B_{11}(x)$ let $2^{m-1}\leqq N-1<2^m$ and $m_k=[\overset{2}{\log}(k+1)]$. Using these notations we have

$$(1.12)\qquad B_{11}(x)\leqq\sum_{k=0}^{N-2}\lambda_k\Omega\left(\mu_k\sum_{n=k+1}^{N-1}\frac{1}{\mu_{n^2}}\bar{\Omega}\left(\frac{1}{n^2\lambda_{n^2}}\right)nx\right)\leqq$$

$$\leqq\sum_{k=0}^{2^m-1}\lambda_k\Omega\left(\sum_{n=2^{m_k}}^{2^m-1}\bar{\Omega}\left(\frac{1}{n^2\lambda_{n^2}}\right)n\frac{\pi}{2^{m-1}}\right)\leqq 2\pi\sum_{k=0}^{2^m-1}\lambda_k\Omega\left(\sum_{l=m_k}^{m-1}\sum_{n=2^l}^{2^{l+1}-1}\bar{\Omega}\left(\frac{1}{n^2\lambda_{n^2}}\right)\frac{n}{2^m}\right)\leqq$$

$$\leqq 2\pi\sum_{k=0}^{2^m-1}\lambda_k\sum_{l=m_k}^{m-1}\Omega\left(\sum_{n=2^l}^{2^{l+1}-1}\bar{\Omega}\left(\frac{1}{n^2\lambda_{n^2}}\right)\frac{n}{2^m}\right)\leqq$$

$$\leqq 2\pi\sum_{l=0}^{m-1}\sum_{m_k\leqq l}\lambda_k\Omega\left(\bar{\Omega}\left(\frac{1}{2^{2l}\lambda_{2^{2l}}}\right)2^{2l+1-m}\right)\leqq 2\pi\sum_{l=0}^{m-1}2^{l+1}\lambda_{2^{l+1}}\Omega\left(\bar{\Omega}\left(\frac{1}{2^{2l}\lambda_{2^{2l}}}\right)2^{2l+1-m}\right)=$$

$$=2\pi\sum_{l=0}^{\frac{m-1}{2}}+2\pi\sum_{l=\frac{m+1}{2}}^{m-1}\leqq 2\pi\sum_{l=0}^{\frac{m-1}{2}}2^{l+1}\lambda_{2^{l+1}}\Omega\left(\bar{\Omega}\left(\frac{1}{2^{2l}\lambda_{2^{2l}}}\right)\right)+$$

$$+2\pi\sum_{l=\frac{m+1}{2}}^{m-1}2^{l+1}\lambda_{2^{l+1}}2^{2l+1-m}\Omega\left(\bar{\Omega}\left(\frac{1}{2^{2l}\lambda_{2^{2l}}}\right)\right)\leqq 12\pi+4\pi\frac{\lambda_2}{\lambda_1}.$$

(1.8)—(1.12) verify the assertion.

Lemma 6. *Let* $r\geqq 1$ *and* Ω *concave. If*

$$\left\|\sum_{k=0}^{\infty} k^r \Omega(|s_k(f)-f|)\right\| < \infty \tag{1.13}$$

then

$$\left\|\sum_{k=0}^{\infty} k^{r-1} \Omega(|s_k(\tilde{f}')-\tilde{f}'|)\right\| < \infty. \tag{1.14}$$

Proof. Let $f\sim\sum_{k=0}^{\infty} A_k(x)$. Taking into account the concavity of Ω and $r\geqq 1$, (1.13) gives that

$$\sum_{k=0}^{\infty} k|A_k(x)| \leqq \sum_{k=0}^{\infty} k(|s_{k-1}(x)-f(x)|+|s_k(x)-f(x)|) = O\left(\sum_{k=0}^{\infty} k^r\,\Omega(|s_k(x)-f(x)|)\right),$$

i.e. $\sum_{k=0}^{\infty} kA_k(x)$ is absolutely convergent. From this it follows that $\tilde{f}'(x)=\sum_{k=0}^{\infty} kA_k(x)$, and hence

$$\sum_{k=0}^{\infty} k^{r-1}\Omega(|s_k(\tilde{f}';\,x)-\tilde{f}'(x)|) = \sum_{k=0}^{\infty} k^{r-1}\Omega\left(\left|\sum_{n=k+1}^{\infty} nA_n(x)\right|\right) =$$

$$= \sum_{k=0}^{\infty} k^{r-1}\Omega\left(\left|k(s_k(x)-f(x))+\sum_{n=k}^{\infty}(s_n(x)-f(x))\right|\right) \leqq \sum_{k=0}^{\infty} k^{r-1}\Omega(k|s_k(x)-f(x)|)+$$

$$+\sum_{k=0}^{\infty} k^{r-1}\sum_{n=k}^{\infty}\Omega(|s_n(x)-f(x)|) \leqq \sum_{k=0}^{\infty} k^r\Omega(|s_k(x)-f(x)|)+\sum_{n=0}^{\infty}\Omega(|s_n(x)-f(x)|)\sum_{k=0}^{n} k^{r-1},$$

from which, using (1.13), we obtain (1.14).

Lemma 7. *Let* $R_n(r,f)=R_n(r,f;x)=\sum_{k=0}^{n}\left(1-\left(\frac{k}{n+1}\right)^r\right)A_k(x)$, *where* $f(x)\sim\sim\sum_{k=0}^{\infty} A_k(x)$. *If* $|f|\leqq\delta$ *and* $r\geqq 1$, *then*

$$|R_n(r,f)| \leqq C_r\delta$$

where C_r *depends only on* r.

Proof. Denote $D_k(t)$ and $K_k(t)$ the k-th Dirichlet and Fejér kernel, respectively. Using the nonnegativity of $K_k(t)$ we get by an Abel rearrangement

$$|R_n(r, f; x)| = \frac{1}{(n+1)^r}\left|\sum_{k=0}^{n} s_k(x)\big((k+1)^r-k^r\big)\right| =$$

$$= \frac{1}{(n+1)^r}\left|\frac{1}{\pi}\int_{-\pi}^{\pi} f(x+u)\left\{\sum_{k=0}^{n} D_k(u)\big((k+1)^r-k^r\big)\right\}du\right| =$$

$$= \frac{1}{(n+1)^r}\frac{1}{\pi}\left|\int_{-\pi}^{\pi} f(x+u)\left\{\sum_{k=0}^{n-1}(k+1)K_k(u)\big(2(k+1)^r-k^r-(k+2)^r\big)+\right.\right.$$

$$\left.\left.+(n+1)K_n(u)\big((n+1)^r-n^r\big)\right\}du\right| \leqq$$

$$\leqq \frac{\delta}{(n+1)^r}\frac{1}{\pi}\left\{\sum_{k=0}^{n-1}(k+1)\big(k^r+(k+2)^r-2(k+1)^r\big)+\big((n+1)^r-n^r\big)(n+1)\right\} =$$

$$= O\left(\frac{\delta}{(n+1)^r}\left\{\sum_{k=1}^{n-1}(k+1)k^{r-2}+(n+1)^r\right\}\right) = O(\delta),$$

and this proves our lemma.

Lemma 8. *For*

$$\tau_n(r, f) = \tau_n(r, f; x) = \frac{2^r R_{2n-1}(r, f; x)-R_{n-1}(r, f; x)}{2^r-1} \qquad (r \geqq 1)$$

we have

$$|\tau_n(r, f)-f| \leqq c_r' E_n(f).$$

Proof.

$$|\tau_n(r, f)-f| = \left|\frac{\sum_{k=n}^{2n-1}(s_k-f)\big((k+1)^r-k^r\big)}{n^r(2^r-1)}\right| \leqq \frac{\sum_{k=n}^{2n-1}|s_k-f|\big((k+1)^r-k^r\big)}{n^r(2^r-1)} \leqq$$

$$\leqq \frac{r}{n}\sum_{k=n}^{2n-1}|s_k-f| = O\big(E_n(f)\big).$$

In the last step we used one of the results of LEINDLER [3].

Lemma 9. *Let Ω be a convex function, for which*

$$\int_0^1 \frac{\bar{\Omega}(x)}{x}dx < \infty,$$

and let $a_n \geqq 0$ *such that*

$$\sum_{k=1}^{\infty} \Omega(ka_k) \leqq K \quad \textit{for some} \quad K \geqq 1.$$

Then

$$\sum_{k=n}^{\infty} a_k \leqq K\bar{\Omega}^*\left(\frac{1}{n}\right)$$

($\bar{\Omega}^*(\delta)$ was defined in the Definition).

Proof. It is enough to prove Lemma 9 for $K=1$, namely if $K>1$ we can apply the case $K=1$ to the sequence $\frac{a_n}{K}$, using the inequality $\Omega\left(\frac{x}{K}\right)\leqq\frac{\Omega(x)}{K}$.

For $K=1$ the proof is very simple:

$$\Omega\left(\sum_{k=2^s n}^{2^{s+1}n-1} a_k\right) \leqq \Omega\left(\frac{\sum_{k=2^s n}^{2^{s+1}n-1} ka_k}{2^s n}\right) \leqq \frac{\sum_{k=2^s n}^{2^{s+1}n-1} \Omega(ka_k)}{2^s n} := \frac{\varepsilon_s}{2^s n}$$

i.e.

$$\sum_{k=2^s n}^{2^{s+1}n-1} a_k \leqq \bar{\Omega}\left(\frac{\varepsilon_s}{2^s n}\right);$$

and if we sum these inequalities for $s=0, 1, \ldots$ we get the required inequality.

Lemma 10. *If ω is concave and $E_n(f)=O\left(\omega^*\left(\frac{1}{n}\right)\right)$, then*

$$\omega(f;\delta) = O\left(\delta\int_\delta^1 \frac{\omega(x)}{x^2}\,dx+\omega^*(\delta)\right). \tag{1.15}$$

Proof. It is enough to prove (1.15) for $\delta=\frac{1}{2^m}$. We shall use the following inequality (see [9], page 333).

$$\omega\left(f;\frac{1}{n}\right) \leqq K\left(\frac{\sum_{k=0}^{n} E_k(f)}{n+1}\right). \tag{1.16}$$

From the definition of ω^* it follows that there are sequences $\{\varepsilon_s^{(r)}\}_{s=0}^{\infty}$ $(r=0, 1, \ldots, m-1)$, for which

$$\omega\left(f;\frac{1}{2^m}\right) = O\left(2^{-m}\sum_{k=1}^{2^m}\omega^*\left(\frac{1}{k}\right)\right) = O\left(2^{-m}\sum_{r=0}^{m-1}2^r\omega^*\left(\frac{1}{2^r}\right)\right) =$$

$$= O\left(2^{-m}\sum_{r=0}^{m-1}2^r\sum_{s=0}^{\infty}\omega\left(\frac{\varepsilon_s^{(r)}}{2^{r+s}}\right)\right) = O\left(2^{-m}\sum_{r=0}^{m-1}2^r\left\{\sum_{s=0}^{m-r-1}\omega\left(\frac{\varepsilon_s^{(r)}}{2^{r+s}}\right)+\omega^*\left(\frac{1}{2^m}\right)\right\}\right) =$$

$$= O\left(2^{-m}\sum_{r=0}^{m-1}2^r\sum_{s=0}^{m-r-1}\omega\left(\frac{1}{2^{r+s}}\right)+\omega^*\left(\frac{1}{2^m}\right)\right) = O\left(2^{-m}\sum_{t=0}^{m-1}2^{t+1}\omega\left(\frac{1}{2^t}\right)+\omega^*\left(\frac{1}{2^m}\right)\right) =$$

$$= O\left(2^{-m}\int_{2^{-m}}^{1}\frac{\omega(x)}{x^2}\,dx+\omega^*\left(\frac{1}{2^m}\right)\right)$$

and this proves (1.15).

§ 2. Proof of the theorems

Proof of Theorem 1. It is enough to prove the theorem for convex Ω, namely if Ω is concave, then (2) implies

$$\left\|\sum_{k=0}^{\infty}|s_k-f|\right\|<\infty,$$

and if we apply the second part of Theorem 1 to the convex function $\Omega(x)=x$ we get that

$$\omega(\tilde{f};\delta)=O\left(\int_0^\delta \frac{x}{x}\,dx\right)=O(\delta)$$

i.e. $\tilde{f}\in \mathrm{Lip}\,1$.

Let thus Ω be convex. First we prove (7). Let us denote by $\sigma_n(f)=\sigma_n(f;x)$ the n-th $(C,1)$-mean of the Fourier series of f, and let

$$\tau_n(f)=\tau_n(f;x)=2\sigma_{2n-1}(f;x)-\sigma_{n-1}(f;x)=\frac{\sum_{k=n}^{2n-1}s_k(x)}{n}.$$

From (2), using the convexity of Ω we get

$$\text{(2.1)}\qquad |\sigma_n(f)-f|=\bar{\Omega}(\Omega(|\sigma_n(f)-f|))\leqq\bar{\Omega}\left(\Omega\left(\frac{\sum_{k=0}^{n}|s_k-f|}{n+1}\right)\right)\leqq$$

$$\leqq\bar{\Omega}\left(\frac{\sum_{k=0}^{n}\Omega(|s_k-f|)}{n+1}\right)=O\left(\bar{\Omega}\left(\frac{1}{n}\right)\right).$$

With the notation

$$f-\sigma_n(f)=g_n(f)$$

we have

$$\text{(2.2)}\qquad \sigma_n(f)-f=(\sigma_n(\sigma_n(f))-\sigma_n(f))+(\sigma_n(g_n(f))-g_n(f)).$$

We can write (2.1) in the form $g_n(f)=O\left(\bar{\Omega}\left(\frac{1}{n}\right)\right)$, from which $\sigma_n(g_n(f))=$ $=O\left(\bar{\Omega}\left(\frac{1}{n}\right)\right)$, and so (2.2) implies

$$\text{(2.3)}\qquad \sigma_n(\sigma_n(f))-\sigma_n(f)=O\left(\bar{\Omega}\left(\frac{1}{n}\right)\right).$$

If we keep in view the expression of $\sigma_n(f)$, it is easy to see that

$$\sigma_n(\sigma_n(f))-\sigma_n(f)=-\frac{(\tilde{\sigma}_n(f))'}{n+1},$$

so (2.3) implies $\tilde{\sigma}_n'(f)=O\left[n\bar{\Omega}\left(\frac{1}{n}\right)\right]$, and together with this

$$(2.4)\qquad (\tilde{\tau}_n(f))' = (\tau_n(\tilde{f}))' = O\left[n\bar{\Omega}\left(\frac{1}{n}\right)\right].$$

Now (2.1) gives $E_n(f)=O\left[\bar{\Omega}\left(\frac{1}{n}\right)\right]$, from which by Lemma 2 (ii) it follows $E_n(\tilde{f})=O\left(\int\limits_0^{1/n}\frac{\bar{\Omega}(x)}{x}\,dx\right)$. It is known (see e.g. Lemma 8) that $|\tau_n(g)-g|\leqq \leqq KE_n(g)$; and hence, also using the previous estimation, we get

$$(2.5)\qquad |\tau_n(\tilde{f})-\tilde{f}| = O\left(\int\limits_0^{1/n}\frac{\bar{\Omega}(x)}{x}\,dx\right).$$

Now we are ready to prove (7). If $|h|\leqq\frac{1}{n}$, then (2.4) and (2.5) give

$$|\tilde{f}(x)-\tilde{f}(x+h)|\leqq|\tilde{f}(x)-\tau_n(\tilde{f};x)|+|\tau_n(\tilde{f};x)-\tau_n(\tilde{f};x+h)|+|\tau_n(\tilde{f};x+h)-\tilde{f}(x+h)|=$$

$$=O\left(\int\limits_0^{1/n}\frac{\bar{\Omega}(x)}{x}\,dx+|h\tau_n'(\tilde{f};\,x+\vartheta h)|\right)=$$

$$=O\left(\int\limits_0^{1/n}\frac{\bar{\Omega}(x)}{x}\,dx+|h|\,n\bar{\Omega}\left(\frac{1}{n}\right)\right)=O\left(\int\limits_0^{1/n}\frac{\bar{\Omega}(x)}{x}\,dx\right),$$

and this is equivalent to (7).

By Lemma 4, (2) is satisfied by the function

$$f_0(x)=\sum_{n=1}^{\infty}\frac{1}{\gamma_n}\bar{\Omega}\left(\frac{1}{n}\right)\sin nx.$$

Then,

$$\tilde{f}_0(x)=-\sum_{=1}^{\infty}\frac{1}{8n}\bar{\Omega}\left(\frac{1}{n}\right)\cos nx,$$

and here the right hand side is the Fourier series of a continuous function only if

$$\sum_{n=1}^{\infty}\frac{1}{n}\bar{\Omega}\left(\frac{1}{n}\right)<\infty$$

(for the $(C, 1)$ means of this series must then be bounded), and this is the same as (6). The statement, that in case (6) $\tilde{f}$ is continuous is a direct consequence of (7), proved above.

Let $h=\frac{\pi}{2^{k+1}}$; then

$$\tilde{f}_0(h)-\tilde{f}_0(0)=\sum_{n=2^k+1}^{\infty}\frac{1}{8n}\bar{\Omega}\left(\frac{1}{n}\right)-\sum_{n=2^k+1}^{\infty}\frac{1}{8n}\bar{\Omega}\left(\frac{1}{n}\right)\cos nh+\sum_{n=1}^{2^k}\frac{1}{8n}\bar{\Omega}\left(\frac{1}{n}\right)2\sin^2 n\frac{h}{2}.$$

It is easy to see that

$$\sum_{n=2^k+1}^{\infty} \frac{1}{8n} \bar{\Omega}\left(\frac{1}{n}\right) \cos nh \leqq 0,$$

and so

$$\tilde{f}_0(h)-\tilde{f}_0(0) \geqq \sum_{n=2^k+1}^{\infty} \frac{1}{8n} \bar{\Omega}\left(\frac{1}{n}\right) \geqq c \int_0^{1/2^k} \frac{\bar{\Omega}(x)}{x}\, dx,$$

and hence (8) follows by a standard argument.

We have completed our proof.

Proof of Theorem 2. We have to consider two cases separately

Case I: Ω is convex. Let

$$\sum_{k=0}^{\infty} \lambda_k \Omega(\mu_k |s_k(x)-f(x)|) \leqq K.$$

We have

$$\Omega(\mu_n E_{2n}) \leqq \Omega\left(\mu_n \left\| \frac{\sum_{k=n+1}^{2n} |s_k-f|}{n} \right\|\right) \leqq \Omega\left(\left\| \frac{\sum_{k=n+1}^{2n} \mu_k |s_k-f|}{n} \right\|\right) =$$

$$= \left\| \Omega\left(\frac{\sum_{k=n+1}^{2n} \mu_k |s_k-f|}{n} \right) \right\| \leqq \left\| \frac{\sum_{k=n+1}^{2n} \Omega(\mu_k |s_k-f|)}{n} \right\| \leqq$$

$$\leqq \left\| \frac{\sum_{k=n+1}^{2n} \lambda_k \Omega(\mu_k |s_k-f|)}{n\lambda_n} \right\| \leqq \frac{K}{n\lambda_n},$$

i.e.

$$E_{2n}(f) = O\left(\frac{1}{\mu_n} \bar{\Omega}\left(\frac{1}{n\lambda_n}\right)\right),$$

and hence, using the inequality (1.16),

$$\omega\left(f; \frac{1}{n}\right) = O\left(\frac{\sum_{k=0}^{n} E_k}{n+1}\right) = O\left(\frac{\sum_{k=0}^{n} E_{2k}}{n+1}\right) = O\left(\frac{1}{n} \sum_{k=1}^{n} \frac{1}{\mu_k} \bar{\Omega}\left(\frac{1}{k\lambda_k}\right)\right)$$

and this is (12).

Let

$$f_0(x) = \sum_{n=1}^{\infty} \frac{1}{8n\mu_n} \bar{\Omega}\left(\frac{1}{n\lambda_n}\right) \sin nx.$$

By Lemma 4, f_0 satisfies (11). Now applying Lemma 1 to f_0 we get

$$f_0\left(\frac{\pi}{n}\right)-f_0(0) \geqq \frac{1}{2}\frac{1}{8}\frac{1}{n}\sum_{k=1}^{n}\frac{1}{\mu_k}\bar{\Omega}\left(\frac{1}{k\lambda_k}\right),$$

and this proves (13).

Case II: Ω is concave. By Lemma 3 we have

$$E_{4n}(f) = O\left(\log n\left(n^2\lambda_n^2\mu_n\Omega\left(\frac{\log n}{n\lambda_n}\right)\right)^{-1}\right). \tag{2.6}$$

Let m_k resp. n_k the least and the greatest n (if any), for which

$$\frac{1}{(k+1)\lambda_{k+1}} < \Omega\left(\frac{\log n}{n\lambda_n}\right) \leqq \frac{1}{k\lambda_k}. \tag{2.7}$$

Ω is concave, so there is a $c>0$ for which

$$\Omega\left(\frac{\log k}{k\lambda_k}\right) \geqq c\frac{\log k}{k\lambda_k} > \frac{1}{k\lambda_k}$$

if k is large enough. From this and (2.7) it follows at once for $k \geqq k_0$ that

$$m_k \geqq k+1, \quad \lambda_{m_k} \geqq \lambda_{k+1}, \quad \mu_{m_k} \geqq \mu_k \tag{2.8}$$

$$\frac{\log m_k}{m_k\lambda_{m_k}} \leqq \bar{\Omega}\left(\frac{1}{k\lambda_k}\right), \quad \frac{\log n_k}{n_k\lambda_{n_k}} \geqq \bar{\Omega}\left(\frac{1}{(k+1)\lambda_{k+1}}\right). \tag{2.9}$$

We shall show that for $k \geqq k_0$

(2.10)

$$\sum_{n=m_k}^{n_k}\log n\left(n^2\lambda_n^2\mu_n\Omega\left(\frac{\log n}{n\lambda_n}\right)\right)^{-1} = O\left((k+1)\left(\frac{1}{\mu_k}\bar{\Omega}\left(\frac{1}{k\lambda_k}\right)-\frac{1}{\mu_{k+1}}\bar{\Omega}\left(\frac{1}{(k+1)\lambda_{k+1}}\right)\right)\right).$$

First we consider the case $n_k=m_k=n$. Using the inequalities

$$\bar{\Omega}\left(\frac{1}{(k+1)\lambda_{k+1}}\right) \leqq \bar{\Omega}\left(\frac{1}{(k+1)\lambda_k}\right) \leqq \frac{k}{k+1}\bar{\Omega}\left(\frac{1}{k\lambda_k}\right), \quad \frac{\Omega(x)}{x} \geqq C$$

coming from the concavity of Ω, we obtain for $k \geqq k_0$

$$\log n\left(n^2\lambda_n^2\mu_n\Omega\left(\frac{\log n}{n\lambda_n}\right)\right)^{-1} = O\left(\frac{1}{n\lambda_n\mu_n}\right) = O\left(\frac{\log n}{n\lambda_n\mu_n}\right) = O\left(\frac{1}{\mu_k}\bar{\Omega}\left(\frac{1}{k\lambda_k}\right)\right) =$$

$$= O\left(\frac{(k+1)}{\mu_k}\left(\bar{\Omega}\left(\frac{1}{k\lambda_k}\right)-\bar{\Omega}\left(\frac{1}{(k+1)\lambda_{k+1}}\right)\right)\right) =$$

$$= O\left((k+1)\left(\frac{1}{\mu_k}\bar{\Omega}\left(\frac{1}{k\lambda_k}\right)-\frac{1}{\mu_{k+1}}\bar{\Omega}\left(\frac{1}{(k+1)\lambda_{k+1}}\right)\right)\right).$$

If, however, $n_k > m_k$ and $k \geqq k_0$, then

$$\sum_{n=m_k}^{n_k} \log n \left[n^2\lambda_n^2\mu_n\Omega\left(\frac{\log n}{n\lambda_n}\right)\right]^{-1} = O\left(\frac{(k+1)\lambda_{k+1}}{\lambda_{m_k}^2\mu_{m_k}} \sum_{n=m_k}^{n_k} \frac{\log n}{n^2}\right) =$$

$$= O\left(\frac{(k+1)}{\lambda_{m_k}\mu_{m_k}} \int_{m_k}^{n_k} \frac{\log x - 1}{x^2}\,dx\right) = O\left(\frac{(k+1)}{\lambda_{m_k}\mu_{m_k}}\left(\frac{\log m_k}{m_k} - \frac{\log n_k}{n_k}\right)\right) =$$

$$= O\left(\frac{(k+1)}{\mu_k}\left(\frac{\log m_k}{m_k\lambda_{m_k}} - \frac{\log n_k}{n_k\lambda_{n_k}}\right)\right) = O\left(\frac{(k+1)}{\mu_k}\left(\bar\Omega\left(\frac{1}{k\lambda_k}\right) - \bar\Omega\left(\frac{1}{(k+1)\lambda_{k+1}}\right)\right)\right) =$$

$$= O\left((k+1)\left(\frac{1}{\mu_k}\bar\Omega\left(\frac{1}{k\lambda_k}\right) - \frac{1}{\mu_{k+1}}\bar\Omega\left(\frac{1}{(k+1)\lambda_{k+1}}\right)\right)\right).$$

Thus we have proved (2.10) for $k \geqq k_0$.

Let now $m_i \leqq m \leqq n_i$. Using (2.6) and (2.10) we get

$$\omega\left(f; \frac{1}{m}\right) = O\left(\frac{1}{m}\sum_{k=0}^{m} E_k(f)\right) = O\left(\frac{1}{m}\sum_{k=0}^{m} E_{4k}(f)\right) =$$

$$= O\left(\frac{1}{m}\sum_{k=1}^{i}\sum_{n=m_k}^{n_k} \log n \left(n^2\lambda_n^2\mu_n\Omega\left(\frac{\log n}{n\lambda_n}\right)\right)^{-1}\right) = O\left(\frac{1}{m}\left(\sum_{k=1}^{k_0-1} + \sum_{k=k_0}^{i}\right)\right) =$$

$$= O\left(\frac{1}{m} + \frac{1}{m}\sum_{k=k_0}^{i}(k+1)\left(\frac{1}{\mu_k}\bar\Omega\left(\frac{1}{k\lambda_k}\right) - \frac{1}{\mu_{k+1}}\bar\Omega\left(\frac{1}{(k+1)\lambda_{k+1}}\right)\right)\right) =$$

$$= O\left(\frac{1}{m}\sum_{k=1}^{i}\frac{1}{\mu_k}\bar\Omega\left(\frac{1}{k\lambda_k}\right)\right) = O\left(\frac{1}{m}\sum_{k=1}^{m}\frac{1}{\mu_k}\bar\Omega\left(\frac{1}{k\lambda_k}\right)\right),$$

which proves (12). Let

$$f_0(x) = \sum_{n=1}^{\infty}\frac{1}{\mu_{n^2}}\bar\Omega\left(\frac{1}{n^2\lambda_{n^2}}\right)\sin nx.$$

By Lemma 5, f_0 satisfies (11). Applying again Lemma 1 (it is easy to see that it is applicable), we obtain

$$f_0\left(\frac{\pi}{n}\right) - f_0(0) \geqq \frac{1}{2}\frac{1}{n}\sum_{k=1}^{n} k\frac{1}{\mu_{k^2}}\bar\Omega\left(\frac{1}{k^2\lambda_{k^2}}\right) \geqq \frac{1}{6}\frac{1}{n}\sum_{k=1}^{n}(2k+1)\frac{1}{\mu_{k^2}}\bar\Omega\left(\frac{1}{k^2\lambda_{k^2}}\right) \geqq$$

$$\geqq \frac{1}{6}\frac{1}{n}\sum_{k=1}^{n^2-1}\frac{1}{\mu_k}\bar\Omega\left(\frac{1}{k\lambda_k}\right) \geqq \frac{1}{6}\frac{1}{n}\sum_{k=1}^{n}\frac{1}{\mu_k}\bar\Omega\left(\frac{1}{k\lambda_k}\right),$$

and this is (13). — The proof of Theorem 2 is thus completed.

Proof of Theorem 3. We shall consider only the case when r is odd, the other case could be treated similarly.

If we apply Lemma 6 r-times, we gett hat (14) implies

$$\left\| \sum_{k=0}^{\infty} k^{\beta-r} \Omega\left(|s_k(\tilde{f}^{(r)}) - \tilde{f}^{(r)}|\right) \right\| < \infty,$$

and hence, using the assertion (i) of Theorem 1, we get $\tilde{\tilde{f}}^{(r)} = f^{(r)} \in \operatorname{Lip} 1$, while using Corollary 2 of Theorem 2 we obtain

$$\omega(\tilde{f}^{(r)}; \delta) = O\left(\delta \int_{\delta}^{1} \frac{\bar{\Omega}(x^{1+\beta-r})}{x^2}\, dx\right)$$

as it was proposed in (16).

Let

$$f_\beta(x) = \sum_{n=1}^{\infty} \bar{\Omega}\left(\frac{1}{n^{2+\beta}}\right) \sin nx.$$

If we run through the proof of Lemma 5 we can see that its proof equally works for f_β, so f_β satisfies (14). Keeping in mind that $\bar{\Omega}$ is convex, we have

$$n^{\beta+2} \bar{\Omega}\left(\frac{1}{n^{\beta+2}}\right) \geqq (n+1)^{\beta+2} \bar{\Omega}\left(\frac{1}{(n+1)^{\beta+2}}\right),$$

and this implies that

$$n^{r+1} \bar{\Omega}\left(\frac{1}{n^{\beta+2}}\right) \geqq (n+1)^{r+1} \bar{\Omega}\left(\frac{1}{(n+1)^{\beta+2}}\right),$$

so we can apply Lemma 1 to $\tilde{f}_\beta^{(r)}$, and this gives

$$\text{(2.11)} \quad \left| \tilde{f}_\beta^{(r)}\left(\frac{\pi}{n}\right) - \tilde{f}_\beta(0) \right| \geqq \frac{1}{2} \frac{1}{n} \sum_{k=1}^{n} k^{r+1} \bar{\Omega}\left(\frac{1}{k^{\beta+2}}\right) \geqq c \frac{1}{n} \int_{1/n}^{1} \frac{\bar{\Omega}(x^{\beta+2})}{x^{r+3}}\, dx =$$

$$= c' \frac{1}{n} \int_{n^{-\frac{2+\beta}{1+\beta-r}}}^{1} \frac{\bar{\Omega}(u^{1+\beta-r})}{u^{\gamma}}\, du,$$

where

$$\gamma = \frac{1+\beta-r}{2+\beta}\left(r+3+\frac{1+r}{1+\beta-r}\right) \geqq 2.$$

Also $\dfrac{2+\beta}{1+\beta-r} \geqq 1$, so we get from (2.11) that

$$\left| \tilde{f}_\beta^{(r)}\left(\frac{\pi}{n}\right) - \tilde{f}_\beta^{(r)}(0) \right| \geqq c \frac{1}{n} \int_{1/n}^{1} \frac{\bar{\Omega}(x^{1+\beta-r})}{x^2}\, dx,$$

which was to be proved.

Thus we have completed our proof.

Proof of Theorem 4. Let $f(x)\sim\sum_{k=0}^{\infty}A_k(x)$ and

$$R_n(\beta, f; x) = \sum_{k=0}^{n}\left(1-\left(\frac{k}{n+1}\right)^{\beta}\right)A_k(x).$$

Using an Abel rearrangement we get from (18)

$$\Omega\left(\frac{(n+1)^{\beta}}{2^{\beta}(\beta+1)}|R_n(\beta+1,f)-f|\right) = \Omega\left(\frac{(n+1)^{\beta}}{2^{\beta}(\beta+1)}\left|\frac{\sum_{k=0}^{n}s_k((k+1)^{\beta+1}-k^{\beta+1})}{(n+1)^{\beta+1}}-f\right|\right) =$$

$$= \Omega\left(\frac{(n+1)^{\beta}}{2^{\beta}(\beta+1)}\left|\frac{\sum_{k=0}^{n}(s_k-f)((k+1)^{\beta+1}-k^{\beta+1})}{(n+1)^{\beta+1}}\right|\right) \leqq \Omega\left(\frac{|s_0-f|+\sum_{k=1}^{n}k^{\beta}|s_k-f|}{n+1}\right) \leqq$$

$$\leqq \frac{\Omega(|s_0-f|)+\sum_{k=1}^{n}\Omega(k^{\beta}|s_k-f|)}{n+1} \leqq \frac{K}{n+1},$$

which implies

$$|R_n(\beta+1,f)-f| = O\left(\frac{1}{n^{\beta}}\bar{\Omega}\left(\frac{1}{n}\right)\right). \tag{2.12}$$

Now $R_n(\beta+1,f)$ is a trigonometric polinomial of order at most n, so (2.12) implies

$$E_n(f) = O\left(\frac{1}{n^{\beta}}\bar{\Omega}\left(\frac{1}{n}\right)\right). \tag{2.13}$$

We shall treat after that the cases (i)—(iii) separately.

Case (i). By Lemma 2 (iii) from (2.13) it follows that $E_n(f^{(r)})=O\left(\frac{1}{n^{\beta-r}}\bar{\Omega}\left(\frac{1}{n}\right)\right)$, and this, connecting with inequality (1.16) gives

$$\omega\left(f^{(r)};\frac{1}{n}\right) = O\left(\frac{1}{n}\sum_{k=1}^{n}\frac{1}{k^{\beta-r}}\bar{\Omega}\left(\frac{1}{k}\right)\right)\quad\left(=O\left(\frac{1}{n}\int_{1/n}^{1}\frac{\bar{\Omega}(x)}{x^{2-\beta+r}}dx\right)\right).$$

From the concavity of $\bar{\Omega}$ it follows that $\bar{\Omega}\left(\frac{1}{k}\right)\leqq\frac{n}{k}\bar{\Omega}\left(\frac{1}{n}\right)$ $(n\geqq k)$ and so

$$\omega\left(f^{(r)};\frac{1}{n}\right) = O\left(\frac{1}{n}\sum_{k=1}^{n}\frac{1}{k^{\beta-r}}\frac{n}{k}\bar{\Omega}\left(\frac{1}{n}\right)\right) = O\left(\bar{\Omega}\left(\frac{1}{n}\right)\right),$$

and this proves (19).

Case (ii) According to Lemma 2 (iv), (2.13) implies

$$E_n(f^{(r-1)}) = O\left(\frac{1}{n}\bar{\Omega}\left(\frac{1}{n}\right)\right),$$

and so

$$\omega\left(f^{(r-1)};\frac{1}{n}\right)=O\left(\frac{1}{n}\sum_{k=1}^{n}\frac{1}{k}\bar{\Omega}\left(\frac{1}{k}\right)\right)=O\left(\frac{1}{n}\int_{1/n}^{1}\frac{\bar{\Omega}(x)}{x}\,dx\right),$$

from which (20) already follows.

Let r e.g. even, and

$$f_0(x)=\sum_{n=1}^{\infty}\frac{1}{n^{r+1}}\bar{\Omega}\left(\frac{1}{n^2}\right)\cos nx$$

(if r is odd then we must take $\sin x$ in place of $\cos x$). f_0 satisfies (18):

$$\sum_{k=0}^{\infty}\Omega\left(k^r\left|\sum_{n=k+1}^{\infty}\frac{1}{n^{r+1}}\bar{\Omega}\left(\frac{1}{n^2}\right)\cos nx\right|\right)\leqq\sum_{k=0}^{\infty}\Omega\left(k^r\frac{1}{(k+1)^r}\bar{\Omega}\left(\frac{1}{(k+1)^2}\right)\right)\leqq\sum_{k=0}^{\infty}\frac{1}{(k+1)^2}.$$

$$f_0^{(r-1)}(x)=(-1)^{r/2}\sum_{n=1}^{\infty}\frac{1}{n^2}\bar{\Omega}\left(\frac{1}{n^2}\right)\sin nx,$$

and so using Lemma 1 we get

$$\left|f_0^{(r-1)}\left(\frac{\pi}{n}\right)-f_0^{(r-1)}(0)\right|\geqq\frac{1}{2}\frac{1}{n}\sum_{k=1}^{n}\frac{1}{k}\bar{\Omega}\left(\frac{1}{k^2}\right)\geqq\frac{1}{6}\frac{1}{n}\sum_{k=1}^{n}(2k+1)\frac{1}{k^2}\bar{\Omega}\left(\frac{1}{k^2}\right)\geqq$$

$$\geqq\frac{1}{6}\frac{1}{n}\sum_{k=1}^{n^2-1}\frac{1}{k}\bar{\Omega}\left(\frac{1}{k}\right)\geqq c\frac{1}{n}\int_{1/n}^{1}\frac{\bar{\Omega}(x)}{x}\,dx,$$

and this proves that (20) is best possible in general.

Lemma 2 (i) and the above proofs show that all of the above statements are true for the conjugate function, too.

Case (iii). We shall consider the case when r is even. Let

$$f=R_n(r+1,f)+g_n(f).$$

With this notation

(2.14)
$$R_n(r+1,f)-f=\big(R_n(r+1,R_n(r+1,f))-R_n(r+1,f)\big)+\big(R_n(r+1,\,g_n(f))-g_n(f)\big).$$

By (2.12) $g_n(f)=O\left(\frac{1}{n^r}\bar{\Omega}\left(\frac{1}{n}\right)\right)$, and this implies by Lemma 7 that $R_n(r+1,g_n(f))=O\left(\frac{1}{n^r}\bar{\Omega}\left(\frac{1}{n}\right)\right)$, and so from (2.14) it follows that

(2.15) $$R_n(r+1,\,R_n(r+1,f))-R_n(r+1,f)=O\left(\frac{1}{n^r}\bar{\Omega}\left(\frac{1}{n}\right)\right).$$

Let

$$R_n(r+1,f)\sim\sum_{k=0}^{n}A_k(x).$$

Then

$$R_n(r+1,\, R_n(r+1,f))-R_n(r+1,f)=\sum_{k=0}^{n}\left(1-\left(\frac{k}{n+1}\right)^{r+1}\right)A_k(x)-\sum_{k=0}^{n}A_k(x)=$$

$$=-\frac{1}{(n+1)^{r+1}}\sum_{k=0}^{n}k^{r+1}A_k(x)=\frac{(-1)^{r/2+1}}{(n+1)^{r+1}}(\tilde{R}_n(r+1,f))^{(r+1)}.$$

This equality together with (2.15) gives

$$\tilde{R}_n^{(r+1)}(r+1,f)=O\left(n\bar{\Omega}\left(\frac{1}{n}\right)\right),$$

from which

$$\tilde{\tau}_n^{(r+1)}(r+1,f)=\tau_n^{(r+1)}(r+1,\tilde{f})=\tau_n'(r+1,\tilde{f}^{(r)})=O\left(n\bar{\Omega}\left(\frac{1}{n}\right)\right) \tag{2.16}$$

follows at once $(\tau_n(r,f)$ was defined in Lemma 8).

(2.13) implies by Lemma 2 (i) and (v) and by Lemma 8 that

$$|\tau_n(r+1,\tilde{f}^{(r)})-\tilde{f}^{(r)}|=O\left(\int_0^{1/n}\frac{\bar{\Omega}(x)}{x}\,dx\right). \tag{2.17}$$

Now we get (22) from (2.16) and (2.17) as we got (7) in Theorem 1 from (2.4) and (2.5).

Before proving (23) we show that $f^{(r)}$ is the sum of its Fourier series. Because of the continuity of $f^{(r)}$ it is enough to prove that its Fourier series everywhere convergent. With the usual notations

$$f^{(r)}(x)\sim(-1)^{r/2}\sum_{k=0}^{\infty}k^rA_k(x),$$

$$\sum_{k=m}^{n}k^rA_k(x)=\sum_{k=m}^{n-1}(k^r-(k+1)^r)s_k(x)-m^rs_{m-1}(x)+n^rs_n(x)= \tag{2.18}$$

$$=O\left(\sum_{k=m}^{n-1}k^{r-1}|s_k(x)-f(x)|+m^r|s_{m-1}(x)-f(x)|+n^r|s_n(x)-f(x)|\right).$$

Lemma 9 shows by (18) that

$$\sum_{k=m}^{n-1}k^{r-1}|s_k(x)-f(x)|\to 0\quad\text{as}\quad n,m\to\infty,$$

moreover $\Omega(n^r|s_n(x)-f(x))\to 0$ $(n\to\infty)$, and this implies $n^r|s_n(x)-f(x)|\to 0$ as $n\to\infty$. Thus (2.18) gives the convergence of $\sum_{k=0}^{\infty}k^rA_k(x)$, and so

$$f^{(r)}(x)=(-1)^{r/2}\sum_{k=0}^{\infty}k^rA_k(x).$$

On this account the following transformations are legitimate, and (2.12), as well as Lemma 9 give

$$(-1)^{r/2}(\sigma_n(f^{(r)})-f^{(r)}) = \sum_{k=0}^{n}\left(1-\frac{k}{n+1}\right)k^r A_k - \sum_{k=0}^{\infty} k^r A_k = (n+1)^r(R_n(r+1,f)-f)-$$

$$-\sum_{k=n+1}^{\infty}(k^r-(n+1)^r)A_k = (n+1)^r(R_n(r+1,f)-f)+\sum_{k=n+2}^{\infty}(s_{k-1}-f)(k^r-(k-1)^r) =$$

$$= O\left((n+1)^r\frac{1}{n^r}\bar{\Omega}\left(\frac{1}{n}\right)+\sum_{k=n+2}^{\infty}|s_{k-1}-f|(k-1)^{r-1}\right) = O\left(\bar{\Omega}\left(\frac{1}{n}\right)+\bar{\Omega}^*\left(\frac{1}{n}\right)\right) =$$

$= O\left(\bar{\Omega}^*\left(\frac{1}{n}\right)\right)$, from which $E_n(f^{(r)})=O\left(\bar{\Omega}^*\left(\frac{1}{n}\right)\right)$ follows at once. Now we can apply Lemma 10, and we get (23).

To prove (24) let

$$f_r(x) = \sum_{n=1}^{\infty}\frac{1}{8n^{r+1}}\bar{\Omega}\left(\frac{1}{n}\right)\sin nx.$$

By Lemma 4, f_r satisfies (18). Now

$$f_r^{(r)}(x) = (-1)^{r/2+1}\sum_{n=1}^{\infty}\frac{1}{8n}\bar{\Omega}\left(\frac{1}{n}\right)\cos nx$$

and in the proof of Theorem 1 we have already seen that for this function (24) is true.

The proof of Theorem 4 is thus completed.

Proof of Remark. Let r be, for example, an odd number.

We separate the proof into two cases.

1. $\delta\int_{\delta}^{1}\frac{\bar{\Omega}(x)}{x^2}dx \neq O(\omega(\delta))$. In this case by the aid of the above defined function

$$f_r(x) = \sum_{n=1}^{\infty}\frac{1}{8n^{r+1}}\bar{\Omega}\left(\frac{1}{n}\right)\sin nx$$

the proof can be easily carried out.

2. If $\delta\int_{\delta}^{1}\frac{\bar{\Omega}(x)}{x^2}dx = O(\omega(\delta))$, then there is a sequence of natural numbers $\{n_m\}$, for whicl.

$$\omega\left(\frac{\pi}{2n_m}\right) < \frac{1}{4^m}\bar{\Omega}^*\left(\frac{\pi}{2n_m}\right). \tag{2.19}$$

Let n be a fixed natural number, and $\varepsilon_0 \geqq \varepsilon_1 \geqq \dots$; $\sum_{k=0}^{\infty} \varepsilon_k \leqq 1$. Let $c_m = \bar{\Omega}\left(\frac{\varepsilon_k}{2^k n}\right)$ if $2^k n \leqq m < 2^{k+1} n$, and

$$f(x) = f_{\{\varepsilon_k\},n}(x) = (-1)^{\frac{r-1}{2}} \sum_{m=n}^{\infty} \frac{1}{2m^{r+1}} c_m \cos mx.$$

With the aid of (21) we get that $\tilde{f}^{(r)}$ exists, and less than a bound independent from $\{\varepsilon_k\}$ and n. We show that f satisfies (18).

$$(2.20) \quad \sum_{k=0}^{\infty} \Omega(k^r |s_k(x) - f(x)|) \leqq \sum_{k=1}^{n-1} \Omega\left(k^r \sum_{m=n}^{\infty} \frac{1}{2m^{r+1}} c_m\right) + \sum_{k=n}^{\infty} \Omega\left(k^r \sum_{m=k+1}^{\infty} \frac{1}{2m^{r+1}} c_m\right) \leqq$$

$$\leqq \sum_{k=1}^{n-1} \Omega\left(k^r \frac{1}{2n^r} c_n\right) + \sum_{k=n}^{\infty} \Omega\left(k^r \frac{1}{2k^r} c_k\right) \leqq \frac{1}{2} \sum_{k=1}^{n-1} \Omega\left(\bar{\Omega}\left(\frac{\varepsilon_0}{n}\right)\right) +$$

$$+ \frac{1}{2} \sum_{k=0}^{\infty} 2^k n \Omega\left(\bar{\Omega}\left(\frac{\varepsilon_k}{2^k n}\right)\right) \leqq \frac{\varepsilon_0}{2} + \frac{\sum_{k=0}^{\infty} \varepsilon_k}{2} \leqq 1.$$

Now

$$\tilde{f}^{(r)}(0) - \tilde{f}^{(r)}\left(\frac{\pi}{2n}\right) = \sum_{m=n}^{\infty} \frac{1}{m} c_m - \sum_{m=n}^{\infty} \frac{1}{m} c_m \cos m \frac{\pi}{2n},$$

and from the monotonicity of $\{c_m\}$ it follows that

$$\sum_{m=n}^{\infty} \frac{c_m}{m} \cos m \frac{\pi}{2n} \leqq 0,$$

and so

$$\tilde{f}^{(r)}(0) - \tilde{f}^{(r)}\left(\frac{\pi}{2n}\right) \geqq \frac{1}{2} \sum_{k=0}^{\infty} \bar{\Omega}\left(\frac{\varepsilon_k}{2^k n}\right) \sum_{m=2^k n}^{2^{k+1} n - 1} \frac{1}{m} \geqq \frac{1}{4} \sum_{k=0}^{\infty} \bar{\Omega}\left(\frac{\varepsilon_k}{2^k n}\right).$$

Consequently, by a suitable choice of $\{\varepsilon_k\}$ one can attain that the above defined function $f_{\{\varepsilon_k\},n}(x) = f_n(x)$ satisfies

$$(2.21) \qquad \tilde{f}_n^{(r)}(0) - \tilde{f}_n^{(r)}\left(\frac{\pi}{2n}\right) \geqq \frac{1}{10} \bar{\Omega}^*\left(\frac{1}{n}\right) \geqq \frac{1}{20} \bar{\Omega}^*\left(\frac{\pi}{2n}\right),$$

for in the definition of $\bar{\Omega}^*$ we could have supposed that $\{\varepsilon_k\}$ is monotone.

Let now

$$f(x) = \sum_{m=1}^{\infty} \frac{1}{2^m} f_{n_m}(x).$$

This function f satisfies (18); indeed, using the convexity of Ω, we get from (2.20)

$$\sum_{k=0}^{\infty}\Omega(k^r|s_k(f;x)-f(x)|)\leqq\sum_{k=0}^{\infty}\Omega\left(k^r\sum_{m=1}^{\infty}\frac{1}{2^m}|s_k(f_{n_m};x)-f_{n_m}(x)|\right)\leqq$$

$$\leqq\sum_{k=0}^{\infty}\sum_{m=1}^{\infty}\frac{1}{2^m}\Omega(k^r|s_k(f_{n_m};x)-f_{n_m}(x)|)=$$

$$=\sum_{m=1}^{\infty}\frac{1}{2^m}\sum_{k=0}^{\infty}\Omega(k^r|s_k(f_{n_m};x)-f_{n_m}(x)|)\leqq\sum_{m=1}^{\infty}\frac{1}{2^m}=1.$$

From the remark made after the definition of the functions $f_{\{\varepsilon_k\},n}(x)$, it follows that $\tilde{f}^{(r)}$ exists; and using (2.19), (2.21) we obtain

$$\tilde{f}^{(r)}(0)-\tilde{f}^{(r)}\left(\frac{\pi}{2n_m}\right)\geqq\frac{1}{2^m}\left(\tilde{f}_{n_m}^{(r)}(0)-\tilde{f}_{n_m}^{(r)}\left(\frac{\pi}{2n_m}\right)\right)\geqq\frac{1}{2^m}\frac{1}{20}\bar{\Omega}^*\left(\frac{\pi}{2n_m}\right)\geqq\frac{2^m}{20}\omega\left(\frac{\pi}{2n_m}\right),$$

so

$$\omega(\tilde{f}^{(r)};\delta)\neq O(\omega(\delta)),$$

which proves our Remark.

References

[1] G. Freud, Über die Sättigungsklasse der starken Approximation durch Teilsummen der Fourierschen Reihe, *Acta Math. Acad. Sci. Hung.*, **20** (1969), 275—279.

[2] V. G. Krotov and L. Leindler, On the strong summability of Fourier series and the classes H^ω, *Acta Sci. Math.*, **40** (1978), 93—98.

[3] L. Leindler, Über die Approximation im starken Sinne, *Acta Math. Acad. Sci. Hung.*, **16** (1965), 255—262.

[4] L. Leindler, On strong approximation of Fourier series, *Approximation Theory* (Proceedings of a Conference in Poznań, 1972), 129—140.

[5] L. Leindler, On strong summability of Fourier series. II, *Acta Math. Acad. Sci. Hung.*, **20** (1969), 347—355.

[6] L. Leindler and E. M. Nikišin, Note on strong approximation by Fourier series, *Acta Math. Acad. Sci. Hung.*, **24** (1973), 223—227.

[7] K. I. Oskolkov, On strong summability of Fourier series and differentiability of functions, *Analysis Math.*, **2** (1976), 41—47.

[8] J. Szabados, On a problem of L. Leindler concerning strong approximation by Fourier series, *Analysis Math.*, **2** (1976), 155—161.

[9] A. F. Timan, *Theory of approximation of functions of a real variable* (Delhi, 1966).

[10] V. Totik, On the modulus of continuity in connection with a problem of J. Szabados concerning strong approximation, *Analysis Math.*, **4** (1978), 145—152.

BOLYAI INSTITUTE
UNIVERSITY SZEGED
6720 SZEGED, HUNGARY

Acta Sci. Math., 41 (1979), 253—257

Nilpotent torsion-free rings and triangles of types

M. C. WEBB

1. Introduction. This note modifies an idea of VINSONHALER and WICKLESS [7] concerning associative rings having torsion-free additive group. In [7] the necessary and sufficient conditions for a group to support only trivial rings given by REE and WISNER [4] are generalised in such a way that certain groups supporting only nilpotent rings are characterised. In fact more precise information can be obtained giving a bound on the nilstufe of a group. The nilstufe, a notion due to SZELE [6], $n(X)$, of a group X is the largest integer n such that there is an associative ring on X with a non-zero product of n elements. If no such largest integer exists then $n(X)=\infty$. Several authors [1], [3], [5], [8], [9] have obtained bounds for the nilstufe of a group in certain circumstances. The bound obtained here applies in quite general circumstances. In this note the basic tool of [7] is modified and then used to prove results based on two of the main theorems from [7], in one case giving a considerable generalisation.

From now on all groups are torsion-free abelian groups and all undefined concepts are standard from FUCHS [2]. In particular the product of a pair of types $\mathbf{t}_1, \mathbf{t}_2$ is written $\mathbf{t}_1\mathbf{t}_2$ not $\mathbf{t}_1+\mathbf{t}_2$ as in [7], and $T(X)=\{\mathbf{t}(x)\colon x\in X\}$ is the type-set of the group X.

2. **The results.** The following is a modification of a definition of [7].

Definition. A collection of types $\{\mathbf{t}_{ij}\colon 1\leqq j\leqq n-i+1,\ 1\leqq i\leqq n\}$ is a *triangle of base size* n if for any three integers i, j, k satisfying $1\leqq k\leqq n-i-j+1$, $1\leqq j\leqq n-i+1$, $1\leqq i\leqq n$,

$$\mathbf{t}_{i,j}\mathbf{t}_{k,(i+j)}\leqq \mathbf{t}_{(i+k),j}.$$

This is strongly related to the definition in [7] but, as will become apparent below, it is easy to handle. For example, the following based on Theorem 2.1 of [7] has a quite straightforward proof.

Received July 27, 1977, in revised form April 17, 1978.

Theorem 1. *For any torsion-free group X and any associative ring $(X, \circ)$, a non-zero product of n elements in the ring defines a triangle of base size n from the type-set of X.*

Proof. Suppose that for the elements $x_1, \ldots, x_n$ from X the product $x_1\circ\ldots\circ x_n$ is non-zero. Then denoting by $\langle x\rangle^*$ the pure subgroup generated by an element x of X, the following pure rank 1 subgroups of X are defined: for each pair of integers i, j satisfying $1\leqq j\leqq n-i+1$, $1\leqq i\leqq n$,

$$X_{i,j} = \langle x_j\circ\ldots\circ x_{j+i-1}\rangle^*.$$

As a consequence, for the integers i, j, k satisfying $1\leqq k\leqq n-i-j+1$, $1\leqq j\leqq \leqq n-i+1$, $1\leqq i\leqq n$,

$$\mathbf{t}(X_{(i+k),j}) = \mathbf{t}(x_j\circ\ldots\circ x_{j+i+k-1}) \geqq \mathbf{t}(x_j\circ\ldots\circ x_{j+i-1})\mathbf{t}(x_{j+i}\circ\ldots\circ x_{j+i+k-1}).$$

However,

$$\mathbf{t}(X_{i,j})\mathbf{t}(X_{k,(i+j)}) = \mathbf{t}(x_j\circ\ldots\circ x_{j+i-1})\mathbf{t}(x_{j+i}\circ\ldots\circ x_{j+i+k-1})$$

implying $\mathbf{t}(X_{(i+k),j})\geqq\mathbf{t}(X_{i,j})\mathbf{t}(X_{k,(i+j)})$. So if $\mathbf{t}_{i,j}=\mathbf{t}(X_{i,j})$ for $1\leqq j\leqq n-i+1$, $1\leqq i\leqq n$ the set of types so defined form a triangle of base size n.

Recalling the definition of nilstufe there is the following corollary.

Corollary. *For any torsion-free group X, its nilstufe $n(X)$ is bounded by the maximum base size of triangles that can be formed from $T(X)$.*

Note that in the Corollary there are no restraints on the group X. Other authors [1], [5], [8] have obtained bounds for $n(X)$ but for groups whose type-sets satisfy some form of chain condition. In most cases the Corollary gives a better bound on the nilstufe and an illustrative example is given following Theorem 2. A second important result of [7] that can now be more easily proved is

Theorem 2. *If for $X=\bigoplus_{i\in I} X_i$ where each X_i has rank 1 a triangle of base size n can be formed from the set $\{\mathbf{t}(X_i)\colon i\in I\}$ without using the same summand twice, then $n(X)\geqq n$.*

Proof. Suppose that the triangle is $\{\mathbf{t}_{i,j}\colon 1\leqq j\leqq n-i+1,\ 1\leqq i\leqq n\}$ where the $\mathbf{t}_{i,j}$ are types of distinct summands X_i. Then,

$$\mathbf{t}_{(i+k),j} \geqq \mathbf{t}_{i,j}\mathbf{t}_{k,(i+j)}$$

where $1\leqq k\leqq n-j-i+1$, $1\leqq j\leqq n-i+1$, $1\leqq i\leqq n$. The characteristic of an element x in X is denoted by $\chi(x)$ and so the above inequalities imply that elements $x_{i,j}$ from the corresponding rank one summands can be found such that

$$\chi(x_{(i+k),j}) \geqq \chi(x_{i,j})\chi(x_{k,(i+j)}).$$

If Y is the direct sum of the summands of X used in defining the types then a ring $(Y, *)$ can be defined using the following products. For integers i, j, k satisfying $1 \leqq k \leqq n-l+1$, $1 \leqq j \leqq n-i+1$, $1 \leqq i \leqq n$, $1 \leqq l \leqq n$.

$$x_{i,j} * x_{k,l} = \begin{cases} x_{(i+k),j} & \text{if } \ l = i+j \\ 0 & \text{if } \ \text{not}. \end{cases}$$

These products and the distributive law define a ring which can be shown to be associative.

Take three subscripts (k_1, k_2), (l_1, l_2), (m_1, m_2); then

$$x_{k_1,k_2} * x_{l_1,l_2} = \begin{cases} 0 & \text{if } \ l_2 \neq k_1+k_2 \\ x_{(k_1+l_1),k_2} & \text{if } \ l_2 = k_1+k_2 \end{cases}$$

and

$$x_{(k_1+l_1),k_2} * x_{m_1,m_2} = \begin{cases} 0 & \text{if } \ m_2 \neq k_1+l_1+k_\lambda \\ x_{(k_1+l_1+m_1),k_2} & \text{if } \ m_2 = k_1+l_1+k_\lambda. \end{cases}$$

Also,

$$x_{l_1,l_2} * x_{m_1,m_2} = \begin{cases} 0 & \text{if } \ m_2 \neq l_1+l_2 \\ x_{(l_1+m_1),l_2} & \text{if } \ m_2 = l_1+l_2 \end{cases}$$

and

$$x_{k_1,k_2} * x_{(l_1+m_1),l_2} = \begin{cases} 0 & \text{if } \ l_2 \neq k_1+k_2 \\ x_{(k_1+l_1+m_1),k_2} & \text{if } \ l_2 = k_1+k_2. \end{cases}$$

So that the products $(x_{k_1,k_2} * x_{l_1,l_2}) * x_{m_1,m_2}$, $x_{k_1,k_2} * (x_{l_1,l_2} * x_{m_1,m_2})$ are non-zero if and only if $l_2 = k_1+k_2$ and $m_2 = l_1+l_2$ in which case they both equal $x_{(k_1+l_1+m_1),k_2}$. Furthermore, $x_{1,1} * x_{1,2} * \dots * x_{1,n} = x_{n,1}$ is non-zero.

To define an associative ring on X merely take the ring direct sum of $(Y, *)$ and the trivial ring on the complement of Y in X and so $n(X) \geqq n$.

Finally an example is given to illustrate that the bounds obtained using Theorem 1 can be lower than those obtained using other available results.

Example. Begin by partitioning the set of all primes into two disjoint infinite subsets P_1 and P_2 where $P_2 = \{p_1, p_2, p_3, \dots\}$. Then define the following subgroups of the rationals.

$$A_1 = A_2 = gp\left\{\frac{1}{p} : p \in P_1 \cup P_2\right\}, \quad B = gp\left\{\frac{1}{p^n}, \frac{1}{q^2} : p \in P_1, \ n \in \mathbf{Z}^+, \ q \in P_2\right\},$$

and for each integer $i \geqq 1$,

$$C_i = gp\left\{\frac{1}{p}, \frac{1}{p_1^n}, \frac{1}{p_2^n}, \dots, \frac{1}{p_i^n} : p \in P_1, \ n \in \mathbf{Z}^+\right\}.$$

For each positive integer m let

$$X_m = A_1 \oplus A_2 \oplus B \oplus C_1 \oplus \dots \oplus C_m.$$

Then $T(X_m) = \{\mathbf{t}(A_1), \mathbf{t}(B), \mathbf{t}(C_1), \dots, \mathbf{t}(C_m)\}$ contains a chain of length $m+1$, but no chain of greater length.

A type $\mathbf{t}_1$ in $T(X)$ absorbs a type $\mathbf{t}_2$ in $T(X)$ if $\mathbf{t}_1\mathbf{t}_2 = \mathbf{t}_1$. It is clear that $\mathbf{t}(X_m)$ contains no absorbing types so that Proposition 1.2 of [5] implies $n(X_m) \leqq m+1$. However, a triangle of base length two can be formed from $T(X_m)$, namely $\mathbf{t}_{21} = \mathbf{t}(B)$, $\mathbf{t}_{1i} = t(A_i)$ $i = 1, 2$. So, by Theorem 2, $n(X_m) \geqq 2$. The following lemma will show that no larger triangles can be formed.

Lemma. *If X is a group such that $T(X)$ contains no absorbing types the apex of a triangle of base size n formed from $T(X)$ has a chain of length n descending from it in $T(X)$.*

Proof. The proof is by induction on n, the result being trivial for $n = 1$.

Suppose $n > 1$. Let $\mathbf{t}_{k,1}$ be the apex of a triangle of base size k, then the apex of any triangle of base size $(n-1)$ from $T(X)$ has a chain of length $(n-1)$ descending from it in $T(X)$. If the triangle is $\{\mathbf{t}_{i,j}\colon 1 \leqq j \leqq n-i+1,\ 1 \leqq i \leqq n\}$ then

$$\mathbf{t}_{n,1} \geqq \mathbf{t}_{(n-1),1}\mathbf{t}_{1,n}.$$

Now $\mathbf{t}_{(n-1),1}$ is the apex of a triangle of base size $(n-1)$ from $T(X)$ so is the maximal type of a chain of length $(n-1)$ in $T(X)$. Suppose $\mathbf{t}_{n,1} = \mathbf{t}_{(n-1),1}$ then $\mathbf{t}_{(n-1),1}\mathbf{t}_{1,n} = \mathbf{t}_{(n-1),1}$ and $\mathbf{t}_{(n-1),1}$ is an absorbing type in $T(X)$. Thus $\mathbf{t}_{n,1} > \mathbf{t}_{(n-1),1}$ and so has a chain of length n descending from it in $T(X)$.

Returning to the example it should be noted that $\mathbf{t}(B)$ has only chains of at most length 2 descending from it in $T(X_n)$. Furthermore by considering the summands of X_m it is clear that in any ring on X_m, $X_m^2 \subseteqq B$ since $P_1 \cap P_2$ is empty. Since any product of more than two elements in any ring on X is in B, $\mathbf{t}(B)$ is the apex of all triangles formed by applying Theorem 1. Hence $n(X_m) \leqq 2$ which with the earlier work gives $n(X_m) = 2$ for all $m \geqq 1$.

In conclusion it is noted that this technique fails to work for non-associative rings as the bracketing of a product may be quite arbitrary. Essential to the above is that for a product in an associative ring $\mathbf{t}((ab)c) = \mathbf{t}(a(bc))$.

References

[1] S. FEIGELSTOCK, The nilstufe of rank two torsion-free groups, *Acta Sci. Math.*, **36** (1974), 29—32.
[2] L. FUCHS, *Infinite Abelian Groups*, Vol. II, Academic Press (New York, 1973).
[3] B. J. GARDNER. Rings on completely decomposable torsion-free abelian groups, *Comment. Math. Univ. Carolinae*, **15** (1974), 381—392.
[4] R. REE and R. J. WISNER, A note on torsion-free nil groups, *Proc. Amer. Math. Soc.*, **7** (1956), 6—8.
[5] A. E. STRATTON and M. C. WEBB, Type-sets and nilpotent multiplications, *Acta Sci. Math.*, **41** (1979), 209—213.
[6] T. SZELE, Gruppentheoretische Beziehungen bei gewissen Ringkonstruktionen, *Math. Z.*, **54** (1951), 168—180.
[7] C. VINSONHALER and W. J. WICKLESS, Completely decomposable groups which admit only nilpotent multiplications, *Pac. J. Math.*, **53** (1974), 273—280.
[8] M. C. WEBB, A bound for the nilstufe of a group, *Acta Sci. Math.*, **39** (1977), 185—188.
[9] W. J. WICKLESS, Abelian groups which admit only nilpotent multiplications, *Pac. J. Math.*, **40** (1972), 251—259.

UNIVERSITY OF GHANA
P.O. BOX 62
LEGON, GHANA

Acta Sci. Math., 41 (1979), 259—266

Hyperinvariant subspaces of weak contractions

PEI YUAN WU

Introduction

The aim of this paper is to study Hyperlat T, the hyperinvariant subspace lattice, of a completely non-unitary (c.n.u.) weak contraction T with finite defect indices. The work here is a continuation of the investigations of Hyperlat T which we made in [14] and [15]. There we only considered c.n.u. C_{11} contractions with finite defect indices. Now we shall generalize the results of [14] and [15]. Among other things, we shall show that for the contractions considered, (i) if T_1 is quasi-similar to T_2, then Hyperlat T_1 is (lattice) isomorphic to Hyperlat T_2 (Corollary 3.4) and (ii) Hyperlat T is (lattice) generated by subspaces of the forms $\overline{\operatorname{ran} S}$ and $\ker V$ where S, V are operators in $\{T\}''$, the double commutant of T (Theorem 3.8). We also give necessary and sufficient conditions, in terms of the characteristic function and the Jordan model of T, that Lat T, the invariant subspace lattice of T, be equal to Hyperlat T.

Preliminaries and results

We follow the notations and terminologies used in [14] and [15]. Only the concepts concerning weak contractions will be reviewed here.

A contraction T is called a *weak contraction* if (i) its spectrum $\sigma(T)$ does not fill the open unit disc, and (ii) $I-T^*T$ is of finite trace. Examples of weak contractions are $C_0(N)$ contractions and c.n.u. C_{11} contractions with finite defect indices. The characteristic function Θ_T of every weak contraction T admits a *scalar multiple*, that is, there exist a contractive analytic function Ω and a scalar valued analytic function $\delta\not\equiv 0$ such that $\Omega\Theta_T=\Theta_T\Omega=\delta$. For a c.n.u. weak contraction

Received October 14, 1977.

This research was partially supported by the National Science Council of Taiwan.

T on H we can consider its C_0-C_{11} decomposition. Let $H_0, H_1 \subseteq H$ be the invariant subspaces for T such that $T_0 \equiv T|H_0$ and $T_1 \equiv T|H_1$ are the C_0 and C_{11} parts of T, respectively. Indeed, T_0 and T_1 are equal to those appearing in the triangulations

$$T = \begin{bmatrix} T_0 & X \\ 0 & T_1' \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} T_1 & Y \\ 0 & T_0' \end{bmatrix}$$

on $H = H_0 \oplus H_0^{\perp}$ and $H = H_1 \oplus H_1^{\perp}$ corresponding to the $*$-canonical factorization $\Theta_T = \Theta_{*e}\Theta_{*i}$ and the canonical factorization $\Theta_T = \Theta_i \Theta_e$, respectively. H_0 and H_1 are even hyperinvariant for T and satisfy $H_0 \vee H_1 = H$ and $H_0 \cap H_1 = \{0\}$. For the details the readers are referred to [4], Chap. VIII.

* *

It was shown in [4], p. 334 that $H_0 = \ker m(T)$ and $H_1 = \overline{\operatorname{ran} m(T)}$, where m is the minimal function of T_0. Note that $m(T) \in \{T\}''$. Now we have the following supplementary result.

Theorem 1. *If T is a c.n.u. weak contraction on H and H_0, H_1 are subspaces of H such that $T_0 = T|H_0$ and $T_1 = T|H_1$ are the C_0 and C_{11} parts of T, respectively, then there exists an operator S in $\{T\}''$ such that $H_0 = \overline{\operatorname{ran} S}$ and $H_1 = \ker S$.*

Proof. We consider T being defined on $H \equiv [H^2_{\mathfrak{D}} \oplus \overline{\Delta L^2_{\mathfrak{D}}}] \ominus \{\Theta_T w \oplus \Delta w : w \in H^2_{\mathfrak{D}}\}$ by $T(f \oplus g) = P(e^{it} f \oplus e^{it} g)$ for $f \oplus g \in H$, where Θ_T is the characteristic function of T, $\Delta(t) = (I_{\mathfrak{D}} - \Theta_T(t)^* \Theta_T(t))^{1/2}$ and P denotes the (orthogonal) projection onto H. Since Θ_T admits a scalar multiple, the same is true for its outer factor Θ_e and inner factor Θ_i (cf. [4], p. 217). Let $\delta_1 \not\equiv 0$ and $\delta_2 \not\equiv 0$ be their respective scalar multiples, and let Ω_1 and Ω_2 be contractive analytic functions such that $\Omega_1 \Theta_e = \Theta_e \Omega_1 = \delta_1 I_{\mathfrak{D}}$ and $\Omega_2 \Theta_i = \Theta_i \Omega_2 = \delta_2 I_{\mathfrak{D}}$. We may assume that δ_1 is outer and δ_2 is inner (cf. [4], p. 217). Let $\delta = \delta_1 \delta_2$ and $\Omega = \Omega_1 \Omega_2$. Then $\Omega \Theta_T = \Theta_T \Omega = \delta I_{\mathfrak{D}}$. Consider the operator $S = P \begin{bmatrix} \delta_1 & 0 \\ \bar{\delta}_2 \Delta \Omega & 0 \end{bmatrix}$. We prove $H_1 = \ker S$ and $H_0 = \overline{\operatorname{ran} S}$ in the following steps. In each step the first statement is proved.

(1) $S \in \{T\}''$. Let $V = P \begin{bmatrix} A & 0 \\ B & C \end{bmatrix}$ be an operator in $\{T\}'$, where A is a bounded analytic function while B and C are bounded measurable functions satisfying the conditions $A\Theta_T = \Theta_T A_0$ and $B\Theta_T + C\Delta = \Delta A_0$ a.e., where A_0 is another bounded analytic function (cf. [5]). An easy calculation shows that

$$SV = P \begin{bmatrix} \delta_1 A & 0 \\ \bar{\delta}_2 \Delta \Omega A & 0 \end{bmatrix} \quad \text{and} \quad VS = P \begin{bmatrix} A\delta_1 & 0 \\ B\delta_1 + C\bar{\delta}_2 \Delta \Omega & 0 \end{bmatrix}.$$

We have $\bar{\delta}_2 \Delta\Omega A\delta = \bar{\delta}_2 \Delta\Omega A\Theta_T \Omega = \bar{\delta}_2 \Delta\Omega\Theta_T A_0 \Omega = \bar{\delta}_2 \Delta\delta A_0 \Omega = \delta_1 \Delta A_0 \Omega = \delta_1(B\Theta_T + C\Delta)\Omega = B\delta_1\delta + C\bar{\delta}_2\Delta\Omega\delta = (B\delta_1 + C\bar{\delta}_2\Delta\Omega)\delta$. Since $\delta \not\equiv 0$, we conclude that $\bar{\delta}_2\Delta\Omega A = B\delta_1 + C\bar{\delta}_2\Delta\Omega$. Hence $SV = VS$ and we have $S \in \{T\}''$.

(2) $H_1 \subseteq \ker S$. It was shown in [6] that $H_1 = \{f \oplus g \in H\colon f \in \Theta_i H^2_{\mathfrak{D}}\}$. For $\Theta_i u \oplus g \in H_1$, $S(\Theta_i u \oplus g) = P(\delta_1 \Theta_i u \oplus \bar{\delta}_2 \Delta\Omega\Theta_i u) = P(\Theta_i \Theta_e \Omega_1 u \oplus \bar{\delta}_2 \Delta\Omega_1 \Omega_2 \Theta_i u) = P(\Theta\Omega_1 u \oplus \Delta\Omega_1 u) = 0$, which shows that $H_1 \subseteq \ker S$.

(3) $\ker S \subseteq H_1$. For $f \oplus g \in \ker S$, $S(f \oplus g) = P(\delta_1 f \oplus \bar{\delta}_2 \Delta\Omega f) = (\delta_1 f - \Theta_T w) \oplus (\bar{\delta}_2 \Delta\Omega f - \Delta w) = 0$ for some $w \in H^2_{\mathfrak{D}}$. Hence $\delta_1 f = \Theta_T w$. Note that $\frac{1}{\delta_1}\Theta_e w = \Theta_i^* f$ is an element of $L^2_{\mathfrak{D}}$. However $\frac{1}{\delta_1}\Theta_e w$ is also analytic in the open unit disc, and therefore belongs to $H^2_{\mathfrak{D}}$. We conclude that $f = \Theta_i w'$, where $w' = \frac{1}{\delta_1}\Theta_e w \in H^2_{\mathfrak{D}}$. This shows that $f \oplus g \in H_1$, and hence $\ker S \subseteq H_1$.

(2) and (3) imply that $H_1 = \ker S$. Next we prove that $H_0 = \overline{SH}$.

(4) $\overline{SH} \subseteq H_0$. It was shown in [6] that $H_0 = \{f \oplus g \in H\colon \Theta_T g = \Delta_* f\}$, where $\Delta_* = (I_{\mathfrak{D}} - \Theta_T \Theta_T^*)^{1/2}$. For any $f \oplus g \in H$, $S(f \oplus g) = (\delta_1 f - \Theta_T w) \oplus (\bar{\delta}_2 \Delta\Omega f - \Delta w)$ for some $w \in H^2_{\mathfrak{D}}$. Note that $(I_{\mathfrak{D}} - \Theta_T^* \Theta_T)\Omega = \Omega - \Theta_T^* \delta = \Omega(I_{\mathfrak{D}} - \Theta_T \Theta_T^*)$, whence $\Delta\Omega = \Omega\Delta_*$. Similarly, $\Theta_T \Delta = \Delta_* \Theta_T$. Thus $\Theta_T(\bar{\delta}_2 \Delta\Omega f - \Delta w) = \bar{\delta}_2 \Delta_* \Theta_T \Omega f - \Delta_* \Theta_T w = \bar{\delta}_2 \Delta_* \delta f - \Delta_* \Theta_T w = \Delta_*(\delta_1 f - \Theta_T w)$, which shows that $S(f \oplus g) \in H_0$, and hence $\overline{SH} \subseteq H_0$.

(5) $S|H_0 = \delta_1(T_0)$. Since H_0 is the invariant subspace corresponding to $\Theta_T = \Theta_{*e}\Theta_{*i}$ and Θ_{*i} is inner from both sides, $H_0 = \{\Theta_{*e} u \oplus Z^{-1}(\Delta_2 u)\colon u \in H^2_{\mathfrak{D}}\} \ominus \{\Theta_T w \oplus \Delta w\colon w \in H^2_{\mathfrak{D}}\}$, where $\Delta_2 = (I_{\mathfrak{D}} - \Theta_{*e} * \Theta_{*e})^{1/2}$ and Z is the unitary operator from $\overline{\Delta L^2_{\mathfrak{D}}}$ onto $\overline{\Delta_2 L^2_{\mathfrak{D}}}$ such that $Z(\Delta v) = \Delta_2 \Theta_{*i} v$ for $v \in L^2_{\mathfrak{D}}$ (cf. [4], p. 288). For any $\Theta_{*e} u \oplus Z^{-1} (\Delta_2 u) \in H_0$, we have $S(\Theta_{*e} u \oplus Z^{-1}(\Delta_2 u)) = (\delta_1 \Theta_{*e} u - \Theta_T w) \oplus (\bar{\delta}_2 \Delta\Omega\Theta_{*e} u - \Delta w)$ for some $w \in H^2_{\mathfrak{D}}$. Since Θ_T, along with Θ_{*e} and Θ_{*i}, admits a scalar multiple, $\Theta_T(t)^{-1} = \Theta_{*i}(t)^{-1}\Theta_{*e}(t)^{-1}$ exists for almost all t. Therefore, $\Omega = \delta\Theta_T^{-1} = \delta\Theta_{*i}^{-1}\Theta_{*e}^{-1}$ a.e. We have $Z(\bar{\delta}_2 \Delta\Omega\Theta_{*e} u) = \Delta_2 \Theta_{*i} \bar{\delta}_2 \Omega\Theta_{*e} u = \Delta_2 \Theta_{*i} \bar{\delta}_2 \delta\Theta_{*i}^{-1}\Theta_{*e}^{-1}\Theta_{*e} u = \delta_1 \Delta_2 u$, and it follows that $S(\Theta_{*e} u \oplus Z^{-1}(\Delta_2 u)) = (\delta_1 \Theta_{*e} u - \Theta_T w) \oplus (\delta_1 Z^{-1}(\Delta_2 u) - \Delta w)$. This shows that $S|H_0 = \delta_1(T_0)$.

(6) $\overline{SH} = H_0$. Since δ_1 is outer, $\delta_1(T_0)$ is a quasi-affinity (cf. [4], p. 118). Hence $\overline{\delta_1(T_0) H_0} = H_0$. By (4) and (5), this implies that $\overline{SH} = H_0$.

The next lemma is needed in the proof of Theorem 3.3.

Lemma 2. *Let T be a c.n.u. weak contraction on H and let H_0, H_1 be subspaces of H such that $T_0 = T|H_0$ and $T_1 = T|H_1$ are the C_0 and C_{11} parts of T, respectively. If H_0', $H_0' \subseteq H$ are invariants subspaces for T such that $H_0' \vee H_1' = H$ and $T|H_0' \in C_0$, $T|H_1' \in C_{11}$, then $H_0 = H_0'$ and $H_1 = H_1'$.*

Proof. The maximality property of H_0 and H_1 implies that $H_0' \subseteq H_0$ and $H_1' \subseteq H_1$ (cf. [4], p. 331). Now we show that $H_0 \subseteq H_0'$. Since $H_0 = \overline{\operatorname{ran} S}$ where S is the operator defined in Theorem 1, for any $h \in H_0$ and $\varepsilon > 0$ there exists some k in H such that $\|h - Sk\| < \varepsilon$. The hypothesis $H = H_0' \vee H_1'$ implies that $\|k - k_0 - k_1\| < \varepsilon$ holds for some $k_0 \in H_0'$ and $k_1 \in H_1'$. Hence $\|Sk - Sk_0 - Sk_1\| = \|Sk - Sk_0\| < \|S\|\varepsilon$, and it follows that $\|h - Sk_0\| < (1 + \|S\|)\varepsilon$. Since $Sk_0 = \delta_1(T_0)k_0 = \delta_1(T)k_0 \in H_0'$ and ε is arbitrary, we conclude that $h \in H_0'$ and hence $H_0' = H_0$. $H_1' = H_1$ can be proved in a similar fashion by noting that $H_1 = \overline{\operatorname{ran} m(T)}$ and $H_0 = \ker m(T)$, where m denotes the minimal function of T_0.

Now we have the following main theorem.

Theorem 3. *Let T be a c.n.u. weak contraction on H and let H_0, H_1 be subspaces of H such that $T_0 = T|H_0$ and $T_1 = T|H_1$ are the C_0 and C_{11} parts of T, respectively. Then the following lattices are isomorphic:*

$$\operatorname{Hyperlat} T, \quad \operatorname{Hyperlat} T_0 \oplus \operatorname{Hyperlat} T_1, \quad \textit{and} \quad \operatorname{Hyperlat}(T_0 \oplus T_1).$$

Proof. Since T_0 and T_1 are of class C_{00} and of class C_{11}, respectively, Hyperlat $T_0 \oplus \operatorname{Hyperlat} T_1 \cong \operatorname{Hyperlat}(T_0 \oplus T_1)$ follows from Prop. 3 and Lemma 4 of [2].

Next we show that a subspace $K \subseteq H$ is hyperinvariant for T if and only if $K = K_0 \vee K_1$ where $K_0 \subseteq H_0$ and $K_1 \subseteq H_1$ are hyperinvariant for T_0 and T_1, respectively. To prove one direction, let $K \subseteq H$ be hyperinvariant for T and let $K_0 = K \cap H_0$, $K_1 = K \cap H_1$. Note that $\sigma(T|K) \subseteq \sigma(T)$ [1] and hence $T|K$ is also a weak contraction. Thus K_0 and K_1 are subspaces of K on which the C_0 and C_{11} parts of $T|K$ act (cf. [4], p. 332). We have $K = K_0 \vee K_1$. Now we show the hyperinvariance of K_0 and K_1. Note that $H_0 = \overline{SH}$, where S is the operator defined in Theorem 1. For any $S_0 \in \{T_0\}'$, consider the operator $S_0 S$ on H. It is easily seen that $S_0 S \in \{T\}'$. Since $K_0 = K \cap H_0$ is hyperinvariant for T, $S_0 S K_0 \subseteq K_0$. As proved in Theorem 1, $S|H_0 = \delta_1(T_0)$ for some outer function δ_1. Thus $\overline{SK_0} = \overline{\delta_1(T|K_0)K_0} = K_0$. It follows that $S_0 K_0 \subseteq K_0$ and hence K_0 is hyperinvariant for T_0. That K_1 is hyperinvariant for T_1 can be proved similarly by noting that $H_1 = \overline{m(T)H}$ where m is the minimal function of T_0 and $m(T|K_1)$, being an analytic function of a c.n.u. C_{11} contraction, is a quasi-affinity (cf. [4], p. 123).

To prove the converse, let $S \in \{T\}'$ and $S_0 = S|H_0$, $S_1 = S|H_1$. It is obvious that $S_0 \in \{T_0\}'$ and $S_1 \in \{T_1\}'$. If $K_0 \subseteq H_0$ and $K_1 \subseteq H_1$ are hyperinvariant for T_0 and T_1, respectively, then $S_0 K_0 \subseteq K_0$ and $S_1 K_1 \subseteq K_1$. Hence $S(K_0 \vee K_1) \subseteq K_0 \vee K_1$, which shows that $K_0 \vee K_1$ is hyperinvariant for T and proves our assertion.

That K_0 and K_1 are uniquely determined by K follows from Lemma 2, and it is easily seen that Hyperlat $T \cong$ Hyperlat $T_0 \oplus$ Hyperlat T_1.

In [11] a specific description of the elements in Hyperlat T for a special class of c.n.u. weak contractions is given.

Corollary 4. *Let T_1, T_2 be c.n.u. weak contractions with finite defect indices. If T_1 is quasi-similar to T_2, then* Hyperlat T_1 *is isomorphic to* Hyperlat T_2.

Proof. Let T_{10}, T_{20} be the C_0 parts of T_1, T_2 and T_{11}, T_{21} be their C_{11} parts, respectively. If T_1 is quasi-similar to T_2, then T_{10}, T_{11} are quasi-similar to T_{20}, T_{21}, respectively (cf. [10]). Since T_1, T_2 have finite defect indices, T_{10}, T_{20} are of class $C_0(N)$ and the defect indices of T_{11}, T_{21} are also finite. Thus Hyperlat $T_{10} \cong$ Hyperlat T_{20} and Hyperlat $T_{11} \cong$ Hyperlat T_{21} (cf. [7] and [14], resp.). Now Hyperlat $T_1 \cong$ $\cong$ Hyperlat T_2 follows from Theorem 3.

Recall that a c.n.u. weak contraction T is *multiplicity-free* if T admits a cyclic vector and that T is multiplicity-free if and only if its C_0 part and C_{11} part are (cf. [12]).

Corollary 5. *Let T be a c.n.u. multiplicity-free weak contraction on H with defect indices $n < +\infty$. Let $K \subseteq H$ be an invariant subspace for T with the corresponding regular factorization $\Theta_T = \Theta_2 \Theta_1$. Then the following are equivalent to each other:*

(1) $K \in$ Hyperlat T;

(2) *the intermediate space of $\Theta_T = \Theta_2 \Theta_1$ is of dimension n.*

Proof. (1)$\Rightarrow$(2). If $K \in$ Hyperlat T, then, as proved before, $T|K$ is a weak contraction. Hence its characteristic function admits a scalar multiple, which implies that the intermediate space of $\Theta_T = \Theta_2 \Theta_1$ is of dimension n.

(2)$\Rightarrow$(1). The hypothesis implies that $T|K$ has equal defect indices. It is easily seen that a c.n.u. contraction S with finite equal defect indices is a weak contraction if and only if $\det \Theta_S \not\equiv 0$. Since $\det \Theta_T \not\equiv 0$ implies that $\det \Theta_1 \not\equiv 0$, it follows that $T|K$ is a weak contraction. Let K_0, K_1 be subspaces of K on which the C_0 and C_{11} parts of $T|K$ act. We have $K = K_0 \vee K_1$. It follows from the proof of Theorem 3 that we have only to show that K_0 and K_1 are hyperinvariant for $T_0 = T|H_0$ and $T_1 = T|H_1$, the C_0 and C_{11} parts of T, respectively. Since $K_0 \subseteq H_0$ is invariant for the multiplicity-free $C_0(N)$ contraction T_0, it is hyperinvariant for it (cf. [8], Corollary 4.4). On the other hand, T_1 is a multiplicity-free C_{11} contraction on H_1 with finite defect indices and $K_1 \subseteq H_1$ is such that $T_1|K_1 \in C_{11}$. It follows easily from Theorem 1 of [14] that K_1 is hyperinvariant for T_1, completing the proof.

The next corollary gives necessary and sufficient conditions that Lat T be equal to Hyperlat T for the operators we considered.

Corollary 6. *Let T be a c.n.u. weak contraction on H with defect indices $n<+\infty$. Let $T_0=T|H_0$ and $T_1=T|H_1$ be its C_0 and C_{11} parts, respectively, and let Θ_e be the outer factor of the characteristic function Θ_T of T. Then the following conditions are equivalent:*

(1) *Lat* $T=$Hyperlat T;

(2) *Lat* $T_0=$Hyperlat T_0 *and Lat* $T_1=$Hyperlat T_1;

(3) *T_0 and T_1 are multiplicity-free and $\Theta_e(t)$ is isometric on a set of positive Lebesgue measure;*

(4) *T is multiplicity-free and $\Theta_T(t)$ is isometric on a set of positive Lebesgue measure.*

Proof. (1)$\Rightarrow$(2). We only show that Lat $T_0=$Hyperlat T_0; Lat $T_1=$Hyperlat T_1 can be proved similarly. To this end, let $K_0\subseteqq H_0$ be an invariant subspace for T_0. It is obvious that $K_0\in$Lat $T=$Hyperlat T. Let S be the operator defined in Theorem 1. Then $H_0=\overline{SH}$ and $S|H_0=\delta_1(T_0)$ for some outer function δ_1. For any $S_0\in\{T_0\}'$, S_0S is an operator in $\{T\}'$. Hence $\overline{S_0SK_0}=\overline{S_0\delta_1(T|K_0)K_0}=\overline{S_0K_0}\subseteqq K_0$, which shows that $K_0\in$Hyperlat T_0 and proves our assertion.

(2)$\Rightarrow$(3). This follows from Corollary 4.4 of [8] and Theorem 4.3 of [15].

(3)$\Rightarrow$(4). This follows from the remark before Corollary 5 and the fact that $\Theta_T(t)$ is isometric if and only if $\Theta_e(t)$ is.

(4)$\Rightarrow$(1). Let $K\in$Lat T with the corresponding regular factorization $\Theta_T=\Theta_2\Theta_1$. In light of Corollary 5 it suffices to show that the intermediate space of $\Theta_T=\Theta_2\Theta_1$ is of dimension n. Note that rank $\Delta(t)=$rank $\Delta_1(t)+$rank $\Delta_2(t)$ a.e., where $\Delta(t)=(I-\Theta_T(t)^*\Theta_T(t))^{1/2}$ and $\Delta_j(t)=(I-\Theta_j(t)^*\Theta_j(t))^{1/2}$, $j=1,2$. The hypothesis implies that $\Delta(t)=0$ on a set of positive Lebesgue measure, say α. It follows that $\Delta_1(t)=\Delta_2(t)=0$ on α, and hence $\Theta_1(t)$ and $\Theta_2(t)$ are isometric for t in α. Therefore, the intermediate space of $\Theta_T=\Theta_2\Theta_1$ is of dimension n, as asserted.

We remark that the preceding corollary generalizes part of the main result in [9].

Corollary 7. *Let T be a c.n.u. multiplicity-free weak contraction with finite defect indices. If $K_1, K_2\in$Hyperlat T and $T|K_1$ is quasi-similar to $T|K_2$, then $K_1=K_2$.*

Proof. Since $K_1, K_2\in$Hyperlat T, $T|K_1$, $T|K_2$ are weak contractions. Considering the C_0 and C_{11} parts of $T|K_1$ and $T|K_2$ and using the corresponding results for multiplicity-free $C_0(N)$ contractions and C_{11} contractions, we can deduce that $K_1=K_2$ (cf. [3], Theorem 2 and [14], Corollary 3). We leave the details to the interested readers.

The next theorem, being another application of Theorem 3, is interesting in itself.

Theorem 8. *Let T be a c.n.u. weak contraction on H with finite defect indices. Then* Hyperlat T *is (lattice) generated by subspaces of the forms* $\overline{\operatorname{ran} S}$ *and* $\ker V$, *where* $S, V \in \{T\}''$.

Proof. Let $T_0 = T|H_0$ and $T_1 = T|H_1$ be the C_0 and C_{11} parts of T, respectively, and let $K \in$ Hyperlat T. Since $T|K$ is a c.n.u. weak contraction, we may consider its C_0 part $T|K_0$ and C_{11} part $T|K_1$. By Theorem 1, $H_0 = \overline{SH}$ for some $S \in \{T\}''$. Since $K_0 \subseteq H_0$ is hyperinvariant for the $C_0(N)$ contraction T_0 (by Theorem 3), it follows from [13] that $K_0 = \bigvee_{i=1}^{n} [\ker \psi_i(T_0) \cap \overline{\xi_i(T_0) H_0}] = \bigvee_{i=1}^{n} [\ker \psi_i(T_0) \cap \overline{\xi_i(T) SH}]$, where ψ_i, ξ_i are inner functions, $i = 1, \dots, n$. On the other hand, since $K_1 \subseteq H_1$ is hyperinvariant for T_1 (by Theorem 3 again), Theorem 3.6 of [15] implies that $K_1 = \overline{VH_1}$ for some $V \in \{T_1\}''$. Hence $K_1 = \overline{Vm(T)H}$, where m denotes the minimal function of T_0. We claim that $K = \bigvee_{i=1}^{n} [\ker \psi_i(T) \cap \overline{\xi_i(T) SH}] \vee \overline{Vm(T)H}$. Indeed, this follows from $K = K_0 \vee K_1$ and the fact that $\ker \psi(T_0) = \ker \psi(T)$ for any $\psi \in H^\infty$. Since it is easily seen that $\psi_i(T)$, $\xi_i(T) S \in \{T\}''$ for all i and $Vm(T) \in \{T\}''$, the proof is complete.

Corollary 9. *Let T be a c.n.u. multiplicity-free weak contraction on H with finite defect indices and let K be a subspace of H. Then the following are equivalent:*

(1) $K \in$ Hyperlat T;

(2) $K = \overline{\operatorname{ran} S}$ *for some* $S \in \{T\}''$;

(3) $K = \ker V$ *for some* $V \in \{T\}''$.

Proof. The equivalence of (2) and (3) is easily established by considering T^* and $K^\perp$. (2)⇒(1) is trivial.

(1)⇒(2) is proved by following the same line of arguments in the proof of Theorem 8 and noting that any hyperinvariant subspace for a multiplicity-free $C_0(N)$ contraction T is of the form $\overline{\operatorname{ran} \xi(T)}$ for some inner function ξ.

References

[1] C. Apostol, Spectral decompositions and functional calculus, *Rev. Roum. Math. Pures et Appl.*, **13** (1968), 1481—1528.

[2] J. B. Conway and P. Y. Wu, The splitting of $\mathscr{A}(T_1 \oplus T_2)$ and related questions, *Indiana Univ. Math. J.*, **26** (1977), 41—56.

[3] B. Sz.-Nagy and C. Foiaş, Opérateurs sans multiplicité, *Acta Sci. Math.*, **30** (1969), 1—18.

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

[5] B. SZ.-NAGY and C. FOIAŞ, On the structure of intertwining operators, *Acta Sci. Math.*, **35** (1973), 225—254.

[6] R. I. TEODORESCU, Sur les décompositions directes C_0—C_{11} des contractions, *Acta Sci. Math.*, **36** (1974), 181—187.

[7] M. UCHIYAMA, Hyperinvariant subspaces of operators of class $C_0(N)$, *Acta Sci. Math.*, **39** (1977), 179—184.

[8] P. Y. WU, Commutants of $C_0(N)$ contractions, *Acta Sci. Math.*, **38** (1976), 193—202.

[9] P. Y. WU, On contractions satisfying $\mathrm{Alg}\, T=\{T\}'$, *Proc. Amer. Math. Soc.*, **67** (1977), 260—264.

[10] P. Y. WU, Quasi-similarity of weak contractions, *Proc. Amer. Math. Soc.*, **69** (1978), 277—282.

[11] P. Y. WU, Hyperinvariant subspaces of the direct sum of certain contractions,s *Indiana Univ. Math. J.*, **27** (1978), 267—274.

[12] P. Y. WU, Jordan model for weak contractions, *Acta Sci. Math.*, **40** (1978), 189—196.

[13] P. Y. WU, The hyperinvariant subspace lattice of contractions of class $C_{\cdot 0}$, *Proc. Amer. Math. Soc.*, **72** (1978), 527—530.

[14] P. Y. WU, Hyperinvariant subspaces of C_{11} contractions, *Proc. Amer. Math. Soc.*, to appear.

[15] P. Y. WU, Hyperinvariant subspaces of C_{11} contractions. II, *Indiana Univ. Math. J.*, **27** (1978), 805—812.

DEPARTMENT OF APPLIED MATHEMATICS
NATIONAL CHIAO TUNG UNIVERSITY
HSINCHU, TAIWAN

Bibliographie

Arthur L. Besse, Manifolds all of whose geodesics are closed (Ergebnisse der Mathematik und ihrer Grenzgebiete, 93), IX+262 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1978.

It had been a long-standing open problem, originating from W. Blaschke, whether an oriented "Wiedersehensfläche" is necessarily isometric to a sphere. (A Riemannian 2-manifold M is a "Wiedersehensfläche" if for every $x \in M$ there exists a $y \in M$ such that each geodesic starting from x passes through y. The question was solved affirmatively by L. Green in 1963, but many interesting problems, closely related to the previous one, were left open. The aim of this book is to give a detailed introduction and a comprehensive survey of the results and open questions in this topic.

In the first two chapters the author gives a short introduction to Riemannian geometry, geodesic flows and manifolds of geodesics. Chapter 3 presents a complete treatment of the geometric properties of compact symmetric Riemannian spaces of rank one, which are the basic examples of manifolds all of whose geodesics are closed. Chapter 4 deals with the geometry of Zoll and Tannery surfaces, which are further examples of such maifolds. Chapter 5 is devoted to the proof of Blaschke's "Wiedersehensfläche" conjecture and related questions. In Chapter 6 the geometry of geodesics in a harmonic manifold is studied. In Chapters 7—8 several results concerning topological invariants and the spectrum of the Laplace operator on a maifold all of whose geodesics are closed, are proved. The book closes with 5 Appendices written by D. V. A. Epstein, J. B. Bourguignon, L. B. Bergery, M. Berger, and J. L. Kazdan.

The book is well organized. It contains a detailed description of various Riemannian manifolds which are very close to the standard non-euclidean spaces from the geometric view-point and which have not been considered in earlier monographs. The book is highly recommended to anyone interested in the geometry of Riemannian manifolds.

P. T. Nagy (Szeged)

Eugen Blum—Werner Oettli, Mathematische Optimierung. Grundlagen und Verfahren (Ökonometrie und Unternehmensforschung, Bd. 20), XIV+413 Seiten, Berlin—Heidelberg—New York, Springer-Verlag, 1975.

Der Band gibt einen sehr guten Überblick über dieses in dem letzten Jahrzehnt sich rapid entwickelnde Wissensgebiet. Einen bedeutenden Teil seines Umfangs nimmt die Beschreibung der Methoden und Verfahren der nichtlinearen Optimierung ein, wenn auch das zweite Kapitel die Zusammenfassung der grundlegenden Ergebnisse der linearen Programmierung enthält. Die Verfahren der nichtlinearen Programmierung, die auf gleichem Grundprinzip liegen, bilden je eine Verfahren-Familie. Jedes Kapitel des Buches stellt, nach der Mitteilung der diesbezüglichen theoretischen Kenntnisse, je eine Familie durch ein oder zwei konkrete Verfahren vor.

Die behandelten Verfahrenstypen (gleichzeitig Kapitelaufschriften) sind wie folgt: 5. Optimierung ohne Restriktionen; 6. Projektions- und Kontraktionsverfahren; 7. Einzelschrittverfahren; 8.

Schnittverfahren; 9. Dekompositionsverfahren; 10. Strafkostverfahren; 11. Verfahren der zulässigen Richtungen; 12. Das Verfahren der projizierten Gradienten.

Die Verfasser hatten besonderen Wert auf die theoretische Begründung der behandelten Verfahren gelegt und darauf, dass die Verfahren, die im Buch vorkommen, auch für Computers gut verwendbar sind. — Es gibt zwei Kapitel die ausschliesslich der theoretischen Begründung dienen: das eine Kapitel beschäftigt sich mit den Optimalitätsbedingungen, das andere mit der Dualitätstheorie.

Diese Monographie, die sowohl umfassende theoretische Kenntnisse als auch praktische Verfahren darbietet, kann jenen zum Studieren empfohlen werden, die sich im Themenkreis mathematische Optimierung weitläufige Kenntnisse erwerben möchten. Dem Textteil ist eine umfangreiche Bibliographie der nichtlinearen Optimierung beigefügt, die für die Spezialisten die weitere Orientierung ermöglicht.

L. Megyesi (Szeged)

J. C. Burkill, A First Course in Mathematical Analysis, Vi+186 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1978.

This is the first paperback edition of this textbook. The previous editions were published in 1962 (first edition), and in 1964, 1967, 1970, 1974 (reprints). The exposition is very clear and straightforward, the symbolism is very simple. The chapter on functions of several variables is quite short, it does not include the integration of such functions. There are many valuable exercises. The book can be recommended to undergraduate students of mathematics or to anyone who knows high school mathematics and wishes to start studying mathematical analysis.

József Szűcs (Szeged)

J. S. R. Chisholm, Vectors in three-dimensional space, XII+293 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1978.

This book deals with vector algebra and analysis and their application to three-dimensional geometry and to the analysis of fields in 3-space. Both the "pure" and "applied" aspects are considered. The text starts with the algebra of vectors based on the axioms of vector space algebra. When the axioms are introduced, their geometrical interpretation is given, so that they can be understood intuitively. The axiomatic scheme is extended to provide a definition of Euclidean space. The scalar and outer products in 3-space are also introduced in a geometric way. Descriptions of coordinate transformations, congruence and general linear transformations in terms of matrices are given and tensors in 3-space are defined in a classical manner by means of transformation laws. Another part of the text deals with vector analysis. This part begins with the definition and differentiation of curves and surfaces, and with a short account of the differential geometry of curves. Surface and volume integrals are also defined. At the end of the text the differential calculus of scalar and vector fields are investigated and two versions of Stokes' theorem are proved. All chapters contain a large number of problems, some of them are solved at the end of the book.

The book can serve as a textbook for undergraduate students.

L. Gehér (Szeged)

P. M. Cohn, Skew field constructions (London Mathematical Society Lecture Note Series, 27), XII+253 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1977.

This book is based on courses and lectures given by the author at numerous universities all over the world in the years 1971—1976. The purpose is to describe some methods of constructing skew fields (also called division rings), the starting point being the "coproduct construction", the author's famous result (1971) on the existence of a universal field of fractions of any coproduct of skew fields. This construction and the powerful coproduct theorems of G. M. Bergman (1974) form the background of the subsequent topics: a general discussion of skew field extensions in terms of presentations, the word problem for free fields and the solving of equations.

Only familiarity with the material of a standard algebra course is supposed on the reader's part, as the first three chapters summarize the classical results, e.g. Ore' s method of skew polynomials, skew power series and extensions of finite degree. The book is recommended first of all to research workers and postgraduate students, who want to get acquainted with this comparatively new branch of algebra which developed extremely rapidly during the past decade.

Á. Szendrei (Szeged)

Pierre Collet—Jean-Pierre Eckmann, A Renormalization Group Analysis of the Hierarchical Model in Statistical Mechanics (Lecture Notes in Physics, 74), 199 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1978.

The present book deals with one of the most interesting methods of statistical physics, the so-called renormalizition group method. More precisely, the hierarchical models are investigated in detail. In this case the renormalization group method leads to a relatively simple non-linear integral equation. The investigation of this equation makes it possible to obtain a rigorous description of critical phenomena in the hierarchical models. Let us remark that the (rigorous) description of critical phenomena is one of the most difficult problems in statistical physics, and it is solved only in special cases.

The authors give a bibliography of the most important works on hierarchical models, and also prove several new results.

The book consists of two parts. In the first part the main definitions and theorems are given, and the different aspects of the renormalization group technique are discussed. The second part contains the detailed proofs.

The notation of hierarchical models is introduced in Sections 1.1 and 1.2, and a probabilistic interpretation of the renormalization group is given here also. The basic non-linear equation of the theory of hierarchical models is deduced, the critical models are investigated, and in particular the critical indices are computed. In Sec. 1.3 a very important theorem is given about the existence of non-gaussian solution of the basic non-linear equation.

Sections 1.4—1.6 contain several difficult theorems, connected with the computation of the critical indices. From a probabilistic point of view the problem is to determine the limit distribution of the average spin with an appropriate norming factor when the temperature is in a small neighbourhood of the critical temperature T_0. If the temperature T is above the critical temperature T_0 then the limit distribution is gaussian with variance $\sigma=\sigma(T)$ tending to infinity as $T\to T_0$. In case $T<T_0$ this distribution is the mixture of two gaussian distributions. If $T=T_0$, then the distribution is the solution of the basic nonlinear equation investigated in Sec. 1.3. We remark that the first results of this type were obtained in the works of Blecher and Sinai, but the proofs given in the second part of this book are considerably different. Sec. 1.7 contains a proof about the existence of the thermodynamical limit of free energy, of magnetisation, and other observations, and the critical indices of the hierarchical models are directly computed.

The book is written in a clear and concise form. It is an excellent introduction to this rapidly developing field. It may be very useful both for mathematicians and physicists.

Péter Major (Budapest)

B. Davis, Integral transforms and their applications (Applied Mathematical Sciences, 25), XII+ 411 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1978.

The book is intended to serve as introductory and reference material for the application of integral transforms to the solution of mathematical problems in the physical, chemical, engineering and related sciences. The material involved is rather selective than encyclopedic. There are many facets of subject, which are omitted or only outlined. On the other hand, the material is treated in various aspects and illustrated by appropriately chosen application examples.

The book is divided into four parts, supplemented by three Appendices, a Bibliography and an Index. Part I is devoted to the study of the Laplace transform. The inversion theorem is presented in detail, then various applications are made to the solution of ordinary differential equations, partial differential equations (diffusion, wave propagation, etc.) and integral equations (of Volterra type, equation for hard rods etc.).

Part II deals with the fundamental properties of the Fourier transform (up to the Kramers—König relations) and its application to potential and wave problems. Then the treatment proceeds to the theory of generalized functions. There is a considerable amount of "pure mathematics" associated with the understanding and use of generalized functions, because their use adds essentially to the power of the Fourier transform as a tool. Fourier transforms in two or more variables are also included.

Part III contains other important transforms: (*i*) Mellin transforms, (*ii*) Hankel transforms, and (*iii*) integral transforms generated by Green's functions. Special techniques are collected in Part IV. The Wiener—Hopf technique is developed in relation to some instructive problems like reflection and diffraction of waves, radiative processes in astrophysics etc. Then the presentation of the Laplace method for ordinary differential equations follows, by which one can produce integral transform solutions using Hermite-, Bessel- etc. functions. This part ends with different numerical inversion forms of Laplace transforms.

Two more remarks on the presentation. (*i*) The author lays great stress on the use of complex variable techniques, which is frequently of great power. (*ii*) Each section is followed by a rich collection of appropriate problems serving as exercises for the reader.

The book is warmly recommended to everybody wishing to get acquainted with integral transforms, the applications of which outside mathematics, both directly and through differential equations, meet an ever increasing demand of natural sciences.

F. Móricz (Szeged)

W. G. Dixon, Special relativity. The foundation of macroscopic physics, VIII+261 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1978.

The macroscopic physics treated in special relativity seems irrelevant to many physicists, because these macroscopic phenomena under terrestrial conditions are described by the Newtonian theory in a simpler way and with a negligibly small error. This book proves that the macroscopic physics discussed in the framework of special relativity is significant not only from a theoretical point of view. It shows that an understanding of the basic laws of macroscopic systems can be gained more easily within relativistic physics than within Newtonian physics.

The first two chapters contain an introduction to Newtonian physics, to the spacetime structure of special relativity and to tensor algebra. After this introduction the book is devoted to three subjects of special relativity: dynamics, thermodynamics, and electromagnetism.

The theory of dynamics contains both the point particle dynamics and continuum dynamics. The most important topics of this theory are the momentum, angular-momentum and energy conservation laws. The part on thermodynamics describes the entropy law, the equilibrum thermodynamics of relativistic simple fluids and the thermodynamics of irreversible processes. The same techniques is applied to the study of the interaction of simple fluids with an electromagnetic field.

The corresponding Newtonian results of these theories are obtained by taking the Newtonian limit.

The book is not directed towards any particular university course, and it should be accessible to any undergraduate in mathematics or physics.

Z. I. Szabó (Szeged)

V. Dudley, Elementary Number Theory, 2^{nd} edition, IX+249 pages, W. H. Freeman and Co., Reading — San Francisco, 1978.

The second edition of this outstanding undergraduate textbook is a slightly extended version of the first one. Some errors have been removed (and as the author asserts in the preface, some new ones have been added). One of the main merits of the book is that — in contrast with most university textbooks — it can be used with success not only by the best students but also by the average ones, and indeed they can rather deeply understand the topic from it. This is achieved, beside a very clear treatment, by well-chosen exercises inserted in the basic text.

The first twelve chapters give a standard course on divisibility and on congruences, including quadratic reciprocity. Sections 14—15 deal with arithmetic in different place-value systems, sections 16—20 with different non-linear diophantine equations, and sections 21—22 with primes. The last section contains 100 additional problems of different levels. There are three appendices: one on proof by induction, another on problems for computers, and a factor table for integers $< 10\,000$.

There is a number of misprints (whether old or new) and sometimes they emerge at the most inconvenient moments — in numerical examples which ought to inform the reader; and puzzle him instead. This excellent book would have deserved a more careful printing.

G. Pollák (Szeged)

Herman H. Goldstine, A history of Numerical Analysis from the 16th through the 19th century (Studies in the History of Mathematics and Physical Sciences, 2.), XIV+348 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1977.

The author worked together with Professors von Neumann and Murray on the problem of determining the eigenvalues and vectors of real symmetric matrices. They rediscovered, among other results, Jacobi's method. The author presented this in 1951 at a meeting on numerical analysis at UCLA. After the presentation Professor Ostrowski of Basel told him this had been done a century earlier by Jacobi. Partly this event had caused the author to get more comprehensive and thorough information from the history of ideas and methods of numerical analysis as a result of which this excellent book came into being.

The author attempts to trace the development of numerical analysis during the period in which the foundations of the modern theory were being laid. He chooses the most famous mathematicians of the period in question and concentrates on their major works in numerical analysis.

The book is divided into five chapters and ends with a rich Bibliography containing about 300 items and a detailed Index.

Chapter 1 is entitled „The Sixteenth and Early Seventeenth Centuries". During this period mathematical notation began to improve quite markedly and the reasonable symbolisms contributed greatly to the development of mathematics. One of the great discoveries of the sixteenth century was that of logarithms made independently by Bürgi and Napier.

Chapter 2 ("The Age of Newton") discusses Newton's contributions to numerical techniques such as his method for solving equations iteratively, his interpolation and numerical integral formulas as well as his ideas on calculating tables of logarithms and of sines and cosines. Newton's friends and contemporaries Halley, Cotes, Stirling, Maclaurin, Gregory, Moivre, and James Bernoulli, among others, quickly took up his ideas and published a great deal of work which is of interest in numerical analysis.

Chapter 3 ("Euler and Lagrange"): The invention of classical analysis is very largely due to Euler. Even a glance through a volume of his enormous collectedworks shows how Euler differed from Newton. There are no geometrical figures present, he worked with functions and studied their properties in the modern manner. He layed down at least the groundwork in virtually all topics of modern numerical analysis, especially the basic notions for the numerical integration of differential equations. Lagrange worked on linear difference equations and elaborated his famous method of variation of parameters in this connection. He was interested in interpolation theory and introduced some quite elegant formalistic procedures which enabled him to develop many important results.

Chapter 4 ("Laplace, Legendre, and Gauss") begins with the presentation of Laplace's work, who used and developed the method of generating functions to study difference equations which came up in his study of probability theory. Using this apparatus, he was also able to develop various interpolation functions and to produce a calculus of finite differences. Gauss wrote much on numerical matters and obviously enjoyed calculating. The Method of Least Squares was published by Legendre in 1805 but had been used much earlier by Gauss. Also, Gauss took the Newton-Cotes method of numerical integration and showed that by viewing the position of the ordinates as parameters to be chosen one can materially improve convergence. Later Jacobi reconsidered this result and gave a very elegant exposition of it. Gauss wrote penetratingly on interpolation, and particularly on trigonometric interpolation. In fact he developed the entire subject of finite Fourier series, including what is now called the Cooley-Tukey algorithm or the fast Fourier transform.

After Gauss there were a considerable number of excellent mathematicians who either continued his ideas on numerical analysis or utilized their own discoveries to make more elegant what earlier mathematicians had done. Thus, for example, we find on the one hand Jacobi reconsidering some of Gauss' work and on the other hand Cauchy using his Residue Theorem to obtain polynomial approximations to a function. Chapter 5 ("Other Nineteenth Century Figures") is mainly devoted to the presentation of their work in this subject. Among others, Jacobi wrote a paper on finding the characteristic values of a real symmetric matrix mentioned at the beginning which has given rise to the modern Jacobi method and its variants. One of Cauchy's most significant discoveries was a method for finding a rational function which passed through a sequence of given points. This idea of approximation by rational, rather than polynomial, functions is still important and in another connection — Padé approximations — is also used today. Another great advance that Cauchy made was his method for showing the existence of the solutions of differential equations. This so-called Cauchy-Lipschitz method, as well as that of Picard, form the basis for some very important techniques for the numerical integration of such equations. These theoretical methods were exploited by Adams, Bashforth, and Moulton. In a quite different direction Heun, Kutta, and Runge developed a very pretty method for numerical integration of differential equations. One of the first problems run on ENIAC was done using Heun' s method. Their ideas are current today.

A listing of the contents could hardly give a right impression of the richness of the book. It will certainly be a very instructive and profitable reading for everyone interested in numerical analysis.

F. Móricz (Szeged)

H. B. Griffiths— P. J. Hilton, A comprehensive textbook of classical mathematics, XXIX + 637 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1978.

This second edition differs from the original one published in 1970, by Van Nostrand Reinhold Co. in the elimination of some errors and the addition of a few calculus exercises. Since *these Acta* have not reviewed the first edition, we have a closer look at the book here. It is written to those already familiar, at a certain level, with the subject matter they choose from it. Roughly speaking, a one year specialized mathematical study is sufficient to read the material presented without difficulty. The main purpose of the authors has been "to encourage the reader to look at rather familiar ideas a second time, with a view to fitting them into the framework of present-day mathematical thought; and thus to enable the reader to see how certain key ideas recur again and again and give a real unity to apparently separate parts of his early mathematical experience". The presentation follows the "spiral approach": ideas are introduced informally; then precise proofs and definitions are given, after which informality comes again. To give the reader some information of the material covered in the book we list the titles of the eight parts as follows: The Language of Mathematics, Further Set Theory, Arithmetic, Geometry of R^3, Algebra, Number Systems and Topology, Calculus, Additional Topics in the Calculus, Foundations.

We recommend this book to anyone, even to the well-educated mathematician, who wishes to brush up on his basic mathematics.

József Szűcs (Szeged)

P. R. Halmos — V. S. Sunder, Bounded integral operators on L^2 spaces (Ergebnisse der Mathematik und ihrer Grenzgebiete, 96), XV + 132 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1978.

The phrase "integral operator" used in this book is the natural "continuous" generalization of the operators induced by matrices, and the only integrals that appear are the Lebesgue-Stieltjes integrals on classical non-pathological measure spaces. To be more concrete, let X and Y be σ-finite and separable measure spaces. A kernel $k=k(x, y)$ is a complex-valued measurable function on the Cartesian product $X\times Y$. The domain of k is the set dom k of $g\in L^2(Y)$ that satisfies the following two conditions:

(i) $k(x, \cdot)g\in L^1(Y)$ for almost every x in X,

(ii) if $f(x)=\int k(x, y)g(y)\, dy$, then $f\in L^2(X)$.

It may happen that dom $k=0$ (the identically 0 function), for example, this is the case if $k(x, y)= =1/(x-y)$ on $R\times R$ (the kernel that defines the Hilbert-transform). In any event, whatever its domain might be, a kernel always induces an operator, denoted by Int k, that maps dom k (in $L^2(Y)$) into $L^2(X)$; the image under Int k of a function g in dom k is the function f in $L^2(X)$ given by (ii). The integral operator Int k is linear but not necessarily bounded.

The book does not strive for maximum generality. The study is restricted mostly to bounded integral operators as indicated in the title. Even in this special setting the authors do not answer all the questions about integral operators. Frequently, when the systematic treatment encounters unanswered questions, the authors point out, where such questions arise, how they are connected with others, and what partial information about them is available. The emphasis in the treatment is on the basic implication relations on which the subject rests, rather than on its mechanical techniques.

The main prerequisite for an uninterrupted reading of the book is familiarity with the standard facts of measure theory and operator theory.

The book consists of 17 sections. The first five contain the definitions and the examples that are needed throughout. Sections 6—9 describe what can and what cannot be done with integral operators. The topics are the possibility of transforming integral operators by measure-preserving isomorphisms, the correspondence from kernels to operators, and the extent to which that correspondence preserves the algebraic operations on kernels. Sections 10 and 11 treat two important classes of kernels: absolutely bounded kernels and Carleman kernels.

Sections 12—14 provide some necessary tools from operator theory for the subsequent sections: a discussion of two different kinds of compactness, and the properties of the essential spectrum, culminating in the celebrated Weyl-von Neumann theorem on the possibility of a kind of generalized diagonalization for Hermitian operators on infinite-dimensional Hilbert spaces. The last three sections deal with the most interesting and up-to-date questions:

(i) Which operators can be integral operators? In precise terms, it asks for a characterization of those operators A on $L^2(X)$ for which there exists a unitary operator U on $L^2(X)$ such that UAU^* is an integral operator? This question has a complete answer.

(ii) Which operators must be integral operators? In precise terms: under what conditions on an operator A on $L^2(Y)$ does it happen that UAU^* is an integral operator for every unitary U on $L^2(X)$? This question has a satisfactory partial answer, but some special questions (e.g., about absolutely bounded kernels) remain open.

(iii) Which operators are integral operators? The problem is one of recognition: if an integral operator on $L^2(X)$ is given in some manner other than by its kernel, how do its operational and measure-theoretic properties reflect the existence of a kernel that induces it? Here various useful sufficient conditions are available, but none of them are necessary.

The writing of the book was mainly motivated by the fact, as the authors admit in Preface, that the theory of integral operators is the source of all modern functional analysis and remains to this day a rich source of non-trivial examples. Since the major obstacle to progress in many parts of operator theory is the dearth of concrete examples whose properties can be explicitly determined, a systematic theory of integral operators offers new hope for new insights. And the programme of the authors is completely materialized in this book, which makes a very essential contribution to the systematization of the theory of integral operators.

This book is indispensable for all specialists of functional analysis, but it is also warmly recommended to everybody who wants to keep pace with up-to-date developments in analysis.

F. Móricz (Szeged)

Herbert Heyer, Probability measures on locally compact groups, X+537 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1977.

Probability measures on algebraic topological structures and especially on locally compact topological groups have become of increasing importance in recent years. The main purpose of the present book is to give a systematic presentation of the most developed part of the work done in this field. The text is divided into 6 Chapters. To make the book as self-contained as possible the first two chapters have been devoted to general tools from the harmonic analysis of locally compact groups and from the elementary convergence theory of convolution sequences of probability measures on the group. In Chapter 3 the general embedding problem is posed. The most important step on the way to central limit theorem is the embedding of an infinitely divisible measure into a continuous one parameter convolution semigroup. Since the embedding theorem does not hold in a general locally compact group, the question arises what classes of groups yield the validity of an embedding theorem. Establishing these classes of groups is the aim of this chapter. Chapter 4 includes an extensive discussion of the canonical representation of all continuous convolution semigroups in the sense of a Lévy — Khintchine formula. There is a natural connection between convolution semigroups on a group G and contraction semigroups of operators on certain function spaces E on G such that the problem of generating a convolution semigroup becomes a problem of determining the existence of the infinitesimal generator of the corresponding contraction semigroup on E. Using the solution of Hilbert's V. problem and the ideas of Bruhat a differentiable structure can be introduced in any locally compact group and the problem will be reduced via Lie projectivity to the Lie group case. The aim of Chapter 5 is twofold: to motivate the broad discussion of the central limit theorem in the special case of an Abelian group and to give certain auxiliary results which will be needed for the treatment of the problems for more general groups. Most of this material is applied to a detailed treatment of additive stochastic processes with values on a locally compact Abelian group having a countable topological basis. Chapter 6 is devoted to the central limit problem in the general case. Many facts discussed at an earlier stage will be combined here for a detailed study of Poisson and Gauss measures on arbitrary locally compact groups as well as for the study of the convergence behavior of triangular systems of probability measures in the sense of a Lindeberg-Feller central limit problem.

The book is highly recommended to research workers taking interest in modern probability theory and having certain knowledge of representation theory.

L. Gehér (Szeged)

M. Karoubi, *K*-theory (Grundlehren der mathematischen Wissenschaften — A Series of Comprehensive Studies in Mathematics, 226), XVIII+308 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1978.

It's well-known that ordinary cohomology theory is defined uniquely (on the category of polyhedra) by the Eilenberg-Steenrod axioms. However, if we omit the so-called dimension axiom: $H^k(P)=0$ if P is a point and $k>0$, then there will exist infinitely many functors from the category of polyhedra into the category of abelian groups satisfying the other axioms. These functors are called extraordinary cohomology theories. One of the extraordinary theories is K-theory. The advantage of extraordinary theories is that they usually give much more information about the topological situation considered. The particular advantage of K-theory is that it appears very naturally in consideration of differential manifolds and fibre bundles, because its elements are — roughly speaking — the vector bundles themselves. The exact definition of the (topological) K-functor is the following: Let us consider all vector bundles over a space X. They form a semiring under the Whitney sum and the tensor product. This semiring — as well as any other one — defines a "minimal" ring (the Grothendick ring of the semiring). This ring is the $K(X)$ ring for the space X. Starting from the K-functor one can define a cohomology theory in the following way: For $i>0$ let $K^{-i}(X)$ equal $K(S^iX)$ where S^iX is the i-fold suspension on X. The sequence $K^{-i}(X)$ turns out to be periodic modulo 2 in the complex case and modulo 8 in the real case. This enables us to extend the definition of $K^i(X)$ for $i>0$. As well-known, the characteristic classes serve for the description of vector bundles by means of the classical cohomology theory. Characteristic classes can be defined in K-theory, too. The present book is the first monograph dealing with characteristic classes in K-theory, as well. There exist three ways to define characteristic classes:

1) the axiomatic way;
2) using the cohomology ring of the Grassman manifold;
3) by the Thom isomorphism theorem.

In this book all the three definitions are presented. The particularity of K-theory is that in it the characteristic classes are the same as the cohomology operations. The cohomology operations are describe at the end of Chapter IV. Chapter V deals with interesting applications. First, it presents Adams' proof for the statement that on the sphere S^n there exists an H-structure iff $n=1$ or 3 or 7. (We remind that the H-structure is a generalization of the topological group structure.) The second application is the solution of the following question: How many continuous linearly independent vector fields do there exist on the u-sphere? The third interesting question in this chapter is the so-called Chern character which is an isomorphism between the groups $K_c(X)\,Q$ and $H^{even}(X; Q)$ for any compact space. At the end of the book are the most interesting questions: the Riemann-Roch theorem and the integrability theorems for the characteristic classes. The book is recommended to anybody wishing to study this new exciting and powerful mathematical theory.

András Szűcs (Szeged)

W. Klingenberg, Lectures on Closed Geodesics (Grundlehren der mathematischen Wissenchaften, 230), IX+227 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1978.

The question about the existence of closed geodesics on a simply connected compact Riemannian manifold has been in the centre of investigations in global differential geometry since Jacobi's description of geodesics on an ellipsoid in 1842. In 1905 Poincaré claimed that this problem was closely related to the question whether there existed a periodic solution of the restricted three body problem.

The greatest advances in the theory of closed geodesics were the results of L. A. Lusternik and L. G. Schnirelmann in 1929 and of L. A. Lusternik and A. I. Fet in 1951. They showed that on a simply connected compact surface there exist at least three closed geodesics without self intersection and that at least one closed geodesic exists on every compact Riemannian manifold.

In the last 15 years Prof. Klingenberg worked out two very effective new approaches to the existence problem of closed geodesics: the Morse theory on an infinite dimensional Hilbert-Riemann manifold and the method of Hamiltonian systems and geodesic flows. The starting problem is completely solved at the present stage of investigations by the main theorem of this monograph: On a compact Riemannian manifold with finite fundamental group there exist infinitely many closed geodesics. This fundamental result of Klingenberg, published in detail here for the first time, gives essential new information even in the case of convex surfaces in euclidean 3-space.

The aim of this book is to give an up-to-date and detailed introduction to the new methods of investigation on the geometry of closed geodesics and to give self-contained proofs of the essential new results in this theory.

In Chapter 1 the notion of a Hilbert manifold is introduced and a canonical Hilbert manifold structure is defined on the space of closed curves in a compact Riemannian manifold. The question about the existence of closed geodesics can be translated into a question about the critical values of the energy function on the Hilbert manifold of closed curves.

Chapter 2 is devoted to the development of the Lusternik-Schnirelmann and Morse theory on the manifold of closed curves.

In Chapter 3 the theory of Hamiltonian systems is discussed from the aspects of geodesic flows on a Riemannian manifold. This proves to be a very effective tool for the study of periodic geodesics in a neighborhood of a given one.

Chapter 4 contains the main result on the existence of infinitely many closed geodesics on a manifold with finite fundamental group. It concludes with generic existence theorems derived from the properties of geodesic flows.

In Chapter 5 an n-dimensional generalization of the classical Lusternik-Schnirelmann theorem and a number of miscellaneous results about closed geodesics on special Riemannian manifolds are proved.

In an Appendix, an elementary treatment of the Lusternik-Schnirelmann theory is given independently of the previous parts of the book.

The book contains fundamental and new information about central problems of global differential geometry. Chapters 1—3 can serve as an excellent introduction into the new methods of investigation of geometry of geodesics, Chapters 4—5 contain the main results of the theory. The latter part is not very easily readable, because a great variety of analytical and topological methods is used. It is suggested to the reader that though the presented results solve the starting problems of the theory, a great many interesting questions are left open which can be studied with these new methods only.

The book is an indispensable monograph on the subject. It is warmly recommended to research workers in differential geometry, the global theory of dynamical systems, and nonlinear functional analysis.

P. T. Nagy (Szeged)

Wilhelm Klingenberg, A course in differential geometry (Graduate Texts in Mathematics, 51), XII + 178 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1978.

English translation of the original "Eine Vorlesung über Differentialgeometrie" (Heidelberger Taschenbücher, 1973; reviewed in *these Acta* **36** (1974)). It contains an excellent introduction to elementary differential geometry for undergraduate students. The present edition is more detailed and a number of figures is added.

P. T. Nagy (Szeged)

Hans Kurzweil, Endliche Gruppen. Eine Einführung in die Theorie der endlichen Gruppen (Hochschultext), XII + 187 Seiten, Berlin—Heidelberg—New York, Springer-Verlag, 1977.

Dieses Buch, das durch seine leichte Lesbarkeit und seinen didaktisch guten Aufbau überwiegend für Studenten zusammengestellt ist, "möchte — nach seiner Zielsetzung — den Leser mit den Grundlagen und Methoden der Theorie der endlichen Gruppen vertraut machen und ihn bis an aktuelle Ergebnisse heranführen". Sein Lesen benötigt nur elementare Kenntnisse der linearen Algebra. Für diejenigen, die sich späterhin mit dem Problem der Bestimmung einfacher Gruppen eingehend befassen möchten, ist das Lesen dieses Lehrbuches besonders vom Nutzen, denn sie finden hier zahlreiche grundlegende Kenntnisse, Begriffe, die zum Studieren des genannten Themas unentbehrlich sind.

L. Megyesi (Szeged)

W. S. Massey, Algebraic topology: An introduction (Graduate Texts in Mathematics, 56) XXI + 261 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1977.

The book is the 4th corrected printing of the excellent textbook, the earlier printings of which were published by Harcourt Brace and World, New York, 1967. It is a very elegant introduction to algebraic topology, concerning three classical subjects: 2-dimensional manifolds, fundamental groups and covering spaces.

Chapter 1 discusses 2-dimensional manifolds with numerous examples and exercises. The classification theorem for compact surfaces is also proved.

Chapters 2—4 deal with fundamental groups. Besides their basic properties, the Brouwer fixed point theorem and the Seifert — Van Kampen theorem on the fundamental group of the union of two spaces are discussed.

The covering spaces and the relationship between covering spaces and the fundamental groups are described in Chapter 5. Chapters 6—7 present topological proofs of several well-known theorems of group theory, namely, the Nielsen-Schreier theorem on subgroups of a free group, the Kurosh theorem on subgroups of a free product, and the Grushko theorem on the decomposition of a finitely generated group as a free product. Chapter 8 gives an outlook to algebraic topology.

Z. I. Szabó (Szeged)

Th. Meis und U. Marcowitz, Numerische Behandlung partieller Differentialgleichungen (Hochschultext), VIII+452 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1978.

A good introduction to the study of numerical methods of partial differential equations. The authors confine themselves chiefly to methods of solving linear second-order partial differential equations with two independent variables. However, differential equations with more variables than two as well as non-linear partial differential equations are also treated.

The book consists of three parts and an Appendix. Part I is devoted to the study of initial value problems in differential equations of hyperbolic or parbolic type. Part II proceeds to the solution of boundary value problems in differential equations of elliptic type. The concepts of consistency, stability and convergence of a method are in the central place of the treatment. The most widely used numerical methods of solving partial differential equations are the finite difference methods, and they are presented in detail. The use of the Fourier method for standard problems in mathematical physics, the variational methods and collocation methods of solving boundary value problems are dealt with also in detail.

Part III systematically develops a substantial portion of the theory of iterative methods for solving systems of (linear or non-linear) algebraic equations that arise in the numerical solution of boundary value problems by finite difference methods. The focal point is an analysis of the convergence properties of the successive overrelaxation method (SOR method) in the linear case, and that of the Newton-Raphson method with some of its variants in the non-linear case. Some techniques for solving large systems of linear algebraic equations with sparse matrices are also included.

The forth part called Appendix contains the FORTRAN programs of certain well-chosen problems with the necessary explanation and documentation.

The presentation is always clear and well-readable. The theoretical background is given in detail, the methods are illuminated in a many-sided manner. The textbook is highly recommended to students in numerical analysis as well as to experts in physics, chemistry and engineering interested in the solution of partial differential equations.

F. Móricz (Szeged)

David Mumford, Algebraic Geometry. I. Complex Projective Varieties (Grundlehren der mathematischen Wissenschaften, 221), X+186 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1976.

From the thirties on, Algebraic Geometry turned from Geometry towards Algebra. The algebraic methods enriched the machinery with extremely powerful tools, but they need a very long, systematic, and — at least for the beginners — boring foundation. On the other hand, even if one works with schemes, a good geometric intuition is needed in order to "see" the problem. However, most recent books on algebraic geometry emphasise the algebra part, and say very little (if any) about the geometric sources.

The present book provides an introduction to the subject, emphasising the geometric part. It shows the deep connections between algebraic and analytic geometry and topology. First the analytic structure of an algebraic variety is investigated. Then the Zariski and Euclidean topologies are compared, and Chow's theorem is proved which characterises the projective algebraic varieties as closed analytic submanifolds of complex projective spaces. Another fascinating characterisation is given later when projective varieties are described as compact oriented differentiable submanifolds of the complex projective spaces having minimal volume in a certain natural sense. Finally the Euclidean topology of curves is determined.

The use of topological methods throughout the book enables the author to make considerable shortcuts in the proofs and makes the definitions clearer (this concerns mainly the notions connected with multiplicity). The last paragraph deals with the 27 lines on a cubic surface, one of the most astonishing facts in geometry.

The algebraic part of the theory is also developed in a very efficient way that leads quickly to interesting theorems.

The whole book is written in a very clear and concise style. All this makes the book an excellent introduction, especially suitable for mathematicians who are not primarily interested in algebraic geometry.

János Kollár (Budapest)

R. K. Sachs — H. Wu, General Relativity for Mathematicians (Graduate Texts in Mathematics, 48) XII+291 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1977.

Many recent monographs on general relativity treat the subject in the frames of modern differential geometry. The present book also gives a clear and geometrical description of general relativity, using this terminology. The reader is only supposed to have familiarity with tensor algebra, differential topology, and the rudiments of Riemannian geometry.

After some mathematics and physics background on Lorentzian manifolds the book gives a systematic description of particle dynamics, electromagnetism and several matter models. In the second part it treats several cosmological questions: Einstein's field equation, the theory of photons and photon gases, the Einstein — de Sitter and Schwarzschild model of space-time, black holes, etc.

The book is a fundamental monograph on the subject. It is especially well-organized; its didactical value is greatly enhanced also by the great number of examples worked out.

Z. I. Szabó (Szeged)

Mathematics Today. Twelve Informal Essays, Edited by Lynn Arthur Steen, VI+367 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1978.

This book is intended to be a popularizing work for the intelligent non-mathematicians, particulary for those who have already met mathematics in their scientific research. It contains 12 well-readable essays by 12 authors extraordinarily well selected from distinct fields of pure and applied mathematics. The essays are the following: Mathematics—Our Invisible Culture (Allen L. Hammond), Number Theory (Ian Richards), Groups and Symmetry (Jonathan Alperin), The Geometry of the Universe (Roger Penrose), The Mathematics of Meteorology (Philip Thompson), The Four Color Problem (Kenneth Appel and Wolfgang Haken), Combinatorial Scheduling Theory (Ronald Graham), Statistical Analysis of Experimental Data (David S. Moore), What is a Computation? (Martin Davis), Mathematics as a Tool for Economic Understanding (Jacob Schwartz), Mathematical Aspects of Population Biology (Frank C. Hoppensteadt), The Relevance of Mathematics (Felix E. Browder and Saunders Mac Lane).

The book is printed in an aesthetic format. It is highly recommended to professional mathematicians as well as to laymen.

L. Gehér (Szeged)

Şerban Strătilă—László Zsidó, Lectures on von Neumann Algebras, 478 pages, Editura Academiei (Bucureşti, Romania) — Abacus Press (Tunbridge Wells, Kent, England), 1978.

This book is a revised and updated English version of the original Roumanian "Lecţii de algebre von Neumann" published in 1975 by the above Roumanian publisher. Both authors made important contributions to the theory of von Neumann algebras. This theory was initiated by J. von Neumann and F. J. Murray in the thirties in connection with infinite group representations and theoretical physics, etc. The first systematic treatment of the subject was given by J. Dixmier in 1957, which followed I. Kaplansky's lecture notes "Rings of operators" published in 1955 (reprinted in 1968). It was Dixmier who used the term "von Neumann algebras" as an equivalent of "rings of operators" or "operator algebras". His monograph included almost all the important results known in the field by then. The theory of von Neumann algebras has developed very rapidly and extensively since that time. Besides the many research papers, a number of expository works have also been published, such as Sakai's excellent monograph "C^*-algebras and W^*-algebras" (1971), which followed his lecture notes "The theory of W^*-algebras" (1962), J. R. Ringrose's "Lecture notes on von Neumann algebras" (1967), "Lectures on Banach algebras and spectral theory", and "Lectures on the trace in finite von Neumann algebras" (1972), Takesaki's lecture notes "The theory of operator algebras" (1970), and "Lecture notes on operator algebras" (1973/1974), Topping's "Lectures on von Neumann algebras" (1971), a new edition of Dixmier's classic in 1969, and the Roumanian version of the present book in 1975.

A turning point in the development of the theory of von Neumann algebras came with M. Tomita's discovery of modular Hilbert algebras in 1967. His results were published in Takesaki's lecture notes "Tomita's theory of modular Hilbert algebras and its applications" in 1970. Tomita devised canonical forms for arbitrary von Neumann algebras.

At the present stage of development it cannot be expected that a single volume expounds all features of the existing theory. The book under review develops the theory of standard von Neumann algebras (or, in other words, Tomita's theory), but it does not discuss reduction theory, the isomorphism theory of factors, non-commutative harmonic analysis and ergodic theory, applications to operator theory and theoretical physics, the generation of von Neumann algebras. C^*-algebras are only incidentally referred to. Just as Dixmier's classic, it develops the spatial theory of von Neumann algebras, i.e., von Neumann algebras are considered sub-algebras of the full operator algebra, in contrast with Sakai's abstract treatment, where they are considered as C^*-algebras with a predual. The material presented covers the results contained in M. Takesaki's work on Tomita's theory and in Dixmier's classic, except reduction theory and examples of factors. The contributions of van Daele and the second author of this book made it possible to simplify the proof given by Takesaki in his lecture notes. Following I. Cuculescu and S. Sakai, the commutation theorem for tensor products is proved independently of Tomita's theory.

The book is clearly written. It only assumes knowledge of the rudiments of functional analysis. There are many valuable exercises, some of them borrowed from Dixmier's classic. Very valuable comments supplement the text proper at the end of each chapter. A very thorough 20 page bibliography on operator algebras and related topics is included. We very warmly recommend this book to beginners in operator algebras or to research workers who will find that this book gives very thorough bibliographical information or serves very well as a reference book.

József Szűcs (Szeged)

Livres reçus par la rédaction

C. de Boor, A practical guide to splines (Applied Mathematical Sciences, Vol. 27), XXIV+392 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1978. — DM 32,50.

Shing-Shen Chern, Selected papers, XXI+476 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1978. — DM 52,—.

Yuan Shih Chow—Henry Teicher, Probability theory. Independence, interchangability, martingales, XV+455 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1978. — DM 55,—.

H. Cohn, A classical invitation to algebraic numbers and class fields with two appendices by Olga Taussky. "Artins 1932 Göttingen lectures on class field theory" and "Connection between algebraic number theory and integral matrices" (Universitext), XIII+328 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1978. — DM 32,—.

F. J. Fritz—B. Huppert—W. Willems, Stochastische Matrizen (Hochschultext), 2., neubearbeitete und erweiterte Auflage, VIII+193 Seiten, Springer-Verlag, Berlin—Heidelberg—New York, 1979. — DM 32,—.

W. Gähler, Grundstrukturen der Analysis, Bd. II (Mathematische Lehrbücher und Monographien. II. Abteilung: Mathematische Monographien, Bd. 42), VIII+ 623 Seiten, Akademie-Verlag, Berlin, 1978, — 78,— M.

T. W. Gamelin, Uniform algebras and Jensen measures (London Mathematical Society Lecture Note Series, 32), VI+162 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1978. — Ł 6.50.

W. Greub, Multilinear Algebra (Universitext), 2nd edition, VII+294 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1978. — DM 43,—.

I. A. Ibragimov—Y. A. Rozanov, Gaussian random processes (Applications of Mathematics, Vol. 9), Translated from the Russian by A. B. Aries, X+275 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1978. — DM 49,60.

B. A. Koslow—I. A. Uschakow, Handbuch zur Berechnung der Zuverlässigkeit in Elektronik und Automatentechnik (Elektronisches Rechnen und Regeln, Sonderband 25), In deutscher Sprache herausgegeben und ergänzt von K. Reinschke, XIX+598 Seiten, Akademie-Verlag, Berlin, 1978. — 120,— M.

S. Kovalevskaya, A Russian Childhood. Translated and introduced by B. Stillman, XIV+250 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1978. — DM 33,—.

W. Köhnen, Einführung in die Theorie der metrischen Räume (Beiträge zur Fachdidaktik), 392 Seiten, Aloys Henn Verlag, Kastellaun Hunsrück, 1978. — DM 39,80.

H. J. Kushner—D. S. Clark, Stochastic approximation methods for constrained and unconstrained systems (Applied Mathematical Sciences, Vol. 26), X+261 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1978. — DM 26,40.

S. Lang, Elliptic curves — Diophantine analysis (Grundlehren der mathematischen Wissenschaften — A Series of Comprehensive Studies in Mathematics, 231), XI+261 pages, Springer-Verlag, Berlin—Heidelberg—NewYork, 1978. — DM 68,—.

S. Lang, Cyclotomic fields (Graduate Texts in Mathematics, Vol. 59), XI+253 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1978.

E. H. Lockwood—R. H. MacMillen, Geometric symmetry, X+228 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1978. — Ł 10.50.

D.P. Parent, Exercices de théorie des nombres, X+307 pages, Gauthier-Villars, Paris, 1978. — 59,— F.

J. F. Pommaret, Systems of partial differential equations and Lie pseudogroups (Mathematics and its Applications, Vol. 14), XIV+411 pages, Gordon and Breach, London—New York, 1978.— Ł 21. 90.

K. Prachar, Primzahlverteilung (Grundlehren der mathematischen Wissenschaften, Bd. 91), Reprint der Erstauflage Berlin/Göttingen/Heidelberg 1957. Ergänzt um ein Vorwort des Autors und ein Verzeichnis mit Ergänzungen, Verbesserungen und neuere Literatur, X+451 Seiten, Springer-Verlag, Berlin—Heidelberg—New York, 1978. — DM 86,—.

F. Rehbock, Geometrische Perspektive, IX+155 Seiten, Springer-Verlag, Berlin—Heidelberg—New York, 1979, — DM 29,50.

L. A. Steen—J. A. Seebach, Jr., Counterexamples in topology, 2nd edition (Originally published by Holt, Rinehart and Winston, New York. 1970) XI+244 pages Springer-Verlag, New York—Heidelberg—Berlin, 1978. — DM 26,—.

O. Steinfeld, Quasi-ideals in rings and semigroups (Disquisitiones Mathematicae Hungaricae, 10), XI+154 pages, Akadémiai Kiadó, Budapest, 1978.

Á. Szabó, The beginning of Greek mathematics, 358 pages. Akadémiai Kiadó, Budapest, 1978.

R. J. Walker, Algebraic curves (Originally published in 1950 by Princeton University Press), X+201 p., Springer-Verlag, New York—Heidelberg—Berlin, 1978. — DM 22,—.

H. Werner—R. Schaback, Praktische Mathematik II. Methoden der Analysis (Hochschultext), 2., neubearbeitete und erweiterte Auflage ,VIII+388 Seiten, Springer-Verlag, Berlin—Heidelberg New York, 1979. — DM 38,—.

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

79-356 — Szegedi Nyomda — F. v.: Dobó József igazgató

Felelős szerkesztő és kiadó: Szőkefalvi-Nagy Béla
A kézirat a nyomdába érkezett: 1979. Január 19.
Megjelenés: 1979. Június 30.

Példányszám: 1200. Terjedelem: 24,5 (A/5) ív
Készült monószedessel, íves magasnyomással,
az MSZ 5601-24 és az MSZ 5602-55 szabvány szerint