ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 38
FASC. 3–4

SZEGED, 1976

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNY

ACTA SCIENTIARUM MATHEMATICARUM

38. KÖTET
FASC. 3—4

SZEGED, 1976

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

László Kalmár

(1905—1976)

It was an irreplaceable loss to Hungarian science that Professor László Kalmár, a member of the Hungarian Academy of Sciences unexpectedly died on August 2, 1976.

He was born in a small Hungarian village, Alsó-Bogátpuszta, on March 27, 1905. He started to be interested in mathematics as a high school student. In 1922 he won the first prize of a mathematical contest organized by the Loránd Eötvös Mathematical and Physical Society for the high school graduates of that year.

After studying mathematics and physics at Budapest University, he received his doctor's degree in mathematics at the same university. For a few months he worked as a physicist in a radio tube factory.

From 1930 to 1947 he was an assistant of F. Riesz and A. Haar in Szeged, and in 1947 he was appointed professor of mathematics. Until his retirement in 1975 he headed the Department of Computer Science and the Laboratory of Cybernetics of Szeged University as well as the Research Group on Mathematical Logic and Automata Theory of the Hungarian Academy of Sciences. Besides, he was a member of several national and international committees and was on the editorial board of numerous international periodicals. He taught and researched at Szeged University until his death.

Professor Kalmár had been among the editors of our Acta for thirty years and in his youth he had the duties of a technical editor. Many of the present editors were his students.

He was elected a corresponding member of the Hungarian Academy of Sciences in 1948, and a member in 1961. He was awarded the Kossuth prize in 1950 and the State prize in 1975.

The preponderant part of his mathematical activity falls in the field of mathematical logic, in several branches of which he achieved basic results. Much of his work is related to the decision problem of logic. For instance, he proved that Church's theorem is just a special case of Gödel's theorem on relative undecidability. Another significant result of his is a counterexample to a hypothesis of Schröter that intended to support Church's thesis.

His ability to see the basic points of a newly acquainted proof led him more than once to essential simplifications of the original reasoning. It is enough to mention the ingenious simplification of Gentzen's proof of the consistency of the arithmetics of integers. This result was included in the fundamental work of Hilbert and Bernays: *Grundlagen der Mathematik.*

He was the first in Hungary to realize the use of mathematical logic in sciences and in practice. In the middle 50's he initiated the teaching and research of computer science and cybernetics in Hungary. His own results in these fields contributed to the theory of programming languages. He also obtained interesting results that have applications in medical diagnostics and linguistics. With his manifold ability to conceive the new he also won others for computer science and cybernetics. He raised numerous problems which he did not elaborate himself but made it possible for others to start working in computer science.

Besides his main research areas he obtained many results in analytic number theory, analysis, algebra, and the theory of games.

He always felt obliged to popularize mathematical logic and computer science. He wrote several papers and gave lectures to achieve this goal, and much helped to organize the scientific life in Hungary; in particular we owe him for the foundation of the Laboratory of Cybernetics and of the Research Group on Mathematical Logic and Automata Theory at Szeged University.

We honour the memory of Professor László Kalmár, the mathematician, the teacher and the man.

The Editors

Acta Sci. Math., 38 (1976), 223—238

Diagonalization of matrices over H^∞

BÉLA SZ.-NAGY

Homage to the memory of F. Riesz (1880—1956)

By H^∞ we mean the Banach algebra of bounded holomorphic functions $u(\lambda)$ on the disc $|\lambda|<1$, with the sup-norm $\|u\|_\infty$. For the relevant fundamental notions and facts (inner functions and their canonical representation, inner factor of a non-zero $u \in H^\infty$, largest common inner divisor $\bigwedge_\alpha u_\alpha$, and least common inner multiple $\bigvee_\alpha u_\alpha$ (if it exists), of a family $\{u_\alpha\}$ of inner functions, etc.) we refer e.g. to [4], Chapter III. It is convenient to define $\bigwedge_\alpha v_\alpha$ for *any* family $\{v_\alpha\}$ of elements of H^∞: this is the largest common inner divisor of the v_α whenever not all v_α are zero, and 0 otherwise. Note that the operations $\wedge$, $\vee$ are defined up to constant factors of modulus 1.

Matrices over H^∞ naturally occur in the theory of unitary equivalence, similarity, or quasi-similarity models of certain types of operators on Hilbert space, as made clear e.g. by the investigations of SZ.-NAGY—FOIAŞ [4], [5], [7]. It was in particular the paper [5] first establishing a Jordan model theory for operators of class C_0 which pointed out the need for a diagonalization theory of matrices over H^∞. This task was achieved, for finite rectangular matrices over H^∞, by NORDGREN [3]. The classical equivalence theory cannot be applied here since the algebra H^∞ does not possess all properties required. However, by introducing a convenient generalization of the notion of equivalence for matrices, called *quasi-equivalence,* Nordgren was able to extend the classical results to this case. SZŰCS [10] gave an analysis of the abstract algebraic background of Nordgren's theory.

The results of [3] were applied in [1], [2], [8] to obtain Jordan models for some classes of operators on Hilbert space, namely to contractions T with finite defect indices and of class C_0 (i.e. such that $T^{*n} \to 0$).

The aim of the present paper is to extend the Nordgren diagonalization theory. The key to this extension is the Main Lemma (Sec.2) which establishes a remarkable property of H^∞. It can be applied to solve the diagonalization problem for finite

Received November 1, 1975.

and semi-finite matrices over H^∞ as well, and also to get some insight into the case of (doubly) infinite matrices (Sec.3). The full solution of the problem of infinite matrices would, however, require further study because of the convergencc difficulties which there arise.

The concluding Sec. 4 indicates how the matrix diagonalization results can be applied to obtain a Jordan model of operators $T\in C_{\cdot 0}$ with at least one finite defect index, thus generalizing the results of [8] and [2].

1. Preliminaries

For convenience of reference we begin with some more or less known lemmas.

Lemma 1. *Let ω be an inner function and let p_α, $q_\alpha(\alpha\in A)$ be inner divisors of ω such that $p_\alpha\cdot q_\alpha=\omega$ for each $\alpha\in A$. Then,*

$$\bigwedge_\alpha p_\alpha\cdot\bigvee_\alpha q_\alpha=\omega.$$

Poof. $q^\vee=\bigvee_\alpha q_\alpha$ is divisible by each q_α; hence there exist inner functions v_α such that $q^\vee p_\alpha=\omega v_\alpha$. Since $q^\vee$ is a divisor of ω, we have $p_\alpha=(\omega/q^\vee)\cdot v_\alpha$ for all α. Then $\omega/q^\vee$ is a divisor of $p^\wedge=\bigwedge_\alpha p_\alpha$ also. Therefore, we have

$$\omega/p^\wedge \mid q^\vee.$$

On the other hand, we have $\omega/p^\wedge=(\omega/p_\beta)(p_\beta/p^\wedge)=q_\beta\cdot(p_\beta/p^\wedge)$, and hence $q_\beta\mid(\omega/p^\wedge)$ for every $\beta\in A$, and therefore,

$$q^\vee\mid\omega/p^\wedge.$$

The two relations yield the result we wished to prove.

Corollary. *Under the hypotheses of Lemma* 1 *we have*

$$\bigvee_\alpha q_\alpha=\omega \quad \textit{if and only if} \quad \bigwedge_\alpha p_\alpha=1.$$

Lemma 2. (M. Sherman, *cf.* [6]) *Let $f_1, f_2\in H^\infty$ and let ω be an inner function. Then for every complex number t, with the exception of at most countable many values, we have*

$$\omega\wedge(f_1+tf_2)=\omega\wedge f_1\wedge f_2.$$

Proof. Let $g_1, g_2\in H^\infty$ be any pair linearly equivalent (with constant coefficients) to the pair f_1, f_2. Then $g_1\wedge g_2=f_1\wedge f_2$, and hence

$$\omega\wedge g_1\wedge g_2=\omega\wedge f_1\wedge f_2. \tag{1.1}$$

Applying Lemma 1 to the inner function ω and to its inner divisors $p_\alpha=\omega\wedge g_\alpha$ and $q_\alpha=\omega/p_\alpha$ ($\alpha=1, 2$) we get, taking account of (1.1),

$$q_1\vee q_2 = \frac{\omega}{p_1\wedge p_2} = \frac{\omega}{\omega\wedge g_1\wedge g_2} = \frac{\omega}{\omega\wedge f_1\wedge f_2}\,(=\Omega).$$

By the corollary of Lemma 1, applied with this Ω in place of ω, we obtain that

$$\Omega/q_1\wedge\Omega/q_2 = 1.$$

Consider now the one parameter family of functions $h_t=f_1+tf_2$. For $t_1\neq t_2$ the pair h_{t_1}, h_{t_2} is linearly equivalent to the pair f_{t_1}, f_{t_2}. Hence, the family of functions

$$\Omega\Big/\frac{\omega}{\omega\wedge h_t} = \frac{\omega\wedge h_t}{\omega\wedge f_1\wedge f_2} \quad (t \text{ complex parameter})$$

consists of pairwise prime inner divisors of Ω.

Now, it follows from the canonical representation of the inner function Ω (by its zeros in the unit disc and the corresponding singular measure on the unit circle) that no family of pairwise prime inner divisors of Ω can contain more than countably many non-constant elements. Thus, for all values of the parameter t, with the possible exception of a countable set of values, we have

$$\omega\wedge h_t = \omega\wedge f_1\wedge f_2.$$

This concludes the proof.

Lemma 3. *Let $\mathcal{J}$ be a family of inner functions such that*

(i) $u_1, u_2\in\mathcal{J}$ *imply* $u_1\vee u_2\in\mathcal{J}$,

(ii) $\inf_{u\in\mathcal{J}} |u(\lambda_0)|>0$ *for some point* λ_0, $|\lambda_0|<1$.

Then $u^\vee = \bigvee_{u\in\mathcal{J}} u$ *exists and every sequence* u_n *minimizing* $|u(\lambda_0)|$ *has a subsequence converging to* $u^\vee$ *in the unit disc* $|\lambda|<1$.

For a proof, based on the Vitali—Montel theorem, *cf.* [6] or [7], Lemma 1.

2. Main Lemma

The following lemma on functions in H^∞ is related to a theorem on Hilbert space operators, proved in [7] (Theorem 1). We present here a direct proof, using elements of the proof of the operator theoretic theorem in [7]. (Although we shall only use in this paper the case when $\omega_i=\omega$ for all i, the general case is considered in view of possible further applications.)

Main Lemma. *Let* $f_{ik}\in H^\infty$, $\|f_{ik}\|_\infty\leqq M$ ($i, k=1, 2, \ldots$), *and let* ω_i ($i=1, 2, \ldots$) *be inner functions. Suppose that*

(2.1) $$\omega_i\wedge f_{i1}\wedge f_{i2}\wedge\ldots = 1 \quad (i = 1, 2, \ldots).$$

Then there exists a numerical sequence $\langle x_2, x_3, \dots\rangle$, *with* $\Sigma|x_k|$ *as small as we wish, such that*

$$\omega_i\wedge(f_{i1}+x_2f_{i2}+x_3f_{i3}+\dots) = 1 \quad (i = 1, 2, \dots). \tag{2.2}$$

Proof. a) Consider the linear transformations

$$r_i : l^1 \to H^\infty \quad (i = 1, 2, \dots)$$

defined for $x=\langle x_1, x_2, \dots\rangle \in l^1$ by

$$r_i x = \sum_1^\infty x_k f_{ik};$$

clearly, $\|r_i x\|_\infty \leqq M\,\|x\|_1$. Denote by R_i the range of r_i in H^∞.

Condition (2.1) is obviously equivalent to

$$\bigwedge_{g\in R_i} (\omega_i\wedge g) = 1 \quad (i = 1, 2, \dots)$$

and this in its turn is equivalent, by the corollary of Lemma 1, to

$$\bigvee_{g\in R_i} \frac{\omega_i}{\omega_i\wedge g} = \omega_i \quad (i = 1, 2, \dots). \tag{2.3}$$

Choose a point λ_0, $|\lambda_0|<1$, different from the zeros of the functions $\omega_1, \omega_2, \dots$; thus

$$|\omega_i(\lambda_0)| = \mu_i > 0; \tag{2.4}$$

and define

$$\nu_i = \inf_{g\in R_i} \left|\frac{\omega_i}{\omega_i\wedge g}(\lambda_0)\right| \quad (i = 1, 2, \dots). \tag{2.5}$$

Clearly, $\nu_i \geqq |\omega_i(\lambda_0)| = \mu_i$; thus the family of functions

$$\mathscr{I}_i = \left\{\frac{\omega_i}{\omega_i\wedge g} : g\in R_i\right\}$$

satisfies condition (ii) in Lemma 3. It also satisfies condition (i). For, if $g_1, g_2 \in \mathscr{I}_i$ then by linearity of l^1 and r_i we have $g_1+tg_2\in R_i$ for all values of the complex parameter t. Now, by Lemma 2 we have $\omega_i\wedge(g_1+tg_2)=(\omega_i\wedge g_1)\wedge(\omega_i\wedge g_2)$ for all t with the possible exception of countable many, and for a non-exceptional value of t we have by Lemma 1

$$\frac{\omega_i}{\omega_i\wedge g_1} \vee \frac{\omega_i}{\omega_i\wedge g_2} = \frac{\omega_i}{\omega_i\wedge(g_1+tg_2)};$$

thus condition (i) holds true for each $\mathscr{I}_i$.

Fix i and consider a sequence $\{g_n\}$ minimizing in (2.5); by virtue of Lemma 3 we can choose this sequence even so that

$$(2.6) \qquad \frac{\omega_i}{\omega_i \wedge g_n} \to \bigvee_{g \in R_i} \frac{\omega_i}{\omega_i \wedge g} \quad \text{pointwise in} \quad |\lambda| < 1, \quad \text{as} \quad n \to \infty.$$

By Lemma 1 and by (2.3), this limit equals

$$\omega_i / \bigwedge_{g \in R_i} (\omega_i \wedge g), \quad \text{i.e.} \quad \omega_i.$$

Thus we have, in particular,

$$(2.7) \qquad v_i = \lim_{n \to \infty} \left| \frac{\omega_i}{\omega_i \wedge g_n} (\lambda_0) \right| = |\omega_i(\lambda_0)| = \mu_i \quad \text{for all} \quad i.$$

b) Next we assert that the infimum v_i in (2.5) is attained for every value of i. Moreover, we assert that there exists an $x = \langle x_1, x_2, \ldots \rangle \in l^1$, independent of i, such that, for every i, the infimum v_i is attained for $g_i = r_i x$, that is,

$$\left| \frac{\omega_i}{\omega_i \wedge r_i x} (\lambda_0) \right| = v_i = \mu_i = |\omega_i(\lambda_0)|, \quad |(\omega_i \wedge r_i x)(\lambda_0)| = 1 \quad (i = 1, 2, \ldots).$$

By the maximum principle this implies $\omega_i \wedge r_i x = 1$, i.e.

$$(2.8) \qquad \omega_i \wedge (x_1 f_{i1} + x_2 f_{i2} + \ldots) = 1 \quad (i = 1, 2, \ldots).$$

To prove our assertion suppose the contrary, i.e., that for every $x \in l^1$ we have

$$\left| \frac{\omega_i}{\omega_i \wedge r_i x} (\lambda_0) \right| > \mu_i$$

for at least one subscript i, or equivalently, that l^1 is the union of the subsets

$$\sigma_{ij} = \left\{ x : x \in l^1, \left| \frac{\omega_i}{\omega_i \wedge r_i x} (\lambda_0) \right| \geqq \mu_i + \frac{1}{j} \right\} \quad (i, j = 1, 2, \ldots).$$

Let us show that each of these subsets is *closed.*

To this effect consider a sequence of vectors $x_n \in \sigma_{ij}$ (i, j fixed), converging in l^1 to a limit x; then

$$g_n = r_i x_n, \quad g = r_i x \quad \text{satisfy} \quad \|g_n - g\|^\infty \leqq \|r_i\| \, \|x_n - x\|_1 \to 0 \quad (n \to \infty)$$

and therefore we have, in particular,

$$(2.9) \qquad g_n \to g \quad \text{pointwise in} \quad |\lambda| < 1.$$

Passing, if necessary, to a subsequence we can also assume, by virtue of the Vitali—Montel theorem, that

$$(2.10) \qquad \frac{\omega_i}{\omega_i \wedge g_n} \to p, \quad \frac{g_n}{\omega_i \wedge g_n} \to q \quad \text{pointwise for } |\lambda| < 1 \quad \text{as} \quad n \to \infty,$$

where p and q are analytic for $|\lambda|<1$; clearly $\|p\|_\infty \leqq 1$ and $\|q\|_\infty \leqq M \cdot \sup \|x_n\|_1$. Note that, in particular,

$$|p(\lambda_0)| = \lim_{n\to\infty} \left| \frac{\omega_i}{\omega_i \wedge g_n}(\lambda_0) \right| \geqq \mu_i + \frac{1}{j}. \tag{2.11}$$

From (2.9) and (2.10) we infer

$$\frac{\omega_i}{\omega_i \wedge g_n} g_n \to pg, \quad \omega_i \frac{g_n}{\omega_i \wedge g_n} \to \omega_i q \quad \text{pointwise, as} \quad n \to \infty,$$

and hence, $pg = \omega_i q$, $p^\circ g^\circ = \omega_i q^\circ$, where the superscript $^\circ$ indicates inner factor. It follows that $\dfrac{\omega_i}{\omega_i \wedge g}$ is an inner divisor of p°, and hence

$$\left| \frac{\omega_i}{\omega_i \wedge g}(\lambda_0) \right| \geqq |p^\circ(\lambda_0)| \geqq |p(\lambda_0)|, \tag{2.12}$$

because the outer factor $p' = p/p^\circ$ has the same norm $\|\cdot\|_\infty$ as p, thus $|p'(\lambda)| \leqq 1$ for $|\lambda|<1$. From (2.11) and (2.12) we infer that $x \in \sigma_{ij}$: σ_{ij} is closed.

Thus l^1 is the union of the closed subsets σ_{ij} $(i, j = 1, 2, \ldots)$. By virtue of the Baire category theorem, at least one of the sets σ_{ij} must contain a ball

$$\mathscr{B} = \{x : \|x - x_0\| < \varrho\} \quad \text{in} \quad l^1,$$

that is, there exist a subscript i and a number μ_i' greater than μ_i, such that

$$\left| \frac{\omega_i}{\omega_i \wedge g}(\lambda_0) \right| \geqq \mu_i' \quad \text{for all} \quad g \in r_i \mathscr{B}. \tag{2.13}$$

On the other hand, on account of the equality $\nu_i = \mu_i$ we have $\nu_i < \mu_i'$, and therefore there exists $y \in l^1$ such that

$$\left| \frac{\omega_i}{\omega_i \wedge h}(\lambda_0) \right| < \mu_i' \quad \text{for} \quad h = r_i y. \tag{2.14}$$

Set $f_0 = r_i x_0$ and apply Lemma 2 to obtain that there exists t, $0 < t < \varrho/\|y\|_1$, such that

$$\frac{\omega_i}{\omega_i \wedge (f_0 + th)} = \frac{\omega_i}{\omega_i \wedge f_0} \vee \frac{\omega_i}{\omega_i \wedge h}. \tag{2.15}$$

As we have $f_0 + th = r_i(x_0 + ty) \in r_i \mathscr{B}$, the function at the left hand side of (2.15) has, by (2.13), absolute value $\geqq \mu_i'$. The function at the right hand side of (2.15), being an inner multiple of the function $\dfrac{\omega_i}{\omega_i \wedge h}$, is majorized in absolute value by the latter function everywhere in the unit disc; thus by (2.14) the function at the right hand side of (2.15) has at the point λ_0 absolute value $< \mu_i'$.

So we have arrived at a contradiction. This proves our assertion stated at the beginning of part b) of the proof, namely that there exists an $x \in l^1$ satisfying (2.8).

c) In the last step of our proof we shall again refer to the (Sherman) Lemma 2. Let $x = \langle x_1, x_2, \ldots \rangle \in l^1$ be any vector for which (2.8) holds, i.e. such that

$$\omega_i \wedge \varphi_i = 1 \quad \text{for} \quad \varphi_i = x_1 f_{i1} + x_2 f_{i2} + \ldots \quad (i = 1, 2, \ldots).$$

Then by Lemma 2 we also have

$$\omega_i \wedge (\varphi_i + t f_{i1}) = \omega_i \wedge \varphi_i \wedge f_{i1} = 1 \wedge f_{i1} = 1 \quad (i = 1, 2, \ldots)$$

for all values of the complex parameter t, with the possible exception of countably many values. Given $\varepsilon > 0$, if we choose t not exceptional for any i, and moreover different from $-x_1$ and sufficiently large, we will have

$$\omega_i \wedge (f_{i1} + x_2' f_{i2} + x_3' f_{i3} + \ldots) = 1 \quad (i = 1, 2, \ldots),$$

with $x_k' = x_k/(x_1 + t)$ and $\sum_2^\infty |x_k'| < \varepsilon$.

This completes the proof of the Main Lemma.

When referring to the Main Lemma we shall mean its following direct corollary:

Let a_{ik} be a (finite, semi-finite, or infinite) rectangular matrix over H^∞, with $\|a_{ik}\|_\infty \leqq M$, and let ω be an inner function. Then there exists a numerical sequence $\langle x_2, x_3, \ldots \rangle$, with $\sum |x_k|$ as small as we wish, such that, for every value of i, we have

$$a_{i1} + x_2 a_{i2} + x_3 a_{i3} + \ldots = h_i \cdot (a_{i1} \wedge a_{i2} \wedge a_{i3} \wedge \ldots),$$

where $h_i \in H^\infty$, $h_i \wedge \omega = 1$.

3. Quasi-equivalence and diagonalization of matrices over H^∞

1. Let $\mathscr{M}(n, m)$ $(1 \leqq n \leqq \infty,\ 1 \leqq m \leqq \infty)$ be the set of $n \times m$ matrices $A = [a_{ik}]$ over H^∞, for which

$$\sum_i \Big| \sum_k \xi_k a_{ik}(\lambda) \Big|^2 \leqq M^2 \sum_k |\xi_k|^2 \quad (M \geqq 0) \tag{3.1}$$

holds for $|\lambda| < 1$ and for any square-summable sequence of complex numbers ξ_k, i.e. whose values $A(\lambda)$ $(|\lambda| < 1)$ are operators from (complex euclidean) m-space E_m into n-space E_n, bounded by the constant M,

$$\|A\|_\infty = \sup_{|\lambda| < 1} \|A(\lambda)\| \leqq M.$$

By $\mathscr{N}(n)$ $(1 \leqq n \leqq \infty)$ we denote the set of matrices $X = X(\lambda)$ in $\mathscr{M}(n, n)$ for which $X(\lambda)^{-1}$ exists $(|\lambda| < 1)$ and also belongs to $\mathscr{M}(n, n)$.

Finally, for a given inner function ω we denote by $\mathcal{N}_\omega(n)$ the set of matrices $X\in\mathcal{M}(n, n)$ which have a scalar multiple φ prime to ω, that is, for which there exists $X^a\in\mathcal{M}(n, n)$ such that

$$X^a X = XX^a = \varphi\cdot I_n, \quad \varphi\in H^\infty, \quad \varphi\not\equiv 0, \quad \varphi\wedge\omega = 1$$

(I_n is the unit matrix of order n).

It is clear that $\mathcal{N}(n)\subset\mathcal{N}_\omega(n)$, and that a finite product of elements of $\mathcal{N}_\omega(n)$ also belongs to $\mathcal{N}_\omega(n)$.

Let $A, B\in\mathcal{M}(n, m)$. We call A, B *equivalent* if there exist matrices $X\in\mathcal{N}(n)$, $Y\in\mathcal{N}(m)$ such that

$$XA = BY, \tag{3.2}$$

and we call them *ω-equivalent* if there exist $X\in\mathcal{N}_\omega(n)$, $Y\in\mathcal{N}_\omega(m)$ satisfying (3.2).

Equivalence implies ω-equivalence, but not conversely. Both are *symmetric*. This is obvious for equivalence, while for ω-equivalence it can be shown as follows: If $X^aX=XX^a=\varphi\cdot I_n$, $Y^aY=YY^a=\psi\cdot I_m$, $\varphi\wedge\omega=1$, $\psi\wedge\omega=1$, then (3.2) implies:

$$A\cdot\varphi Y^a = \varphi AY^a = X^aXAY^a = X^aBYY^a = X^aB\psi I_m = \psi X^a\cdot B,$$

where $\varphi Y^a\in\mathcal{N}_\omega(m)$ and $\psi X^a\in\mathcal{N}_\omega(n)$ because

$$\varphi Y^a\cdot Y = \varphi\psi I_m = Y\cdot\varphi Y^a, \quad \varphi\psi\wedge\omega = 1,$$

and similarly for ψX^a. — Clearly, both kinds of equivalence are *transitive*.

In case A, B are ω-equivalent for *every* inner ω, they are called *quasi-equivalent*.

These concepts were introduced by NORDGREN [3]; see also SZŰCS [10].

"Determinant divisors" $\mathcal{D}_k$ and "invariant factors" $\mathcal{E}_k$ of a matrix $A\in\mathcal{M}(n, m)$ are defined, for all (finite) integers k, $1\leq k\leq\min\{n, m\}$, as in the classical case, namely:

$\mathcal{D}_k=\wedge\det A^{(k)}$, where $A^{(k)}$ runs over the set of all square submatrices of A of order k (thus $\mathcal{D}_k=0$ iff all these submatrices have determinant 0, and $\mathcal{D}_k=0$ implies $\mathcal{D}_{k+1}=0$);

$\mathcal{E}_k=\mathcal{D}_k/\mathcal{D}_{k-1}$, with the conventions $\mathcal{D}_0=1$ and $\mathcal{E}_k=0$ if $\mathcal{D}_{k-1}=0$.

Lemma 4. *If $A, B\in\mathcal{M}(n, m)$ are ω-equivalent, then*

$$\mathcal{D}_k(A)|\alpha_k\mathcal{D}_k(B), \quad \mathcal{D}_k(B)|\beta_k\mathcal{D}_k(A) \quad (k = 1, 2, \ldots), \tag{3.3}$$

where α_k, β_k are inner functions prime to ω. If A, B are even quasi-equivalent, then

$$\mathcal{D}_k(A) = \mathcal{D}_k(B) \quad (k = 1, 2, \ldots). \tag{3.4}$$

Proof. Suppose $X\in\mathcal{N}_\omega(n)$ and $Y\in\mathcal{N}_\omega(m)$ satisfy (3.2). If φ and ψ are their corresponding scalar multiples, prime to ω, then we deduce from (3.2) that

$$X^aBY = \varphi\cdot A, \quad XAY^a = \psi\cdot B. \tag{3.5}$$

As the Cauchy—Binet multiplication rule for minors extends to the present case, we get from (3.5), first, that $\mathcal{D}_k(A)=0$ iff $\mathcal{D}_k(B)=0$. Next, if $\mathcal{D}_k(A)$ and $\mathcal{D}_k(B)$ are non-zero, and therefore inner functions, we deduce that

$$(3.6) \qquad \mathcal{D}_k(B)\,|\,\varphi^k\mathcal{D}_k(A) \quad \text{and} \quad \mathcal{D}_k(A)\,|\,\psi^k\mathcal{D}_k(B),$$ [1]

and we have only to observe that φ^k and ψ^k are also prime to ω.

If A, B are quasi-equivalent and if for fixed k such that $\mathcal{D}_k(A)$ and $\mathcal{D}_k(B)$ are non-zero we choose $\omega=\mathcal{D}_k(A)\,\mathcal{D}_k(B)$, then φ^k and ψ^k are prime to $\mathcal{D}_k(A)$ and $\mathcal{D}_k(B)$, so that (3.6) implies $\mathcal{D}_k(B)|\mathcal{D}_k(A)$ and $\mathcal{D}_k(A)|\mathcal{D}_k(B)$, i.e. (3.4).

2. For later use we introduce the following notations:

Let $u=\langle 0, u_1, u_2, \ldots\rangle$ be a sequence of length n (finite or infinite) of functions in H^∞ satisfying

$$\|u\|_\infty = \sup_{|\lambda|<1} \Big(\sum_k |u_k(\lambda)|^2\Big)^{1/2} < \infty$$

and let $C(u)$ and $R(u)$ respectively denote the square matrices whose first column or the first row is given by this sequence and all other entries are 0. These matrices obviously belong to $\mathcal{M}(n, m)$, with $\|C(u)\|_\infty=\|R(u)\|_\infty=\|u\|_\infty$; moreover the matrices $I\pm C(u)$, $I\pm R(u)$ belong to $\mathcal{N}(n)$ because

$$(I\pm C(u))(I\mp C(u)) = I, \quad (I\pm R(u))(I\mp R(u)) = I.$$

Every diagonal (square) matrix $D=\operatorname{diag}(w_1, w_2, \ldots)$ of order n whose diagonal entries are inner functions and have a common inner multiple w, belongs to $\mathcal{N}_\omega(n)$ for every ω such that $w\wedge\omega=1$; indeed,

$$\|D\|_\infty = 1 \quad \text{and} \quad D^a D = DD^a = wI, \quad \text{where} \quad D^a = \operatorname{diag}\left(\frac{w}{w_1}, \frac{w}{w_2}, \ldots\right).$$

Finally, observe that if A_0, A_1 are ω-equivalent to A_0', A_1', then $A=A_0\oplus A_1$ is ω-equivalent to $A'=A_0'\oplus A_1'$. Indeed, if X_0, Y_0 and X_1, Y_1 are operators for A_0, A_0' and A_1, A_1', with the respective scalar multiples φ_0, ψ_0, and φ_1, ψ_1, prime to ω, then $X=X_0\oplus X_1$, $Y=Y_0\oplus Y_1$ will correspond to the pair A, A', and setting

$$X^a = \varphi_1 X_0^a \oplus \varphi_0 X_1^a, \quad Y^a = \psi_1 Y_0^a \oplus \psi_0 Y_1^a$$

we see that X, Y have the scalar multiples $\varphi_0\cdot\varphi_1$, and $\psi_0\cdot\psi_1$, respectively, which are also prime to ω.

3. We are now able to prove:

---

[1] Here we use the fact that if $\{u_\alpha\}$ is a system of inner functions and f is a function in L^∞ such that $fu_\alpha\in H^\infty$ for all α then $f\cdot\bigwedge_\alpha u_\alpha\in H^\infty$; *cf.* Proposition III. 1. 5 in [4]. This fact implies, namely, that if w is inner, if v is in H^∞, and if $w|vu_\alpha$ for all α, then $w\big|(v\cdot\bigwedge_\alpha u_\alpha)$ (set $f=\bar{w}v$).

Theorem 1. *Let $A=[a_{ik}]\in\mathcal{M}(n, m)$, $1\leqq n\leqq\infty$, $1\leqq m\leqq\infty$, and let r be an integer, $1\leqq r\leqq\min\{n, m\}$. Then, for any given inner function ω, A is ω-equivalent to a matrix of the form*

$$\operatorname{diag}[\mathcal{E}_1(A), \dots, \mathcal{E}_r(A), A_r]$$

where $A_r\in\mathcal{M}(n_r', m_r)$ $(r+n_r=n,\ r+m_r=m)$, and we have

$$\mathcal{E}_1(A)\,|\,\mathcal{E}_2(A)\,|\dots|\,\mathcal{E}_r(A)\,|\,A_r.$$

Proof. The case $A=0$ being trivial we can assume $A\neq 0$ so that $\mathcal{D}_1(A)$ is an inner function. From (3.1) it follows, in particular, that $\|a_{ik}\|_\infty\leqq\|A\|_\infty$ $(=M)$.

Denote by ω_r the product of the given inner function ω by the non-zero (and hence, inner) terms of the sequence $\mathcal{D}_1(A), \dots, \mathcal{D}_r(A)$. Then any $h\in H^\infty$ prime to ω_r is prime to ω as well as to each of these determinant divisors of A.

By virtue of the Main Lemma there exists a numerical sequence $\langle x_1, x_2, \dots\rangle$ of length m, with $x_1=1$ and $\sum_2^m |x_k|$ as small as we wish, such that

$$(3.7)\qquad (a_i =)\sum_{k=1}^m x_k a_{ik} = h_i\cdot\bigwedge_{k=1}^m a_{ik},\quad h_i\in H^\infty,\quad h_i\wedge\omega_r = 1\quad (i = 1, 2, \dots, n).$$

Then

$$\sum_i |a_i|^2 = \sum_i \Big|\sum_k x_k a_{ik}\Big|^2 \leqq M^2\sum_k |x_k|^2 \leqq M'^2$$

for some M' (as close to M as we wish) and for all λ, $|\lambda|<1$. Hence, $\|a_i\|_\infty\leqq M'$.

Appling the Main Lemma again we can choose a numerical sequence $\langle y_1, y_2, \dots\rangle$ of length n, with $y_1=1$ and $\sum_2^m |y_i|$ as small as we wish, such that

$$\sum_{i=1}^n y_i a_i = h'\cdot\bigwedge_{i=1}^n a_i,\quad h'\in H^\infty,\quad h'\wedge\omega_r = 1.$$

Observe that there is an inner function h'' such that $h''\wedge\omega_r=1$ and

$$\bigwedge_i a_i = \bigwedge_i\Big(h_i\cdot\bigwedge_k a_{ik}\Big) = h''\cdot\bigwedge_{i,k} a_{ik}.$$ [2)]

We have, therefore,

$$(3.8)\qquad \sum_i y_i a_i = h\cdot\mathcal{D}_1(A),\quad\text{where}\quad h = h'h'',\quad h\wedge\omega_r = 1.$$

[2)] Set $b_i=\bigwedge_k a_{ik}$ and $b=\bigwedge_i b_i$: then $b=\mathcal{D}_1(A)$ and $\bigwedge_i (b_i/b)=1$. We have

$$\bigwedge_i a_i = \bigwedge_i (h_i b_i) = \left(\bigwedge_i\left(h_i\frac{b_i}{b}\right)\right)\cdot b \equiv h''\cdot b.$$

Since $h_i\wedge\omega_r=1$, we have

$$h''\wedge\omega_r = \bigwedge_i\left(\left(h_i\frac{b_i}{b}\right)\wedge\omega_r\right) = \bigwedge_i\left(\frac{b_i}{b}\wedge\omega_r\right) = 1.$$

The author is indepted to Prof. T. Ando for this proof and also for some other useful remarks he has made when reading the manuscript.

Form the matrices $C_m(x)$ and $R_n(y)$ associated with the sequences $x=\langle 0, x_2, x_3, \ldots\rangle$ and $y=\langle 0, y_2, y_3, \ldots\rangle$ according to Subsection 2. From (3.7) and (3.8) we deduce that the matrix

$$A' = [a'_{ik}] = (I_n+R_n(y))A(I_m+C_m(x)) \tag{3.9}$$

has the leading entry $a'_{11}=h\cdot\mathscr{D}_1(A)$, while $a'_{ik}=a_{ik}$ for $i, k\geqq 2$. As $I+R_n$ and $I+C_m$ are invertible, A is equivalent to A', and therefore, by Lemma 4,

$$\mathscr{D}_k(A) = \mathscr{D}_k(A') \quad \text{for every} \quad k,$$

in particular $\mathscr{D}_1(A)|A'$.

Now, we set

$$A'' = \begin{bmatrix} \mathscr{D}_1(A) & a'_{12} & a'_{13} & \cdots \\ a'_{21} & ha_{22} & ha_{32} & \cdots \\ a'_{31} & ha_{32} & ha_{33} & \cdots \\ \vdots & \vdots & & \ddots \end{bmatrix}$$

and observe that

$$A'\cdot\operatorname{diag}(1, h, h, \ldots)_m = \operatorname{diag}(h, 1, 1, \ldots)_n\cdot A'';$$

as a consequence, A' is ω_r-equivalent to A''.

Next, form the matrices $C_n(u)$ and $R_m(v)$ associated with the sequences $u=\langle 0, u_2, u_3, \ldots\rangle$ and $v=\langle 0, v_2, v_3, \ldots\rangle$, where $u_i=a'_{i1}/\mathscr{D}_1(A)$ and $v_k=a'_{1k}/\mathscr{D}_1(A)$. Because $\mathscr{D}_1(A)$ is an inner function, we have $\|u\|_\infty=\|a'_{.1}\|_\infty$ and $\|v\|_\infty=\|a'_{1.}\|_\infty$, where $a'_{.1}=\langle 0, a'_{21}, a'_{31}, \ldots\rangle$ and $a'_{1.}=\langle 0, a'_{12}, a'_{13}, \ldots\rangle$. Hence, the matrices $I_n-C_n(u)$ and $I_m-R_m(v)$ belong to $\mathscr{N}(n)$ and $\mathscr{N}(m)$, respectively, so that A'' is equivalent to

$$A''' = (I_n-C_n(u)A''(I_m-R_m(v)).$$

This matrix has the form

$$A''' = \begin{bmatrix} \mathscr{D}_1(A) & 0 \\ 0 & A_1 \end{bmatrix} (= \operatorname{diag}[\mathscr{D}_1(A), A_1]),$$

where $A_1\in\mathscr{M}(n_1, m_1)$ $(n=1+n_1, m=1+m_1)$. Note that $\mathscr{D}_1(A)$, which divides A', also divides A'' (see the explicit form of A'') and therefore will divide A''' as well. We conclude that A is ω_r-equivalent to diag $[\mathscr{D}_1(A), A_1]$, and $\mathscr{D}_1(A)|A_1$.

Now apply the same argument to A_1 in place of A, and continue this procedure r times. Recalling the last remark in Subsection 2 we conclude that A is ω_r-equivalent to a matrix of the form

$$A^{(r)} = \operatorname{diag}(\delta_1, \delta_2, \ldots, \delta_r, A_r), \tag{3.10}$$

where $A_r\in\mathscr{M}(n_r, m_r)$ $(r+n_r=n, r+m_r=m)$, and

$$\delta_1\,|\,\delta_2\,|\ldots|\,\delta_r\,|\,A_r, \quad \text{each } \delta_k \text{ inner or zero.} \tag{3.11}$$

The concluding arguments are essentially the same as in [3], p.308. By (3.6), ω_r-equivalence of A and $A^{(r)}$ implies

$$\mathscr{D}_k(A)\,|\,\varphi_k\cdot\mathscr{D}_k(A^{(r)}), \quad \mathscr{D}_k(A^{(r)})\,|\,\psi_k\cdot\mathscr{D}_k(A), \quad \varphi_k, \psi_k \text{ prime to } \omega_r.$$

Since φ_k is then prime to $\mathscr{D}_k(A)$ for $k=1, \ldots, r$, we infer $\mathscr{D}_k(A) \mid \mathscr{D}_k(A^{(r)})$, and hence $\mathscr{D}_k(A^{(r)})=\alpha_k \cdot \mathscr{D}_k(A)$ with α_k inner, $k=1, \ldots, r$. Thus $\alpha_k \cdot \mathscr{D}_k(A) \mid \psi_k \cdot \mathscr{D}_k(A)$, and therefore, $\alpha_k | \psi_k$ whenever $\mathscr{D}_k(A) \neq 0$.

Let j denote the largest among the integers $k=1, 2, \ldots, r$ for which $\mathscr{D}_k(A)$ is non-zero. Then we have for $k=1, \ldots, j$:

$$\mathscr{D}_k(A^{(r)}) = \alpha_k \cdot \mathscr{D}_k(A), \quad \alpha_k \text{ inner}, \quad \alpha_k \wedge \omega_r = 1, \tag{3.12}$$

and hence,

$$\alpha_{k-1} \cdot \mathscr{D}_{k-1}(A) \mid \alpha_k \cdot \mathscr{D}_k(A), \quad \text{with} \quad \alpha_0 = 1.$$

Now, α_{k-1} is prime to $\mathscr{D}_k(A)$ so we infer $\alpha_{k-1} | \alpha_k$, i.e. α_k/α_{k-1} is inner. From (3.12) we have

$$\mathscr{E}_k(A^{(r)}) = (\alpha_k/\alpha_{k-1}) \mathscr{E}_k(A) \quad (k = 1, \ldots, j). \tag{3.13}$$

On the other hand, it readily follows from (3.10) and (3.11) that $\mathscr{E}_k(A^{(r)})=\delta_k$ $(k=1, \ldots, r)$, and therefore, by (3.13) and (3.11),

$$(\alpha_k/\alpha_{k-1}) \mathscr{E}_k(A) \mid (\alpha_{k+1}/\alpha_k) \mathscr{E}_{k+1}(A) \quad (k = 1, \ldots, j-1). \tag{3.14}$$

Since α_{k+1} is prime to ω_r, α_{k+1}/α_k is prime to $\mathscr{E}_k(A)$. Therefore, (3.14) implies

$$\mathscr{E}_k(A) \mid \mathscr{E}_{k+1}(A)$$

for $k=1, \ldots, j-1$ (and then for all k).

Finally, combining (3.10) and (3.14) we see that

$$A^{(r)} = Z \cdot \operatorname{diag}[\mathscr{E}_1(A), \ldots, \mathscr{E}_r(A), A_r],$$

where $Z=\operatorname{diag}[\alpha_1/\alpha_0, \alpha_2/\alpha_1, \ldots, \alpha_j/\alpha_{j-1}, 1, 1, \ldots]$ (n terms); note that Z has α_j as a scalar multiple, $\alpha_j \wedge \omega_r = 1$. Also note that $(\alpha_j/\alpha_{j-1}) \mathscr{E}_j(A) = \mathscr{E}_j(A^{(r)}) = \delta_j \mid A_r$, and hence $\mathscr{E}_k(A) \mid A_r$ for $k=1, \ldots, r$.

This concludes the proof of Theorem 1.

4. Consider now the case of $A \in \mathscr{M}(n, m)$, where at least one of n, m is finite; it is no restriction of generality to suppose that m is finite and $m \leqq n \leqq \infty$.

Applying Theorem 1 with $r=m$ we obtain that A is ω-equivalent to the diagonal $n \times m$ matrix formed from the invariant factors of A. Now, this matrix does not depend on the choice of ω. Therefore, we have:

Theorem 2. *Every matrix $A \in \mathscr{M}(n, m)$, with m finite and with $m \leqq n \leqq \infty$, is quasi-equivalent to the diagonal $n \times m$ matrix*

$$\operatorname{diag}[\mathscr{E}_1(A), \ldots, \mathscr{E}_m(A)],$$

and we have $\mathscr{E}_1(A) \mid \mathscr{E}_2(A) \mid \ldots \mid \mathscr{E}_m(A)$.

4. Jordan models of operators of class $C_{\cdot 0}$

1. Let A, B be $n\times m$ matrix valued *inner* functions over H^∞, [3]) with *m finite* and *n possibly infinite*, $m\leqq n\leqq\infty$, and suppose A, B are quasi-equivalent. The condition for A, B to be inner implies that all determinant divisors are non-zero; in particular,

$$\omega = \mathscr{D}_m(A) = \mathscr{D}_m(B)$$

is a (scalar valued) inner function.

Choose Φ, $\Phi^a\in\mathscr{M}(n, n)$ and Ψ, $\Psi^a\in\mathscr{M}(m, m)$ such that

(4.1) $\Phi A = B\Psi, \quad \Phi^a\Phi = \Phi\Phi^a = \varphi I_n, \quad \Psi^a\Psi = \Psi\Psi^a = \psi I_m, \quad \varphi, \psi$ prime to ω.

Let $S(A)$, $S(B)$ be the operators defined on the Hilbert spaces $\mathfrak{H}(A)=H^2_n\ominus AH^2_m$, $\mathfrak{H}(B)=H^2_n\ominus BH^2_m$ [4]) by

$$S(A)u = P_{\mathfrak{H}(A)}(\chi u) \quad \text{for } u\in\mathfrak{H}(A), \qquad S(B)u = P_{\mathfrak{H}(B)}(\chi u) \quad \text{for } u\in\mathfrak{H}(B),$$

and set

(4.2) $$Xu = P_{\mathfrak{H}(B)}\Phi u \quad \text{for} \quad u\in\mathfrak{H}(A).$$

Then the operator X: $\mathfrak{H}(A)\to\mathfrak{H}(B)$ satisfies

(4.3) $$S(B)X = XS(A),$$

and is *injective*. These facts follow by the same arguments as in [8], Sec. 2, by giving the role of $\Psi^\wedge$ and $\det\Psi$ to Ψ^a and ψ, respectively.

Using the relation $\Phi A=B\Psi$ we get

(4.4) $$X\mathfrak{H}(A) = P_{\mathfrak{H}(B)}\Phi\mathfrak{H}(A) = P_{\mathfrak{H}(B)}\Phi H^2_n.$$

Since $\Phi H^2_n\supset\Phi\Phi^a H^2_n=\varphi H^2_n$, (4.4) implies

(4.5) $$X\mathfrak{H}(A)\supset P_{\mathfrak{H}(B)}(\varphi H^2_n).$$

Set now $\omega_1=\omega\cdot\varphi^\circ$, φ° being the inner factor of φ, and choose Φ_1, Ψ_1, etc., correspondingly. So we get X_1 such that

$(4.5)_1$ $$X_1\mathfrak{H}(A)\supset P_{\mathfrak{H}(B)}(\varphi_1 H^2_n).$$

As φ_1 is prime to φ, by Beurling's theorem φH^2_n and $\varphi_1 H^2_n$ together span H^2_n. As a result, the *ranges of* X *and* X_1 *together span* $\mathfrak{H}(B)$.

2. In some special cases (but not always, *cf.* [8], Sec. 3) we can choose X such that its range alone spans $\mathfrak{H}(B)$; i.e. that X be a *quasi-affinity*. Such is the case if $n=m(<\infty)$, or more generally, if $B=\begin{bmatrix}B_1\\0\end{bmatrix}$, where B_1 is a square matrix of order m, and 0 is the $l\times m$ zero matrix, where

$$n = m+l, \quad 0\leqq l\leqq\infty.$$

[3]) That is, A and B are isometry valued a. e. on the unit circle.

[4]) $H^2_n=H^2(E_n)$ is the Hardy—Hilbert space of E_n-vector valued analytic functions in the unit disc; and $\chi(\lambda)\equiv\lambda$.

For l finite, *cf.* [2]. The following generalization of the proof given in [2] applies to l infinite as well.

Choose Φ, Ψ to satisfy (4.1) with $\omega=\mathcal{D}_m(B)=\det B_1$ and partition the matrices Φ and Φ^a in the form

$$\Phi=\underbrace{\begin{bmatrix}\Phi_1\\ \Phi_2\end{bmatrix}}_{n}\begin{matrix}\}m\\ \}l\end{matrix},\quad \Phi^a=[\underbrace{\Omega_1}_{m}\ \underbrace{\Omega_2}_{l}]\}n.$$

Equations (4.1) are equivalent to the following ones:

$$(4.6)\quad\begin{cases}\Phi_1A=B_1\Psi,\quad \Phi_2A=0,\\ \Omega_1\Phi_1+\Omega_2\Phi_2=\varphi I_n,\ \Phi_1\Omega_1=\varphi I_m,\ \Phi_2\Omega_2=\varphi I_l,\ \Phi_1\Omega_2=0,\ \Phi_2\Omega_1=0,\ \varphi\wedge\omega=1,\\ \Psi^a\Psi=\Psi\Psi^a=\psi I_m,\quad \psi\wedge\omega=1.\end{cases}$$

Clearly, $\Phi_2\in\mathcal{M}(l,n)$, with $\|\Phi_2\|_\infty\leqq\|\Phi\|_\infty$. Let

$$\Phi_2=\Phi_2^\circ\Phi_2'$$

be the canonical factorization of the bounded analytic function $\{E_n, E_l, \Phi_2(\lambda)\}$ into its outer factor $\{E_n, \mathfrak{F}, \Phi_2'(\lambda)\}$ and inner factor $\{\mathfrak{F}, E_l, \Phi_2^\circ(\lambda)\}$, where $\mathfrak{F}$ is some auxiliary Hilbert space; *cf.* [4], Chapter V. By taking $d=\dim\mathfrak{F}$ we can assume $\mathfrak{F}=E_d$; then

$$\Phi_2'\in\mathcal{M}(d,n)\quad\text{and}\quad \Phi_2^\circ\in\mathcal{M}(l,d).$$

As Φ_2' is outer, $\Phi_2'H_n^2$ is dense in H_d^2, and therefore $\Phi_2H_n^2=\Phi_2^\circ\Phi_2'H_n^2$ is dense in $\Phi_2^\circ H_d^2$. On the other hand we have, by (4.6), $\Phi_2H_n^2\supset\Phi_2\Omega_2H_l^2=\varphi H_l^2$. Therefore,

$$(4.7)\qquad \Phi_2^\circ H_d^2\supset\varphi^\circ H_l^2\quad(\varphi^\circ\text{ is the inner factor of }\varphi).$$

On account of this inclusion, for every $u\in H_l^2$ there exists a $v\in H_d^2$ such that $\Phi_2^\circ v=\varphi^\circ u$; the map $u\to v$ defines an isometry $W\colon H_l^2\to H_d^2$ which intertwines the natural unilateral shifts on these spaces, i.e.

$$(4.8)\qquad S_dW=WS_l.$$

This implies that $l\leqq d$; *cf.* [8], Theorem 5/6.

The inclusion

$$(4.9)\qquad \Phi_2^\circ H_d^2=\Phi_2^\circ\overline{\Phi_2'H_n^2}=\overline{\Phi_2H_n^2}\subset H_l^2$$

shows that Φ_2° is an isometry from H_d^2 into H_l^2, which obviously intertwines S_d and S_l in the reverse order, and therefore, $d\leqq l$.

Thus $d=l$, and hence $\Phi_2'\in\mathcal{M}(l,n)$, $\Phi_2^\circ\in\mathcal{M}(l,l)$, and $\Omega_2\Phi_2^\circ\in\mathcal{M}(m,l)$. Therefore, both

$$\tilde\Phi=\begin{bmatrix}\Phi_1\\ \Phi_2'\end{bmatrix}\quad\text{and}\quad\tilde\Phi^a=[\Omega_1\ \ \Omega_2\Phi_2^\circ]$$

are in $\mathcal{M}(n, n)$. Moreover, it easily follows from (4.6) that (4.1) holds true for $\tilde{\Phi}$, $\tilde{\Phi}^a$ in place of Φ, Φ^a. Indeed, we have, e.g.

$$\Omega_1\Phi_1+\Omega_2\Phi_2^\circ\cdot\Phi_2'=\Omega_1\Phi_1+\Omega_2\Phi_2=\varphi I_n,$$

$$\Phi_2'\cdot\Omega_2\Phi_2^\circ=(\Phi_2^\circ)^*\cdot\Phi_2\Omega_2\cdot\Phi_2^\circ=(\Phi_2^\circ)^*\cdot\varphi\cdot\Phi_2^\circ=\varphi I_l, \quad \text{etc.}$$

The rest of the argument is similar to the one in [2]. We regard H_n^2 as the direct sum $H_m^2\oplus H_l^2$ and set $\mathfrak{N}=\tilde{\Phi}H_n^2+BH_m^2$. Since we have

$$\tilde{\Phi}H_n^2\supset\tilde{\Phi}\tilde{\Phi}^aH_n^2=\varphi H_n^2\supset\varphi H_m^2\oplus 0 \quad\text{and}\quad BH_m^2=B_1H_m^2\oplus\{0\}\supset(\det B_1)H_m^2\oplus\{0\},$$

and since φ is prime to $\det B_1$, it follows from Beurling's theorem that

$$\overline{\mathfrak{N}}\supset H_m^2\oplus\{0\}.$$

From the fact that $\mathfrak{N}$ includes $\tilde{\Phi}H_n^2=\{\Phi_1u\oplus\Phi_2'u\colon u\in H_n^2\}$ it now follows that $\overline{\mathfrak{N}}$ also includes $\{0\}\oplus\Phi_2'H_m^2$, and hence,

$$\overline{\mathfrak{N}}\supset\{0\}\oplus H_l^2.$$

Thus, $\overline{\mathfrak{N}}=H_n^2$.

Now, for the operator $\tilde{X}$ associated with $\tilde{\Phi}$ in the sense of (4.1) we have, by (4.4), $\tilde{X}\mathfrak{H}(A)=P_{\mathfrak{H}(B)}\mathfrak{N}$, and hence the closure of the range of $\tilde{X}$ equals $P_{\mathfrak{H}(B)}H_n^2$, i.e. $\mathfrak{H}(B)$. Thus $\tilde{X}$ is a quasi-affinity.

3. Applying Theorem 2 and the above results to the characteristic matrix function $\Theta\in\mathcal{M}(n, m)$ of a contraction T on $\mathfrak{H}$, of class $C_{\cdot 0}$, with defect indices

$$\dim[I-T^*T)^{1/2}\mathfrak{H}]^-=m, \quad \dim[(I-TT^*)^{1/2}\mathfrak{H}]^-=n,$$

where $m<\infty$ while $(m\leqq)n\leqq\infty$, and to the diagonal $n\times m$ matrix formed by $e_k=\mathscr{E}_k(\Theta)$ $(k=1, \ldots, m)$, we conclude as in [8] and [2]:

Theorem 3. *The "Jordan operator" $J=S(e_m)\oplus\ldots\oplus S(e_1)\oplus S_l$ on $\mathfrak{H}_J=\mathfrak{H}(e_m)\oplus\ldots\oplus\mathfrak{H}(e_1)\oplus H_l^2$ $(l=n-m)$ is completely injection-similar to T. More precisely, there exist injections*

$$X:\mathfrak{H}\to\mathfrak{H}_J, \quad Y_i:\mathfrak{H}_J\to\mathfrak{H} \quad (i=1,2)$$

such that

$$JX=XT, \quad TY_i=Y_iJ \quad (i=1,2),$$

and the range of X is dense in $\mathfrak{H}_J$ while the ranges of Y_1 and Y_2 together span $\mathfrak{H}$.

The problem concerning uniqueness of the model can be dealt with as in [8].

Problems. 1. In [9], the existence of a unique quasi-similar Jordan model $\bigoplus_k S(m_k)$ (m_k inner, $m_{k+1}|m_k$, $k=1, 2, \ldots$) has been proved for every contraction $T\in C_0$ with minimal function $m_T=m_1$. In the general case the relation of the functions m_k to the invariant factors of the characteristic matrix of T remains to be elucidated.

2. It also remains to be investigated under which conditions Theorem 1 can be sharpened so that quasi-equivalence is established to the diagonal matrix formed only by the invariant factors.

References

[1] B. MOORE, III — E. A. NORDGEN, On quasi-equivalence and quasi-similarity, *Acta Sci. Math.*, **34** (1973), 311—316.

[2] B. MOORE, III — E. A. NORDGREN, Remark on the Jordan model for contractions of class $C_{\cdot 0}$, *Acta Sci. Math.*, **37** (1975), 307—312.

[3] E. A. NORDGREN, On quasi-equivalence of matrices over H^∞, *Acta Sci. Math.*, **34** (1973), 301—310.

[4] B. SZ.-NAGY—C. FOIAŞ, *Harmonic analysis of operators on Hilbert space*, North Holland/ Akadémiai Kiadó (Amsterdam/Budapest, 1970).

[5] 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.

[6] B. SZ.-NAGY—C. FOIAŞ, Local characterization of operators of class C_0, *J. Functional Anal.*, **8** (1971), 76—81.

[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] 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.

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

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Acta Sci. Math., 38 (1976), 239—251

On differentiation

LEON W. COHEN

Homage to the memory of F. Riesz

The ideas developed by F. RIESZ in his proof [1] that a monotonic function is almost everywhere differentiable are used here to prove:

Theorem 1. *If f and φ increase on an open interval (a, b) then $df/d\varphi$ is finite except on a subset of (a, b) of μ_φ-measure zero.*

Theorem 2. *If the increasing function f is absolutely continuous relative to the increasing function φ on (a, b) then*

$$f(b-)-f(a+) = \int_{(a,b)} df/d\varphi \, d\mu_\varphi. \text{ [1]}$$

This closes a gap left by the Radon—Nikodym theorem. The obvious definition

$$(1) \qquad df/d\varphi|_x = \lim_{y\to x} \frac{f(y)-f(x)}{\varphi(y)-\varphi(x)}$$

can not be used for Theorem 1 as the following example shows. Let $f(x)$ be -1 for $x<0$, 0 for $x=0$, 1 for $x>0$, and let $\varphi(x)$ be -1 for $x<0$ and 1 for $x\geqq 0$. Then $df/d\varphi|_0$, by (1), does not exist and $\mu_\varphi(\{0\})=2$. However

$$\lim_{h\downarrow 0,\, k\uparrow 0} \frac{f(h)-f(k)}{\varphi(h)-\varphi(k)} = 1.$$

This suggests that $df/d\varphi$ be defined as the common value, if it exists, of the upper and lower derivates of f relative to φ.

For any real function f on (a, b) and all $I=(u, v)\subset(a, b)$ let $f(I)=f(v)-f(u)$.

Received September 19, 1975.

[1]) These theorems seem to by be a part of the oral mathematical tradition but diligent inquiry by the author did not disclose any written record of their proofs.

The author is indebted to the referee for refinements and improvements of his manuscript.

Definition. Let f and φ be real functions on (a, b), $x\in(a, b)$ and assume that $\varphi(I)\neq 0$ for sufficiently small I containing x. Set

$$D_\varphi f(x) = \sup_{x\in J}\ \inf_{x\in I\subset J} f(I)/\varphi(I), \quad D^\varphi f(x) = \inf_{x\in J}\ \sup_{x\in I\subset J} f(I)/\varphi(I).$$

If $D_\varphi f(x)=d(x)=D^\varphi f(x)$ let $df/d\varphi|_x=d(x)$.

In the manner of Riesz, we consider the Dini derivates of f relative to φ.

Definition. If f and φ are functions on (a, b) and $x\in(a, b)$ let

$$D_l^\varphi f(x) = \sup_{\alpha<x}\ \inf_{\alpha<y<x} \frac{f(y)-f(x)}{\varphi(y)-\varphi(x)}, \quad D_L^\varphi f(x) = \inf_{\alpha<x}\ \sup_{\alpha<y<x} \frac{f(y)-f(x)}{\varphi(y)-\varphi(x)},$$

$$D_r^\varphi f(x) = \sup_{x<\beta}\ \inf_{x<y<\beta} \frac{f(y)-f(x)}{\varphi(y)-\varphi(x)}, \quad D_R^\varphi f(x) = \inf_{x<\beta}\ \sup_{x<y<\beta} \frac{f(y)-f(x)}{\varphi(y)-\varphi(x)},$$

provided that the denominators do not vanish. If the four derivates have a common value let if be $d_\varphi f(x)$. The following two statements are immediate consequences of the definitions.

Proposition 1. *$df/d\varphi|_x=d(x)$ if and only if for all sequences of open intervals (x_k, y_k) containing x such that $y_k-x_k\to 0$*

$$\lim_k \frac{f(y_k)-f(x_k)}{\varphi(y_k)-\varphi(x_k)} = d(x).$$

Corollary. (a) *If $f(x+)$, $f(x-)$, $\varphi(x+)$, $\varphi(x-)$ are finite and $\varphi(x+)\neq\varphi(x-)$ then $df/d\varphi|_x$ is finite.* (b) *If f and φ increase on (a, b) and φ is not continuous at $x\in(a, b)$ then $0\leq df/d\varphi|_x<+\infty$.*

Proposition 2. $\lim\limits_{y\to x}\dfrac{f(y)-f(x)}{\varphi(y)-\varphi(x)}=d(x)$ *if and only if $d_\varphi f(x)=d(x)$.*

Proposition 3. *If φ increases on (a, b) and $d_\varphi f(x)$ is finite then $df/d\varphi|_x= =d_\varphi f(x)$.*

Proof. For any $\varepsilon>0$ there is some $\delta>0$ such that if $x-\delta<y'<x<y''<x+\delta$ then

$$d_\varphi f(x)-\varepsilon < \frac{f(y')-f(x)}{\varphi(y')-\varphi(x)}, \quad \frac{f(y'')-f(x)}{\varphi(y'')-\varphi(x)} < d_\varphi f(x)+\varepsilon. \tag{1}$$

Consider the points $P'(\varphi(y'), f(y'))$, $P(\varphi(x), f(x))$, $P''(\varphi(y''), f(y''))$ in the (φ, f)-plane and the slopes S', S, S'' of $P'P$, $P'P''$, PP'' respectively. Since φ increases on (a, b) it follows from (1) that the strict inequalities $\varphi(y')<\varphi(x)<\varphi(y'')$ hold. Hence

$$\min\{S', S''\}\leq S\leq \max\{S', S''\}.$$

Consequently

$$d_\varphi f(x)-\varepsilon \leqq D_\varphi f(x) \leqq D^\varphi f(x) \leqq d_\varphi f(x)+\varepsilon \quad \text{for all} \quad \varepsilon > 0.$$

The conclusion follows from the definition of $df/d\varphi|_x$.

It is convenient to fix some notation. We use f and φ for increasing functions on a closed interval $[a, b]$. For $x \in (a, b)$

$$\varphi^\lambda(x) = \sup_{y<x} \varphi(y), \quad \varphi^\varrho(x) = \inf_{y>x} \varphi(y), \quad E(\varphi) = \{x|\varphi^\lambda(x) < \varphi^\varrho(x)\}.$$

Then on (a, b), φ^λ and φ^ϱ increase, $\varphi^\lambda \leqq \varphi \leqq \varphi^\varrho$, $\varphi^{\lambda\lambda}=\varphi^\lambda$, $\varphi^{\varrho\varrho}=\varphi^\varrho$ and, if $(x, y) \neq \emptyset$, $(x, y)-E(\varphi)$ is uncountable since $E(\varphi)$ is the countable set of discontinuities of φ.

The exceptional set $E(f, \varphi)$

The sets

$$E_{l,R}^{\varphi\lambda}(f^\varrho) = \{x \in (a, b)|D_l^{\varphi^\lambda} f^\varrho(x) < D_R^{\varphi^\lambda} f^\varrho(x)\},$$

$$E_{r,L}^{\varphi\varrho}(f^\lambda) = \{x \in (a, b)|D_r^{\varphi^\varrho} f^\lambda(x) < D_L^{\varphi^\varrho} f^\lambda(x)\},$$

$$E_{R,\infty}^{\varphi\lambda}(f^\varrho) = \{x \in (a, b)|D_R^{\varphi^\lambda} f^\varrho(x) = +\infty\},$$

modeled on the similar sets in [1], are called the *Riesz sets*.

The set $C(\varphi)$, next to be defined, is determined by the intervals on which φ is constant. Let

$$C_x = \{y|\varphi(y) = \varphi(x)\} \quad \text{and} \quad \lambda_x = \inf C_x, \quad \varrho_x = \sup C_x \quad \text{for} \quad x \in (a, b).$$

The sets C_x are disjoint and contain x. The set of non-empty (λ_x, ϱ_x) is countable. Let these open intervals be (λ_n, ϱ_n) and let $[\lambda_n, \varrho_n]$ be their closures, and set

$$C(\varphi) = \bigcup_n [\lambda_n, \varrho_n] \cap (a, b).$$

Proposition 4. *If* $x \in (a, b)-C(\varphi)$ *and* $a<x'<x<x''<b$, $\varphi(x')<\varphi(x)<\varphi(x'')$.

Proof. Otherwise $x' \in C_x$ or $x'' \in C_x$. In either case $(\lambda_x, \varrho_x) \neq \emptyset$ and $x \in [\lambda_x, \varrho_x] \subset \subset C(\varphi)$, contrary to hypothesis.

The exceptional set for f and φ on $[a, b]$ is

$$E(f, \varphi) = E(f) \cup E(\varphi) \cup C(\varphi) \cup E_{l,R}^{\varphi\lambda}(f^\varrho) \cup E_{r,L}^{\varphi\varrho}(f^\lambda) \cup E_{R,\infty}^{\varphi\lambda}(f^\varrho).$$

Proposition 5. *If* $x \in (a, b)-(E(f, \varphi)-E(\varphi))$, *then* $0 \leqq df/d\varphi|_x < +\infty$.

Proof. Consider $x \in (a, b) - E(f, \varphi)$ and $a < x' < x < x'' < b$. Since $x \notin E(f) \cup \cup E(\varphi) \cup C(\varphi)$ we infer from Proposition 4

$$f^\lambda(x') \leqq f(x') \leqq f^\varrho(x') \leqq f^\lambda(x) = f(x) = f^\varrho(x) \leqq f^\lambda(x'') \leqq f(x'') \leqq f^\varrho(x''),$$

$$\varphi^\lambda(x') \leqq \varphi(x') \leqq \varphi^\varrho(x') < \varphi^\lambda(x) = \varphi(x) = \varphi^\varrho(x) < \varphi^\lambda(x'') \leqq \varphi(x'') \leqq \varphi^\varrho(x'');$$

and hence,

$$0 \leqq \frac{f^\varrho(x) - f^\varrho(x')}{\varphi^\lambda(x) - \varphi^\lambda(x')} \leqq \frac{f(x) - f(x')}{\varphi(x) - \varphi(x')} \leqq \frac{f^\lambda(x) - f^\lambda(x')}{\varphi^\varrho(x) - \varphi^\varrho(x')} < +\infty,$$

$$0 \leqq \frac{f^\lambda(x'') - f^\lambda(x)}{\varphi^\varrho(x'') - \varphi^\varrho(x)} \leqq \frac{f(x'') - f(x)}{\varphi(x'') - \varphi(x)} \leqq \frac{f^\varrho(x'') - f^\varrho(x)}{\varphi^\lambda(x'') - \varphi^\lambda(x)} < +\infty.$$

Therefore,

$$(1)\qquad \begin{gathered} 0 \leqq D_l^{\varphi^\lambda} f^\varrho(x) \leqq D_l^\varphi f(x) \leqq D_L^\varphi f(x) \leqq D_L^{\varphi^\varrho} f^\lambda(x) \leqq +\infty, \\ 0 \leqq D_r^{\varphi^\varrho} f^\lambda(x) \leqq D_r^\varphi f(x) \leqq D_R^\varphi f(x) \leqq D_R^{\varphi^\lambda} f^\varrho(x) \leqq +\infty. \end{gathered}$$

Since the Riesz sets exclude x it follows from their defining inequalities and (1) that

$$0 \leqq D_l^\varphi f(x) = D_L^\varphi f(x) = D_r^\varphi f(x) = D_R^\varphi f(x) = D_R^{\varphi^\lambda} f^\varrho(x) < +\infty.$$

By Proposition 3,

$$(2)\qquad 0 \leqq df/d\varphi|_x < +\infty \quad \text{for} \quad x \in (a, b) - E(f, \varphi).$$

By the Corollary to Proposition 1

$$(3)\qquad 0 \leqq df/d\varphi|_x = \frac{f^\varrho(x) - f^\lambda(x)}{\varphi^\varrho(x) - \varphi^\lambda(x)} < +\infty \quad \text{for} \quad x \in E(\Phi).$$

The conclusion follows from (2), (3).

Toward $\mu_\varphi(E(f, \varphi) - E(\varphi)) = 0$

We summarize the properties of measure which play a role in what follows. For an increasing function φ defined on an open interval I of $\mathbf{R}$ and any $A \subset I$, let

$$\mu_\varphi(A) = \inf\{\sum_n \varphi(I_n) | A \subset \cup I_n,\ I_n = (a_n, b_n) \subset [a_n, b_n] \subset I\}.$$

Proposition 6. *For A, $[a_n b_n]$, (x, y), $(x, y]$, $[x, y]$, $\{x\}$ and A_n subsets of I we have:*

(a) $\mu_\varphi(A) = \inf\{\sum_n \varphi(I_n) | A \subset \bigcup_n I_n,\ I_n = (a_n, b_n),\ a_n, b_n \notin E(\varphi)\}$.

(b) $\mu_\varphi((x, y)) = \varphi^\lambda(y) - \varphi^\varrho(x)$, $\mu_\varphi((x, y]) = \varphi^\varrho(y) - \varphi^\varrho(x)$,

$\mu_\varphi([x, y]) = \varphi^\varrho(y) - \varphi^\lambda(x)$.

(c) $\mu_\varphi(\{x\}) = \varphi^\varrho(x) - \varphi^\lambda(x)$.

(d) If $\mu_\varphi(A_n) = 0$ for $n \in \mathbf{N}$, $\mu_\varphi(\bigcup_n A_n) = 0$.

Proposition 7. *If φ, ψ increase on I then $\mu_\varphi(A)=\mu_\psi(A)$ for all $A\subset I$ if and only if*

$$(1) \qquad E(\varphi) = E(\psi) \quad \text{and} \quad \varphi(x)-\psi(x) \quad \text{is constant on} \quad I-E(\varphi).$$

Proof. Assume (1). Then, by Proposition 6(*a*), $\mu_\varphi(A)=\mu_\psi(A)$ for $A\subset I$. Conversely, the latter equality implies $E(\varphi)=E(\psi)$ by Proposition 6(*c*) and then, choosing $a\in I-E(\varphi)$, $\varphi(x)-\varphi(a)=\mu_\varphi([a, x])=\mu_\psi([a, x])=\psi(x)-\psi(a)$ for $x\in I-E(\varphi)$, $x>a$, and a similar argument applies if $x\in I-E(\varphi)$, $x<a$, by Proposition 6(*b*).

Corollary. *For all $A\subset I$, $\mu_{\varphi^\lambda}(A)=\mu_\varphi(A)=\mu_{\varphi^\varrho}(A)$.*

Proposition 8. $\mu_\varphi((E(f)\cup C(\varphi))-E(\varphi))=0$.

Proof. By the definition of $C(\varphi)$,

$$(E(f)\cup C(\varphi))-E(\varphi)\subset(E(f)-E(\varphi))\cup\left(\bigcup_n(\lambda_n, \sigma_n)\cup(\{\lambda_n, \varrho_n|n\in\mathbf{N}\}-E(\varphi))\right).$$

The first and last sets are countable and φ is continuous at each of their points. Since for each n, φ is constant on (λ_n, ϱ_n), $\varphi^\varrho(\lambda_n)=\varphi^\lambda(\varrho_n)$ for all n. The result now follows from Proposition 6(*d*).

The 'rising sun' theorem [1] is used as a lemma to show that the three Riesz sets are of μ_φ-measure zero.

Lemma. *If g is a real function on $[a, b]$, $g(a)\geqq g(a+)$, $g(b)\geqq g(b-)$, and $g(x)\geqq \geqq\max\{g(x+), g(x-)\}$ for $a<x<b$, then there are sequences (a_n, b_n), (c_n, d_n) of disjoint subintervals of (a, b) such that*

$$\{x\in(a, b)|g(y)>g(x) \text{ for some } y\in(a, x)\}=\bigcup_n(a_n, b_n),$$

$$\{x\in(a, b)|g(y)>g(x) \text{ for some } y\in(x, b)\}=\bigcup_n(c_n, d_n),$$

$$g(a_n)\geqq g(b_n-), \quad g(c_n+)\leqq g(d_n) \quad \text{for all} \quad n.$$

Proposition 9. *If f, φ increase on $[a, b]$, $f(a)=f(a+)$, $f=f^\varrho$, $\varphi(b)=\varphi(b-)$, $\varphi=\varphi^\lambda$, $t>0$, and $g=f-t\varphi$ then g satisfies the hypotheses of the Lemma.*

Proof. Since $\varphi^\lambda=\varphi\leqq\varphi^\varrho$, $f^\lambda\leqq f=f^\varrho$ on (a, b), we have for $x\in(a, b)$

$$g(x+)=f^\varrho(x)-t\varphi^\varrho(x)\leqq f(x)-t\varphi(x)=g(x),$$

$$g(x-)=f^\lambda(x)-t\varphi^\lambda(x)\leqq f(x)-t\varphi(x)=g(x).$$

A similar argument applies for $x=a$ and $x=b$.

In applying the Lemma to the Riesz sets we use Proposition 9 and the fact that $f^{\varrho\varrho}=f^\varrho$, $\varphi^{\lambda\lambda}=\varphi^\lambda$. The next proposition may be called the *Riesz covering theorem*,

Proposition 10. *If $f=f^\varrho$, $\varphi=\varphi^\lambda$ on $J=(\alpha,\beta)\subset(\alpha,\beta)$ and*

$$E=\{x\in J \mid D_L^\varrho f(x)<u<v<D_R^\varrho f(x)\}$$

then there are $N\subset J$ and a countable set S of disjoint subintervals of J such that

$$\mu_\varphi(N)=0,\quad S \text{ covers } E-N,\quad \sum_{I\in S}\varphi(I)\leqq\frac{u}{v}\varphi(J).$$

Proof. If $x\in E$ there is some $y\in(\alpha,x)$ such that $(f(x)-f(y))/(\varphi(x)-\varphi(y))<u$. Hence

$$g_u(y)=f(y)-u\varphi(y)>f(x)-u\varphi(x)=g_u(x).$$

Since $g_u=f-u\varphi$ satisfies the hypothesis of Proposition 9, it follows from the Lemma that there are disjoint $I_n=(a_n,b_n)\subset J$, $n\in\mathbf{N}$, such that, since $\varphi=\varphi^\lambda$ and $f(b_n-)=f^\lambda(b_n)$,

$$(1)\qquad E\subset\bigcup_n I_n,\quad g_u(a_n)=f(a_n)-u\varphi(a_n)\geqq f^\lambda(b_n)-u\varphi(b_n)=g_u(b_n-).$$

Hence

$$(2)\qquad f^\lambda(b_n)-f(a_n)\leqq u(\varphi(b_n)-\varphi(a_n))=u\varphi(I_n),\quad n\in\mathbf{N}.$$

For each n there is a sequence $b_{n,p}\in I_n-E(\varphi)$ such that $b_{n,p}\uparrow b_n$. Let $b_{n,0}=a_n$, $I_{n,p}=(b_{n,p-1},b_{n,p})$, $N'=\{b_{n,p}\mid n,p\in\mathbf{N}\}$. Then

$$(3)\qquad \mu_\varphi(N')=0,\ I_{n,p},\quad n,p\in\mathbf{N},\quad \text{are disjoint},\quad E-N'\subset\bigcup_{n,p}I_{n,p}\subset\bigcup_n I_n\subset J.$$

Since f increases and $b_{n,0}=a_n$ for all n

$$\sum_p f(I_{n,p})=\sum_p(f(b_{n,p})-f(b_{n,p-1}))=\lim_p f(b_{n,p})-f(a_n)=f^\lambda(b_n)-f(a_n).$$

By (2), (3), since φ increases,

$$(4)\qquad \sum_{n,p}f(I_{n,p})=\sum_n(f^\lambda(b_n)-f(a_n))\leqq u\sum_n\varphi(I_n)\leqq u\varphi(J).$$

For each n,p if $x\in E\cap I_{n,p}$ there is some $y\in(x,b_{n,p})$ such that $(f(y)-f(x))/(\varphi(y)-\varphi(x))>v$. Now

$$g_v(y)=f(y)-v\varphi(y)>f(x)-v\varphi(x)=g_v(x).$$

Since $g_v=f-v\varphi$ satisfies the hypothesis of Proposition 9 it follows from the Lemma that there is a sequence of disjoint $I_{n,p,m}=(c_{n,p,m},d_{n,p,m})\subset I_{n,p}$ such that, since $f=f^\varrho$ and $\varphi(c_{n,p,m}+)=\varphi^\varrho(c_{n,p,m})$,

$$E\cap I_{n,p}\subset\bigcup_m I_{n,p,m},\quad f(c_{n,p,m})-v\varphi^\varrho(c_{n,p,m})\leqq f(d_{n,p,m})-v\varphi(d_{n,p,m}).$$

Hence

$$(5)\qquad v(\varphi(d_{n,p,m})-\varphi^\varrho(c_{n,p,m}))\leqq f(I_{n,p,m}),\quad n,p,m\in\mathbf{N}.$$

For all n, p, m there is a sequence $c_{n,p,m,q} \in I_{n,p,m} - E(\varphi)$ such that $c_{n,p,m,q} \downarrow c_{n,p,m}$. Let $c_{n,p,m,0} = d_{n,p,m}$, $I_{n,p,m,q} = (c_{n,p,m,q}, c_{n,p,m,q-1})$ and

$$N'' = \{c_{n,p,m,q} | n, p, m, q \in \mathbf{N}\}.$$

Then

$$\text{(6)} \qquad \begin{aligned} &\mu_\varphi(N'') = 0, \quad I_{n,p,m,q}, \quad n, p, m, q \in \mathbf{N}, \quad \text{are disjoint,} \\ &E - (N' \cup N'') \subset \bigcup_{n,p,m,q} I_{n,p,m,q} \subset \bigcup_{n,p,m} I_{n,p,m} \subset \bigcup_{n,p} I_{n,p}. \end{aligned}$$

Since $c_{n,p,m,q} \downarrow c_{n,p,m}$ and $c_{n,p,m,0} = d_{n,p,m}$

$$\text{(7)} \qquad \begin{aligned} \sum_q \varphi(I_{n,p,m,q}) &= \sum_q (\varphi(c_{n,p,m,q-1}) - \varphi(c_{n,p,m,q})) \\ &= \varphi(d_{n,p,m}) - \lim_q \varphi(c_{n,p,m,q}) = \varphi(d_{n,p,m}) - \varphi^\varrho(c_{n,p,m}). \end{aligned}$$

Since f increases it follows from (4), (5), (6), (7) that

$$\text{(8)} \qquad v \sum_{n,p,m,q} \varphi(I_{n,p,m,q}) \leqq \sum_{n,p,m} f(I_{n,p,m}) \leqq \sum_{n,p} f(I_{n,p}) \leqq u\varphi(J).$$

Let $N = N' \cup N''$ and $S = \{I_{n,p,m,q} | n, p, m, q \in \mathbf{N}\}$. By (3), (6), (8), N and S satisfy the required conditions.

Proposition 11. $\mu_{\varphi^\lambda}(E^{\varphi^\lambda}_{l,R}(f^\varrho)) = 0$.

Proof. $E^{\varphi^\lambda}_{l,R}(f^\varrho)$ is the union of the countable set of

$$E^J_{u,v} = \{x \in J = (a, b) | D^{\varphi^\lambda}_l f^\varrho(x) < u < v < D^{\varphi^\lambda}_R f^\varrho(x)\}, \quad u, v \text{ rational.}$$

We note that $f^\varrho = f^{\varrho\varrho}$, $\varphi^\lambda = \varphi^{\lambda\lambda}$ and show that for $k \in \mathbf{N}$ there are $N_k \subset J$ and a countable set S_k of disjoint open subintervals of J such that

$$\{k\} \qquad \mu_{\varphi^\lambda}(N_k) = 0, \quad S_k \text{ covers } E^J_{u,v} - N_k, \quad \sum_{I \in S_k} \varphi(I) \leqq \left(\frac{u}{v}\right)^k \varphi(J).$$

By Proposition 10 with $(\alpha, \beta) = (a, b)$ there are N_1, S_1 satisfying $\{1\}$. Assume that N_k and S_k satisfy $\{k\}$. Let I_p, $p \in \mathbf{N}$, be the intervals of S_k. By Proposition 10 with $(\alpha, \beta) = I_p$ there are $M_p \subset I_p$ and a countable set T_p of disjoint open subintervals of I_p such that

$$\mu_{\varphi^\lambda}(M_p) = 0, \quad T_p \text{ covers } E^J_{u,v} \cap I_p - M_p, \quad \sum_{I \in T_p} \varphi(I) \leqq \frac{u}{v} \varphi(I_p), \quad p \in \mathbf{N}.$$

Let $N_{k+1} = N_k \cup (\bigcup_p M_p)$ and $S_{k+1} = \bigcup_p T_p$. Then $\mu_{\varphi^\lambda}(N_{k+1}) = 0$, S_{k+1} covers $E^J_{u,v} - N_{k+1}$ and

$$\sum_{I \in S_{k+1}} \varphi(I) = \sum_p \sum_{I \in T_p} \varphi(I) \leqq \sum_p \frac{u}{v} \varphi(I_p) \leqq \left(\frac{u}{v}\right)^{k+1} \varphi(J).$$

Thus N_{k+1}, S_{k+1} satisfy $\{k+1\}$, and therefore, $\{k\}$ is satisfied for all $k \in N$.

Let $N=\bigcup_k N_k$. Then $\mu_{\varphi^\lambda}(N)=0$, S_k covers $E^J_{u,v}-N$ for all k and, since $\lim_k (u/v)^k$ $\varphi(J)=0$, $\mu_{\varphi^\lambda}(E^J_{u,v})=0$ for all rational u, v. Hence $\mu_{\varphi^\lambda}(E^{\varphi\lambda}_{l,R}(f^\varrho))=0$.

Proposition 12. $\mu_{\varphi^\varrho}(E^{\varphi\varrho}_{r,L}(f^\lambda))=0$.

Proof. Let $T(x)=-x$ for $x\in\mathbf{R}$. Let $h(T(x))=-f(x)$, $\psi(T(x))=-\varphi(x)$. Then h, ψ increase on $(T(b), T(a))$ $h^\varrho=-f^\lambda$, $\psi^\lambda=-\varphi^\varrho$, and for all $A\subset(T(b), T(a))$, $\mu_{\varphi^\varrho}(T^{-1}(A))=\mu_{\psi^\lambda}(A)$. Since $T(y)<T(x)$ if and only if $x<y$,

$$\frac{h(T(y))-h(T(x))}{\psi(T(y))-\psi(T(x))}=\frac{f(x)-f(y)}{\varphi(x)-\varphi(y)}$$

if either difference quotient is finite. Hence

$$E^{\varphi\varrho}_{r,L}(f^\lambda)=T^{-1}(E^{\psi\lambda}_{l,R}(h^\varrho)).$$

By Proposition 11, $\mu_{\psi^\lambda}(E^{\psi\lambda}_{l,R}(h^\varrho))=0$. Hence $\mu_{\varphi^\varrho}(E^{\varphi\varrho}_{r,L}(f^\lambda))=0$.

Proposition 13. $\mu_{\varphi^\lambda}(E^{\varphi\lambda}_{R,\infty}(f^\varrho))=0$.

Proof. For each $m\in\mathbf{N}$ let

$$E_m=\{x\in(a,b)\,|\,D^{\varphi\lambda}_R f^\varrho(x)>m\}.$$

Then $E_{m+1}\subset E_m\subset(a,b)$ for all m. If $x\in E_m$ there is some $y\in(x,b)$ such that

$$g_m(y)=f^\varrho(y)-m\varphi^\lambda(y)>f^\varrho(x)-m\varphi^\lambda(x)=g_m(x).$$

By Proposition 9 and the Lemma there is a sequence of disjoint $I_p=(c_p, d_p)\subset(a,b)$ such that, since $f^\varrho(c_p+)=f^\varrho(c_p)$,

$$E_m\subset\bigcup_p I_p,\quad f^\varrho(c_p)-m\varphi^\lambda(c_p+)\leqq f^\varrho(d_p)-m\varphi^\lambda(d_p),\quad p\in\mathbf{N}.$$

For each p there is a sequence $c_{p,q}\in I_p-E(\varphi)$ such that $c_{p,q}\downarrow c_p$. Let $c_{p,0}=d_p$, $I_{p,q}=(c_{p,q}, c_{p,q-1})$ and $N=\{c_{p,q}|p, q\in\mathbf{N}\}$. Then $\mu_{\varphi^\lambda}(N)=0$, $E_m-N\subset\bigcup_{p,q} I_{p,q}\subset\bigcup_p I_p\subset(a,b)$ for all m,

$$m\sum_{p,q}\varphi^\lambda(I_{p,q})=m\sum_p\sum_q(\varphi^\lambda(c_{p,q-1})-\varphi^\lambda(c_{p,q}))=m\sum_p(\varphi^\lambda(d_p)-\varphi^\lambda(c_p+))\leqq$$
$$\leqq\sum_p(f^\varrho(d_p)-f^\varrho(c_p))\leqq f^\varrho((a,b))<+\infty.$$

Hence, $\mu_\varphi(E_m)\leqq f^\varrho((a,b))/m$ for all m. Since $E^{\varphi\lambda}_{R,\infty}(f^\varrho)\subset\bigcap_m E_m\subset(a,b)$,

$$0\leqq\mu_{\varphi^\lambda}(E^{\varphi\lambda}_{R,\infty}(f^\varrho))\leqq\lim_m\mu_{\varphi^\lambda}(E_m)=0.$$

Theorem 1. *If f and φ increase on (a,b) there is some $A\subset I$ such that*

$$0\leqq df/d\varphi|_x<+\infty\quad\textit{for}\quad x\in A\quad\textit{and}\quad\mu_\varphi((a,b)-A)=0.$$

Proof. By representing (a, b) as a union of countably many closed subintervals, we may consider one of them and assume that f and φ increase on $[a, b]$. By the definition of the exceptional set $E(f, \varphi)$

$$E(f, \varphi)-E(\varphi)\subset((E(f)\cup C(\varphi))-E(\varphi))\cup E^{\varphi^\lambda}_{l,R}(f^\varrho)\cup E^{\varphi^\varrho}_{r,L}(f^\lambda)\cup E^{\varphi^\lambda}_{R,\infty}(f^\varrho).$$

Since $E(\varphi)$ is the set of discontinuities of φ, φ^λ, φ^ϱ and $\varphi=\varphi^\lambda=\varphi^\varrho$ on $(a, b)-E(\varphi)$ it follows from Proposition 7 that μ_φ, μ_{φ^λ}, μ_{φ^ϱ} are identical measures.

Let $A=(a, b)-\big(E(f, \varphi)-E(\varphi)\big)$. The conclusion follows from Propositions 6, 8, 11, 12, 13.

Toward Theorem 2

FUBINI's theorem [2] on the derivative of a function represented by a convergent series of increasing functions is extended in the following proposition.

Proposition 14. *If f_n, $n\in\mathbf{N}$, and φ increase on (a, b) and*

$$\sum_n f_n(x) = f(x) \quad \text{is finite on} \quad (a, b)$$

then there is some $A\subset(a, b)$ such that $\mu_\varphi\big((a, b)-A\big)=0$ and

$${}'\sum_n df_n/d\varphi|_x = df/d\varphi|_x \quad for \quad x\in A.$$

The proof is so close to that of Fubini for the case where $\varphi(x)=x$ that it is omitted.

Similarly, Lebesgue's density theorem may be generalized. It is convenient to say that a sequence of open intervals (x_k, y_k) determines x if $x\in(x_k, y_k)$ for all k and $\lim_k (x_k-y_k)=0$.

Definition. Let φ increase on an open interval $I\subset\mathbf{R}$. The *μ_φ-density* of a set A at $x\in I$ is $\Delta(A, x)$ if for all sequences (x_k, y_k) which determine x

$$\lim_k \frac{\mu_\varphi\big(A\cap(x_k, y_k]\big)}{\mu_\varphi\big((x_k, y_k]\big)} = \Delta(A, x).$$

Proposition 15. *If $A\subset(a, b)\subset[a, b]\subset I$ is a μ_φ-measurable set then there is some $D\subset A$ such that*

$$\Delta(x, A) = 1 \quad for \quad x\in D \quad and \quad \mu_\varphi(A-D) = 0.$$

Proof. There is by Proposition 6(d) an open set G_n for $n\in\mathbf{N}$ such that $A\subset G_n\subset(a, b)$ and $\mu_\varphi(G_n)<\mu_\varphi(A)+1/2^n$. Let

$$f(x) = \mu_\varphi\big(A\cap(a, x]\big), \quad \psi(x) = \varphi^\varrho(x)-\varphi^\varrho(a), \quad x\in(a, b),$$
$$f_n(x) = \mu_\varphi\big(G_n\cap(a, x]\big) \quad x\in(a, b), \quad n\in\mathbf{N}.$$

Then f, ψ, f_n increase on (a, b). Since

$$0 \leqq f_n(x)-f(x) = \mu_\varphi((G_n-A)\cap(a, x]) \leqq \mu_\varphi(G_n-A) < 1/2^n$$

and

$$f_n(y)-f(y)-(f_n(x)-f(x)) = \mu_\varphi((G_n-A)\cap(x, y]) \geqq 0 \quad \text{for} \quad x<y$$

it follows from Theorem 1 and Proposition 14 that there is some $D\subset A$ such that

$$(1) \qquad 0 \leqq \sum_n (df_n/d\psi|_x - df/d\psi|_x) < +\infty, \quad x\in D, \quad \text{and} \quad \mu_\psi(A-D)=0.$$

For $x\in D$ and any sequence (x_k, y_k) which determines x, there is some $k_{n,x}$ such that $(x_k, y_k]\subset G_n$ for $k\geqq k_{n,x}$. Then by Theorem 1, Proposition 1 and Proposition 6

$$df_n/d\psi|_x = \lim_k \frac{f_n(y_k)-f_n(x_k)}{\psi(y_k)-\psi(x_k)} = \lim_k \frac{\mu_\varphi(G_n\cap(x_k, y_k])}{\mu_\varphi((x_k, y_k])} = 1, \quad x\in D, \quad n\in\mathbf{N}.$$

By (1),

$$0 \leqq \sum_n (1-df/d\psi|_x) < +\infty \quad \text{for} \quad x\in D.$$

Hence

$$(2) \qquad df/d\psi|_x = 1 \quad \text{for} \quad x\in D.$$

Since $E(\varphi)=E(\varphi^\varrho)=E(\psi)$ and $\varphi(x)-\psi(x)=\varphi^\varrho(a)$ for $x\in(a,b)-E(\varphi)$, it follows from Proposition 7 that $\mu_\varrho(A-D)=\mu_\psi(A-D)=0$. Hence by (2)

$$1 = df/d\psi|_x = \lim_k \frac{f(y_k)-f(x_k)}{\psi(y_k)-\psi(x_k)} = \lim_k \frac{\mu_\varphi(A\cap(x_k, y_k])}{\mu_\varphi((x_k, y_k])} = \Delta(x, A),$$

for $x\in D$, $\mu_\varphi(A-D)=0$.

Proposition 16. *If φ increases on an open interval $I\subset\mathbf{R}$, f is μ_φ-integrable on $[a, b]\subset I$ and*

$$F(x) = \int_{(a,x]} f\, d\mu_\varphi \quad \textit{for} \quad x\in(a,b)$$

there is some $A\subset(a,b)$ such that

$$dF/d\varphi|_x = f(x) \quad \textit{for} \quad x\in A \quad \textit{and} \quad \mu_\varphi((a,b)-A)=0.$$

Proof. It is assumed, without loss of generality, that f is positive. There is a sequence of compact $C_n\subset(a,b)$ such that

$$C_n\subset C_{n+1} \quad \text{and } f \text{ is continous on } C_n, \text{ for } n\in\mathbf{N},$$

$$\lim_n \mu_\varphi((a,b)-C_n)=0, \quad \lim_n \int_{C_n} f\,d\mu_\varphi = \int_{(a,b)} f\,d\mu_\varphi < +\infty.$$

For $n\in\mathbf{N}$ let $f_n(x)=f(x)$ for $x\in C_n$, and $f_n(x)=0$, for $x\in(a,b)-C_n$, and set $A_1=\bigcup_n C_n$. Then

$$f_n\uparrow f \quad \text{on } A_1, \quad \mu_\varphi((a,b)-A_1) = \lim_n \mu_\varphi(A_1-C_n)=0.$$

Let

$$F_n(x) = \int_{(a,x]} f_n \, d\mu_\varphi \quad \text{for} \quad n \in \mathbf{N}, \quad x \in (a, b).$$

Since f_n and $f_{n+1}-f_n$ are positive on A_1, F_n and $F_{n+1}-F_n$ increase on (a, b). By the monotonic convergence theorem

$$\begin{aligned} F_1(x) &+ \sum_n (F_{n+1}(x) - F_n(x)) \\ &= \int_{(a,x]} f_1 \, d\mu_\varphi + \sum_n \Big(\int_{(a,x]} f_{n+1} \, d\mu_\varphi - \int_{(a,x]} f_n \, d\mu_\varphi \Big) \\ &= \lim_n \int_{(a,x]} f_n \, d\mu_\varphi = \int_{(a,x]} \lim_n f_n \, d\mu_\varphi = F(x) < +\infty, \quad x \in (a, b). \end{aligned}$$

Hence by Theorem 1 and the generalized Fubini theorem, Proposition 14, there is some $A_2 \subset A_1$ such that

$$0 \leqq dF_n/d\varphi|_x, \quad dF/d\varphi|_x < +\infty,$$

(1) $$\lim_n dF_n/d\varphi|_x = dF/d\varphi|_x \quad \text{for} \quad x \in A_2 \quad \text{and} \quad \mu_\varphi(A_1 - A_2) = 0.$$

Consider $x \in A_2$. There is a sequence (x_k, y_k) which determines x such that $x_k, y_k \notin E(\varphi)$ for all k. Then

(2) $$dF_n/d\varphi|_x = \lim_k \frac{F_n(y_k) - F_n(x_k)}{\varphi(y_k) - \varphi(x_k)} \quad \text{for} \quad n \in \mathbf{N}.$$

Since φ is continuous at each x_k, y_k, by Proposition 6

(3) $$\mu_\varphi((x_k, y_k]) = \varphi(y_k) - \varphi(x_k) \quad \text{for all} \quad k.$$

On the compact set $C_n \cap [x_k, y_k]$, f is continuous and $f = f_n$. Hence there are $x_{n,k}$, $y_{n,k} \in C_n \cap [x_k, y_k]$, such that

(4) $$f(x_{n,k}) \leqq f(z) \leqq f(y_{n,k}) \quad \text{for} \quad z \in C_n \cap [x_k, y_k], \quad n, k \in \mathbf{N}.$$

Since $y_k - x_k \to 0$

(5) $$\lim_k f(x_{n,k}) = f(x) = \lim_k f(y_{n,k}) \quad \text{for} \quad x \in C_n.$$

By (3), (4)

$$\begin{aligned} f(x_{n,k}) \frac{\mu_\varphi(C_n \cap (x_k, y_k])}{\mu_\varphi((x_k, y_k])} &\leqq (\varphi(y_k) - \varphi(x_k))^{-1} \int_{C_n \cap (x_k, y_k]} f \, d\mu_\varphi \\ &= \frac{F_n(y_k) - F_n(x_k)}{\varphi(y_k) - \varphi(x_k)} \leqq f(y_{n,k}) \frac{\mu_\varphi(C_n \cap (x_k, y_k])}{\mu_\varphi((x_k, y_k])}, \quad n, k \in \mathbf{N}. \end{aligned}$$

By the density theorem, Proposition 15, for each n there is some $D_n \subset C_n$ such that

(6) $$\lim_k \frac{\mu_\varphi(C_n \cap (x_k, y_k])}{\mu_\varphi((x_k, y_k])} = 1 \quad \text{for} \quad x \in D_n \quad \text{and} \quad \mu_\varphi(C_n - D_n) = 0.$$

By (2), (5), (6)

$$dF_n/d\varphi|_x = f(x) \quad \text{for} \quad x \in A_2 \cap D_n, \quad n \in \mathbf{N}. \tag{7}$$

Since $\mu_\varphi(A_1 - C_n) \to 0$, there are n_j such that $\mu_\varphi(A_1 - C_{n_j}) < 1/2^j$ for $j \in \mathbf{N}$. Let

$$D = \bigcup_k \bigcap_{j \geqq k} D_{n_j}.$$

Since $D_{n_j} \subset C_{n_j} \subset A_1$ and $\mu_\varphi(C_{nj} - D_{nj}) = 0$ for all j,

$$A_1 - D = \bigcap_k \bigcup_{j \geqq k} (A_1 - D_{n_j}) \subset \bigcup_{j \geqq k} (A_1 - C_{n_j}) \cup \bigcup_{j \geqq k} (C_{n_j} - D_{n_j}),$$

$$\mu_\varphi(A_1 - D) \leqq \sum_{j \geqq k} \mu_\varphi(A_1 - C_{n_j}) < \sum_{j \geqq k} 1/2^j = 1/2^{k-1}, \quad k \in \mathbf{N}.$$

Hence $\mu_\varphi(A_1 - D) = 0$. Let $A = A_2 \cap D$. If $x \in A$ then, for some k and all $j \geqq k$, $x \in A_2 \cap D_{nj}$. By (1), (7)

$$dF/d\varphi|_x = \lim_j dF_{n_j}/d\varphi|_x = f(x) \quad \text{for} \quad x \in A. \tag{8}$$

Since $A = A_2 \cap D \subset A_2 \subset A_1 \subset (a, b)$

$$\begin{aligned} 0 &\leqq \mu_\varphi((a, b) - A) \leqq \mu_\varphi((a, b) - A_1) + \\ &+ \mu_\varphi(A_1 - A_2) + \mu_\varphi(A_2 - A) \leqq \mu_\varphi(A_1 - D) = 0. \end{aligned} \tag{9}$$

By (8), (9), A satisfies the required conditions.

Theorem 2. *Let f, φ increase on an open interval $I \subset \mathbf{R}$ and let f be absolutely continuous with respect to φ, i.e., $\mu_f(A) = 0$ for all $A \subset I$ such that $\mu_\varphi(A) = 0$. Then*

$$f(b-) - f(a+) = \int_{(a,b)} df/d\varphi|_x \, d\mu_\varphi \quad \text{for all} \quad (a, b) \subset I.$$

Proof. Consider the measures μ_f, μ_φ. By the theorem SAKS ([3], p. 33) calls the Lebesgue decomposition theorem there are, for any (a, b)$\subset I$, some $H \subset (a, b)$ such that $\mu_\varphi(H) = 0$ and a positive function g, μ_φ-integrable on (a, b), such that

$$\mu_f((a, x]) = \int_{(a,x]} g \, d\mu_\varphi + \mu_f(H \cap (a, x]) \quad \text{for all} \quad x \in (a, b).$$

Since f is absolutely continuous with respect to φ and $\mu_\varphi(H) = 0$, $\mu_f(H \cap (a, x]) = 0$ for all $x \in (a, b)$. Hence

$$\psi(x) = \mu_f((a, x]) = \int_{(a,x]} g \, d\mu_\varphi \quad \text{for} \quad x \in (a, b).$$

By Proposition 16 there is some $A_1 \subset (a, b)$ such that

$$d\psi/d\varphi|_x = g(x) \quad \text{for} \quad x \in A_1 \quad \text{and} \quad \mu_\varphi((a, b) - A_1) = 0.$$

Since f increases on I there is, by Theorem 1, some $A_2 \subset (a, b)$ such that

$$0 \leqq df/d\varphi|_x < +\infty \quad \text{for} \quad x \in A_2 \quad \text{and} \quad \mu_\varphi((a, b) - A_2) = 0.$$

Let $A = A_1 \cap A_2$. For $x \in A$ there is a sequence $(x_k, y_k) \subset (a, b)$, determining x and such that $x_k, y_k \in (a, b) - (E(\varphi) \cup E(f))$. By Proposition 6

$$f(y_k) - f(x_k) = \mu_f((x_k, y_k]) = \psi(y_k) - \psi(x_k) \quad \text{for all} \quad k.$$

By Proposition 1

$$df/d\varphi|_x = \lim_k \frac{f(y_k) - f(x_k)}{\varphi(y_k) - \varphi(x_k)} = \lim_k \frac{\psi(y_k) - \psi(x_k)}{\varphi(y_k) - \varphi(x_k)} = d\psi/d\varphi|_x = g(x), \quad x \in A,$$

and

$$0 \leqq \mu_\varphi((a, b) - A) \leqq \mu_\varphi((a, b) - A_1) + \mu_\varphi((a, b) - A_2) = 0.$$

Hence

$$\mu_f((a, x]) = \int_{(a, x]} df/d\varphi|_x \, d\mu_\varphi \quad \text{for} \quad x \in (a, b).$$

There are sequences $a_k, b_k \in (a, b) - E(f)$ such that $a_1 < b_1$ and $a_k \downarrow a$, $b_k \uparrow b$. Now

$$f(b_k) - f(a_k) = \mu_f((a_k, b_k]) = \int_{(a_k, b_k]} df/d\varphi|_x \, d\mu_\varphi \quad \text{for all} \quad k.$$

Hence

$$f(b-) - f(a+) = \lim_k \int_{(a_k, b_k]} df/d\varphi|_x \, d\mu_\varphi = \int_{(a, b)} df/d\varphi|_x \, d\mu_\varphi .$$

References

[1] F. Riesz, Sur l'existence de la dérivée des fonctions monotones ..., *Acta Sci. Math.*, **5** (1932), 208—221.

[2] G. Fubini, Sulla derivazione per serie, *Atti Accad, Naz. Lincei, Rend.*, **24** (1915), 204—206.

[3] S. Saks, *Theory of the Integral*, 2nd ed. English translation by L. C. Young, Stechert & Co. (New York, 1937).

[4] J. Radon, Theorie und Anwendungen der absolut additiven Mengenfunktionen, *S.-B. Akad. Wiss. Wien*, **122** (1913), 1295—1438.

[5] O. Nikodym, Sur une généralisation des intégrales de M. Radon, *Fundam. Math.*, **10** (1930), 131—179.

UNIVERSITY OF MARYLAND
DEPARTMENT OF MATHEMATICS
COLLEGE PARK, MARYLAND 20742

Acta Sci. Math., 38 (1976), 253—260

Lebesgue-type decomposition of positive operators

T. ANDO

1. Introduction

Our main concerns in this paper are bounded (linear) *positive*, i.e. non-negative definite, operators on a Hilbert space $\mathfrak{H}$. Given a positive operator A, we say a positive operator C to be *A-absolutely continuous* if there exists a sequence $\{C_n\}$ of positive operators such that $C_n \uparrow C$ and $C_n \leqq \alpha_n A$ for some $\alpha_n \geqq 0$ $(n=1, 2, \ldots)$. Here $C_n \uparrow C$ means that $C_1 \leqq C_2 \leqq C_3 \leqq \ldots$ and C_n converges strongly to C. A positive operator C is said to be *A-singular* if $0 \leqq D \leqq A$ and $0 \leqq D \leqq C$ imply $D=0$. These definitions are motivated by the corresponding notions in measure theory (cf. [3]). In accordance with a well-known theorem of measure theory (cf. [3] § 32), by an *A-Lebesgue decomposition* of a positive operator B we shall mean a decomposition $B=B_c+B_s$ into positive operators such that B_c and B_s are A-absolutely continuous and A-singular, respectively.

In a recent paper [1] ANDERSON and TRAPP proved that given a (closed) subspace $\mathfrak{G}$, each positive operator B is written uniquely as a sum of two positive operators $B=C+D$ such that $\operatorname{ran}(C^{1/2}) \subseteqq \mathfrak{G}$ and $\operatorname{ran}(D^{1/2}) \cap \mathfrak{G} = \{0\}$. Here $C^{1/2}$ is the positive square-root of C, and "ran" stays for "range". If $\operatorname{ran}(A)=\mathfrak{G}$, that is, if A has closed range, then $\operatorname{ran}(C^{1/2}) \subseteqq \mathfrak{G}$ implies $C \leqq \alpha A$ for some $\alpha \geqq 0$ while $\operatorname{ran}(D^{1/2}) \cap \mathfrak{G} = \{0\}$ is equivalent to the A-singularity of D (see § 3). The above cited result shows that A-Lebesgue decomposition is always guaranteed and is unique in case A has closed range.

Received April 2, 1976.

The work was done while the author stayed at Bolyai Institute of Mathematics, Szeged University. The research was supported jointly by the Hungarian Institute of Cultural Relations and the Japan Society for the Promotion of Science. The author wishes to express his gratitude to Professor B. Sz.-Nagy and the Bolyai Institute for their hospitality.

The purpose of this paper is to construct an A-Lebesgue decomposition for each positive operator and to find a condition for the uniqueness of A-Lebesgue decompositions.

2. Lebesgue decomposition

Let us recall a useful binary operation in the class $\mathscr{P}$ of all positive operators, which is defined and called parallel addition by ANDERSON and TRAPP [1]. The *parallel sum* $A:B$ of two positive operators A and B is determined by the formula:

$$((A:B)h, h) = \inf_{g \in \mathfrak{H}} \{(Ag, g)+(B(h-g), h-g)\}.$$

The expression on the right side defines really a positive operator. For, define a new scalar product on the direct sum $\mathfrak{H}\oplus\mathfrak{H}$ by

$$\langle g\oplus k, g'\oplus k'\rangle = (Ag, g')+(Bk, k').$$

Let $\mathfrak{K}$ be the associated Hilbert space and $\mathfrak{G}$ the closure of the manifold $\{g\oplus k : g+k=0\}$. The expression is equal to $\langle (I-P)(0\oplus h), 0\oplus h\rangle$ where P is the projection from $\mathfrak{K}$ onto $\mathfrak{G}$.

Obviously, $A, B \geqq A:B \geqq 0$, and $A_1 \geqq A_2$ implies $A_1:B \geqq A_2:B$. Now since $(nA):B$ increases along with n and is bounded by B from above, we can introduce an operation $[A]$ in the class $\mathscr{P}$ by the formula:

$$[A]B = \lim_{n\to\infty} (nA):B,$$

where lim means strong limit. Since $(nA):B \uparrow [A]B$ and $(nA):B \leqq nA$, by definition $[A]B$ is A-absolutely continuous and $[A]B \leqq B$. Remark that the operation $[A]$ is *monotone* in the sense that $B_1 \leqq B_2$ implies $[A]B_1 \leqq [A]B_2$. This operation is not additive.

The above definition is motivated by a consideration of ANDERSON and TRAPP ([1]; Theorem 12) as well as a proof of the Lebesgue decomposition theorem in measure theory (cf. [3]; § 32).

Lemma 1. *Let A and B be positive operators. Then B is A-absolutely continuous if and only if $[A]B=B$.*

Proof. As remarked above, $[A]B$ is always A-absolutely continuous. Suppose that B is A-absolutely continuous. Then by definition there exists a sequence $\{B_m\}$

such that $B_m \uparrow B$ and $B_m \leqq \alpha_m A$ for some $\alpha_m > 0$. The definition of parallel addition yields, with the convention $0/0=0$, that

$$\begin{aligned}((nA):B_m)h, h) &= \inf_{g\in\mathfrak{H}} \{(nAg, g)+(B_m(h-g), h-g)\} \\ &= (B_m h, h)+\inf_{g\in\mathfrak{H}} \{((nA+B_m)g, g)-2|(B_m g, h)|\} \\ &= (B_m h, h)+\inf_{g\in\mathfrak{H}} \inf_{\lambda>0} \{\lambda^2((nA+B_m)g, g)-2\lambda|(B_m g, h)|\} \\ &= (B_m h, h)-\sup_{g\in\mathfrak{H}} \frac{|(B_m g, h)|^2}{((nA+B_m)g, g)},\end{aligned}$$

hence

$$\begin{aligned}0 &\leqq (B_m h, h)-(((nA):B_m)h, h) \\ &\leqq \sup_{g\in\mathfrak{H}} \frac{(B_m g, g)(B_m h, h)}{(n\alpha_m^{-1}+1)(B_m g, g)} \leqq \frac{\alpha_m}{n+\alpha_m}(Bh, h).\end{aligned}$$

This implies

$$B_m = \lim_{n\to\infty} (nA):B_m \equiv [A]B_m.$$

Now since by the monotonity of the operation $[A]$

$$B \geqq [A]B \geqq [A]B_m = B_m,$$

taking the limit of B_m we have $B=[A]B$. This completes the proof.

Theorem 2. *Let A be a positive operator. Then for each positive operator B the decomposition*

$$B = [A]B+(B-[A]B)$$

is an A-Lebesgue decomposition with A-absolutely continuous $[A]B$ and A-singular $B-[A]B$. Moreover $[A]B$ is the maximum of all A-absolutely continuous positive operators C with $C\leqq B$.

Proof. The operator $[A]B$ is A-absolutely continuous and $[A]B\leqq B$. If a positive operator C is A-absolutely continuous and $C\leqq B$, the monotonity of $[A]$ and Lemma 1 imply that $C=[A]C\leqq[A]B$. Therefore $[A]B$ has the maximum property in question. It remains to show the A-singularity of $B-[A]B$. Suppose that $0\leqq D\leqq A$ and $0\leqq D\leqq B-[A]B$. Since D is obviously A-absolutely continuous, by definition so is the sum $[A]B+D$. On the other hand, the maximum property of $[A]B$ implies $[A]B+D\leqq[A]B$, hence $D=0$. Thus $B-[A]B$ is A-singular by definition. This completes the proof.

Corollary 3. *Let A and B be positive operators. Then B is A-singular if and only if $[A]B=0$.*

3. Characterization of absolute continuity

Some order relations between two positive operators can be expressed in terms of their range spaces. Here a basic tool is supplied by the following lemma due to Douglas ([2] Theorem 2.1).

Lemma 4. *For bounded linear operators S and T the following conditions are mutually equivalent:*
(a) $\operatorname{ran}(S) \subseteqq \operatorname{ran}(T)$,
(b) *There exists $\alpha \geqq 0$ such that $SS^* \leqq \alpha TT^*$,*
(c) *There exists a bounded linear operator R such that $S = TR$. Here R is uniquely determined under the additional requirement that R^* vanishes on the orthocomplement of* $\operatorname{ran}(T^*)$.

When applied to the square roots of positive operators A and B, Lemma 4 yields that $\operatorname{ran}(B^{1/2}) \subseteqq \operatorname{ran}(A^{1/2})$ is equivalent to the existence of $\alpha \geqq 0$ such that $B \leqq \alpha A$, a condition stronger than the A-absolute continuity of B. Lemma 4 shows further that $\operatorname{ran}(A^{1/2}) \cap \operatorname{ran}(B^{1/2}) = \{0\}$ implies the A-singularity of B. Conversely, in view of the general formula

$$\operatorname{ran}(A^{1/2}) \cap \operatorname{ran}(B^{1/2}) = \operatorname{ran}\left((A:B)^{1/2}\right)$$

([1] Theorem 11) and the inequality $0 \leqq A:B \leqq A, B$, the A-singularity of B implies $\operatorname{ran}(A^{1/2}) \cap \operatorname{ran}(B^{1/2}) = \{0\}$. Our purpose in this section is to find a characterization of A-absolute continuity in this direction.

Theorem 5. *Let A and B be positive operators. Then B is A-absolutely continuous if and only if the linear manifold $\{h: B^{1/2}h \in \operatorname{ran}(A^{1/2})\}$ is dense in $\mathfrak{H}$.*

Proof. Suppose that the linear manifold $\mathfrak{D} \equiv \{h: B^{1/2}h \in \operatorname{ran}(A^{1/2})\}$ is dense in $\mathfrak{H}$. Since the orthocomplement of the kernel of $A^{1/2}$ coincides with $\operatorname{ran}(A^{1/2})^-$, the closure of $\operatorname{ran}(A^{1/2})$, the correspondence $h \mapsto g$ from $\mathfrak{D}$ to $\operatorname{ran}(A^{1/2})^-$, defined by $B^{1/2}h = A^{1/2}g$, determines a linear operator T with domain $\mathfrak{D}$. As easily follows from the boundedness of $A^{1/2}$ and $B^{1/2}$ ([2] Theorem 2.1), T is closed. Now since T is a densely defined closed operator, its adjoint T^* is a densely defined closed operator (cf. [4]; V, § 3.1). Since $A^{1/2}T \subseteqq B^{1/2}$ by definition, the boundedness of $A^{1/2}$ and $B^{1/2}$ yields $T^*A^{1/2} = B^{1/2}$. Let $T^* = VS$ be the polar decomposition of T^* (cf. [4]; VI, § 2,7); S is an (unbounded) positive self-adjoint operator whose domain coincides with that of T^* and V is a partial isometry with initial space $\operatorname{ran}(S)^-$ and final space $\operatorname{ran}(T^*)^-$. Then $\operatorname{ran}(A^{1/2})$ is included in the domain of S, and for all $h \in \mathfrak{H}$

$$\|SA^{1/2}h\|^2 = (Bh, h).$$

Consider the spectral representation

$$S = \int_0^\infty \lambda \, dE(\lambda) \quad \text{and let} \quad S_n = \int_0^n \lambda \, dE(\lambda) \qquad (n = 1, 2, \ldots).$$

Then we can readily verify that $A^{1/2} S_n^2 A^{1/2} \uparrow B$ and $A^{1/2} S_n^2 A^{1/2} \leqq n^2 A$, hence B is A-absolutely continuous.

Suppose conversely that B is A-absolutely continuous. Then by definition there exists a sequence $\{B_n\}$ such that $B_n \uparrow B$ and $B_n \leqq \alpha_n A$ for some $\alpha_n \geqq 0$. By Lemma 4 for each n there exists a bounded linear operator R_n such that $B_n^{1/2} = A^{1/2} R_n$ and R_n^* vanishes on the orthocomplement of $\operatorname{ran}(A^{1/2})$. Then $B_n \leqq B_{n+1}$ implies $R_n R_n^* \leqq R_{n+1} R_{n+1}^*$. Let $\mathfrak{D}$ denote the linear manifold of all g with $\sup_n \|R_n^* g\| < \infty$, and define a functional φ on $\mathfrak{D}$ by the formula

$$\varphi(g) \equiv \sup_n \|R_n^* g\|^2 = \lim_{n \to \infty} \|R_n^* g\|^2.$$

The functional φ is closed in the sense that if $g_n \in \mathfrak{D}$, $\lim_{n \to \infty} g_n = h$ and if $\lim_{n, m \to \infty} \varphi(g_n - g_m) = 0$, then $h \in \mathfrak{D}$ and $\lim_{n \to \infty} \varphi(h - g_n) = 0$. Further, since, by definition of $\{B_n\}$, for all $h \in \mathfrak{D}$

$$\sup_n \|R_n^* A^{1/2} h\|^2 = \sup_n \|B_n^{1/2} h\|^2 = (Bh, h) < \infty$$

and since every R_n^* vanishes on the orthocomplement of $\operatorname{ran}(A^{1/2})$, the linear manifold $\mathfrak{D}$ includes the dense set $\operatorname{ran}(A^{1/2}) + (\mathfrak{H} \ominus \operatorname{ran}(A^{1/2}))$. Thus φ is densely defined, closed and expressed as the limit of the bounded quadratic forms $\|R_n^* g\|^2$. Now in view of a theorem on quadratic forms ([4]; VI, § 2,6) there exists an (unbounded) positive self-adjoint operator S such that its domain coincides with $\mathfrak{D}$ and $\|Sg\|^2 = \varphi(g)$. Then we have for all $h \in \mathfrak{H}$

$$\|SA^{1/2} h\|^2 = (Bh, h) = \|B^{1/2} h\|^2,$$

hence there exists a partial isometry V with initial space $\operatorname{ran}(B^{1/2})^-$ such that $SA^{1/2} = VB^{1/2}$. This implies $A^{1/2} S \subseteqq B^{1/2} V^*$, and consequently

$$V^*(\mathfrak{D}) \subseteqq \{h : B^{1/2} h \in \operatorname{ran}(A^{1/2})\}.$$

Since $\mathfrak{D}$ is dense in $\mathfrak{H}$ and V is a partial isometry with initial space $\operatorname{ran}(B^{1/2})^-$, we can conclude

$$\operatorname{ran}(B^{1/2})^- \subseteqq \{h : B^{1/2} h \in \operatorname{ran}(A^{1/2})\}^-.$$

Finally since $B^{1/2}$ vanishes on the orthocomplement of $\operatorname{ran}(B^{1/2})$, the subspace $\{h : B^{1/2} h \in \operatorname{ran}(A^{1/2})\}^-$ includes this orthocomplement, too, hence coincides with the whole space $\mathfrak{H}$. This completes the proof.

4. Uniqueness condition

Let A be a positive operator. Then A-absolute continuity is *additive* in the sense that the sum of two positive operators is A-absolutely continuous whenever both summands are so. A-singularity is not always additive while it is *hereditary* in the sense that A-singularity of the sum of two positive operators implies A-singularity of both summands. A-absolute continuity is not always hereditary. These discrepancies can cause non-uniqueness in A-Lebesgue decomposition.

Let us say a positive operator B to be *A-strongly continuous* if $B\leqq\alpha A$ for some $\alpha\geqq 0$, or equivalently, as is remarked in § 3, if $\operatorname{ran}(B^{1/2})\subseteqq\operatorname{ran}(A^{1/2})$. Then *$A$-strong continuity* is additive as well as hereditary.

Theorem 6. *Let A be a positive operator. Then a positive operator B admits a unique A-Lebesgue decomposition if and only if $[A]B$ is A-strongly continuous, that is, $[A]B\leqq\alpha A$ for some $\alpha\geqq 0$.*

Proof. Suppose that $[A]B$ is A-strongly continuous and take an arbitrary A-Lebesgue decomposition $B\doteq C+D$ with A-absolutely continuous C and A-singular D. Theorem 2 implies $D\geqq[A]B-C\geqq 0$. The positive operator $[A]B-C$ is A-strongly continuous as well as A-singular so that it must be equal to 0. Therefore B admits a unique A-Lebesgue decomposition.

Suppose conversely that $[A]B$ is not A-strongly continuous. Then by Lemma 1, Lemma 4 and Theorem 5 the linear manifold $\mathfrak{D}\equiv\{h;\ ([A]B)^{1/2}h\in\operatorname{ran}(A^{1/2})\}$ is dense in $\mathfrak{H}$ but not closed. As in the proof of Theorem 5 there exists a closed operator with domain $\mathfrak{D}$, so that there exists a (bounded) positive operator S with $\operatorname{ran}(S)=\mathfrak{D}$ (cf. [2]; Theorem 1.1). We may assume $S^2\leqq\frac{1}{2}I$. Since $\operatorname{ran}(S)$ is not closed and $[A]B\neq 0$ by assumption, there exists a separable (closed) subspace $\mathfrak{G}$ such that $SP=$ $=PS$, $([A]B)\cdot P=P\cdot([A]B)\neq 0$ and $\operatorname{ran}(SP)$ is not closed, where P is the ortho-projection onto $\mathfrak{G}$. Then in view of a theorem of von Neumann ([2] Theorem 3.6) there exists a unitary operator U_0 on the separable Hilbert space $\mathfrak{G}$ such that

$$\operatorname{ran}(SP)\cap\operatorname{ran}(U_0SP)=\{0\}.$$

Let us define a unitary operator U on $\mathfrak{H}$ by $U=U_0P+(I-P)$. Then it follows from the properties of $\mathfrak{G}$ and U_0 that

$$\mathfrak{D}\cap U^*(\mathfrak{D})\subseteqq\mathfrak{H}\ominus\mathfrak{G}.$$

Consider the positive operators defined by

$$D\equiv([A]B)^{1/2}U^*S^2U([A]B)^{1/2}\quad\text{and.}\quad C\equiv[A]B-D.$$

First we shall show that C is A-absolutely continuous. Since

$$[A]B\geqq C=([A]B)^{1/2}U^*(I-S^2)U([A]B)^{1/2}\geqq\frac{1}{2}[A]B,$$

by Lemma 4 (cf. [2]; Corollary 2.1.1) there exists a bounded invertible operator R such that $C^{1/2}R=([A]B)^{1/2}$. Then we have

$$\{h:C^{1/2}h\in\operatorname{ran}(A^{1/2})\}=R(\mathfrak{D}).$$

Since $\mathfrak{D}$ is dense in $\mathfrak{H}$ and R is invertible, $R(\mathfrak{D})$ is dense in $\mathfrak{H}$ too, so that the above relation implies the A-absolute continuity of C by Theorem 5.

Let us prove that D is not A-absolutely continuous. Suppose the contrary. Then Theorem 5 implies that $\operatorname{ran}(D^{1/2})\cap\operatorname{ran}(A^{1/2})$ is dense in $\operatorname{ran}(D^{1/2})$. On the other hand, by Lemma 4 and definition of D we have

$$\operatorname{ran}(D^{1/2})\cap\operatorname{ran}(A^{1/2})=\operatorname{ran}\big(([A]B)^{1/2}U^*S\big)\cap\operatorname{ran}(A^{1/2}).$$

Take an arbitrary h such that $([A]B)^{1/2}U^*Sh\in\operatorname{ran}(A^{1/2})$. This requirement is equivalent to $U^*Sh\in\mathfrak{D}$ by the definition of $\mathfrak{D}$. Since $\operatorname{ran}(S)=\mathfrak{D}$, it follows that

$$([A]B)^{1/2}U^*Sh\in([A]B)^{1/2}(\mathfrak{D}\cap U^*(\mathfrak{D}))\subseteqq([A]B)^{1/2}(\mathfrak{H}\ominus\mathfrak{G}).$$

Since $\mathfrak{H}\ominus\mathfrak{G}$ reduces $[A]B$, we can conclude

$$\operatorname{ran}(D^{1/2})\cap\operatorname{ran}(A^{1/2})\subseteqq\mathfrak{H}\ominus\mathfrak{G}.$$

Finally since P commutes with S, U and $[A]B$, the subspace $\mathfrak{G}$ reduces $D^{1/2}$ and $D^{1/2}(\mathfrak{G})\neq\{0\}$ according to $([A]B)P\neq 0$. Therefore the above inclusion relation leads to a contradiction that $\operatorname{ran}(D^{1/2})\cap\operatorname{ran}(A^{1/2})$ is not dense in $\operatorname{ran}(D^{1/2})$.

Now consider a decomposition $B=C_1+D_1$, where $C_1=C+[A]\{D+(B-[A]B)\}$ and $D_1=B-C_1$. This is an A-Lebesgue decomposition. In fact, obviously C_1 is positive A-absolutely continuous while D_1 is positive A-singular by Theorem 2, because

$$D_1=\{D+(B-[A]B)\}-[A]\{D+(B-[A]B)\}.$$

Finally C_1 does not coincide with $[A]B$. For otherwise the relation

$$[A]\{D+(B-[A]B)\}=[A]B-C=D$$

would imply the A-absolute continuity of D by Theorem 2, which is a contradiction. Thus B admits an A-Lebesgue decomposition different from the one given in Theorem 2. This completes the proof of the theorem.

Corollary 7. *The following conditions for a positive operator A are mutually equivalent:*

(a) *$\operatorname{ran}(A)$ is closed,*

(b) *A-absolute continuity is hereditary,*

(c) *Each positive operator admits a unique A-Lebesgue decomposition.*

Proof. (a)⇒(b) is immediate, because under the closedness of $\operatorname{ran}(A)$ it is easy to prove the equivalence of A-absolute continuity and A-strong continuity. (b)⇒(c) is proved just as in the first part of the proof of Theorem 6. (c)⇒(a): Let P be the orthoprojection onto the closure of $\operatorname{ran}(A)$. Then obviously P is A-absolutely continuous. Now (c) implies by Theorem 6 that $P\leqq\alpha A$ for some $\alpha\geqq 0$, which is equivalent to the closedness of $\operatorname{ran}(A)$. This completes the proof.

References

[1] W. N. ANDERSON and G. E. TRAPP, Shorted operators. II, *SIAM J. Appl. Math.*, **28** (1975), 60—71.

[2] P. A. FILLMORE and J. P. WILLIAMS, On operator ranges, *Advances in Math.*, **7** (1971), 254—281.

[3] P. R. HALMOS, *Measure theory*, Springer (Berlin—Heidelberg—New York, 1974).

[4] T. KATO, *Perturbation theory for linear operators*, Springer (Berlin—Heidelberg—New York, 1966).

DIVISION OF APPLIED MATHEMATICS
RESEARCH INSTITUTE OF APPLIED ELECTRICITY
HOKKAIDO UNIVERSITY
SAPPORO, JAPAN

Acta Sci. Math., 38 (1976), 261—263

Strongly reductive operators are normal

CONSTANTIN APOSTOL, CIPRIAN FOIAŞ and DAN VOICULESCU

An operator on a Hilbert space $\mathfrak{H}$ is called reductive if every subspace $\mathfrak{L}$ [1]) invariant for T is also invariant for T^* (i.e. $\mathfrak{L}$ is reducing T). By a theorem of DYER, PEDERSEN and PORCELLI [4] every reductive operator is normal if and only if every operator has a (non-trivial) invariant subspace. Therefore the study of reductive operators might shed some light into the intricate structure of general operators. In particular it looked instructive to study a natural subclass of reductive operators [6], [2]. Let us recall that an operator T on $\mathfrak{H}$ is called *strongly reductive* if

$$\varepsilon_T(\delta) = \sup\{\|(I-P)T^*P\| : \|(I-P)TP\| < \delta\}$$

tends to 0 for $\delta \searrow 0$; P runs through the family $\mathscr{P}_{\mathfrak{H}}$ of orthogonal projections in $\mathfrak{H}$. Concerning this concept, the following was proved by HARRISON [6] (Cor. 2.4. and Thm. 3.8).

Proposition. *If T is strongly reductive then its spectrum $\sigma(T)$ neither divides the (complex) plane nor has interior (in the plane). These conditions on $\sigma(T)$ imply, in case T is normal, that T is strongly reductive.*

The aim of this short Note is to supplement these results with the following.

Theorem. *Every strongly reductive operator is normal.*

We will divide the proof of this theorem in several steps:

1. Lemma. *Let T be a strongly reductive operator on $\mathfrak{H}$ and let X be an operator on some space $\mathfrak{K}$ such that $\|X-U_jTU_j^{-1}\| \to 0$ $(j\to\infty)$ where U_j $(j=1, 2, \ldots)$ are unitary operators from $\mathfrak{H}$ onto $\mathfrak{K}$. Then X is also strongly reductive.*

Proof. For $\delta > 0$ let $P \in \mathscr{P}_{\mathfrak{K}}$ be such that $\|(I-P)XP\| < \delta$. Denote $T_j = U_jTU_j^{-1}$ and take j large enough such that $\|X-T_j\| < \delta - \|(I-P)XP\|$. Then for $P_j = U_j^{-1}PU_j$

Received February 28, 1976.

1) All the spaces involved are complex Hilbert spaces; the subspaces will be always considered linear and closed. Also all operators will be linear, continuous, and mapping Hilbert spaces into Hilbert spaces.

we have $P_j \in \mathscr{P}_{\mathfrak{H}}$ and $\|(I-P_j)TP_j\| < \delta$ so that

$$\|(I-P)X^*P\| \leqq \|X^*-T_j^*\| + \|(I-P)T_j^*P\| =$$
$$= \|X-T_j\| + \|(I-P_j)T^*P_j\| \leqq \|X-T_j\| + \varepsilon_T(\delta),$$

whence (letting $j\to\infty$), $\|(I-P)X^*P\| \leqq \varepsilon_T(\delta)$.

2. Lemma. *Let T be a strongly reductive operator on a separable space $\mathfrak{H}$. Then T^*T-TT^* is compact.*

Proof. Let $\mathscr{B}$ be the C^*-algebra with unity, generated in the Calkin algebra $C(\mathfrak{H})$ [2]) by the image $\tilde{T}$ of T. Let moreover ϱ be a faithful C^*-representation of $\mathscr{B}$ on a separable Hilbert space $\mathfrak{H}_\varrho$. [3]) By virtue of [8], Thm. 1.3, we can take the operator X in Lemma 1 of the form $X = T \oplus \varrho(\tilde{T}) \oplus \varrho(\tilde{T})$; therefore this operator is strongly reductive, henceforth reductive. But if P denotes the orthogonal projection of $\mathfrak{H}\oplus\mathfrak{H}_\varrho\oplus\mathfrak{H}_\varrho$ onto $\{0\oplus h\oplus\varrho(\tilde{T})h : h\in\mathfrak{H}_\varrho\}$ then $(I-P)XP=0$, thus also $\|(I-P)X^*P\|=0$. Whence we easily infer that $\varrho(\tilde{T})^*\varrho(\tilde{T})h = \varrho(\tilde{T})\varrho(\tilde{T})^*h$ for all $h\in\mathfrak{H}_\varrho$, i.e. $\varrho(\tilde{T}^*\tilde{T}-\tilde{T}\tilde{T}^*)=0$, $\widetilde{T^*T-TT^*} = \tilde{T}^*\tilde{T}-\tilde{T}\tilde{T}^*=0$.

3. Lemma. *Let T be a strongly reductive operator on $\mathfrak{H}$. Then, if* $\dim\mathfrak{H}>1$, *there exists a (non-trivial) subspace of $\mathfrak{H}$, invariant for T (thus also reducing T).*

Proof. Since, if $\dim\mathfrak{H}<\infty$ then T is obviously normal and if $\dim\mathfrak{H}>\aleph_0$ then T is obviously reduced by separable subspaces of $\mathfrak{H}$, it remains to consider only the case $\dim\mathfrak{H}=\aleph_0$. In this case, the properties of $\sigma(T)$ (yielded by Harrison's Proposition) together with the spectral characterization of quasitriangular operators [3], Thm. 5.4, imply that T is quasi-triangular. Therefore if $\|p(T)\| \neq \|\widetilde{p(T)}\|$ for some polynomial $p(\lambda)$, the existence of (non-trivial) subspaces reducing T is already established in [2]. Thus we can assume that

$$\|p(T)\| = \|\widetilde{p(T)}\| = \|p(\tilde{T})\| \tag{1}$$

for all polynomials $p(\lambda)$. But in virtue of Lemma 2, T is normal in $C(\mathfrak{H})$, thus

$$\|p(T)\| = \|p\|_{C(\sigma(\tilde{T}))} (:= \max\{|p(\lambda)| : \lambda\in\sigma(\tilde{T})\}), \tag{2}$$

where $\sigma(\tilde{T})(\subset\sigma(T))$ neither separates the plane nor has interior. By (1), (2) and by virtue of the classical theorem of LAVRENTIEV [5], Ch. II, 8.7, the map $p|_{\sigma(\tilde{T})}\mapsto p(T)$ extends to an isometric algebraic map of $C(\sigma(\tilde{T}))$ in $L(\mathfrak{H})$. Consequently, if $\sigma(\tilde{T})$

---

[2]) This is the quotient C^*-algebra $C(\mathfrak{H}) = L(\mathfrak{H})/K(\mathfrak{H})$, where $L(\mathfrak{H})$ denotes the algebra of all operators on $\mathfrak{H}$ while $K(\mathfrak{H})$ denotes the ideal of all compact operators on $\mathfrak{H}$. We shall denote the element $X+K(\mathfrak{H})$ $(X\in L(\mathfrak{H}))$ in $C(\mathfrak{H})$ by $\tilde{X}$.

[3]) The existence of such a representation follows easily from the separability of $\mathscr{B}$ and the classical Gelfand — Naĭmark theorem [7], Ch. V., § 24, Sec. 2.

reduces to a single point λ, then $T=\lambda$, if not then taking two continuous functions f and g on $\sigma(\tilde{T})$ not vanishing identically and such that $fg=0$ we have $f(T)\neq 0$, $g(T)\neq 0$, $f(T)g(T)=0$, and T leaves invariant the (non-trivial) null-spaces of $f(T)$ and $g(T)$.

4. We are now in state to achieve the proof of the theorem. As in the proof of Lemma 3 we can assume that $\mathfrak{H}$ is separable. Also we can discard from $\mathfrak{H}$ the largest reducing subspace $\mathfrak{L}$ of $\mathfrak{H}$ on which $T|\mathfrak{L}$ is normal (see [1]). Therefore, in case T is not normal we can assume that for any subspace $\mathfrak{L}\subset\mathfrak{H}$, reducing T, the operator $T|\mathfrak{L}$ is not normal; it follows that for such subspaces $\mathfrak{L}$ we have $\dim \mathfrak{L}=\aleph_0$. Using these facts together with Lemma 3 we can prove that for any maximal totally ordered family $\mathcal{F}$ of invariant subspaces $\mathfrak{K}$ for T and for every $\mathfrak{K}_0\in\mathcal{F}$ the continuity properties

$$\bigvee\{\mathfrak{K} : \mathfrak{K} \subsetneqq \mathfrak{K}_0,\ \mathfrak{K}\in\mathcal{F}\} = \mathfrak{K}_0 = \bigcap\{\mathfrak{K} : \mathfrak{K} \supsetneqq \mathfrak{K}_0,\ \mathfrak{K}\in\mathcal{F}\}$$

hold. Moreover, $\{0\}$ and $\mathfrak{H}$ belong to $\mathcal{F}$. As T is (strongly) reductive the subspaces $\mathfrak{K}$ reduce T, and therefore, $C=T^*T-TT^*$ too. Since T is not normal, $C\neq 0$. On the other hand, by Lemma 2 the operator C is compact so that it has a finite dimensional non-zero eigen-subspace $\mathfrak{L}$. Then the corresponding orthogonal projection $P_{\mathfrak{L}}$ is reduced by each $\mathfrak{K}\in\mathcal{F}$. Consequently, $\mathcal{F}'=\{\mathfrak{K}\cap\mathfrak{L}:\ \mathfrak{K}\in\mathcal{F}\}$ has the same continuity properties as $\mathcal{F}$. This contradicts the finite dimensionality of $\mathfrak{L}$.

References

[1] C. Apostol, Sur la partie normale d'un ensemble d'opérateurs de l'espace de Hilbert, *Acta Math. Acad. Sci. Hung.* **17** (1966), 1—4.

[2] C. Apostol and C. K. Fong, Invariant suspaces for algebras generated by strongly reductive operators, *Duke Math. J.*, **42** (1975), 495—498.

[3] C. Apostol, C. Foiaş and D. Voiculescu, Some results on non-quasitriangular operators, *Revue Roum. Math. Pures et Appl.*, **18** (1973), 487—514.

[4] J. A. Dyer, E. A. Pedersen, and P. Porcelli, An equivalent formulation of the invariant subspace conjecture, *Bull. Amer. Math. Soc.*, **78** (1972), 1020—1023.

[5] T. W. Gamelin, *Uniform Algebras*, Prentice-Hall (Englewood Cliffs, N. J., 1969).

[6] K. J. Harrison, Strongly reductive operators, *Acta Sci. Math.*, **37** (1975), 205—212.

[7] M. A. Naimark, *Normed Rings*, Gosud. Izd. Tehn.-Teor. Lit. (Moscow, 1956). (Russian)

[8] D. Voiculescu, A non-commutative Weyl — von Neumann theorem, *Revue Roum. Math. Pures et Appl.*, **21** (1976), 97—113.

FACULTY OF MATHEMATICS
UNIVERSITY BUCHAREST
STR. ACADEMIEI 14
BUCHAREST, ROMANIA

Acta Sci. Math., 38 (1976), 265—274

On thin operators relative to an ideal in a von Neumann algebra

BRUCE A. BARNES

§ 1. Introduction

Let A be a von Neumann algebra, let Z be the center of A, and let K be a proper closed ideal of A with the property that if $T \in A$ and $TK = \{0\}$, then $T=0$. The set of thin operators of A relative to K, denoted $\mathfrak{I}_K$, is the set of operators of the form $X+T$ where $X \in Z$ and $T \in K$. In the case where $A = B(\mathfrak{H})$, the algebra of all bounded linear operators on a Hilbert space $\mathfrak{H}$, and $K = K(\mathfrak{H})$, the closed ideal of compact operators in $B(\mathfrak{H})$, this definition is due to R. DOUGLAS and C. PEARCY [6]. Let θ_K be the collection of all projections in K. If $P, Q \in \theta_K$, then $P \vee Q \in \theta_K$. This follows from [11, Lemma 2.1] where the proof is given for the more general case when A is an AW^*-algebra. Thus θ_K is upward directed in the usual ordering of projections ($P \leqq Q$ means $PQ = QP = P$). In [6], DOUGLAS and PEARCY characterized the thin operators in $B(\mathfrak{H})$ relative to $K(\mathfrak{H})$ as the set of all operators T that satisfy

$$\lim_{P \in \theta_K} \|PTP - TP\| = 0$$

[6, Theorem 2]. Also in [6], they related the η function of A. BROWN and C. PEARCY [4], [10], to

$$\lim_{P \in \theta_K} \sup \|PTP - TP\|.$$

They asked if there is a suitable extension of these results to the case where A is a general von Neumann algebra.

In a series of papers [7], [8] C. OLSEN proved the Douglas—Pearcy characterization of the thin operators in the general case. Also, she conjectured [8, p. 572]. that the distance from $T \in A$ to $\mathfrak{I}_K$ is given by

$$\lim_{P \in \theta_K} \sup \|PTP - TP\|.$$

Received July 29, 1975, revised March 20, 1976.

That this conjecture holds when $A=B(\mathfrak{H})$ and $K=K(\mathfrak{H})$ was proved by C. Apostol, C. Foiaş, and L. Zsidó in [1].

In [2], for A a von Neumann algebra or a C^*-factor, C. Apostol and L. Zsidó made a systematic study of the relationship between the distance of an element $T\in A$ from $\mathfrak{J}_K$, the η function evaluated at T, and the norm of the inner derivation induced on A by T.

In this paper we make three contributions to this circle of ideas. First, in § 2 we give a new proof that when A is a von Neumann algebra, then $T\in A$ is in $\mathfrak{J}_K$ if and only if

$$\lim_{P\in\theta_K} \|TP-PT\| = 0.$$

We note in this connection that C. Olsen proves [8, Theorem 2] that it is always the case that

$$\lim_{P\in\theta_K} \sup \|PTP-TP\| = \lim_{P\in\theta_K} \sup \|TP = PT\|.$$

Our proof depends only on elementary arguments, and is considerably shorter than the proof by Olsen in [7], [8]. Second, in § 3 we introduce a nonspatial form of the η function of Brown and Pearcy [4], [10]. The generalized function η is defined on A using pure states of A, and is completely independent of any particular representation of A as a von Neumann algebra of operators on a Hilbert space. We prove some of the elementary properties of η in § 3. Then in § 4 we prove that $\eta(T)$ measures the distance from T to $\mathfrak{J}_K$. This is a generalization of [1, Lemma 1.1]. Third, in § 4 we prove the conjecture of C. Olsen that the distance from T to $\mathfrak{J}_K$ is given by

$$\lim_{P\in\theta_K} \sup \|TP-PT\| = \lim_{P\in\theta_K} \sup \|PTP-TP\|.$$

This result provides another proof of the Douglas—Pearcy—Olsen characterization of $\mathfrak{J}_K$.

At this point we introduce some notation. Throughout this paper A, Z, K, θ_K, and $\mathfrak{J}_K$ will be as stated at the beginning of this §. The identity operator in A is denoted by I. If B is a subalgebra of A and P is a projection in A, then $B_P=PBP$. Also, if $T\in A$, then $T_P=PTP$. The distance of $T\in A$ from a subspace $B\subset A$ is denoted $d(T, B)$, i.e.,

$$d(T, B) = \inf\{\|T+S\| : S\in B\}.$$

The set of pure states of A is denoted P_A. If $\alpha\in P_A$, then let Φ_α be the irreducible representation determined by α, and let $\mathfrak{H}_\alpha$ be the corresponding representation space. The inner product of vectors $\xi, \tau\in\mathfrak{H}_\alpha$ is denoted by $\langle\xi, \tau\rangle$.

If $P\in\theta_K$, then let

$$\Delta(P) = \{\alpha\in P_A : \alpha(K) \neq \{0\} \quad\text{and}\quad \alpha(P) = 0\}.$$

The collection $\Delta(P)$ plays an important role in later sections. Now we verify that $\Delta(P)$ is nonempty. For assume that $P\in\theta_K$. If $AP=K$, then $K(I-P)=\{0\}$. This implies that $P=I$, a contradiction. Thus, $AP\subset K$ and $AP\neq K$. By [5, Théorème 2.9.5] there exists a maximal left ideal M of A such that $AP\subset M$ and $K\not\subset M$. By this same result it follows that there exists $\alpha\in P_A$ such that $\alpha(AP)=\{0\}$ and $\alpha(K)\neq\{0\}$. Therefore $\alpha\in\Delta(P)$.

§ 2. The characterization of the thin operators

In this § we give a new proof of the Douglas—Pearcy—Olsen characterization of $\mathfrak{I}_K$ [6], [7], [8]. The main tool in the proof is a result of the present author [3, Lemma 6.1]. Before proving the characterization, we state this result.

2.1. Assume $\alpha\in P_A$ and $T_k\in A$, $1\leqq k\leqq m$. Then there exists a sequence of nonzero projections $\{E_n\}\subset A$ such that for $1\leqq k\leqq m$,

$$\lim_{n\to\infty}\|E_nT_kE_n-\alpha(T_k)E_n\|=0.$$

This result is established in [3] using completely elementary arguments.

Theorem 2.2. *$T\in\mathfrak{I}_K$ if and only if* $\lim_{P\in\theta_K}\|TP-PT\|=0.$

Proof. If $T\in\mathfrak{I}_K$, then it is straigthforward to prove

$$\lim_{P\in\theta_K}\|TP-PT\|=0; \tag{1}$$

see the proof of [7, Proposition 2.1]. We prove the converse. Assume that (1) holds. Let $\varepsilon>0$ be arbitrary. Choose $Q\in\theta_A$ such that $P\in\theta_A$, $P\geqq Q$ implies that $\|TP-PT\|<\varepsilon$. Assume $R\in\theta_K$ and $R\leqq(I-Q)$. Then $R+Q\in\theta_A$ and $R+Q\geqq Q$. Thus, by the choice of Q, we have $\|T(R+Q)-(R+Q)T\|<\varepsilon$ and $\|TQ-QT\|<\varepsilon$. Therefore, $\|TR-RT\|<2\varepsilon$. This proves

$$\text{if}\quad R\in\theta_K\quad\text{and}\quad R\leqq I-Q,\quad\text{then}\quad\|TR-RT\|<2\varepsilon. \tag{2}$$

Let α be any pure state of A such that $\alpha(K)=\{0\}$. Then α restricts to a pure state of A_{I-Q}. Let S be any operator in A. Consider the elements of A_{I-Q}, $T_1=T_{I-Q}$, $T_2=S_{I-Q}$, and $T_3=(TS)_{I-Q}$. Applying (2.1) to the operators $T_k\in A_{I-Q}$, $1\leqq k\leqq 3$, we have that there exists a sequence of nonzero projections $\{E_n\}$ in A_{I-Q} such that for $k=1,2,3$

$$\|E_nT_kE_n-\alpha(T_k)E_n\|\to 0\quad\text{as}\quad n\to\infty.$$

Note that since $\alpha(Q)=0$, we have $\alpha(R_{I-Q})=\alpha(R)$ for all $R\in A$. Therefore,

$$\|E_nTE_n-\alpha(T)E_n\|\to 0,\quad \|E_nSE_n-\alpha(S)E_n\|\to 0,\quad \|E_nTSE_n-\alpha(TS)E_n\|\to 0. \tag{3}$$

Now E_nKE_n is a nonzero closed ideal in the von Neumann algebra E_nAE_n. Therefore for each n we can choose a nonzero projection $F_n \in E_nKE_n \subset K$. Thus $F_n \leqq E_n \leqq I-Q$ for each $n \geqq 1$. It follows immediately from (2) that

$$\|TF_n - F_nT\| < 2\varepsilon \qquad (n \geqq 1). \tag{4}$$

Since $F_n \leqq E_n$ for all n, we have by (3) that

$$\|F_nTF_n - \alpha(T)F_n\| \to 0, \quad \|F_nSF_n - \alpha(S)F_n\| \to 0, \quad \|F_nTSF_n - \alpha(TS)F_n\| \to 0. \tag{5}$$

Now,

$$|\alpha(T)\alpha(S) - \alpha(TS)| = \|\alpha(T)\alpha(S)F_n - \alpha(TS)F_n\| \leqq$$
$$\leqq \|F_nTSF_n - \alpha(TS)F_n\| + \|F_nTSF_n - F_nTF_nSF_n\| + \|F_nTF_nSF_n - \alpha(T)\alpha(S)F_n\|.$$

The first and third terms of the sum on the right hand side of this inequality approach zero by (5). Also,

$$\|F_nTSF_n - F_nTF_nSF_n\| = \|F_nT(I-F_n)SF_n\| \leqq \|F_nT - TF_n\|\,\|S\| \leqq 2\varepsilon\|S\|$$

for all $n \geqq 1$, by (4). Therefore, $|\alpha(T)\alpha(S) - \alpha(TS)| < 2\varepsilon\|S\|$, and since $\varepsilon > 0$ is arbitrary,

$$\alpha(TS) = \alpha(T)\alpha(S).$$

A similar proof shows that for all $S \in A$,

$$\alpha(ST) = \alpha(S)\alpha(T) = \alpha(TS).$$

Thus $\alpha(ST - TS) = 0$ for all $S \in A$ and all $\alpha \in P_A$ with $\alpha(K) = \{0\}$. Therefore T commutes with A modulo K, i.e. the natural quotient map of A onto A/K maps T into the center of A/K. Then by [5, Exercise 7, p. 259], $T \in \mathfrak{I}_K$.

§ 3. The nonspatial from of the η function

In [4], A. Brown and C. Pearcy define a function η on the von Neumann algebra $A = B(\mathfrak{H})$ relative to the ideal K of compact operators by the formula

$$\eta(T) = \inf_{P \in \theta_K} (\sup \{\|T\xi - (T\xi, \xi)\xi\| : \xi \in \mathfrak{H},\ \|\xi\| = 1,\ P\xi = 0\}) \tag{1}$$

If $\xi \in \mathfrak{H}$, $\|\xi\| = 1$, then let ω_ξ be the pure state of $B(H)$ given by $\omega_\xi(T) = (T\xi, \xi)$. Observe that

$$\|T\xi - (T\xi, \xi)\xi\|^2 = \omega_\xi(T^*T) - |\omega_\xi(T)|^2.$$

In this case, $\{\omega_\xi : \xi \in H, \|\xi\| = 1\}$ is exactly the set of pure states α of A with the property that $\alpha(K) \neq \{0\}$. If $\alpha \in P_A$ and $P \in \theta_K$, then we use the notations

$$\gamma(\alpha, T) = (\alpha(T^*T) - |\alpha(T)|^2)^{1/2} \quad (T \in A), \quad \Delta(P) = \{\alpha \in P_A : \alpha(K) \neq \{0\}, \alpha(P) = 0\}. \tag{2}$$

Recall from the Introduction that $\Delta(P)$ is nonempty. With the notation above the formula in (1) takes the form

(3) $$\eta(T) = \inf_{P\in\theta_K} (\sup\{\gamma(\alpha, T) : \alpha\in\Delta(P)\}).$$

Now θ_K is an upward directed set. For a fixed $T\in A$, the net

$$P \to \sup\{\gamma(\alpha, T) : \alpha\in\Delta(P)\}$$

is decreasing on θ_K. Thus,

$$\eta(T) = \lim_{P\in\theta_K} (\sup\{\gamma(\alpha, T) : \alpha\in\Delta(P)\}).$$

In general, if A is a von Neumann algebra and K is a closed ideal of A, then the definitions in (2) and (3) make sense. In particular, (3) is a generalized nonspatial expression of the useful η function of Brown and Pearcy. At times, in order to indicate the dependence of the function η on the ideal K, we write η_K in place of η. In this § we derive the elementary properties of the function η, while in the next §, we show that $\eta_K(T)$ measures the distance of an operator $T\in A$ from the thin operators relative to K.

Since for any $\alpha\in P_A$ we have $\gamma(\alpha, T)^2 \leqq \alpha(T^*T) \leqq \|T\|^2$, it follows that

(3.1) $$\eta(T) \leqq \|T\| \quad (T\in A).$$

Next we show that

(3.2) $$T \to \eta(T) \quad \text{is a seminorm on} \quad A.$$

That $\eta(\lambda T) = |\lambda|\eta(T)$, $T\in A$, λ a scalar, is obvious. Since α is a positive functional on A, we have

(4) $$\alpha((C+B)^*(C+B))^{1/2} \leqq \alpha(C^*C)^{1/2} + \alpha(B^*B)^{1/2},$$

for all $C, B\in A$. Also, note that

$$\gamma(\alpha, T) = \alpha((T^* - \overline{\alpha(T)}I)(T - \alpha(T)I))^{1/2}.$$

Thus, setting $C = T - \alpha(T)I$ and $B = S - \alpha(S)I$ in (4), we have $\gamma(\alpha, T+S) \leqq \gamma(\alpha, T) + \gamma(\alpha, S)$. Therefore,

$$\sup_{\alpha\in\Delta(P)} \gamma(\alpha, T+S) \leqq \left(\sup_{\alpha\in\Delta(P)} \gamma(\alpha, T) + \sup_{\alpha\in\Delta(P)} \gamma(\alpha, S)\right).$$

Taking limits over $P\in\theta_K$ we have $\eta(T+S) \leqq \eta(T) + \eta(S)$.

(3.3) $$\text{If} \quad T\in A \quad \text{and} \quad S\in K, \quad \text{then} \quad \eta(T+S) = \eta(T).$$

To prove (3.3) first observe that $\eta(P) = 0$ whenever $P\in\theta_K$. Since η is a seminorm, it follows that if L is any finite linear combination of projections in θ_K, then $\eta(L) = 0$.

Now assume that $S \in K$. Let $\varepsilon > 0$ be arbitrary. Choose L, a finite linear combination of projections in θ_K such that $\|S-L\| < \varepsilon$. Then

$$\eta(S) = |\eta(S)-\eta(L)| \leqq \eta(S-L) \leqq \|S-L\| < \varepsilon.$$

Thus, $\eta(S)=0$. Then $\eta(T)-\eta(S) \leqq \eta(T+S) \leqq \eta(T)+\eta(S)=\eta(T)$.

(3.4) $$\eta(T+X) = \eta(T) \quad (T \in A,\ X \in Z).$$

To prove (3.4), assume that $\alpha \in P_A$, $T \in A$, and $X \in Z$. Form the irreducible representation $(\Phi_\alpha, \mathfrak{H}_\alpha)$, and choose $\xi \in \mathfrak{H}_\alpha$ $\|\xi\|=1$, such that

$$\alpha(S) = \langle \Phi_\alpha(S)\xi, \xi \rangle \quad (S \in A).$$

Then $\Phi_\alpha(X)$ is the scalar $\alpha(X)$ times the identity operator on H_α. Therefore

$$\alpha(TX) = \langle \Phi_\alpha(T)\Phi_\alpha(X)\xi, \xi \rangle = \alpha(X)\langle \Phi_\alpha(T)\xi, \xi \rangle = \alpha(X)\alpha(T).$$

Thus,

$$\gamma(\alpha, T+X)^2 = \alpha\big((T^*+X^*)(T+X)\big) - |\alpha(T+X)|^2 =$$

$$= \alpha(T^*T) + \overline{\alpha(T)}\alpha(X) + \alpha(T)\overline{\alpha(X)} + |\alpha(X)|^2 - \big(\overline{\alpha(T)+\alpha(X)}\big)\big(\alpha(T)+\alpha(X)\big) =$$

$$= \alpha(T^*T) - |\alpha(T)|^2 = \gamma(\alpha, T)^2.$$

Therefore $\eta(T+X)=\eta(T)$.

§ 4. The distance from the thin operators

Throughout this §, A is a von Neumann algebra and K is a closed ideal of A with the property that if $T \in A$ and $TK=\{0\}$, then $T=0$. When A is represented spatially, this property of K is equivalent to the property that K is weak operator dense in A. In this § we prove the following theorem.

Theorem 4.1. *Let A and K be as above. Then*

$$\eta_K(T) = \lim_{P \in \theta_K} \sup \|TP-PT\| = d(T, \mathfrak{J}_K).$$

The first equality in this statement generalizes a result of R. Douglas and C. Pearcy in [6], and the second equality is a conjecture of C. Olsen [8. p. 572].

We prove Theorem 4.1 in several steps. The first of these, the next proposition, is a direct generalization of [6, Theorem 1].

Proposition 4.2. $\eta(T) = \lim_{P \in \theta_K} \sup \|PT(I-P)\|$.

Proof. Let μ equal the lim sup on the right hand side of the equality above. Fix $P \in \theta_K$. Then

$$(I-P)T^*PT(I-P) \in K_{I-P}.$$

There exists $\beta \in P_A$ such that $\beta(I-P)=1$, and

(1) $$\beta(T^*PT) = \beta((I-P)T^*PT(I-P)) = \|PT(I-P)\|^2.$$

Note that if $PT(I-P) \neq 0$, then $\beta(K) \neq \{0\}$. Also,

(2) $$\beta(T^*(I-P)T) - |\beta(T)|^2 = \beta(T^*(I-P)T) - |\beta((I-P)T)|^2 \geqq 0.$$

Adding (1) and (2) we have

$$\gamma(\beta, T)^2 = \beta(T^*T) - |\beta(T)|^2 \geqq \|PT(I-P)\|^2.$$

Therefore,

$$\sup\{\gamma(\alpha, T) : \alpha \in \Delta(P)\} \geqq \|PT(I-P)\|.$$

Taking the lim sup over $P \in \theta_K$ on both sides of this inequality, it follows that $\eta(T) \geqq \mu$.

Conversely, let $\delta > 0$ be arbitrary. Fix $P \in \theta_K$. We proceed to find $Q \in \theta_K$ such that $Q \geqq P$ and

$$\|QT(I-Q)\| \geqq \eta(T) - \delta.$$

Then this suffices to prove the inequality $\mu \geqq \eta(T)$.

Assume $\alpha \in \Delta(P)$ is such that

$$\gamma(\alpha, T_{I-P}) > \eta(T_{I-P}) - \delta.$$

Denote by α_0 the restriction of α to A_{I-P}. Then α_0 is a pure state of A_{I-P}. Form the irreducible representation $(\Phi_{\alpha_0}, \mathfrak{H}_{\alpha_0})$ of A_{I-P}. Choose $z \in \mathfrak{H}_{\alpha_0}$, $\|z\|=1$, such that

$$\alpha_0(S) = \langle \Phi_{\alpha_0}(S)z, z\rangle \quad (S \in A_{I-P}).$$

Let $w = \Phi_{\alpha_0}(T_{I-P})z - \alpha_0(T_{I-P})z$. Then

$$\|w\|^2 = \alpha_0((I-P)T^*(I-P)T(I-P)) - |\alpha_0(T_{I-P})|^2 = \gamma(\alpha, T_{I-P})^2.$$

Observe that $w \perp z$ in $\mathfrak{H}_{\alpha_0}$. Then by Kadison's Transitivity Theorem [5, Théorème 2.8.3] there exists a selfadjoint operator $S \in K_{I-P}$ such that $\Phi_{\alpha_0}(S)z=0$ and $\Phi_{\alpha_0}(S)w = w$. Then $\Phi_{\alpha_0}(S^2)z=0$ and $\Phi_{\alpha_0}(S^2)w=w$. Using the spectral resolution of the identity for S^2, it is not difficult to show that there exists a sequence of projections $\{R_n\} \subset K_{I-P}$ such that

$$\Phi_{\alpha_0}(R_n)z = 0 \quad \text{and} \quad \Phi_{\alpha_0}(R_n)w \to w.$$

Then

$$\alpha_0((I-P)T^*R_nT(I-P)) = \langle \Phi_{\alpha_0}((I-P)T^*R_nT(I-P))z, z\rangle$$
$$= \|\Phi_{\alpha_0}(R_n)\Phi_{\alpha_0}(T_{I-P})z\|^2 = \|\Phi_{\alpha_0}(R_n)(\Phi_{\alpha_0}(T_{I-P})z - \alpha_0(T_{I-P})z)\|^2$$
$$= \|\Phi_{\alpha_0}(R_n)w\|^2 \to \|w\|^2.$$

Therefore

$$\alpha_0((I-P)T^*R_nT(I-P)) \to \|w\|^2, \quad \|w\|^2 = \gamma(\alpha, T_{I-P})^2 > (\eta(T_{I-P}) - \delta)^2.$$

Set $R = R_m$ for some m so large that

$$\alpha_0((I-P)T^*R_mT(I-P)) > (\eta(T_{I-P}) - \delta)^2.$$

Now we have

$$\alpha(T^*RT) = \alpha((T^*RT)_{I-P}) = \alpha_0((I-P)T^*RT(I-P)).$$

Also, by (3.3), $\eta(T_{I-P})=\eta(T)$. Thus $PR=RP=0$, and $\alpha(T^*RT)>(\eta(T)-\delta)^2$. Let $Q=P+R$. Then $Q\geqq P$ and $\alpha(Q)=0$. Finally

$$\|QT(I-Q)\|^2 \geqq \alpha((I-Q)T^*QT(I-Q)) = \alpha(T^*QT) \geqq \alpha(T^*RT) > (\eta(T)-\delta)^2.$$

This completes the proof of the proposition.

If $T\in A$, $X\in Z$, and $J\in K$, then by (3.3) and (3.4) we have $\eta(T)=\eta(T+X+J)$. It follows using (3.1) that $\eta(T)\leqq\|T+X+J\|$. Therefore $\eta(T)\leqq d(T, \mathfrak{I}_K)$.

We state this result as a lemma.

Lemma 4.3. $\eta_K(T)\leqq d(T, \mathfrak{I}_K)$.

Our aim now is to prove the reverse of the inequality appearing in Lemma 4.3. First we need a technical result. Let Γ be the set of all primitive ideals B of A such that $K\not\subset B$. For $B\in\Gamma$, let π_B be the natural quotient map of A onto A/B. We show that

(4.4) $$\|S\| = \sup_{B\in\Gamma} \|\pi_B(S)\| \quad (S\in A).$$

Let Φ be the map from A into the C^*-direct product of the C^*-algebras A/B, $B\in\Gamma$, given by

$$\Phi(S) = (\pi_B(S))_{B\in\Gamma}.$$

Since $\bigcap_{B\in\Gamma}(B\cap K)=\{0\}$, Φ is an isomorphism on K. If $S\in A$ and $S\neq 0$, then there exists $J\in K$ such that $SJ\neq 0$. Then $\Phi(SJ)\neq 0$, so $\Phi(S)\neq 0$. Thus Φ, is a *-isomorphism of A, and therefore, an isometry. This proves (4.4).

Lemma 4.5. $\eta_K(T)\geqq d(T, \mathfrak{I}_K)$.

Proof. Let Δ be the set of all $\alpha\in P_A$ such that $\alpha(K)\neq\{0\}$. Assume $T\in A$. We prove

(1) $$\sup_{\alpha\in\Delta} \gamma(\alpha, T) \geqq d(T, Z).$$

Assume $\alpha\in\Delta$, and let $(\Phi_\alpha, \mathfrak{H}_\alpha)$ be the irreducible representation of A determined by α. If $\xi\in\mathfrak{H}_\alpha$, $\|\xi\|=1$, let

$$\omega_\xi(S) = \langle\Phi_\alpha(S)\xi, \xi\rangle \quad (S\in A).$$

By definition [9, p. 216], ω_ξ is representable by $(\Phi_\alpha, \mathfrak{H}_\alpha)$. Then by [9, Lemma (4.5.8)] the *-representation of A associated with ω_ξ is unitarily equivalent to $(\Phi_\alpha, \mathfrak{H}_\alpha)$. Thus, ω_ξ is a pure state of A [9, Theorem (4.6.4)]. Since $\Phi_\alpha(K)$ acts irreducibly on $\mathfrak{H}_\alpha$, we have $\omega_\xi\in\Delta$. Let D_T and $D_{\alpha,T}$ be the inner derivations determined by T on A, and by $\Phi_\alpha(T)$ on $B(\mathfrak{H}_\alpha)$, respectively. Observe that

$$\gamma(\omega_\xi, T) = \|\Phi_\alpha(T)\xi-\langle\Phi_\alpha(T)\xi, \xi\rangle\xi\|.$$

Then by [2, Corollary 1.3]

$$\sup_{\xi \in H_\alpha, \|\xi\|=1} \gamma(\omega_\xi, T) = \frac{1}{2} \|D_{\alpha, T}\|. \tag{2}$$

Let $\varepsilon > 0$ be arbitrary. Choose $S \in A$, $\|S\| = 1$, such that

$$\|TS - ST\| \geqq \|D_T\| - \varepsilon.$$

Then by (2)

$$\sup_{\xi \in \mathfrak{H}_\alpha, \|\xi\|=1} \gamma(\omega_\xi, T) \geqq \frac{1}{2} \|\Phi_\alpha(TS - ST)\|. \tag{3}$$

Let B_α be the primitive ideal that is the kernel of Φ_α, and let π_α be the natural quotient map of A onto A/B_α. If $R \in A$, then $\|\Phi_\alpha(R)\| = \|\pi_\alpha(R)\|$. Therefore by (4.4)

$$\|R\| = \sup_{\alpha \in \Delta} \|\pi_\alpha(R)\| = \sup_{\alpha \in \Delta} \|\Phi_\alpha(R)\|.$$

Applying this equality to (3), we have

$$\sup_{\alpha \in \Delta} \gamma(\alpha, T) \geqq \frac{1}{2} \sup_{\alpha \in \Delta} \|\Phi_\alpha(TS - ST)\| = \frac{1}{2} \|TS - ST\| \geqq \frac{1}{2} (\|D_T\| - \varepsilon).$$

This proves that

$$\sup_{\alpha \in \Delta} \gamma(\alpha, T) \geqq \frac{1}{2} \|D_T\|.$$

Then by [12, Corollary, p. 148]

$$\sup_{\alpha \in \Delta} \gamma(\alpha, T) \geqq d(T, Z).$$

This completes the proof of (1).

Now fix $P \in \theta_K$. The center of A_{I-P} is Z_{I-P}. Applying (1) to the algebra A_{I-P} and the element $(I-P)T(I-P)$, we have

$$\sup_{\alpha \in \Delta(P)} \gamma(\alpha, T) \geqq d((I-P)T(I-P), Z_{I-P}).$$

Also,

$$d((I-P)T(I-P), Z_{I-P}) = \inf_{X \in Z} \|(I-P)T(I-P) + (I-P)X(I-P)\|$$

$$\geqq d(T, \mathfrak{J}_K).$$

Therefore, $\eta_K(T) \geqq d(T, \mathfrak{J}_K)$.

By [8, Theorem 2]

$$\lim_{P \in \theta_K} \sup \|PT(I-P)\| = \lim_{P \in \theta_K} \sup \|TP - PT\|.$$

This equality in conjunction with Proposition 4.2, Lemma 4.3, and Lemma 4.5, proves Theorem 4.1.

Corollary 4.6. *Let A and K be as before. Then the following are equivalent for $T\in A$:*

$$\lim_{P\in\theta_K}\|TP-PT\|=0,\quad \eta_K(T)=0,\quad \textit{and}\quad T\in\mathfrak{I}_K.$$

Acknowledgement. The author acknowledges with thanks the many constructive suggestions made by the referee. These suggestions resulted in significant improvements in this paper.

References

[1] C. Apostol, C. Foiaş, and L. Zsidó, Some results on nonquasitriangular operators, *Indiana Univ. Math. J.*, **22** (1973), 1151—1161.
[2] C. Apostol and L. Zsidó, Ideals in W^*-algebras and the function η of A. Brown and C. Pearcy, *Rev. Roum. Math. Pures et Appl.*, **18** (1973), 1151—1170.
[3] B. Barnes, Closed one-sided ideals in certain B^*-algebras, *Illinois J. Math.*, **15** (1971), 503—514.
[4] A. Brown and C. Pearcy, Structure of commutators of operators, *Annals of Math.*, **82** (1965), 112—127.
[5] J. Dixmier, *Les C^*-algèbres et leurs représentations*, Gauthier-Villars (Paris, 1964).
[6] R. Douglas and C. Pearcy, A characterization of thin operators, *Acta Sci. Math.*, **29** (1968), 295—297.
[7] C. Olsen, Thin operators in a von Neumann algebra, *Acta Sci. Math.*, **35** (1973), 211—216, Correction, **36** (1974), 377—378.
[8] C. Olsen, A characterization of thin operators in a von Neumann algebra, *Proc. Amer. Math. Soc.*, **39** (1973), 571—577.
[9] C. Rickart, *General theory of Banach algebras*, D. Van Nostrand, Inc. (Princeton, N. J., 1960).
[10] N. Salinas, On the η function of Brown and Pearcy and the numerical function of an operator, *Canad. J. Math.*, **23** (1971), 565—578.
[11] F. Wright, A reduction for algebras of finite type, *Annals of Math.* (2) **60** (1954), 560—570.
[12] L. Zsidó, The norm of a derivation in a W^*-algebra, *Proc. Amer. Math. Soc.*, **38** (1973), 147—150.

THE UNIVERSITY OF OREGON,
EUGENE, OREGON 97403

(Received July 29, 1975, revised March 20, 1976)

Jordan model for some operators

HARI BERCOVICI

The aim of this Note is to find the Jordan model of a C_0 operator whose characteristic function coincides with $e_A(z)=\exp\left(A\frac{z+1}{z-1}\right)$, where A is a bounded positive operator acting on a separable Hilbert space $\mathfrak{K}$. This problem was proposed by C. Foiaş for $\mathfrak{K}=L^2(0, 1)$ and the operator A defined by $(Af)(x)=xf(x)$, $f\in L^2(0, 1)$.

1. Preliminaries

We will frequently use the following assertion. If T, T' are two quasisimilar completely non-unitary contractions, $m\in H^\infty$, $\mathfrak{N}=(\operatorname{ran} m(T))^-$ and $\mathfrak{N}'=(\operatorname{ran} m(T'))^-$, then $T|\mathfrak{N}$ and $T'|\mathfrak{N}'$ are also quasisimilar (cf. [2]).

Let us recall that if the operator T is acting on $\mathfrak{H}$, its multiplicity μ_T is defined as the minimum cardinality of a subset $\mathfrak{M}\subset\mathfrak{H}$ such that $\bigvee_{n=0}^{\infty} T^n\mathfrak{M}=\mathfrak{H}$. If T and T' are quasisimilar, then $\mu_T=\mu_{T'}$ (cf. [3]).

Proposition A. (cf. [4], [5], [1]) *Let T be a C_0 operator acting on a separable Hilbert space. Then there exists a sequence $\{m_j\}_{j=1}^n$ of inner functions such that:*

(1) *m_{j+1} divides m_j for each j;*

(2) *T is quasisimilar to $\bigoplus_{j=1}^{n} S(m_j)$;*

(3) $m_1=m_T$;

(4) $n=\mu_T$ $(\leqq\infty)$.

The sequence $\{m_j\}_{j=1}^n$ is uniquely determined by conditions (1) *and* (2).

The operator $\bigoplus_{j=1}^{n} S(m_j)$ is called the Jordan model of T. An operator of the form $\bigoplus_{j=1}^{n} S(m_j)$, for which (1) holds, is called a Jordan operator.

Received December 15, 1975.

Let us recall that with each inner function $\{\mathfrak{K}, \mathfrak{K}, \Theta(z)\}$ in the unit disc we can associate the operator $S(\Theta)$ acting on the space

$$\mathfrak{H}(\Theta) = H^2(\mathfrak{K}) \ominus \Theta H^2(\mathfrak{K}), \tag{1.1}$$

defined by

$$S(\Theta)u = P_{\mathfrak{H}(\Theta)}(zu(z)), \quad u \in \mathfrak{H}(\Theta). \tag{1.2}$$

If the function $\{\mathfrak{K}, \mathfrak{K}, \Theta(z)\}$ is pure, then it coincides with the characteristic function of the contraction $S(\Theta)$ (cf. [2]).

If is obvious that if I is an at most countable set and for each $i \in I$, $\{\mathfrak{K}_i, \mathfrak{K}_i, \Theta_i(z)\}$ is an inner function in the unit disc, then the function $\{\mathfrak{K}, \mathfrak{K}, \Theta(z)\}$, where $\mathfrak{K} = \bigoplus_{i \in I} \mathfrak{K}_i$ and $\Theta(z) = \bigoplus_{i \in I} \Theta_i(z)$, is also inner and we have

$$S(\Theta) = \bigoplus_{i \in I} S(\Theta_i). \tag{1.3}$$

2. The Jordan model of $S(e_A)$

Let A be a positive operator on the separable Hilbert space $\mathfrak{K}$, with spectral measure E. We can then define an inner function $\{\mathfrak{K}, \mathfrak{K}, e_A(z)\}$ by the formula:

$$e_A(z) = \exp\left(A \frac{z+1}{z-1}\right) = \int_0^a e_t(z)\, dE_t, \quad a = \|A\|, \tag{2.1}$$

where we use the notation:

$$e_t(z) = \exp\left(t \frac{z+1}{z-1}\right). \tag{2.2}$$

As $e_A(0) = \exp(-A)$, it is easy to see that the function e_A is pure if and only if $\ker A = \{0\}$.

Lemma 1. *The characteristic function of*

$$S(e_A) | (\operatorname{ran} e_t(S(e_A)))^-, \quad t \geqq 0,$$

is $\{\mathfrak{K}_t, \mathfrak{K}_t, e_{A_t}(z)\}$, *where* $\mathfrak{K}_t = E((t, \|A\|])\mathfrak{K}$ *and* $A_t = (A - tI)|\mathfrak{K}_t$. *Thus* $S(e_A)$ *is a* C_0 *operator and its minimal function is* $e_{\|A\|}$.

Proof. We first show that

$$(\operatorname{ran} e_t(S(e_A)))^- = e_{A_t'} H^2(\mathfrak{K}) \ominus e_A H^2(\mathfrak{K}) \tag{2.3}$$

where

$$A_t' = AE((0, t]) + tE((t, \|A\|]). \tag{2.4}$$

Indeed we have

$$(2.5)\qquad \begin{aligned}(\operatorname{ran} e_t(S(e_A)))^- &= (P_{\mathfrak{H}(e_A)} e_t \mathfrak{H}(e_A))^- = (P_{\mathfrak{H}(e_A)} e_t H^2(\mathfrak{K}))^- = \\ &= (e_t H^2(\mathfrak{K}) + e_A H^2(\mathfrak{K}))^- \ominus e_A H^2(\mathfrak{K}).\end{aligned}$$

The operator of multiplication by e_t on $H^2(\mathfrak{K})$ may be represented as a product $e_{A_t'} e_{A_t''}$, where $A_t'' = (tI - A)E((0, t])$, thus $e_t H^2(\mathfrak{K}) \subset e_{A_t'} H^2(\mathfrak{K})$ and from (2.5) we infer

$$(2.6)\qquad (\operatorname{ran} e_t(S(e_A)))^- \subset e_{A_t'} H^2(\mathfrak{K}) \ominus e_A H^2(\mathfrak{K}).$$

Now, for $u \in H^2(\mathfrak{K})$ we have

$$e_{A_t'} u = e_A E((0, t]) u + e_t E((t, \|A\|]) u,$$

thus $e_{A_t'} H^2(\mathfrak{K}) \subset e_A H^2(\mathfrak{K}) + e_t H^2(\mathfrak{K})$ and from (2.5) we infer

$$e_{A_t'} H^2(\mathfrak{K}) \ominus e_A H^2(\mathfrak{K}) \subset (\operatorname{ran} e_t(S(e_A)))^-.$$

This inclusion and (2.6) prove the equality (2.3).

Now let us remark that the operator $R\colon \mathfrak{H}(e_{A_t}) \to \mathfrak{H}(e_A)$ defined by $Ru = e_t u$ is isometric,

$$R\mathfrak{H}(e_{A_t}) = e_t H^2(\mathfrak{K}_t) \ominus e_A H^2(\mathfrak{K}_t) = e_{A_t'} H^2(\mathfrak{K}) \ominus e_A H^2(\mathfrak{K}) = (\operatorname{ran} e_t(S(e_A)))^-$$

and $RS(e_{A_t}) = S(e_A)R$. Thus $S(e_A)|(\operatorname{ran} e_t(S(e_A)))^-$ is unitarily equivalent so $S(e_{A_t})$ and the lemma follows if we remark that $\ker A_t = \{0\}$, that is e_{A_t} is pure.

Lemma 2. *We have* $\mu_{S(e_A)} = \operatorname{Rank} A$.

Proof. We may suppose without loss of generality that $\ker A = \{0\}$. If $\operatorname{Rank} A = n < \infty$, A is represented, for an adequate choice of the basis in $\mathfrak{K}$, by the matrix

$$\begin{pmatrix} t_1 & 0 & \dots & 0 \\ 0 & t_2 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & t_n \end{pmatrix}, \quad t_1 \geqq t_2 \geqq \dots \geqq t_n > 0.$$

It follows that $S(e_A)$ is unitarily equivalent to the Jordan operator $\bigoplus_{j=1}^{n} S(e_{t_j})$; thus $S(e_A)$ is of multiplicity n.

Conversely, let us suppose that $S(e_A)$ is of multiplicity $n < \infty$. We show first that the spectrum $\sigma(A)$ consists of at most n points. If $\sigma(A)$ contains more than n points, we can find $0 = t_0 < t_1 < \dots < t_{n+1} = \|A\|$ such that $E((t_i, t_{i+1}]) \neq 0$, $i = 0, 1, \dots, n$. Because $A = \bigoplus_{i=0}^{n} A|E((t_i, t_{i+1}])\mathfrak{K} = \bigoplus_{i=0}^{n} A_i$, we have $S(e_A) = \bigoplus_{i=0}^{n} S(e_{A_i})$. From Lemma 1 and Proposition A it follows that $S(e_{A_i})$ is quasisimilar to a Jordan operator

$$S(e_{s_i}) \oplus \dots, \quad \text{where} \quad s_i = \|A_i\| \in (t_i, t_{i+1}].$$

Thus $S(e_A)$ is quasisimilar to

$$T = S(e_{s_n}) \oplus S(e_{s_{n-1}}) \oplus \dots \oplus S(e_{s_0}) \oplus \dots,$$

$s_n > s_{n-1} > \ldots > s_0 > 0$ (T may not be a Jordan operator). It is clear that $\mu_T \geqq n+1$ and this contradicts the equality $\mu_T = \mu_{S(e_A)} = n$. Thus $\sigma(A)$ consits of at most n points, say

$$\sigma(A) = \{\tau_1, \tau_2, \ldots, \tau_k\}, \quad \tau_1 > \tau_2 > \ldots > \tau_k > 0 \quad (k \leqq n).$$

Each τ_i is an eigenvalue of A say of multiplicity $n_i (\leqq \infty)$. Because $A = \bigoplus_{i=1}^{k} A|E(\{\tau_i\})\mathfrak{R}$, it follows that $S(e_A)$ is unitarily equivalent to

$$\bigoplus_{i=1}^{k} \left(\bigoplus_{j=1}^{n_i} S(e_{\tau_i}) \right). \tag{2.7}$$

Now, the operator (2.7) is of finite multiplicity if and only if $n_i < \infty$, $i = 1, \ldots, k$, and then its multiplicity equals $n_1 + n_2 + \ldots + n_k = \text{Rank } A$. The lemma follows.

Lemma 3. *Let $S = \bigoplus_{j=1}^{\infty} S(m_j)$ be a Jordan operator of infinite multiplicity and let T be a C_0 operator acting on a separable Hilbert space with the property that m_T divides m_j for each j. Then the Jordan model of $T \oplus S$ is S.*

Proof. Let $S' = \bigoplus_{j=1}^{\infty} S(m'_j)$ be the Jordan model of $T \oplus S$. For each j, $(T \oplus S)|(\text{ran}\, m'_j(T \oplus S))^-$ is quasisimilar to $S'|(\text{ran}\, m'_j(S'))^-$, thus it has finite multiplicity. It follows that, for sufficiently large i, $m'_j(S(m_i)) = 0$, thus m_i divides m'_j. From the hypothesis it follows that m_T divides m'_j for each j. Now, $(T \oplus S)|(\text{ran}\, m_T(T \oplus S))^-$ and $S'|(\text{ran}\, m_T(S'))^-$ are quasisimilar. Because $(T \oplus S)|(\text{ran}\, m_T(T \oplus S))^-$, $S'|(\text{ran}\, m_T(S'))^-$ are unitarily equivalent to $\bigoplus_{j=1}^{\infty} S(m_j/m_T)$, $\bigoplus_{j=1}^{\infty} S(m'_j/m_T)$ respectively, from the uniqueness assertion of Proposition A it follows that $m_j/m_T = m'_j/m_T$, $m_j = m'_j$ for each j.

The lemma is proved.

Let us put

$$t_0 = \inf\{t : \dim E((t, \|A\|])\mathfrak{R} < \infty\}. \tag{2.8}$$

Then $\sigma(A) \cap (t_0, \|A\|]$ contains only eigenvalues of finite multiplicity. Let $\{t_j\}_{j=1}^{n_A}$, $n_A = \dim E((t_0, \|A\|])\mathfrak{R} \leqq \infty$, $t_1 \geqq t_2 \geqq \ldots$, be these eigenvalues, each one being counted according its multiplicity. So we are able to state the main result of this paper:

Theorem. *The Jordan model of $S(e_A)$ is:*

(a) $\bigoplus_{j=1}^{\infty} S(e_{t_j})$ *if* $n_A = \dim E((t_0, \|A\|])\mathfrak{R} = \infty$;

(b) $\left(\bigoplus_{j=1}^{n_A} S(e_{t_j}) \right) \oplus \left(\bigoplus_{i=1}^{\infty} S(e_{t_0}) \right)$ *if* $n_A < \infty$.

Proof. We have the relation $A=A'\oplus\left(\bigoplus_{j=1}^{n_A} t_j\right)$ (here t_j is considered as a multiplication operator on a 1-dimensional Hilbert space), thus $S(e_A)=S(e_{A'})\oplus \oplus\left(\bigoplus_{j=1}^{n_A} S(e_{t_j})\right)$. If $n_A=\infty$, the conditions of Lemma 3 are satisfied for $T=S(e_{A'})$ and $S=\bigoplus_{j=1}^{\infty} S(e_{t_j})$, thus (a) follows.

Let us suppose that $n_A<\infty$. Then, if E' denotes the spectral measure of A', we have $\dim \operatorname{ran} E'((t, t_0])=\infty$ for each $t<t_0=\|A'\|$. From Lemmas 1 and 2 it follows that for each $t<t_0=\|A'\|$ the operator $S(e_{A'})|(\operatorname{ran} e_t(S(e_{A'})))^-$ is of infinite multiplicity. Let $S=S(e_{t_0})\oplus(\bigoplus_{j=1}^{\infty} S(e_{t^j}))$, $t_0\geqq t^1\geqq t^2\geqq\ldots$, be the Jordan model of $S(e_{A'})$. If $t^j=t<t_0$ for some j, it follows that $S|(\operatorname{ran} e_t(S))^-$ is of finite multiplicity, thus $S(e_{A'})|(\operatorname{ran} e_t(S(e_{A'})))^-$ is of finite multiplicity, a contradiction. It follows that $t^j=t_0$ for each j, thus $S(e_A)$ is quasisimilar to

$$\left(\bigoplus_{j=1}^{n_A} S(e_{t_j})\right)\oplus\left(\bigoplus_{i=1}^{\infty} S(e_{t_0})\right).$$

The last operator is a Jordan operator and the theorem follows from the uniqueness assertion of Proposition A.

Remark. If A acts on a finite dimensional Hilbert space we have $n_A=\operatorname{Rank} A$, $t_0=0$, and the Jordan model has the form $\bigoplus_{j=1}^{n_A} S(e_{t_j})$. Thus our theorem is verified in this case also.

Example. Let A be defined by $(Af)(x)=x\cdot f(x)$ on $\mathfrak{K}=L^2(0, 1)$. Then $\|A\|=1$ and A has no eigenvalues. It follows that the Jordan model of $S(e_A)$ is $\bigoplus_{i=1}^{\infty} S(e_1)$.

References

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

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

[3] B. Sz.-Nagy, C. Foiaş, Vecteurs cycliques et quasiaffinités, *Studia Math.*, **31** (1968), 35—42.

[4] 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.

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

FACULTY OF MATHEMATICS
UNIVERSITY BUCHAREST
STR. ACADEMIEI 14
BUCHAREST, ROMANIA

Acta Sci. Math., 38 (1976), 281—290

On intertwining dilations

ZOIA CEAUŞESCU

Introduction. Let T, T' be two contractions on the Hilbert space $\mathfrak{H}$ and $\mathfrak{H}'$, and U, U' their isometric dilations on $\mathfrak{K}$ and $\mathfrak{K}'$, respectively. For an operator $A \in L(\mathfrak{H}', \mathfrak{H})$ (the space of all bounded operators from $\mathfrak{H}'$ into $\mathfrak{H}$) intertwining T and T' (i.e. $TA = AT'$) let us call an *intertwining dilation* of A any operator $B \in L(\mathfrak{K}'; \mathfrak{K})$ satisfying: $P_{\mathfrak{H}} B | \mathfrak{H}' = A$, $UB = BU'$ and $B(\mathfrak{K}' \ominus \mathfrak{H}') \subset \mathfrak{K} \ominus \mathfrak{H}$. If, moreover, B satisfies $\|B\| = \|A\|$ it will be called an *exact* intertwining dilation of A. It is known that for any operator A intertwining T and T' there exists at least one exact intertwining dilation (see Th. 2. 3 of [5]).

In the present paper we are concerned with the problem of uniqueness of such an exact intertwining dilation. We reduce this problem to the similar problem for the Hahn—Banach extensions of continuous functionals on some adequate quotient spaces of projective tensor products.[1])

Our main result is contained in Section 3. Thus we show that if an operator intertwining two contractions has a unique exact intertwining dilation, then all the operators which are "dominated" (in the sense of Definition 3.1) by it have the same property (see Th. 3.2). As an illustrative example, in the last section, an application of the above theorem to Hankel operators is given.

I take this opportunity to express my gratitude to Prof. C. Foiaş, for many helpful discussions. Also I thank Prof. B. Sz.-Nagy for his useful remarks on the first version of this paper.

1. Let $\mathfrak{K}$ and $\mathfrak{G}$ be two Hilbert spaces. We shall denote by $\mathfrak{K}^* \otimes \mathfrak{G}$ the subspace of $L(\mathfrak{K}; \mathfrak{G})$ consisting of operators τ which admit a representation of the form

$$(1) \qquad \tau = \sum_{j=1}^{n} k_j^* \otimes g_j, \quad \text{where} \quad k_j \in \mathfrak{K},\ g_j \in \mathfrak{G},\ 1 \leqq j \leqq n,$$

that is,

$$(2) \qquad \tau(k) = \sum_{h=1}^{n} (k, k_j) g_j \quad (k \in \mathfrak{K}).$$

Received December 18, 1975, revised March 5, 1976.

[1]) This reduction already was done in some more or less particular cases (see for instance [6]).

We shall use the notation $\|\cdot\|_\pi$ for the nuclear norm on $\mathfrak{K}^*\otimes\mathfrak{G}$:

(3) $$\|\tau\| = \inf\left\{\sum_{j=1}^{n}\|k_j\|\,\|g_j\| : \tau = \sum_{j=1}^{n} k_j^*\otimes g_j\right\}.$$

The space $\mathfrak{K}^*\otimes\mathfrak{G}$ endowed with this norm will be denoted by $\mathfrak{K}^*\underset{\pi}{\otimes}\mathfrak{G}$.

An immediate result is expressed by the following

Lemma 1.1. *For a subspace $\mathfrak{H}$ of $\mathfrak{K}$ the space $\mathfrak{H}^*\underset{\pi}{\otimes}\mathfrak{G}$ can be identified with the subspace $\mathfrak{L}$ of $\mathfrak{K}^*\underset{\pi}{\otimes}\mathfrak{G}$ consisting of those $\tau\in\mathfrak{K}^*\underset{\pi}{\otimes}\mathfrak{G}$ for which*

(4) $$\tau|\mathfrak{K}\ominus\mathfrak{H} = 0.$$

On account of Lemma 1.1 we may and will identify $\mathfrak{H}^*\underset{\pi}{\otimes}\mathfrak{G}$ with the subspace $\mathfrak{L}$ defined by (4), of $\mathfrak{K}^*\underset{\pi}{\otimes}\mathfrak{G}$. We shall denote by $\mathfrak{K}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G}$ and $\mathfrak{H}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G}$ the completions of $\mathfrak{K}^*\underset{\pi}{\otimes}\mathfrak{G}$ and $\mathfrak{H}^*\underset{\pi}{\otimes}\mathfrak{G}$, respectively.

Let us recall some well known properties (see [7]) of the completion of projective tensor product.

(i) Every element τ of $\mathfrak{K}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G}$ is the sum of an absolutely convergent series:

$$\tau = \sum_{n=0}^{\infty} k_n^*\otimes g_n, \quad \text{and} \quad \|\tau\|_\pi = \inf\left\{\sum_{n=0}^{\infty}\|k_n\|\,\|g_n\| : \tau = \sum_{n=0}^{\infty} k_n^*\otimes g_n\right\}.$$

(ii) The dual of $\mathfrak{K}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G}$ is realized as the space $L(\mathfrak{G};\mathfrak{K})$.

Also, we shall consider operators U on $\mathfrak{K}$, T on $\mathfrak{H}$, and Z on $\mathfrak{G}$, and assume that $\mathfrak{H}$ is a subspace of $\mathfrak{K}$ invariant for U^*, and $U^*|\mathfrak{H}=T^*$.

We denote by $[Z, U]$ the operator on $L(\mathfrak{K};\mathfrak{G})$, defined by

(5) $$[Z, U]V = ZV - VU \quad \text{for} \quad V\in L(\mathfrak{K};\mathfrak{G}).$$

Note that $\mathfrak{K}^*\otimes\mathfrak{G}$ and $\mathfrak{H}^*\otimes\mathfrak{G}$ are invariant for $[Z, U]$, and in virtue of the condition $T^*=U^*|\mathfrak{H}$ we have

$$[Z, U]|\mathfrak{H}^*\otimes\mathfrak{G} = [Z, T]|\mathfrak{H}^*\otimes\mathfrak{G}$$

(where $[Z, T]$ is defined on $L(\mathfrak{H};\mathfrak{G})$ in the same way as $[Z, U]$ is on $L(\mathfrak{K};\mathfrak{G})$). The operators $[Z, T]$ and $[Z, U]$ can be extended continuously to $\mathfrak{H}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G}$ and $\mathfrak{K}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G}$, respectively. Now, denote

(6) $$\mathfrak{N}_U = ([Z, U](\mathfrak{K}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G}))^-, \quad \mathfrak{N}_T = ([Z, T](\mathfrak{H}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G}))^-$$

where the closures are taken in the spaces $\mathfrak{K}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G}$ and $\mathfrak{H}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G}$, respectively. We shall consider the quotients modulo $\mathfrak{N}_U$ and $\mathfrak{N}_T$ of the nuclear norms on $\mathfrak{K}^*\underset{\pi}{\otimes}\mathfrak{G}$ and $\mathfrak{H}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G}$, respectively; thus, if ψ and φ denote the canonical epimorphism

$$\psi : \mathfrak{K}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G} \to (\mathfrak{K}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G})/\mathfrak{N}_U, \quad \varphi : \mathfrak{H}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G} \to (\mathfrak{H}^*\hat{\underset{\pi}{\otimes}}\mathfrak{G})/\mathfrak{N}_T$$

then

$$\|\psi(\tau)\| = \inf_{\tau_1 \in \mathfrak{R}_U} \|\tau+\tau_1\|_\pi (\tau \in \mathfrak{K}^* \hat{\otimes}_\pi \mathfrak{G}) \quad \text{and} \quad \|\varphi(\tau)\| = \inf_{\tau_1 \in \mathfrak{R}_T} \|\tau+\tau_1\|_\pi (\tau \in \mathfrak{H}^* \hat{\otimes}_\pi \mathfrak{G}).$$

Since, $\mathfrak{R}_U \supset \mathfrak{R}_T$, we infer that

$$\|\psi(\tau)\| \leqq \|\varphi(\tau)\| \quad \text{for} \quad \tau \in \mathfrak{H}^* \hat{\otimes}_\pi \mathfrak{G}. \tag{7}$$

Lemma 1.2. (i) *The dual of the Banach space* $(\mathfrak{K}^* \hat{\otimes}_\pi \mathfrak{G})/\mathfrak{R}_U$ *is isometric-isomorphic to the subspace*

$$\{B \in L(\mathfrak{G}; \mathfrak{K}) : UB = BZ\} \quad \textit{of} \quad L(\mathfrak{G}; \mathfrak{K}),$$

(ii) *The dual of the Banach space* $(\mathfrak{H}^* \hat{\otimes}_\pi \mathfrak{G})/\mathfrak{R}_T$ *is isometric-isomorphic to the subspace*

$$\{A \in L(\mathfrak{G}; \mathfrak{H}) : TA = AZ\} \quad \textit{of} \quad L(\mathfrak{G}; \mathfrak{H}).$$

Proof. (i): Firstly, let us observe that $\{B \in L(\mathfrak{G}; \mathfrak{K})$: $UB=BZ\}$ is isometric-isomorphic to $\mathfrak{R}_U^\perp$, where we denote by $\mathfrak{R}_U^\perp$ the orthogonal of $\mathfrak{R}_U$ i.e.

$$\mathfrak{K}_U^\perp = \{f \in (\mathfrak{K}^* \hat{\otimes}_\pi \mathfrak{G})' : f|\mathfrak{R}_U = 0\}.$$

Indeed, since $L(\mathfrak{G}; \mathfrak{K})$ is isometric-isomorphic to $(\mathfrak{K}^* \hat{\otimes}_\pi \mathfrak{G})'$, for any $B \in L(\mathfrak{G}; \mathfrak{K})$ with the property $UB=BZ$ there is a unique f from $(\mathfrak{K}^* \hat{\otimes}_\pi \mathfrak{G})'$ with the properties

(a) $f(k^* \otimes g) = (Bg, k)$ $(k \in \mathfrak{K},\ g \in \mathfrak{G})$ and (b) $\|f\| = \|B\|$.

But, for this f and for any $k \in \mathfrak{K}$, $g \in \mathfrak{G}$, we also have:

$$f([Z, U](k^* \otimes g)) = (BZg, k) - (UBg, k) = 0.$$

Since the set $\{[Z, U](k^* \otimes g) : k \in \mathfrak{K},\ g \in \mathfrak{G}\}$ spans $\mathfrak{R}_U$, it results readily $f|\mathfrak{R}_U = 0$.

Conversely, since $L(\mathfrak{G}; \mathfrak{K}) \cong (\mathfrak{K}^* \hat{\otimes}_\pi \mathfrak{G})'$, for any $f \in (\mathfrak{K}^* \hat{\otimes}_\pi \mathfrak{G})'$ with $f|\mathfrak{R}_U = 0$, there exists a unique $B \in L(\mathfrak{G}; \mathfrak{K})$ satisfying conditions (a), (b) above; moreover, we have

$$((UB-BZ)g, k) = f([Z, U](k^* \otimes g)) = 0 \quad \text{for any} \quad k \in \mathfrak{K},\ g \in \mathfrak{G}.$$

Thus, the operator B has also the property $UB=BZ$.

Now, statement (i) of the Lemma results from the following general fact: If $\mathfrak{X}$ is a Banach space and $\mathfrak{Y}$ is a subspace of $\mathfrak{X}$, then the orthogonal $\mathfrak{Y}^\perp$ of $\mathfrak{Y}$ is isometric-isomorphic to the dual of the quotient space $\mathfrak{X}/\mathfrak{Y}$.

(ii): The proof is analogous to that of (i), due to the similar definition for the space $\mathfrak{H}^* \hat{\otimes}_\pi \mathfrak{G}$, and thus for $(\mathfrak{H}^* \hat{\otimes}_\pi \mathfrak{G})/R_T$ too.

Lemma 1.3. *The following two statements are equivalent:*

(P_1) *For any $A\in L(\mathfrak{G},\mathfrak{H})$ satisfing the condition $TA=AZ$, there exists at least one exact intertwining dilation $B\in L(\mathfrak{G};\mathfrak{K})$ of A.*

(P_2) *For any $\tau\in\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}$, we have $\|\psi(\tau)\|=\|\varphi(\tau)\|$.*

Proof. First, we notice that, on account of Lemma 1.2, (P_1) is equivalent to:

(P_1') *For any $f\in((\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_T)'$ there exists an "extension" $\tilde{f}\in((\mathfrak{K}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_U)'$ of f (i.e. $\tilde{f}\psi(\tau)=f\varphi(\tau)$ for all $\tau\in\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}$) such that:*

$$\|\tilde{f}\|=\|f\| \quad (\text{or equivalently}, \quad \|\tilde{f}\psi\|=\|f\varphi\|).$$

Indeed, if (P_1) holds then, in virtue of Lemma 1.2, for $f\in((\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_T)'$ there is $\tilde{f}\in((\mathfrak{K}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_U)'$ such that $\|\tilde{f}\|=\|f\|$ and $\tilde{f}\psi(h^*\otimes g)=f\varphi(h^*\otimes g)$ for all $h\in\mathfrak{H}$ and $g\in\mathfrak{G}$. Since, for $\tau\in\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}$ there are the representations $\tau=\sum_{n\in N}h_n^*\otimes g_n$ where the series $\sum_{n\in N}h_n^*\otimes g_n$ is absolutely convergent, and since f, $\tilde{f}$, φ, ψ, are continuous, we also have

$$f\varphi(\tau)=\tilde{f}\psi(\tau) \quad \text{for all} \quad \tau\in\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}.$$

The converse implication $(P_1')\Rightarrow(P_1)$ is, by Lemma 1.2, even more obvious.

Now, we assume that (P_1') holds. Let us take $\tau_0\in\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}$ with $\varphi(\tau_0)\neq0$. There exists $f\in((\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_T)'$ with the properties:

$$\|f\|=\|f\varphi\|=1, \quad f\varphi(\tau_0)=\|\varphi(\tau_0)\|.$$

For this f there exists, according to (P_1'), $\tilde{f}\in((\mathfrak{K}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_U)'$ such that

$$\|\tilde{f}\|=\|f\|=1 \quad \text{and} \quad \tilde{f}\psi(\tau)=f\varphi(\tau) \quad (\tau\in\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}).$$

Thus, by (7),

$$\|\varphi(\tau_0)\|=\tilde{f}\psi(\tau_0)\leqq\|\tilde{f}\|\,\|\psi(\tau_0)\|=\|\psi(\tau_0)\|\leqq\|\varphi(\tau_0)\|.$$

If $\varphi(\tau_0)=0$ then, by (7), $0\leqq\|\psi(\tau_0)\|\leqq\|\varphi(\tau_0)\|=0$. Consequently, we obtain $\|\varphi(\tau)\|=\|\psi(\tau)\|$ for all $\tau\in\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}$.

Let us now assume that $\|\varphi(\tau)\|=\|\psi(\tau)\|$ for all $\tau\in\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}$. This means that the continuous canonical epimorphism

$$\varphi(\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})=(\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_T\to(\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_U=\psi(\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})$$

is an isometry. Therefore, we can identify $(\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_T$ with the subspace $(\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_U$

of $(\mathfrak{K}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_U$. Now, the implication $(P_2)\Rightarrow(P_1')$ follows from the Hahn—Banach Theorem.

It is known that if T is a contraction on $\mathfrak{H}$, U a minimal isometric dilation of T on $\mathfrak{K}$, and Z an isometry on $\mathfrak{G}$, then assertion (P_1) of Lemma 1.3 is true (cf. [5] Prop. II 2.2). Thus we have

Theorem 1.1. *Let T be a contraction on $\mathfrak{H}$, U a minimal isometric dilation of T, and Z an isometry on $\mathfrak{G}$. Then,*

$$(\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/([Z,T](\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}))^-$$

is linear canonically isometric to the image of $\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}$ in

$$(\mathfrak{K}^*\hat{\otimes}_\pi\mathfrak{G})/([Z,U](\mathfrak{K}^*\hat{\otimes}_\pi\mathfrak{G}))^-.$$

2. In the sequel we shall only treat the case considered in Theorem 1.1; that is, T is a contraction on $\mathfrak{H}$, U is a minimal isometric dilation of T on $\mathfrak{K}$, and Z is an isometry on $\mathfrak{G}$.

Remark 2.1. Let $A\in L(\mathfrak{G};\mathfrak{H})$ satisfy $TA=AZ$. In order that A should have a unique intertwining dilation $B\in L(\mathfrak{G};\mathfrak{K})$ with $\|B\|=\|A\|$ it is necessary and sufficient that the functional $f\in((\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_U)'$ (where $(\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_U$ is identified with $(\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_T$, in virtue of Theorem 1.1), corresponding to A by: $f\psi(h^*\otimes g)=(Ag,h)$, have a unique norm-preserving extension to the space $(\mathfrak{K}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_U$. On the other hand, a well-known consequence of the classical proof of the Hahn—Banach Theorem is that a functional $f\in((\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_U)'$ of norm 1 has a unique norm-preserving extension to $(\mathfrak{K}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_U$ if an only if for any $\tau\notin\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}$,

$$\sup\{\operatorname{Re}f(\dot{\tau}_1)-\|\dot{\tau}_1-\dot{\tau}\|:\dot{\tau}_1\in(\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_U\}=$$

$$=\inf\{\|\dot{\tau}_2+\dot{\tau}\|-\operatorname{Re}f(\dot{\tau}_2):\dot{\tau}_2\in(\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})/\mathfrak{N}_U\}.$$

(Here, as in the sequel, we set $\dot{\tau}=\psi(\tau)$ for $\tau\in\mathfrak{K}^*\hat{\otimes}_\pi\mathfrak{G}$). Hence, we easily infer the following sufficient and necessary condition for that an $A\in L(\mathfrak{G};\mathfrak{H})$, $\|A\|=1$, satisfying $TA=AZ$ have a unique exact intertwining dilation.

For any $\varepsilon>0$ and $\tau_0\in(\mathfrak{K}^*\hat{\otimes}_\pi\mathfrak{G})\setminus(\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})$ there exists $\tau_1,\tau_2\in\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}$ satisfying

$$\|\dot{\tau}_1+\dot{\tau}_2\|\leqq\|\dot{\tau}_1-\dot{\tau}_0\|+\|\dot{\tau}_2+\dot{\tau}_0\|<\operatorname{Re}f(\dot{\tau}_1+\dot{\tau}_2)+\varepsilon. \tag{8}$$

3. We introduce the following definition for contractions on Hilbert spaces:

Definition 3.1. Let $A_1, A_2 \in L(\mathfrak{H}_1; \mathfrak{H}_2)$ be two contractions. We say that A_1 *Harnack-dominates* A_2 if for some positive constants C, C' we have:

$$\|D_{A_2}h\| \leqq C\|D_{A_1}h\| \quad \text{and} \quad \|(A_2-A_1)h\| \leqq C'\|D_{A_1}h\| \tag{9}$$

for all $h \in \mathfrak{H}_1$. Here D_{A_1}, D_{A_2} are the defect operators of A_1, A_2, i.e. $D_{A_i}=(1-A_i^*A_i)^{1/2}$ (i=1, 2).

Remark 3.1. Let us introduce, for the contractions $A_1, A_2 \in L(\mathfrak{H}_1, \mathfrak{H}_2)$, the following isometries:

$$\hat{A}_i = \begin{pmatrix} A_i \\ D_{A_i} \end{pmatrix} : \mathfrak{H}_1 \to \begin{matrix} \mathfrak{H}_2 \\ \oplus \\ \mathfrak{D}_{A_i} \end{matrix} \quad (i = 1, 2),$$

where $\mathfrak{D}_{A_i}=\overline{D_{A_i}\mathfrak{H}_1}$ $(i=1, 2)$. Then, conditions (9) of Definition 3.1 are plainly equivalent to the following: There exists a bounded operator

$$K : \begin{matrix} \mathfrak{H}_2 \\ \oplus \\ \mathfrak{D}_{A_1} \end{matrix} \to \begin{matrix} \mathfrak{H}_2 \\ \oplus \\ \mathfrak{D}_{A_2} \end{matrix}$$

such that

$$K\begin{pmatrix} h_2 \\ 0 \end{pmatrix} = \begin{pmatrix} h_2 \\ 0 \end{pmatrix} \quad \text{for all} \quad h_2 \in \mathfrak{H}_2, \quad \text{and} \quad \hat{A}_2 = K\hat{A}_1. \tag{10}$$

Remark 3.2. We note that, if $\mathfrak{H}_1$ and $\mathfrak{H}_2$ coincide, then the equivalence relation for contractions on $\mathfrak{H}$, defined by: A_1 Harnack-dominates A_2, and A_2 Harnack-dominates A_1 coincides with the Harnack-equivalence as defined in [4], p. 362.

For two operators $A_1, A_2 \in L(\mathfrak{G}; \mathfrak{H})$, intertwining T and Z, denote by f_{A_1}, f_{A_2} the functionals $\in((\mathfrak{H}^* \hat{\otimes}_\pi \mathfrak{G})/\mathfrak{N}_U)'$, corresponding to A_1 and A_2, respectively, and by F_{A_1}, F_{A_2} the functionals $\in(\mathfrak{H}^* \hat{\otimes}_\pi \mathfrak{G})'$, satisfying $F_{A_1}|\mathfrak{N}_U=F_{A_2}|\mathfrak{N}_U=0$, which correspond to f_{A_1}, f_{A_2} by virtue of the isometric-isomorphism

$$((\mathfrak{H}^* \hat{\otimes}_\pi \mathfrak{G})/\mathfrak{N}_U)' \cong \mathfrak{N}_U^\perp.$$

Lemma 3.1. *Let* $A_1, A_2 \in L(\mathfrak{G}; \mathfrak{H})$ *be two operators intertwining* T *and* Z, $\|A_1\|=\|A_2\|=1$, *and such that* A_1 *Harnack-dominates* A_2. *Then,*

$$\|\tau\|_\pi - \operatorname{Re} F_{A_1}(\tau) \leqq \varepsilon \quad (\textit{for some } \varepsilon > 0 \textit{ and } \tau \in \mathfrak{H}^* \hat{\otimes}_\pi \mathfrak{G})$$

implies

$$\operatorname{Re} F_{A_1}(\tau) \leqq \operatorname{Re} F_{A_2}(\tau) + 2\varepsilon(\|K\|^2 - 1).$$

(K is the bounded operator satisfying (10), which exists by Remark 3.1.)

Proof. Let $\tau \in \mathfrak{H}^* \hat{\otimes}_\pi \mathfrak{G}$ be such that $\|\tau\|_\pi - \mathrm{Re}\, F_{A_1}(\tau) \leqq \varepsilon$ for some $\varepsilon > 0$. There exists a representation of τ, say

$$\tau = \sum_{n \in N} h_n^* \otimes g_n,$$

with

$$\|g_n\| = 1, \quad \sum_{n \in N} \|h_n\| < \infty, \quad \text{and} \quad \|\tau\|_\pi \leqq \sum_{n \in N} \|h_n\| < \|\tau\|_\pi + \varepsilon.$$

Since $F_{A_i}(h_n^* \otimes g_n) = (A_i g_n, h_a)$ $(i=1, 2)$, and since F_{A_i} are continuous it result that the series $\sum_{n \in N} (A_i g_n, h_n)$ $(i=1, 2)$ are absolutely convergent, and

$$F_{A_i}(\tau) = \sum_{n \in N} (A_i g_n, h_n)$$

Consequently,

$$\sum_{n \in N} \|h_n\| - \sum_{n \in N} \mathrm{Re}\,(A_1 g_n, h_n) \leqq 2\varepsilon.$$

Now let us notice that

$$1 - \mathrm{Re}\,(A_i g_n, f_n) = \frac{1}{2} \left\| \begin{bmatrix} A_i g_n - f_n \\ D_{A_i} g_n \end{bmatrix} \right\|^2 = \frac{1}{2} \|\hat{A}_i g_n - \hat{f}_n\|^2$$

where $f_n = \dfrac{h_n}{\|h_n\|}$ and $\hat{f}_n = \begin{pmatrix} f_n \\ 0 \end{pmatrix}$ $(n \in N)$. Since A_1 Harnack-dominates A_2 in virtue of Remark 3.1 we also have

$$\|\hat{A}_2 g_n - \hat{f}_n\|^2 = \|K(\hat{A}_1 g_n - \hat{f}_n)\|^2 \leqq \|K\|^2 \|\hat{A}_1 g_n - \hat{f}_n\|^2$$

Therefore

$$\mathrm{Re}\,(A_1 g_n, h_n) - \mathrm{Re}\,(A_2 g_n, h_n) \leqq \frac{1}{2} (\|K\|^2 - 1) \|\hat{A}_1 g_n - \hat{f}_n\|^2 \|h_n\| \quad (n \in N).$$

Whence,

$$\mathrm{Re}\, F_{A_1}(\tau) - \mathrm{Re}\, F_{A_2}(\tau) \leqq (\|K\|^2 - 1) \sum_{n \in N} \frac{1}{2} \|\hat{A}_1 g_n - \hat{f}_n\|^2 \|h_n\| =$$

$$= (\|K\|^2 - 1) \sum_{n \in N} [\|h_n\| - \mathrm{Re}\,(A_1 g_n, h_n)] < 2\varepsilon (\|K\|^2 - 1).$$

We may now state and prove our main theorem concerning the uniqueness of exact intertwining dilation.

Theorem 3.1. *Let $A_1, A_2 \in L(\mathfrak{G}; \mathfrak{H})$ be operators with the properties: $TA_1 = A_1 Z$, $TA_2 = A_2 Z$, $\|A_1\| = \|A_2\| = 1$, A_1 Harnack-dominates A_2. Then, if A_1 has a unique exact intertwining dilation so has A_2.*

Proof. By Remark 2.1, we must show that if the functional $f_{A_1} \in ((\mathfrak{H}^* \hat{\otimes}_\pi \mathfrak{G})/\mathfrak{N}_U)'$ defined by A_1 satisfies condition (8), then the functional $f_{A_2} \in ((\mathfrak{H}^* \hat{\otimes}_\pi \mathfrak{G})/\mathfrak{N}_U)'$ defined by A_2, also satisfies it.

Assume that for $\varepsilon>0$ and $\tau_0\in(\mathfrak{K}^*\hat{\otimes}_\pi\mathfrak{G})\setminus(\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G})$ we have

$$\|\dot{\tau}_1+\dot{\tau}_2\|\leqq\|\dot{\tau}_1-\dot{\tau}_0\|+\|\dot{\tau}_2+\dot{\tau}_0\|<\operatorname{Re}f_{A_1}(\dot{\tau}_1+\dot{\tau}_2)+\varepsilon \tag{11}$$

for some $\tau_1,\tau_2\in\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}$. Since $\|\dot{\tau}\|=\|\varphi(\tau)\|=\|\psi(\tau)\|$ for all $\tau\in\mathfrak{H}^*\hat{\otimes}_\pi\mathfrak{G}$, there exists $\tau'\in\mathfrak{N}_T$ such that

$$\|\tau_1+\tau_2+\tau'\|_\pi<\|\varphi(\tau_1+\tau_2)\|+\varepsilon=\|\dot{\tau}_1+\dot{\tau}_2\|+\varepsilon.$$

Denote $\tau_2'=\tau_2+\tau'$ and note that

$$\|\dot{\tau}_1+\dot{\tau}_2'\|=\|\dot{\tau}_1+\dot{\tau}_2\| \quad\text{and}\quad f_{A_i}(\dot{\tau}_1+\dot{\tau}_2')=f_{A_i}(\dot{\tau}_1+\dot{\tau}_2).$$

Then, from (11) we readily infer that

$$\|\tau_1+\tau_2'\|_\pi<\operatorname{Re}f_{A_1}(\dot{\tau}_1+\dot{\tau}_2')+2\varepsilon=\operatorname{Re}F_{A_1}(\tau_1+\tau_2')+2\varepsilon.$$

Consequently, in virtue of Lemma 3.1, it follows

$$\operatorname{Re}F_{A_1}(\tau_1+\tau_2')\leqq\operatorname{Re}F_{A_2}(\tau_1+\tau_2')+2\varepsilon(\|K\|^2-1)$$

or, equivalently,

$$\operatorname{Re}f_{A_1}(\dot{\tau}_1+\dot{\tau}_2)\leqq\operatorname{Re}f_{A_2}(\dot{\tau}_1+\dot{\tau}_2)+2\varepsilon(\|K\|^2-1).$$

Whence it results that f_{A_2} satisfies the condition

$$\|\dot{\tau}_1-\dot{\tau}_0\|+\|\dot{\tau}_2+\dot{\tau}_0\|<\operatorname{Re}f_{A_2}(\dot{\tau}_1+\dot{\tau}_2)+2\varepsilon(\|K\|^2-1).$$

Thus, we can conclude that f_{A_2} satisfies (8) too.

As a corollary of the previous theorem we have the following more general result:

Theorem 3.2. *Let T, T' be two contractions on the Hilberts spaces $\mathfrak{H}$ and $\mathfrak{H}'$, respectively. Moreover let $A_1, A_2\in L(\mathfrak{H}';\mathfrak{H})$ satisfy the conditions:*

$TA_1=A_1T'$, $TA_2=A_2T'$, $\|A_1\|=\|A_2\|=1$, A_1 Harnack-dominates A_2 Then, if A_1 has a unique exact intertwining dilations so has A_2.

Indeed, denoting by Z the minimal isometric dilation of T' it is known (see [5], Th. 2.3). that all exact intertwining dilations of A_i $(i=1,2)$ are obtained as exact intertwining dilations of the operators $B_i=A_iP_{\mathfrak{H}'}$ $(i=1,2)$ intertwining T and Z.

4. Let T, T' be two contractions on the Hilbert space $\mathfrak{H}$ and $\mathfrak{H}'$, and let U, U' be their minimal isometric dilations on the spaces $\mathfrak{K}$ and $\mathfrak{K}'$, respectively.

Theorem 4.1. *Let $B_1, B_2\in L(\mathfrak{K}';\mathfrak{K})$ have the properties:* $\|B_1\|=\|B_2\|=1$, $UB_i=B_iU'$, $PB_i(I-P')=0$ $(i=1,2)$ where, $P=P_{\mathfrak{H}}$, $P'=P_{\mathfrak{H}'}$ *B_1 Harnack-dominates B_2, and let $A_1, A_2\in L(\mathfrak{H}';\mathfrak{H})$ be the operators $A_i=PB_i|\mathfrak{H}'$ $(i=1,2)$. Then, if B_1 is an exact intertwining dilation of A_1, then A_2 is an exact intertwining dilation of A_2; moreover, if B_1 is the unique exact intertwining dilation for A_1, so is B_2 for A_2.*

Proof. First, by hypothesis we observe that $PB_i=A_iP'$ and A_i is intertwining T and T'. Thus, B_i is an intertwining dilation of A_i $(i=1,2)$.

Now, in order to prove that B_2 is an exact intertwining dilation for A_2 if B_1 is so for A_1, it suffices to show that $\|A_2\|=1$.

Clearly, we have (by definition of A_2) $\|A_2\|\leqq 1$.

For the converse inequality we observe that, since B_1 Harnack-dominates B_2, i.e. $\|D_{B_2}k'\|\leqq C\|D_{B_1}k'\|$ and $\|(B_2-B_1)k'\|\leqq C'\|D_{B_1}k'\|$ with C, $C'>0$, we have for $h'\in\mathfrak{H}'$

$$\|(1-P)B_2h'\| \leqq \|(1-P)B_1h'\|+\|(1-P)(B_2-B_1)h'\| \leqq \|D_{A_1}h'\|+\|(B_2-B_1)h'\| \leqq$$
$$\leqq \|D_{A_1}h'\|+C'\|D_{B_1}h'\| \leqq (1+C')\|D_{A_1}h'\|$$

and therefore,

$$\|D_{A_2}h'\|^2 = \|D_{B_2}h'\|^2+\|(1-P)B_2h'\|^2 \leqq \left(C^2+(1+C')^2\right)\|D_{A_1}h'\|^2 = C''\|D_{A_1}h'\|^2,$$

for any $h'\in\mathfrak{H}$.

Since $\|A_1\|=1$, we infer from this inequality that $\|A_2\|=1$ too, thus B_2 is an exact intertwining dilation of A_2.

The above relation with the following one:

$$\|(A_2-A_1)h'\| \leqq \|(B_2-B_1)h'\| \leqq C'\|D_{B_1}h'\| \leqq C'\|D_{A_1}h'\| \quad (h'\in\mathfrak{H}')$$

means that A_1 Harnack-dominates A_2. Now the second statement of this theorem can be obtained by referring to Theorem 3.2.

Lemma 4.1. *Let $B_1, B_2\in L(\mathfrak{K}';\mathfrak{K})$, $\|B_1\|=\|B_2\|=1$ be of the form $B_i=B_0\oplus S_i$ where S_i are strict contractions $(i=1, 2)$. Then B_1, B_2 Harnack-dominate each other.*

Proof. Consider the decomposition $\mathfrak{K}'=\mathfrak{K}_0'\oplus\mathfrak{K}_1'$ for which

$$B_1P_{\mathfrak{K}_0'} = B_2P_{\mathfrak{K}_0'} = B_0 \quad\text{and}\quad S_i = B_iP_{\mathfrak{K}_1'} = B_i(1-P_{\mathfrak{K}_0'})$$

and note that

$$\|D_{B_i}k'\|^2 = (\|k_0'\|^2-\|B_0k_0'\|^2)+(\|k_1'\|^2-\|S_ik_1'\|^2) \geqq$$
$$\geqq \|k_1'\|^2-\|S_ik_1'\|^2 \geqq (1-\|S_i\|^2)\|k_1'\|^2, \quad\text{where}\quad k_0' = P_{\mathfrak{K}_0'}k', \quad k_1' = P_{\mathfrak{K}'}k'.$$

Whence, by taking $C=\max\{(1-\|S_1\|^2)^{-1/2}, (1-\|S_2\|^2)^{-1/2}\}$ it follows

$$\|P_{K_1'}k'\| \leqq C\|D_{B_i}k'\| \quad\text{for all}\quad k'\in\mathfrak{K}'.$$

Therefore, we have $\|(B_2-B_1)k'\|\leqq\|S_2-S_1\|\|k_1'\|\leqq C'\|D_{B_i}k'\|$ and also

$$\|D_{B_2}k'\|^2 = \|k'\|^2-\|B_0k_0'\|^2-\|S_2k_1'\|^2 = \|D_{B_1}k'\|^2+(\|S_1k_1'\|-\|S_1k_1'\|)(\|S_2k_1'\|+\|S_2k_1'\|)$$
$$\leqq \|D_{B_1}k'\|^2+\|S_1-S_2\|(\|S_1\|+\|S_2\|)\|k_1'\|^2;$$

hence $\|D_{B_2}k'\|\leqq C''\|D_{B_1}k'\|$ for all $k'\in\mathfrak{K}$, where C', C'' are constants.

Thus B_1 Harnack-dominates B_2. By symmetry B_2 also Harnack-dominates B_1.

Theorem 4.1 and Lemma 4.1 have the following

Corollary 4.1. *Let $B_1, B_2\in L(\mathfrak{K}';\mathfrak{K})$ be two operators as in Lemma 4.1, intertwining U and U' and such that: $B_i(\mathfrak{K}'\ominus\mathfrak{H}')\subset\mathfrak{K}\ominus\mathfrak{H}$ $(i=1, 2)$. Then, B_1 is an* *exact intertwining dilation of* $A_1=P_{\mathfrak{H}}B_1|\mathfrak{H}'$, *if and only if* B_2 *is an exact intertwining dilation of* $A_2=P_{\mathfrak{H}}B_2|\mathfrak{H}'$; *moreover,* B_1 *is the unique exact intertwining dilation for* A_1 *if and only if* B_2 *is so for* A_2.

In virtue of Theorems 2 and 5 of [2], we also have the following corollary of Theorem 4.1, concerning the Hankel operators.[2])

Corollary 4.2. *Let* $F_1, F_2 \in L^\infty(\mathfrak{E}, \mathfrak{F})$ ($\mathfrak{E}, \mathfrak{F}$-separable Hilbert spaces) *have the properties:*

$\|F_1\|=\|F_2\|=1$,

$F_1(t)=F_2(t)$ *whenever* $\max\{\|F_1(t)\|, \|F_2(t)\|\}>1-\theta$ *for some fixed* θ, $0<\theta<1$.

Then, if one of these functions is a minifunction for its Hankel operator, then so is the other. Moreover, if one of them is the unique minifunction of its Hankel operator so is the other.

Proof. Set $\sigma=\{t\in[0,1]: \max\{\|F_1(t)\|, \|F_2(t)\|\}>1-\theta\}$, and $\mathfrak{L}_0=\chi_\sigma L^2(\mathfrak{E})$, $\mathfrak{L}_1=\chi_{[0,1]\setminus\sigma}L^2(\mathfrak{E})$ where χ_σ is the characteristic function of σ. Then $L^2(\mathfrak{E})=\mathfrak{L}_0\oplus\mathfrak{L}_1$. Also, denoting by B_i the operators: $f\to F_i f$ from $L^2(\mathfrak{E})$ to $L^2(\mathfrak{F})$ ($i=1,2$), we observe that

$$B_1P_{\mathfrak{L}_0}=B_2P_{\mathfrak{L}_0},\quad B_i\mathfrak{L}_0\subset\chi_\sigma L^2(\mathfrak{F})\quad\text{and}\quad B_i\mathfrak{L}_1\subset\chi_{[0,1]\setminus\sigma}L^2(\mathfrak{F}).$$

Thus the operators B_i can be written $B_i=B_0\oplus S_i$ where

$$B_0=B_iP_{\mathfrak{L}_0},\quad S_i=B_iP_{\mathfrak{L}_1}\quad\text{and}\quad \|S_i\|<1\quad(i=1,2).$$

Now Corollary 4.2 follows at once by Corollary 4.1.

References

[1] V. M. Adamyan, D. Z. Arov, M. G. Krein, Бесконечные блочно-Ганкелевы матрицы и связанные с ними проблемы продолжения, *Изв. Акад. Наук Армянской ССР*, 6 (1971), 87—112.

[2] L. B. Page, Bounded and compact vectorial Hankel operators, *Trans. Amer. Math. Soc.*, **150** (1970), 529—539.

[3] I. Suciu, *Function algebras*, Ed. Acad. and Nordhoff Inter. Publ. (Bucureşti and Leyden, 1975).

[4] I. Suciu, Analytic relations between functional models for contractions, *Acta Sci. Math.*, **34** (1973), 359—365.

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

[6] B. Sz.-Nagy, C. Foiaş, Commutants de certains opérateurs, *Acta Sci. Math.*, **29** (1968), 1—17.

[7] F. Trèves, *Topological vector spaces, distributions and kernels*, Acad. Press (New York/London, 1967).

„INCREST"
CALEA VICTORIEI 114
BUCHAREST 22, ROMANIA

[2]) This corollary can be also obtained as a consequence of Theorems 1.3 and 3.1 of [1].

Acta Sci. Math., 38 (1976), 291—300

Universal quasinilpotent operators

DOMINGO A. HERRERO

1. Introduction. Let $\mathfrak{H}$ be a complex Hilbert space of (topological) dimension h and let $\mathcal{L}(\mathfrak{H})$ be the algebra of all (bounded linear) operators in $\mathfrak{H}$. Given T in $\mathcal{L}(\mathfrak{H})$, let $\mathcal{S}(T)=\{WTW^{-1}$: W is invertible in $\mathcal{L}(\mathfrak{H})\}$ ("similarity orbit" of T). What is $\mathcal{S}(T)^-$, the norm-closure of $\mathcal{S}(T)$? In this note it will be shown that the similarity orbit of a quasinilpotent perator could be surprisingly large. The norm-closure of the set $\mathcal{N}(\mathfrak{H})=\{Q\in\mathcal{L}(\mathfrak{H})$: Q is nilpotent$\}$ was completely characterized in [1] (separable case) and [11] (non-separable case); it was shown, in particular, that every quasinilpotent operator belongs to $\mathcal{N}(\mathfrak{H})^-$. Since $\mathcal{N}(\mathfrak{H})^-$ is invariant under similarities, it readily follows that $\mathcal{S}(Q)^-$ must be contained in $\mathcal{N}(\mathfrak{H})^-$ for every quasinilpotent operator Q. The main result says that the converse inclusion is also true for a suitably chosen Q.

First of all, consider the finite dimensional case. Assume that T is a nilpotent operator on a Hilbert space $\mathfrak{H}$ of dimension n ($0<n<\infty$). Then there exists an orthonormal basis $\{e_1, \ldots, e_n\}$ with respect to which T can be written as a matrix $T=(t_{jk})_{j,k=1}^n$, where $t_{jk}=0$ for all $j\geq k$ (i.e., an upper triangular matrix with 0's in the diagonal). Given $\varepsilon>0$, let $T_\varepsilon=(t_{jk,\varepsilon})_{j,k=1}^n$, where $t_{jk,\varepsilon}=t_{jk}$ if $k\neq j+1$ or $t_{j,j+1}\neq 0$ and $t_{j,j+1,\varepsilon}=\varepsilon$ if $k=j+1$ and $t_{j,j+1}=0$. Clearly, $\|T-T_\varepsilon\|\leq\varepsilon$ and T_ε is similar to its Jordan form, given by the matrix $Q_{un}=(\delta_{j+1,k})$, where δ_{jk} denotes the Kronecker delta. Since ε can be chosen arbitrarily small, we have arrived to the following result:

Lemma 1. *Let $\mathfrak{H}$ be an n-dimensional Hilbert space* ($0<n<\infty$) *and let* $Q_{un}=(\delta_{j+1,k})$ *(with respect to some ONB). Then $\mathcal{S}(Q_{un})^-$ coincides with the set of all nilpotent operators in $\mathfrak{H}$.*

2. The ideal of compact operators. Let $\mathcal{K}(\mathfrak{H})$ denote the ideal of compact operators on a Hilbert space $\mathfrak{H}$ of infinite dimension h.

Lemma 2. *The compact quasinilpotent operator* $K_{uh}\approx\left(\bigoplus_{n=1}^{\infty} 1/nQ_{un}\right)\oplus 0$, *where* 0 *is the zero operator acting on a subspace of dimension* h ($\approx$ *means* "*unitarily equivalent to*") *has the property:* $\mathcal{S}(K_{uh})^- = \{K\in\mathcal{K}(\mathfrak{H})$: K is quasinilpotent$\}$.

Received September 13, 1975.

Proof. Let K be a compact quasinilpotent operator. Then $\mathfrak{H}=\mathfrak{H}_0\oplus\mathfrak{H}_1$, where $\mathfrak{H}_0, \mathfrak{H}_1$ reduce K, $\dim\mathfrak{H}_0=\aleph_0$ and $K|\mathfrak{H}_1=0$ (the vertical bar denotes restriction). Now it is clear that, by a trivial modification of the proof given by R. G. Douglas in [8] for the case when $\mathfrak{H}$ is separable, it can be shown that K is a norm limit of finite rank nilpotents. On the other hand, we already know that the set of all compact quasinilpotents is closed in $\mathscr{L}(\mathfrak{H})$ (see, e.g., [12]). Thus, in order to complete the proof we only have to show that $\mathscr{S}(K_{uh})^-$ actually contains every finite rank nilpotent.

Let F be a finite rank nilpotent in $\mathscr{L}(\mathfrak{H})$. Then there exists a finite dimensional subspace $\mathfrak{H}_n$ of dimension n, $0<n<\infty$, reducing F such that $F|\mathfrak{H}_n^\perp=0$. Up to a unitary transformation (of $\mathfrak{H}$ onto itself) we can obviously assume that $\mathfrak{H}_n$ is the space of Q_{un}. Hence, $F|\mathfrak{H}_n\in\mathscr{S}(Q_{un})^-$ (use Lemma 1).

Since $K_{uh}=(1/n)Q_{un}\oplus K_n''$ (with respect to the decomposition $\mathfrak{H}=\mathfrak{H}_n\oplus\mathfrak{H}_n^\perp$), where K_n'' is a quasinilpotent operator acting on $\mathfrak{H}_n^\perp$, it follows from [16] that $(1/n)Q_n\oplus 0\in\mathscr{S}(K_{uh})^-$. Since Q_n and $(1/n)Q_n$ are similar, we conclude that $F\in\mathscr{S}(K_{uh})^-$. □

This result suggests the following

Definition 1. A (necessarily quasinilpotent, but not nilpotent) operator $Q_u(\mathscr{J})$ satisfying the equality $\mathscr{S}[Q_u(\mathscr{J})]^-=\{Q\in\mathscr{J}\colon Q \text{ is quasinilpotent}\}$ for a given closed bilateral ideal $\mathscr{J}$ of $\mathscr{L}(\mathfrak{H})$ will be called a *universal quasinilpotent for the ideal* $\mathscr{J}$.

Let K be an arbitrary compact quasinilpotent, but not nilpotent, operator. Then ([8]) there exists a vector $x\in\mathfrak{H}$ such that $K^n x\neq 0$ for all $n=0, 1, 2, \ldots$. Let $\mathfrak{H}$ be the (closed) subspace spanned by $\{K^n x\}_{n=0}^\infty$ and let

$$K=\begin{bmatrix} K_{11} & K_{12} \\ 0 & K_{22}\end{bmatrix}$$

be the matrix representation of K with respect to the orthogonal decomposition $\mathfrak{H}=\mathfrak{H}_x\oplus\mathfrak{H}_x^\perp$. Clearly, K_{11} and K_{22} are quasinilpotent operators, so that we can proceed as in [12] in order to show that $K_{11}\oplus 0\in\mathscr{S}(K)^-$. Assuming that K_{11} is similar to a compact weighted shift with non-zero weights, it is not difficult to prove (by using the arguments of [12] and the proof of Lemma 2) that K_{11} and, a fortiori, K are compact universal quasinilpotents. This suggests the following

Conjecture 1. *A compact quasinilpotent operator is either nilpotent or a compact universal quasinilpotent.*

The above observations reduce this conjecture to the analysis of those compact quasinilpotents having a cyclic vector.

3. Similarity orbits of certain normal operators. Our next step will be a partial characterization of the set $\mathcal{S}(N)^-$ for the case when N is a normal operator. (A more complete description of this case will be given in an oncoming article [13].)

The closed bilateral ideals of $\mathcal{L}(\mathfrak{H})$ have been completely characterized by several authors ([3; 6; 14]): Let α be a cardinal number such that $\aleph_0 \leqq \alpha \leqq h = \dim \mathfrak{H}$ and let $\mathcal{J}_\alpha$ be the norm-closure of the set of all operators T in $\mathcal{L}(\mathfrak{H})$ such that $\dim (T\mathfrak{H})^- < \alpha$. Then $\mathcal{J}_\alpha$ is a closed bilateral ideal of $\mathcal{L}(\mathfrak{H})$ and every such proper (non-zero) ideal has this form. The *weighted spectrum* of $A \in \mathcal{L}(\mathfrak{H})$ corresponding to $\mathcal{J}_\alpha$ is the spectrum $\Lambda_\alpha(A)$ of the canonical projection of A in the quotient algebra $\mathcal{L}(\mathfrak{H})/\mathcal{J}_\alpha$; namely, $\Lambda_{\aleph_0}(A) = E(A)$ is the usual Calkin essential spectrum of A, and $\Lambda_h(A)$ is the *heavy spectrum* (i.e., the one corresponding to the largest ideal). For the analysis of these weighted spectra, as well as for the definition and properties of the *approximate nullity* $\delta(A)$ of an operator A, the reader is referred to [4; 11]. We recall that, in the separable case, the condition $\delta(\lambda - A) = \delta(\bar{\lambda} - A^*)$ (where A^* denotes the adjoint of the operator A) for all complex λ is equivalent to saying that if $(\lambda - A)$ is a semi-Fredholm operator, then its index is 0, i.e., A is a bi-quasitriangular operator in the sense of [1; 2].

Theorem 1. *Let N be a normal operator such that $\Lambda(N)$ (the spectrum of N) is a perfect set and coincides with $\Lambda_h(N)$. Then $\mathcal{S}(N)^-$ contains every operator $A \in \mathcal{L}(\mathfrak{H})$ such that $\Lambda(A) = \Lambda_h(A) = \Lambda(N)$ and $\delta(\lambda - A) = \delta(\bar{\lambda} - A^*)$* for all complex λ.

Let A be as in Theorem 1. By using the results of [2, Theorem 2.2] and [11] we can see that, given $\varepsilon > 0$, there exists an operator A' satisfying the same hypotheses as A such that $\|A - A'\| < \varepsilon$ and

$$A' \approx \begin{bmatrix} N & 0 & T_1 \\ 0 & N & L_1 \\ 0 & 0 & L_2 \end{bmatrix} = \begin{bmatrix} N & T \\ 0 & L \end{bmatrix}, \quad T = [0 \ T_1], \quad L = \begin{bmatrix} N & L_1 \\ 0 & L_2 \end{bmatrix}.$$

(All these matrices of operators are referred to suitable orthogonal direct sum decompositions of the underlying spaces.) It readily follows that L also satisfies the hypotheses of Theorem 1. Therefore, by [11; 18], L is a norm limit of algebraic operators with spectra contained in $\Lambda(N)$; furthermore, by an easy approximation argument, L can be actually approximated in the norm by operators which are similar to normal operators with *finite spectrum* contained in $\Lambda(N)$. Thus, in order to complete the proof of Theorem *1* it will be enough to prove the following weaker version of it:

Theorem 1′. *Let N be a normal operator in $\mathcal{L}(\mathfrak{H})$ such that $\Lambda(N) = \Lambda_h(N)$ is a perfect set, let $T: \mathfrak{H}' \to \mathfrak{H}$ be an arbitrary continuous linear mapping from a Hilbert space $\mathfrak{H}'$, $\dim \mathfrak{H}' = h' \leqq h$, and let $M, W \in \mathcal{L}(\mathfrak{H}')$, where M is normal with a finite* *spectrum contained in $\Lambda(N)$ and W is invertible. Then $\mathcal{S}(N)^-$ contains every operator in $\mathcal{L}(\mathfrak{H})$ unitarily equivalent to*

$$\begin{bmatrix} N & T \\ 0 & WMW^{-1} \end{bmatrix}$$

(with respect to the orthogonal direct sum decomposition $\mathfrak{H}\oplus\mathfrak{H}'$).

The proof will be given in a series of lemmas.

Lemma 3. *Let N be as in Theorem* 1 *and let $\lambda\in\Lambda(N)$. If*

$$A \approx \begin{bmatrix} N & T \\ 0 & \lambda I' \end{bmatrix}$$

(I' = identity on $\mathfrak{H}'$), then $A\in\mathcal{S}(N)^-$.

Proof. Clearly, we can translate N by a multiple of the identity and assume that $\lambda=0$. According to the characterization of the norm closure of $\mathcal{U}(N)=\{UNU^{-1}: U$ is unitary$\}$ given in [12] (see also [7]), $\mathcal{U}(N)^-$ (which is obviously contained in $\mathcal{S}(N)^-$) contains every normal operator $N'\approx N\oplus 0'$, where $0'$ denotes the zero operator in $\mathcal{U}'$.

Case I: h' is finite.

In this case A is a compact perturbation of an operator N' as above and the result follows from [10, Lemma 1].

Case II: $\aleph_0\leqq h'<h$.

Proceeding as in [11], it is possible to find an orthogonal direct sum decomposition $\mathfrak{H}=\mathfrak{H}_0\oplus\mathfrak{H}''$, such that $\dim\mathfrak{H}''=h'$, $\dim\mathfrak{H}_0=h$ and

$$\begin{bmatrix} N & T \\ 0 & 0' \end{bmatrix} = N_0\oplus\begin{bmatrix} N'' & T'' \\ 0'' & 0' \end{bmatrix}$$

with respect to $\mathfrak{H}_0\oplus\mathfrak{H}''\oplus\mathfrak{H}'$, where $N_0\in\mathcal{L}(\mathfrak{H}_0)$, $N''\in\mathcal{L}(\mathfrak{H}'')$ are normal operators satisfying $\Lambda(N_0)=\Lambda_h(N_0)=\Lambda(N'')=\Lambda_{h'}(N'')=\Lambda(N)$.

This reduces our problem to

Case III: $h'=h$.

Given $\varepsilon>0$, we can find an ε', $0<\varepsilon'<\min\{\varepsilon, 1\}$ such that if $\Delta_0=\{\lambda: |\lambda|\geqq\varepsilon'\}$ and $\Delta_0'=\{\lambda: |\lambda|<\varepsilon\}$, then $\Delta_0\cap\Lambda(N)$ and $[\Delta_0'\cap\Lambda(N)]^-$ are nonempty perfect sets. To simplify the notation, we can directly assume that $\varepsilon'=\varepsilon$ and $0<\varepsilon<1$. Let $E(\cdot)$ be the spectral measure of N; then $E(\Delta_0)\mathfrak{H}=\mathfrak{H}_0$ and $E(\Delta_0')\mathfrak{H}=\mathfrak{H}_0'$ are complementary h-dimensional orthogonal reducing subspaces of N and N can be written as $N=$

$=N_0 \oplus N_0'$, where $N_0 \in \mathscr{L}(\mathfrak{H}_0)$ and $N_0' \in \mathscr{L}(\mathfrak{H}_0')$, with respect to this decomposition. Then we can also write

$$B = \begin{bmatrix} N & T \\ 0 & 0' \end{bmatrix} = \begin{bmatrix} N_0 & 0 & T_1 \\ 0 & N_0' & T_2 \\ 0 & 0 & 0' \end{bmatrix}$$

with respect to $\mathfrak{H}_0 \oplus \mathfrak{H}_0' \oplus \mathfrak{H}'$.

Combining T_2 with an isometry V from $\mathfrak{H}_0'$ onto $\mathfrak{H}'$ and using the polar decomposition of VT_2, it is not difficult to see that $\mathfrak{H}_0'$ and $\mathfrak{H}'$ can be written as orthogonal direct sums $\mathfrak{H}_0' = \mathfrak{H}_{0a}' \oplus \mathfrak{H}_{0b}'$ and $\mathfrak{H}' = \mathfrak{H}_a' \oplus \mathfrak{H}_b'$, where $\dim \mathfrak{H}_{0a}' = \dim \mathfrak{H}_{0b}' = \dim \mathfrak{H}_a' = \dim \mathfrak{H}_b' = h$ and $T_2 \mathfrak{H}_a' \subset \mathfrak{H}_{0a}'$ and $T_2 \mathfrak{H}_b' \subset \mathfrak{H}_{0b}'$. Therefore, we can write $T_2 = T_{2a} \oplus T_{2b}$, where $T_{2a}(T_{2b}) = T_2 | \mathfrak{H}_a'$ ($\mathfrak{H}_b'$, resp.) and

$$B = \begin{bmatrix} N_0 & 0 & T_1 \\ 0 & N_0' & T_{2a} \oplus T_{2b} \\ 0 & 0 & 0 \oplus 0 \end{bmatrix}.$$

Let $\Delta_j = \{\lambda: \varepsilon_{j+1} \leqq |\lambda| < \varepsilon_j\}$, $j=1, 2, 3, 4$, be such that $[\Delta_j \cap \Lambda(N_0')]^-$ is perfect for all j and $0 = \varepsilon_5 < \varepsilon_4 < \varepsilon_3 < \varepsilon_2 < \varepsilon^2 < \varepsilon_1 = \varepsilon$. Proceeding as in the first part of the proof, we van decompose $\mathfrak{H}_0' = \bigoplus_{j=1}^{4} \mathfrak{H}_j$ and $N_0' = \bigoplus_{j=1}^{4} N_j$ in such a way that $N_j \in \mathscr{L}(\mathfrak{H}_j)$ and $\Lambda(N_j) = [\Delta_j \cap \Lambda(N_0')]^-$. Now choose arbitrary normal operators $M_1 \in \mathscr{L}(\mathfrak{H}_{0a}')$, $M_2 \in \mathscr{L}(\mathfrak{H}_{0b}')$, $M_3 \in \mathscr{L}(\mathfrak{H}_a')$ and $M_4 \in \mathscr{L}(\mathfrak{H}_b')$ such that $M_j \approx N_j$, $j=1, 2, 3, 4$. Since $\Lambda(M_1) \cap \Lambda(M_3) = \Lambda(M_2) \cap \Lambda(M_4) = \emptyset$, it follows from ROSENBUML's Corollary ([15, Corollary 0.15]) that the operators $M_1 \oplus M_3$ and $M_2 \oplus M_4$ are similar to

$$\begin{bmatrix} M_1 & T_{2a} \\ 0 & M_3 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} M_2 & T_{2b} \\ 0 & M_4 \end{bmatrix},$$

respectively. Hence,

$$R = N_0 \oplus \begin{bmatrix} M_1 & T_{2a} \\ 0 & M_3 \end{bmatrix} \oplus \begin{bmatrix} M_2 & T_{2b} \\ 0 & M_4 \end{bmatrix} = \begin{bmatrix} N_0 & 0 & 0 \\ 0 & M_1 \oplus M_2 & T_{2a} \oplus T_{2b} \\ 0 & 0 & M_3 \oplus M_4 \end{bmatrix} =$$

$$= \begin{bmatrix} N_0 & 0 & 0 \\ 0 & M_1 \oplus M_2 & T_2 \\ 0 & 0 & M_3 \oplus M_4 \end{bmatrix}$$

is similar to N. Thus, if $X = -N_0^{-1} T_1$ and

$$W = \begin{bmatrix} I & 0 & X \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}, \quad \text{then} \quad WRW^{-1} = \begin{bmatrix} N_0 & 0 & T_1 - X(M_3 \oplus M_4) \\ 0 & M_1 \oplus M_2 & T_2 \\ 0 & 0 & M_3 \oplus M_4 \end{bmatrix}.$$

Since $\|B-WRW^{-1}\| \leqq \|X(M_3 \oplus M_4)\| + \|N_0' - M_1 \oplus M_2\| + \|M_3 \oplus M_4\| \leqq \varepsilon^2 \|N_0^{-1}\| \cdot \|T\| + 2\varepsilon + \varepsilon^2 \leqq \varepsilon\|T\| + 2\varepsilon + \varepsilon^2 < (3+\|T\|)\varepsilon$ and WRW^{-1} is similar to N, we conclude that dist $[A, \mathcal{S}(N)] < (3+\|T\|)\varepsilon$, whence the result follows. □

Lemma 4. *Lemma 3 remains true if N is replaced by WNW^{-1}, for some invertible W.*

Proof. Clearly, $\mathcal{S}(N)^- = \mathcal{S}(WNW^{-1})^-$ and therefore it is enough to show that if

$$A \approx \begin{bmatrix} WNW^{-1} & T \\ 0 & \lambda I' \end{bmatrix},$$

then $A \in \mathcal{S}(N)^-$.

By Lemma 3, every operator $A' \in \mathcal{L}(\mathfrak{H})$ such that

$$A' \approx \begin{bmatrix} N & W^{-1}T \\ 0 & \lambda I' \end{bmatrix}$$

belongs to $\mathcal{S}(N)^-$.

On the other hand,

$$\begin{bmatrix} W & 0 \\ 0 & I' \end{bmatrix} \begin{bmatrix} N & W^{-1}T \\ 0 & \lambda I' \end{bmatrix} \begin{bmatrix} W & 0 \\ 0 & I' \end{bmatrix}^{-1} = \begin{bmatrix} WNW^{-1} & T \\ 0 & \lambda I' \end{bmatrix}.$$

Since $\mathcal{S}(N)^-$ is invariant under similarities ([12]), it readily follows that $A \in \mathcal{S}(N)^-$. □

Lemma 5. *Let N be as in Theorem 1, let $\{\lambda_1, \ldots, \lambda_m\}$ be a finite subset of $\Lambda(N)$, let I_j be the identity operator on a Hilbert space $\mathfrak{H}_j$ of dimension $h_j \leqq h$, and let $M = \bigoplus_{j=1}^{m} \lambda_j I_j \in \mathcal{L}(\mathfrak{H}')$, where $\mathfrak{H}' = \bigoplus_{j=1}^{m} \mathfrak{H}_j$. Then $\mathcal{S}(N)^-$ contains every operator $A \in \mathcal{L}(\mathfrak{H})$ unitarily equivalent to*

$$\begin{bmatrix} N & T \\ 0 & M \end{bmatrix}.$$

(With respect to the orthogonal direct sum $\mathfrak{H} \oplus \mathfrak{H}'$.)

Proof. This follows by induction over m. For $m=1$, it is the result of Lemma 3. Assume that the result is true for $m=n$ and let $m=n+1$. Set $M = M_n \oplus \lambda_{n+1} I_{n+1}$, where $M_n = \bigoplus_{j=1}^{n} \lambda_j I_j$; then

$$\begin{bmatrix} N & T \\ 0 & M \end{bmatrix} = \begin{bmatrix} N & T_n & T_{n+1} \\ 0 & M_n & 0 \\ 0 & 0 & \lambda_{n+1} I_{n+1} \end{bmatrix} = \begin{bmatrix} N_n & T_{n+1} \\ 0 & \lambda_{n+1} I_{n+1} \end{bmatrix}, \quad \text{where} \quad N_n = \begin{bmatrix} N & T_n \\ 0 & M_n \end{bmatrix}.$$

(The first matrix corresponds to the decomposition $\mathfrak{H}\oplus\mathfrak{H}'$, the second one to $\mathfrak{H}\oplus\oplus\left(\bigoplus_{j=1}^{n}\mathfrak{H}_j\right)\oplus\mathfrak{H}_{n+1}$ and the third one to $\left[\mathfrak{H}\oplus\left(\bigoplus_{j=1}^{n}\mathfrak{H}_j\right)\right]\oplus\mathfrak{H}_{n+1}$; the matrix of N_n corresponds to the decomposition $\mathfrak{H}\otimes\left(\bigoplus_{j=1}^{n}\mathfrak{H}_j\right)\Big)$.

By our inductive hypothesis, there exists an operator $N_n'\in\mathcal{L}\left[\mathfrak{H}\oplus\left(\bigoplus_{j=1}^{n}\mathfrak{H}_j\right)\right]$, similar to N, such that $\|N_n-N_n'\|$ is smaller than an arbitrarily small given $\varepsilon>0$. On the other hand, by Lemma 4,

$$\begin{bmatrix} N_n' & T_{n+1} \\ 0 & \lambda_{n+1}I_{n+1}\end{bmatrix}$$

can be approximated in the norm by operators similar to N_n'.

Since

$$\begin{bmatrix} N_n & T_{n+1} \\ 0 & \lambda_{n+1}I_{n+1}\end{bmatrix}-\begin{bmatrix} N_n' & T_{n+1} \\ 0 & \lambda_{n+1}I_{n+1}\end{bmatrix}=(N_n-N_n')\oplus 0_{n+1},$$

dist $[A, \mathcal{S}(N)]\leqq\|N_n-N_n'\|<\varepsilon$, whence the result follows. □

Proof of Theorem 1′. The last step of the proof is very similar to that of Lemma 4. Indeed, observe that if M is chosen as in Lemma 5 and W is an invertible operator in $\mathcal{L}(\mathfrak{H}')$, then

$$\begin{bmatrix} I & 0 \\ 0 & W\end{bmatrix}\begin{bmatrix} N & TW \\ 0 & M\end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & W\end{bmatrix}^{-1}=\begin{bmatrix} N & T \\ 0 & WMW^{-1}\end{bmatrix}.$$

Since

$$\begin{bmatrix} N & TW \\ 0 & M\end{bmatrix}$$

can be uniformly approximated by operators similar to N (Lemma 5) and $\mathcal{S}(N)^-$ is invariant under similarities ([12]), we are done. □

4. The main result. The following result is our goal.

Theorem 2. *For every dimension $h\geqq\aleph_0$ there exists a universal quasinilpotent operator $Q_{uh}\in\mathcal{L}(\mathfrak{H})$*, dim $\mathfrak{H}=h$.

Proof. The proof combines the result of Theorem 1 with an argument due to N. Salinas ([5, Theorem 3.2]). Let $H_k\in\mathcal{L}(\mathfrak{H})$ be an hermitian operator such that $\Lambda(H_k)=\Lambda_h(H_k)=[0, 1/k]$ $(k=1, 2, \dots)$. According to [9; 11], there exists a sequence $\{R_{kn}\}_{n=1}^{\infty}$ of nilpotent operators such that $\|H_k-R_{kn}\|<1/n$, $n=1, 2, \dots$. By [16], there also exist nilpotent operators R_{kn}' similar to R_{kn}, such that $\|R_{kn}'\|<1/(k\cdot n)$.

Let Q_{uh} be an arbitrary quasinilpotent operator in $\mathcal{L}(\mathfrak{H})$, unitarily equivalent to $\bigoplus_{k,n=1}^{\infty} R_{kn}'$. Proceeding as in the proof of Lemma 2, we can see that $\mathcal{S}(Q_{uh})^-$ contains every operator unitarily equivalent to $R'_{kn}\oplus 0$ (for every fixed pair of values k and n). A fortiori, every $H'_k\approx H_k\oplus 0$ belongs to $\mathcal{S}(Q_{uh})^-$.

Let Q be an arbitrary quasinilpotent operator in $\mathcal{L}(\mathfrak{H})$. It follows from [2; 10; 11] that there exists an operator Q_k unitarily equivalent to

$$\begin{bmatrix} H'_k & T \\ 0 & L \end{bmatrix},$$

where $\Lambda(L)\subset\Lambda(H'_k)=\Lambda_h(H'_k)=[0, 1/k]$, such that $\|Q-Q_k\|<2/k$. Since, by Theorem 1, $Q_k\in\mathcal{S}(H'_k)^-\subset\mathcal{S}(Q_{uh})^-$ for $k=1, 2, \ldots$, it is easy to see that Q belongs to $\mathcal{S}(Q_{uh})^-$ too. □

5. Universal quasinilpotents for other closed bilateral ideals of $\mathcal{L}(\mathfrak{H})$. Let $\mathcal{J}_\alpha$ be a non-zero proper closed bilateral ideal of $\mathcal{L}(\mathfrak{H})$. Does there always exist a universal quasinilpotent for $\mathcal{J}_\alpha$? The answer is NO. Indeed, the existence of such universal operator depends on the cardinal α. Following [4; 6], we shall say that α is $\aleph_0$-regular if it cannot be written in the form $\alpha=\sum_{n=1}^{\infty}\alpha_n$ $(=\sup_n \alpha_n)$ for a sequence $\{\alpha_n\}_{n=1}^{\infty}$ of cardinal numbers strictly smaller than α; α is called $\aleph_0$-irregular in the converse case. Now the complete answer to the above question is given by the following

Theorem 3. *Let* $\dim\mathfrak{H}=h\geqq\aleph_0$ *and let* $\mathcal{J}_\alpha$, $\aleph_0\leqq\alpha\leqq h$, *be a proper closed bilateral ideal of* $\mathcal{L}(\mathfrak{H})$. *If neither*

(i) $\alpha=\aleph_{\nu+1}$ *for some ordinal* ν, *or*

(ii) α *is* $\aleph_0$*-irregular,*

then there exists a universal quasinilpotent operator $K_u=K_u(\alpha; h)$ *for* $\mathcal{J}_\alpha$.

On the other hand, if α *is an* $\aleph_0$*-regular limit cardinal, then* $\mathcal{S}(K)^-\subset\mathcal{J}_\beta$ *for some cardinal* β *strictly smaller than* α, *and therefore there is no universal quasinilpotent operator for* $\mathcal{J}_\alpha$.

Proof. Lemma 2 takes care of the case when $\alpha=\aleph_0$, so we can restrict our attention to the case $\alpha>\aleph_0$. We shall need the following auxiliary result.

Lemma 6. *Let* $\aleph_0<\alpha\leqq h=\dim\mathfrak{H}$. *Then the closure of the set of all nilpotent operators in* $\mathcal{J}_\alpha$ *coincides with* $\mathcal{J}_\alpha\cap\mathcal{N}(\mathfrak{H})^-$. *In particular, this set contains every quasinilpotent element of* $\mathcal{J}_\alpha$.

Proof. Let $T\in\mathcal{J}_\alpha\cap\mathcal{N}(\mathfrak{H})^-$. Then there exist two sequences of operators, $\{T_n\colon \dim(T_n\mathfrak{H})^-=\alpha_n<\alpha\}_{n=1}^{\infty}$ and $\{Q_n\colon Q_n\in\mathcal{N}(\mathfrak{H})\}$ such that $\|T-T_n\|+\|T-Q_n\|<1/n$. Proceeding as in [11] we can find a subspace $\mathfrak{H}_n$ of dimension $\alpha'_n=\max\{\alpha_n, \aleph_0\}$ reducing T_n and Q_n, such that $T_n|\mathfrak{H}_n^\perp=0$. Clearly, $\|T_n|\mathfrak{H}_n-Q_n|\mathfrak{H}_n\|\leqq\|T_n-Q_n\|<2/n$.

Let $R_n=(Q_n|\mathfrak{H}_n)\oplus(0|\mathfrak{H}_n^\perp)$. It readily follows that $R_n\in\mathcal{J}_\alpha$ and that $R_n^{k_n}=0$ if $Q_n^{k_n}=0$, i.e., R_n is a nilpotent element of $\mathcal{J}_\alpha$. Moreover, $\|T-R_n\|\leq\|T-T_n\|+\|T_n-R_n\|<3/n$. Hence T is a norm limit of nilpotent elements of $\mathcal{J}_\alpha$. Therefore, $\mathcal{J}_\alpha\cap\mathcal{N}(\mathfrak{H})^-\subset\{Q\in\mathcal{J}_\alpha\colon Q \text{ is nilpotent}\}^-$. Since the converse inclusion is trivial, we have proved the first statement, the second one follows from [11]. □

Now we are in a position to finish the proof of Theorem 3. By Lemma 6, it will be enough to show that if $\alpha>\aleph_0$, then $\mathcal{S}(K_u)^-$ contains $\mathcal{J}_\alpha\cap\mathcal{N}(\mathfrak{H})^-$, for a suitable $K_u\in\mathcal{J}_\alpha$.

If α satisfies (i), $\mathcal{J}_\alpha=\{T\in\mathcal{L}(\mathfrak{H})\colon \dim(T\mathfrak{H})^-\leq\aleph_v\}$ ([6; 14]) and the result follows as in Theorem 2; in fact, if $K\in\mathcal{J}_\alpha$ is nilpotent, then $\mathfrak{H}=\mathfrak{H}_0\oplus\mathfrak{H}_1$, where $\mathfrak{H}_0$, $\mathfrak{H}_1$ reduce K, $\dim\mathfrak{H}_0=\aleph_v$ and $K|\mathfrak{H}_1=0$. If $Q_u\in\mathcal{L}(\mathfrak{H}_0)$ is the operator defined in Theorem 2, then it readily follows that $K\in\mathcal{S}(Q_u\oplus 0)^-$, and $K_u=Q_u\oplus 0\in\mathcal{J}_\alpha$ is the solution to our problem. It α satisfies (ii), write $\mathfrak{H}=\bigoplus_{n=1}^{\infty}\mathfrak{H}_n$, where $\dim\mathfrak{H}_n=\alpha_n<\alpha$ and $\sum_{n=1}^{\infty}\alpha_n=\alpha$, and define $K_u=[\bigoplus_{n=1}^{\infty}(1/n)Q_{u\alpha_n}]$, where $Q_{u\alpha_n}$ is the universal quasinilpotent of Theorem 2 in dimension α_n. Clearly, K_u is a quasinilpotent element of $\mathcal{J}_\alpha$. Now the arguments of the proof of Theorem 2 and the results of [11] show that $\mathcal{S}(K_u)^-$ actually contains every nilpotent operator of $\mathcal{J}_\beta$ for every cardinal $\beta<\alpha$, and Lemma 3 and its proof show that $\mathcal{S}(K_u)^-$ also contains every nilpotent of $\mathcal{J}_\alpha$.

Let α be an $\aleph_0$-regular limit cardinal. Then, $\mathcal{J}_\alpha=\{T\in\mathcal{L}(\mathfrak{H})\colon \dim(T\mathfrak{H})^-<\alpha\}$ and, given $K\in\mathcal{J}_\alpha$, there exists a cardinal $\beta<\alpha$ such that $\dim(K\mathfrak{H})^-<\beta$ ([4]). Hence, $\mathcal{S}(K)^-\subset\mathcal{J}_\beta$, and this ideal is properly contained in $\mathcal{J}_\alpha$. Thus, if $T\in\mathcal{J}_\alpha\setminus\mathcal{J}_\beta$ and $A\in\mathcal{J}_\alpha$ is unitarily equivalent to

$$\begin{bmatrix}0 & T\\ 0 & 0\end{bmatrix},$$

then $A^2=0$, and A cannot belong to $\mathcal{S}(K)^-$. Therefore, there is no universal quasinilpotent operator for $\mathcal{J}_\alpha$. □

References

[1] C. Apostol, C. Foiaş and D. Voiculescu, On the norm-closure of nilpotents, *Rev. Roum. Math. Pures Appl.*, **19** (1974), 549—577.

[2] C. Apostol, C. Foiaş and D. Voiculescu, Some results on non-quasitriangular operators. IV, *Rev. Roum. Math. Pures Appl.*, **18** (1973), 487—514.

[3] L. A. Coburn and A. Lebow, Components of invertible elements in quotient algebras of operators, *Trans. Amer. Math. Soc.*, **130** (1968), 359—366.

[4] G. Edgar, J. Ernest and S. G. Lee, Weighing operator spectra, *Indiana Univ. Math. J.*, **21** (1971), 61—80.

[5] L. A. Fialkow, A note on direct sums of quasinilpotent operators, *Proc. Amer. Math. Soc.*, **47** (1975), 125—131.

[6] B. Gramsch, Eine Idealstruktur Banachscher Operatoralgebren, *J. reine angew. Math.* **225** (1967), 97—115.

[7] D. W. Hadwin, Closures of unitary equivalence classes, *Notices Amer. Math. Soc.*, **21** (1974), 74T—B55, A—306.

[8] P. R. Halmos, Ten problems in Hilbert space, *Bull. Amer. Math. Soc.*, **76** (1970), 887—933.

[9] D. A. Herrero, Normal limits of nilpotent operators, *Indiana Univ. Math. J.*, **23** (1974), 1097—1108.

[10] D. A. Herrero, Toward a spectral characterization of the set of norm limits of nilpotent operators, *Indiana Univ. Math. J.*, **24** (1975), 847—864.

[11] D. A. Herrero, Norm limits of nilpotent operators and weighted spectra in non-separable Hilbert spaces, *Rev. Un. Mat. Argentina* **27** (1975), 83—105.

[12] D. A. Herrero, Clausura de las órbitas de similaridad de operadores en espacios de Hilbert, *Rev. Un. Mat. Argentina* (to appear).

[13] D. A. Herrero, Closure of similarity orbits of Hilbert space operators. II: Normal operators, *J. London Math. Soc.* (2), **13** (1976), 299—316.

[14] E. Luft, The two-sided closed ideals of the algebra of bounded linear operators of a Hilbert space, *Czechoslovak Math. J.*, **18** (1968), 595—605.

[15] H. Radjavi and P. Rosenthal, *Invariant subspaces*, Ergebnisse der Math. und ihrer Grenzgebiete, 77, Springer-Verlag (New York—Heidelberg—Berlin, 1973).

[16] G. C. Rota, On models for linear operators, *Comm. Pure Appl. Math.*, **13** (1960), 469—472.

[17] N. Salinas, Subnormal limits of nilpotent operators, *Acta Sci. Math.*, **37** (1970), 117—124.

[18] D. Voiculescu, Norm-limits of algebraic operators, *Rev. Roum. Math. Pures Appl.*, **19** (1974), 371—378.

UNIVERSIDAD SIMÓN BOLÍVAR
DEPARTAMENTO DE MATEMÁTICAS
APARTADO POSTAL 80659
CARACAS 108, VENEZUELA

Acta Sci. Math., 38 (1976), 301—303

A remark on convergence systems in measure

I. JOÓ

1. Preliminaries

Denote by $S=S(0, 1)$ the set of Lebesgue measurable almost everywhere finite functions on the interval (0, 1) with the complete metrizable topology of convergence in measure. In this paper "lim" will mean "limit in measure", unless stated otherwise explicitly.

Let $T=\|t_{i,j}\|$ be a matrix, not necessarily with rows of finite length, such that

$$(1)\qquad |t_{i,j}| \leqq K \quad (i, j = 1, 2, \ldots), \quad \lim_{i\to\infty} t_{i,j} = 1 \quad (j = 1, 2, \ldots).$$

Here and in the sequel K will denote some absolute constant not necessarily the same at each occurrence.

Finally, let B be a Banach space of sequences $a=\{a_1, a_2, \ldots\}$ of real numbers such that for $a\in B$ we have

$$(2)\qquad \begin{gathered} |a_i| \leqq \|a\|_B, \quad a(N_1, N_2) = \{0, \ldots, 0, a_{N_1}, a_{N_1+1}, \ldots, a_{N_2}, 0, 0, \ldots\}\in B, \\ \lim_{N_1\to\infty} \|a(N_1, N_2)\|_B = 0; \end{gathered}$$

furthermore, if ε_j^i is a bounded double sequence of reals $(i, j=1, 2, \ldots)$ such that $\lim_{i\to\infty} \varepsilon_j^i=1$ $(j=1, 2, \ldots)$ then we have

$$(3)\qquad a^i = \{a_1\varepsilon_1^i, a_2\varepsilon_2^i, \ldots\}\in B, \quad \lim_{i\to\infty} \|a^i-a\|_B = 0 \quad \text{for all} \quad a\in B.$$

For example l_p is such a space for $1\leqq p<\infty$.

The sequence $\{f_n\}\subset S$ is called a *T convergence system in measure for B* if the limit

$$(4)\qquad \hat{T}(a) = \lim_{i\to\infty} \lim_{N\to\infty} \tau_i^N(a) \quad \text{of} \quad \tau_i^N(a) = \sum_{j=1}^{N} t_{i,j}\cdot a_j f_j$$

Received October 16, 1975.

exists for all $a \in B$. In the special case, when $t_{i,j}=0$ for $i<j$ and $=1$ otherwise, $\{f_n\}$ is simply called a *convergence system in measure for B.*

Furthermore, the sequence $\{f_n\} \subset S$ is said to be *almost orthonormal* on the interval (0, 1) if for every $\varepsilon>0$ there exist a Lebesgue measurable set $E_\varepsilon \subset (0, 1)$, a constant M_ε depending only on ε, and an orthonormal system $\{\psi_n(\varepsilon, x)\}$ on (0, 1) such that mes $E_\varepsilon \geqq 1-\varepsilon$ and

$$f_n(x) = M_\varepsilon \psi_n(\varepsilon, x) \quad (x \in E_\varepsilon,\ n = 1, 2, \ldots).$$

It is obvious that an almost orthonormal system is a convergence system in measure for l_2. In [2] Nikišin proved the converse statement.

In [3] Tandori proved the following generalization of Nikišin's result: If $\{f_n\} \subset S$ is a $(C, 1)$ convergence system in measure for l_2, that is if

$$\sum_{j=1}^{i} \left(1-\frac{j-1}{i}\right) a_j f_j = \sum_{j=1}^{i} t_{i,j} a_j f_j$$

converges in measure on (0, 1) as $n \to \infty$, for every $a \in l_2$, then $\{f_n\}$ is almost orthonormal.

Later on Tandori [4] generalized this statement even to any summation method generated by a matrix $\|t_{i,j}\|$ having rows of finite lengths and satisfying conditions in (1).

In this paper we prove a theorem by which Tandori's general result follows from Nikišin's. Namely, in section 2 we are going to prove:

Theorem. *Under conditions* (1), (2), (3) *the system* $\{f_n\}$ $(\subset S)$ *is a T convergence system in measure for B if and only if it is a convergence system in measure for B.*

2.

We need the following Banach—Steinhaus type result.

Lemma. (See, e.g. [1] p. 52.) *Let E be a Banach space, F a metrizable topological vector space, and L_n continuous linear operators on E with values in F, converging at all points of E. Then the limit operator is also continuous and linear.*

In proving the Theorem first suppose $\{f_n\} \subset S$ is a T convergence system in measure for B. Apply the Lemma twice, first for fixed i to the sequence $\{\tau_i^N\}$ of operators in (4), which are continuous on account (3). Denoting by τ_i the limit operators and applying the Lemma to this sequence we obtain that the linear operator $\hat{T}$ in (4) is continuous. Let $a \in B$ be arbitrary. We have to prove the existence of the limit

$$\lim_{N\to\infty} \sum_{j=1}^{N} a_j f_j.$$

According to the completeness of S it is enough to prove that

$$\lim_{N_1, N_2 \to \infty} \sum_{j=N_1}^{N_2} a_j f_j = 0.$$

But this follows from the continuity of $\hat{T}$ at the zero element of B, using

$$\hat{T}(a(N_1, N_2)) = \sum_{j=N_1}^{N_2} a_j f_j.$$

Conversely, suppose $\{f\}$ is a convergence system in measure for B. Then we obtain similarly that the linear operator

$$L(a) = \lim_{N \to \infty} \sum_{j=1}^{N} a_j f_j$$

is continuous. Using (3) for $\varepsilon_j^i = t_{i,j}$, this shows that

$$\lim_{i \to \infty} L(a^i) = L(\lim_{i \to \infty} a^i) = L(a).$$

This completes the proof.

References

[1] N. DUNFORD and J. T. SCHWARZ, *Linear operators*. I, Interscience (New York, 1958).

[2] E. M. Никищин, О системах сходимости по мере для l_2, *Матем. заметки*, **13** (1973), 337—340.

[3] K. TANDORI, Remark on a theorem of Alexits and Sharma, *Analysis Math.*, **1** (1975), 223—230.

[4] K. TANDORI, Über die Summierbarkeit der Funktionenreihen, *Periodica Math.*, **7** (1976), 161—168.

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1.
6720 SZEGED, HUNGARY

Acta Sci. Math., 38 (1976), 305—309

A remark on convergence of orthogonal series

I. JOÓ and K. TANDORI

1. Let S denote the set of Lebesgue measurable, almost everywhere finite functions on the interval (0, 1). Let $T=\|t_{i,j}\|_0^\infty$ be a matrix such that

(1) $$|t_{i,j}| \leqq K(<\infty) \quad (i, j=0, 1, \ldots), \quad \lim_{i\to\infty} t_{i,j}=1 \quad (j=0, 1, \ldots),$$

and let $f=\{f_k(x)\}_0^\infty$ be a sequence of functions belonging to S. A series

(2) $$\sum_{k=0}^{\infty} c_k f_k(x)$$

is said to be T summable in measure (almost everywhere) if the series

$$t_i(x)=\sum_{k=0}^{\infty} t_{i,k} c_k f_k(x) \quad (i=0, 1, \ldots)$$

converge in measure (almost everywhere) and the sequence $\{t_i(x)\}_0^\infty$ converges in measure (almost everywhere) to a function belonging to S.

The system f is said to be a T convergence system in measure (T convergence system) for l_2 if for every $c=\{c_k\}_0^\infty \in l_2$ the series (2) is T summable in measure (T summable almost everywhere).

The system f is said to be a convergence system in measure (almost everywhere) for l_2 if $c\in l^2$ implies the convergence of the series (2) in measure (almost everywhere).

Joó [3] proved a general theorem which contains the following statement as a special case:

Let T be a matrix satisfying conditions (1). *If the system f is a T convergence system in measure for l_2, then it is also a convergence system in measure for l_2.*

2. A natural question is whether a similar statement is true for almost everywhere convergence.

In this note we give a negative answer to this question.

Received January 3, 1976.

Let $\nu=\{\nu_n\}_0^\infty$ be a strictly increasing sequence of non-negative integers, $\nu_0=0$. We call T_ν the summation process generated by a matrix $\|t_{i,k}\|$ of the form

$$t_{ik}=1 \quad (k=0,1,\ldots,\nu_i), \quad t_{i,k}=0 \quad (k=\nu_i+1,\nu_i+2,\ldots) \quad (i=0,1,\ldots).$$

The T summation is said to be equivalent to T_ν summation if for every $c\in l_2$ and for every orthonormal system $\varphi=\{\varphi_k(x)\}_0^\infty$ on (0, 1) the orthogonal series

$$\sum_{k=0}^{\infty} c_k\varphi_k(x) \tag{3}$$

is T summable almost everywhere if and only if it is T_ν summable almost everywhere. (We recall the fact that, e.g., $(C, 1)$ summability is equivalent to $T_{\{2^\nu\}}$ summability; see e.g. ALEXITS [1], p. 118.)

After this preparation, our statement is:

Theorem. *Let ν be a sequence of indices such that $\overline{\lim\limits_{n\to\infty}}(\nu_{n+1}-\nu_n)=\infty$. Let T be a summation process equivalent to T_ν. Then there exists an orthonormal system $\Phi=\{\varphi_k(x)\}_0^1$ on (0, 1), which is a T convergence system for l_2 but is not a convergence system for l_2, indeed there exists a sequence $c\in l_2$ such that the series (3) diverges almost everywhere.*

We remark that the system Φ in our Theorem is obtained by a rearrangement of the Walsh system $\{w_n(x)\}_0^\infty$. Using ideas of F. MÓRICZ [4] it is easy to see that one can obtain an orthonormal system, with similar properties, also by rearrangement of the trigonometrical system $\{1, \cos 2\pi x, \sin 2\pi x, \ldots\}$.

3. The proof of the Theorem. Let $r_n(x)=\operatorname{sign}\sin 2^n\pi x$ be the n^{th} Rademacher function ($n=0,1,\ldots$). The Walsh functions are defined as follows. Let $w_0(x)=r_0(x)$. If n is a natural number and $n=2^{k_1}+\ldots+2^{k_m}$ ($0\leqq k_1<\ldots<k_m$; k_i integers) is its diadic expansion then define

$$w_n(x)=r_{k_1+1}(x)\ldots r_{k_m+1}(x).$$

We shall use a Theorem of BILLARD [2] which states that the Walsh system is a convergence system for l_2. We also need the following lemma which is proved essentially in TANDORI [5].

Lemma. *Let $m\geqq 2$ be an arbitrary natural number. Then there exists a sum of the form*

$$S_m(x)=\sum_{k=1}^{l(m)} a_k(m)w_k(x) \quad (l(m)<\mu(m+1)),$$

where $\mu(m)=2^{2^m}$, such that

$$\int_0^1 S_m^2(x)\,dx\leqq\frac{5}{m}, \tag{4}$$

furthermore, it has a rearrangement

$$S_m^*(x) = \sum_{l=1}^{l(m)} a_{k_l(m)}(m)\, w_{k_l(m)}(x)$$

such that

(5) $$\max_{1 \leqq j \leqq l(m)} \left| \sum_{l=0}^{j} a_{k_l(m)}(m)\, w_{k_l(m)}(x) \right| \geqq 1 \quad (x \in (0, 1/4) \setminus D),$$

where D denotes the set of diadic numbers.

Consider the sum

$$\sigma_m(x) = r_{2^{m+1}+1}(x) \cdot S_m(x).$$

According to the definition of Walsh functions, $\sigma_m(x)$ has the form

$$\sigma_m(x) = \sum_{k=\mu(m+1)+1}^{\mu(m+1)+l(m)} b_k(m)\, w_k(x) \quad (l(m) < \mu(m+1)).$$

Our Lemma shows that

(6) $$\int_0^1 \sigma_m^2(x)\, dx \leqq \frac{5}{m},$$

furthermore, $\sigma_m(x)$ has a rearrangement

$$\sigma_m^*(x) = \sum_{l=1}^{l(m)} b_{k_l(m)}(m)\, w_{k_l(m)}(x),$$

such that

(7) $$\max_{1 \leqq j \leqq l(m)} \left| \sum_{l=1}^{j} b_{k_l(m)}(m)\, w_{k_l(m)}(x) \right| \geqq 1 \quad (x \in (0, 1/4) \setminus D).$$

Now we define the system Φ in our Theorem. First let $\{n_m\}_2^\infty$ be a strictly increasing sequence of indices such that

$$\nu_{n_m+1} - \nu_{n_m} \geqq \mu(m^2+1) \quad (m = 2, 3, \ldots);$$

such a sequence exists according to our assumption concerning ν. For all $m (\geqq 2)$ consider the sum $\sigma_{m^2}(x)$. It is obvious by the definition that in the case $m \neq \overline{m}$ the same Walsh functions do not occur in both $\sigma_{m^2}(x)$ and $\sigma_{\overline{m}^2}(x)$ with coefficients different from zero. Further it is easy to see that the sum $\sigma_{2^2}(x), \ldots, \sigma_{m^2}(x)$ are built from Walsh functions $w_1(x), \ldots, w_{2\mu(m^2+1)-1}(x)$.

Consider the rearrangement of the sum $\sigma_{m^2}(x)$:

$$\sigma_{m^2}^*(x) = \sum_{l=1}^{l(m^2)} b_{k_l(m^2)}(m^2)\, w_{k_l(m^2)}(x).$$

Let

$$\varphi_{\nu_{n_m}+l}(x) = w_{k_l(m^2)}(x) \quad (l = 1, \ldots, l(m^2)).$$

Let

$$\Omega_1 = \bigcup_{m=2}^{\infty} \{k_l(m^2) : l = 1, \ldots, l(m^2)\}, \quad \Omega_2 = \{0, 1, \ldots\} \setminus \Omega_1,$$

and denote the elements of Ω_2 in the order of magnitude by $q_1, q_2, \dots$. At last let $r_1, r_2, \dots$ be those indices, in order of magnitude, for which the function $\varphi_k(x)$ are not yet defined. Next let

$$\varphi_{r_i}(x) = w_{q_i}(x) \quad (i = 1, 2, \dots).$$

So we have defined an orthonormal system $\Phi = \{\varphi_k(x)\}_0^\infty$ on (0, 1), which is a rearrangement of the Walsh system $\{w_k(x)\}_0^\infty$.

Let $c = \{c_k\}_0^\infty \in l_2$ be arbitrary. According to the definition of the functions $\varphi_k(x)$,

$$\text{(8)} \quad \begin{aligned} \sum_{k=0}^{\infty} c_k \varphi_k(x) &= \sum_{i=1}^{\infty} c_{r_i} \varphi_{r_i}(x) + \sum_{m=2}^{\infty} \sum_{j=\nu_{n_m}+1}^{\nu_{n_m}+l(m^2)} c_j \varphi_j(x) = \\ &= \sum_{i=1}^{\infty} c_{r_i} w_{q_i}(x) + \sum_{m=2}^{\infty} \sum_{l=1}^{l(m^2)} c_{\nu_{n_m}+l} w_{k_l(m^2)}(x) = \sum\nolimits_1 + \sum\nolimits_2. \end{aligned}$$

The sum $\sum_1$ is a Walsh expansion in l_2 thus, according to Billard's theorem, it converges almost everywhere on (0, 1).

On the other hand, for all m

$$\sum_{l=1}^{l(m^2)} c_{\nu_{n_m}+l} w_{k_l(m^2)}(x) = \sum_{l=\mu(m^2+1)+1}^{\mu((m+1)^2+1)} \bar{c}_l w_l(x),$$

where

$$\sum_{l=1}^{l(m^2)} c^2_{\nu_{n_m}+l} = \sum_{l=\mu(m^2+1)+1}^{\mu((m+1)^2+1)} \bar{c}_l^2.$$

Now set

$$d_k = c_{\nu^2_{n_m}+j} \quad \text{for} \quad k = \nu_{n_m}+j; \quad j = 1, \dots, l(m^2), \quad \text{and} \quad d_k = 0 \quad \text{otherwise}.$$

Obviously, $d = \{d_k\}_0^\infty \in l_2$ and the ν_n^{th} partial sum of the series

$$\sum\nolimits_2 = \sum_{k=0}^{\infty} d_k \varphi_k(x)$$

is equal to the $\mu(m^2+1)+l(m^2)^{\text{th}}$ partial sum of the series

$$\sum_{k=0}^{\infty} \bar{c}_k w_k(x)$$

for some m. Apply Billard's theorem again to obtain that the sequence of the ν_n^{th} partial sums of the series $\sum_2$ converges almost everywhere. Using (8) we obtain thath the sequence of the ν_n^{th} partial sums of the series (3) also converges almost everywhere. This shows that the system Φ is a T convergence system for l_2. (We use our assumption for T that it is equivalent to T_ν.)

On the other hand, consider the series

$$\text{(9)} \quad \sum_{m=2}^{\infty} \sigma^2_{m^2}\left(x - \frac{m}{4}\right).$$

From the definition of the system Φ and from (6) it follows that (9) is an l_2-expansion in Φ:

$$\sum_{m=2}^{\infty} \sigma_{m^2}^{*}\left(x-\frac{m}{4}\right) = \sum_{k=0}^{\infty} a_k \varphi_k(x) \quad (\{a_k\}_0^\infty \in l_2).$$

But it is clear from (7) that this series diverges almost everywhere on (0, 1). So our Theorem is proved.

References

[1] G. Alexits, *Convergence Problems of Orthogonal Series*, Akadémiai Kiadó (Budapest, 1961).

[2] P. Billard, Sur la convergence presque partout des séries de Fourier—Walsh des functions de l'espace $L^2(0, 1)$, *Studia Math.*, **28** (1967), 363—388.

[3] I. Joó, A remark on convergence systems in measure *(to appear)*.

[4] F. Móricz, On the order of magnitude of the partial sums of rearranged Fourier series of square integrable functions, *Acta Sci. Math.*, **28** (1967), 155—167.

[5] K. Tandori, Über die Divergenz der Walshschen Reihen, *Acta Sci. Math.*, **27** (1966), 261—263.

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1.
6720 SZEGED, HUNGARY

Acta Sci. Math., 38 (1976), 311—315

Commutants and bicommutants of operators of class C_0

BÉLA SZ.-NAGY and CIPRIAN FOIAŞ

Dedicated to P. R. Halmos on his 60th birthday

Introduction

By operator we mean a linear and bounded one. For any operator T on a Hilbert space $\mathfrak{H}$ we consider the following weakly (or equivalently, strongly) closed subalgebras of $\mathscr{B}(\mathfrak{H})$:

$\mathscr{A}_T$: the subalgebra generated by I and T;

$\{T\}'$: the commutant of T;

$\{T\}''$: the bicommutant of T;

$\mathscr{L}_T$: the subalgebra consisting of those $X\in\{T\}'$ for which $\mathrm{Lat}\, X\supset \mathrm{Lat}\ T$ (i.e. X leaves invariant every subspace of $\mathfrak{H}$ invariant for T).

If T is a completely non-unitary contraction on $\mathfrak{H}$ we also define:

$\mathscr{N}_T$: the set of operators on $\mathfrak{H}$ which admit a representation $X=v(T)^{-1}u(T)$ with functions $u, v\in H^\infty$ such that $v(T)$ is a quasi-affinity (i.e. an operator with zero kernel and dense range).

From this definition it readily follows:

$$\mathscr{N}_T\subset\{T\}'', \quad cf. \text{ [H], Chapter IV.} \tag{0}$$

We shall consider operators T of class C_0, i.e. completely non-unitary contractions such that $w(T)=0$ for some inner function w; among these functions w there is a minimal one, denoted by m_T. For $T\in C_0$ and $v\in H^\infty$ the operator $v(T)$ is a quasi-affinity if and only if $v\wedge m_T=1$ (i.e. if v and m_T have no non-constant inner divisor); *cf.* [H], Proposition III. 4.7.

For $T\in C_0$ we have equality in (0), i.e.

$$\mathscr{N}_T = \{T\}'' \quad \text{for} \quad T\in C_0. \tag{1}$$

Received October 30, 1975.

This was proved in [2] if the underlying space $\mathfrak{H}$ is separable, by using the "Jordan model" of operators of class C_0. A subsequent extension of the Jordan model to the non-separable case, given in [3], yields, by the same proof, the validity of (1) for non-separable $\mathfrak{H}$ also.

In Sections 1 and 2 of the present paper we shall prove the inclusions

(2) $$\mathcal{N}_T \subset \mathcal{A}_T \quad \text{for} \quad T \in C_0,$$

(3) $$\mathcal{L}_T \subset \mathcal{N}_T \quad \text{for} \quad T \in C_0.$$

As a consequence of (1), (2), (3), and of the trivial inclusion $\mathcal{A}_T \subset \mathcal{L}_T$ we deduce

$$\{T\}'' = \mathcal{N}_T \subset \mathcal{A}_T \subset \mathcal{L}_T \subset \mathcal{N}_T = \{T\}'' \quad \text{for} \quad T \in C_0.$$

So we establish the following:

Theorem. *For any operator T of class C_0 we have*

$$\mathcal{A}_T = \mathcal{L}_T = \{T\}'' = \mathcal{N}_T.$$

For operators T of class C_0 with *finite* defect indices (classes $C_0(N)$; $N=1, 2, \ldots$) these results were proved in the recent paper [4] by Wu (Theorems 3.2 and 3.3). It was this paper that suggested the present investigation. The proofs we are going to give for the general case employ quite different arguments as those in [4].

1. Proof of $\mathcal{N}_T \subset \mathcal{A}_T$

Let $T \in C_0$ on $\mathfrak{H}$. Suppose there is an $X \in \mathcal{N}_T$ which is not contained in $\mathcal{A}_T$. This means that there exist $h_1, \ldots, h_r \in \mathfrak{H}$ and $\varepsilon > 0$ such that

(1.1) $$\sum_{j=1}^{r} \|Xh_j - p(T)h_j\|^2 \geqq \varepsilon^2 \quad \text{for all polynomials } p.$$

Setting

$$\mathbf{H} = \bigoplus_1^r \mathfrak{H}, \quad \mathbf{T} = \bigoplus_1^r T, \quad \mathbf{X} = \bigoplus_1^r X, \quad \mathbf{h} = \bigoplus_1^r h_j,$$

(1.1) can also be expressed as

(1.2) $$\|\mathbf{X}\mathbf{h} - p(\mathbf{T})\mathbf{h}\| \geqq \varepsilon \quad \text{for all polynomials } p.$$

As $X \in \mathcal{N}_T$ there exist $u, v \in H^\infty$ such that $v \wedge m_T = 1$, $v(T)X = u(T)$, and hence,

(1.3) $$v(\mathbf{T})\mathbf{X} = u(\mathbf{T}).$$

Denote by $\mathbf{H_h}$ the cyclic subspace for $\mathbf{T}$ generated by $\mathbf{h}$ and define

(1.4) $$\mathbf{K} = \{\mathbf{k} \in \mathbf{H} : v(\mathbf{T})\mathbf{k} \in \mathbf{H_h}\}.$$

Clearly, $\mathbf{K}$ is invariant for $\mathbf{T}$ and $\mathbf{T}_0=\mathbf{T}|\mathbf{K}$ is of class C_0. Its minimal function is a divisor of $m_{\mathbf{T}}$ $(=m_T)$ so we also have $v\wedge m_{\mathbf{T}_0}=1$. Thus, $v(\mathbf{T}_0)$ is a quasi-affinity on $\mathbf{K}$ and so it has dense range in $\mathbf{K}$. As by definition (1.4)

$$v(\mathbf{T}_0)\mathbf{K} = v(\mathbf{T})\mathbf{K}\subset \mathbf{H}_{\mathbf{h}}$$

we infer that

$$\mathbf{K}\subset \mathbf{H}_{\mathbf{h}}. \tag{1.5}$$

Now, by (1.3) we have $v(\mathbf{T})\mathbf{X}\mathbf{h}=u(\mathbf{T})\mathbf{h}\in \mathbf{H}_{\mathbf{h}}$, and therefore $\mathbf{X}\mathbf{h}\in\mathbf{K}$; thus, by (1.5),

$$\mathbf{X}\mathbf{h}\in \mathbf{H}_{\mathbf{h}}.$$

This implies that there is a polynomial p such that

$$\|\mathbf{X}\mathbf{h}-p(\mathbf{T})\mathbf{h}\| < \varepsilon.$$

This contradicts (1.2), and hence achieves the proof.

2. Proof of $\mathscr{L}_T\subset\mathscr{N}_T$

Let $T\in C_0$ on $\mathfrak{H}$. By [2], Proposition 2, we have

$$T \succ S(m)\oplus G,$$

where $m=m_T$ and G is the restriction of T to some invariant subspace $\mathfrak{G}$, i.e. there exists a quasi-affinity

$$A: \mathfrak{H}(m)\oplus \mathfrak{G} \to \mathfrak{H}$$

such that

$$TA = A(S(m)\oplus G). \tag{2.1}$$

Here, as usual, $S(m)$ denotes the compression of the canonical shift on H^2 to the subspace $\mathfrak{H}(m)=H^2\ominus mH^2$.

Consider the cyclic vector e for $S(m)$, given by $e=1-\overline{m(0)}m$, and an arbitrarily chosen vector $g\in\mathfrak{G}$, and set

$$h_t = A((1-t)e\oplus tg), \tag{2.2}$$

t being a numerical parameter to be fixed later. Further, set

$$\mathfrak{H}_t = \bigvee_{n\geqq 0} T^n h_t, \quad T_t = T|\mathfrak{H}_t, \quad \text{and} \quad m_t = m_{T_t}.$$

From (2.1) and (2.2) we deduce

$$w(T)h_0 = A(w(S(m))e\oplus 0) \quad \text{for all} \quad w\in H^\infty;$$

hence T_0 has the same minimal function as $S(m)$, i.e. $m_0=m=m_T$.

While it may happen that m_1 is a proper inner divisor of m_T, it follows from a lemma due to M. SHERMAN that the values t for which m_t is a proper divisor of m_T are exceptional, that is, countable many; *cf.* [1]. Let τ be a non-exceptional value of t, different from 0 and 1; thus $m_\tau = m_T$, $0 \neq \tau \neq 1$.

Let $X \in \mathscr{L}_T$. Then $X\mathfrak{H}_t \subset \mathfrak{H}_t$ and $X|\mathfrak{H}_t \in \{T_t\}'$, for all t. Since T_t is a C_0 class operator with cyclic vector h_t, every operator in its commutant is a function of T_t of the "Nevanlinna class" $\mathscr{N}_{T_t}$ (*cf.* [H], Chapter IV, and [1], Théorème 2). Thus there exist functions $u_t, v_t \in H^\infty$ such that

(2.3) $$v_t \wedge m_t = 1 \quad \text{and} \quad v_t(T)Xh_t = u_t(T)h_t;$$

in particular,

(2.4) $$v_0 \wedge m_T = 1, \quad v_\tau \wedge m_T = 1.$$

Set

(2.5) $$X' = v_0(T)X - u_0(T);$$

X' also belongs to $\mathscr{L}_T$ and by (2.3) we have

(2.6) $$X'h_0 = 0 \quad \text{and} \quad v_t(T)X'h_t = u_t'(T)h_t \quad \text{for} \quad u_t' = v_0 u_t - u_0 v_t.$$

Hence, $X'h_t = X'((1-t)h_0 + th_1) = tX'h_1$ and

$$v_t(T)v_1(T)X'h_t = \begin{cases} v_t(T)\cdot tu_1'(T)h_1 = t(v_t u_1')(T)h_1 \\ v_1(T)v_t(T)X'h_t = v_1(T)u_t'(T)h_t = (v_1u_t')(T)((1-t)h_0 + th_1) \end{cases}$$

so we have

$$(1-t)(v_1u_t')(T)h_0 = t(v_tu_1' - v_1u_t')(T)h_1.$$

By (2.1) and (2.2), and since A is injective, this implies

$$(1-t)\cdot(v_1u_t')(S(m))e \oplus 0 = 0 \oplus t\cdot(v_tu_1' - v_1u_t')(G)g;$$

so we have for any $t \neq 0$, 1, and in particular for $t = \tau$:

(2.7) $$(v_1u_\tau')(S(m))e = 0, \quad (v_\tau u_1' - v_1u_\tau')(G)g = 0.$$

The first equation (2.7) implies $v_1u_\tau' \in mH^\infty$. Since $m_G | m$ we infer $(v_1u_\tau')(G) = 0$. Comparing this with the second equation (2.7) we deduce

$$v_\tau(G)u_1'(G)g = (v_\tau u_1')(G)g = 0.$$

On account of (2.4), $v_\tau(T)$ is a quasi-affinity so its restriction $v_\tau(G)$ is injective; thus $u_1'(G)g = 0$. Hence,

$$u_1'(T)h_1 = u_1'(T)A(0\oplus g) = Au_1'(S(m)\oplus G)(0\oplus g) = A(0\oplus u_1(G)g) = 0,$$

and therefore, by (2.6), $v_1(T)X'h_1 = 0$. Now, the subspace $\mathfrak{H}_1$ being invariant for T is also invariant for X'; thus $X'h_1 \in \mathfrak{H}_1$. But $v_1 \wedge m_1 = 1$ by (2.3), and thus $v_1(T_1) = v_1(T)|\mathfrak{H}_1$ is a quasi-affinity on $\mathfrak{H}_1$, and in particular injective, so we conclude $X'h_1 = 0$.

Combining this result with the equation $X'h_0=0$, see (2.6), and recalling that by its definition (2.5) the operator X' is independent of the choice of g in $\mathfrak{G}$ we readily conclude that $X'A=0$, $X'=0$, and therefore

$$v_0(T)X-u_0(T)=0 \quad (v_0 \wedge m_T=1),$$

that is, $X \in \mathcal{N}_T$.

This concludes the proof.

References

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

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

[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, On the Jordan model of C_0 operators, *Studia Math.* (to appear).

[4] P. Y. Wu, Commutants of $C_0(N)$ contractions, *Acta Sci. Math.*, **38** (1976), 193—202.

B. SZ.-NAGY
BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

C. FOIAŞ
FACULTY OF MATH., UNIV. BUCHAREST
STR. ACADEMIEI 14
BUCHAREST, ROMANIA

Acta Sci. Math., 38 (1976), 317—324

On the strong approximation of Fourier series

L. LEINDLER

1. Let $f(x)$ be a continuous and 2π-periodic function and let

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{1.1}$$

be its Fourier series. Let $s_n(x) = s_n(f; x)$ and $\sigma_n^\alpha(x) = \sigma_n^\alpha(f; x)$ denote the n-th partial sum and the (C, α)-mean of (1.1), and let $\tilde{f}(x)$, $\tilde{s}_n(x)$, $\tilde{\sigma}_n^\alpha(x)$ denote the conjugate functions, respectively.

In [2] we investigated among others the means

$$V_n(f, \lambda, p; x) = \left\{ \frac{1}{\lambda_n} \sum_{k=n-\lambda_n}^{n-1} |s_k(x) - f(x)|^p \right\}^{1/p},$$

where $\lambda = \{\lambda_n\}$ is a nondecreasing sequence of integers such that $\lambda_1 = 1$ and $\lambda_{n+1} - \lambda_n \leqq 1$, and $p > 0$. Such a mean is called a "generalized strong de la Vallée Poussion mean", or briefly, a *strong (V, λ)-mean.*

In [2] we proved the following theorems:

Theorem A. *If $n = O(\lambda_n)$ and $p > 0$, then*

$$V_n(f, \lambda, p; x) = O(E_{n-\lambda_n}) \tag{1.2}$$

polds uniformly, where $E_n = E_n(f)$ denotes the best approximation of f by trigonometric holynomials of order at most n.

Theorem B. *Suppose that $f(x)$ r times derivable and $f^{(r)} \in Lip\, \alpha$ $(0 < \alpha \leqq 1)$, and that $n = 0(\lambda_n)$. Then for any $p > 0$*

$$V_n(f, \lambda, p; x) = \begin{cases} O\left(\dfrac{1}{n^{r+\alpha}}\right) & \text{for} \quad (r+\alpha)p < 1, \\ O\left(\dfrac{1}{n^{r+\alpha}}\left(1 + \log \dfrac{n}{n - \lambda_n + 1}\right)^{1/p}\right) & \text{for} \quad (r+\alpha)p = 1, \end{cases} \tag{1.3}$$

Received January 13, 1976.

uniformly. The same estimate holds for $V_n(\tilde{f}, \lambda, p; x)$. *Furthermore, if* $(r+\alpha)p=1$ $(0<\alpha<\leqq 1)$, *then there exist functions* $f_1(x)$ *and* $f_2(x)$ *such that their r-th derivatives exist and belong to* Lip α, *moreover, both*

$$\varlimsup_{n\to\infty} V_n(f_1, \lambda, p; 0) \quad and \quad \varlimsup_{n\to\infty} V_n(\tilde{f}_2; \lambda, p; 0) \quad are \quad \geqq \frac{c}{n^{r+\alpha}}\left(1+\log\frac{n}{n-\lambda_n+1}\right)^{1/p},$$

where $c(>0)$ *is independent of n.*

In this paper we generalize these results. Among others we omit the restriction $n=O(\lambda_n)$, but then the estimations will not be necessarily best possible, and show that there exists a function f_0 such that both $f_0^{(r)}$ and $\tilde{f}_0^{(r)}$ belong to the class Lip 1 and the estimations (1.3) are best possible for the means $V_n(f_0, \lambda, p; x)$ also. Furthermore we show that if $0<\alpha<1$ then the partial sums in the means $V_n(f, \lambda, p; x)$ can be replaced by (C, β)-means of negative order.

More precisely we prove the following theorems:

Theorem 1. *For any positive p we have*

$$V_n(f, \lambda, p; x) = O\left(\left(\frac{n}{\lambda_n}\right)^{1/p} E_{n-\lambda_n}\right) \tag{1.4}$$

uniformly.

Theorem 2. *If* $f^{(r)}\in$ Lip α $(0<\alpha\leqq 1)$, *then for any* $p>0$

$$V_n(f, \lambda, p; x) = \begin{cases} O\left(\left(\frac{n}{\lambda_n}\right)^{1/p}\frac{1}{n^{r+\alpha}}\right) & \text{for} \quad (r+\alpha)p<1, \\ O\left(\frac{1}{\lambda_n^{r+\alpha}}\left(1+\log\frac{n}{n-\lambda_n+1}\right)^{1/p}\right) & \text{for} \quad (r+\alpha)p=1, \\ O\left(\lambda_n^{-1/p}(n-\lambda_n+1)^{\frac{1}{p}-r-\alpha}\right) & \text{for} \quad (r+\alpha)p>1, \end{cases} \tag{1.5}$$

holds uniformly. The same estimate also holds for $V_n(\tilde{f}, \lambda, p; x)$.

Theorem 3. *Suppose that* $0<\alpha\leqq 1$, $p>0$, *and* $n=O(\lambda_n)$. *Then there exists* f_0 *such that* $f_0^{(r)}$ *and* $\tilde{f}_0^{(r)}$ *belong to the class* Lip α, *and still*

$$\varlimsup_{n\to\infty} V_n(f_0, \lambda, p; 0) \geqq \begin{cases} dn^{-r-\alpha} & \text{if} \quad (r+\alpha)p<1, \\ dn^{-r-\alpha}\left(1+\log\frac{n}{n-\lambda_n+1}\right)^{1/p} & \text{if} \quad (r+\alpha)p=1, \\ dn^{-1/p}(n-\lambda_n+1)^{1/p-r-\alpha} & \text{if} \quad (r+\alpha)p>1, \end{cases} \tag{1.6}$$

where $d=d(\lambda, p)>0$.

Theorem 4. *Suppose that $f \in \mathrm{Lip}\,\alpha$ for some $0<\alpha<1$, that $\beta>-1/2$ and that the positive number p satisfies the inequality $p\beta>-1$. Then we have, uniformly,*

$$(1.7)\qquad \left[\frac{1}{\lambda_n}\sum_{k=n-\lambda_n}^{n}|\sigma_k^\beta(x)-f(x)|^p\right]^{1/p}=\begin{cases} O\left(\left(\dfrac{n}{\lambda_n}\right)^{1/p}\dfrac{1}{n^\alpha}\right)\\ O\left(\dfrac{1}{\lambda_n^\alpha}\left(1+\log\dfrac{n}{n-\lambda_n+1}\right)^{1/p}\right)\\ O\left(\lambda_n^{-1/p}(n-\lambda_n+1)^{1/p-\alpha}\right)\end{cases}$$

according as αp is <1, $=1$, or >1.

In what follows $\|\cdot\|$ and $[\cdot]$ denote supremum norm and integral part, respectively, and $\omega(f;\delta)$ denotes the modulus of continuity of f.

Finally we improve one part of the following theorem of SZABADOS [7]:

Theorem C. *If $0<p<1$ and $r=[1/p]$, then the condition*

$$(1.8)\qquad \left\|\sum_{n=0}^{\infty}|s_n(x)-f(x)|^p\right\|\leqq K$$

implies that $f^{(r-1)}(x)$ is continuous and

$$\omega(f^{(r-1)};h)=\begin{cases} O\left(h\left(\log\dfrac{1}{h}\right)^{1/p}\right) & \text{if } \dfrac{1}{p}=r,\\ O(h) & \text{otherwise.}\end{cases}$$

We have the following

Theorem 5. *If $0<p<1$ and $1/p-r=\alpha>0$, then condition (1.8) implies that $f^{(r)}$ is continuous and*

$$(1.9)\qquad \omega(f^{(r)},h)=O\left(h^\alpha\left(\log\frac{1}{h}\right)^{1/p-1}\right).$$

In connection with these results we formulate the following

Conjecture. *) *If $0<p<1$ and $1/p=r+\alpha$, then condition (1.8) implies that*

$$(1.10)\qquad \omega(f^{(r-1)};h)=O\left(h\log\frac{1}{h}\right)\quad \text{if}\quad \alpha=0,$$

and

$$(1.11)\qquad \omega(f^{(r)};h)=O(h^\alpha)\quad \text{if}\quad \alpha>0.$$

---

*) Added in proof: This conjecture has been verified by the author.

Finally we remark that the estimations (1.10) and (1.11) are, in general, best possible. Namely, if $1/p=r+\alpha$ and r is an odd integer, then the function

$$f_0(x)=\sum_{n=1}^{\infty}\frac{\sin nx}{n^{1+1/p}}$$

has $(r-1)$-th and r-th derivatives such that if $\alpha=0$ then

$$\left|f_0^{(r-1)}\left(\frac{\pi}{2^n}\right)-f_0^{(r-1)}(0)\right|>\frac{1}{8}\frac{\pi}{2^n}\log\frac{2^n}{\pi}\quad\text{for all}\quad n\geqq 6,$$

(see [5], pp. 224—227); and since

$$f_0^{(r)}(x)=\pm\sum_{n=1}^{\infty}\frac{\cos nx}{n^{1+\alpha}}\quad(\alpha>0),$$

the inequality $\omega(f_0^{(r)},h)\geqq c\,h^{\alpha}\,(c>0)$ is obvious. Furthermore a standard computation (see e.g. [5], pp. 225—226) shows that for this function f_0 (1.8) holds.

2. To prove our theorems we require three lemmas.

Lemma 1. ([2], Lemma 2) *If $g\in L(0,2\pi)$ and $|g(x)|\leqq M$ for all x, then, for any $q>0$, we have*

$$\frac{1}{m}\sum_{k=1}^{m}|s_k(g;x)|^q\leqq C_q^qM^q.$$

Lemma 2. ([3], Lemma) *If $f\in\text{Lip}\,\gamma$, $0<\gamma<1$, $\delta>-1/2$, and if the positive number p satisfies the inequality $p\delta>-1$, then we have for any $n(\geqq 1)$*

$$\frac{1}{n}\sum_{\nu=n}^{2n}|\sigma_\nu^{\delta}(f;x)-\sigma_\nu^{\delta+1}(f;x)|^p=O(n^{-\gamma p}).$$

Lemma 3. ([2], estimate (6), p. 150) *We have for any $q>0$ and n*

$$h_n(f,q;x)\equiv\left(\frac{1}{n}\sum_{\nu=n}^{2n}|s_\nu(f,x)-f(x)|^q\right)^{1/q}=O(E_n).$$

3. Proof of Theorem 1. Let T_m^* denote the trigonometric polynomial of best approximation to f of order at most m. From the definition of s_n it is clear that if $\nu\geqq m$ then $s_\nu(f-T_m^*;x)=s_\nu(f;x)-T_m^*(x)$. Using this we have

$$\left(\frac{1}{\lambda_n}\sum_{\nu=n-\lambda_n}^{n-1}|s_\nu(x)-f(x)|^p\right)^{1/p}\leqq\left[\frac{2^p}{\lambda_n}\sum_{\nu=n-\lambda_n}^{n-1}\left(|s_\nu(f-T_{n-\lambda_n}^*;x)|^p+|T_{n-\lambda_n}^*(x)-f(x)|^p\right)\right]^{1/p}\leqq$$

$$\leqq 2^{1+1/p}\left(\left\{\frac{n}{\lambda_n}\cdot\frac{1}{n}\sum_{\nu=n-\lambda_n}^{n-1}|s_\nu(f-T_{n-\lambda_n}^*;x)|^p\right\}^{1/p}+E_{n-\lambda_n}\right).\tag{3.1}$$

Applying Lemma 1 (with $g=f-T_{n-\lambda_n}^*$ and $q=p$) we immediately obtain the statement of Theorem 1.

Proof of Theorem 2. By the well-known theorem of Jackson the assumption $f^{(r)} \in \text{Lip}\,\alpha$ $(0<\alpha\leqq 1)$ implies that

$$E_n(f) = O(n^{-r-\alpha}) \quad \text{and} \quad E_n(\tilde{f}) = O(n^{-r-\alpha}).$$

Hence, by Lemma 3, we obtain that

$$h_n(f, p; x) = O(n^{-r-\alpha}) \quad \text{and} \quad h_n(\tilde{f}, p; x) = O(n^{-r-\alpha}). \tag{3.2}$$

If $2^{m_1} \leqq n-\lambda_n < 2^{m_1+1}$ and $2^{m_2} < n \leqq 2^{m_2+1}$ then, by (3.2), we have

$$\frac{1}{\lambda_n}\sum_{\nu=n-\lambda_n}^{n-1} |s_\nu(x)-f(x)|^p \leqq \frac{1}{\lambda_n}\sum_{m=m_1}^{m_2}\sum_{\nu=2^m}^{2^{m+1}-1} |s_\nu(x)-f(x)|^p \leqq$$

$$\leqq \frac{O(1)}{\lambda_n}\sum_{m=m_1}^{m_2} 2^{m(1-p(r+\alpha))} \equiv \sum\nolimits_1.$$

Now,

$$\sum\nolimits_1 \leqq O(1)\frac{1}{\lambda_n} 2^{m_2(1-p(r+\alpha))} = O\left(\frac{n}{\lambda_n}\cdot\frac{1}{n^{p(r+\alpha)}}\right), \quad \text{if} \quad p(r+\alpha) < 1,$$

$$\sum\nolimits_1 \leqq O(1)\frac{1}{\lambda_n}(m_2-m_1) = O\left(\frac{1}{\lambda_n}\left(1+\log\frac{n}{n-\lambda_n+1}\right)\right), \quad \text{if} \quad p(r+\alpha) = 1,$$

$$\sum\nolimits_1 = O(\lambda_n^{-1}(n-\lambda_n+1)^{1-p(r+\alpha)}), \quad \text{if} \quad p(r+\alpha) > 1.$$

Whence (1.5) obviously follows.

The proof for $\tilde{f}$ runs similarly.

Proof of Theorem 3. Set

$$f_0(x) = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{2^{n\alpha}}\sum_{l=2^{n-1}+1}^{2^n}\left(\frac{\cos(5\cdot 2^n-l)x}{(5\cdot 2^n-l)^r l} - \frac{\cos(5\cdot 2^n+l)x}{(5\cdot 2^n+l)^r l}\right).$$

In [4] (Theorem 1) it is proved that $f_0^{(r)}$ and $\tilde{f}_0^{(r)}$ belong to the class $\text{Lip}\,\alpha$ if $\alpha=1$, furthermore in [1] this statement in the case $\alpha<1$ with an odd r is verified. Thus we only have to show that $f_0^{(r)} \in \text{Lip}\,\alpha$ if r is an even integer and $0<\alpha<1$. In this case

$$f_0^{(r)}(x) = \sum_{n=1}^{\infty}\frac{(-1)^{n+\frac{r}{2}+1}}{2^{n\alpha}}\sum_{l=2^{n-1}+1}^{2^n}\left(\frac{\cos(5\cdot 2^n-l)x}{l} - \frac{\cos(5\cdot 2^n+l)x}{l}\right) \equiv$$

$$\equiv \sum_{n=1}^{\infty}\frac{(-1)^{n+\frac{r}{2}+1}}{2^{n\alpha}} R_n(x),$$

where $\|R_n(x)\| \leqq 2$. Thus, if $4\cdot 2^m \leqq n < 4\cdot 2^{m+1}$, then

$$E_n(f_0^{(r)}) \leqq \|f_0^{(r)}(x) - s_n(f_0^{(r)}; x)\| \leqq 2\sum_{k=m}^{\infty}\frac{1}{2^{k\alpha}} = O\left(\frac{1}{n^\alpha}\right),$$

which implies $f_0^{(r)} \in \text{Lip}\,\alpha$ $(0<\alpha<1)$.

In the proof of (1.6) we distinguish two cases according as the sequence $\left\{\frac{n}{n-\lambda_n}\right\}$ is bounded or not. First we investigate the bounded case. Let $n=12\cdot 2^m$ and let $m_1=\max(n-\lambda_n, 22\cdot 2^{m-1})$, $m_2=\max(m_1, 23\cdot 2^{m-1})$ and $m_3=\max\left(m_2, n-\left[\frac{\lambda_n+1}{2}\right]\right)$. Then

$$V_n(f_0, \lambda, p; 0)=\left\{\frac{1}{\lambda_n}\sum_{\nu=n-\lambda_n}^{n-1}|s_\nu(0)-f_0(0)|^p\right\}^{1/p}\geqq$$

$$\geqq\left\{\frac{1}{\lambda_n}\left(\sum_{\nu=m_1}^{m_2-1}+\sum_{\nu=m_2}^{m_3}\right)\left|\frac{1}{n^\alpha}\sum_{l=\nu-10\cdot 2^m+1}^{2^{m+1}}\frac{1}{n^r l}\right|^p\right\}^{1/p}.$$

Hence, by $n=O(\lambda_n)$, it follows that

$$\sum_{\nu=m_1}^{m_2-1}\left|\frac{1}{n^\alpha}\sum_{l=\nu-10\cdot 2^m+1}^{2^{m+1}}\frac{1}{n^r l}\right|^p\geqq(m_2-m_1)\left|\frac{1}{n^\alpha}\sum_{l=m_2-10\cdot 2^m+1}^{2^{m+1}}\frac{1}{n^r l}\right|^p\geqq$$

$$\geqq(m_2-m_1)\left|\frac{1}{n^{\alpha+r+1}}(n-m_2)\right|^p\geqq d_1(p,\lambda)(m_2-m_1)\frac{1}{n^{(\alpha+r)p}},$$

and

$$\sum_{\nu=m_2}^{m_3}\left|\frac{1}{n^\alpha}\sum_{l=\nu-10\cdot 2^m+1}^{2^{m+1}}\frac{1}{n^r l}\right|^p\geqq(m_3-m_2)\left|\frac{1}{n^{\alpha+r+1}}(n-m_3)\right|^p\geqq d_2(p,\lambda)(m_3-m_2)\frac{1}{n^{(\alpha+r)p}}.$$

Thus we obtain that

$$V_n(f_0, \lambda, p; 0)\geqq d_3(p,\lambda)\left[(m_3-m_1)\frac{1}{\lambda_n}\cdot\frac{1}{n^{(\alpha+r)p}}\right]^{1/p}\geqq d_4(p,\lambda)\frac{1}{n^{r+\alpha}},$$

which proves the statements of (1.6) under the assumption that the sequence $\left\{\frac{n}{n-\lambda_n}\right\}$ is bounded.

If $\left\{\frac{n}{n-\lambda_n+1}\right\}$ is not bounded, then we may suppose that there exist infinitely many n with $4\cdot 2^m<n\leqq 4\cdot 2^{m+1}$ and $4\cdot 2^\mu\leqq n-\lambda_n+4<4\cdot 2^{\mu+1}$ such that $m>\mu+2$. Then

$$V_n(f_0, \lambda, p; 0)^p\geqq\frac{1}{\lambda_n}\sum_{i=\mu+1}^{m-1}\sum_{\nu=4\cdot 2^i+1}^{4\cdot 2^{i+1}}|s_\nu(0)-f_0(0)|^p\geqq$$

(3.3)

$$\geqq\frac{1}{\lambda_n}\sum_{i=\mu+1}^{m-1}\sum_{\nu=11\cdot 2^{i-1}+1}^{12\cdot 2^{i-1}}|s_\nu(0)-f_0(0)|^p\equiv\frac{1}{\lambda_n}\sum_{i=\mu+1}^{m-1}I_i.$$

I_i can be estimated as follows

$$I_i \geqq \sum_{\nu=11\cdot 2^{i-1}}^{23\cdot 2^{i-2}} \left(\frac{1}{2^{i\alpha}} \sum_{l=\nu-10\cdot 2^{i-1}+1}^{2^i} \frac{1}{6^r 2^{ir} l} \right)^p \geqq \sum_{\nu=11\cdot 2^{i-1}}^{23\cdot 2^{i-2}} \left(\frac{1}{2^{i\alpha}} \sum_{l=3\cdot 2^{i-2}+1}^{2^i} \frac{1}{6^r 2^{ir} l} \right)^p \geqq$$

$$\geqq d_1(p, r) 2^{i-2} \frac{1}{2^{i(r+\alpha)p}} = d_2(p, r) 2^{i(1-(r+\alpha)p)}.$$

Hence and from (3.3) we obtain that

$$V_n(f_0, \lambda, p; 0) \geqq d_3(p, r) \left(\frac{1}{\lambda_n} \sum_{i=\mu+1}^{m-1} 2^{i(1-(r+\alpha)p)} \right)^{1/p},$$

whence (1.6) can be deduced by an easy calculation.

The proof of Theorem 3 is thus completed.

Proof of Theorem 4. It is clear that

$$(3.4)\quad \frac{1}{\lambda_n} \sum_{k=n-\lambda_n}^{n} |\sigma_k^\beta(x) - f(x)|^p \leqq \frac{K}{\lambda_n} \sum_{k=n-\lambda_n}^{n} \left(|\sigma_k^\beta(x) - \sigma_k^{\beta+1}(x)|^p + |\sigma_k^{\beta+1}(x) - f(x)|^p \right) \equiv \equiv \sum\nolimits_1 + \sum\nolimits_2 .$$

It is known (see e.g. [1] Theorem 3) that $f(x) \in \mathrm{Lip}\,\alpha$ implies

$$|\sigma_k^{\beta+1}(x) - f(x)| = O(k^{-\alpha}) \quad (\beta > -\tfrac{1}{2}),$$

whence

$$(3.5)\quad \sum\nolimits_2 = O\left(\frac{1}{\lambda_n} \sum_{k=n-\lambda_n}^{n} k^{-\alpha p} \right).$$

Furthermore,

$$(3.6)\quad \sum\nolimits_1 = \frac{1}{\lambda_n} \left(\sum_{k=n-\lambda_n}^{n/2} + \sum_{k=n/2}^{n} \right) |\sigma_k^\beta(x) - \sigma_k^{\beta+1}(x)|^p = \sum\nolimits_3 + \sum\nolimits_4 .$$ [1]

By Lemma 2

$$(3.7)\quad \sum\nolimits_4 = O\left(\frac{1}{\lambda_n} n^{1-\alpha p} \right)$$

and if $2^\mu \leqq n - \lambda_n < 2^{\mu+1}$ and $2^{\mu_1} < n/2 \leqq 2^{\mu_1+1}$, then

$$(3.8)\quad \sum\nolimits_3 \leqq \frac{1}{\lambda_n} \sum_{m=\mu}^{\mu_1} \sum_{k=2^m}^{2^{m+1}} |\sigma_k^\beta(x) - \sigma_k^{\beta+1}(x)|^p \leqq \frac{1}{\lambda_n} \sum_{m=\mu}^{\mu_1} 2^{m(1-\alpha p)}.$$

Collecting the estimates (3.4), (3.5), (3.6), (3.7) and (3.8) an easy calculation gives the statements of (1.7), which is the required proof.

---

[1] $\sum_{n=a}^{b}$, where a and b are not integers, means a sum over all integers between a and b; if $b < a$ then the sum means zero.

Proof of Theorem 5. The proof runs on analogous lines as that of Szabados. Using the Lebesgue's estimate and (1.8) we obtain

$$E_{2n} \leqq \left\| \frac{1}{n+1} \sum_{k=n}^{2n} s_k(x) - f(x) \right\| \leqq$$

$$\leqq \frac{1}{n} \left\| \sum_{k=n}^{2n} |s_k(x)-f(x)|^p \, |s_k(x)-f(x)|^{1-p} \right\| \leqq K_1 \frac{1}{n} (E_n \log n)^{1-p},$$

whence, by a standard computation (see inequality (8) in [7]),

$$E_n^p = O\left(n^{-1}(\log n)^{1-p}\right) \tag{3.9}$$

follows. Using the estimate ([6], Theorem 8, p. 61)

$$E_n(f^{(r)}) \leqq K(r) \sum_{k=[n/2]}^{\infty} k^{r-1} E_k(f),$$

(3.9) implies that

$$E_n(f^{(r)}) = O\left(\frac{(\log n)^{1/p-1}}{n^\alpha}\right),$$

whence, according to the inequality ([6], Theorem 4, p. 59)

$$\omega(f, h) \leqq Kh \sum_{n=0}^{1/h} E_n(f)$$

we get

$$\omega(f^{(r)}, h) \leqq Kh \sum_{n=1}^{1/h} \frac{(\log n)^{1/p-1}}{n^\alpha} \leqq K_1 h^\alpha \left(\log \frac{1}{h}\right)^{1/p-1}$$

which completes the proof.

References

[1] L. Leindler, Über die Approximation im starken Sinne, *Acta Math. Acad. Sci. Hung.*, **16** (1965), 255—262.

[2] L. Leindler, On summability of Fourier series, *Acta Sci. Math.*, **29** (1968), 147—162.

[3] L. Leindler, On strong approximation of Fourier series, *Periodica Math. Hung.*, **1** (1971), 157—162.

[4] L. Leindler, On strong approximation of Fourier series, *Approximation Theory*, Proceedings of a Conference in Poznań 1972, 129—140.

[5] L. Leindler and E. M. Nikišin, Note on strong approximation by Fourier series, *Acta Math. Acad. Sci. Hung.*, **24** (1973), 223—227.

[6] G. G. Lorentz, *Approximation of Functions*, Holt, Rinehart and Winston (1966).

[7] J. Szabados, On problem of L. Leindler concerning strong approximation by Fourier series, *Analysis Math.*, **2** (1976), 155—161.

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1.
6720 SZEGED, HUNGARY

Acta Sci. Math., 38 (1976), 325—341

Probability inequalities of exponential type and laws of the interated logarithm

F. MÓRICZ

Introduction

Let $\xi_1, \xi_2, \ldots, \xi_n$ be random variables (in abbreviation: rv); they need not be independent or identically distributed. Set

$$S_k = \sum_{i=1}^{k} \xi_i \quad \text{and} \quad M_n = \max_{1 \leq k \leq n} |S_k|.$$

Further, for each vector $(\xi_{b+1}, \xi_{b+2}, \ldots, \xi_{b+k})$ of k consecutive ξ_i's, let $F_{b,k}$ denote the joint distribution function and let

$$S_{b,k} = \sum_{i=b+1}^{b+k} \xi_i = S_{b+k} - S_b \quad (S_{b,0} = 0)$$

and

$$M_{b,k} = \max \{|S_{b,1}|, |S_{b,2}|, \ldots, |S_{b,k}|\}.$$

Thus $S_k = S_{0,k}$ and $M_n = M_{0,n}$. Set $F_n = F_{0,n}$. The concern of this paper is to provide bounds on $E\{\exp(\lambda M_n)\}$ in terms of given bounds on $E\{\exp(\lambda |S_{b,k}|)\}$, where $\lambda > 0$.

We emphasize that it is *not* assumed that the ξ_i's are independent. The only restrictions on the dependence will be those imposed on the assumed bounds for $E\{\exp(\lambda |S_{b,k}|)\}$. In point of fact, these assumed bounds are guaranteed under a suitable dependence restriction (e.g., mutual independence, martingale differences, weak multiplicativity, or the like).

Bounds on $E\{\exp(\lambda M_n)\}$ are of use in deriving convergence properties of S_n as $n \to \infty$. For development of such results under various dependence restrictions, the theorems of this paper reduce the problem of placing appropriate bounds on $E\{\exp(\lambda M_n)\}$ to the typically easier problem of placing appropriate bounds on $E\{\exp(\lambda |S_{b,k}|)\}$.

Received October 1, 1975.

The proof of our main result (Theorem 1) is based on the "bisection" technique of BILLINGSLEY [1; p. 102] and the treatment is in a setting close to that of SERFLING [9]. The use of Theorem 1 simplifies and extends the method of SERFLING [10] to obtain results such as laws of the iterated logarithm, convergence rates thereof, etc. under probability inequalities of exponential type. For generalities concerning different convergence properties the reader is sent to our main reference [10].

Another extension of Serfling's method based on the study of the moment inequalities of type $E|S_{b,n}|^{\nu}$ with a fixed $\nu>0$ is dealt with in [6].

§ 1. The main result

In the following the function $g(F_{b,k})$ denotes a non-negative functional depending on the joint distribution function of $\xi_{b+1}, \xi_{b+2}, \dots, \xi_{b+k}$. Examples are: $g(F_{b,k})=k^{\alpha}$ where $\alpha>0$, or $g(F_{b,k})=\sum_{i=b+1}^{b+k} a_i^2$ where $\{a_i\}$ is a sequence of numbers. (In most cases a_i^2 is the finite variance of ξ_i, but this plays no role in our results.) In the sequel $C, C_1, C_2, \dots$ denote positive constants; b, k, l, n non-negative integers and λ a positive real number.

Theorem 1. *Suppose that there exists a non-negative function $g(F_{b,k})$ satisfying*

$$g(F_{b,k})+g(F_{b+k,l}) \leqq g(F_{b,k+l}) \quad (all\ b \geqq 0,\ k\geqq 1,\ l\geqq 1) \tag{1.1}$$

such that

$$E\{e^{\lambda|S_{b,k}|}\} \leqq Ce^{\lambda^2 g(F_{b,k})} \quad (all\ b \geqq 0,\ k \geqq 1, \lambda > 0). \tag{1.2}$$

Then

$$E\{e^{\lambda M_n}\} \leqq 8Ce^{12\lambda^2 g(F_n)} \quad (all\ n \geqq 1, \lambda > 0). \tag{1.3}$$

In Theorem 1 the bounds may involve parameters of the joint distribution function of $\xi_1, \xi_2, \dots, \xi_n$, a flexibility particularly useful with non-identically distributed rv.

Proof. We are to find two constants C_1 and C_2 not less then 1, for which

$$E\{e^{\lambda M_n}\} \leqq C_1 e^{C_2\lambda^2 g(F_n)} \quad (n\geqq 1, \lambda > 0). \tag{1.4}$$

The proof goes by induction on n. The result is trivial for $n=1$. Assume now as induction hypotheses that the result holds for each integer less than n. The function $g(F_n)$ being non-negative and non-decreasing in n, we may assume $g(F_n)>0$. There exists an integer h, $1\leqq h\leqq n$, such that

$$g(F_{h-1}) \leqq \tfrac{1}{2}g(F_n) < g(F_h), \tag{1.5}$$

where $g(F_{h-1})$ on the left is 0 if $h=1$. Then (1.1) and (1.5) imply

$$g(F_{h,n-h}) \leqq g(F_{0,n})-g(F_{0,h}) < \frac{1}{2} g(F_n). \tag{1.6}$$

It is obvious that for $1 \leqq k < h$ we have

$$|S_k| \leqq M_{0,h-1},$$

and for $h \leqq k \leqq n$

$$|S_k| \leqq |S_h| + M_{h,n-h}.$$

Also, for $1 \leqq k \leqq n$ and $\lambda > 0$ we have

$$\lambda |S_k| \leqq \lambda |S_h| + \log (e^{\lambda M_{0,h-1}} + e^{\lambda M_{h,n-h}}).$$

Therefore,

$$\lambda M_n \leqq \lambda |S_h| + \log (e^{\lambda M_{0,h-1}} + e^{\lambda M_{h,n-h}}),$$

whence

$$e^{\lambda M_n} \leqq e^{\lambda |S_h|} (e^{\lambda M_{0,h-1}} + e^{\lambda M_{h,n-h}})$$

for all $\lambda > 0$. Let p and q be positive numbers with $1/p + 1/q = 1$, whose values will be determined later on. Using Hölder's and then Minkowski's inequalities, we find that

$$\begin{aligned} E\{e^{\lambda M_n}\} &\leqq E\{e^{p\lambda |S_h|}\}^{1/p} E\{(e^{\lambda M_{0,h-1}} + e^{\lambda M_{h,n-h}})^q\}^{1/q} \leqq \\ &\leqq E\{e^{p\lambda |S_h|}\}^{1/p} (E\{e^{q\lambda M_{0,h-1}}\}^{1/q} + E\{e^{q\lambda M_{h,n-h}}\}^{1/q}). \end{aligned} \tag{1.7}$$

Since $h-1 < n$, we may apply the induction hypothesis to the rv $\xi_1, \xi_2, \ldots, \xi_{h-1}$ and conclude by (1.4) that

$$E\{e^{q\lambda M_{0,h-1}}\}^{1/q} \leqq C_1^{1/q} e^{qC_2\lambda^2 g(F_{h-1})} \leqq C_1^{1/q} \exp\left[\frac{1}{2} qC_2 \lambda^2 g(F_n)\right], \tag{1.8}$$

the last inequality following by (1.5). We note that if $h=1$, then (1.8) is obvious.

If the indices in (1.2) are restricted to $b \geqq h$, $1 \leqq k \leqq n-b$, then only the rv $\xi_{h+1}, \xi_{h+2}, \ldots, \xi_n$ are involved. Since $n-h < n$, the induction hypothesis applies to $\xi_{h+1}, \xi_{h+2}, \ldots, \xi_n$. Hence (1.4) yields

$$E\{e^{q\lambda M_{h,n-h}}\}^{1/q} \leqq C_1^{1/q} e^{qC_2\lambda^2 g(F_{h,n-h})} \leqq C_1^{1/q} \exp[\tfrac{1}{2} qC_2\lambda^2 g(F_n)], \tag{1.9}$$

where the last inequality follows by (1.6). (If $h=n$, (1.9) is trivial.)

Finally, (1.2) implies

$$E\{e^{p\lambda |S_h|}\}^{1/p} \leqq C^{1/p} e^{p\lambda^2 g(F_h)} \leqq C^{1/p} e^{p\lambda^2 g(F_n)}. \tag{1.10}$$

Combining inequalities (1.7)—(1.10), we arrive at

$$E\{e^{\lambda M_n}\} \leqq 2C^{1/p} C_1^{1/q} \exp[(p + \tfrac{1}{2} qC_2) \lambda^2 g(F_n)].$$

Assuming $1 < q < 2$, and consequently $p > 2$, we have

$$2C^{1/p} C_1^{1/q} \leqq C_1 \quad \text{and} \quad p + \tfrac{1}{2} qC_2 \leqq C_2,$$

provided

$$C_1 \geqq 2^p C \quad \text{and} \quad C_2 \geqq \frac{2p}{2-q}. \tag{1.11}$$

Choosing, for example, $q=3/2$ and $p=3$, the smallest C_1 and C_2 satisfying (1.11) are given by $C_1=8C$ and $C_2=12$, as they are given in (1.3). This completes the induction step and the proof of Theorem 1.

Although the specific values of C_1 and C_2 will have no importance for us, the best value (provided by the above proof) of C_2 may be taken as $C_2=6+4\sqrt{2}$. (Namely, the expression $2p/(2-q)$ attains its minimum on $(2, \infty)$ at $p=2+\sqrt{2}$.)

The extension of the validity of Theorem 1, when λ^2 in the exponents on the right of (1.2) and (1.3) is replaced by a polynomial in λ, say $r(\lambda)$, is of interest in itself and may be of use in some applications.

Theorem 2. *Suppose that there exist a non-negative function $g(F_{b,k})$ satisfying (1.1) and a polynomial*

$$r(\lambda) = \sum_{i=l}^{m} \alpha_i \lambda^i$$

of at least first degree, strictly positive for $\lambda>0$, such that

$$E\{e^{\lambda|S_{b,k}|}\} \leqq Ce^{r(\lambda)g(F_{b,k})} \quad (all\ b \geqq 0,\ k \geqq 1,\ \lambda > 0). \tag{1.12}$$

Then

$$E\{e^{\lambda M_n}\} \leqq C_1 e^{C_2 r(\lambda) g(F_n)} \quad (all\ n \geqq 1,\ \lambda > 0), \tag{1.13}$$

where C_1 and C_2 are constants depending only on $r(\lambda)$.

Proof. The proof of Theorem 2 runs along the same lines as that of Theorem 1. The same sort of argument that yielded (1.8)—(1.10) shows that

$$E\{e^{q\lambda M_{0,h-1}}\}^{1/q} \leqq C_1^{1/q} \exp\left[\frac{1}{2q} C_2 r(q\lambda) g(F_n)\right],$$

$$E\{e^{q\lambda M_{h,n-h}}\}^{1/q} \leqq C_1^{1/q} \exp\left[\frac{1}{2q} C_2 r(q\lambda) g(F_n)\right],$$

and

$$E\{e^{p\lambda|S_h|}\}^{1/p} \leqq C^{1/p} \exp\left[\frac{1}{p} r(p\lambda) g(F_n)\right].$$

Combining inequality (1.7) with the last three ones, we arrive at

$$E\{e^{\lambda M_n}\} \leqq 2C^{1/p} C_1^{1/q} \exp\left(\left[\frac{1}{p} r(p\lambda) + \frac{1}{2q} C_2 r(q\lambda)\right] g(F_n)\right). \tag{1.14}$$

Now we have to choose $q<2$ $(p=q/(q-1))$ and the constants C_1, C_2 in such a way that

$$2C^{1/p}C_1^{1/q} \leqq C_1 \tag{1.15}$$

and

$$\frac{1}{p}r(p\lambda)+\frac{1}{2q}C_2 r(q\lambda) \leqq C_2 r(\lambda) \tag{1.16}$$

hold for all $\lambda>0$. Condition (1.15) does not cause any difficulty. On the other hand, (1.16) requires some arguments. Writing

$$s(\lambda) = C_2\left[r(\lambda)-\frac{1}{2q}r(q\lambda)\right]-\frac{1}{p}r(p\lambda),$$

we will prove the existence of q and C_2 such that $s(\lambda)\geqq 0$ for all $\lambda>0$.

First we notice that from the assumption on $r(\lambda)$ it immediately follows that $\alpha_m>0$ and $\alpha_l>0$. Then we show that

$$r(\lambda)-\frac{1}{2q}r(q\lambda) \geqq \frac{1}{4}r(\lambda) \tag{1.17}$$

for all $\lambda>0$, provided q is sufficiently close to 1. Inequality (1.17) is equivalent to

$$t(\lambda) = 3r(\lambda)-\frac{2}{q}r(q\lambda) \geqq 0 \tag{1.18}$$

for all $\lambda>0$. We consider only those q's for which $q^{m-1}\leqq 3/2$ minus a small positive number, say let $q^{m-1}\leqq 5/4$. A simple reasoning gives that if

$$\lambda \geqq \max\left(1, \frac{1}{2\alpha_m}\sum_{i=l}^{m-1}|\alpha_i|\right)$$

or

$$0 < \lambda \leqq \min\left(1, \frac{\alpha_l}{2\sum_{i=l+1}^{m}|\alpha_i|}\right)$$

then (1.18) is true. Since

$$\lim_{q\to 1+0} t(\lambda) = r(\lambda)$$

uniformly on each finite segment, hence we can choose q, $1<q$ and $q^{m-1}\leqq 5/4$, such that (1.18) holds for all $\lambda>0$. Thus we can and do fix $q>1$ for which (1.17) is satisfied. Let $p=q/(q-1)$ and return to the study of $s(\lambda)$.

The behaviour of $s(\lambda)$ for λ large enough is determined by the coefficient of λ^m. Hence we have to choose C_2 such that

$$\alpha_m\left(C_2-\frac{1}{2}C_2q^{m-1}-p^{m-1}\right)>0,$$

i.e.,

$$C_2>\frac{2p^{m-1}}{2-q^{m-1}}. \tag{1.19}$$

This choice implies $s(\lambda)\geqq 0$ for sufficiently large λ, say $\lambda\geqq\Lambda_0$.

In case when λ is small enough, the coefficient of λ^l is decisive for the sign of $s(\lambda)$. In order to ensure that $s(\lambda)\geqq 0$ for sufficiently small λ, say $0<\lambda\leqq\lambda_0$, we have to require that

$$C_2>\frac{2p^{l-1}}{2-q^{l-1}}.$$

But condition (1.19) implies this, it suffices to keep in mind only that $m\geqq l$, $p>2$, $q>1$, and $q^{m-1}<2$.

Thus it remains to deal with the case $\lambda_0\leqq\lambda\leqq\Lambda_0$. Since the polynomial $r(\lambda)$ has no zero on $0<\lambda<\infty$, it follows that

$$r_1=\min_{\lambda_0\leqq\lambda\leqq\Lambda_0} r(\lambda)$$

is a positive number. Further, set

$$R_p=\max_{\lambda_0\leqq\lambda\leqq\Lambda_0}\frac{1}{p}r(p\lambda).$$

Taking into account that (1.17) holds for all $\lambda>0$, we have

$$s(\lambda)\geqq\frac{1}{4}C_2r(\lambda)-\frac{1}{p}r(p\lambda)\geqq\frac{1}{4}C_2r_1-R_p\geqq 0$$

for every λ in $[\lambda_0, \Lambda_0]$ provided $C_2\geqq 4R_p/r_1$. If, in addition, C_2 fulfills (1.19) then we can conclude that $s(\lambda)\geqq 0$, and consequently, (1.16) is satisfied for all $\lambda>0$. Finally, if $C_1=2^pC$ then (1.15) is also satisfied.

Continuing our reasoning with (1.14), by (1.15) and (1.16) we arrive at the desired (1.13). Thus we finished the proof of Theorem 2.

Before coming to the applications, we make a remark on the validity of Theorems 1 and 2. Viewing the proofs, it is striking that we use no full power of a probability space. In fact, Hölder's and Minkowski's inequalities were applied only, which are available in any measure space (X, A, μ). Hence Theorems 1 and 2 are valid on (X, A, μ) taking integrals over X with respect to μ in place of the expectations on the left-hand sides of the corresponding inequalities.

§ 2. Laws of the iterated logarithm as consequences of a probability inequality of exponential type for $S_{b,n}$

Now we will discuss the stochastic convergence properties of S_n under restrictions of type (1.2). The following result, which expresses a form of the law of the iterated logarithm, certainly has a broad scope of application.

Theorem 3. *Suppose that there exist a positive number K and a sequence* $\{a_i\}$ *of numbers such that*

$$E\{e^{\lambda|S_{b,k}|}\} \leqq C\exp\left(\tfrac{1}{2}K\lambda^2A_{b,k}^2\right) \quad (\text{all } b \geqq 0,\ k \geqq 1,\ \lambda > 0), \tag{2.1}$$

where

$$A_{b,k} = \left(\sum_{i=b+1}^{b+k} a_i^2\right)^{1/2} \quad \text{and} \quad A_n = A_{0,n} \to \infty \quad (n \to \infty). \tag{2.2a}$$

Then it follows a law of the iterated logarithm with K, i.e.

$$P\left\{\limsup_{n\to\infty}\frac{|S_n|}{(2KA_n^2\log\log A_n)^{1/2}} \leqq 1\right\} = 1. \tag{2.3}$$

We note that the conclusion of Theorem 3 in the special case $a_i \equiv 1$, $A_n^2 = n$ was proved by Serfling [10, Theorem 4.1] for *uniformly bounded rv,* $|\xi_i| \leqq B$, *having the following properties:*

(i) *for any* $\nu > 2$ *there exists a constant* C_ν *such that*

$$E|S_{b,n}|^\nu \leqq C_\nu n^{\nu/2} \quad (\text{all } b \geqq 0,\ n \geqq 1), \tag{2.4}$$

(ii) *the inequality*

$$P\{|S_n| > y\} \leqq 2\exp\left\{-\frac{y^2}{2B^2n}\right\} \quad (\text{all } n \geqq 1)$$

holds for any $y > 0$.

The following theorem provides information on the rate of convergence in (2.3).

Theorem 4. *Suppose that* (2.1) *holds, where*

$$A_n \to \infty \quad \text{and} \quad a_n = o(A_n) \quad (n \to \infty). \tag{2.2b}$$

Then, for each $\theta > 2K$, *we have*

$$\sum_n \frac{a_n^2}{A_n^2\log A_n} P\left\{\sup_{k\geqq n}\frac{|S_k|}{(\theta A_k^2\log\log A_k)^{1/2}} \geqq 1\right\} < \infty. \tag{2.5}$$

If the factor $(\theta\log\log A_k)^{1/2}$ in the expression (2.5) is replaced by a rougher factor $(\log A_k)^\alpha$ with an $\alpha > 0$, then an essentially better rate of convergence depending on α can be achieved, as the following theorem shows.

Theorem 5. *Suppose that* (2.1) *and* (2.2b) *hold. Then setting*

$$P_n = P\left\{\sup_{k\geqq n} \frac{|S_k|}{A_k(\log A_k)^\alpha} \geqq 1\right\}$$

we have for each choice of $0<\alpha<1/2$ *and* $\beta>0$

$$\sum_n \frac{a_n^2(\log A_n)^\beta}{A_n^2} P_n < \infty,$$

for $\alpha=1/2$ *and* $\beta>0$

$$\sum_n \frac{a_n^2}{A_n^{\beta+(2K-1)/K}} P_n < \infty,$$

and for $\alpha>1/2$ *and* $\beta>0$

$$\sum_n a_n^2 A_n^\beta P_n < \infty.$$

It is instructive to compare Theorem 5 with a result of Serfling [10, Corollary 5.3.1], which reads as follows: *Suppose that in the special case* $a_i\equiv 1$, $A_n^2=n$, *we have* (2.4) *for all* $v>2$. *Then*

$$\sum_n \frac{1}{n(\log n)^{1-\beta}} P\left\{\sup_{k\geqq n} \frac{|S_k|}{k^{1/2}(\log k)^\alpha} > 1\right\} < \infty$$

holds for each choice of $\alpha>0$ *and* $0<\beta<1$.

The results stated in Theorems 3—5 are obtained by adaption of more or less standard arguments [2], [4], and [7] making use of Theorem 1. More precisely, bounds on $E\{\exp(\lambda M_{b,k})\}$ are of use in deriving bounds on the tail distribution of $M_{b,k}$. By Chebyshev's inequality, (2.1) implies

$$(2.6)\quad P\{|S_n|\geqq y\} = P\{e^{\lambda|S_n|}\geqq e^{\lambda y}\} \leqq C\exp\left(\frac{1}{2}K\lambda^2A_n^2-\lambda y\right) = C\exp\left(-\frac{y^2}{2KA_n^2}\right),$$

if λ is chosen as $\lambda=y/KA$. Here and in the sequel y denotes a positive number. Further, also by Chebyshev's inequality, (2.1) implies via Theorem 1 that

$$(2.7)\qquad P\{M_{b,k}\geqq y\} \leqq 8C\exp\left(-\frac{y^2}{24KA_{b,k}^2}\right).$$

The proofs below are based on the bounds (2.7) on the tail distribution of $M_{b,k}$, which is of interest in its own right, too. An extra factor of 8 in the coefficient on the right-hand side of (2.7) will not matter for our purposes, and the bounds we derive will decrease with increasing y slowly enough that passing from y^2 to $y^2/12$ in the exponent will have no important effect.

Proof of Theorem 3. We have to prove that, for any $\theta>2K$, with probability 1 we have

$$|S_n| \leqq (\theta A_n^2 \log\log A_n^2)^{1/2}$$

for all n large enough. It is clear that this implies (2.3).

Let $\delta>1$ be a fixed number and define a sequence of integers $1 \leqq n_1 \leqq n_2 \leqq \dots$ in the following way:

$$A_{n_k-1}^2 \leqq \delta^k < A_{n_k}^2 \quad (k=1,2,\dots;\ A_0=0). \tag{2.8}$$

This is possible by (2.2a), and obviously $n_k \to \infty$ as $k \to \infty$.

Set

$$\gamma = \frac{\theta}{2K} \quad \text{and} \quad \mu(n) = (\theta A_n^2 \log\log A_n^2)^{1/2}.$$

By the above assumption $\gamma>1$. Then (2.6) provides

$$P\{|S_{n_k}| \geqq \mu(n_k)\} \leqq C\exp(-\gamma \log\log A_{n_k}^2) = \frac{C}{(\log A_{n_k}^2)^\gamma}.$$

By (2.8) we get

$$\sum_k{}' P\{|S_{n_k}| \geqq \mu(n_k)\} \leqq \frac{C}{(\log\delta)^\gamma}\sum_{k=1}^\infty \frac{1}{k^\gamma} < \infty,$$

where $\sum_k{}'$ means that the summation is taken only once for equal n_k's. In virtue of the Borel—Cantelli lemma, this yields with probability 1 that

$$|S_{n_k}| \leqq (\theta A_{n_k}^2 \log\log A_{n_k}^2)^{1/2} \tag{2.9}$$

for all k large enough.

For an arbitrary n, either $n=n_k$ or $n_k<n<n_{k+1}$ for some k. If $n_k<n<n_{k+1}$, consider

$$\frac{S_n}{\mu(n)} = \frac{S_{n_k}}{\mu(n_k)}\frac{\mu(n_k)}{\mu(n)} + \frac{|S_n - S_{n_k}|}{\bar\mu(n_k)}\frac{\bar\mu(n_k)}{\mu(n)},$$

where

$$\bar\mu(n_k) = (12\,\theta A_{n_k,\nu_k-1}^2 \log\log A_{n_k}^2)^{1/2} \quad \text{and} \quad \nu_k = n_{k+1}-n_k.$$

Since $\mu(n)$ is non-decreasing, it follows that

$$\frac{|S_n|}{\mu(n)} \leqq \frac{|S_{n_k}|}{\mu(n_k)} + \frac{|S_n - S_{n_k}|}{\bar\mu(n_k)}\frac{\bar\mu(n_k)}{\mu(n)}. \tag{2.10}$$

We will show that with probability 1

$$\max_{n_k<n<n_{k+1}} \frac{|S_n - S_{n_k}|}{\bar\mu(n_k)} = \frac{M_{n_k,\nu_k-1}}{\bar\mu(n_k)} \leqq 1 \tag{2.11}$$

for all k large enough. To this effect, utilize (2.7). Then

$$P\{M_{n_k,\nu_k-1} \geqq \bar{\mu}(n_k)\} \leqq 8C\exp(-\gamma \log\log A_{n_k}^2).$$

As above, this implies

$$\sum_k{}'' P\{M_{n_k,\nu_k-1} \geqq \bar{\mu}(n_k)\} < \infty,$$

where $\sum_k{}''$ means that the summation is extended to such k's that $n_k < n_{k+1}-1$. By the Borel—Cantelli lemma we get the wanted (2.11).

Owing to (2.8) we have $A_{n_k}^2 > \delta^k$ and

$$A_{n_k,\nu_k-1}^2 = A_{n_{k+1}-1}^2 - A_{n_k}^2 \leqq \delta^k(\delta-1).$$

Thus

$$\frac{\bar{\mu}(n_k)}{\mu(n_k)} = \frac{\sqrt{12}\,A_{n_k,\nu_k-1}}{A_{n_k}} \leqq [12(\delta-1)]^{1/2}.$$

The right-most member here can be made as small as needed if $\delta \to 1$. Hence, combining (2.9)—(2.11) it follows that, for any $\varepsilon > 0$, with probability 1

$$|S_n| \leqq [(\theta+\varepsilon)A_n^2 \log\log A_n^2]^{1/2}$$

holds for all n large enough. Since $\theta+\varepsilon$ may be chosen arbitrarily close to $2K$, the conclusion of Theorem 3 is proved.

Proof of Theorem 4. Let $\delta > 1$ be a fixed number. We will show that (2.2b) implies the existence of a strictly increasing sequence $\{n_k\}$ of positive integers such that

$$\delta^k \leqq A_{n_k}^2 < \delta^{k+1} \tag{2.12}$$

for all k large enough. Otherwise, for infinitely many n's, we have

$$A_n^2 < \delta^{k+1} \quad \text{and} \quad A_{n+1}^2 \geqq \delta^{k+2}$$

with suitable k's. This gives that

$$\frac{a_{n+1}^2}{A_{n+1}^2} = 1 - \frac{A_n^2}{A_{n+1}^2} \geqq 1 - \frac{\delta^{k+1}}{\delta^{k+2}} = \frac{\delta-1}{\delta}$$

for infinitely many n's, which contradicts (2.2b).

In proving the convergence of the series (2.5), we make use of the convergence part of the following assertion, applied widely in the theory of numerical series: *Let $d_i \geqq 0$ be the terms of a divergent series with partial sums D_n. Then the series*

$$\sum_n \frac{d_n}{D_n(\log D_n)^{1+\varepsilon}}$$

converges or diverges according as $\varepsilon>0$ *or* $\varepsilon\leqq 0$. Hence it is enough to demonstrate that

$$(2.13)\qquad P_n = P\left\{\sup_{l\geqq n}\frac{|S_l|}{(\theta A_l^2\log\log A_l^2)^{1/2}}\geqq 1\right\}\leqq\frac{C_3}{(\log A_n^2)^{\varepsilon}}$$

with an appropriate $\varepsilon>0$.

To this effect, let us fix a number θ_1 so that

$$(2.14)\qquad 2K<\theta_1<\theta.$$

Let $k_0=k_0(n)$ be defined by $n_{k_0}<n\leqq n_{k_0+1}$. We may assume that n, and consequently k, are large enough, so that (2.12) is satisfied. It is obvious that

$$(2.15)\qquad P_n\leqq\sum_{k=k_0}^{\infty}Q_k\quad\text{where}\quad Q_k=P\left\{\max_{n_k<l\leqq n_{k+1}}\frac{|S_l|}{(\theta A_l^2\log\log A_l^2)^{1/2}}\geqq 1\right\}.$$

It can be easily checked that

$$(2.16)\quad Q_k\leqq P\left\{\frac{|S_{n_k}|}{[\theta_1\sigma(n_k)]^{1/2}}\geqq 1\right\}+P\left\{\max_{n_k<l\leqq n_{k+1}}\frac{|S_l-S_{n_k}|}{[2K\sigma(n_k)]^{1/2}}\geqq\eta\right\}=Q_{1,k}+Q_{2,k},$$

where, for the sake of brevity, we put

$$\sigma(n)=A_n^2\log\log A_n^2\quad\text{and}\quad\eta=\left[1-\left(\frac{\theta_1}{\theta}\right)^{1/2}\right]\left(\frac{\theta}{2K}\right)^{1/2}.$$

Repeating the argument that yielded (2.9) in the proof of Theorem 3, we can establish with ease by (2.6) that

$$Q_{1,k}\leqq C\exp(-\gamma_1\log\log A_{n_k}^2)=\frac{C}{(\log A_{n_k}^2)^{\gamma_1}},$$

where $\gamma_1=\theta_1/2K$. By (2.14) we have $\gamma_1>1$. Thus, using (2.12), we find that

$$(2.17)\qquad \begin{aligned}\sum_{k=k_0}^{\infty}Q_{1,k}&\leqq\frac{C}{(\log\delta)^{\gamma_1}}\sum_{k=k_0}^{\infty}\frac{1}{k^{\gamma_1}}\leqq\frac{C}{(\gamma_1-1)(\log\delta)^{\gamma_1}(k_0-1)^{\gamma_1-1}}\leqq\\ &\leqq\frac{2^{\gamma_1-1}C}{(\gamma_1-1)(\log\delta)^{\gamma_1}(k_0+2)^{\gamma_1-1}}\leqq\frac{2^{\gamma_1-1}C}{(\gamma_1-1)\log\delta\,(\log A_n^2)^{\gamma_1-1}},\end{aligned}$$

provided $k_0+2\leqq 2(k_0-1)$, i.e., $k_0\geqq 4$, which we may assume without loss of generality.

Let us now deal with the series $\sum_{k=k_0}^{\infty}Q_{2,k}$. By (2.7) it is bounded from above by the series

$$8C\sum_{k=k_0}^{\infty}\exp\left(-\frac{\eta^2A_{n_k}^2\log\log A_{n_k}^2}{12(A_{n_{k+1}}^2-A_{n_k}^2)}\right),$$

and therefore also by

$$8C\sum_{k=k_0}^{\infty}\exp\left(-\frac{\eta^2\log\log A_{n_k}^2}{12(\delta^2-1)}\right)=8C\sum_{k=k_0}^{\infty}\frac{1}{(\log A_{n_k}^2)^{\gamma_2}}$$

with $\gamma_2=\eta^2/12(\delta^2-1)$, since by (2.12)

$$\frac{A_{n_k}^2}{A_{n_{k+1}}^2-A_{n_k}^2}\geqq\frac{\delta^k}{\delta^{k+2}-\delta^k}=\frac{1}{\delta^2-1}.$$

Since δ may be chosen arbitrary close to 1, fix $\delta>1$ in such a way that $\gamma_2>1$. Then the same sort of argument that yielded (2.17) shows that

$$\sum_{k=k_0}^{\infty}Q_{2,k}\leqq\frac{2^{\gamma_2+2}C}{(\gamma_2-1)\log\delta\,(\log A_n^2)^{\gamma_2-1}}. \tag{2.18}$$

Putting together (2.15)—(2.18), we arrive at (2.13) with $\varepsilon=\min(\gamma_1,\gamma_2)-1$. This completes the proof of Theorem 4.

The proof of Theorem 5 runs along the same lines as that of Theorem 4. We only notice that after the application of (2.6) and (2.7) we have to use the following elementary inequalities:

$$\exp\{-\gamma(\log x)^{2\alpha}\}\leqq\begin{cases}C(\log x)^{-\beta} & \text{if}\quad 0<\alpha<\frac{1}{2}\quad\text{and}\quad\beta>0,\\ x^{-\gamma} & \text{if}\quad \alpha=\frac{1}{2},\\ Cx^{-\beta} & \text{if}\quad \alpha>\frac{1}{2}\quad\text{and}\quad\beta>0,\end{cases}$$

where $x\geqq2$ and C depend only on α, β and $\gamma>0$.

In the sequel as a particular case, consider a sequence $\{\varphi_i\}$ of *weakly multiplicative* rv, i.e., we assume that

$$W_r=\Big(\sum_{1\leqq i_1<i_2<\ldots<i_r}E^2\{\varphi_{i_1}\varphi_{i_2}\ldots\varphi_{i_r}\}\Big)^{1/2}<\infty\quad(r=4,6,\ldots), \tag{2.19}$$

where the summation is extended over all integers satisfying only the condition $1\leqq i_1<i_2<\ldots<i_r$, and further

$$W_r^{1/r}=O(1)\quad(r\to\infty).$$

This is a generalization of the concept of *multiplicativity* defined by

$$E\{\varphi_{i_1}\varphi_{i_2}\ldots\varphi_{i_r}\}=0\quad(1\leqq i_1<i_2<\ldots<i_r;\ r=4,6,\ldots). \tag{2.20}$$

The condition (2.20) is stronger than (2.19). Even the former includes the case of a sequence of martingale differences and the case of mutually independent rv and special varieties thereof (see Révész [7]).

We proved in [5, Lemma 3] that (2.1) is valid with a definite K for uniformly bounded sequences of weakly multiplicative rv. More precisely, the following result holds: *Let* $\{\varphi_i\}$ *be a sequence of rv such that*

$$|\varphi_i| \leqq B(<\infty) \quad (i=1,2,\ldots) \tag{2.21}$$

and

$$\limsup_{r\to\infty} W_r^{1/r} = W(<\infty). \tag{2.22}$$

Then for every $\gamma>0$ *there exists a constant* C_γ *such that for every sequence* $\{a_i\}$ *of numbers we have*

$$E\{e^{\lambda|S_{b,k}|}\} \leqq C_\gamma \exp\left[\tfrac{1}{2}(B^2+W^2+\gamma)\lambda^2 A_{b,k}^2\right] \quad (\text{all } b\geqq 0,\ k\geqq 1,\ \lambda>0),$$

where

$$S_{b,k} = \sum_{i=b+1}^{b+k} a_i\varphi_i \quad \text{and} \quad A_{b,k}^2 = \sum_{i=b+1}^{b+k} a_i^2.$$

Hence, via Theorems 3—5, we obtain

Corollary 1. *Let* $\{\varphi_i\}$ *be a sequence of rv satisfying* (2.21) *and* (2.22). *Let* $\{a_i\}$ *be a sequence of numbers with* (2.2a). *Then there follows a law of the iterated logarithm for* $\{\xi_i = a_i\varphi_i\}$ *with* $K=B^2+W^2$, *i.e.*,

$$P\left\{\limsup_{n\to\infty} \frac{\left|\sum_{i=1}^{n} a_i\varphi_i\right|}{[2(B^2+W^2)A_n^2\log\log A_n]^{1/2}} \leqq 1\right\} = 1.$$

Corollary 2. *Let* $\{\varphi_i\}$ *be a sequence of rv satisfying* (2.21) *and* (2.22). *Let* $\{a_i\}$ *be a sequence of numbers with* (2.2b). *Then, for each* $\theta>2(B^2+W^2)$, *we have*

$$\sum_n \frac{a_n^2}{A_n^2\log A_n} P\left\{\sup_{k\geqq n} \frac{\left|\sum_{i=1}^{k} a_i\varphi_i\right|}{(\theta A_k^2\log\log A_k)^{1/2}} \geqq 1\right\} < \infty.$$

Corollary 3. *Under the conditions of Corollary 2 we have*

$$\sum_n \frac{a_n^2(\log A_n)^\beta}{A_n^2} P\left\{\sup_{k\geqq n} \frac{\left|\sum_{i=1}^{k} a_i\varphi_i\right|}{A_k(\log A_k)^\alpha} \geqq 1\right\} < \infty$$

for each choice of $\alpha>0$ *and* $\beta>0$.

Corollaries 1 and 2 were proved by the present author [5] in another way, and the latter one under somewhat more restricted conditions stipulated on $\{a_i\}$. Laws of the iterated logarithm, convergence rates in them was proved for multiplicative rv in the special case $a_i\equiv 1$, $A_n^2=n$, by Serfling [8].

§ 3. Strong convergence and complete convergence

A trivial consequence of the laws of the iterated logarithm is the strong law of large numbers, i.e., under conditions (2.1) and (2.2a) it follows that

$$P\left\{\frac{S_n}{A_n^2} \to 0\right\} = 1. \tag{3.1}$$

It is of interest to obtain information on the rate of convergence in (3.1). Besides, we will give a condition on the sequence $\{c_n\}$ of numbers that

$$\sum_{n=1}^{\infty} P\left\{\frac{|S_n|}{c_n} \geqq \varepsilon\right\}$$

converge for every $\varepsilon>0$, which is referred to as $\{S_n/c_n\}$ *converges completely to zero* in the sense of HSU and ROBBINS [3].

Theorem 6. *Suppose that there exist a positive number K and a sequence $\{a_i\}$ of numbers such that* (2.1) *holds. Furthermore, suppose that with some $\beta>0$ we have*

$$A_n \geqq C_4 n^\beta \quad (n \geqq n_0) \quad \text{and} \quad a_n = o(A_n) \quad (n \to \infty). \tag{3.2}$$

Then, for each $\varepsilon>0$, we have

$$\sum_n \varrho^{A_n^2} P\left\{\sup_{k\geqq n} \frac{|S_k|}{A_k^2} \geqq \varepsilon\right\} < \infty \tag{3.3}$$

for any positive $\varrho<\exp(\varepsilon^2/2K)$; in particular,

$$\sum_n A_n^\alpha P\left\{\sup_{k\geqq n} \frac{|S_k|}{A_k^2} \geqq \varepsilon\right\} < \infty$$

for any $\alpha>0$.

Proof. We use the following elementary inequalities:

(i) If $0<u<1$, $\delta>1$, and k is a positive integer, then

$$u^{\delta^k} + u^{\delta^{k+1}} + u^{\delta^{k+2}} + \ldots \leqq u^{\delta^k}(1-u^{\delta^k(\delta-1)})^{-1}. \tag{3.4}$$

Indeed, if we substitute u^{δ^k} by v then (3.4) becomes

$$v + v^\delta + v^{\delta^2} + \ldots \leqq v(1-v^{\delta-1})^{-1},$$

where $0<v<1$. Now, if $\delta = 1+\eta$ with an $\eta>0$, then

$$v + v^\delta + v^{\delta^2} + \ldots \leqq v + v^{1+\eta} + v^{1+2\eta} + \ldots = v(1-v^\eta)^{-1},$$

which makes (3.4) evident.

(ii) If $0<w<1$ and $\beta>0$ then the series

$$w+w^{2^\beta}+w^{3^\beta}+\ldots$$

is convergent. This is clear by Bernoulli's inequality, according to which $n^\beta \geqq \beta(n-1)$.

After these preliminaries, let us fix $\varepsilon_1<\varepsilon$ so that $\varrho<\exp(\varepsilon_1^2/2K)$ and fix $\delta>1$ in such a way that

$$\varrho < \exp\left(\frac{\varepsilon_1^2}{2K\delta^2}\right) \quad \text{and} \quad \varepsilon_1 \leqq \frac{\varepsilon-\varepsilon_1}{[12(\delta^2-1)]^{1/2}}. \tag{3.5}$$

Then define a strictly increasing sequence $\{n_k\}$ of integers by (2.12) as we did in the proof of Theorem 4.

By (ii) and (3.5) it is enough to prove that

$$I_n = P\left\{\sup_{k\geqq n} \frac{|S_k|}{A_k^2} \geqq \varepsilon\right\} \leqq C_5 \exp\left(-\frac{\varepsilon_1^2}{2K\delta^2} A_n^2\right) \tag{3.6}$$

for all n large enough. Towards this end, let $n_{k_0}<n\leqq n_{k_0+1}$. We obviously have

$$I_n \leqq \sum_{k=k_0}^{\infty} P\left\{\max_{n_k<l\leqq n_{k+1}} \frac{|S_l|}{A_l^2} \geqq \varepsilon\right\} \leqq \sum_{k=k_0}^{\infty} P\left\{\frac{|S_{n_k}|}{A_{n_k}^2} \geqq \varepsilon_1\right\}+$$

$$+\sum_{k=k_0}^{\infty} P\left\{\max_{n_k<l\leqq n_{k+1}} \frac{|S_l-S_{n_k}|}{A_{n_k}^2} \geqq \varepsilon-\varepsilon_1\right\} = J_1+J_2.$$

Applying (2.6) with $y=\varepsilon_1 A_{n_k}^2$ gives

$$J_1 \leqq C \sum_{k=k_0}^{\infty} \exp\left(-\frac{\varepsilon_1^2}{2K} A_{n_k}^2\right) \leqq C \sum_{k=k_0}^{\infty} \exp\left(-\frac{\varepsilon_1^2 \delta^k}{2K}\right),$$

while the application of (2.7) with $y=(\varepsilon-\varepsilon_1)A_{n_k}^2$ and (3.5) leads us to

$$J_2 \leqq 8C \sum_{k=k_0}^{\infty} \exp\left(-\frac{(\varepsilon-\varepsilon_1)^2 A_{n_k}^4}{24K(A_{n_{k+1}}^2-A_{n_k}^2)}\right) \leqq$$

$$\leqq 8C \sum_{k=k_0}^{\infty} \exp\left(-\frac{(\varepsilon-\varepsilon_1)^2 \delta^k}{24K(\delta^2-1)}\right) \leqq 8C \sum_{k=k_0}^{\infty} \exp\left(-\frac{\varepsilon_1^2\delta^k}{2K}\right),$$

where we used that by (2.12)

$$\frac{A_{n_k}^4}{A_{n_{k+1}}^2-A_{n_k}^2} \geqq \frac{\delta^{2k}}{\delta^{k+2}-\delta^k} = \frac{\delta^k}{\delta^2-1}.$$

To sum up,

$$I_n \leqq J_1+J_2 \leqq 9C \sum_{k=k_0}^{\infty} \exp\left(-\frac{\varepsilon_1^2\delta^k}{2K}\right).$$

Now making use of (3.4) with $v=\exp(-\varepsilon_1^2/2K)$ and of (2.12), we get that

$$J_n \leqq 9C\exp\left(-\frac{\varepsilon_1^2}{2K}\delta^{k_0}\right)\left(1-\exp\left[-\frac{\varepsilon_1^2}{2K}\delta^{k_0}(\delta-1)\right]\right)^{-1} \leqq \tag{3.7}$$
$$\leqq 18C\exp\left(-\frac{\varepsilon_1^2}{2K\delta^2}A^2_{n_{k_0+1}}\right) \leqq 18C\exp\left(-\frac{\varepsilon_1^2}{2K\delta^2}A_n^2\right),$$

provided

$$\exp\left[-\frac{\varepsilon_1^2}{2K}\delta^{k_0}(\delta-1)\right] \leqq \frac{1}{2},$$

which is the case if n (and a fortiori k_0) is large enough.

Observe that (3.6) and (3.7) coincide if C_5 is taken to $18C$. This completes the proof of Theorem 6.

Finally, we consider the question of norming S_n in such a way that S_n/c_n converge completely to zero. The following theorem may be derived.

Theorem 7. *Suppose that there exist a positive number K and a sequence $\{a_i\}$ of numbers such that* (2.1) *holds. Furthermore, suppose that with some $\beta>0$ we have* (3.2). *Then $M_n/(A_n^2\log A_n)^{1/2}g(n)$, and hence also $S_n/(A_n^2\log A_n)^{1/2}g(n)$, converges completely to 0 if $g(n)\to\infty$ as $n\to\infty$.*

Proof. Let $\varepsilon>0$ be given. Then we obtain immediately by (2.7) that

$$\Sigma = \sum_n P\left\{\frac{M_n}{(A_n^2\log A_n)^{1/2}g(n)} \geqq \varepsilon\right\} \leqq 8C\sum_n \exp\left(-\frac{\varepsilon^2 g^2(n)\log A_n}{24K}\right) = 8C\sum_n A_n^{-\nu_n},$$

where $\nu_n=\varepsilon^2 g^2(n)/24K$. Taking into account (3.2), it follows that

$$\Sigma \leqq 8C\sum_n n^{-\beta\nu_n} < \infty,$$

since $\beta\nu_n$ with $g(n)$ tends to ∞ as $n\to\infty$. Here we suppose that $C_4\geqq 1$, but this does not bother generality. The proof of Theorem 7 is ready.

Condition (3.2) stipulated on the growth of A_n, plays a crucial role in the proofs of Theorems 6 and 7. Namely, (3.2) ensures the convergence of the series $\sum q^{A_n}$ for $0<q<1$ (in the proof of Theorem 6) and that of the series $\sum A_n^{-g(n)}$ for $g(n)\to\infty$ (in the proof of Theorem 7), which fail if, for example, $A_n=\log n$, $q=1/2$, and $g(n)=\log\log n$. Of course, it might be some relaxation of (3.2) using another technique, but we are unable to do so.

Confining attention to a uniformly bounded sequence of weakly multiplicative rv, we get the following

Corollary 4. *Let* $\{\varphi_i\}$ *be a sequence of rv satisfying* (2.21) *and* (2.22). *Let* $\{a_i\}$ *be a sequence of numbers with* (3.2). *Then, for each* $\varepsilon>0$, *we have*

$$\sum_n \varrho^{A_n^2} P\left\{\sup_{k\geqq n} \frac{1}{A_k^2}\left|\sum_{i=1}^k a_i\varphi_i\right| \geqq \varepsilon\right\} < \infty$$

for any $\varrho<\exp[\varepsilon^2/2(B^2+W^2)]$.

Corollary 5. *Let* $\{\varphi_i\}$ *be a sequence of rv satisfying* (2.21) *and* (2.22). *Under conditions* (3.2) *we have*

$$\sum_n P\left\{\frac{1}{(A_n^2\log A_n)^{1/2}g(n)} \max_{1\leqq k\leqq n}\left|\sum_{i=1}^k a_i\varphi_i\right| \geqq \varepsilon\right\} < \infty,$$

provided $g(n)\to\infty$ *as* $n\to\infty$.

We note that Theorem 6 in the special case $a_i\equiv 1$, $A_n^2=n$, was proved by SERFLING [10, Theorem 5.2]. Furthermore, Corollaries 4 and 5 were proved also by SERFLING [8] for sequences of uniformly bounded multiplicative rv and for $a_i\equiv 1$. The proofs given above essentially differ from those of Serfling, since in the case of general sequences $\{a_i\}$ (satisfying merely (3.2)) not only (2.6) but also (2.7) are employed.

References

[1] P. BILLINGSLEY, *Convergence of probability measures*, Wiley (New York, 1968).
[2] J. L. DOOB, *Stochastic processes, Wiley* (New York, 1953).
[3] P. L. HSU and H. ROBBINS, Complete convergence and the law of large numbers, *Proc. Nat. Acad. Sci. USA*, **33** (1947), 25—31.
[4] M. LOÈVE, *Probability theory*, Van Nostrand (New York, 1955).
[5] F. MÓRICZ, The law of the iterated logarithm and related results for weakly multiplicative systems, *Analysis Math.*, **2** (1976), 211—229.
[6] F. MÓRICZ, Moment inequalities and the strong laws of large numbers, *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **35** (1976), 299—314.
[7] P. RÉVÉSZ, *The laws of large numbers*, Academic Press (New York, 1968).
[8] R. J. SERFLING, Probability inequalities and convergence properties for sums of multiplicative random variables, *Technical Report, Florida State University, 1969.*
[9] R. J. SERFLING, Moment inequalities for the maximum cumulative sum, *Ann. Math. Statist.*, **41** (1970), 1227—1234.
[10] R. J. SERFLING, Convergence properties of S_n under moment restrictions, *Ann. Math. Statist.*, **41** (1970), 1235—1248.

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Acta Sci. Math., 38 (1976), 343—351

Spectral mapping theorems for semigroups of operators

B. NAGY

1. Introduction and notations

Spectral mapping theorems for essential spectra have been investigated by B. GRAMSCH and D. LAY [1] even in the case if T is a closed unbounded linear operator with nonvoid resolvent set. However, their results do not apply to the essential spectra of semigroups of linear operators, for in this case the mapping f is not locally holomorphic on a neighborhood of the extended spectrum of T. The aim of this paper is to extend the results of [1] to semigroups of linear operators in Banach spaces.

Let X, Y denote complex Banach spaces, $B(X, Y)$ the space of bounded linear operators from X to Y and set $B(X)=B(X, X)$. We shall always assume that the semigroup $\{T(t),\ t>0\}\subset B(X)$ is of class (A), and additional restrictions will be explicitly stated (cf. [2, pp. 321—323]). A will denote the infinitesimal generator of $T(t)$.

Let V be a closed linear operator with domain $D(V)\subset X$ and range $R(V)\subset X$. Suppose that the resolvent set $\varrho(V)$ of V is nonvoid. The nullity of V, $n(V)$ is the dimension of the kernel $N(V)$. The defect of V, $d(V)$ is the algebraic dimension of the quotient vector space $X/R(V)$. The index of V, ind (V) is $n(V)-d(V)$, where $\infty-\infty$ is undefined. The ascent of V, $a(V)$ is the smallest nonnegative integer p such that $N(V^p)=N(V^{p+1})$. The descent of V, $e(V)$ is the smallest nonnegative integer q with $R(V^q)=R(V^{q+1})$. If no such p or q exist, set $a(V)=\infty$ or $e(V)=\infty$, respectively.

A comprehensive survey of the essential spectra of V has been given in [1]. To unify notation, we shall define them by means of regularity sets G_i ($i=1, 2, \ldots, 11$) as follows. $V\in G_i$ if and only if

Received October 8, 1975.

G_1 : $V^{-1}\in B(X)$,
G_2 : ind $(V)=0$ and $a(V)=e(V)<\infty$,
G_3 : ind $(V)=0$,
G_4 : ind (V) is finite,
G_5 : $n(V)<\infty$ and $R(V)=R(P)$ for some $P=P^2\in B(X)$,
G_6 : $d(V)<\infty$ and $N(V)=R(P)$ for some $P=P^2\in B(X)$,
G_7 : $n(V)<\infty$ and $R(V)$ is closed,
G_8 : $d(V)<\infty$,
G_9 : $G_7\cup G_8$,
G_{10}: $R(V)$ is closed,
G_{11}: $a(V)<\infty$ and $e(V)<\infty$.

We shall omit nomenclature, for it is not unified in the literature. It is clear that the following relations hold:

$$\begin{array}{c} G_1\subset G_2\subset G_3\subset G_4\subset \begin{Bmatrix} G_6\subset G_8 \\ G_5\subset G_7\end{Bmatrix}\subset G_9\subset G_{10}. \\ \cap \qquad\qquad\qquad\qquad\qquad\qquad \\ G_{11} \qquad\qquad\qquad\qquad\qquad\qquad \end{array}$$

We remark that the following example shows that in general we do not have $B(X)\cap G_{11}\subset G_{10}$.

Example. Put $X=l_2$ and for $x=(x_1, x_2, \ldots)\in X$ define $Vx=(0, x_1, 0, 1/3x_3, 0, 1/5x_5, 0, \ldots)$. Then V is compact and $\dim R(V)=\infty$, hence $R(V)$ is not closed. Further $V^2=0$, thus $a(V)=2$, and $e(V)=2$, contrary to the required containment relation.

The essential spectrum $s_i(V)$ is the set of complex numbers c such that $V-c=V-cI\notin G_i$ ($i=2, 3, \ldots, 11$; for $i=1$ we get the spectrum of V). We emphasize that $s_i(V)$ is a subset of the proper complex plane ($i=1, 2, \ldots, 11$), contrary to the definitions of [1, pp. 30—31]. In what follows we intend to prove mapping theorems of the type

$$\exp[ts_i(A)]\subset s_i[T(t)],$$

which is well-known for $i=1$. Theorem 10 will show that the converse relations as a rule cannot be expected to be true.

We remark that a projection operator will always be understood to belong to $B(X)$, X^* will denote the adjoint space of X and V^* the adjoint of the operator V.

From the method of the proofs it will be seen that, according to the results of [9, pp. 285—286], some spectral mapping theorems for essential spectra of cosine operator functions can be proved by a similar method. In this connection we take the opportunity to note that Theorem 3 in [9, p. 285] has been misstated, and the following should be substituted for it. The author apologizes for the error.

Theorem 3. [9] *If C is a cosine operator function, A its generator and $s \in R$, then* $\operatorname{ch}\{s\sqrt{\sigma(A)}\} \subset \sigma\{C(s)\}$. *Further, if a is complex number and $a^2 - A$ has the spectral property P_v ($v = 1, 2, 3$), then so does* $\operatorname{ch}(as) - C(s)$.

The proof remains unchanged.

2. Spectral mapping theorems

Let $T(t)$ be a semigroup of class (A), and A its infinitesimal generator. In what follows we will heavily rely on the definitions and results of the operational calculus and spectral theory as developed in [2, Chapters 15, 16]. The most relevant results are summarized in the following lemma (cf. [2, Theorem 16.6.1]).

Lemma 1. *Let φ be a real-valued Borel measurable submultiplicative function on $[0, \infty)$ with $\varphi(0) = 1$. Suppose $g \in \mathcal{S}(\varphi)$, $A \prec \varphi$, then the linear operator $F(g; A)$ defined by*

$$F(g; A)x = \int_0^\infty T(t)x\,dg(t)$$

for $x \in X_1(A) = \{x \in X;\ \lim_{t \to 0+} T(t)x = x\}$ has a unique bounded linear extension $F(g) \in \in B(X)$. Moreover, the complex function

$$f(g; c) = \int_0^\infty e^{ct}\,dg(t)$$

is defined and holomorphic for $\operatorname{Re} c < w_0 = \lim_{t \to \infty} t^{-1} \log \varphi(t)$. *Suppose $a \in \mathcal{S}(\varphi)$, $A \prec \varphi$, $c \in s_1(A)$. Then there exist a submultiplicative φ' such that $A \prec \varphi'$, $\mathcal{S}(\varphi) \subset \mathcal{S}(\varphi')$, and an element $b \in \mathcal{S}(\varphi')$ such that*

(1) $$F(a) - f(a; c) = (A - c)\,F(b),$$

(2) $$[F(a) - f(a; c)]x = F(b)(A - c)x \quad \textit{for} \quad x \in D(A).$$

The most important special case is described in

Lemma 2. *For every $t > 0$ and c complex,*

(3) $$T(t) - e^{ct} = (A - c)\,F,$$

(4) $$[T(t) - e^{ct}]x = F(A - c)x \quad \textit{for} \quad x \in D(A),$$

where $F \in B(X)$, and for $x \in X_1(A)$

(5) $$Fx = F_c^t x = e^{ct} \int_0^t e^{-cs}\,T(s)x\,ds.$$

Further, if $T(t)$ is of class $(1, A)$, then (5) holds for every $x \in X$.

Proof. A closer inspection of the proof of [2, Theorem 16.6.1] shows that if $a=e_t$, i.e. $F(a)=T(t)$, $f(a; c)=e^{ct}$, then (1) and (2) hold for every complex c, and $F=F(b)$ is given by (5) for $x\in X_1(A)$. Moreover, if $T(t)$ is of class $(1, A)$, then the right side of (5) is defined for every $x\in X$, and

$$\left\|\int_0^t e^{-cs}T(s)x\,ds\right\| \leqq \int_0^t |e^{-cs}|\,\|T(s)\|\,ds\,\|x\| \leqq K\|x\|.$$

Thus the operators on both sides of (5) are bounded and coincide on the dense set $X_1(A)$, hence on all of X.

Remark. In what follows we will prove theorems of the following type: $c\in s_i(A)$ implies $e^{ct}\in s_i(T(t))$ for $t>0$, or equivalently,

$$T(t)-e^{ct}\in G_i \quad \text{implies} \quad A-c\in G_i \quad (t>0). \tag{6}$$

Since for every complex number c the operator $B=A-c$ is the infinitesimal generator of the semigroup $S(t)=e^{-ct}T(t)$ of the same class (see [2, pp. 357—359]), we may and will restrict ourselves in the statements and proofs to the case $c=0$ in (6). For a fixed $t>0$ we shall often write, for the sake of brevity,

$$V=T(t)-I, \quad V_0=FA. \tag{7}$$

Theorem 1. *If $T(t)-I\in G_7$, then $A\in G_7$.*

Proof. To avoid trivialities we assume $\dim X=\infty$. Since $V\supset FA$, therefore $N(A)\subset N(V)$, hence $n(A)<\infty$. By assumption, there is a projection P of X onto $N(V)$, i.e.

$$X=PX\oplus(I-P)X=N(V)\oplus X',$$

where $P\neq I$. $R(V)$ is closed, thus for $x\in X$

$$\|Vx\|\geqq q\cdot\operatorname{dist}\left(x, N(V)\right)=q\|x-n\|,$$

where $q>0$ and $n\in N(V)$. Hence

$$\|Vx\|\geqq\frac{q}{\|I-P\|}\|(I-P)x\|=q'\|(I-P)x\|.$$

For $x\in D(A)$ we get $\|F\|\cdot\|Ax\|\geqq\|Vx\|\geqq q'\|(I-P)x\|$. The equality $F=0$ would imply $N(V)=X$, a contradiction, thus for $x\in D(A)\cap X'$ we have

$$\|Ax\|\geqq r\|x\|\geqq r\cdot\operatorname{dist}\left(x, N(A|X')\right) \quad (r>0), \tag{8}$$

where $A|X'$ denotes the restriction of A to $D(A)\cap X'$. The set $X'=(I-P)X$ is a closed subspace of X, hence $A|X'$ is a closed operator. By (8), $A|X'$ has closed range, and again [3, Lemma 333] yields that $R(A)$ is closed, hence $A\in G_7$.

Theorem 2. *$T(t)-I\in G_8$ implies $A\in G_8$.*

Proof. Since $R(V)=R(AF)\subset R(A)$, we obtain codim $R(A)\leqq$ codim $R(V)$, i.e. $d(A)\leqq d(V)$, and the assertion follows immediately.

From these results we obtain the following

Corollary 1. *$T(t)-I\in G_i$ implies $A\in G_i$ $(i=4, 9)$.*

Theorem 3. *If $T(t)$ is of class $(1, A)$ and $T(t)-I\in G_{11}$, then $A\in G_{11}$.*

Proof. Since $D(V)=X$, the assumption implies that $a(V)=e(V)=p$, by [5, Theorem 5.41—E]. If $p=0$, then $1\in\varrho(T(t))$, hence $A\in G_1\subset G_{11}$, by [2, Theorem 16.7.1]. If $p>0$, then [4, Theorem 2.1] yields that 1 is a pole of the resolvent operator $R(c; T(t))$ of order p. Then there exists a deleted neighborhood U of 0 in the complex plane such that $c\in U$ implies $e^{ct}\in\varrho(T(t))$. The relations (3) and (4) yield then for $c\in U$ that

$$R(c; A) = R(e^{ct}; T(t))F_c. \tag{9}$$

Here we have emphasized that $F_c=F$ depends on c, by (5), and made use of the fact that $R(e^{ct}; T(t))$ commutes with F_c.

Since (5) holds for every $x\in X$, it is easily seen that F_c is holomorphic on the whole complex plane, and (9) gives that $R(c; A)$ is holomorphic in U. Moreover, since $\lim_{c\to 0}\left|\frac{c}{e^{ct}-1}\right|^{p+1} = t^{-p-1}>0$, there is a positive number q such that in a deleted neighborhood $U_0\subset U$ of zero

$$|c|^{p+1}\|R(c; A)\| < q|e^{ct}-1|^{p+1}\left\|R(e^{ct}; T(t))\right\|\cdot\|F_c\|. \tag{10}$$

1 is a pole of the resolvent of $T(t)$ of order p and $\|F_c\|$ is locally bounded, hence the left side of (10) converges to 0, if $c\to 0$. But then $c=0$ is a regular point or a pole of order $\leqq p$ of $R(c; A)$, and we obtain from [5, Theorem 5.8-A] that $a(A)=e(A)=$ $=m\leqq p$. Thus $A\in G_{11}$, and the proof is complete.

The following related result may also be interesting.

Theorem 4. *If $T(t)$ is of class (A) and $a(T(t)-I)=k<\infty$, then $a(A)\leqq k$.*

Proof. From (3) and (4) $AF\supset FA$, hence $F^rA^rx=V^rx$ for every $x\in D(A^r)$, $r\geqq 1$. Suppose now $x\in D(A^{k+1})$, $A^{k+1}x=0$, then $V^{k+1}x=F^{k+1}A^{k+1}x=0$. By assumption, we obtain $F^kA^kx=V^kx=0$, and we have to show that $y=A^kx=0$.

We know that $Ay=0$, and [2, Corollary 3, p. 347] yields that $T(s)y=y$ for every $s>0$, hence $y\in X_1(A)$ and, by (5), $Fy=\int_0^t T(s)y\,dx=t\cdot y$. We obtain similarly that $t^ky=F^ky=F^kA^kx=0$, hence $A^kx=0$, thus $a(A)\leqq k$.

Theorem 5. *If $T(t)$ is of class $(1, A)$ and $T(t)-I\in G_2$, then $A\in G_2$.*

Proof. By assumption, $V \in G_{11}$, hence the proof of Theorem 3 yields that $a(A)=e(A)=m<\infty$. Moreover, if $m=0$, then $A \in G_1 \subset G_2$, and if $m>0$, then 0 is a boundary point of $s_1(A)$. Supposing the latter, we also establish that $V \in G_4$, hence $A \in G_4$, by Corollary 1. Consequently, [4, Theorem 2.9] yields that $n(A)=d(A)<\infty$, hence $A \in G_2$.

Concerning the regularity set G_5, our result is not quite general. We shall call a projection $P \in B(X)$ an A-projection, if $P[D(A)] \subset D(A)$.

Theorem 6. *If $n(T(t)-I)<\infty$ and there is an A-projection P of X onto $R(T(t)-I)$, then $A \in G_5$.*

Proof. By assumption, with the notation $C=(I-P)X$ we have

$$X = PX \oplus (I-P)X = R(V) \oplus C.$$

Since P is an A-projection, we obtain

$$D(A) = [R(V) \cap D(A)] \oplus [C \cap D(A)], \tag{11}$$

where the members of the direct sum are closed sets in the induced topology of the subset $D(A) \subset X$. Since A is closed, $D(A)$ becomes a Banach space D under the norm

$$|x| = \|x\| + \|Ax\| \quad (x \in D).$$

It is easily seen that each set closed in the induced topology of $D(A)$ is also closed in D.

From (3) we see that $R(F) \subset D$, hence $R(V_0)=R(FA) \subset R(V) \cap D$. Further, if $y \in R(V) \cap D$, i.e. $y=Vx \in D$, then we can construct a sequence $\{x_k\} \subset D$ such that $V_0 x_k \xrightarrow{D} Vx$ (here $\xrightarrow{D}$ denotes convergence in D, and $\to$ will denote convergence in X). Indeed, for $k>\omega_1$ put $x_k=kR(k;A)x \in D$, then $x_k \to x$ $(k \to \infty)$, hence $V_0 x_k \to Vx$ because $V_0 \subset V \in B(X)$. On the other hand $AV_0 x_k = kAVR(k;A)x = kAR(k;A)y = kR(k;A)Ay \to AVx$, as asserted, hence $R(V_0)$ is D-dense in $R(V) \cap D$.

It is clear that $A \in B(D, X)$ and, since for $x \in X$ $|Fx| = \|Fx\| + \|Vx\| \leqq (K_1+K_2)\|x\|$, we establish that $F \in B(X, D)$. By assumption, $V \in B(X)$ has property (A) as defined by B. Yood [6, p. 600]: $R(V)$ is closed and $n(V)<\infty$. Since $AF=V$, [6, Theorem 3.5] yields that F has property (A). Since $V \in G_7$, Theorem 1 gives that A also has property (A). Since $V_0=FA$, we have $V_0 \in B(D)$, and [6, Theorem 3.4] implies that V_0 has property (A), hence $R(V_0)$ is closed in D, consequently $R(V_0)=R(V) \cap D$.

We obtain from (11)

$$D = R(V_0) \oplus [C \cap D] \tag{12}$$

where the members of the direct sum are closed sets in D. Hence there exists a projection $Q \in B(D)$ of D onto $R(V_0)$. Since $V_0=FA$, [6, Theorem 5.1] yields that there exists a projection $R \in B(X)$ of X onto $R(A)$ and $n(A)<\infty$, thus the proof is finished.

Theorem 7. *$T(t)-I \in G_6$ implies $A \in G_6$.*

Proof. By assumption, there exists a projection Q of X onto $N(V)$; here $Q\in B(X, N(V))$. An inspection of the proof of [2, Theorem 16.7.2] yields that there always exists a projection $P\in B(N(V))$ of $N(V)$ onto $N(A)$, hence $R=PQ\in B(X)$, and the range of R is $N(A)$. For every $x\in X$ we have $R^2x=PQ(PQx)=P^2Qx=Rx$, hence R is a projection of X onto $N(A)$. Further, $V\in G_8$, thus Theorem 2 implies $A\in G_6$.

Concerning the essential spectrum s_{10} we have a positive result merely in the case $A\in B(X)$. It is nevertheless remarkable, because in general there is no containment relation between $s_{10}(f(A))$ and $f(s_{10}(A))$, if f is a complex-valued function which is locally holomorphic on an open set containing $s_1(A)$, see [1, p. 29]. (In our case $f(z)=e^{tz}$.)

Theorem 8. *Suppose $T(t)$ is a uniformly continuous group of operators, i.e. $A\in B(X)$. If $T(t)-I\in G_{10}$ for some $t\neq 0$, then $A\in G_{10}$.*

Proof. Clearly we may and will assume $t>0$. Since $A\in B(X)$, thus $V=FA$ and $F(A(X))$ is closed. Since F is continuous, the inverse image $A(X)+N(F)= =R(A)+N(F)$ is closed in X. We show that $N(F)\subset R(A)$.

Let M denote $N(V)$ and, according to the proof of [2, Theorem 16.7.2], define the projections $J_r\in B(M)$ by

$$J_r x = t^{-1}\int_0^t e^{-2\pi irs/t}T(s)x\,ds \quad (x\in M).$$

Then $J_r(M)=N(A-c_r)$, where $c_r=2\pi irt^{-1}$ (r integer). Since $s_1(A)$ is compact, there is a positive integer k such that $J_r=0$ for $|r|>k$, thus formula (16.7.5) of [2, p. 468] reduces to

$$x = \sum_{r=-k}^{k} J_r x \quad \text{for} \quad x\in M. \tag{13}$$

By (5) (with $c=0$), $Fx=\int_0^t T(s)x\,ds=tJ_0x$ for $x\in M$, thus the fact that $N(F)\subset M$ implies $N(F)=N(J_0)$. Hence for $x\in N(F)$, (13) yields

$$x = \sum_{r=-k}^{k} J_r x \quad (r\neq 0). \tag{14}$$

For $r\neq 0$ we have $J_r(M)=N(A-c_r)=\{x\in X;\, x=Ac_r^{-1}x\}\subset R(A)$, thus from (14) we obtain $N(F)\subset R(A)$, hence $A\in G_{10}$ and the proof is completed.

It is remarkable that in general no similar mapping theorem holds for the essential spectrum s_3. More exactly, we have

Theorem 9. *There is a Banach space X and an $A\in B(X)$ such that $c\in s_3(A)$ for some complex c, while $e^{ct}\notin s_3(T(t))$ for some $t>0$ (here $T(t)$ denotes the group generated by A).*

Proof. For every real number s define

$$K(s) = \frac{-s^3+(1+i)s}{s^4+3}; \tag{15}$$

K is the Fourier transform of some $k \in L_1(-\infty, \infty)$, i.e.

$$K(s) = \int_{-\infty}^{\infty} e^{ist} k(t)\,dt$$

(see, e.g., [8, pp. 13—14]). Let X denote $L_p(0, \infty)$ $(p \geqq 1)$, or any other of the spaces in (6.4) of [8, p. 38]. Define $A \in B(X)$ by

$$[Ax](t) = \int_0^{\infty} k(t-s)x(s)\,ds,$$

then $\|A\| \leqq \|k\|_1$. If z is a complex number such that $z \neq K(s)$ for $-\infty \leqq s \leqq \infty$, then

$$v = v(z) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} d_s \arg(K(s)-z) = \operatorname{ind}(A-z), \tag{16}$$

moreover, $v=0$ implies $z \in \varrho(A)$, $v>0$ implies $n(A-z)=v$ and $d(A-z)=0$, while $v<0$ implies $n(A-z)=0$ and $d(A-z)=-v$ (see [8, p. 61] and [7, p. 109]).

Put $c_k = \frac{i}{8}(2k-1)$ (k integer), $r=8\pi$, then $\{c_k\}$ is the set of all complex solutions of the equation $e^{cr}=-1$. From (15) and (16) we see that because of the properties of $K(s)$

$$n(A-c_k) = \delta_{0k} \quad \text{and} \quad d(A-c_k) = \delta_{1k}, \tag{17}$$

where δ is the Kronecker symbol.

Let $T(t)$ be the group generated by A, then $T(t)$ is continuous in the uniform operator topology. [2, Theorem 16.7.2] yields that $N(T(r)+1)$ is the closed linear subspace generated by $\{N(A-c_k)\}$, hence by (17)

$$n(T(r)+1) = n(A-c_0) = 1. \tag{18}$$

From (17) we see that $R(A-c_k)$ is closed, because $d(A-c_k)<\infty$ for every k. A result of Gramsch and Lay ([1, p. 22] for σ_5 and $f(z)=e^{rz}$) then yields that $d(T(r)+1)<\infty$, hence $R(T(r)+1)$ is closed. Then we have

$$n(A^*-c_k) = d(A-c_k) \quad \text{and} \quad d(T(r)+1) = n(T(r)^*+1). \tag{19}$$

$A^* \in B(X^*)$, hence it generates the uniformly continuous group $\{T(t)^*\} \subset B(X^*)$. Applying [2, Theorem 16.7.2] now to the adjoint group, we obtain from (17) and (19) that

$$d(T(r)+1) = n(T(r)^*+1) = n(A^*-c_1) = 1, \tag{20}$$

hence, by (18), ind $(T(r)+1)=0$. From (17) we obtain ind $(A-c_0)=1$, thus $c_0 \in s_3(A)$, though $e^{c_0 r}=-1 \notin s_3(T(r))$. The proof is complete.

Remark. Some of the theorems and proofs obviously extend to the more general situation described in Lemma 1. Others apparently do not.

According to the results of GRAMSCH and LAY [1], if $A \in B(X)$, then some of the theorems above admit a converse in the well-known sense. However, we have

Theorem 10. *There exist a strongly continuous group $T(t)$ and a complex p such that $p \in s_i(T(1))$ for $i=1, 2, \ldots, 11$, whereas $c \in \varrho(A)$ for every complex c with $e^c=p$.*

Proof. We can take the example of [2, p. 469], and put $X=l_2$, $T(t)\{b_n\}= =\{e^{int}b_n\}$. Then $T(t)$ is a strongly continuous group. It is shown there that if $p \in C\sigma(T(1))$ (the nonvoid continuous spectrum), then every c is an element of $\varrho(A)$. Moreover, with the notation $U=T(1)-p$ the set $R(U)$ is not closed, hence $p \in s_i(T(1))$ for $i=1, 2, \ldots, 10$. Since U is 1—1 and $R(U) \neq X$, we have $a(U)=0$, $e(U) \neq 0$. But then $e(U)=\infty$ (see [5, pp. 272—273]), hence $p \in s_{11}(T(1))$.

References

[1] B. GRAMSCH—D. LAY, Spectral mapping theorems for essential spectra, *Math. Ann.*, **192** (1971), 17—32.

[2] E. HILLE—R. PHILLIPS, *Functional analysis and semi-groups* (Providence, 1957).

[3] T. KATO, Perturbation theory for nullity, deficiency and other quantities of linear operators, *J. Analyse Math.*, **6** (1958), 261—322.

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

[5] A. E. TAYLOR, *Introduction to functional analysis* (New York, 1958).

[6] B. YOOD, Properties of linear transformations preserved under addition of a completely continuous transformation, *Duke Math. J.*, **18** (1951), 599—612.

[7] И. Ц. Гохберг—М. Г. Крейн, Основные положения о дефектных числах, корневых числах и индексах линейных операторов, *Успехи матем. наук*, **12** (1957), вып. 2, 43—118.

[8] М. Г. Крейн, Интегральные уравнения на полупрямой с ядром, зависящим от резности аргументов, *Успехи матем. наук*, **13** (1958), вып. 5, 3—120.

[9] B. NAGY, On cosine operator functions in Banach spaces, *Acta Sci. Math.*, **36** (1974), 281—289.

DEPARTMENT OF MATH.
FACULTY OF CHEMICAL ENGINEERING
UNIVERSITY OF TECHNOLOGY
STOCZEK U. 2. H. II
1111 BUDAPEST

Acta Sci. Math., 38 (1976), 353—358

Note on an embedding theorem

J. NÉMETH

Let $\varphi\equiv\varphi_p$ $(p>1)$ be a nonnegative increasing function on $[0, \infty)$ with the following properties:

$$\frac{\varphi(x)}{x}\uparrow \quad\text{and}\quad \frac{\varphi(x)}{x^p}\downarrow \quad\text{as}\quad x\to\infty.$$

The set of measurable functions f on $[0, 1]$ for which $\int\limits_0^1 \varphi(|f(x)|)\,dx<\infty$ will be denoted by $\varphi(L)$.

If $f\in\varphi(L)$, the "modulus of continuity of f with respect to φ" will be defined by

$$\omega_\varphi(\delta; f) = \sup_{0\leqq h\leqq\delta} \bar\varphi\Big(\int\limits_0^{1-h} \varphi(|f(x+h)-f(x)|)\,dx\Big) \quad (0\leqq\delta\leqq 1),$$

where $\bar\varphi(x)$ denotes the inverse function of $\varphi(x)$. Given a function φ and a nondecreasing continuous function ω with $\omega(0)=0$, $H_\varphi^\omega\equiv H_\varphi^{\omega(\delta)}$ will denote the collection of functions $f(x)$ satisfying the condition

$$\omega_\varphi(\delta, f) = O(\omega(\delta)).$$

LEINDLER [2] gave a sufficient condition for $H_\varphi^{\omega(\delta)}\subset\varphi(L)\Lambda(L)$, where $\Lambda(x)$ is a "slowly increasing" function. Namely he proved the following:

Theorem A. ([2], Theorem 1) *Let* $f\in\varphi(L)$ $(\varphi=\varphi_p, p\geqq 1)$ *and let* $\{\lambda_k\}$ *be a nonnegative monotonic sequence of numbers such that*

$$\sum_{k=m}^{\infty}\frac{\lambda_k}{k^{1+\varepsilon}}\leqq K(\lambda)\frac{\lambda_m}{m^\varepsilon},^{1)}$$

Received September 5, 1975.

1) K and K_i denote either absolute constants or constants depending on certain functions and numbers which are not necessary to specify; $K(\alpha, \beta)$ and $K_i(\alpha, \beta)$ denote positive constants depending only on the indicated parameters. These constants are not necessarily the same at each occurrence.

where $\varepsilon=(4[p+1]+2)^{-1}$; [2]) *and let* $\Lambda(x)=\sum_{k=1}^{x}\frac{\lambda_k}{k}$. [3]) *Then*

(1) $$\sum_{n=1}^{\infty}\frac{\lambda_n}{n}\cdot\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right)<\infty$$

implies $f\in\varphi(L)\Lambda(L)$ *and*

$$\int_0^1\varphi(|f(x)|)\Lambda(|f(x)|)\,dx\leqq K(\varphi,\lambda)\left\{\sum_{n=1}^{\infty}\frac{\lambda_n}{n}\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right)+\int_0^1\varphi(|f(x)|)\,dx\right\}.$$

In the present paper we are going to prove that for certain functions $\omega(\delta)$ condition (1) is also a necessary for

$$H_\varphi^{\omega(\delta)}\subset\varphi(L)\Lambda(L).$$

More precisely, we prove the following

Theorem. *Let* $\omega(\delta)$ *be a nondecreasing, continuous function with* $\omega(0)=0$, *for which the limit*

(2) $$\lim_{h\to 0}\frac{\omega\left(\frac{h}{2}\right)}{\omega(h)}\ {}^{4)}$$

exists, and let $\{\lambda_k\}$ *be a nonnegative monotonic sequence of numbers satisfying* $\lambda_{k^2}\leqq K\lambda_k$ *for any k. Then a necessary and sufficient condition for*

(3) $$H_\varphi^{\omega(\delta)}\subset\varphi(L)\Lambda(L)$$

is that

(4) $$\sum_{n=1}^{\infty}\frac{\lambda_n\varphi\left(\omega\left(\frac{1}{n}\right)\right)}{n}<\infty,$$

where $\Lambda(x)$ *means the same as in Theorem A.*

1. We make use of the following:

Lemma ([3], Lemma 13). *Let* $A(u)$ *be a nonnegative nondecreasing function on* $[0,\infty)$ *such that* $A(u^2)\leqq KA(u)$ *for any* $u\in[0,\infty)$ *and let* $B(u)$ *be a nonnegative function on* [0, 1]. *Then*

$$\int_0^1 B(u)A(B(u))\,du<\infty\quad\textit{implies}\quad\int_0^1 B(u)A\left(\frac{1}{u}\right)du<\infty.$$

[2]) $[y]$ denotes the integral part of y.

[3]) $\sum_a^b$, where a and b are not necessarily integers, means a sum over all integers between a and b.

[4]) In the proof we shall use instead of (2) only the condition $\frac{1}{\sqrt[p]{2}}<\lim_{h\to 0}\frac{\omega(h/2)}{\omega(h)}$, where p is from the definition of the function $\varphi=\varphi_p$.

2. Proof of the Theorem

The sufficiency of (4) was proved in LEINDLER [2].

The necessity of (4) will be proved indirectly.

Suppose that

$$\sum_{n=1}^{\infty} \frac{\lambda_n \varphi\left(\omega\left(\frac{1}{n}\right)\right)}{n} = \infty. \tag{5}$$

but (3) holds. Then we can construct a function f_0 leading to a contradiction.

The construction of this function is similar to that of LEINDLER [1] made in the case $\varphi(x)=x^p$. We define $f_0(x)$ as follows:

$$f_0(x) = \begin{cases} \varrho_n, & \text{if} \quad x = 3\cdot 2^{-n-2}, \\ 0 & \text{if} \quad x = 0, \quad x\in[\frac{1}{2}, 1], \quad x = 2^{-n}, \\ \text{linear on} & [2^{-n-1}, 3\cdot 2^{-n-2}], \quad [3\cdot 2^{-n-2}, 2^{-n}], \end{cases}$$

$(n=1, 2, \ldots)$, where $\varrho_n=\bar{\varphi}\left(2^{n+1}\left(\varphi\left(\omega\left(\frac{1}{2^n}\right)\right)-\varphi\left(\omega\left(\frac{1}{2^{n+1}}\right)\right)\right)\right)$. First we show that $f_0(x)\in H_\varphi^{\omega(\delta)}$. Let

$$h\in(2^{-k-3}, 2^{-k-2}], \quad k \geqq 2. \tag{6}$$

Then

$$\int_0^{1-h} \varphi(|f_0(t+h)-f_0(t)|)\, dt = \left(\int_0^{3h} + \int_{3h}^{1-h}\right) \varphi(|f_0(t+h)-f_0(t)|)\, dt = I_1+I_2.$$

We have

$$I_1 \leqq K(\varphi) \int_0^{4h} \varphi(|f_0(x)|)\, dx \leqq K \int_0^{2^{-k}} \varphi(|f_0(x)|)\, dx \leqq$$

$$\leqq \sum_{n=k}^{\infty} \int_{2^{-n-1}}^{2^{-n}} \varphi(|f_0(x)|)\, dx \leqq K_1 \sum_{n=k}^{\infty} \varphi(\varrho_n)2^{-n-1} =$$

$$= K_1 \sum_{n=k}^{\infty}\left[\varphi\left(\omega\left(\frac{1}{2^n}\right)\right)-\varphi\left(\omega\left(\frac{1}{2^{n+1}}\right)\right)\right] = K_1\varphi\left(\omega\left(\frac{1}{2^k}\right)\right) \leqq K_2\varphi(\omega(h)).$$

Next we prove that for any k:

$$\sum_{n=0}^{k} 2^{-n}\varphi\left(\frac{2^n}{2^k}\bar{\varphi}\left(2^n\varphi\left(\omega\left(\frac{1}{2^n}\right)\right)\right)\right) \leqq K\varphi\left(\omega\left(\frac{1}{2^k}\right)\right). \tag{7}$$

To prove (7) we mention first of all that by (2) and (5)

$$\lim_{h\to 0} \frac{\omega\left(\frac{h}{2}\right)}{\omega(h)} = 1 \tag{8}$$

follows. For, if $\lim_{h\to 0}\frac{\omega(h/2)}{\omega(h)}<q<1$, then we have

$$\varphi\left(\omega\left(\frac{1}{2^{n+1}}\right)\right)\leqq q\varphi\left(\omega\left(\frac{1}{2^{n}}\right)\right)$$

which by $\lambda_{k^2}\leqq K_1\lambda_k$ implies the contrary of (5).

By (8) we may assume that there exists a positive number α such that $0<\alpha<1$ and that for any $n>n_0$

$$\omega\left(\frac{1}{2^{n-1}}\right)\leqq \sqrt[p]{2}\cdot\alpha\omega\left(\frac{1}{2^{n}}\right). \tag{9}$$

Hence by $\varphi(kx)\leqq k^p\varphi(x)$ $(k>1)$, we have

$$\varphi\left(\omega\left(\frac{1}{2^{n-1}}\right)\right)\leqq 2\alpha^p\varphi\left(\omega\left(\frac{1}{2^{n}}\right)\right), \tag{10}$$

or

$$2^{n-1}\varphi\left(\omega\left(\frac{1}{2^{n-1}}\right)\right)\leqq \alpha^p 2^n\cdot\varphi\left(\omega\left(\frac{1}{2^{n}}\right)\right). \tag{11}$$

Since $\bar\varphi(kx)\leqq\sqrt[p]{k}\,\bar\varphi(x)$ for $k\leqq 1$ we have by (11)

$$\bar\varphi\left(2^{n-1}\varphi\left(\omega\left(\frac{1}{2^{n-1}}\right)\right)\right)\leqq \alpha\cdot\bar\varphi\left(2^{n}\varphi\left(\omega\left(\frac{1}{2^{n}}\right)\right)\right), \tag{12}$$

and consequently

$$\frac{2^{n-1}}{2^k}\bar\varphi\left(2^{n-1}\varphi\left(\omega\left(\frac{1}{2^{n-1}}\right)\right)\right)\leqq \frac{\alpha}{2}\frac{2^n}{2^k}\bar\varphi\left(2^{n}\varphi\left(\omega\left(\frac{1}{2^{n}}\right)\right)\right). \tag{13}$$

Since $\varphi(kx)\leqq k\varphi(x)$ for $k\leqq 1$, we obtain by (13),

$$\varphi\left(\frac{2^{n-1}}{2^k}\bar\varphi\left(2^{n-1}\varphi\left(\omega\left(\frac{1}{2^{n-1}}\right)\right)\right)\right)\leqq \frac{\alpha}{2}\varphi\left(\frac{2^n}{2^k}\bar\varphi\left(2^{n}\varphi\left(\omega\left(\frac{1}{2^{n}}\right)\right)\right)\right).$$

Hence,

$$2^{-n+1}\varphi\left(\frac{2^{n-1}}{2^k}\bar\varphi\left(2^{n-1}\varphi\left(\omega\left(\frac{1}{2^{n-1}}\right)\right)\right)\right)\leqq \alpha\cdot 2^{-n}\varphi\left(\frac{2^n}{2^k}\bar\varphi\left(2^{n}\varphi\left(\omega\left(\frac{1}{2^{n}}\right)\right)\right)\right),$$

which implies (7), since $0<\alpha<1$.

Having (7) we can estimate I_2. Since

$$|f_0(t+h)-f_0(t)|\leqq h\cdot 2^{n+2}(\varrho_n+\varrho_{n-1})\quad\text{if}\quad 2^{-n-1}\leqq t\leqq 2^{-n},\quad 1\leqq n\leqq k-1,$$

we have

$$I_2\leqq\int_{2^{-k}}^{2^{-1}}\varphi(|f_0(t+h)-f_0(t)|)\,dt=\sum_{n=1}^{k-1}\int_{2^{-n-1}}^{2^{-n}}\varphi(|f_0(t+h)-f_0(t)|)\,dt\leqq$$

$$\leqq K(\varphi)\sum_{n=0}^{k}2^{-n}\varphi\left(\frac{2^n}{2^k}\varrho_n\right)\leqq K_1(\varphi)\sum_{n=0}^{k}2^{-n}\varphi\left(\frac{2^n}{2^k}\bar\varphi\left(2^n\varphi\left(\omega\left(\frac{1}{2^n}\right)\right)\right)\right)\leqq$$

$$\leqq K_2(\varphi)\cdot\varphi\left(\omega\left(\frac{1}{2^k}\right)\right)\leqq K_3(\varphi)\cdot\varphi(\omega(h));$$

and hence,

$$f_0(x)\in H_\varphi^\omega.$$

Finally we prove that

$$f_0(x)\notin \varphi(L)\Lambda(L).$$

By (5)

(14) $$\sum_{n=1}^{N}\frac{\lambda_n\varphi\left(\omega\left(\frac{1}{n}\right)\right)}{n}\to\infty,\quad\text{as}\quad N\to\infty.$$

Using (14) and $\lambda_{2n}\leqq K_1\lambda_n$, furthermore that for any N there exists an integer N_1 such that $\varphi\left(\omega\left(\frac{1}{N_1}\right)\right)\leqq\frac{1}{4K_1}\varphi\left(\omega\left(\frac{1}{N}\right)\right)$, an easy computation gives that

(15) $$\sum_{n=1}^{\mu}\Lambda(2^n)\varphi(\varrho_n)2^{-n}\to\infty\quad\text{as}\quad\mu\to\infty.$$

Indeed, if $2^\mu>N_1$, we have

$$\sum_{k=1}^{N}\lambda_k k^{-1}\varphi\left(\omega\left(\frac{1}{k}\right)\right)\leqq 2\sum_{k=1}^{N}\lambda_k k^{-1}\left[\varphi\left(\omega\left(\frac{1}{k}\right)\right)-2K_1\varphi\left(\omega\left(\frac{1}{N_1}\right)\right)\right]\leqq$$

$$\leqq 2\sum_{k=1}^{2^\mu}\lambda_k k^{-1}\left[\varphi\left(\omega\left(\frac{1}{k}\right)\right)-2K_1\varphi\left(\omega\left(\frac{1}{2^\mu}\right)\right)\right]\leqq$$

$$\leqq 2\left[\sum_{n=1}^{\mu}\sum_{k=2^{n-1}+1}^{2^n}\lambda_k k^{-1}\varphi\left(\omega\left(\frac{1}{k}\right)\right)-2K_1\Lambda(2^\mu)\varphi\left(\omega\left(\frac{1}{2^\mu}\right)\right)\right]+K_2\leqq$$

$$\leqq 2\left[\sum_{n=1}^{\mu}\varphi\left(\omega\left(\frac{1}{2^{n-1}}\right)\right)\sum_{k=2^{n-1}+1}^{2^n}\lambda_k k^{-1}-2K_1\Lambda(2^\mu)\varphi\left(\omega\left(\frac{1}{2^\mu}\right)\right)\right]+K_2\leqq$$

$$\leqq 2\left[\sum_{n=2}^{\mu}2K_1\varphi\left(\omega\left(\frac{1}{2^{n-1}}\right)\right)\sum_{k=2^{n-2}+1}^{2^{n-1}}\lambda_k k^{-1}-2K_1\Lambda(2^\mu)\varphi\left(\omega\left(\frac{1}{2^\mu}\right)\right)\right]+K_3\leqq$$

$$\leqq K_4\left[\sum_{i=1}^{\mu-1}\varphi\left(\omega\left(\frac{1}{2^i}\right)\right)\sum_{k=2^{i-1}+1}^{2^i}\lambda_k k^{-1}-\Lambda(2^\mu)\varphi\left(\omega\left(\frac{1}{2^\mu}\right)\right)\right]+K_3\leqq$$

$$\leqq K_4\left[\sum_{i=1}^{\mu-1}\varphi\left(\omega\left(\frac{1}{2^i}\right)\right)(\Lambda(2^i)-\Lambda(2^{i-1}))-\Lambda(2^\mu)\varphi\left(\omega\left(\frac{1}{2^\mu}\right)\right)\right]+K_5\leqq$$

$$\leqq K_4\sum_{n=1}^{\mu-1}\Lambda(2^n)\left[\varphi\left(\omega\left(\frac{1}{2^n}\right)\right)-\varphi\left(\omega\left(\frac{1}{2^{n+1}}\right)\right)\right]+K_5\leqq$$

$$\leqq K_4\sum_{n=1}^{\mu}\Lambda(2^n)\varphi(\varrho_n)\cdot 2^{-n}+K_5,$$

which proves (15) by (14).

It is clear that for any m

$$\int_{1/2^{m+1}}^{1} \varphi(|f_0(x)|)\Lambda\left(\frac{1}{x}\right)dx = \sum_{n=0}^{m} \int_{2^{-n-1}}^{2^{-n}} \varphi(|f_0(x)|)\Lambda\left(\frac{1}{x}\right)dx \geqq$$

$$\geqq \sum_{n=0}^{m} \Lambda(2^n) \int_{2^{-n-1}}^{2^{-n}} \varphi(|f_0(x)|)\,dx \geqq K_6 \sum_{n=0}^{m} \Lambda(2^n)\varphi(\varrho_n)2^{-n},$$

and thus, by (15), we get

$$\int_0^1 \varphi(|f_0(x)|)\Lambda\left(\frac{1}{x}\right)dx = \infty. \tag{16}$$

Since $\lambda_{k^2} \leqq K_1\lambda_k$, we have

$$\Lambda(u^2) \leqq K_2\Lambda(u), \tag{17}$$

thus, by (16) and applying our Lemma, we obtain

$$\int_0^1 \varphi(|f_0(x)|)\Lambda(\varphi(|f_0(x)|))\,dx = \infty. \tag{18}$$

Using (17) and the properties of the function φ, we have

$$\Lambda(\varphi(x)) \leqq K_3\Lambda(x), \tag{19}$$

whence by (18) and (19)

$$\int_0^1 \varphi(|f_0(x)|)\Lambda(|f_0(x)|)\,dx = \infty$$

follows, that is,

$$f_0 \notin \varphi(L)\Lambda(L).$$

The proof of our Theorem is completed.

References

[1] L. Leindler, On embedding of classes H_p^ω, *Acta Sci. Math.*, **31** (1970), 13—31.
[2] L. Leindler, On imbedding theorems, *Acta Sci. Math.*, **34** (1973), 231—244.
[3] П. Л. Ульянов, Вложение некоторых классов функций H_p^ω, *Изв. АН. СССР*, сер. матем.. **32** (1968), 649—686.

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1.
6720 SZEGED, HUNGARY

Acta Sci. Math., 38 (1976), 359—363

Derivations and translations on lattices

JUHANI NIEMINEN

1. Introduction. Let S be a meet-semilattice and φ a single-valued mapping of S into itself. φ is called a *meet-translation* on S [3], if $\varphi(x\wedge y)=\varphi(x)\wedge y$ for each pair of elements $x, y\in S$. If $S=L$ is a lattice and φ a single-valued mapping of L into L such that

$$\varphi(x\vee y)=\varphi(x)\vee\varphi(y) \quad \text{and} \quad \varphi(x\wedge y)=(\varphi(x)\wedge y)\vee(\varphi(y)\wedge x)$$

for each pair $x, y\in L$, then φ is called a *derivation* on L [5]. As shown by SZÁSZ in [5], a single-valued mapping on a lattice L is a derivation on L if and only if it is a meet-translation as well as an endomorphism on L.

Each meet-translation φ on S has the following properties [3]: $\varphi(x)\leqq x$, $\varphi(\varphi(x))=\varphi(x)$, and $x\leqq y\Rightarrow\varphi(x)\leqq\varphi(y)$. Moreover, in a lattice L the fixedelements of φ, i.e. the elements $t=\varphi(t)$, constitute an ideal K_φ of L [4]. As shown in [4], K_φ determines φ uniquely.

In this note we shall illuminate the dependence of φ from the properties of the ideal K_φ.

A single-valued mapping φ of a join-semilattice V into itself is called a *join-translation* on V, if $\varphi(x\vee y)=\varphi(x)\vee y$ for each pair $x, y\in V$. The results on translations in the papers [1]—[4] are given in terms of join-translations. As we shall consider here meet-translations, we always use the dual of the corresponding result obtained in the papers [1]—[4].

2. Derivations on lattices. We denote by $\mathcal{I}(L)$ the lattice of all ideals of a lattice L; $(z]=\{x|x\leqq z,\ x, z\in L\}$.

Theorem 1. *An ideal I of a lattice L generates a meet-translation φ on L such that $I=K_\varphi$ if and only if for each $y\in L$ there is an element $k\in L$ such that $I\wedge(y]=(k]$.*

Proof. If $I=K_\varphi$ for a meet-translation φ on L, then $I\wedge(y]=(\varphi(y)]$ for each $y\in L$.

Received September 25, 1975.

Conversely, let $I\wedge(y]=(k]$ for each $y\in L$. We put $\varphi(y)=k$ and show that φ is a meet-translation on L. Obviously φ is single-valued and $K_\varphi=I$. $I\wedge(x\wedge y]=(I\wedge(x])\wedge(y]$; thus $\varphi(x\wedge y)=\varphi(x)\wedge y$ and the theorem follows.

Theorem 2. *Let D be an ideal of a lattice L generating a meet-translation φ on L. Then φ is a derivation on L if and only if $D\wedge((y]\vee(x])=(D\wedge(y])\vee(D\wedge(x])$ for each pair of elements $x, y\in L$.*

Proof. As D generates a meet-translation φ on L, $D\wedge(y]=(k]$ for each $y\in L$. Let the condition of the theorem be valid for the elements $x, y\in L$. Then $D\wedge(x\vee y]=(D\wedge(x])\vee(D\wedge(y])$, whence $\varphi(x\vee y)=\varphi(x)\vee\varphi(y)$. Furthermore, $D\wedge(x\wedge y]=(D\wedge(x])\wedge(y]=(D\wedge(y])\wedge(x]=\{(D\wedge(x])\wedge(y]\}\vee\{(D\wedge(y])\wedge(x]\}$ which implies that $\varphi(x\wedge y)=(\varphi(x)\wedge y)\vee(\varphi(y)\wedge x)$.

Conversely, let φ be a derivation on L and K_φ the ideal generating it. According to the properties of φ, $K_\varphi\wedge(x]=(\varphi(x)]$. So $\varphi(x\vee y)=\varphi(x)\vee\varphi(y)$ implies that $K_\varphi\wedge(x\vee y]=(\varphi(x\vee y)]=(\varphi(x)]\vee(\varphi(y)]=(K_\varphi\wedge(x])\vee(K_\varphi(y])$. This completes the proof.

An element x of a lattice L is called *distributive*, if $x\wedge(y\vee z)=(x\wedge y)\vee(x\wedge z)$ for each pair $y, z\in L$. The following lemma shows that the condition of Theorem 2 reduces to the distributivity of D in the lattice $\mathcal{I}(L)$.

Lemma 1. *Let T be an ideal of a lattice L such that $T\wedge((x]\vee(y])=(T\wedge(x])\vee(T\wedge(y])$ for each two elements $x, y\in L$. Then $T\wedge(I\vee K)=(T\wedge I)\vee(T\wedge K)$ for each two elements $I, K\in\mathcal{I}(L)$.*

Proof. As is well known, it is sufficient to show that $T\wedge(I\vee K)\subseteq(T\wedge I)\vee(T\wedge K)$ Let $x\in T\wedge(I\vee K)$, i.e. $x\in T$ and $x\leqq i\vee k$ for some $i\in I$ and $k\in K$. Then $(x]\subseteq(i]\vee(k]$ and $x\in(x]=T\wedge(x]\subseteq(T\wedge(i])\vee(T\wedge(k])\subseteq(T\wedge I)\vee(T\wedge K)$, and the lemma follows.

The lattice $\mathcal{I}(L)$ of a modular lattice L is modular. Already the relation $T\wedge(I\vee K)=(T\wedge I)\vee(T\wedge K)$ implies the neutrality of T in a modular lattice [6, Thm. 103 and its corollary]. So we can write

Corollary 1. *A meet-translation φ on a modular lattice L is a derivation on L if and only if K_φ is a neutral element of the lattice $\mathcal{I}(L)$.*

By *the join of two derivations* φ and λ on a lattice L we mean the mapping $\varphi(x)\vee\lambda(x)$ on L and by *the meet* the mapping $\varphi(x)\wedge\lambda(x)$. In the following we consider some conditions under which the join and meet defined above are also derivations on L.

Theorem 3. *The meet of two derivations φ and λ on a lattice L is always a derivation on L. Moreover, the join of φ and λ is a derivation on L if K_φ and K_λ are neutral ideals of L.*

Proof. $(K_\varphi\wedge K_\lambda)\wedge(x]=(K_\varphi\wedge(x])\wedge(K_\lambda\wedge(x])=(\varphi(x)\wedge\lambda(x)]$ and so $K_\lambda\wedge K_\varphi$ generates a meet-translation which is $\varphi(x)\wedge\lambda(x)$. Further, $(K_\varphi\wedge K_\lambda)\wedge(x\vee y]=$

$=K_\varphi \wedge \{K_\lambda \wedge (x \vee y]\}$, and by applying now K_λ and K_φ sequently, $(K_\varphi \wedge K_\lambda) \wedge (x \vee y] = \{(K_\varphi \wedge K_\lambda) \wedge (x]\} \vee \{(K_\varphi \wedge K_\lambda) \wedge (y]\}$, whence $\varphi(x \vee y) \wedge \lambda(x \vee y) = (\varphi(x) \wedge \lambda(x)) \wedge \vee (\varphi(y) \wedge \lambda(y))$. This means that the meet of λ and φ is a join-endomorphism, too, and the first assertion follows.

Let the ideals K_φ and K_λ be neutral and let us consider the ideal $K_\varphi \vee K_\lambda$. $(K_\varphi \vee K_\lambda) \wedge (x] = (K_\varphi \wedge (x]) \vee (K_\lambda \wedge (x]) = (\varphi(x)] \vee (\lambda(x)] = (\varphi(x) \vee \lambda(x)]$. Thus the ideal $K_\varphi \vee K_\lambda$ generates a meet-translation $\beta(x) = \lambda(x) \vee \varphi(x)$ on L. The join of two neutral ideals is also a neutral ideal, and so $(K_\varphi \vee K_\lambda) \wedge (x \vee y] = \{(K_\varphi \vee K_\lambda) \wedge (x]\} \vee \{(K_\varphi \vee K_\lambda) \wedge \wedge (y]\}$. Hence $\beta(x)$ is a join-endomorphism on L and also a derivation on L.

In [5, Thm. 3] Szász has shown that the product $\varphi\lambda$ of two derivations on a lattice L is always a derivation, and moreover, $\varphi\lambda(x) = \varphi(\lambda(x)) = \varphi(x) \wedge \lambda(x)$.

As shown by Szász [5, Thm. 2], the derivations of a lattice L are exactly those meet-translations of L that are also endomorphisms on L. As immediate corollary of the construction of Kolibiar in [1, Thm. 1], we can write

Theorem 4. *On a modular lattice L there is a one-to-one correspondence between meet-translations φ and congruence relations θ_φ having the property*

(i) *There is in L a neutral ideal T such that every rest class modulo θ_φ contains exactly one element of T.*

The congruence relation θ_φ relating to the meet-translation φ and the meet-translation φ_θ relating to the congruence relation θ_φ are characterized by (ii) *and* (iii), *respectively:*

(ii) $x\theta_\varphi y \Leftrightarrow \varphi(x) = \varphi(y)$, $x, y \in L$;

(iii) $\varphi_\theta(x) = x'' \in T$ *for which* $x\theta_\varrho x''$.

Now we can prove an extension of [2, Thm. 1]

Theorem 5. *Let L be a modular lattice. The set of all congruence relations θ_φ relating to the derivations φ on L constitutes a sublattice of the lattice $\theta(L)$ of all congruence relations on L.*

Proof. According to Theorem 4, $x\theta_\varphi y \Leftrightarrow (x] \wedge K_\varphi = (y] \wedge K_\varphi$ for each derivation φ on L. As L is modular, for each derivation φ on L the ideal K_φ is a neutral element of $\mathcal{J}(L)$ (Corollary 1). Hence, for any two derivations φ and λ on L the mappings $\varphi(x) \vee \lambda(x)$ and $\varphi(x) \wedge \lambda(x)$ are derivations on L, too (Theorem 3). Let $\beta(x) = \varphi(x) \wedge \lambda(x)$. We prove $\theta_\beta = \theta_\varphi \vee \theta_\lambda$ by showing that 1) $\theta_\varphi \vee \theta_\lambda \leqq \theta_\beta$, and 2) $\theta_\varphi \vee \theta_\lambda \geqq \theta_\beta$.

1) $x\theta_\varphi y \Leftrightarrow (x] \wedge K_\varphi = (y] \wedge K_\varphi \Rightarrow (x] \wedge (K_\varphi \wedge K_\lambda) = (y] \wedge (K_\varphi \wedge K_\lambda) \Leftrightarrow x\theta_\beta y$, and so $\theta_\varphi \leqq \theta_\beta$. Similarly we see that $\theta_\lambda \leqq \theta_\beta$, whence $\theta_\varphi \vee \theta_\lambda \leqq \theta_\beta$.

2) Let $x\theta_\beta y \Leftrightarrow (x] \wedge K_\varphi \wedge K_\lambda = (y] \wedge K_\varphi \wedge K_\lambda \Leftrightarrow x \wedge \varphi(x) \wedge \lambda(x) = y \wedge \varphi(y) \wedge \lambda(y)$. On the other hand, $x \wedge \varphi(x) \theta_\lambda x \wedge \varphi(x) \wedge \lambda(x)$, and moreover, $x\theta_\varphi x \wedge \varphi(x)$. Hence,

$x(\theta_\varphi \vee \theta_\lambda)x \wedge \varphi(x) \wedge \lambda(x)$. Similarly we see that $y(\theta_\varphi \vee \theta_\lambda)y \wedge \varphi(y) \wedge \lambda(y)$, and by combining these results we obtain $x(\theta_\varphi \vee \theta_\lambda)y$. Thus $\theta_\varphi \vee \theta_\lambda \geqq \theta_\beta$.

Let $\alpha(x) = \varphi(x) \vee \lambda(x)$; we prove that $\theta_\alpha = \theta_\varphi \wedge \theta_\lambda$ by showing that 3) $\theta_\alpha \geqq \theta_\varphi \wedge \theta_\lambda$ and 4) $\theta_\alpha \leqq \theta_\varphi \wedge \theta_\lambda$.

3) Let $x(\theta_\varphi \wedge \theta_\lambda)y \Leftrightarrow x\theta_\varphi y$ and $x\theta_\lambda y \Leftrightarrow (x] \wedge K_\varphi = (y] \wedge K_\varphi$ and $(x] \wedge K_\lambda = (y] \wedge K_\lambda \Rightarrow \Rightarrow (x] \wedge (K_\varphi \vee K_\lambda) = (y] \wedge (K_\varphi \vee K_\lambda) \Leftrightarrow x\theta_\alpha y$. Thus $\theta_\alpha \geqq \theta_\varphi \wedge \theta_\lambda$.

4) Let $x\theta_\alpha y \Leftrightarrow (x] \wedge (K_\varphi \vee K_\lambda) = (y] \wedge (K_\varphi \vee K_\lambda) \Rightarrow (x] \wedge (K_\varphi \vee K_\lambda) \wedge K_\varphi = (x] \wedge K_\varphi = = (y] \wedge (K_\varphi \vee K_\lambda) \wedge K_\varphi = (y] \wedge K_\varphi$, and so $x\theta_\varphi y$. Similarly we set that $x\theta_\lambda y$, too. Consequently, $x(\theta_\varphi \wedge \theta_\lambda)y$, which implies the desired result.

A meet-translation φ on a lattice L is called a *weak derivation* on L, if $\varphi(\varphi(x) \vee y) = = \varphi(x) \vee \varphi(y)$ for each two elements $x, y \in L$.

Theorem 6. *Let M be an ideal of a lattice L generating a meet-translation φ on L. Then φ is a weak derivation on L if and only if $M \wedge ((x] \vee (y]) = (M \wedge (x]) \vee \vee (M \wedge (y])$ for each two elements $x, y \in L$ and $x \in M$.*

The proof follows the lines of that of Theorem 2, and hence we omit it. Further, the proof of the following lemma is analogous to that of Lemma 1, and hence it is omitted.

Lemma 2. *Let T be an ideal of a lattice L such that $T \wedge ((x] \vee (y]) = (T \wedge (x]) \vee \vee (T \wedge (y])$ for each two elements $x, y \in L$, $x \in T$. Then $T \wedge (I \vee K) = (T \wedge I) \vee (T \wedge K)$ for each two elements $I, K \in \mathscr{I}(L)$, $I \subseteqq T$.*

As shown by Szász [4, Thms. 4 and 5], the distributivity and modularity of a lattice L can be characterized by derivations and weak derivations of L, respectively. It is interesting to see that these characterizations reduce the distributivity (the modularity) of L to the distributivity (the modularity) of $\mathscr{I}(L)$, as one can deduce from Theorem 2 and Lemma 1, and from Theorem 6 and Lemma 2, respectively.

3. Meet-translations on meet-semilattices. In this section we shall show a connection between meet-translations on meet-semilattices and lattices. We shall consider meet-semilattices only, and hence we shall use the brief expression semilattice instead of meet-semilattice. Note that in S a nonvoid set I is an *ideal* if (i) $x \in I$ and $r \geqq x$ imply $r \in I$, and (ii) $x, y \in I$ imply $x \wedge y \in I$. S is *up-directed* if for each pair $x, y \in S$ there is an element $k \in S$ such that $k \geqq x, y$. In particular, if S is up-directed, then $I \wedge J$ is an ideal of S for each two ideals I and J of S.

Theorem 7. *Let S be an up-directed semilattice and φ a meet-translation on S. Then φ generates a meet-translation φ^g on the lattice $\mathscr{I}(L)$ of all ideals of S defined as follows: $\varphi^g(I) = \{x | x \geqq \varphi(y); y \in I \in \mathscr{I}(S)\}$.*

Proof. At first we show that $\varphi^g(I)$ is an ideal of S. Let $x \in \varphi^g(I)$ and $r \geqq x$. Then there exists an $y \in I$ such that $r \geqq x \geqq \varphi(y)$, and so $r \in \varphi^g(I)$. Let $a, b \in \varphi^g(I)$. Thus $a \wedge b \geqq \varphi(y_a) \wedge \varphi(y_b) = \varphi(y_a \wedge y_b)$, where $y_a \wedge y_b \in I$; therefore $a \wedge b \in \varphi^g(I)$.

Clearly φ^g is a single-valued mapping on $\mathscr{I}(S)$; thus it remains to show that $\varphi^g(I\wedge J)=\varphi^g(I)\wedge J$. Let $x\in\varphi^g(I\wedge J)$. Then there is an element $y\in I\wedge J$ such that $x\geqq\varphi(y)$. On the other hand, $y\geqq i\wedge j$ with some $i\in I$ and $j\in J$, and $\varphi(y)\geqq\varphi(i\wedge j)=$ $=\varphi(i)\wedge j$. Thus $x\geqq\varphi(i)\wedge j$ with $\varphi(i)\in\varphi^g(I)$ and $j\in J$, whence $x\in\varphi^g(I)\wedge J$. This shows that $\varphi^g(I\wedge J)\subseteqq\varphi^g(I)\wedge J$.

Let now $x\in\varphi^g(I)\wedge J$. Then $x\geqq r\wedge j$ for some $r\in\varphi^g(I)$ and $j\in J$. Furthermore, there exists an $i\in I$ such that $r\geqq\varphi(i)$, and so $x\geqq\varphi(i)\wedge j=\varphi(i\wedge j)$, where $i\wedge j\in I\wedge J$. Therefore, $x\in\varphi^g(I\wedge J)$, and the relation $\varphi^g(I)\wedge J\subseteqq\varphi^g(I\wedge J)$ holds. Consequently, $\varphi^g(I\wedge J)=\varphi^g(I)\wedge J$, and the theorem follows.

Let $[z)=\{x|x\geqq z,\ x, z\in S\}$. The validity of the following assertion is obvious.

Theorem 8. *A meet-translation φ on $\mathscr{I}(S)$ is generated by a meet-translation λ on S, i.e. $\varphi=\lambda^g$, if and only if for each $x\in S$ there is an element $k\in S$ such that $\varphi([x))=[k)$.*

Acknowledgement. The author wishes to thank the referee for his valuable comments and suggestions.

References

[1] M. Kolibiar, Bemerkungen über Translationen der Verbände, *Acta Fac. R. N. Univ. Comenianane Math.*, 5 (1961), 455—458.

[2] J. Nieminen, A note on relations of distributive lattices, *Acta Fac. R. N. Univ. Comenianae Math.*, to appear.

[3] G. Szász, Die Translationen der Halbverbände, *Acta Sci. Math.*, 17 (1956), 165—169.

[4] G. Szász, Translationen der Verbände, *Acta Fac. R. N. Univ. Comenianae Math.*, 5 (1961), 449—453.

[5] G. Szász, Derivations of lattices, *Acta Sci. Math.*, 36 (1975), 149—154.

[6] G. Szász, *Introduction to lattice theory*, Academic. Press and Akadémiai Kiadó (New York—London and Budapest, 1963).

DEPARTMENT OF TECHNICAL SCIENCES, FINNISH ACADEMY
LAUTTASAARENTIE 1
00200 HELSINKI 20, FINLAND

Acta Sci. Math., 38 (1976), 365—374

On the volume function of parallel sets

L. L. STACHÓ

1. Introduction

In 1959 B. SZ.-NAGY [1] proved the following theorem and its corollary:

Sz.-Nagy's Theorem. *Given an arbitrary compact set G in the plane with k connected components, if G_t denotes the parallel set of G of radius t then the function* $\text{area}(G_t)-\pi k t^2$ *is concave on* $(0, \infty)$.

Corollary. *For any bounded plane set A the function* $\text{area}\, A_t$ *is everywhere differentiable on* $(0, \infty)$ *except for a countable set of values of t. This means that the length of the parallel curves exists in the Minkowski sense for all $t>0$ outside of some countable subset of* $(0, \infty)$.

The above geometrical interpretation is based on

Pucci's Theorem. *For any subset S of Euclidean n-space E^n derivability of the function $V(t)=\text{vol}\,(S_t)$ at the point $r>0$ implies that the $n-1$ dimensional surface area of the boundary of S_r exists in the Minkowski sense and equals $V'(r)$.*

We remark that Sz.-Nagy's Theorem and its Corollary played a central role in proving the estimations of E. MAKAI [3] and L. E. PAYNE—H. F. WEINBERGER [4] for the foundamental frequency of planar membranes; [4] points also to the connections between Sz.-Nagy's Theorem and the isoperimetric theorem in 2 dimensions.

It is a natural problem to find generalizations of Sz.-Nagy's Theorem to higher dimensions that enables us to extend the Corollary and the results in mathematical physics mentioned above. The question is by no means trivial on account of difficulties of global differential geometrical type.

In the present paper we shall show in Theorem 1 of Section 2 that an inequality of M. KNESER [5] concerning parallel sets directly yields a simple integral representation of the volume function of parallel sets, which makes it possible to gen-

Received December 1, 1975.

eralize in some sense the Corollary to n dimensions and very likely opens a way of obtaining estimations concerning 3 or more dimensional vibrating bodies analogous to, but probably weaker than, those for the 2 dimensional case treated in [3] and [4].

However Theorem 1 in Section 2 does not imply the isoperimetric theorem. The main reason is the strongly local character of Kneser's inequality as shown by Lemma 5 in Section 3. Nevertheless Theorem 1 gives an idea for a new proof of less local type and a generalization of Kneser's inequality, and is suitable to extend Pucci's Theorem too. This will be the subject of Theorem 4 in Section 4 and Theorem 2 in Section 3, respectively.

2. Concavity properties of the volume function of parallel sets

Throughout this work we consider bounded subsets of E^n for an arbitrary fixed n. Let d denote the distance function[1]). Recall that the parallel set of radius t of any set A in E^n is defined by $A_t=\{p\in E^n: d(p,A)<t\}$ for $t>0$. For A fixed, the volume of A_t is a non-negative monotone increasing continuous function on $(0, \infty)$.

Our fundamental point is the following inequality

Kneser's Lemma. [5] *If* $A\subset E^n$, $b\geqq a>0$, *and* $\lambda\geqq 1$ *then*

$$\mathrm{vol}\,(A_{\lambda b}\setminus A_{\lambda a}) \leqq \lambda^n \,\mathrm{vol}\,(A_b\setminus A_a).$$

(For a new proof, also applying to a more general case, see Theorem 4 in Section 4.)

Definition. We say that a continuous function f defined on some subinterval I of $(0, \infty)$ is of *Kneser type* (or a *Kneser function*) if it satisfies

$$f(\lambda b)-f(\lambda a) \leqq \lambda^n[f(b)-f(a)] \tag{1}$$

for all $a, b\in I$ with $b\geqq a$ and for $\lambda\geqq 1$.

Lemma 1. *Let f be a Kneser function on I and let a, b be two fixed points of I with $a<b$ and $f(a)\geqq f(b)$. Then the restriction of f to the interval $[b, \infty)\cap I$ is concave and monotone decreasing.*

Proof. Let $\lambda>1$, $x_0\in I$, and $x_k=\lambda^k x_0$ for $k=1, 2, \ldots$. Examine the behaviour of f restricted to the sequence $\{x_0, x_1, \ldots\}\cap I$. Let

$$\gamma_k = [f(x_k)-f(x_{k-1})]/(x_k-x_{k-1}) \quad (k = 1, 2, \ldots).$$

[1]) I.e. for $p, q\in E^n$ and $A\subseteq E^n$ the values $d(p, q)$ and $d(p, A)$ are the distances between the points p, q and between the point p and the set A, respectively.

Then by (1) we have

$$\gamma_{k+1} \leqq \lambda^{n-1}\gamma_k \quad (k = 1, 2, \ldots).$$

In particular, if $f(x_i) \leqq f(x_{i-1})$ holds for some i then

$$0 \geqq \gamma_i \geqq \gamma_{i+1} \geqq \gamma_{i+2} \geqq \ldots.$$

This means that the function $f|_{\{x_{i-1}, x_i, x_{i+1}, \ldots\}}$ is monotone decreasing and concave. Now let $x_0 = a$ and $\lambda = (b/a)^{2^{-m}}$ for some natural number m. Since $f(a) \geqq f(b)$, there exists at least one index i with $1 \leqq i \leqq 2^m$ for which $f(x_i) \leqq f(x_{i-1})$. Therefore with the notation

$$Q_m = \{a^{j2^{-m}} b^{k2^{-m}} : j \geqq 0,\ j+k = 2^m\} \cap I \quad (m = 1, 2, \ldots)$$

we obtain that for any m the function $f|_{Q_m}$ is monotone decreasing and concave. Since $Q_1 \subseteqq Q_2 \subseteqq \ldots$ and $\bigcup_{m=1}^{\infty} Q_m$ is dense in $[b, \infty) \cap I$, we have by the continuity of f that the statement of the lemma holds.

Lemma 2. *For any Kneser function f we have that*

(i) *f is absolutely continuous,*

(ii) *$f'(t)$ exists outside of a countable subset of* dom f,

(iii) *the left and right hand side derivatives of f ($f^{(-)}$ and $f^{(+)}$) exist at every inner point of* dom f, *and $f^{(-)} \geqq f^{(+)}$,*

(iv) *$f^{(-)}$ and $f^{(+)}$ are continuous from the left and from the right, respectively.*

Proof. Let a_0 and b_0 be arbitrarily chosen inner points of dom f with $a_0 < b_0$. Clearly, it suffices to prove that the function g defined by

$$g(t) = f(t) - t^n [f(b_0) - f(a_0)]/(b_0^n - a_0^n)$$

is concave on $[b_0, \infty) \cap \operatorname{dom} f$.

Observe that $g(a_0) = g(b_0)$ and that g also satisfies (1). Then the previous lemma shows that $g|_{[b_0, \infty)}$ is concave, which completes the proof.

Theorem 1. *If f is a function of Kneser type and $a \in \operatorname{dom} f$ then there exists a monotone decreasing function α such that*

(2) $$f(t) = \int_a^t \tau^{n-1}\alpha(\tau)\,d\tau + f(a) \quad \textit{for all} \quad t \in \operatorname{dom} f.$$

Or, which is the same, there exists a concave function $\varkappa$ such that (2) *holds with $d\varkappa(\tau)$ in place of $\alpha(\tau)\,d\tau$.*

Proof. By Lemma 2 we have $f(t) - f(a) = \int_a^t f^{(+)}(\tau)\,d\tau$. Therefore the only thing we have to prove is that the function $f^{(+)}(t) \cdot t^{1-n}$ is monotone decreasing.

Let $t \in \operatorname{dom} f$, $\lambda \geqq 1$ and $h > 0$. Then (1) implies that

$$f(t+h)-f(t) \geqq \lambda^{-n}[f(\lambda t+\lambda h)-f(\lambda t)],$$

i.e.
$$[f(t+h)-f(t)]/h \geqq \lambda^{-n+1}[f(\lambda t+\lambda h)-f(\lambda t)]/(\lambda h).$$

Thus for $h \searrow 0$ we have $f^{(+)}(t) \geqq \lambda^{1-n} f^{(+)}(\lambda t)$ which estableshes $f^{(+)}(t)t^{1-n} \geqq \geqq f^{(+)}(\lambda t)(\lambda t)^{1-n}$. The proof is complete.

Remark. Relation (2) characterizes the functions of Kneser type i.e., as it can be easily seen, if any function f defined on a subinterval of $(0, \infty)$ is of the form (2), with α monotone decreasing, then f is a Kneser function.

Corollary. *For all monotone increasing Kneser functions we have*

(3)
$$f(a+\lambda y)-f(a+\lambda x) \leqq \lambda^n [f(a+y)-f(a+x)]$$

if $a+x$, $a+\lambda x$, $a+y$, $a+\lambda y \in \operatorname{dom} f$ *with* $a>0$, $\lambda \geqq 1$ *and* $y \geqq x \geqq 0$.

Proof. By Theorem 1 there exists a monotone decreasing function α such that

$$f(a+y)-f(a+x) = \int_{a+x}^{a+y} \tau^{n-1}\alpha(\tau)\,d\tau = \int_0^1 [\tau_1(\sigma)]^{n-1}\alpha_1(\sigma)(y-x)\,d\sigma$$

where $\tau_1(\sigma)=\sigma\cdot(a+y)+(1-\sigma)\cdot(a+x)$ and $\alpha_1(\sigma)=\alpha(\tau_1(\sigma))$.

Similarly, with the same function α,

$$f(a+\lambda y)-f(a+\lambda x) = \int_0^1 [\tau_2(\sigma)]^{n-1}\alpha_2(\sigma)\lambda\cdot(y-x)\,d\sigma$$

where $\tau_2(\sigma)=\sigma\cdot(a+\lambda y)+(1-\sigma)\cdot(a+\lambda x)$ and $\alpha_2(\sigma)=\alpha(\tau_2(\sigma))$.

Since $a, x, y \geqq 0$ and $\lambda \geqq 1$, we have $\tau_2(\sigma) \geqq \tau_1(\sigma)$ if $\sigma \in [0, 1]$. Therefore $\alpha_1(\sigma) \leqq \geqq \alpha_2(\sigma)$ for $\sigma \in [0, 1]$. But on the other hand we have $\alpha_1, \alpha_2, \lambda\cdot\tau_2 \geqq 0$, consequently

$$\lambda^n[\tau_1(\sigma)]^{n-1}\alpha_1(\sigma)(y-x) \geqq [\tau_2(\sigma)]^{n-1}\alpha_2(\sigma)\cdot\lambda\cdot(y-x)$$

for all $\sigma \in [0, 1]$, which implies the statement.

Lemma 3. *Let $f_k \to f_0$ be a convergent sequence of Kneser functions defined on a common interval I. Then for any $t \in I$ we have*

$$f_0^{(-)}(t) \geqq \overline{\lim}_k f_k^{(-)}(t) \geqq \underline{\lim}_k f_k^{(+)}(t) \geqq f_0^{(+)}(t).$$

Proof. The relation $\overline{\lim}_k f_k^{(-)}(t) \geqq \underline{\lim}_k f_k^{(+)}(t)$ is trivial.

Proof of $f_0^{(-)}(t) \geqq \overline{\lim}_k f_k^{(-)}(t)$: We know that the functions

(4)
$$\alpha_k(t) = f_k^{(-)}(t)t^{1-n} \quad (k = 0, 1, \ldots)$$

are monotone decreasing on I and satisfy

(5) $$f_k(t) = \int_a^t \tau^{n-1}\alpha_k(\tau)\,d\tau + f_k(a), \quad k = 0, 1, \ldots.$$

Now assume the contrary of the statement, i.e. that for some $\varepsilon > 0$ and for a subsequence $k_1, k_2, \ldots$ of subscripts we have $\lim_i \alpha_{k_i}(t) - \alpha_0(t) > \varepsilon$ for some $t \in I$. Since the left hand side derivatives of Kneser functions are continuous from the left, by the definitions of the functions α_k and since they are monotone decreasing, we obtain that there exists $\delta > 0$ such that

$$\alpha_{k_i}(\tau) \geqq \alpha_0(\tau) + \varepsilon/2 \quad \text{for} \quad \tau \in [t-\delta, t] \quad \text{and} \quad i = 1, 2, \ldots.$$

Therefore for every subscript i we have

$$[f_{k_i}(t) - f_{k_i}(t-\delta)] - [f_0(t) - f_0(t-\delta)] = \int_{t-\delta}^{t} \tau^{n-1}[\alpha_{k_i}(\tau) - \alpha_0(\tau)]\,d\tau \geqq (\varepsilon/2) \int_{t-\delta}^{t} \tau^{n-1}\,d\tau =$$
$$= \text{const} > 0$$

in contradiction to the fact that $f_k \to f_0$.

The proof of $\underline{\lim}_k f_k^{(+)} \geqq f_0^{(+)}$ goes analogously.

Lemma 4. *Suppose that $f_1, f_2, \ldots$ are Kneser functions on the domain I and suppose that the series $\sum_{k=1}^{\infty} f_k(t)$ converge for all $t \in I$. Then, if $f_0 = \sum_{k=1}^{\infty} f_k$, we have $f_0^{(+)}(t) = \sum_{k=1}^{\infty} f_k^{(+)}(t)$ and $f_0^{(-)}(t) = \sum_{k=1}^{\infty} f_k^{(-)}(t)$ for all inner points t of I.*

Remark. Since obviously f_0 is now also a Kneser function on I, the derivate numbers $f_0^{(-)}(t)$ and $f_0^{(+)}(t)$ exist for all inner points t of I.

Proof. As in the proof of Lemma 3 the functions $f_0, f_1, \ldots$ can be represented in the form (5) where $\alpha_0, \alpha_1, \ldots$ are defined by (4). Since the functions $\alpha_0, \alpha_1, \ldots$ are monotone decreasing and continuous from the left, then if $\sum_{k=1}^{\infty} \alpha_k(t)$ also exists on I the function $\beta(t) = \sum_{k=1}^{\infty} \alpha_k(t)$ is also monotone decreasing and continuous from the left, which shows by (5) that $\beta(t) = \alpha_0(t)$ in the interior of I. Now let t be any inner point of I. By our Remark and Lemma 2 we can choose a pair of points $a, b \in I$ with $a < t < b$ where $f_0'(a)$ and $f_0'(b)$ exist. Then we have

(6) $$0 \leqq \sum_{k=1}^{m} [\alpha_k(a) - \alpha_k(t)] \leqq \sum_{k=1}^{m} [\alpha_k(a) - \alpha_k(b)] \quad (m = 1, 2, \ldots).$$

On the other hand we have by Lemma 3 that $\sum_{k=1}^{\infty} \alpha_k(a)$ and $\sum_{k=1}^{\infty} \alpha_k(b)$ exist. This fact and (6) ensure the existence of $\sum_{k=1}^{\infty} \alpha_k(t)$ which completes the proof of Lemma 4.

3. An extension of Pucci's Theorem

In this and the next section we shall discuss some geometrical applications of the above results on Kneser functions. Recall that the $n-1$ dimensional Minkowski measure of any set $S\subset E^n$ is defined to equal $\lim_{t\searrow 0} \mathrm{vol}\,(S_t)/(2t)$ if this limit exists. (In the contrary case we say that S is not Minkowski measurable in $n-1$ dimensions.) We shall denote the $n-1$ dimensional Minkowski measure simply by μ.

Definition. Let X and A be subsets of E^n. We say that X is metrically associated with A if for any $p\in X$ there exists a point $q\in \bar{A}$ (the closure of A) so that $d(p, q)= =d(p, A)$ and all inner points of the straight line segment joining p with q belong to X.

Remark. It is obvious that the parallel sets of a set A are metrically associated with A. Unions and intersections of sets metrically associated with A are also metrically associated with A.

Lemma 5. *Let $A\subset E^n$ and let X be a measurable set metrically associated with A. Then the function $f(t)=\mathrm{vol}\,(A_t\cap X)$ is of Kneser type.*

Remark. We can omit the proof of Lemma 5 since its statement was essentially proved by M. Kneser ([5] p. 254).

Theorem 2. *Let A be any bounded subset of E^n. Then $\mu(\partial A_t)$ [2]) esists for all $t>0$, and denoting $V(t)=\mathrm{vol}\,(A_t)$ we have*

$$\mu(\partial A_t) = \frac{1}{2}[V^{(-)}(t)+V^{(+)}(t)].$$

Proof. It is enough to consider the case $t=1$ i.e. it suffices to see that

$$\mu(\partial A_1) = \frac{1}{2}[V^{(-)}(1)+V^{(+)}(1)].$$

Introduce the extended real valued function $h\colon E^n\to[0, \infty]$ which is defined as follows: For any point $x\in E^n$ let $h(x)$ be the least upper bound of all numbers l for which there exist points $p\in\bar{A}$ and $q\in E^n$ such that $l=d(p, q)=d(q, A)$ and the point x lies on the closed straight line segment joining p with q.

It follows directly from this definition that the inverse images $h^{-1}(a)$ for any $a\in[-\infty, \infty]$ are metrically associated with A. Furthermore, it is easy to observe that the sets $h^{-1}([a, \infty])$ are closed, and therefore if B is any Borel subset of $[-\infty, \infty]$ then $h^{-1}(B)$ is measurable and metrically associated with A.

Let us define the following functions on $(0, \infty)$:

[2]) For any set $S\subseteq E^n$ the symbol ∂S denotes its boundary.

For any Borel subset B of $[-\infty, \infty]$ let V_B be the function

$$V_B(t) = \operatorname{vol}(A_t \cap h^{-1}(B)).$$

Now by Lemma 5 we have that all the functions V_B are of Kneser type.

Next, let us examine the behavior of vol $((\partial A_1)_t)$ for $t \searrow 0$.

It is well-known that the sets $(\partial A_1)_t$ can be represented in the form

$$(\partial A_1)_t = [A_{1+t} \setminus \overline{A_{1-t}}] \setminus Y(t) \quad (t \in (0, 1))$$

where

$$Y(t) = \{p : 1 > d(p, A) > 1-t \quad \text{and} \quad d(p, \partial A_1) > t\}.$$

By Lemma 2 the only thing we have to prove is that

$$\lim_{t \searrow 0} t^{-1} \operatorname{vol}(Y(t)) = 0. \tag{6}$$

For this we only need to observe that

$$Y(t) \subseteqq h^{-1}([0, 1)) \cap (A_1 \setminus A_{1-t}) \quad \text{for} \quad t \in (0, 1). \tag{7}$$

The inclusions $Y(t) \subseteqq A_1 \setminus A_{1-t}$ are obvious. Now suppose that for some point $x \in Y(t)$ we have $h(x) \geqq 1$. This means by definition of $h(x)$ that for some $q \in E^n$ and $p \in \bar{A}$ the point x lies on the closed segment between p and q and $d(p, q) = d(q, A) \geqq 1$ holds. Therefore there is a point $\tilde{q}$ on the closed segment pq lying at a distance 1 from p, and we have

$$1 = d(\tilde{q}, A) = d(\tilde{q}, p) \tag{8}$$

$$d(\tilde{q}, x) = d(\tilde{q}, p) - d(x, p) = 1 - d(x, A) \leqq t. \tag{9}$$

But (9) contradicts the fact implied by (8) that $\tilde{q} \in \partial A_1$, since by $x \in Y(t)$ we have $d(x, \partial A_1) > t$. Thus we have proved (7).

By (7) we have

$$0 \leqq \operatorname{vol}(Y(t)) \leqq V_{[0,1)}(1) - V_{[0,1)}(1-t) = \sum_{k=1}^{\infty} \left\{ V_{\left[1-\frac{1}{k}, 1-\frac{1}{k+1}\right)}(1) - V_{\left[1-\frac{1}{k}, 1-\frac{1}{k+1}\right)}(1-t) \right\}.$$

Consequently, by Lemma 4,

$$0 \leqq \lim_{t \searrow 0} t^{-1} \operatorname{vol}(Y(t) \leqq \sum_{k=1}^{\infty} V^{(-)}_{\left[1-\frac{1}{k}, 1-\frac{1}{k+1}\right)}(1) \tag{10}$$

holds. However, any function $V_{[a, b)}$ is constant for $t > b$, therefore the right hand side of inequality (10) equals 0 which proves (6) and the theorem itself.

Beside this generalization of Pucci's Theorem we mention here as a consequence of Section 2 concerning the Minkowski measurability of the boundary of parallel sets the following approximation theorem:

Theorem 3. *Let* $\{A^k\}_{k=1}^{\infty}$ *be a sequence of non-empty bounded subsets of* E^n *tending in Hausdorff distance to a bounded set* A_0.[3]) *Then the relation* $\lim_k \mu(\partial A_t^k) = \mu(\partial A_t^0)$ *holds for all* $t \in (0, \infty)$ *except for a countable subset of* $(0, \infty)$.

[3]) The Hausdorff distance between $X, Y \subseteqq E^n$ is defined by $\inf\{\delta > 0 : X \subseteqq Y_\delta \text{ and } Y \subseteqq X_\delta\}$.

Proof. For $k=0, 1, 2, \ldots$ let $V_k(t)$ denote the volume function of the parallel sets of the set A^k and let ε_k be the Hausdorff distance of A^k from A^0. Since obviously $A^0_{t-(\varepsilon_k+1/k)} \subseteqq A^k_t \subseteqq A^0_{t+(\varepsilon_k+1/k)}$ whenever $t > \varepsilon_k + 1/k$, by the continuity of V_0 we have

$$V_k(t) \to V_0(t) \quad \text{for} \quad t > 0 \quad \text{and} \quad k \to \infty.$$

Then Lemma 3 implies that for all points t where $V'(t)$ exists,

$$\mu(\partial A^k_t) = \frac{1}{2}[V_k^{(-)}(t) + V_k^{(+)}(t)] \to V_0'(t) = \mu(\partial A^0_t)$$

holds if $k \to \infty$ which completes the proof.

4. A new proof and a generalization of Kneser's Lemma

Theorem 1 has a simple geometrical interpretation which enables us to give a new proof to Kneser's Lemma.

Let A be an arbitrary bounded subset of E^n and let $f(t) = \mathrm{vol}\,(A_t)$. We have to prove that f is a Kneser function.

Observe that it suffices to prove Kneser's Lemma for sets A consisting of merely finitely many points, since the general case can be obtained from here by the following simple approximation procedure: Choose any countable subset $\{p_1, p_2, \ldots\}$ of A, dense in A, and take the functions $f_k(t) = \mathrm{vol}\,(\{p_1, p_2, \ldots, p_k\}_t)$ $(k=1, 2, \ldots)$. Since obviously $f_k \to f$ for $k \to \infty$, we have that if $f_1, f_2, \ldots$ are functions of Kneser type then so is f too.

Thus let $A = \{p_1, \ldots, p_k\}$. In order to simplify the notations, we consider throughout this section a fixed point z as the origin of E^n and all the points p of the space E^n will be identified with the vector of the directed line segment $\overrightarrow{zp}$. Further let K^0 denote the open unit ball of centre z in E^n.

Then A_t can be written in the form of the following Minkowski sum:

$$A_t = A + tK^0 = \bigcup_{i=1}^{k} (p_i + tK^0) = \bigcup_{i=1}^{k} [D_i \cap (p_i + tK^0)] \tag{11}$$

where D_i denotes the Dirichlet cell of p_i with respect to $\langle p_1, \ldots, p_n \rangle$ i.e.

$$D_i = \{p : d(p, p_j) \geqq d(p, p_i) \;\; \text{if} \;\; j \leqq i \;\; \text{and} \;\; d(p, p_j) > d(p, p_i) \;\; \text{if} \;\; j > i\}$$
$$(i = 1, 2, \ldots, k).$$

Since $D_1, \ldots, D_k$ are pairwise disjoint convex figures (not necessarily bounded polyhedra), (11) implies that

$$\mathrm{vol}\,(A_t) = \sum_{i=1}^{k} \mathrm{vol}\,[D_i \cap (p_i + tK^0)] = \sum_{i=1}^{k} \int_0^t \mathrm{area}\,[D_i \cap \partial(p_i + \tau K^0)]\, d\tau. \tag{12}$$

Observe that any cell D_i is starshaped from the point p_i (implied by convexity of D_i), and therefore the angles consisting of the rays issued from p_i and joining p_i with the points of the figure $D_i \cap \partial(p_i+tK^0)$ on the sphere give a monotone decreasing set valued function of the variable t. Consequently, the functions α_i defined by

$$\alpha_i(t) = t^{1-n}\,\text{area}\,[D_i \cap \partial(p_i+tK^0)] \quad (i = 1, \dots, k)$$

are monotone decreasing. Thus for $\alpha(\tau)=\sum_{i=1}^{k}\alpha_i(\tau)$ we have by (12) that $f(t)= =\text{vol}\,(A_t)=\int_0^t \tau^{n-1}\alpha(\tau)\,d\tau$ which means that f is a Kneser function. Qu.e.d.

The application of Dirichlet cells enables us to extend Kneser's Lemma as follows:

Theorem 4. *Let K be an arbitrary open bounded central symmetrical convex figure of E^n and let $A \subset E^n$ be also bounded. Then the function $V(t)=\text{vol}\,(A+tK)$ is of Kneser type.*

Proof. It is easy to see that it suffices to restrict our attention to the case of $A=\{p_1, \dots, p_k\}$ as above. We may assume without any loss of generality that z is the centre of K. Introduce the function $\varrho\colon E^n\times E^n \to [0, \infty)$ defined as follows: For $x, y \in E^n$ let $\varrho(x, y)$ be equal to the unique coefficient σ for which the inclusion $y \in \partial(x+\sigma K)$ holds.

Since now we have that $(-1)K=K$, the function ϱ will be a translation invariant metric on E^n, i.e.

(13) $\varrho(x, y) = 0$ if and only if $x = y$,

(14) $\varrho(x, y)+\varrho(y, u) \geqq \varrho(x, u)$,

(15) $\varrho(x, y)+\varrho(y, u) = \varrho(x, u)$ if y belongs to the closed segment xu.

In this case it is convenient to consider

$$D_i = \{p : \varrho(p_i, p) \leqq \varrho(p, p_j) \text{ if } j \leqq i \text{ and } \varrho(p_j, p) > \varrho(p, p_i) \text{ if } j > i\}$$

$(i=1, \dots, k)$. Then for the same reason as by which (12) was obtained we have

$$V(t) = \sum_{i=1}^{k} \text{vol}\,[D_i \cap (p_i+tK)].$$

On the other hand, one can prove that any figure D_i is starshaped with respect to the point p_i.

In fact. Fix an arbitrary index i, and let $p \in D_i$, $\beta \in [0, 1]$, and $q=p_i+\beta\cdot(p-p_i)$. We have to point out that $q \in D_i$, i.e.

$$\varrho(q, p_j) \geqq \varrho(p_i, q) \quad \text{if } j \leqq i \tag{16}$$

$$\varrho(p_j, q) > \varrho(q, p_i) \quad \text{if } j > i. \tag{17}$$

Let e.g. $j \leqq i$. Then by (14) and (15) we have

(18) $$\varrho(q, p)+\varrho(p, p_j) \geqq \varrho(q, p_j),$$

(19) $$\varrho(p_i, q)+\varrho(q, p) = \varrho(p_i, p).$$

By the definition of D_i, relation $p \in D_i$ implies that

(20) $$\varrho(p_j, p) \geqq \varrho(p_i, p).$$

But (18), (19) and (20) immediately yield (16). The way to obtain (17) is similar.

Now the fact that D_i is a starshaped domain with respect to p_i can be formulated in terms of Minkowski sums as

(21) $$(1-\beta)\cdot p_i+\beta D_i \subseteqq D_i \quad \text{for any} \quad \beta \in [0, 1].$$

From here it easily follows that the function $f(t)=\operatorname{vol}[D_i \cap (p_i+tK)]$ is of Kneser type. In order to prove this let $b \geqq a \geqq 0$ and $\lambda \geqq L$. We have to see that

$$\operatorname{vol}[D_i \cap \{p_i+(\lambda bK \setminus \lambda aK)\}] \leqq \lambda^n \operatorname{vol}[D_i \cap \{p_i+(bK \setminus aK)\}].$$

For this it suffices to prove that the homothetic image of the set $D_i \cap \{p_i+(\lambda bK \setminus \lambda aK)\}$ from the point p_i with coefficient λ^{-1} is included in $D_i \cap \{p_i+(bK \setminus aK)\}$. Or which is the same, we have to prove

$$[\beta D_i+(1-\beta)p_i] \cap \{p_i+(bK \setminus aK) \subseteqq D_i \cap \{p_i+(bK \setminus aK)\}$$

for $\beta=\lambda^{-1}(\in[0, 1])$. But this is a direct corollary of (21).

Remark. It is not hard to see that no analogue of Lemma 5 holds in this generality if we replace A_t by $A+tK$ where K denotes a central symmetrical convex figure and if we replace the metric d of E^n by the metric ϱ defined in the above proof in terms of K. This fact clearly shows the essential differences between the original and the present proof of Kneser's Lemma.

References

[1] B. Sz.-Nagy, Über Parallelmengen nichtkonvexer ebener Bereiche, *Acta Sci. Math.*, **20** (1959), 36—47.

[2] C. Pucci, A proposito di un theorema riguardante la misura di involucri di insiemi, *Boll. U. M. I.*, **12** (1957), 420—421.

[3] A. Makai, Bounds for the principal frequency of a membrane and the torsional rigidity of a beam, *Acta Sci. Math.*, **20** (1959), 33—35.

[4] L. E. Payne and H. F. Weinberger, Some isoperimetric inequalities for membrane frequencies, *J. Math. Analysis and Appl.*, **2** (1961), 210—216.

[5] *M.* Kneser, Über den Rand von Parallelkörpern, *Math. Nachr.*, **5** (1951), 251—258.

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1.
6720 SZEGED, HUNGARY

Acta Sci. Math., 38 (1976), 375—378

Decomposable elements and ideals in semigroups

G. SZÁSZ

1. Introduction. An element d [or an ideal D] of a semigroup S is called *decomposable* if there exist elements a, b [ideals A, B] in S such that $d=ab$ [$D=AB$]. In particular, an ideal D of S is called *idempotent* if $D^2=D$; it is said to be *left- [right-] reproduced* if $D=SD$ [$D=DS$] and it is said to be *reproduced* if $SD=D=DS$ ([3]). A semigroup in which every element is decomposable will be called a *semigroup with decomposable elements* and the analogous terminology will be used for the semigroups in which every ideal (or principal ideal) is decomposable or reproduced, and so on (cf. [5]).

Let $\mathcal{D}_e$ [$\mathcal{D}_p$, $\mathcal{D}_i$] denote the class of semigroups with decomposable elements [principal ideals, ideals]. Then $S\in\mathcal{D}_i$ implies $S\in\mathcal{D}_p$ and the latter implies $S\in\mathcal{D}_e$, obviously. Concerning the converse implications, our earlier investigations give, as direct consequences, the following results:

(i) *$S\in\mathcal{D}_e$ implies $S\in\mathcal{D}_p$ if S is commutative* ([4], Lemma 7); [1])

(ii) *$S\in\mathcal{D}_p$ implies $S\in\mathcal{D}_i$ if S is finite and commutative* ([6], Theorem 2).

It will be shown in Section 2 that neither (i) nor (ii) remains true if we omit (any one) of the conditions written there; consequently, $\mathcal{D}_e\supset\mathcal{D}_p\supset\mathcal{D}_i$.

An ideal A of a semigroup S is called *I-pure* ([2]) if

$$(1)\qquad A\cap XS = XA \quad\text{and}\quad A\cap SX = AX$$

for any ideal X of S and it is said to be *weakly prime* if $XY\subseteqq A$ implies $X\subseteqq A$ or $Y\subseteqq A$ for each pair X, Y of ideals of S. Let $\mathcal{P}$ [$\mathcal{R}$, $\mathcal{I}$] denote the class of semigroups with *I*-pure [reproduced, idempotent] ideals. By Theorems 9—11 of [2], $\mathcal{P}\cap\mathcal{R}=\mathcal{I}$. In Section 3 we improve this result by showing $\mathcal{P}\cap\mathcal{D}_e=\mathcal{I}$. Finally, in Section 4 we prove that any weakly prime decomposable ideal is reproduced at least from one side.

For the notations and concepts not defined here, see [1].

Received December 1, 1975, revised March 25, 1976.

[1]) The analogous problem for prime ideals has been solved in [3], Satz 1.

2. On the classes $\mathscr{D}_e$, $\mathscr{D}_p$ and $\mathscr{D}_i$ of semigroups. The following example was constructed by András Botos (Szeged). Consider the semigroup S generated by the set $\{g_0, g_1, g_2, g_3\}$ and subject to the generating relations

$$g_0 = g_0^2, \quad g_1 = g_1 g_3, \quad g_2 = g_3 g_2, \quad g_3 = g_3^2.$$

S is obviously decomposable. Let A and B be any ideals of S such that $J(g_1 g_2) \subseteqq AB$. Then $g_1 g_2 \in AB$ and the generating relations imply $g_1 \in A$, $g_2 \in B$. It follows that $g_1 g_0 g_2 \in AB$, too. But $g_1 g_0 g_2 \notin J(g_1 g_2)$ whence $J(g_1 g_2) \subset AB$. Thus we have got

Proposition 1. *There exist semigroups with decomposable elements that are not semigroups with decomposable principal ideals.*

It remains to solve the problem whether the class $\mathscr{D}_e \setminus \mathscr{D}_p$ contains also finite semigroups or not.

Let C denote the additive semigroup of all complex numbers $a+bi$ with $a \geqq 0$, $b \geqq 0$ and $a+b \neq 0$. Then every element and, consequently, every principal ideal of C is decomposable. The set

$$I = \{u+vi : u \geqq 1 \text{ or } v \geqq 1\}$$

is an ideal of C. Let A and B be ideals of C such that $I \subseteqq A+B$. Then $1 \in A+B$. Since the number 1 can be decomposed in C only into the sum of two positive real numbers less than 1, there exists an $a_0 \in A$ with $a_0 < 1$. Similarly, $i \in A+B$ implies the existence of a $b_0 i \in B$ with $b_0 < 1$. It follows that $A+B$ contains an element $a_0 + b_0 i$ of C with $a_0, b_0 < 1$. Hence $I \subset A+B$ and we have got

Proposition 2. *There exist (infinite) commutative semigroups with decomposable principal ideals that are not semigroups with decomposable ideals.*

Consider, finally, the semigroup $F = \{0, a, b, c\}$ in which

$$bc = b, \quad ca = a, \quad cc = c \quad \text{and} \quad xy = 0 \quad \text{for any other pairs} \quad x, y \in F.$$

It is a semigroup with decomposable principal ideals:

$$J(0) = \{0\} = J(0) \cdot J(0), \qquad J(b) = \{0, b\} = J(b) \cdot F,$$

$$J(a) = \{0, a\} = F \cdot J\{a\}, \qquad J(c) = F = F^2.$$

The set $P = \{0, a, b\}$ is an ideal of F, too. Let A, B be any ideals of F such that $P \subseteqq AB$. Then $a \in AB$ and $b \in AB$, implying $c \in A$ and $c \in B$, respectively. It follows, by $J(c) = F$, that $A = B = F$. Hence $P \subset AB$ and we have proved:

Proposition 3. *There exist (non-commutative) finite semigroups with decomposable principal ideals that are not semigroups with decomposable ideals.*

Remark. A semigroup N with 0 is called *nilpotent* if there exists a positive integer r such that $N^r=\{0\}$. Let S be a semigroup with decomposable elements. Then $S=S^2=S^3=\dots$. It follows that S cannot be nilpotent if $|S|>1$.

3. I-pure ideals in semigroups with decomposable elements. In order to improve the result

$$\mathscr{P}\cap\mathscr{R}=\mathscr{I}, \tag{2}$$

mentioned in the introduction, we begin with

Theorem 1. *Any I-pure ideal of a semigroup with decomposable elements is idempotent.*

Proof. Let A be an I-pure ideal of the semigroup. Applying the first equation in (1) for $X=S$ and the second one for $X=A$ we get $A\cap S^2=SA$ and $SA=A^2$, i.e.

$$A\cap S^2=A^2$$

(without making any restriction for S). If, in particular, $S^2=S$, then $A=A\cap S= =A\cap S^2=A^2$. Thus the theorem is proved.

Remark. Zero semigroups Z with $|Z|>1$ furnish trivial examples for semigroups in which every ideal is I-pure but none of the elements except the 0 is decomposable.

Theorem 2. *The classes $\mathscr{P}$, $\mathscr{D}_e$ and $\mathscr{I}$ of semigroups satisfy the equation $\mathscr{P}\cap\mathscr{D}_e=\mathscr{I}$.*

Proof. Clearly, $\mathscr{P}\cap\mathscr{D}_e\subseteqq\mathscr{I}$ by Theorem 1. Thus, (2) implies $\mathscr{I}=\mathscr{P}\cap\mathscr{R}\subseteqq \subseteqq\mathscr{P}\cap\mathscr{D}_i\subseteqq\mathscr{P}\cap\mathscr{D}_e\subseteqq\mathscr{I}$, i.e. $\mathscr{I}=\mathscr{P}\cap\mathscr{D}_e$, as asserted.

4. On decomposable ideals. In this section we prove

Theorem 3. *Let A be a decomposable ideal of a semigroup S. If A is weakly prime, too, then it is left- or right-reproduced.*

Proof. Let X, Y be ideals of S such that $A=XY$. Then $A\subseteqq X$ and $A\subseteqq Y$. If A is weakly prime, too, then at least one of the converse inclusions $X\subseteqq A$ and $Y\subseteqq A$ is true as well. In the first case $X=A$, whence we get

$$AS\subseteqq A=AY\subseteqq AS,$$

i.e. $A=AS$. Similarly, in the second case we get $A=SA$.

References

[1] A. H. CLIFFORD—G. B. PRESTON, *The algebraic theory of semigroups*. I, Math. Surveys **7** (Providence, 1961).

[2] N. KUROKI, Note on purity of ideals on semigroups, *Comm. Math. Univ. St. Pauli*, **22** (1973), 13—17.

[3] G. SZÁSZ, Über Halbgruppen, die ihre Ideale reproduzieren, *Acta Sci. Math.*, **27** (1966), 141—146.

[4] G. SZÁSZ, Eine Characteristik der Primidealhalbgruppen, *Publ. Math.*, **17** (1970), 209—214.

[5] G. SZÁSZ, On semigroups with T-ideals, *Mat. Vesnik*, **11** (26) (1974), 293—296.

[6] G. SZÁSZ, On semigroups in which $a \in Sa^k S$ for any element, *Math. Japonicae*, **20** (1976), 283—284.

TECHNOLOGICAL UNIVERSITY
1111 BUDAPEST
MŰEGYETEM RKP. 3
„H" ÉPÜLET

Acta Sci. Math., 38 (1976), 379—381

Unitary subsemigroups in commutative semigroups

G. SZÁSZ

1. Introduction. We use the terminology and notations of [1]. In particular, a subset U of a semigroup S will be called *left [right] unitary* if, for each $u \in U$ and $s \in S$, $us \in U$ $[su \in U]$ implies $s \in U$; a subset which is both left and right unitary will be called *unitary*.

In this paper we deal only with commutative semigroups. Clearly, the terms "left unitary", "right unitary" and "unitary" have the same meaning in this case.

2. Connections with a special congruence relation. Let S be a commutative semigroup and R a subsemigroup of S. Define $a\varrho_R b$ $(a, b \in S)$ to mean that there exists an $x \in R$ such that $ax=bx$. It is well-known that ϱ_R is a congruence on S. T. TAMURA and H. B. HAMILTON discussed in [4] the case when R is cofinal in S (that is, to each $s \in S$ there exists an $r \in R$ such that $sr \in R$). A part of their results can be formulated as follows: *If R is a cofinal subsemigroup of the commutative semigroup S, then*

(i) *R is included in a ϱ_R-class (i.e., $x\varrho_R y$ for each $x, y \in R$), but*

(ii) *R is itself a ϱ_R-class if and only if it is unitary.*

Now we show that (i) and (ii) remain true if cofinality is replaced by the condition that R is a subsemilattice of S. We recall that a *semilattice* is a commutative semigroup every element of which is idempotent.

Theorem 1. *Let S be a commutative semigroup and R a subsemilattice of S. Then $x\varrho_R y$ for each pair $x, y \in R$.*

Proof. For any elements x, y of R we have $x \cdot xy = y \cdot xy$ and $xy \in R$. Hence $x\varrho_R y$ indeed.

Before formulating the analogue of (ii) we prove a more general proposition:

Theorem 2. *Let S be a commutative semigroup and R a unitary subsemigroup in S. Then $u\varrho_R a$ $(a \in R)$ implies $u \in R$ (i.e., R is the union of some ϱ_R-classes).*

Received December 1, 1975, revised February 15, 1965.

Proof. Let $a \in R$, $u \in S$ and $u \varrho_R a$. Then there exists an $x \in R$ such that $xu = xa \in R$. Since R is unitary, $u \in R$.

Theorem 3. *Let S be a commutative semigroup and R a subsemilattice of S. Then R is a ϱ_R-class if and only if it is unitary.*

Proof. If R is unitary, then it is a ϱ_R-class by Theorems 1 and 2. Conversely, suppose that R is a ϱ_R-class and $ax=b$ with $a, b \in R$. Then $ax = a^2x = ab$ and therefore $x \varrho_R b$. Since R is a ϱ_R-class, we conclude that $x \in R$. This means that R is unitary, indeed.

3. Unitary subsemilattices in semilattices. A subsemilattice F of a semilattice S is called a *filter* if, for any elements $e \in F$ and $s \in S$, $es=e$ implies $s \in F$. By the following theorem the filters and the unitary subalgebras will be identified in semilattices:

Theorem 4. *The following assertions concerning a subsemilattice R of a semilattice S are equivalent:*

(A) *R is a ϱ_R-class;*

(B) *R is a filter;*

(C) *R is unitary.*

Proof. Since (A) and (C) are equivalent by Theorem 3, we have only to show that (B) and (C) are also equivalent.

Let $ax=b$ with $a, b \in R$. Then $b = ax^2 = bx$. Assuming (B), we get $x \in R$. This means that (B) implies (C).

Let $a = as$ with $a \in R$, $s \in S$. Assuming (C), we get $s \in R$. This means that (C) implies (B), too.

In the rest of this paper we point to a prominent role of unitary subsemilattices. Let S and Σ be semilattices with identity elements e and ε, respectively. Let, further, a^b $(a, b \in S)$ denote a mapping of $S \times S$ into Σ. Define a multiplication in $S \times \Sigma$ by the rule

$$(1) \qquad (a, \alpha) \circ (b, \beta) = (ab, a^b \alpha\beta).$$

The resulting grupoid, denoted by $S \circ \Sigma$, is a (degenerated) Rédeian skew product of S and Σ in the sense of [2]. It was shown in [3] that $S \circ \Sigma$ is a semilattice if and only if $a^b = b^a$ and

$$(2) \qquad a^a = \varepsilon$$

for each $a, b \in S$. Now we prove

Theorem 5. *Let S and Σ be semilattices with the identity elements e and ε, respectively. If their Rédeian skew product $S \circ \Sigma$ is a semilattice, too, then the set*

$$\Gamma = \{(e, \alpha):\ \alpha \in \Sigma\}$$

is a subsemilattice of $S \circ \Sigma$ such that

(i) *Γ is unitary and isomorphic with Σ;*
(ii) *$S \circ \Sigma / \varrho_\Gamma$ is isomorphic with S.*

Proof. By (1), Γ is a subalgebra of $S \circ \Sigma$. Property (i) can be derived immediately from (1) and (2). As for (ii), $(a, \alpha)\, \varrho_\Gamma\, (b, \beta)$ means that there exists an (e, γ) such that $(a, \alpha) \circ (e, \gamma) = (b, \beta) \circ (e, \gamma)$ which implies $a = b$. Conversely, $a = b$ implies $(a, \alpha)\, \varrho_\Gamma (b, \beta)$ for arbitrary $\alpha, \beta \in \Sigma$ because $(a, \alpha) \circ (e, \alpha\beta) = (a, a^e \alpha\beta) = (b, b^e \alpha\beta) = (b, \beta) \circ (e, \alpha\beta)$ in this case. Thus (ii) is proved, too.

References

[1] A. H. Clifford—G. B. Preston, *The algebraic theory of semigroups.* I—II, Math. Surveys 7 (Providence, 1961, 1967).

[2] R. Kochendörffer, Zur Theorie der Rédeischen schiefen Produkte, *J. reine u. angew. Math.*, **192** (1953), 96—101.

[3] G. Szász, Rédeische schiefe Produkte von Halbverbänden, *Acta Math. Acad. Sci. Hung.*, 7 (1957), 441—461.

[4] T. Tamura—H. B. Hamilton, Commutative semigroups with greatest group-homomorphism, *Proc. Japan Acad.*, 47 (1971), 671—675.

TECHNOLOGICAL UNIVERSITY
1111 BUDAPEST
MŰEGYETEM RKP. 3
„H" ÉPÜLET

Acta Sci. Math., 38 (1976), 383—398

Sur la connexion naturelle à torsion nulle

J. SZENTHE

L'étude systématique des espaces homogènes réductifs et de leurs connexions invariantes a été lancée par K. NOMIZU dans son travail fondamental [7]. Parmi les possibles connexions affines invariantes d'un espace homogène réductif, la connexion naturelle à torsion nulle est d'importance particulière pour ses propriétés favorables. Soit en effet $M=G/H$ un espace homogène, $\pi: G\to M$ la projection canonique et $\mathfrak{g}=\mathfrak{m}\oplus\mathfrak{h}$ une structure réductive de M. Une connexion affine invariante à torsion nulle de M est dite naturelle, si toutes les trajectoires d'origine $o=\pi(H)$ des sous-groupes à 1 paramètre qui sont définis par éléments de $\mathfrak{m}$ sont des géodésiques de la connexion ([5], II. p. 197—200). En général, cette définition n'est pas simplifiable. En effet, il y a des espaces homogènes qui n'admettent pas des structures réductives, mais qui ont des connexions affines invariantes à torsion nulle dont toutes les géodésiques sont des trajectoires ([8] et [4], p. 102—115). Le but de ce travail est de montrer qu'une simplification de la définition est pourtant possible dans un cas important. En effet, le théorème suivant sera prouvé en supposant quelque condition de differentiabilité: *Soit G un groupe de Lie connexe, $H\subset G$ un sous-groupe compact connexe et $M=G/H$ l'espace homogène correspondant. Soit donnée une connexion affine invariante à torsion nulle de M telle que toutes ses géodésiques sont des trajectoires. Il y a alors une structure réductive de M telle que la connexion donnée est sa connexion naturelle.* Pour démontrer ce théorème quelques préparations semblent être convenables. Donc une classification des trajectoires d'un espace homogène sera donnée d'abord et quelques observations générales seront faites ensuite sur les connexions affines invariantes dont toutes les géodésiques sont des trajectoires. Le théorème ci-dessus sera une conséquence directe des résultats ainsi obtenus.

Reçu le 28. janvier 1975, revu le 13. février 1976.

1. Classification des trajectoires

Soient G un groupe de Lie connexe, H. $H \subset G$ un sous-groupe fermé connexe et $M = G/H$ l'espace quotient correspondant formé par les classes à gauche de H dans G. Soit $\pi: G \to M$ la projection canonique et $\alpha: G \times M \to M$ l'action naturelle de G sur M. On considère M muni de la structure unique de variété analytique pour que π et α sont des applications analytiques. L'algèbre de Lie $\mathfrak{g}$ de G sera identifiée avec l'espace tangent $T_e G$ de G en l'élément neutre $e \in G$ et par conséquent l'algèbre de Lie $\mathfrak{h}$ de H sera identifiée avec le sous-espace correspondant de $T_e G$.

Si $\varphi: R \to G$ est un sous-groupe à 1 paramètre et si $m \in M$, on appelle l'application $\tau \mapsto \alpha(\varphi(\tau), m)$, $\tau \in \mathbf{R}$ la *trajectoire d'origine* m *de* φ. En particulier, si $o = \pi(e)$, on a évidemment $\alpha(\varphi(\tau), o) = \pi \circ \varphi(\tau)$, $\tau \in \mathbf{R}$. Une trajectoire est banale, si $m = \alpha(\varphi(\tau), m)$ pour tout $\tau \in \mathbf{R}$.

Étant donnée une trajectoire non-banale $\pi \circ \varphi$ d'origine o, il existe évidemment un $\varepsilon > 0$ tel que $\pi \circ \varphi$ est injective sur $[-2\varepsilon, 2\varepsilon]$ et par suite $\pi \circ \varphi([-\varepsilon, \varepsilon]) = C \subset \subset C' = \pi \circ \varphi([-2\varepsilon, 2\varepsilon])$ sont des arcs de M. Soit $L(\varphi; \varepsilon)$ l'ensemble des éléments $g \in G$ tels que si $\psi: \mathbf{R} \to G$ est un sous-groupe à 1 paramètre avec $g = \psi(\tau_0)$ les éléments $\psi(\tau) \in G$ pour $|\tau| \leqq |\tau_0|$ transforment C en un arc qui est contenu dans l'intérieur de C'. Soit ensuite $\mathfrak{B}$ le filtre des voisinages de e dans G et $\mathfrak{B}'$ le système des ensembles $V \cap L(\varphi; \varepsilon)$ où $V \in \mathfrak{B}$.

Proposition 1. *L'ensemble P des éléments de G, qui sont engendrés par éléments de $L(\varphi; \varepsilon)$ est un sous-groupe de G. Il existe exactement une topologie sur P qui rend P un groupe topologique et $\mathfrak{B}'$ une base de filtre des voisinages de e dans P.*

Démonstration. Il résulte de la définition de $L(\varphi; \varepsilon)$ que $L(\varphi; \varepsilon)^{-1} = L(\varphi; \varepsilon)$. Alors, l'ensemble P des éléments de G qui sont engendrés par éléments de $L(\varphi; \varepsilon)$ est un sous-groupe de G.

La deuxième assertion de la proposition sera prouvée en montrant que les conditions pour une base de filtre ([3], p. 4—5) sont satisfaites par $\mathfrak{B}'$.

1. Quel que soit $U' \in \mathfrak{B}'$, il existe $V' \in \mathfrak{B}'$ tel que $V' \cdot V' \subset U'$. En effet $U' = U \cap L(\varphi; \varepsilon)$ où $U \in \mathfrak{B}$ et par conséquent il y a un $\tilde{V} \in \mathfrak{B}$ tel que $\tilde{V} \cdot \tilde{V} \subset U$. En vertu de sa définition $L(\varphi; \varepsilon)$ est l'union de sours-arcs de sous-groupes à 1 paramètre et par suite il a y un voisinage W de e dans G formé également par sous-arcs de sous-groupes à 1 paramètre et tel que $L(\varphi; \varepsilon) = W \cap L(\varphi; \varepsilon)$. On voit facilement que W peut être tellement choisi que $L(\varphi; \varepsilon) = W \cap P$ soit aussi valable. Soit $\tilde{W}$ un voisinage de e dans G tel que $\tilde{W} \cdot \tilde{W} \subset W$. Alors on a $(\tilde{W} \cap L(\varphi; \varepsilon)) \cdot (\tilde{W} \cap L(\varphi; \varepsilon)) \subset W \cap P = L(\varphi; \varepsilon)$. Il en résulte que pour $V = \tilde{V} \cap \tilde{W}$ et pour $V' = V \cap L(\varphi; \varepsilon)$ on a $V' \cdot V' \subset U \cap L(\varphi; \varepsilon) = U'$.

2. Quel que soit $U' \in \mathfrak{B}'$, il existe $V' \in \mathfrak{B}'$ tel que $V'^{-1} \subset U'$. En effet, $U' = U \cap L(\varphi; \varepsilon)$ où $U \in \mathfrak{B}$ et par suite il y a un voisinage symétrique V de e dans G tel que $V \subset U$. Alors, pour $V' = V \cap L(\varphi; \varepsilon)$ on a $V'^{-1} = V' \subset U'$.

3. L'élément neutre e appartient évidemment à tout ensemble de $\mathfrak{B}'$.

4. Quels que soient $a \in P$ et $U' \in \mathfrak{B}'$, il existe $V' \in \mathfrak{B}'$ tel que $V' \subset aUa'^{-1}$. Parce que $U' = U \cap L(\varphi; \varepsilon)$ où $U \in \mathfrak{B}$, il y a un $\tilde{V} \in \mathfrak{B}$ tel que $\tilde{V} \subset aUa^{-1}$. Si $\psi: \mathbf{R} \to G$ est un sous-groupe à 1 paramètre qui a un sous-arc appartenant à $L(\varphi; \varepsilon)$, il existe évidemment un $\tau_0 > 0$ tel que $a^{-1}\psi(\tau)a \in L(\varphi; \varepsilon)$ et $\psi(\tau) \in L(\varphi; \varepsilon)$ pour $|\tau| \leqq |\tau_0|$. Soit L' l'union des tels éléments $\psi(\tau)$ pour tous les sous-groupes ψ. On a alors $L' \subset aL(\varphi; \varepsilon)a^{-1}$. D'autre part, on voit facilement que il y a un voisinage $\tilde{W}$ de e dans G tel que $\tilde{W} \cap L(\varphi; \varepsilon) = L'$. Soit $V = \tilde{V} \cap \tilde{W}$ et $V' = V \cap L(\varphi; \varepsilon)$. Par conséquent, on a $V' = \tilde{V} \cap \tilde{W} \cap L(\varphi; \varepsilon) \subset \tilde{V} \cap L' \subset aUa^{-1} \cap aL(\varphi; \varepsilon)a^{-1} = aU'a^{-1}$.

Proposition 2. *P est un sous-groupe de Lie connexe de G.*

Démonstration. Étant évidemment un groupe localement compact connexe qui n'a pas de sous-groupes petits, P est un groupe de Lie. Comme les sous-groupes à 1 paramètre de P sont aussi eaux de G, on voit que P est un sous-groupe de Lie de G.

Les deux propositions précédentes nous conduisent à une notion fondamentale. En effet, étant donnée une trajectoire non-banale on appelle *P le sous-groupe correspondant à la trajectoire $\pi \circ \varphi$ dans G*. Si le sous-groupe à 1 paramètre $\varphi: \mathbf{R} \to G$ est défini par $X \in \mathfrak{g} - \mathfrak{h}$, l'algèbre de Lie de P qui est une sous-algèbre de $\mathfrak{g}$ sera notée par $\mathfrak{p}_X$. On peut étendre la définition ci-dessus au cas général. En effet, soit $H_m \subset G$ le sous-groupe d'isotropie en $\pi(g) = m \in M$ et soit $\pi_m: G \to G/H_m$ la projection canonique correspondante. En identifiant G/H_m avec M on a $\alpha(\varphi(\tau), m) = \alpha(\varphi(\tau), \pi(g)) = \pi(\varphi(\tau)g) = \pi(\varphi(\tau)gH) = \pi_m(\varphi(\tau)gHg^{-1}) = \pi_m \circ \varphi(\tau)$, $\tau \in \mathbf{R}$. Le sous-groupe correspondant à $\pi_m \circ \varphi$ sera appelé *le sous-groupe correspondant à la trajectoire*

$$\tau \mapsto \alpha(\varphi(\tau), m), \ \tau \in \mathbf{R} \text{ dans } G.$$

Lemme 1. *Si $P \subset G$ est le sous-groupe correspondant à la trajectoire $\pi \circ \varphi$, on a $\pi(P) = \{\pi \circ \varphi(\tau) | \tau \in \mathbf{R}\}$.*

Démonstration. Il est évident que φ est un sous-groupe à 1 paramètre de P et en conséquence on a $\{\pi \circ \varphi(\tau) | \tau \in \mathbf{R}\} \subset \pi(P)$. Par contre, si $g \in P$, il existe un sous-groupe à 1 paramètre de P et en conséquence on a $\{\pi \circ \varphi(\tau) | \tau \in \mathbf{R}\} \subset \pi(P)$. Par contre, si $g \in P$, il existe un sous-groupe à 1 paramètre $\psi: \mathbf{R} \to P$ tel que $g = \psi(\xi_0)$ pour un $\xi_0 \in \mathbf{R}$. En vertu d'assertions ci-dessus il y a un $\delta > 0$ tel que $\psi(\xi) \in L(\varphi; \varepsilon)$ pour $|\xi| \leqq \delta$. Par suite, $\pi \cdot \psi(\xi) = \alpha(\psi(\xi), 0) \in C' \subset \{\pi \circ \varphi(\tau) | \tau \in \mathbf{R}\}$ pour $|\xi| \leqq \delta$. Alors, $\pi(g) \in \{\pi \circ \psi(\xi) | \xi \in \mathbf{R}\} \subset \{\pi \circ \varphi(\tau) | \tau \in \mathbf{R}\}$ en vertu de l'analycité de $\pi \circ \varphi$ et de $\pi \circ \psi$.

Soit $T_e\pi\colon \mathfrak{g}\to T_0M$ l'application linéaire tangente à π en e et $L\subset T_0M$ un sous-espace de dimension 1, alors $\mathfrak{k}_L=\{Y|T_e\pi Y\in L,\ Y\in\mathfrak{g}\}$ est un sous-espace de $\mathfrak{g}$. Si $X\in\mathfrak{g}-\mathfrak{h}$ et $T_e\pi X\in L$, on a $\mathfrak{p}_X\subset\mathfrak{k}_L$ en vertu du lemme précédent.

Lemme 2. *Soit $X\in\mathfrak{k}_L-\mathfrak{h}$ et soit $\mathfrak{a}$ une sous-algèbre de $\mathfrak{g}$ telle que $X\in\mathfrak{a}\subset\mathfrak{k}_L$. On a $\mathfrak{a}\subset\mathfrak{p}_X$.*

Démonstration. Soit $A\subset G$ le sous-groupe connexe défini par $\mathfrak{a}$ et soit fixé un système de coordonnées canoniques de la deuxième espèce ([9], p. 302—307) sur un voisinage V de e dans A de facon que $g=\varphi(\tau_0)\zeta_1(\tau_1)\dots\zeta_k(\tau_k)$ pour $g\in V$, où $(\tau_0,\tau_1,\dots,\tau_k)$ sont les coordonnées de g et $\zeta_1,\dots,\zeta_k\colon\mathbf{R}\to A\cap H$ sont des sous-groupes à 1 paramètre qui définissent le système de coordonnées. Si $\varphi(\tau)$, $g\in V$ sont tels que $\zeta_1(\tau_1)\dots\zeta_k(\tau_k)\varphi(\tau)\in V$, on a $\zeta_1(\tau_1)\dots\zeta_k(\tau_k)\varphi(\tau)=\varphi(\tau')\zeta_1(\tau_1')\dots\zeta_k(\tau_k')$ et par conséquent $\alpha(g,\pi\circ\varphi(\tau))=\pi(\varphi(\tau_0)\zeta_1(\tau_1)\dots\zeta_k(\tau_k)\varphi(\tau))=\pi(\varphi(\tau_0)\varphi(\tau')\zeta_1(\tau_1')\dots\zeta_k(\tau_k'))=$ $=\pi\circ\varphi(\tau_0+\tau')$. Cela montre qu'il y a un voisinage $V'\subset V$ de e dans A tel que $V'\subset$ $\subset L(\varphi;\varepsilon)$. Alors, $A\subset P$ et par conséquent $\mathfrak{a}\subset\mathfrak{p}_X$.

Corollaire. *Soient $\mathfrak{a}',\mathfrak{a}''\subset\mathfrak{k}_L$ sous-algèbres de $\mathfrak{g}$ qui sont maximales dans $\mathfrak{k}_L$ mais ne sont pas des sous-algèbres de $\mathfrak{h}$. Alors on a ou bien $\mathfrak{a}'=\mathfrak{a}''$ ou bien $\mathfrak{a}'\cap\mathfrak{a}''\subset\mathfrak{h}$.*

Démonstration. Il suffit évidemment de considérer le cas où $\mathfrak{a}'\neq\mathfrak{a}''$. Pour un raisonnement indirect supposons qu'il y a un $Y\in\mathfrak{a}'\cap\mathfrak{a}''$ tel que $Y\notin\mathfrak{h}$. Alors, en vertu du lemme précédent on a $\mathfrak{a}',\mathfrak{a}''\subset\mathfrak{p}_Y\subset\mathfrak{k}_L$. Mais $\mathfrak{a}'$, $\mathfrak{a}''$ éntant des sous-algèbres maximales dans $\mathfrak{k}_L$, cela entraîne $\mathfrak{a}'=\mathfrak{a}''=\mathfrak{p}_Y$ ce qui contredit la supposition. Par suite $\mathfrak{a}'\cap\mathfrak{a}''\subset\mathfrak{h}$.

Soit le sous-groupe à 1 paramètre $\varphi\colon\mathbf{R}\to G$ défini par $X\in\mathfrak{g}-\mathfrak{h}$ et soit $P\subset G$ le sous-groupe correspondant à la trajectoire $\pi\cdot\varphi$. Soient $\mathfrak{g}_X$ l'algèbre de Lie du sous-groupe $Q=H\cap P$ et $[X]$ le sous-espace de dimension 1 engendré par X. Alors, en conséquence du Lemme 1 on a la décomposition en somme directe de sous-espaces vectoriels $\mathfrak{p}_X=[X]\oplus\mathfrak{q}_X$ qui sera appelée *la décomposition d'isotropie de* $\mathfrak{p}_X$ *en* o. Il en résulte que pour $Z\in\mathfrak{q}_X$ on a $[Z,X]=\varkappa X+Z'$ où $\varkappa\in\mathbf{R}$ et $Z'\in\mathfrak{q}_X$. L'inverse de cette assertion, exprimé par le lemme suivant, se montrera utile dans la suite.

Lemme 3. *Soit $X\in\mathfrak{g}-\mathfrak{h}$ et $\mathfrak{q}\subset\mathfrak{k}$ une sous-algèbre qui est maximale par rapport à la propriété que pour $Z\in\mathfrak{q}$ on a $[Z,X]=\varkappa X+Z'$ où $\varkappa\in\mathbf{R}$ et $Z'\in\mathfrak{q}$. On a alors $\mathfrak{q}=\mathfrak{q}_X$.*

Démonstration. Soit $\mathfrak{p}=[X]\oplus\mathfrak{q}$ qui est évidemment une sous-algèbre de $\mathfrak{g}$ et soit $L=T_e\pi([X])$. Donc $X\in\mathfrak{p}\subset\mathfrak{k}_L$ et par conséquent on a

$$[X]\oplus\mathfrak{q}=\mathfrak{p}\subset\mathfrak{p}_X=[X]\oplus\mathfrak{q}_X,$$

en vertu du Lemme 2. On en conclut $\mathfrak{q}\subset\mathfrak{q}_X$ et la maximalité de $\mathfrak{q}$ entraîne $\mathfrak{q}=\mathfrak{q}_X$.

Pour $g \in G$, on définit par $\tau \mapsto \alpha(g, \pi_0 \varphi(\tau))$, $\tau \in R$ *la transformée de la trajectoire* $\pi \circ \varphi$ *par* g.

Lemme 4. *La transformée de la trajectoire* $\pi \circ \varphi$ *par* $g \in G$ *est la trajectoire d'origine* $\pi(g)$ *du sous-groupe à* 1 *paramètre* $\psi = \mathrm{ad}(g)\varphi \colon \mathbf{R} \to G$. *Le sous-groupe qui correspond à cette trajectoire dans G est* $\mathrm{ad}\,(g)P$.

Démonstration. Puisque $\psi(\tau) = g\varphi(\tau)g^{-1}$, on a $g\varphi(\tau) = \psi(\tau)g$ et par conséquent $\alpha(g, \pi \circ \varphi(\tau)) = \pi(g\varphi(\tau)) = \pi(\psi(\tau)g) = \alpha(\psi(\tau), \pi(g))$, $\tau \in R$, ce qui prouve la première assertion. Le sous-groupe qui correspond à cette dernière trajectoire est par définition celui qui correspond à $\pi_m \circ \psi$ où $m = \pi(g)$. Si φ est défini par $X \in \mathfrak{g} - \mathfrak{h}$, on a la décomposition $\mathfrak{p}_X = [X] \oplus \mathfrak{g}_X$. Soit $\mathrm{T}_e \mathrm{ad}(g) \colon \mathfrak{g} \to \mathfrak{g}$ la restriction à $\mathfrak{g} = \mathrm{T}_e G$ de l'application tangente linéaire à l'automorphisme $\mathrm{ad}\,(g) \colon G \to G$. Alors, on a $\mathrm{T}_e \,\mathrm{ad}\,(g)\mathfrak{p}_X = \mathrm{T}_e \,\mathrm{ad}\,(g)[X] \oplus \mathrm{T}_e \,\mathrm{ad}\,(g)\mathfrak{q}_X$. Mais ψ est défini par $\mathrm{T}_e \,\mathrm{ad}\,(g)X$ et $\mathrm{T}_e \mathrm{ad}\,(g)\mathfrak{q}_X \subset \mathfrak{h}_m$ où $\mathfrak{h}_m$ dest l'algèbre de Lie de H_m. On en conclut en vertu du Lemme 3 que $\mathrm{T}_e \mathrm{ad}\,(g)\mathfrak{p}_X$ est l'algèbre de Lie du sous-groupe P_m correspondant à $\pi_m \circ \psi$. Puisque le sous-groupe correspondant à une trajectoire est connexe selon la Proposition 2, cela entraîne $P_m = \mathrm{ad}\,(g)P$.

Corollaire. *Si* $g = \varphi(\xi)$ *et* $m = \pi(g)$, *le sous-groupe correspondant à* $\pi_m \circ \varphi$ *est P même, mais la décomposition d'isotropie de* $\mathfrak{p}$ *en* m *est* $\mathfrak{p} = [X] \oplus \mathrm{T}_e \mathrm{ad}\,(g)\mathfrak{q}$.

Démonstration. Parce que $\pi_m \circ \varphi(\tau) = \pi(\varphi(\tau)g) = \pi(g\varphi(\tau)) = \alpha(g, \pi \circ \varphi(\tau))$, $\tau \in \mathbf{R}$ et $g \in P$, le sous-groupe correspondant à $\pi_m \circ \varphi$ est P en vertu du Lemme. En conséquence de $X = \mathrm{T}_e \,\mathrm{ad}\,(g)X$, la décomposition d'isotropie de ${}_X$ en m est ${}_X =$ $= [X] \oplus (\mathfrak{p} \cap \mathfrak{h}_m)$, mais $\mathfrak{p} \cap \mathfrak{h}_m = \mathrm{T}_e \,\mathrm{ad}\,(g)(\mathfrak{p} \cap \mathfrak{h}) = \mathrm{T}_e \,\mathrm{ad}\,(g)\mathfrak{q}$.

Proposition 3. *Si* $g \in P$, *on a* $\alpha(g, \pi \circ \varphi(\tau)) = \pi \circ \varphi(\varkappa(\tau))$, $\tau \in \mathbf{R}$ *où* $\varkappa \colon \mathbf{R} \to \mathbf{R}$ *est une bijection analytique.*

Démonstration. Si $g \in P$, on a $\alpha(g, \pi \circ \varphi(\tau)) = \pi(g\varphi(\tau)) \in \pi(P) \subset$ $\subset \{\pi \circ \varphi(\xi) | \xi \in \mathbf{R}\}$ pour $\tau \in \mathbf{R}$ selon le Lemme 1. La trajectoire $\pi \circ \varphi$ étant non-triviale, soit ξ_0 le moins grand nombre positif tel que $\pi \circ \varphi(\xi_0) = \pi \circ \varphi(0)$, s'il y a des tels nombres; autrement soit $\xi_0 = \infty$. Si $\xi_0 = \infty$, il y a exactement un $\xi \in \mathbf{R}$ à un $\tau \in \mathbf{R}$ tel que $\pi \circ \varphi(\xi) = \alpha(g, \pi \circ \varphi(\tau))$; dans ce cas soit $\varkappa(\tau) = \xi$, $\tau \in \mathbf{R}$. L'application $\varkappa \colon \mathbf{R} \to \mathbf{R}$ ainsi définie est évidemment un homoémorphisme. Si $\xi_0 < \infty$, soit d'abord g dans le voisinage $L(\varphi; \varepsilon)$ de e dans P. Alors il y a une suite strictement croissante $\{\tau_i | i \in \mathbf{Z}\}$ telle que $\alpha(g, \pi \circ \varphi(\tau_i)) = o$ pour tout τ_i, $i \in \mathbf{Z}$ et que l'application $\sigma \mapsto \alpha(g, \pi_0 \varphi(\tau_i + \sigma))$, $\tau_i \leqq \tau_i + \sigma < \tau_{i+1}$ est injective pour $i \in \mathbf{Z}$. Par conséquent, pour tout σ tel que $\tau_i \leqq \tau_i + \sigma < \tau_{i+1}$, il existe exactement un $0 \leqq \xi < \xi_0$ tel que $\pi \circ \varphi(\xi) = \alpha(g, \pi \circ \varphi(\tau_i + \sigma))$; dans ce cas soit $\varkappa(\tau_i + \sigma) = i\xi_0 + \xi$ pour $\tau_i \leqq \tau_i + \sigma < \tau_{i+1}$ et $i \in \mathbf{Z}$. L'application $\varkappa \colon \mathbf{R} \to \mathbf{R}$ ainsi définie est évidement un homéomorphisme. Si $\xi_0 < \infty$ et $g \in P$ est arbitraire, il existe, en vertu de la Proposition 2, un $g_0 \in L(\varphi; \varepsilon)$ et un entier non-négatif l tels que

$g = g_0^i$. On peut évidement montrer dans ce cas l'existence d'un homéomorphisme $\varkappa: \mathbf{R} \to \mathbf{R}$ en utilisant le résultat précédent. On voit facilement que $\varkappa$ est analitique dans tous les cas considérés en vertu du théorème des fonctions implicites.

A compte de la proposition précédente, les éléments de P seront appelés *automorphismes de la trajectoire* $\pi \circ \varphi$. Si pour un $g \in P$ on a en particulier $\alpha(g, \pi \circ \varphi(\tau)) = \pi \circ \varphi(\lambda\tau + \mu)$, $\tau \in \mathbf{R}$ où $\lambda, \mu \in \mathbf{R}$, l'élément g sera appelé un *automorphisme linéaire de la trajectoire* $\pi \circ \varphi$. Dans le cas particulier où $\lambda = 1$ l'élément g sera appelé *un automorphisme affin de la trajectoire* $\pi \circ \varphi$, et dans le cas où $\lambda = 1$ *et* $\mu = 0$ l'élément g sera appelé un *automorphisme identique de la trajectoire* $\pi \circ \varphi$. Si tous les éléments de P sont des automorphismes linéaires de $\pi \circ \varphi$, on dit que $\pi \circ \varphi$ est une *trajectoire linéaire* de l'espace homogène M. Si tous les éléments de P sont des automorphismes affines de $\pi \circ \varphi$, on dit que $\pi \circ \varphi$ est une *trajectoire affine* de M. Si $H \cap P = Q = \{e\}$, la trajectoire $\pi \circ \varphi$ est dite *simple*.

Lemme 5. *La trajectoire* $\pi \circ \varphi$ *est linéaire, si le sous-groupe* $F = \{\varphi(\tau) | \tau \in \mathbf{R}\}$ *de* G *est laissé invariant par tout automorphisme* ad (q): $G \to G$, $q \in Q$.

Démonstration. En effet ad $(q)\varphi: \mathbf{R} \to G$ est un sous-groupe à 1 paramètre de G. En vertu de l'hypothèse du Lemme on a donc ad $(q)\varphi(\tau) = \varphi(\lambda\tau)$, $\tau \in \mathbf{R}$, où λ ne dépend que de $q \in Q$. Si $g \in P$, on a évidemment $g = \varphi(\mu)q$ où $\mu \in \mathbf{R}$ et $q \in Q$. Alors, $g\varphi(\tau) = \varphi(\mu)q\varphi(\tau) = \varphi(\mu)(\mathrm{ad}(q)\varphi(\tau))q = \varphi(\mu)\varphi(\lambda\tau)q$ et par conséquent on a $\alpha(g, \pi \circ \varphi(\tau)) = \pi \circ \varphi(\lambda\tau + \mu)$, $\tau \in \mathbf{R}$. Donc, $\pi \circ \varphi$ est une trajectoire linéaire de M.

Corollaire. *Si l'espace homogène* $M = G/H$ *admet une structure réductive* $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{h}$ *et le sous-groupe à 1 paramètre* $\varphi: \mathbf{R} \to G$ *est défini par* $X \in \mathfrak{m} - \{0\}$, *la trajectoire* $\pi \circ \varphi$ *est linéaire.*

Démonstration. Si $q \in Q$, on a T_e ad $(q)X \in \mathfrak{m}$ parce que $Q \subset H$. D'autre part, on a T_e ad $(q)X \in \mathfrak{p}_X$ parce que $Q \subset P$. Il en résulte que T_e ad $(q)X \in \mathfrak{m} \cap \mathfrak{p}_X = [X]$. Par conséquent, T_e ad $(q)X = \lambda X$ où $\lambda \in \mathbf{R}$. On en conclut que la hypothèse du lemme précédent est satisfaite.

Proposition 4. *Étant donné* $X \in \mathfrak{g} - \mathfrak{h}$, *soit* $\mathfrak{q}_X^0 = \mathfrak{h}$ *et soit* $\mathfrak{q}_X^i$ *défini successivement pour tout i naturel par*

$$\mathfrak{q}_X^i = \{Z | Z \in \mathfrak{q}_X^{i-1} \text{ et } [Z, X] = \lambda X + Z^* \text{ où } \lambda \in \mathbf{R} \text{ et } Z^* \in \mathfrak{q}_X^{i-1}\}.$$

Alors, $\mathfrak{h} = \mathfrak{q}_X^0 \supset \mathfrak{q}_X^1 \supset \ldots \supset \mathfrak{q}_X^i \supset \ldots$ *est une suite de sous-algèbres de* $\mathfrak{g}$. *Si j est le plus petit nombre tel que* $\mathfrak{q}_X^j = \mathfrak{q}_X^{j+1}$, *on a* $\mathfrak{q}_X^j = \mathfrak{q}_X$ *où* $\mathfrak{q}_X = \mathfrak{h} \cap \mathfrak{p}_X$.

Démonstration. En supposant que $\mathfrak{q}_X^{i-1}$ est une sous-algèbre de $\mathfrak{g}$, soient $Z', Z'' \in \mathfrak{q}_X^i$ et $[Z', X] = \lambda' X + Z^*$, $[Z'', X] = \lambda'' X + Z^{**}$ où $\lambda', \lambda'' \in \mathbf{R}$ et $Z^*, Z^{**} \in \mathfrak{q}_X^{i-1}$. Par conséquent pour $\xi, \eta \in \mathbf{R}$ on a $[\xi Z' + \eta Z'', X] = (\xi\lambda' + \eta\lambda'')X + \xi Z^* + \eta Z^{**}$, ce qui montre que $\xi Z' + \eta Z'' \in \mathfrak{q}_X^i$. De plus $[[Z'Z''], X] = [Z', [Z''X]] - [Z'', [Z', X]] =$

$=\lambda'' Z^* - \lambda' Z^{**} + [Z', Z^{**}] - [Z'', Z^*]$ entraîne que $[Z', Z''] \in \mathfrak{q}_X^i$. Alors, $\mathfrak{q}_X^i$ est également une sous-algèbre de $\mathfrak{g}$. Soit $L = T_e\pi([X])$, alors l'hypothèse $\mathfrak{q}_X^j = \mathfrak{q}_X^{j+1}$ entraîne que $[X] \oplus \mathfrak{q}_X^j$ est une sous-algèbre de $\mathfrak{q}$ telle que $X \in [X] \oplus \mathfrak{q}_X^j \subset \mathfrak{k}_L$. Donc, $[X] \oplus \mathfrak{q}_X^j \subset \mathfrak{p}_X$ en vertu du Lemme 2 et par conséquent $\mathfrak{q}_X^j \subset \mathfrak{h} \cap \mathfrak{p}_X = \mathfrak{q}_X$. D'autre part, la définition de $\mathfrak{q}_X^i$ entraîne que $\mathfrak{q}_X \subset \mathfrak{q}_X^i$ pour tout entier non-négatif i. Alors, en particulier $\mathfrak{q}_X \subset \mathfrak{q}_X^j$.

A compte de la proposition précédente $\pi \circ \varphi$ sera appelée *une trajectoire principale* de l'espace homogène M, si $\mathfrak{q}_X = \mathfrak{q}_X^1$. On voit facilement que $\pi \circ \varphi$ est une trajectoire principale si et seulement si tous les éléments $g \in G$ qui laissent fixés le point o et le sous-espace de dimension 1 $[T_e\pi X] \subset T_o M$, sont des automorphismes de $\pi \circ \varphi$.

Lemme 6. *Si l'espace homogène $M = G/H$ admet une structure réductive $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{h}$ et le sous-groupe à 1 paramètre φ est défini par $X \in \mathfrak{m} - \{0\}$, la trajectoire $\pi \circ \varphi$ est principale.*

Démonstration. Si $Z \in \mathfrak{g}_X^1$, on a $[Z, X] = \lambda X + Z^*$ où $\lambda \in \mathbf{R}$ et $Z^* \in \mathfrak{h}$. Mais $[Z, X] \in \mathfrak{m}$, parce que $\mathfrak{q}_X^1 \subset \mathfrak{h}$. Il en résulte que $Z^* = 0$ et par conséquent on a $\mathfrak{q}_X^2 = \mathfrak{q}_X^1$.

Lemme 7. *Soient φ, ψ: $\mathbf{R} \to G$ sous-groupes à 1 paramètre qui sont définis respectivement par $X, Y \in \mathfrak{g} - \mathfrak{h}$ où $Z = Y - X \in \mathfrak{h}$ et soit $\pi \circ \varphi$ une trajectoire affine principale. Alors, $\pi \circ \psi$ est une trajectoire principale si et seulement si Z est un élément du normalisateur de $\mathfrak{q}_X$ dans $\mathfrak{h}$.*

Démonstration. Parce que $Y - X = Z \in \mathfrak{h}$, on a évidemment $\mathfrak{q}_X^1 = \mathfrak{q}_Y^1$. Mais $\pi \circ \varphi$ étant principale, on a $\mathfrak{q}_X^1 = \mathfrak{q}_X$. Par conséquent $\mathfrak{q}_X \supset \mathfrak{q}_Y$. Donc, il suffit de montrer que $\mathfrak{q}_X \subset \mathfrak{q}_Y$, si et seulement si Z est un élément du normalisateur de $\mathfrak{q}_X$ dans $\mathfrak{h}$. Soit $U \in \mathfrak{q}_X$, alors $[U, Y] = [U, X+Z] = Z^* + [U, Z]$ où $Z^* \in \mathfrak{q}_X$. Donc, $\pi \circ \psi$ est une trajectoire principale si et seulement si $[U, Z] \in \mathfrak{q}_X$ pour tout $U \in \mathfrak{q}_X$.

2. Géodésiques de connexions affines inveriantes

Soit $L(M)$ la variété analytique formée par les repères linéaires de l'espace homogène $M = G/H$ et ϱ: $L(M) \to M$ la projection canonique. Soit

$$\beta: G \times L(M) \to -L(M)$$

l'action de G sur $L(M)$, qui est induite par l'action naturelle α: $G \times M \to M$. Si $X \in \mathfrak{g}$, le champ de vecteurs de Killing dans le sens plus general correspondant par l'action α sur M à X sera noté par X' et le champ de vecteurs de Killing correspondant par l'action β sur $L(M)$ à X sera noté par X''. Si $r \in L(M)$ et $m = \varrho(r)$, soit $\varkappa_r$: $T_m M \to \mathbf{R}^n$ l'application qui rend à un vecteur $v \in T_m M$ ses coordonnées par rapport à r où $n = \dim M$. Soit ϑ: $TL(M) \to \mathbf{R}^n$ la 1-formenon caique du fibré

$\varrho: L(M) \to M$. Alors on a $\vartheta(w) = \varkappa_r \circ T_r \varrho(w)$ pour $w \in T_r L(M)$ où $T_r \varrho$ est l'application linéaire tangente à ϱ en r. Si $g \in G$, soit $\alpha_g: M \to M$ transformation définie par $m \mapsto \alpha(g, m)$, $m \in M$ et soit $T\alpha_g$: $TM \to TM$ l'application linéaire tangente à α_g. Soit $r_0 \in L(M)$ fixé de façon que $\varrho(r_0) = o$ et soit $\iota: H \to GL(n; \mathbf{R})$ l'homomorphisme de groupes de Lie défini par $h \mapsto \varkappa_0 \circ T_0 \alpha_h \circ \varkappa_0^{-1}$ pour $h \in H$, où $\varkappa_0 = \varkappa_{r_0}$ et $T_0 \alpha_h$ est la restriction de $T\alpha_h$ à $T_0 M$. Soit $T_e \iota$: $\mathfrak{h} \to \mathfrak{gl}\,(n; \mathbf{R})$ l'homomorphisme d'algèbres de Lie qui est l'application linéaire tangente à ι en e. Comme $\alpha_h \circ \pi = \pi \circ \operatorname{ad}(h)$, on en conclut que $\iota(h) \circ \varkappa_0 \circ T_e \pi = \varkappa_0 \circ T_e \pi \circ T_e \operatorname{ad}(h)$ pour $h \in H$. Mais il en résulte que $T_e \iota(U) \vartheta(V_0'') = \varkappa_0([U, V]_0')$ pour $U \in \mathfrak{h}$ et $V \in \mathfrak{g}$ où V_0'' est la valeur du champ V'' en r_0 et $[U, V]_0'$ est la valeur du champ $[U, V]'$ en o.

Soit ω la 1-forme canonique d'une connexion affine invariante de M et soit $\Lambda: \mathfrak{g} \to \mathfrak{gl}\,(n, \mathbf{R})$ l'application linéaire correspondante qui est définie par $X \mapsto \omega(X_0'') = \Lambda(X)$ pour $X \in \mathfrak{g}$ où X_0'' est la valeur du champ X'' en r_0. On sait que Λ satisfait aux conditions suivantes:

1° $\Lambda(Z) = T_e \iota Z$ pour $Z \in \mathfrak{h}$,

2° $\Lambda([Z, X]) = [\Lambda(Z), \Lambda(X)]$ pour $Z \in \mathfrak{h}$ et pour $X \in \mathfrak{g}$.

De plus, on sait que à toute application linéaire $\Lambda: \mathfrak{g} \to \mathfrak{gl}\,(n; \mathbf{R})$ qui satisfait aux deux conditions précédentes il y a exectement une connexion affine invariante de M qui la définit comme ci-dessus ([5], II, p. 186—190).

S'il y a une connexion affine invariante sur l'espace homogène $M = G/H$ la transformation $\alpha_g: M \to M$ est affine pour tout $g \in G$ et par conséquent la transformée d'une géodésique est également une géodésique. Il en résulte édviemment le

Lemme 8. *Pour qu'une trajectoire de l'espace homogène $M = G/H$ soit une géodésique d'une connexion affine invariante de M, il est nécessaire que cette trajectoire soit linéaire et principale.*

Le lemme suivant reproduit une observation utile de R. Vosylius et A. Dreimanas [9]. La démonstration que nous en allons donner est plus détaillée, mais essentiellement la même que l'originelle.

Lemme 9. *Soit $M = G/H$ un espace homogène qui admet une structure réductive $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{h}$ et soit donnée une connexion affine invariante à torsion nulle de M telle que toutes ses géodésiques sont des trajectoires. De plus, soit $\xi: \mathfrak{m} \to \mathfrak{h}$ une application homogène telle que la trajectoire d'origine o définie par $X + \xi(X)$ est une géodésique pour tout $X \in \mathfrak{m} - \{0\}$. Alors, on a*

$$\Lambda(X)\vartheta(Y_0'') = \frac{1}{2}\varkappa_0 \circ T_e \pi([X, Y] + [X+Y, \xi(X+Y)] - [X, \xi(X)] - [Y, \xi(Y)])$$

pour $X, Y \in \mathfrak{m}$ où $\Lambda: \mathfrak{g} \to \mathfrak{gl}\,(n; \mathbf{R})$ est l'application linéaire correspondante à la connexion donnée.

Démonstration. Soient $\varphi, \psi, \chi: \mathbf{R} \to G$ les sous-groupes à 1 paramètre qui sont définis respectivement par $X+\xi(X)$, $Y+\xi(Y)$, $X+Y+\xi(X+Y)$ où $X, Y \in \mathfrak{m}-\{0\}$. Alors, les trajectoires $\pi\circ\varphi, \pi\circ\psi$ sont des géodésiques et $\pi\circ\chi$ est ou une géodésique ou une application banale. Ensuite, les trajectoires d'origine $r_0 \in L(M)$ de φ, ψ, χ sont des « lifts » de $\pi\circ\varphi, \pi\circ\psi, \pi\circ\chi$. Donc, en vertu d'un théorème fondamental ([1], p, 104—105) on a

$$(X''+\xi(X)''+\omega(X''+\xi(X)''))\vartheta(X'') = 0,$$
$$(Y''+\xi(Y)''+\omega(Y''+\xi(Y)''))\vartheta(Y'') = 0,$$
$$(X''+Y''+\xi(X+Y)''+\omega(X''+Y''+\xi(X+Y)''))\vartheta(X''+Y'') = 0$$

le long des trajectoires correspondantes dans $L(M)$ pour la 1-forme ω de la connexion donnée. Il en résulte qu'au point $r_0 \in L(M)$ on a

$$(X''+\omega(X''))\vartheta(Y'')+(Y''+\omega(Y''))\vartheta(X'')-$$
$$-(\xi(X)''+\omega(\xi(X)''))\vartheta(X'')-(\xi(Y)''+\omega(\xi(Y)''))\vartheta(Y'')+$$
$$+(\xi(X+Y)''+\omega(\xi(X+Y)''))\vartheta(X''+Y'') = 0.$$

Puisque la connexion envisagée est à torsion nulle on a $(X''+\omega(X''))\vartheta(Y'')-$ $-(Y''+\omega(Y''))\vartheta(X'')-(\vartheta[X'', Y''])=0$ partout sur la variété $L(M)$, en vertu de la première équation de structure. Par conséquent, au point $r_0 \in L(M)$ on a

$$2(X''+\omega(X''))\vartheta(Y'')-(\xi(X)''+\omega(\xi(X)''))\vartheta(X'')-(\xi(Y)''+\omega(\xi(Y)''))\vartheta(Y'')+$$
$$+(\xi(X+Y)''+\omega(\xi(X+Y)''))\vartheta(X''+Y'')-\vartheta([X'', Y'']) = 0.$$

Mais en vertvertu de faits fondamentaux ([5],I,p. 225—236) on a $U''\vartheta(U'')=L_{U''}\vartheta(U'')=$ $=(L_{U''}\vartheta)(U'')+\vartheta(L_{U''}U'')=0$ partout sur $L(M)$ pour tout $U\in\mathfrak{g}$. On en conclut que $2\omega(X'')\vartheta(Y'') = Y''\vartheta(X'')-X''\vartheta(Y'')-\omega(\xi(X+Y)'')\vartheta(X''+Y'')+\vartheta([X'', Y''])+$ $+\omega(\xi(X)'')\vartheta(X'')+\omega(\xi(Y)'')\vartheta(Y'')$ subsiste au point $r_0 \in L(M)$. Mais $Y''\vartheta(X'')=$ $-L_{Y''}\vartheta(X'')=(L_{Y''}\vartheta)(X'')+\vartheta(L_{Y''}X'')=\vartheta([Y'', X''])$, et de même, $X''\vartheta(Y'')=$ $=\vartheta([X'', Y''])$. De plus, on utilise le fait déjà cité ci-dessus que $\omega(U_0'')\vartheta(V_0'')=$ $=\mathrm{T}_e\imath(U)\vartheta(V_0'')=\varkappa_0([U, V_0'])$ pour $U\in\mathfrak{h}$ et pour $V\in\mathfrak{g}$. Par conséquent, on a

$$2\omega(X_0'')\vartheta(Y_0'') = \varkappa_0([X, Y]_0'+[X+Y, \xi(X+Y)]_0'-[X, \xi(X)]_0'-[Y, \xi(Y)]_0'),$$

en vertu du fait que $[U'', V'']=[V, U]''$ pour $U, V\in\mathfrak{g}$ ([5], II. p. 189); mais l'égalité ainsi obtenue équivaut évidemment à l'assertion du lemme.

Corollaire 1. *Si l'espace homogène $M=G/H$ admet une structure réductive $\mathfrak{g}=\mathfrak{m}\oplus\mathfrak{h}$ et le sous-groupe à 1 paramètre φ est défini par $X\in\mathfrak{m}-\{0\}$, la trajectoire $\pi\circ\varphi$ est affine.*

Démonstration. En effet, soit ν la 1-forme de la connexion naturelle à torsion nulle correspondant à la structure réductive $\mathfrak{g}=\mathfrak{m}\oplus\mathfrak{h}$. Si $Z\in\mathfrak{q}_X$ est fixé, il y a une application homogène $\xi\colon \mathfrak{m}\to\mathfrak{h}$ qui satisfait à l'hypothèse du lemme précédent pour $\omega=\nu$ et qui est telle que $\xi(X)=Z$. On a alors $\nu(X_0'')\vartheta(X_0'')=\varkappa_0\circ \mathrm{T}_e\pi([X,\xi(X)])$, selon le lemme. D'autre part, on a $\nu(X_0'')\vartheta(X_0'')=0$ d'après la définition de la connexion naturelle à torsion nulle. Il en résulte que $[Z,X]=[\xi(X),X]=0$; mais $Z\in\mathfrak{q}_X$ étant arbitrairement fixé, cela montre que la trajectoire $\pi\circ\varphi$ est affine.

Corollaire 2. *Soit $M=G/H$ un espace homogène qui admet une structure réductive $\mathfrak{g}=\mathfrak{m}\oplus\mathfrak{h}$ et soit donnée une connexion affine invariante à torsion nulle de M telle que toutes ses géodésiques sont des trajectoires. La connexion donnée est la connexion naturelle à torsion nulle d'une structure réductive de M si les conditions suivantes sont satisfaites:*

1. *Il y a une application linéaire $\xi\colon\mathfrak{m}\to\mathfrak{h}$ telle que la trajectoire d'origine o du sous-groupe à 1 paramètre défini par $X+\xi(X)$ est une géodésique pour tout $X\in\mathfrak{m}-\{0\}$.*
2. *On a $\mathrm{T}_e\,\mathrm{ad}\,(h)\xi(X)=\xi(\mathrm{T}_e\,\mathrm{ad}\,(h)X)$ pour $h\in H$ et $X\in\mathfrak{m}$.*

Démonstration. $\mathfrak{m}'=\{X+\xi(X)|X\in\mathfrak{m}\}$ est alors un sous-espace complémentaire à dans $\mathfrak{g}$ et la décomposition

$$\mathfrak{g}=\mathfrak{m}'\oplus\mathfrak{h}$$

est évidemment une structure réductive de M. On démontrera que la connexion donnée est la connexion naturelle à torsion nulle correspondant à la structure réductive $\mathfrak{g}=\mathfrak{m}'\oplus\mathfrak{h}$. Soient $U, V\in\mathfrak{m}'$, alors, il y a $X, Y\in\mathfrak{m}$ tels que $U=X+\xi(X)$ et $V=Y+\xi(Y)$. D'après le lemme précédent et par la linéarité de ξ, on a

$$\Lambda(U-\xi(X))\vartheta((V-\xi(Y))_0'')=\Lambda(U)\vartheta(V_0'')-\Lambda(\xi(X))\vartheta(V_0'')=$$
$$=\frac{1}{2}\varkappa_0\circ T_e\pi([U-\xi(X),V-\xi(Y)]+[U-\xi(X),\xi(Y)]+[V-\xi(Y),\xi(X)])=$$
$$=\frac{1}{2}\varkappa_0\circ T_e\pi([U,V]-2[\xi(X),V]).$$

Il en résulte évidemment que $\Lambda(U)\vartheta(V_0'')=1/2\varkappa_0\circ\mathrm{T}_e\pi([U,V])$ pour $U, V\in\mathfrak{m}'$. Alors, la connexion donnée est la connexion naturelle à torsion nulle de la structure réductive $\mathfrak{g}=\mathfrak{m}'\oplus\mathfrak{h}$ par un résultat fondamental ([4], II. p. 190—200).

Lemme 10. *Soient G un groupe de Lie connexe, $H\subset G$ un sous-groupe compact connexe et $\mathfrak{g}=\mathfrak{m}\oplus\mathfrak{h}$ une structure réductive de l'espace homogène $M=G/H$. De plus, soit donnée une connexion affine invariante de M telle que toutes ses géodésiques sont des trajectoires de sous-groupes à 1 paramètre de G. Alors, il y a une application $\xi\colon\mathfrak{m}\to\mathfrak{h}$ telle que les conditions suivantes sont satisfaites:*

1. *la trajectoire d'origine o du sous-groupe à 1 paramètre défini par $X+\xi(X)$ est une géodésique de la connexion donnée pour tout $X\in\mathfrak{m}-\{0\}$;*
2. *on a $\mathrm{T}_e\,\mathrm{ad}\,(h)\xi(X)=\xi(\mathrm{T}_e\,\mathrm{ad}\,(h)X)$ pour tout $h\in H$ et $X\in\mathfrak{m}$.*

De plus, l'application ξ est linèaire si elle est differentiable en $0\in\mathfrak{m}$.

Démonstration. Soit $K: \mathfrak{h}\times\mathfrak{h}\to\mathfrak{h}$ la forme de Killing de l'algèbre de Lie compacte $\mathfrak{h}$ et soit $\mathfrak{h}$ munie du produit intérieur défini par $-K$. Il y a alors un complément orthogonal $\mathfrak{n}_X$ dans $\mathfrak{h}$ de la sous-algèbre $\mathfrak{q}_X$ correspondant à un $X\in\mathfrak{m}-\{0\}$. En plus, soit $\mathfrak{c}'_X$ le centralisateur de $\mathfrak{q}_X$ dans $\mathfrak{h}$ et soit ϑ_X le complément orthogonal de $\mathfrak{c}_X=\mathfrak{c}'_X\cap\mathfrak{n}_X$ dans $\mathfrak{n}_X$. Donc, on a les décompositions suivantes de $\mathfrak{h}$ en sommes directes de sous-espaces vectoriels:

$$\mathfrak{h} = \mathfrak{n}_X\oplus\mathfrak{q}_X = \mathfrak{c}_X\oplus\vartheta_X\oplus\mathfrak{q}_X.$$

Le fait que le normalisateur de $\mathfrak{q}_X$ dans $\mathfrak{h}$ est $\mathfrak{c}_X\oplus\mathfrak{q}_X$ ([5], p. 66—70) se montrera très substantiel dans ce qui suit. La trajectoire d'origine o du sous-groupe à 1 paramètre défini par $X\in\mathfrak{M}-\{0\}$ est principale et affine selon le Lemme 6 et le Corollaire 1 du Lemme 9. Donc, pour que la trajectoire d'origine o du sous-group à 1 paramètre dèfini par $X+Z$ soit principale il faut et il suffit qu'on ait $Z\in\mathfrak{c}_X\oplus\mathfrak{q}_X$, en vertu du Lemme 7. D'autre part on sait par la Proposition 4 que $Z', Z''\in\mathfrak{c}_X\oplus\mathfrak{q}_X$ definissent la même trajectoire principale si et seulement si Z' et Z'' sont élements de la même classe $C+\mathfrak{q}_X$ pour un $C\in\mathfrak{c}_X$. De plus, les géodésiques de la connexion donnée sont trajectoires principales en conséquence du Lemme 8. On en conclut qu'il y a exactement un $C_X\in\mathfrak{c}_X$ tel que la trajectoire d'origine o du sous-groupe à 1 paramètre est la géodésique de la connexion donnée qui a o pour son origine et $T_e\pi X$ pour vecteur tangent en ce point. Soit $\xi(X)=C_X$ si $X\in\mathfrak{m}-\{0\}$ et soit $\xi(0)=0$. On montrera dans ce qui suit que l'application $\xi: \mathfrak{m}\to\mathfrak{h}$ ainsi définie satisfait à chacun des deux conditions posées ci-dessus.

D'abord, on obtiendra des représentations de trajectoires dans des systèmes de coordonnées convenablement choisis. Alors, il existe un voisinage W' de 0 dans $\mathfrak{m}$ tel que la restriction ϱ de $\pi\circ\exp$ à W' est un diffeomorphisme. Donc, l'application ϱ définit un système de coordonnées de l'espace M. De plus, soient $X\in\mathfrak{m}-\{0\}$ et $Z\in\mathfrak{h}$ fixés; dans ce cas, il y a des fonctions analytiques $U(\tau)\in\mathfrak{m}$, $V(\tau)\in\mathfrak{h}$ définies dans un voisinage de 0 dans $\mathbf{R}$ telles qu'on a

$$\exp(\tau(X+Z)) = \exp(U(\tau))\exp(V(\tau))$$

si τ est dans ce voisinage. Donc, la trajectoire $\pi\circ\exp(\tau(X+Z))$, $\tau\in\mathbf{R}$ est représentée par la fonction $U(\tau)$ dans le système de coordonnées défini par ϱ. On va étudier la dépendance de la fonction $U(\tau)$ du choix de Z dans $\mathfrak{h}$ pour un $X\in\mathfrak{m}-\{0\}$ fixé. On a évidemment, par la formule de Taylor,

$$U(\tau) = D^1U(0)\tau+\frac{1}{2}D^2U(0)\tau^2+\tilde{U}(\tau), \quad \text{où} \quad \tilde{U}(\tau) = o(\tau^2),$$

$$V(\tau) = D^1V(0)\tau+\frac{1}{2}D^2V(0)\tau^2+\tilde{V}(\tau), \quad \text{où} \quad \tilde{V}(\tau) = o(\tau^2)$$

pour un voisinage de 0 dans $\mathbf{R}$. D'autre part, soit l'application

$$\Pi : G \times G \to G$$

définie par la multiplication dans le groupe G. Par suite, on définit une application analytique Φ d'un voisinage de $(0, 0)$ dans $\mathfrak{g} \times \mathfrak{g}$ par

$$(A, B) \to \Phi(A, B) = \exp^{-1} \Pi(\exp A, \exp B).$$

En utilisant la formule de Taylor, on obtient ([9], p. 380—387) que dans un voisinage de $(0, 0)$ dans $\mathfrak{g} \times \mathfrak{g}$ on a

$$\Phi(A, B) = A + B + Q(A, B) + \tilde{\Phi}(A, B)$$

où $Q\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ est une application bilinéaire et l'application $\tilde{\Phi}$ est petite de troisième ordre. Il en résulte en vertu des observations précédentes que pour un voisinage de 0 dans $\mathbf{R}$ on a

$$\tau(X+Z) = U(\tau) + V(\tau) + Q\big(U(\tau), V(\tau)\big) + \tilde{\Phi}\big(U(\tau), V(\tau)\big) =$$

$$= \big(D^1 U(0) + D^1 V(0)\big)\tau + \left(\frac{1}{2} D^2 U(0) + \frac{1}{2} D^2 V(0) + Q\big(D^1 U(0), D^1 V(0)\big)\right)\tau^2 + R(\tau)$$

où $R(\tau) = o(\tau^2)$. En introduisant la décomposition $Q = Q' + Q''$ de Q par rapport à la décomposition $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{h}$, on en conclut en vertu de l'analyticité des fonctions considérées que

$$X = D^1 U(0), \quad Z = D^1 V(0),$$

$$0 = \frac{1}{2} D^2 U(0) + Q'\big(D^1 U(0), D^1 V(0)\big), \quad 0 = \frac{1}{2} D^2 V(0) + Q''\big(D^1 U(0), D^1 V(0)\big),$$

$$0 = R(\tau).$$

Par conséquent, on obtient que l'équation suivante est valable:

$$\frac{1}{2} D^2 U(0) + Q'(X, Z) = 0.$$

Mais cette équation exprime une dépendence de la fonction $U(\tau)$ du choix de Z dans $\mathfrak{h}$ pour un $X \in \mathfrak{m} - \{0\}$ fixé.

En utilisant les observations précedentes, on peut indiquer l'ensemble des $Z \in \mathfrak{h}$ tels que la trajectoire $\pi \circ \exp\big(\tau(X+Z)\big)$, $\tau \in \mathbf{R}$ soit représentée par la fonction

$$U(\tau) = X\tau$$

dans un voisinage de 0 dans $\mathbf{R}$. En effet, la trajectoire considerée est principale en vertu du Lemme 6 et par conséquent, l'ensemble envisagé est $\mathfrak{q}_X$. D'autre part, on a $D^2 U(0) = 0$ et par conséquent les éléments Z de l'ensemble envisagé satifont à l'équation

$$Q'(X, Z) = 0$$

en vertu des observations précédentes. Par contre, tous les $Z\in\mathfrak{h}$ qui satisfont à cette équation sont éléments de l'ensemble envisagé. En effet, $D^2U(0)=0$ entraîne que la dérivée covariante $\nabla_{D^1U(0)}D^1U(0)$ est zéro quand on la calcule par la connexion naturelle à torsion nulle de la structure réductive $\mathfrak{g}=\mathfrak{m}\oplus\mathfrak{h}$ parce que le système de coordonnées défini par ϱ est normal pour cette connexion. Par conséquent, $\nabla_{D^1U(\tau)}D^1U(\tau)=0$ pour tout τ considéré parce que $U(\tau)$ représente une trajectoire. Donc, $U(\tau)$ représente une géodésique de la connexion naturelle à torsion nulle de la structure réductive $\mathfrak{g}=\mathfrak{m}\oplus\mathfrak{h}$. Alors, $U(\tau)=X\tau$ dans un voisinage de 0 dans $\mathbf{R}$. Par suite,

$$\mathfrak{q}_X = \{Z|Q'(X,Z)=0, Z\in\mathfrak{h}\}$$

est valable.

Soit maintenant $U(\tau)$ une fonction qui représente une trajectoire principale quelconque pour un $X\in\mathfrak{m}-\{0\}$ fixé. En ce cas, l'ensemble des $Z\in\mathfrak{h}$ qui conduisent à la même fonction $U(\tau)$ est identique à l'ensemble des solutions de l'équation

$$\frac{1}{2}D^2U(0)+Q'(X,Z) = 0.$$

En effet, l'ensemble des $Z\in\mathfrak{h}$ qui conduisent à la fonction donnée $U(\tau)$ est $C+\mathfrak{q}_X$ où $C\in\mathfrak{c}_X$ est uniquement défini en conséquence du Lemme 7. De plus, l'ensemble des solutions de l'équation envisagée est $C+\mathfrak{q}_X$ puisque l'application Q' est bilinéaire. En particulier, soit $U(\tau)$ la fonction qui représente la trajectoire qui est une géodésique de la connexion donnée. En ce cas, l'ensemble des solutions de l'équation

$$\frac{1}{2}D^2U(0)+Q'(X,Z) = 0$$

est l'ensemble $\xi(X)+\mathfrak{q}_X$ en vertu de la définition de l'application ξ.

Pour obtenir d'autres conséquences des observations précédentes, on considère l'appliation $\varepsilon'\colon T_0M\to M$ qui est la restriction du l'application exponentielle de la connexion donnée à l'espace tangent T_0M. Il y a évidemment un voisinage W de 0 dans T_0M tel que la restriction de $\varrho^{-1}\circ\varepsilon'$ à W est un difféomorphisme

$$\varepsilon\colon W\to W'$$

où W' est le voisinage de 0 dans $\mathfrak{m}$ considéré déjà en ce qui précède. En vertu du fait que la fonction $U(\tau)$ correspondant à $Z=\xi(X)$ représente une géodésique de la connexion donnée il existe un vecteur tangent $v\in T_0M$ tel qu'on a

$$U(\tau) = \varepsilon(\tau v)$$

pour tout τ dans un voisinage de 0 dans $\mathbf{R}$. Donc, par la règle de dérivation des fonctions composées on a

$$X = D^1U(0) = D^1\varepsilon(0)v,\quad D^2U(0) = D^2\varepsilon(0)(v,v).$$

Par conséquent, il y a une application bilinéaire symétrique

$$A : \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$$

telle qu'on a $D^2U(0) = 2A(X, X)$ pour la fonction $U(\tau)$ qui représente la géodésique de la connexion donnée qui a o pour origine et $T_e \pi X$ pour vecteur tangent en o. Alors, l'ensemble des solutions Z de l'équation

$$A(X, X) + Q'(X, Z) = 0$$

pour $X \in \mathfrak{m} - \{0\}$ fixé, est $\xi(X) + \mathfrak{q}_X$. Donc, on a obtenu la suivante conséquence importante des observations précédentes: La fonction

$$Z \mapsto -K(Z, Z), \quad Z \in \mathfrak{h}$$

restreinte à l'ensemble des solutions de l'équation $A(X, X) + Q'(X, Z) = 0$ a exactement une valeur minimale qui est atteinte pour $Z = \xi(X)$.

On choisit une base de l'algèbre de Lie $\mathfrak{g}$ compatible avec la décomposition $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{h}$ et telle que sa partie dans $\mathfrak{h}$ soit orthonormée pour le produit intérieur défini par $-K$. Soient $m = \dim G$ et $n = \dim M$. Pour les coordonnées correspondantes à la base choisie on a alors

$$X = (X_1, \ldots, X_n, 0, \ldots, 0), \quad Z = (0, \ldots, 0, Z_{n+1}, \ldots, Z_m),$$

$$\xi(X) = \big(0, \ldots, 0, \xi_{n+1}(X_1, \ldots, X_n), \ldots, \xi_m(X_1, \ldots, X_n)\big),$$

$$-K(Z, Z) = \sum_{k=n+1}^{m} Z_k^2,$$

$$Q'(X, Z) = \left(\sum_{i=1}^{n} \sum_{k=n+1}^{m} q_{ik}^1 X_i Z_k, \ldots, \sum_{i=1}^{n} \sum_{k=n+1}^{m} q_{ik}^n X_i Z_k\right),$$

$$A(X, X) = \left(\sum_{i,j=1}^{n} a_{ij}^1 X_i X_j, \ldots, \sum_{i,j=1}^{n} a_{ij}^n X_i X_j\right).$$

En remaniant la proposition précédente on obtient donc la suivante: La fonction $\sum_{k=1}^{m} Z_k^2$ assujettie aux conditions

$$F_l(Z_{n+1}, \ldots, Z_m) = \sum_{i,j=1}^{n} a_{ij}^l X_i X_j + \sum_{i=1}^{n} \sum_{k=n+1}^{m} q_{ik}^l X_i Z_k = 0, \qquad l = 1, \ldots, n,$$

où $(X_1, \ldots, X_n) \neq (0, \ldots, 0)$ est fixé, admet exactement une valeur minimale qui est atteinte pour $Z_k = \xi_k(X_1, \ldots, X_n)$, $k = n+1, \ldots, m$. On considère la fonction

$$\Phi(Z_{n+1}, \ldots, Z_m) = \sum_{k=n+1}^{m} Z_k^2 + \sum_{\iota=1}^{n} \lambda_\iota F_\iota(Z_{n+1}, \ldots, Z_m)$$

où $\lambda_1, \ldots, \lambda_n$ sont les multiplicateurs de Lagrange uniquement définis. On sait par la théorie des valeurs extrêmes relatives que le système d'équations

$$\frac{\partial \Phi}{\partial Z_k} = 2Z_k + \sum_{l=1}^{n} \lambda_l \sum_{i=1}^{n} q_{ik}^l X_i = 0 \quad (k = n+1, \ldots, m),$$

$$F_l(Z_{n+1}, \ldots, Z_m) = 0 \quad (l = 1, \ldots, n)$$

admet exactement une solution, donnée par $Z_k = \xi_k(X_1, \ldots, X_n)$, $k = n+1, \ldots, m$.

Par une substitution évidente on obtient le système d'équations

$$\sum_{i,j=1}^{n} a_{ij}^r X_i X_j - \frac{1}{2} \sum_{l=1}^{n} \left(\sum_{i,j=1}^{n} \sum_{k=n+1}^{m} q_{ik}^r q_{jk}^l X_i X_j \right) \lambda_l = 0 \quad (r = 1, \ldots, n).$$

Ce système définit uniquement les λ_l et on voit facilement que si l'on les considère comme fonctions de $(X_1, \ldots, X_n)$, ces fonctions $\lambda_l(X_1, \ldots, X_n)$, $l = 1, \ldots, n$, sont analytiques dans $\mathfrak{m} - \{0\}$. Par conséquent, les fonctions $\xi_k(X_1, \ldots, X_n)$, $k = n+1, \ldots, m$ sont aussi analytiques dans le domaine $\mathfrak{m} - \{0\}$. Mais, substitution dans le système d'équations

$$2Z_k + \sum_{l=1}^{n} \lambda_l \sum_{k=1}^{n} q_{ik}^l X_i = 0 \quad (k = n+1, \ldots, m)$$

montre que les fonctions $\xi_k(X_1, \ldots, X_n)$, $k = n+1, \ldots, m$ sont linéaires. Par suite, l'application $\xi \colon \mathfrak{m} \to \mathfrak{h}$ est linéaire.

Il reste encore à montrer que l'application ξ satisfait à la seconde conditon posée. En effet, l'application

$$\mathrm{T}_e \operatorname{ad}(h) : \mathfrak{g} \to \mathfrak{g}$$

est un automorphisme d'algèbre de Lie $\mathfrak{g}$ pour tout $h \in H$. Par conséquent, on a

$$\mathfrak{c}_{\mathrm{T}_e \operatorname{ad}(h) X} = \mathrm{T}_e \operatorname{ad}(h)(\mathfrak{c}_X)$$

pour $X \in \mathfrak{m} - \{0\}$ et $h \in H$. Il en résulte en particulier que

$$\mathrm{T}_e \operatorname{ad}(h) \xi(X) \in \mathfrak{c}_{\mathrm{T}_e \operatorname{ad}(h) X}.$$

D'autre part, la transformation $\alpha_h \colon M \to M$ applique les géodésiques en des géodésiques et par suite la trajectoire

$$\alpha_h \circ \pi \circ \exp(\tau(X + \xi(X))) = \pi \circ \operatorname{ad}(h) \circ \exp(\tau(X + \xi(X))) =$$
$$= \pi \circ \exp(\tau(\mathrm{T}_e \operatorname{ad}(h)(X + \xi(X)))),$$

est une géodésique. Cela entraîne en vertu des observations précédentes, que

$$\mathrm{T}_e \operatorname{ad}(h) \xi(X) \in \xi(\mathrm{T}_e \operatorname{ad}(h) X) + \mathfrak{q}_{\mathrm{T}_e \operatorname{ad}(h) X}.$$

Par conséquent, on a

$$T_e \operatorname{ad}(h)\xi(X) = \xi(T_e \operatorname{ad}(h)X)$$

pour tout $X \in \mathfrak{m} - \{0\}$ et $h \in H$. Donc, la seconde condition est aussi vérifiée.

Les raisonnements ci-dessus ont été faits en vue d'obtenir le suivant

Theorème. *Soient G un groupe de Lie connexe, H un sous-groupe compact connexe de G et soit donnée une connexion affine invariante à torsion nulle de l'espace homogène $M = G/H$ telle que toutes ses géodésiques sont des trajectoires et que ξ est differentiable en $0 \in \mathfrak{m}$. Alors, il y a une structure réductive de M telle que sa connexion naturelle à torsion nulle est la connexion donnée.*

Démonstration. Puisque le sous-groupe H est compact, l'espace homogène M admet une structure réductive $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{h}$. Donc, en conséquence du Lemme 10 il y a une application $\xi: \mathfrak{m} \to \mathfrak{h}$ qui satisfait à chacune des deux conditions posées dans le Corollaire 2 du Lemme 9. Selon ce corollaire il y a donc une structure réductive $\mathfrak{g} = \mathfrak{m}' \oplus \mathfrak{h}$ de M dont la connexion naturelle à torsion nulle est la connexion donnée.

Bibliographie

[1] R. L. Bishop—R. J. Crittenden, *Geometry of manifolds* (New York, 1964).

[2] N. Bourbaki, *Topologie générale*, Chapitre III. *Groupes topologiques* (Paris, 1951).

[3] F. W. Kamber—P. Tondeur, *Invariant differential operators and the cohomology of Lie algebra sheaves*, Memoirs of the American Mathematical Society, No. 113 (Providence, 1971).

[4] S. Kobayashi—K. Nomizu, *Foundations of differential geometry* (New York, 1963—69).

[5] A. Lichnerowicz, *Géométrie des groupes de transformations* (Paris, 1958).

[6] K. Nomizu, Invariant affine connection on homogeneous spaces, *Amer. J. Math.*, **76** (1954), 33—65.

[7] K. Nomizu, On infinitesimal holonomy and isotropy groups, *Nagoya Math. J.*, **11** (1957), 111—114.

[8] P. Восилюс—А. Дреиманас, О геометрии однородных пространств, *Литовский мат. сб.*, **11** (1971), 773—781.

[9] Л. С. Понтрягин, *Непрерывные группы* (Москва, 1973).

UNIVERSITÉ DE TECHNOLOGIE
1521 BUDAPEST, HONGRIE

Acta Sci. Math., 38 (1976), 399—416

On some recurrence equations in a Banach algebra

LAJOS TAKÁCS

1. Introduction. The aim of this paper is to find the solutions of the recurrence equations

(1) $$f_n = \mathbf{L}\{f_{n-1}g_1\} + \mathbf{L}^*\{g_2 f_{n-1}\}$$

and

(2) $$f_n = \mathbf{L}\{f_{n-1}g_1\} + \mathbf{L}^*\{f_{n-1}g_2\},$$

and the solution of the system of recurrence equations

(3) $$u_n = \mathbf{L}\{u_{n-1}h_1 + v_{n-1}h_2\}$$

(4) $$v_n = \mathbf{L}^*\{u_{n-1}h_3 + v_{n-1}h_4\}$$

where $f_0, g_1, g_2, u_0, v_0, h_1, h_2, h_3, h_4$ are elements of a Banach algebra $\mathbf{R}$, $\mathbf{L}$ is a projection in $\mathbf{R}$, and $\mathbf{L}+\mathbf{L}^*$ is the identity transformation in $\mathbf{R}$. The solutions of these recurrence equations make it possible to determine the stochastic laws of the fluctuations of the partial sums for a sequence of independent and identically distributed real random variables and for a semi-Markov sequence of real random variables.

This paper generalizes and extends some earlier results of the author [11].

2. Preliminaries. Let $\mathbf{R}$ be a Banach algebra of elements $f, f_1, f_2, \ldots$. We denote by θ the zero element and by e the identity element of R. Denote by $\|f\|$ the norm of f and let $\|e\| = 1$.

Throughout this paper we shall consider transformations $\mathbf{T}$ in $\mathbf{R}$ which satisfy the following conditions:

(i) The transformation $\mathbf{T}$ is a bounded linear transformation of $\mathbf{R}$ into itself.

(ii) The transformation $\mathbf{T}$ is a projection, that is,

$$\mathbf{T}^2\{f\} = \mathbf{T}\{f\} \text{ for all } f.$$

Received July 10, 1975.

(iii) If either $\mathbf{T}\{f_i\}=f_i$ or $\mathbf{T}\{f_i\}=\theta$ for $i=1, 2$, then

$$\mathbf{T}\{f_1 f_2\} = \mathbf{T}\{f_1\}\mathbf{T}\{f_2\}.$$

We note that (iii) can be expressed in the following equivalent form:

$$\mathbf{T}\{f_1 f_2\} = \mathbf{T}\{f_1 \mathbf{T}\{f_2\}\}+\mathbf{T}\{\mathbf{T}\{f_1\}f_2\}-\mathbf{T}\{f_1\}\mathbf{T}\{f_2\} \tag{5}$$

for all f_1 and f_2.

The norm of $\mathbf{T}$ is defined as the smallest nonnegative number $\|\mathbf{T}\|$ for which $\|\mathbf{T}\{f\}\| \leqq \|\mathbf{T}\| \, \|f\|$ for all $f \in \mathbf{R}$. If $\mathbf{T}$ is not the zero transformation, then (ii) implies that $\|\mathbf{T}\| \geqq 1$.

We define

$$\mathbf{T}^*\{f\} = f - \mathbf{T}\{f\} \tag{6}$$

for any $\mathbf{T}$ and f. If $\mathbf{T}$ statisfies the conditions (i), (ii), (iii), then $\mathbf{T}^*$ too satisfies these conditions. We have $\|\mathbf{T}^*\| \leqq 1 + \|\mathbf{T}\|$.

It will be convenient to introduce here some useful definitions which we shall need later. Let us suppose that $a_0=b_0=e$ and $a_n=\mathbf{T}\{a_{n-1}g\}$ and $b_n=\mathbf{T}^*\{gb_{n-1}\}$ for $n=1, 2, \ldots$ where $g \in \mathbf{R}$. For a nonzero transformation $\mathbf{T}$ let us define $\mu(\mathbf{T})$ as the largest nonnegative number for which

$$\sum_{n=0}^{\infty} \|a_n\| \, |\varrho|^n < \infty$$

whenever $|\varrho| \, \|g\| < \mu(\mathbf{T})$ and $g \in \mathbf{R}$. Similarly for a nonzero transformation $\mathbf{T}^*$ let us define $\bar{\mu}(\mathbf{T}^*)$ as the largest nonnegative number for which

$$\sum_{n=0}^{\infty} \|b_n\| \, |\varrho|^n < \infty$$

whenever $|\varrho| \, \|g\| < \bar{\mu}(\mathbf{T}^*)$ and $g \in \mathbf{R}$. Obviously

$$\|\mathbf{T}\|^{-1} \leqq \mu(\mathbf{T}) \leqq 1 \quad \text{and} \quad \|\mathbf{T}^*\|^{-1} \leqq \bar{\mu}(\mathbf{T}^*) \leqq 1. \tag{7}$$

If $\|\mathbf{T}\|=0$, then we write $\mu(\mathbf{T})=\infty$ and if $\|\mathbf{T}^*\|=0$, then we write $\bar{\mu}(\mathbf{T}^*)=\infty$. Let us define

$$c(\mathbf{T}) = \min\left(\mu(\mathbf{T}), \bar{\mu}(\mathbf{T}^*)\right), \tag{8}$$

and

$$\gamma(\mathbf{T}) = \min\left(c(\mathbf{T}), c(\mathbf{T}^*)\right). \tag{9}$$

We note that if $\mathbf{R}$ is a commutative Banach algebra, and if $\mathbf{T}$ satisfies (i), (ii), (iii), then $c(\mathbf{T})=1$. If $\mathbf{R}$ is a commutative Banach algebra, then we can prove by mathematical induction that

$$na_n = \sum_{k=1}^{n} a_{n-k} \mathbf{T}\{g^k\}$$

for $n=1, 2, \ldots$. Hence

$$n\|a_n\| \leqq \|\mathbf{T}\| \sum_{k=1}^{n} \|a_{n-k}\| (\|g\|)^k \tag{10}$$

for $n=1, 2, \ldots$. By (10) it follows by induction that

$$\|a_n\| \leqq \binom{\|\mathbf{T}\|+n-1}{n} (\|g\|)^n$$

for $n=0, 1, 2, \ldots$. This implies that $\mu(\mathbf{T}) \geqq 1$. Since $\bar{\mu}(\mathbf{T}^*) \geqq 1$ also holds, by (7) and (8) we obtain that $c(\mathbf{T})=1$.

3. The method of factorization. In solving various recurrence equations in the space $\mathbf{R}$ we shall use the method of factorization. It seems the method of factorization in Banach spaces was used for the first time in 1956 by P. MASANI [6]. See also G. BAXTER [1], [2] and I. C. GOHBERG [4].

Let $h(\varrho)$ be an element of $\mathbf{R}$ for $|\varrho|<r$ where r is some positive real number. We say that the element $h(\varrho)$ can be represented by a Taylor series about $\varrho=0$ in the circle $|\varrho|<r$ if

$$h(\varrho) = \sum_{n=0}^{\infty} h_n \varrho^n$$

and

$$\sum_{n=0}^{\infty} \|h_n\| \, |\varrho|^n < \infty$$

for $|\varrho|<r$.

Let us suppose that $\mathbf{T}$ is a transformation in $\mathbf{R}$ which satisfies (i), (ii) and (iii). We shall consider various elements $h(\varrho)$ of $\mathbf{R}$ for $|\varrho|<r$ which satisfy one of the following two properties.

Property (a). The element $h(\varrho)$ has an inverse $[h(\varrho)]^{-1}$, $h(0)=e$, $\mathbf{T}\{h(\varrho)-e\}= =h(\varrho)-e$, $\mathbf{T}\{[h(\varrho)]^{-1}-e\}=[h(\varrho)]^{-1}-e$, $h(\varrho)$ and $[h(\varrho)]^{-1}$ can be represented by a Taylor series about $\varrho=0$.

Property (b). The element $h(\varrho)$ has an inverse $[h(\varrho)]^{-1}$, $h(0)=e$, $\mathbf{T}^*\{h(\varrho)-e\}= =h(\varrho)-e$, $\mathbf{T}^*\{[h(\varrho)]^{-1}-e\}=[h(\varrho)]^{-1}-e$, $h(\varrho)$ and $[h(\varrho)]^{-1}$ can be represented by a Taylor series about $\varrho=0$.

The method of factorization is based on the following theorem.

Theorem 1. *If* $g \in \mathbf{R}$ *and if* $|\varrho| \, \|g\| < c(\mathbf{T})$, *then there exist two elements* $g^+(\varrho) \in \mathbf{R}$ *and* $g^-(\varrho) \in \mathbf{R}$ *such that*

$$e - \varrho g = g^+(\varrho) g^-(\varrho) \tag{11}$$

where $g^+(\varrho)$ *satisfies* (a) *and* $g^-(\varrho)$ *satisfies* (b). *The elements* $g^+(\varrho)$ *and* $g^-(\varrho)$ *are uniquely determined by* (a), (b) *and* (11).

Proof. First, we shall construct two elements $g^+(\varrho)$ and $g^-(\varrho)$ which satisfy (a), (b) and (11). Let us suppose that $a_0=b_0=e$ and $a_n=\mathbf{T}\{a_{n-1}g\}$ and $b_n=\mathbf{T}^*\{gb_{n-1}\}$ for $n=1, 2, \ldots$. Then

$$a(\varrho) = \sum_{n=0}^{\infty} a_n \varrho^n \in \mathbf{R} \tag{12}$$

for $|\varrho|\,\|g\|<\mu(\mathbf{T})$ and

$$b(\varrho) = \sum_{n=0}^{\infty} b_n \varrho^n \in \mathbf{R} \tag{13}$$

for $|\varrho|\,\|g\|<\bar{\mu}(\mathbf{T}^*)$. From the definitions of $a(\varrho)$ and $b(\varrho)$ it follows immediately that $a(0)=b(0)=e$, $\mathbf{T}\{a(\varrho)-e\}=a(\varrho)-e$, $\varrho\mathbf{T}\{a(\varrho)g\}=a(\varrho)-e$, $\mathbf{T}^*\{b(\varrho)-e\}=b(\varrho)-e$ and $\varrho\mathbf{T}^*\{gb(\varrho)\}=b(\varrho)-e$.

Now we shall prove that

$$(e-\varrho g)b(\varrho)a(\varrho) = b(\varrho)a(\varrho)(e-\varrho g) = e \tag{14}$$

and

$$a(\varrho)(e-\varrho g)b(\varrho) = e \tag{15}$$

for $|\varrho|\,\|g\|<c(\mathbf{T})$. If we take into consideration that $\mathbf{T}\{a(\varrho)(e-\varrho g)\}=\mathbf{T}\{e\}$ and $\mathbf{T}^*\{(e-\varrho g)\,b(\varrho)\}=\mathbf{T}^*\{e\}$, then by (5) it follows that

$$\mathbf{T}\{b(\varrho)a(\varrho)(e-\varrho g)\} = \mathbf{T}\{e\}$$

and

$$\mathbf{T}^*\{(e-\varrho g)b(\varrho)a(\varrho)\} = \mathbf{T}^*\{e\}.$$

If we add these two equations, then we get

$$b(\varrho)a(\varrho) = e+\varrho\mathbf{T}\{b(\varrho)a(\varrho)g\}+\varrho\mathbf{T}^*\{gb(\varrho)a(\varrho)\}. \tag{16}$$

If $|\varrho|\,\|g\|<c(\mathbf{T})$, then $b(\varrho)a(\varrho)\in\mathbf{R}$ and in the above equation we can write that

$$b(\varrho)\,a(\varrho) = \sum_{n=0}^{\infty} \gamma_n \varrho^n$$

where $\gamma_n\in\mathbf{R}$ for $n=0, 1, 2, \ldots$. By forming the coefficient of ϱ^n in (16), we get

$$\gamma_n = \mathbf{T}\{\gamma_{n-1}g\}+\mathbf{T}^*\{g\gamma_{n-1}\} \tag{17}$$

for $n=1, 2, \ldots$. Since $\gamma_0=e$, it follows from (17) by induction that $\gamma_n=g^n$ for $n=1, 2, \ldots$. This implies (14).

By (5) it follows also that

$$\mathbf{T}\{a(\varrho)(e-\varrho g)b(\varrho)\} = \mathbf{T}\{e\} \quad\text{and}\quad \mathbf{T}^*\{a(\varrho)(e-\varrho g)b(\varrho)\} = \mathbf{T}^*\{e\}.$$

If we add these two equations, then we get (15).

We can conclude from (14) and (15) that $[a(\varrho)]^{-1}$ and $[b(\varrho)]^{-1}$ exist and

(18) $$[a(\varrho)]^{-1} = (e-\varrho g)b(\varrho)$$

and

(19) $$[b(\varrho)]^{-1} = a(\varrho)(e-\varrho g)$$

for $|\varrho|\,\|g\| < c(\mathbf{T})$.

If we define

(20) $$g^{+}(\varrho) = [a(\varrho)]^{-1}$$

and

(21) $$g^{-}(\varrho) = [b(\varrho)]^{-1}$$

for $|\varrho|\,\|g\| < c(\mathbf{T})$, then $g^{+}(\varrho)$ and $g^{-}(\varrho)$ satisfy (a), (b) and (11).

It remains to show that $g^{+}(\varrho)$ and $g^{-}(\varrho)$ are uniquely determined by (a), (b) and (11). This fact will be proved as a consequence of Theorem 3.

In exactly the same way as we proved Theorem 1 we can prove the following theorem too.

Theorem 2. *If $g\in\mathbf{R}$ and if $|\varrho|\,\|g\| < c(\mathbf{T}^*)$ then there exist two elements $h^{+}(\varrho)\in\mathbf{R}$ and $h^{-}(\varrho)\in\mathbf{R}$ such that*

(22) $$e-\varrho g = h^{-}(\varrho)h^{+}(\varrho)$$

where $h^{+}(\varrho)$ satisfies (a) *and $h^{-}(\varrho)$ satisfies* (b). *The elements $h^{+}(\varrho)$ and $h^{-}(\varrho)$ are uniquely determined by* (a), (b) *and* (22).

If we suppose that $c_0=d_0=e$, $c_n=\mathbf{T}\{gc_{n-1}\}$ and $d_n=\mathbf{T}^*\{d_{n-1}g\}$ for $n=1, 2, \ldots,$

(23) $$c(\varrho) = \sum_{n=0}^{\infty} c_n \varrho^n$$

and

(24) $$d(\varrho) = \sum_{n=0}^{\infty} d_n \varrho^n,$$

then in Theorem 2 we can write that $h^{+}(\varrho)=[c(\varrho)]^{-1}$ and $h^{-}(\varrho)=[d(\varrho)]^{-1}$.

We note that if $\mathbf{R}$ is a commutative Banach algebra, then (12), (13), (23) and (24) can be expressed in the following explicit forms

$$a(\varrho) = c(\varrho) = \exp\big\{-\mathbf{T}\{\log(e-\varrho g)\}\big\} \quad \text{and} \quad b(\varrho) = d(\varrho) = \exp\big\{-\mathbf{T}^*\{\log(e-\varrho g)\}\big\}$$

where

$$\log(e-\varrho g) = -\sum_{n=1}^{\infty}\frac{g^n\varrho^n}{n} \quad \text{for} \quad |\varrho|\,\|g\| < 1 \quad \text{and} \quad \exp(f) = e+\sum_{n=1}^{\infty}\frac{f^n}{n!}$$

for any $f\in\mathbf{R}$.

4. Some linear transformations in R. In this section we shall consider transformations **L** which satisfy conditions (i), (ii), (iii) and can be represented in the form

$$\mathbf{L}\{f\} = \mathbf{T}\{f\} - \alpha(f)e \tag{25}$$

where **T** is a given transformation satisfying (i), (ii), (iii) and $\alpha(f)$ is a complex (or real) functional on **R**.

We can prove that **L** satisfies the above conditions if and only if $\alpha(f)$ satisfies one of the following three sets of conditions: (1) $\alpha(f)\equiv 0$, (2) $\alpha(cf)=c\alpha(f)$ for any constant c, $\alpha(f_1+f_2)=\alpha(f_1)+\alpha(f_2)$, $\alpha(\mathbf{T}\{f\})=\alpha(f)$, $\alpha(\mathbf{T}\{f_1\}\mathbf{T}\{f_2\})=\alpha(f_1)\alpha(f_2)$, $\alpha(e)e=\mathbf{T}\{e\}$, $|\alpha(f)|\leqq\|\mathbf{T}\|^2\|f\|$, (3) $\alpha(cf)=c\alpha(f)$ for any constant c, $\alpha(f_1+f_2)=$ $=\alpha(f_1)+\alpha(f_2)$, $\alpha(\mathbf{T}^*\{f\})=\alpha(f)$, $\alpha(\mathbf{T}^*\{f_1\}\mathbf{T}^*\{f_2\})=-\alpha(f_1)\alpha(f_2)$, $\alpha(e)e=\mathbf{T}^*\{e\}$, $|\alpha(f)|\leqq\|\mathbf{T}^*\|^2\|f\|$.

Later we shall prove that for any **L** defined by (25) we have

$$c(\mathbf{L}) = c(\mathbf{T}) \tag{26}$$

where $c(\mathbf{T})$ is defined by (8).

We shall state here a few general relations which can be deduced from (5). In agreement with (6) we define $\mathbf{L}^*\{f\}=f-\mathbf{L}\{f\}$ for any f.

For any $f\in\mathbf{R}$ we have

$$\mathbf{T}\{\mathbf{T}\{e\}f\} = \mathbf{T}\{e\}\mathbf{T}\{f\} \quad \text{and} \quad \mathbf{T}\{f\mathbf{T}\{e\}\} = \mathbf{T}\{f\}\mathbf{T}\{e\}. \tag{27}$$

By (25) and (27) it follows that if $f\in\mathbf{R}$, $\gamma\in\mathbf{R}$ and $\mathbf{T}\{\gamma\}=\mathbf{T}\{e\}$, then

$$\mathbf{L}\{f\gamma\} = \mathbf{L}\{\mathbf{L}\{f\}\gamma\} \quad \text{and} \quad \mathbf{L}\{\gamma f\} = \mathbf{L}\{\gamma\mathbf{L}\{f\}\}, \tag{28}$$

and if $f\in\mathbf{R}$, $\gamma\in\mathbf{R}$ and $\mathbf{T}^*\{\gamma\}=\mathbf{T}^*\{e\}$, then

$$\mathbf{L}^*\{f\gamma\} = \mathbf{L}^*\{\mathbf{L}^*\{f\}\gamma\} \quad \text{and} \quad \mathbf{L}^*\{\gamma f\} = \mathbf{L}^*\{\gamma\mathbf{L}^*\{f\}\}. \tag{29}$$

If $f\in\mathbf{R}$, $\gamma_i\in\mathbf{R}$ $(i=1, 2)$ and $\mathbf{T}\{\gamma_i\}=\mathbf{T}\{e\}$ $(i=1, 2)$, then we have

$$\mathbf{L}\{\gamma_1\mathbf{L}\{f\}\gamma_2\} = \mathbf{L}\{\gamma_1 f\gamma_2\} \quad \text{and} \quad \mathbf{L}\{\gamma_1\mathbf{L}^*\{f\}\gamma_2\} = \theta. \tag{30}$$

The first equation follows from the repeated applications of (28). The second follows from the first one.

If $f\in\mathbf{R}$, $\gamma_i\in\mathbf{R}$ $(i=1, 2)$ and $\mathbf{T}^*\{\gamma_i\}=\mathbf{T}^*\{e\}$ $(i=1, 2)$, then we have

$$\mathbf{L}^*\{\gamma_1\mathbf{L}^*\{f\}\gamma_2\} = \mathbf{L}^*\{\gamma_1 f\gamma_2\} \quad \text{and} \quad \mathbf{L}^*\{\gamma_1\mathbf{L}\{f\}\gamma_2\} = \theta. \tag{31}$$

The first equation follows from the repeated applications of (29). The second follows from the first one.

Now we shall consider the solutions of the three recurrence equations stated in the Introduction.

5. The first recurrence equation. Let us consider the recurrence equation (1) for $n=1, 2, \ldots$ where $f_0\in\mathbf{R}$, $g_1\in\mathbf{R}$, $g_2\in\mathbf{R}$ and **L** satisfies the conditions (i), (ii), (iii) and can be represented in the form of (25). Obviously, $f_n \in \mathbf{R}$ for $n=1, 2, \ldots$ and our aim is to determine f_n for $n=1, 2, \ldots$.

Denote by $r(\mathbf{L})$ the largest nonnegative number for which

$$\sum_{n=0}^{\infty} \|f_n\| \, |\varrho|^n < \infty$$

whenever $g_1 \in \mathbf{R}$, $g_2 \in \mathbf{R}$ and

$$|\varrho| \max (\|g_1\|, \|g_2\|) < r(\mathbf{L}). \tag{32}$$

The inequalities

$$(\|\mathbf{L}\| + \|\mathbf{L}^*\|)^{-1} \leqq r(\mathbf{L}) \leqq c(\mathbf{L})$$

obviously hold; however, later we shall prove that

$$r(\mathbf{L}) = c(\mathbf{L}) = c(\mathbf{T}) \tag{33}$$

where $c(\mathbf{T})$ is defined by (8).

If (32) is satisfied, then

$$F(\varrho) = \sum_{n=0}^{\infty} f_n \varrho^n \tag{34}$$

belongs to $\mathbf{R}$, and if we multiply (1) by ϱ^n and add for $n=1, 2, \ldots$, then we obtain that

$$\mathbf{L}\{F(\varrho)(e-\varrho g_1)\} + \mathbf{L}^*\{(e-\varrho g_2) F(\varrho)\} = f_0. \tag{35}$$

Conversely, if

$$F(\varrho) = \sum_{n=0}^{\infty} f_n^* \varrho^n \tag{36}$$

belongs to $\mathbf{R}$ for $|\varrho| < r$ where r is some positive number, and if (36) satisfies (35), then by forming the coefficient of ϱ^n for $n=0, 1, 2, \ldots$, we obtain that $f_0^* = f_0$ and f_n^* $(n=1, 2, \ldots)$ satisfies the same recurrence formula as f_n $(n=1, 2, \ldots)$. Thus necessarily $f_n^* = f_n$ for $n \geqq 0$.

We shall demonstrate that $F(\varrho)$ can always be determined by using the method of factorization. Let us assume that

$$e - \varrho g_i = g_i^+(\varrho) g_i^-(\varrho) \tag{37}$$

for $|\varrho| \, \|g_i\| < c(\mathbf{T})$ and $i=1, 2$ where $g_i^+(\varrho)$ and $g_i^-(\varrho)$ satisfy the properties (a) and (b) respectively. We have already proved that such a factorization always exists. By using the factorization (37) which depends only on $\mathbf{T}$, we can determine $F(\varrho)$ not only for $\mathbf{L} = \mathbf{T}$ but for any $\mathbf{L}$ satisfying (i), (ii), (iii) and (25).

Theorem 3. *If* $f_0 \in \mathbf{R}$, $g_1 \in \mathbf{R}$, $g_2 \in \mathbf{R}$ *and*

$$f_n = \mathbf{L}\{f_{n-1} g_1\} + \mathbf{L}^*\{g_2 f_{n-1}\}$$

for $n=1, 2, \ldots$, *and if* (32) *is satisfied, then* (34) *belongs to* $\mathbf{R}$ *and we have*

$$F(\varrho) = [g_2^-(\varrho)]^{-1}[\mathbf{L}\{g_2^-(\varrho)f_0[g_1^-(\varrho)]^{-1}\}+\mathbf{L}^*\{[g_2^+(\varrho)]^{-1}f_0g_1^+(\varrho)\}][g_1^+(\varrho)]^{-1} \tag{38}$$

where $g_i^+(\varrho)$ *and* $g_i^-(\varrho)$ *satisfy* (a), (b) *and* (37).

Proof. If $F(\varrho)$ is defined by (38), then it can be represented in the form of (36). Since $\mathbf{T}\{g_i^-(\varrho)\}=\mathbf{T}\{[g_i^-(\varrho)]^{-1}\}=\mathbf{T}\{e\}$ and $\mathbf{T}^*\{g_i^+(\varrho)\}=\mathbf{T}^*\{[g_i^+(\varrho)]^{-1}\}=\mathbf{T}^*\{e\}$ for $i=1, 2$, by (30) and (31) we obtain that

$$\mathbf{L}\{F(\varrho)(e-\varrho g_1)\} = \mathbf{L}\{f_0\} \tag{39}$$

and

$$\mathbf{L}^*\{(e-\varrho g_2)F(\varrho)\} = \mathbf{L}^*\{f_0\}. \tag{40}$$

If we add (39) and (40), then we get (35). Thus we can conclude that (34) can be expressed in the form of (38). This completes the proof of the theorem.

We note that if $\mathbf{L}=\mathbf{T}$ and $f_0=e$, then (38) reduces to

$$F(\varrho) = [g_2^-(\varrho)]^{-1}[g_1^+(\varrho)]^{-1}.$$

Now let us suppose that in Theorem 3 we have $g_1=wg$ and $g_2=zg$ where $g\in\mathbf{R}$ and w and z are complex (or real) numbers. In this case by using the factorization in Theorem 1 we can choose $g_1^+(\varrho)=g^+(\varrho w)$, $g_1^-(\varrho)=g^-(\varrho w)$, $g_2^+(\varrho)=g^+(\varrho z)$, and $g_2^-(\varrho)=g^-(\varrho z)$ in Theorem 3. Then by (38) we get

$$F(\varrho) = [g^-(\varrho z)]^{-1}[\mathbf{L}\{g^-(\varrho z)f_0[g^-(\varrho w)]^{-1}\}+\mathbf{L}^*\{[g^+(\varrho z)]^{-1}f_0g^+(\varrho w)\}][g^+(\varrho w)]^{-1} \tag{41}$$

for $|\varrho|\max(|w|, |z|)\|g\|<r(\mathbf{L})$. If, in particular, $\mathbf{L}=\mathbf{T}$ and $f_0=e$, then (41) reduces to

$$F(\varrho) = [g^-(\varrho z)]^{-1}[g^+(\varrho w)]^{-1}. \tag{42}$$

Now we are going to prove that in Theorem 1 $g^+(\varrho)$ and $g^-(\varrho)$ are uniquely determined by the properties (a) and (b) and by (11).

If $w=1$ and $z=0$ in (42), then the right-hand side becomes $[g^+(\varrho)]^{-1}$. On the other hand in this case by (12) we have $F(\varrho)=a(\varrho)$. Accordingly, $g^+(\varrho)=[a(\varrho)]^{-1}$ necessarily holds. In a similar way, if $w=0$ and $z=1$ in (42), then the right-hand side becomes $[g^-(\varrho)]^{-1}$. On the other hand in this case by (13) we have $F(\varrho)=b(\varrho)$. Accordingly, $g^-(\varrho)=[b(\varrho)]^{-1}$ necessarily holds. This proves that in Theorem 1 $g^+(\varrho)$ and $g^-(\varrho)$ are uniquely determined by the properties (a) and (b) and by (11), and that (20) and (21) necessarily hold.

Having been established that $g_i^+(\varrho)$ and $g_i^-(\varrho)$ $(i=1, 2)$ are uniquely determined in (38) we can express $g_i^+(\varrho)$ and $g_i^-(\varrho)$ by formulas (18) and (19) and $[g_i^+(\varrho)]^{-1}$ and $[g_i^-(\varrho)]^{-1}$ by formulas (12) and (13). Proceeding in this way we can conclude from (38) that

$$r(\mathbf{L}) \geqq c(\mathbf{T}) \tag{43}$$

necessarily holds. Since evidently $r(\mathbf{L}) \leqq c(\mathbf{L})$, by (43) we have $c(\mathbf{T}) \leqq c(\mathbf{L})$. If we interchange the roles of $\mathbf{L}$ and $\mathbf{T}$, then it follows that $c(\mathbf{L}) \leqq c(\mathbf{T})$ also holds. This proves that (26) and (33) are indeed true.

In particular, it follows from (26) and (33) that if $\mathbf{L}$ is defined by (25), and if $\mu(\mathbf{T})=1$ and $\bar{\mu}(\mathbf{T}^*)=1$, then $r(\mathbf{L})=c(\mathbf{L})=1$ regardless of the values of $\|\mathbf{L}\|$ and $\|\mathbf{L}^*\|$.

If, instead of (1), we consider the recurrence formula

$$f_n = \mathbf{L}\{g_1 f_{n-1}\} + \mathbf{L}^*\{f_{n-1} g_2\} \tag{44}$$

for $n=1, 2, \ldots$ where $f_0 \in \mathbf{R}$, $g_1 \in \mathbf{R}$, $g_2 \in \mathbf{R}$ and $\mathbf{L}$ satisfies the conditions (i), (ii), (iii) and can be represented in the form of (25), and if $|\varrho| \max(\|g_1\|, \|g_2\|) < r(\mathbf{L}^*)$, then

$$F(\varrho) = \sum_{n=0}^{\infty} f_n \varrho^n$$

belongs to $\mathbf{R}$ and can be determined again by the method of factorization. Let us suppose that

$$e - \varrho g_i = h_i^-(\varrho) h_i^+(\varrho)$$

for $|\varrho| \, \|g\| < c(\mathbf{T}^*)$ and $i=1, 2$ where $h_i^+(\varrho)$ satisfies property (a) and $h_i^-(\varrho)$ satisfies property (b). In this case we have

$$F(\varrho) = [h_1^+(\varrho)]^{-1} [\mathbf{L}\{[h_1^-(\varrho)]^{-1} f_0 h_2^-(\varrho)\} + \mathbf{L}^*\{h_1^+(\varrho) f_0 [h_2^+(\varrho)]^{-1}\}] [h_2^-(\varrho)]^{-1} \tag{45}$$

whenever $|\varrho| \max(\|g_1\|, \|g_2\|) < r(\mathbf{L}^*)$.

Note. If $\mathbf{R}$ is a commutative Banach algebra and if $f_0 = e$, then (38) and (45) reduce to

$$F(\varrho) = \exp\left\{-\mathbf{L}\{\log(e - \varrho g_1)\} - \mathbf{L}^*\{\log(e - \varrho g_2)\}\right\}$$

for $|\varrho| \max(\|g_1\|, \|g_2\|) < 1$. In some particular cases this last result was demonstrated in 1952 by F. Pollaczek [9] and in 1958 by J. G. Wendel [12].

6. The second recurrence equation. Let us consider the recurrence equation (2) for $n=1, 2, \ldots$ where $f_0 \in \mathbf{R}$, $g_1 \in \mathbf{R}$, $g_2 \in \mathbf{R}$ and $\mathbf{L}$ satisfies the conditions (i), (ii), (iii) and can be represented in the form of (25). Obviously $f_n \in \mathbf{R}$ for $n=1, 2, \ldots$ and our aim is to determine f_n for $n=1, 2, \ldots$.

Denote by $r^*(\mathbf{L})$ the largest nonnegative number for which

$$\sum_{n=0}^{\infty} \|f_n\| |\varrho|^n < \infty \tag{46}$$

whenever $g_1 \in \mathbf{R}$, $g_2 \in \mathbf{R}$ and

$$|\varrho| \max(\|g_1\|, \|g_2\|) < r^*(\mathbf{L}). \tag{47}$$

We shall prove that for every **L**

(48) $$\gamma(\mathbf{T})/3 \leqq r^*(\mathbf{L}) \leqq 1$$

where $\gamma(\mathbf{T})$ is defined by (9). Actually, we shall prove that if

(49) $$|\varrho|[\min(\|g_1\|, \|g_2\|)+\|g_1-g_2\|] < \gamma(\mathbf{T})$$

then (46) is satisfied and this implies (48).

If (47) is satisfied, then

(50) $$F(\varrho) = \sum_{n=0}^{\infty} f_n \varrho^n$$

belongs to **R**, and if we multiply (2) by ϱ^n and add for $n=1, 2, \ldots$, then we obtain that

(51) $$\mathbf{L}\{F(\varrho)(e-\varrho g_1)\}+\mathbf{L}^*\{F(\varrho)(e-\varrho g_2)\} = f_0.$$

Conversely, if

(52) $$F(\varrho) = \sum_{n=0}^{\infty} f_n^* \varrho^n$$

belongs to **R** for $|\varrho|<r$ where r is some positive number, and if (52) satisfies (51), then $f_n^*=f_n$ for all $n\geqq 0$.

The generating function (50) can be determined by the method of factorization. Let us apply Theorem 1 to $(e-\varrho g_2)^{-1}(e-\varrho g_1)=e-\varrho(e-\varrho g_2)^{-1}(g_1-g_2)$ and Theorem 2 to $(e-\varrho g_1)^{-1}(e-\varrho g_2)=e-\varrho(e-\varrho g_1)^{-1}(g_2-g_1)$. If (49) is satisfied, then we can write that

(53) $$(e-\varrho g_2)^{-1}(e-\varrho g_1) = g^+(\varrho)g^-(\varrho)$$

where $g^+(\varrho)$ and $g^-(\varrho)$ satisfy the properties (a) and (b) respectively.

Theorem 4. *If* $f_0\in\mathbf{R}$, $g_1\in\mathbf{R}$, $g_2\in\mathbf{R}$ *and*

$$f_n = \mathbf{L}\{f_{n-1}g_1\}+\mathbf{L}^*\{f_{n-1}g_2\}$$

for $n=1, 2, \ldots$, *and if* (49) *is satisfied, then* (50) *belongs to* **R** *and we have*

(54) $$F(\varrho) = [\mathbf{L}\{f_0[g^-(\varrho)]^{-1}\}+\mathbf{L}^*\{f_0 g^+(\varrho)\}][g^+(\varrho)]^{-1}(e-\varrho g_2)^{-1}$$

where $g^+(\varrho)$ *and* $g^-(\varrho)$ *satisfy* (a), (b) *and* (53).

Proof. If $F(\varrho)$ is given by (54), then it can be represented in the form of (52) and by using (30) and (31) we can prove that (54) satisfies (51). This proves the theorem.

In a similar way as we proved (53) we can prove that if (49) is satisfied, then we can write that

(55) $$(e-\varrho g_1)(e-\varrho g_2)^{-1} = h^-(\varrho)h^+(\varrho)$$

where $h^+(\varrho)$ and $h^-(\varrho)$ satisfy the properties (a) and (b) respectively. By using (55) we can prove the following result.

If $f_0 \in \mathbf{R}$, $g_1 \in \mathbf{R}$, $g_2 \in \mathbf{R}$ and

$$f_n = \mathbf{L}\{g_1 f_{n-1}\} + \mathbf{L}^*\{g_2 f_{n-1}\} \tag{56}$$

for $n=1, 2, \ldots$, and if (49) is satisfied, then

$$F(\varrho) = \sum_{n=0}^{\infty} f_n \varrho^n$$

belongs to $\mathbf{R}$ and we have

$$F(\varrho) = (e-\varrho g_2)^{-1}[h^+(\varrho)]^{-1}[\mathbf{L}\{[h^-(\varrho)]^{-1} f_0\} + \mathbf{L}^*\{h^+(\varrho) f_0\}] \tag{57}$$

where $h^+(\varrho)$ and $h^-(\varrho)$ satisfy (a), (b) and (55).

If, in particular, $\mathbf{L}=\mathbf{T}$ and $f_0=e$ in (54), then we get $F(\varrho)=[g^+(\varrho)]^{-1}(e-\varrho g_2)^{-1}$. Thus $g^+(\varrho)$ can also be determined by the recurrence formula (2). If, in particular, $\mathbf{L}=\mathbf{T}$ and $f_0=e$ in (57), then we get $F(\varrho)=(e-\varrho g_2)^{-1}[h^+(\varrho)]^{-1}$ and thus $h^+(\varrho)$ can also be determined by the recurrence formula (56).

7. A system of recurrence equations. In this section we shall demonstrate that the system of recurrence equations (3) and (4) can be solved by using Theorem 4 if we apply it to a new Banach algebra $\mathbf{S}$ associated with $\mathbf{R}$. Let us denote by $\mathbf{S}$ the space of matrices

$$\mathbf{f} = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \tag{58}$$

where $f_{ij} \in \mathbf{R}$ for $i, j=1, 2$. In $\mathbf{S}$ let us define the operations of addition, multiplication and multiplication by a complex (or real) constant according to the rules of matrix algebra and according to the rules established in $\mathbf{R}$. Define the norm of $\mathbf{f}$ either by

$$\|\mathbf{f}\|_S = \max(\|f_{11}\| + \|f_{12}\|, \|f_{21}\| + \|f_{22}\|)$$

or alternately by

$$\|\mathbf{f}\|_S = \max(\|f_{11}\| + \|f_{21}\|, \|f_{12}\| + \|f_{22}\|).$$

We can easily see that $\mathbf{S}$ is a noncommutative Banach algebra with zero element and identity element

$$\begin{bmatrix} \theta & \theta \\ \theta & \theta \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} e & \theta \\ \theta & e \end{bmatrix},$$

respectively.

If $\mathbf{T}$ is a transformation in $\mathbf{R}$ which satisfies (i), (ii), and (iii), then let us extend the definition of $\mathbf{T}$ to $\mathbf{S}$ in such a way that we form $\mathbf{T}$ element by element for an $\mathbf{f}$ given by (58), that is

$$\mathbf{T}\{\mathbf{f}\} = [\mathbf{T}\{f_{ij}\}]_{ij=1,2}.$$

We can easily see that $\mathbf{T}$ satisfies (i), (ii) and (iii) in the space $\mathbf{S}$ too.

Now let us consider the system of recurrence equations (3) and (4) for $n=1, 2, \ldots$ where $u_0 \in \mathbf{R}$, $v_0 \in \mathbf{R}$, $h_i \in \mathbf{R}$ $(i=1, 2, 3, 4)$ and $\mathbf{L}$ satisfies the conditions (i), (ii), (iii) and can be represented in the form of (25). We can express (3) and (4) in the following matrix form

$$\begin{bmatrix} u_n & v_n \\ \theta & \theta \end{bmatrix} = \mathbf{L}\left\{\begin{bmatrix} u_{n-1} & v_{n-1} \\ \theta & \theta \end{bmatrix}\begin{bmatrix} h_1 & \theta \\ h_2 & \theta \end{bmatrix}\right\} + \mathbf{L}^*\left\{\begin{bmatrix} u_{n-1} & v_{n-1} \\ \theta & \theta \end{bmatrix}\begin{bmatrix} \theta & h_3 \\ \theta & h_4 \end{bmatrix}\right\}$$

for $n=1, 2, \dots$. This equation is of type (2). If we apply Theorem 4 to the Banach algebra $\mathbf{S}$, then

$$\sum_{n=0}^{\infty} \begin{bmatrix} u_n & v_n \\ \theta & \theta \end{bmatrix} \varrho^n$$

can be determined by (54).

If, instead of (3) and (4), we consider the recurrence equations

$$u_n = \mathbf{L}\{h_1 u_{n-1} + h_2 v_{n-1}\}$$

and

$$v_n = \mathbf{L}^*\{h_3 u_{n-1} + h_4 v_{n-1}\}$$

for $n=1, 2, \dots$, then we can write that

$$\begin{bmatrix} u_n & \theta \\ v_n & \theta \end{bmatrix} = \mathbf{L}\left\{\begin{bmatrix} h_1 & h_2 \\ \theta & \theta \end{bmatrix}\begin{bmatrix} u_{n-1} & \theta \\ v_{n-1} & \theta \end{bmatrix}\right\} + \mathbf{L}^*\left\{\begin{bmatrix} \theta & \theta \\ h_3 & h_4 \end{bmatrix}\begin{bmatrix} u_{n-1} & \theta \\ v_{n-1} & \theta \end{bmatrix}\right\} \tag{59}$$

for $n=1, 2, \dots$. This equation is of type (56). The solution of (59) can be obtained by (57) if we apply it to the Banach algebra $\mathbf{S}$.

By introducing a Banach algebra of finite or countably infinite matrices with elements belonging to $\mathbf{R}$, we can solve a finite or a countably infinite system of recurrence equations in $\mathbf{R}$.

In the next two sections we shall define two Banach algebras $\mathbf{R}_1$ and $\mathbf{R}_2$, and theree transformations $\mathbf{T}, \mathbf{T}_0, \mathbf{T}_1$ satisfying (i), (ii) and (iii). If we apply Theorem 3 and Theorem 4 to these Banach algebras, then we can determine the distributions of several random variables depending on the partial sums of a sequence of independent and identically distributed random variables and of a semi-Markov sequence of real random variables. In particular, we can find the distributions of the maximal partial sum, the ordered partial sums, the number of positive partial sums, the number of changes of sign in the successive partial sums, and the subscript of the first positive partial sum. These applications will be discussed in a subsequent paper.

8. A commutative Banach algebra $\mathbf{R}_1$. Let us define $\mathbf{R}_1$ as the space of functions $\Phi(s)$ defined for $\operatorname{Re}(s)=0$ on the complex plane which can be represented in the form

$$\Phi(s) = \mathbf{E}\{\zeta e^{-s\eta}\} \tag{60}$$

where η is a real random variable and ζ is a complex (or real) random variable for which $\mathbf{E}\{|\zeta|\} < \infty$. Let us define in $\mathbf{R}_1$ the operations to be the pointwise addition, multiplication and multiplication by a complex (or real) constant. The zero element of $\mathbf{R}_1$ is 0, and the identity element of $\mathbf{R}_1$ is 1. Let us define the norm of $\Phi(s)\in\mathbf{R}_1$ by

$$\|\Phi\| = \inf_{\zeta} \mathbf{E}\{|\zeta|\}$$

by where the infimum is taken for all admissible ζ in the representation (60).

We can easily prove that $\mathbf{R}_1$ is a commutative Banach algebra.

Now we shall consider some transformations in $\mathbf{R}_1$ which satisfy (i), (ii) and (iii).

If $\Phi(s)\in\mathbf{R}_1$ is given by (60), then let us define

$$\Phi^+(s) = \mathbf{E}\{\zeta e^{-s\eta^+}\} \tag{61}$$

for $\mathrm{Re}(s)\geqq 0$ and

$$\Phi^-(s) = \mathbf{E}\{\zeta(e^{-s\eta} - e^{-s\eta^+})\} \tag{62}$$

for $\mathrm{Re}(s)\leqq 0$ where $\eta^+ = \max(0, \eta)$. We have $\Phi^+(s)\in\mathbf{R}_1$, $\Phi^-(s)\in\mathbf{R}_1$ and

$$\Phi(s) = \Phi^+(s) + \Phi^-(s) \tag{63}$$

for $\mathrm{Re}(s)=0$, $|\Phi^+(s)|\leqq\|\Phi\|$ for $\mathrm{Re}(s)\geqq 0$ and $|\Phi^-(s)|\leqq 2\|\Phi\|$ for $\mathrm{Re}(s)\leqq 0$.

The function $\Phi^+(s)$ is regular for $\mathrm{Re}(s)>0$, continuous and bounded for $\mathrm{Re}(s)\geqq 0$ and $\Phi^+(0)=\Phi(0)$.

The function $\Phi^-(s)$ is reqular for $\mathrm{Re}(s)<0$, continuous and bounded for $\mathrm{Re}(s)\leqq 0$ and $\Phi^-(0)=0$.

By Liouville's theorem it follows that the above properties uniquely determine $\Phi^+(s)$ and $\Phi^-(s)$ in the representation (63).

If $\Phi(s)\in\mathbf{R}_1$, then for $\mathrm{Re}(s)>0$ we have

$$\Phi^+(s) = \frac{1}{2}\Phi(0) + \lim_{\varepsilon\to 0}\frac{s}{2\pi i}\int_{L_\varepsilon}\frac{\Phi(z)}{z(s-z)}\,dz$$

where $L_\varepsilon = \{z : z = iy,\ -\infty < y \leqq -\varepsilon < \varepsilon \leqq y < \infty\}$. See reference [11].

For any event A let us define $\delta(A)$ as the indicator variable of A, that is, $\delta(A)=1$ if A occurs and $\delta(A)=0$ if A does not occur.

Now we define three transformations $\mathbf{T}$, $\mathbf{T}_0$, $\mathbf{T}_1$ in $\mathbf{R}_1$ which satisfy the conditions (i), (ii) and (iii). If $\Phi(s)\in\mathbf{R}_1$ is given by (60), then let

$$\mathbf{T}\{\Phi(s)\} = \Phi^+(s) = \mathbf{E}\{\zeta e^{-s\eta^+}\}, \tag{64}$$

$$\mathbf{T}_0\{\Phi(s)\} = \Phi^+(s) - \Phi^+(\infty) = \mathbf{E}\{\zeta e^{-s\eta}\delta(\eta > 0)\} \tag{65}$$

and

$$\mathbf{T}_1\{\Phi(s)\} = \Phi^+(s) + \Phi^-(-\infty) = \mathbf{E}\{\zeta e^{-s\eta}\delta(\eta \geqq 0)\}. \tag{66}$$

We define $\mathbf{T}^*$, $\mathbf{T}_0^*$ and $\mathbf{T}_1^*$ by (6). We can easily see that these transformations satisfy (i), (ii), (iii), $\|\mathbf{T}\| = \|\mathbf{T}_0\| = \|\mathbf{T}_1\| = \|\mathbf{T}_0^*\| = \|\mathbf{T}_1^*\| = 1$ and $\|\mathbf{T}^*\| = 2$.

If $\mathbf{L}$ is any one of the transformations $\mathbf{T}$, $\mathbf{T}_0$, $\mathbf{T}_1$, defined by (64), (65), and (66) respectively, then $\mathbf{L}\{\Phi(s)\}$ can be represented in the form of (25), that is,

$$\mathbf{L}\{\Phi(s)\} = \mathbf{T}\{\Phi(s)\} - \alpha(\Phi)$$

where $\mathbf{T}\{\Phi(s)\}$ is defined by (64), and $\alpha(\Phi)\equiv 0$ for $\mathbf{L}=\mathbf{T}$, $\alpha(\Phi)=\Phi^{+}(\infty)$ for $\mathbf{L}=\mathbf{T}_0$, and $\alpha(\Phi)=-\Phi^{-}(-\infty)$ for $\mathbf{L}=\mathbf{T}_1$. If $\mathbf{L}$ is any one of the transformations (64), (65), (66), then by (7), (8) and (26) we have $c(\mathbf{L})=1$ and $c(\mathbf{L}^*)=1$.

If we assume that $\mathbf{T}$ is given by (64), then we can formulate the following version of Theorem 1.

Theorem 5. *If* $\psi(s)\in\mathbf{R}_1$ *and if* $|\varrho|\,\|\psi\|<1$, *then there exist two functions* $\psi^{+}(s,\varrho)\in\mathbf{R}_1$ *and* $\psi^{-}(s,\varrho)\in\mathbf{R}_1$ *such that*

(67) $$1-\varrho\psi(s) = \psi^{+}(s,\varrho)\psi^{-}(s,\varrho)$$

for $\mathrm{Re}(s)=0$ *where* $\psi^{+}(s,\varrho)$ *satisfies property* (α) *and* $\psi^{-}(s,\varrho)$ *satisfies property* (β) *stated below.*

Property (α). *The function* $\psi^{+}(s,\varrho)$ *is regular for* $\mathrm{Re}(s)>0$, *continuous, bounded and free from zeros for* $\mathrm{Re}(s)\geqq 0$.

Property (β). *The function* $\psi^{-}(s,\varrho)$ *is regular for* $\mathrm{Re}(s)<0$, *continuous, bounded and free from zeros for* $\mathrm{Re}(s)\leqq 0$.

Proof. If $\psi^{+}(s,\varrho)$ satisfies (α), and $\psi^{-}(s,\varrho)$ satisfies (β), then we say that (67) is a factorization of $1-\varrho\psi(s)$. Such a factorization always exists. For example, if

(68) $$\psi^{+}(s,\varrho) = C_1(\varrho)\exp\{\mathbf{T}\{\log[1-\varrho\psi(s)]\}\}$$

for $\mathrm{Re}(s)\geqq 0$ and $|\varrho|\,\|\psi\|<1$, and

(69) $$\psi^{-}(s,\varrho) = C_2(\varrho)\exp\{\mathbf{T}^*\{\log[1-\varrho\psi(s)]\}\}$$

for $\mathrm{Re}(s)\leqq 0$ and $|\varrho|\,\|\psi\|<1$, where $C_1(\varrho)C_2(\varrho)=1$, then ($\alpha$), ($\beta$) and (67) are satisfied. Conversely, it follows from Liouville's theorem that conditions (α), (β) and (67) determine $\psi^{+}(s,\varrho)$ and $\psi^{-}(s,\varrho)$ up to a nonvanishing factor depending only on ϱ. Thus (68) and (69) are the general forms of $\psi^{+}(s,\varrho)$ are $\psi^{-}(s,\varrho)$ respectively.

If in (68) and (69) we choose $C_1(\varrho)$ and $C_2(\varrho)$ in an appropriate way, then we can easily see that $\psi^{+}(s,\varrho)$ and $\psi^{-}(s,\varrho)$ satisfy properties (a) and (b) too.

If we want to solve a recurrence equation of type (1) in the space $\mathbf{R}_1$, then instead of (11) we can use the factorization (67). Since in (38) only the product $C_1(\varrho)C_2(\varrho)=1$ appears, therefore it does not matter how we choose $C_1(\varrho)$ and $C_2(\varrho)$ in (68) and (69).

Let us mention one example specifically. Let

$$U_n(s) = w\mathbf{L}\{U_{n-1}(s)\psi(s)\}+z\mathbf{L}^*\{U_{n-1}(s)\psi(s)\}$$

for $n=1,2,\ldots$ where $U_0(s)\in\mathbf{R}_1$, $\psi(s)\in\mathbf{R}_1$, w and z are complex (or real) numbers, and $\mathbf{L}$ is any one of the transformations (64), (65), (66). If $|\varrho|\max(|w|,|z|)\,\|\psi\|<1$, then

$$U(s,\varrho) = \sum_{n=0}^{\infty} U_n(s)\varrho^n$$

belongs to $\mathbf{R}_1$ and by Theorem 3 we have

$$U(s, \varrho) = \big[\mathbf{L}\{U_0(s)\psi^-(s, \varrho z)[\psi^-(s, \varrho w)]^{-1}\}+ \\ +\mathbf{L}^*\{U_0(s)\psi^+(s, \varrho w)[\psi^+(s, \varrho z)]^{-1}\}\big][\psi^+(s, \varrho w)]^{-1}[\psi^-(s, \varrho z)]^{-1}$$

where $\psi^+(s, \varrho)$ and $\psi^-(s, \varrho)$ are determined by Theorem 5 or by (68) and (69), respectively.

Finally, we note that in properties (α) and (β) the requirement of boundedness can be replaced by the weaker conditions $\lim_{|s|\to\infty} [\log\psi^+(s, \varrho)]/s=0$ $(\mathrm{Re}(s)\geqq 0)$ and $\lim_{|s|\to\infty} [\log\psi^-(s, \varrho)]/s=0$ $(\mathrm{Re}(s)\leqq 0)$, respectively.

9. A noncommutative Banach algebra $\mathbf{R}_2$. Let I be a fixed finite or countably infinite set. We consider complex (or real) matrices $\mathbf{A}=[a_{ij}]$, $i\in I$, $j\in I$, for which

$$\mathbf{M}\{\mathbf{A}\} = \sup_{i\in I} \sum_{j\in I} |a_{ij}| < \infty.$$

We shall denote by $\mathbf{0}$ the zero matrix all of whose elements are zeros, and by $\mathbf{I}$ the identity matrix. ($\mathbf{I}=[\delta_{ij}]$, $i\in I$, $j\in I$, where $\delta_{ij}=1$ for $i=j$ and $\delta_{ij}=0$ for $i\neq j$.) If $\mathbf{M}\{\mathbf{A}\}<\infty$, $\mathbf{M}\{\mathbf{B}\}<\infty$ and $\mathbf{AB}=\mathbf{BA}=\mathbf{I}$, then we say that $\mathbf{A}$ and $\mathbf{B}$ are inverse matrices and write $\mathbf{B}=\mathbf{A}^{-1}$.

We say that a matrix function $\mathbf{A}(s)=[a_{ij}(s)]$, $i\in I$, $j\in I$, is continuous, or regular, or bounded on a set D according to whether every $a_{ij}(s)$ is continuous on D, or every $a_{ij}(s)$ is regular on D, or $\mathbf{M}\{\mathbf{A}(s)\}<K$ for $s\in D$ where K is a positive constant.

Let $\mathbf{R}_2$ be the space of all matrix functions

(70) $$\mathbf{\Phi}(s) = [\Phi_{ij}(s)]_{i,j\in I}$$

defined for $\mathrm{Re}(s)=0$ on the complex plane such that I is a fixed countable set, $\Phi_{ij}(s)\in \in\mathbf{R}_1$ and

(71) $$\|\mathbf{\Phi}\| = \sup_{i\in I} \sum_{j\in I} \|\Phi_{ij}\|_{\mathbf{R}_1} < \infty.$$

We define the norm of $\mathbf{\Phi}(s)$ by (71). Let us define the operations of addition, multiplication and multiplication by a complex (or real) constant in $\mathbf{R}_2$ according to the rules of matrix algebra. We can easily see that $\mathbf{R}_2$ is a noncommutative Banach algebra with zero element $\mathbf{0}$ and identity element $\mathbf{I}$.

If $\mathbf{\Phi}(s)\in\mathbf{R}_2$ is given by (70), then let

$$\mathbf{\Phi}^+(s) = [\Phi_{ij}^+(s)]_{i,j\in I}$$

for $\mathrm{Re}(s)\geqq 0$ and

$$\mathbf{\Phi}^-(s) = [\Phi_{ij}^-(s)]_{i,j\in I}$$

for $\mathrm{Re}(s)\leqq 0$ where $\Phi_{ij}^+(s)$ is defined by (61) and $\Phi_{ij}^-(s)$ by (62).

Obviously, $\mathbf{\Phi}^+(s)\in\mathbf{R}_2$, $\mathbf{\Phi}^-(s)\in\mathbf{R}_2$ and

(72) $$\mathbf{\Phi}(s) = \mathbf{\Phi}^+(s)+\mathbf{\Phi}^-(s)$$

for $\mathrm{Re}(s)=0$. We have $\mathbf{M}\{\mathbf{\Phi}^+(s)\}\leqq\|\mathbf{\Phi}\|$ for $\mathrm{Re}(s)\geqq 0$ and $\mathbf{M}\{\mathbf{\Phi}^-(s)\}\leqq 2\|\mathbf{\Phi}\|$ for $\mathrm{Re}(s)\leqq 0$.

The matrix function $\mathbf{\Phi}^+(s)$ is regular for $\mathrm{Re}(s)>0$, continuous and bounded for $\mathrm{Re}(s)\geqq 0$ and $\mathbf{\Phi}^+(0)=\mathbf{\Phi}(0)$.

The matrix function $\mathbf{\Phi}^-(s)$ is regular for $\mathrm{Re}(s)<0$, continuous and bounded for $\mathrm{Re}(s)\leqq 0$ and $\mathbf{\Phi}^-(0)=\mathbf{0}$.

By Liouville's theorem it follows that the above properties uniquely determine $\mathbf{\Phi}^+(s)$ and $\mathbf{\Phi}^-(s)$ in the representation (72).

Now let us extend the definition of the transformations (64), (65), (66) from the space $\mathbf{R}_1$ to the space $\mathbf{R}_2$ in such a way that we form these transformations element by element for $\mathbf{\Phi}(s)\in\mathbf{R}_2$, that is,

$$\mathbf{T}\{\mathbf{\Phi}(s)\} = \mathbf{\Phi}^+(s), \tag{73}$$

$$\mathbf{T}_0\{\mathbf{\Phi}(s)\} = \mathbf{\Phi}^+(s)-\mathbf{\Phi}^+(\infty), \tag{74}$$

and

$$\mathbf{T}_1\{\mathbf{\Phi}(s)\} = \mathbf{\Phi}^+(s)+\mathbf{\Phi}^-(-\infty). \tag{75}$$

We define $\mathbf{T}^*$, $\mathbf{T}_0^*$, $\mathbf{T}_1^*$ by (6). We can easily see that these transformations satisfy (i), (ii), (iii), $\|\mathbf{T}\|=\|\mathbf{T}_0\|=\|\mathbf{T}_1\|=\|\mathbf{T}_0^*\|=\|\mathbf{T}_1^*\|=1$ and $\|\mathbf{T}^*\|=2$.

If $\mathbf{L}$ is any one of the transformations (73), (74), (75) and if $\mathbf{\Phi}(s)\in\mathbf{R}_2$, then $\mathbf{L}\{\mathbf{C\Phi}(s)\}=\mathbf{CL}\{\mathbf{\Phi}(s)\}$ and $\mathbf{L}\{\mathbf{\Phi}(s)\mathbf{C}\}=\mathbf{L}\{\mathbf{\Phi}(s)\}\mathbf{C}$ for any constant matrix $\mathbf{C}$ for which $\mathbf{M}\{\mathbf{C}\}<\infty$. Furthermore, $\mathbf{L}\{\mathbf{\Phi}(s)\}$ can be represented in the following form

$$\mathbf{L}\{\mathbf{\Phi}(s)\} = \mathbf{T}\{\mathbf{\Phi}(s)\}-\boldsymbol{\alpha}(\mathbf{\Phi}) \tag{76}$$

where $\mathbf{T}\{\mathbf{\Phi}(s)\}$ is defined by (73), $\boldsymbol{\alpha}(\mathbf{\Phi})=\mathbf{0}$ for $\mathbf{L}=\mathbf{T}$, $\boldsymbol{\alpha}(\mathbf{\Phi})=\mathbf{\Phi}^+(\infty)$ for $\mathbf{L}=\mathbf{T}_0$, and $\boldsymbol{\alpha}(\mathbf{\Phi})=-\mathbf{\Phi}^-(-\infty)$ for $\mathbf{L}=\mathbf{T}_1$. If $\mathbf{L}$ is any one of the transformations (73), (74), (75), then by (7), (8) and (26) we have $c(\mathbf{L})=1$ and $c(\mathbf{L}^*)=1$.

If we assume that $\mathbf{T}$ is defined by (73), then we can formulate the following version of Theorem 1.

Theorem 6. *If* $\mathbf{\Psi}(s)\in\mathbf{R}_2$ *and if* $|\varrho|\,\|\mathbf{\Psi}\|<1$, *then there exist two matrices* $\mathbf{\Psi}^+(s,\varrho)\in\mathbf{R}_2$ *and* $\mathbf{\Psi}^-(s,\varrho)\in\mathbf{R}_2$ *such that*

$$\mathbf{I}-\varrho\mathbf{\Psi}(s) = \mathbf{\Psi}^+(s,\varrho)\mathbf{\Psi}^-(s,\varrho) \tag{77}$$

for $\mathrm{Re}(s)=0$ *where* $\mathbf{\Psi}^+(s,\varrho)$ *satisfies property* (α) *and* $\mathbf{\Psi}^-(s,\varrho)$ *satisfies property* (β) *stated below.*

Property (α). *The matrix* $\mathbf{\Psi}^+(s,\varrho)$ *has an inverse* $[\mathbf{\Psi}^+(s,\varrho)]^{-1}$ *for* $\mathrm{Re}(s)\geqq 0$, *and* $\mathbf{\Psi}^+(s,\varrho)$ *and* $[\mathbf{\Psi}^+(s,\varrho)]^{-1}$ *are bounded and continuous for* $\mathrm{Re}(s)\geqq 0$ *and regular for* $\mathrm{Re}(s)>0$.

Property (β). *The matrix* $\mathbf{\Psi}^-(s, \varrho)$ *has an inverse* $[\mathbf{\Psi}^-(s, \varrho)]^{-1}$ *for* $\text{Re}(s) \leqq 0$, *and* $\mathbf{\Psi}^-(s, \varrho)$ *and* $[\mathbf{\Psi}^-(s, \varrho)]^{-1}$ *are bounded and continuous for* $\text{Re}(s) \leqq 0$ *and regular for* $\text{Re}(s) < 0$.

Proof. The factorization (77) satisfying (α) and (β) always exists. By the method described in the proof of Theorem 1 we can construct two matrices $\mathbf{A}(s, \varrho)$ and $\mathbf{B}(s, \varrho)$ such that

$$\mathbf{I} - \varrho\mathbf{\Psi}(s) = [\mathbf{A}(s, \varrho)]^{-1}[\mathbf{B}(s, \varrho)]^{-1}$$

for $\text{Re}(s)=0$ and $\mathbf{A}(s, \varrho)$ satisfies (a) and $\mathbf{B}(s, \varrho)$ satisfies (b).

If we define

$$\mathbf{\Psi}^+(s, \varrho) = [\mathbf{A}(s, \varrho)]^{-1}\mathbf{C}_1(\varrho) \tag{78}$$

for $\text{Re}(s) \geqq 0$ and

$$\mathbf{\Psi}^-(s, \varrho) = \mathbf{C}_2(\varrho)[\mathbf{B}(s, \varrho)]^{-1} \tag{79}$$

for $\text{Re}(s) \leqq 0$ where $\mathbf{M}\{\mathbf{C}_1(\varrho)\} < \infty$, $\mathbf{M}\{\mathbf{C}_2(\varrho)\} < \infty$ and $\mathbf{C}_1(\varrho)\mathbf{C}_2(\varrho) = \mathbf{I}$, then all the properties stated in Theorem 6 are satisfied. Conversely, it follows from Liouville's theorem that conditions (α), (β) and (77) determine $\mathbf{\Psi}^+(s, \varrho)$ and $\mathbf{\Psi}^-(s, \varrho)$ up to a matrix factor independent of s. This implies that (78) and (79) are the general forms of $\mathbf{\Psi}^+(s, \varrho)$ and $\mathbf{\Psi}^-(s, \varrho)$ respectively.

In a similar way as we proved Theorem 6, we can prove a corresponding version of Theorem 2.

If we want to solve recurrence equations of type (1) and (2) in the space $\mathbf{R}_2$, then instead of (11), we can use the factorization (77). Since in (38) and in (54) only the product $\mathbf{C}_1(\varrho)\mathbf{C}_2(\varrho) = \mathbf{I}$ appears, it is immaterial how we choose $\mathbf{C}_1(\varrho)$ and $\mathbf{C}_2(\varrho)$ in (78) and (79). We can easily see that although in (76) $\boldsymbol{\alpha}(\mathbf{\Phi})$ is a matrix, not a scalar, we can use formulas (38) and (54) unchangeably. Recurrence equations of types (44) and (56) in the space $\mathbf{R}_2$ can be solved in a similar way by using an analogous version of Theorem 6.

Let us mention one example specefically. Let

$$\mathbf{U}_n(s) = w\mathbf{L}\{\mathbf{U}_{n-1}(s)\mathbf{\Psi}(s)\} + z\mathbf{L}^*\{\mathbf{U}_{n-1}(s)\mathbf{\Psi}(s)\}$$

for $n=1, 2, \ldots$ where $\mathbf{U}_0(s) \in \mathbf{R}_2$, $\mathbf{\Psi}(s) \in \mathbf{R}_2$, w and z are complex (or real) numbers, and $\mathbf{L}$ is any one of the transformations (73), (74), (75). If $|\varrho|\,[\min(|w|, |z|) + |w-z|] \cdot \cdot \|\mathbf{\Psi}\| < 1$, then

$$\mathbf{U}(s, \varrho) = \sum_{n=0}^{\infty} \mathbf{U}_n(s)\varrho^n$$

belongs to $\mathbf{R}_2$ and by Theorem 4 we have

$$\mathbf{U}(s, \varrho) = \big[\mathbf{L}\{\mathbf{U}_0(s)[\mathbf{\Psi}^-(s, \varrho w, \varrho z)]^{-1}\} + \\ + \mathbf{L}^*\{\mathbf{U}_0(s)\mathbf{\Psi}^+(s, \varrho w, \varrho z)\}\big] \cdot [\mathbf{\Psi}^+(s, \varrho w, \varrho z)]^{-1}[\mathbf{I} - \varrho z\mathbf{\Psi}(\varrho)]^{-1}$$

where

$$[\mathbf{I}-\varrho z\mathbf{\Psi}(s)]^{-1}[\mathbf{I}-\varrho w\mathbf{\Psi}(s)] = \mathbf{\Psi}^{+}(s, \varrho w, \varrho z)\mathbf{\Psi}^{-}(s, \varrho w, \varrho z)$$

for $\mathrm{Re}(s)=0$ and $\mathbf{\Psi}^{+}(s, \varrho w, \varrho z)$ satisfies property (α) and $\mathbf{\Psi}^{-}(s, \varrho w, \varrho z)$ satisfies property (β) in Theorem 6.

We note that in the case of finite matrices the method of matrix factorization has already been used in several fields of mathematics, namely, in the theory of systems of integral equations, in the theory of linear prediction of multivariate stationary stochastic processes and in the theory of Markov chains. We refer to the works of G. D. Birkhoff [3], N. Wiener [13], P. Masani [6], N. Wiener and P. Masani [14], I. C. Gohberg and M. G. Krein [5], M. D. Miller [7], [8] and É. L. Presman [10].

References

[1] G. Baxter, An operator identity, *Pacific J. Math.*, **8** (1958), 649—663.

[2] G. Baxter, A two-dimensional operator identity with application to the change of sign in sums of random variables, *Transactions American Math. Soc.*, **96** (1960), 210—221.

[3] G. D. Birkhoff, A theorem on matrices of analytic functions, *Math. Annalen*, **74** (1913), 122—133.

[4] I. C. Gohberg, The factorization problem for operator functions, *Izvestija Akad. Nauk SSSR, Ser. Math.*, **28** (1964), 1055—1082. (Russian) [English translation: *American Math. Soc. Translations* (2), **49** (1966), 130—161.]

[5] I. C. Gohberg and M. G. Krein, Systems of integral equations on a half line with kernels depending on the difference of arguments, *Uspehi Matem. Nauk*, **13** No. 2 (1958), 3—72. (Russian) [English translation: *American Math. Soc. Translations*, (2) **14** (1960) 217—287.]

[6] P. Masani, The Laurent factorization of operator-valued functions, *Proc. London Math. Soc.*, (3), **6** (1956), 59—69.

[7] H. D. Miller, A matrix factorization problem in the theory of random variables defined on a finite Markov chain, *Proc. Cambridge Phil. Soc.*, **58** (1962), 268—285.

[8] H. D. Miller, Absorption probabilities for sums of random variables defined on a finite Markov chain, *Proc. Cambridge Phil. Soc.*, **58** (1962), 286—298.

[9] F. Pollaczek, Fonctions caractéristiques de certaines rèpartitions définies au moyen de la notion d'ordre. Application à la théorie des attentes, *C. R. Acad. Sci. Paris*, **234** (1952), 2334—2336.

[10] È. L. Presman, Factorization methods and boundary problems for sums of random variables given on Markov chains, *Izvestija Akad. Nauk SSSR, Ser. Mat.*, **33** (1969), 861—900. (Russian), [English translation: *Math. USSR Izvestija*, **3** (1969), 815—852.]

[11] L. Takács, On a linear transformation in the theory of probability, *Acta Sci. Math.*, **33** (1972), 15—24.

[12] J. G. Wendel, Spitzer's formula: A short proof, *Proc. American Math. Soc.*, **9** (1958), 905—908.

[13] N. Weiner, On the factorization of matrices, *Comment. Math. Helv.*, **29** (1955), 97—111.

[14] N. Wiener and P. Masani, The prediction theory of multivariate stochastic processes, I—II, *Acta Math.*, **98** (1957), 111—150 and **99** (1958), 93—137.

CASE WESTERN RESERVE UNIVERSITY
CLEVELAND, OHIO 44106 USA

Acta Sci. Math., 38 (1976), 417—421

On a property of strictly logarithmic concave functions

V. A. TOMILENKO

1. Introduction. In the work [1] by A. PRÉKOPA the following theorem was proved.

Theorem 1. *Let $f(x, y)$ be a function of $n+m$ variables, where x is an n-component and y is an m-component vector. Suppose that f is logarithmic concave in R^{n+m} and let A be a convex subset of R^m. Then the function*

$$I(x) = \int_A f(x, y)\, dy$$

is logarithmic concave in the entire space R^n.

The main result of this work is a similar statement for strictly logarithmic concave functions.

Let f be a non-negative logarithmic concave function in R^{n+m}. We denote

$$D = \{z \in R^{n+m} : f(z) > 0\}, \quad D(x) = \{y \in R^m : f(x, y) > 0\}, \quad B = \{x \in R^n : I(x) > 0\}.$$

The sets $D(x)$ ($x \in R^n$), D and B are convex in R^m, R^{n+m} and R^n, respectively. The relative interior of a convex set $C \subset R^k$ is denoted by $\operatorname{ri} C$ (see [2] p. 57) and the closure of C by $\bar{C}$. The basic theorem of this work is

Theorem 2. *Let $f(x, y)$ be a function of $n+m$ variables where $x \in R^n$, $y \in R^m$. Suppose f is logarithmic concave in R^{n+m} and strictly logarithmic concave in ri D, and let A be convex subset of the space R^m. If the sets $D(x) \subset R^m$ are bounded for every $x \in R^n$, then the function I is logarithmic concave in the entire space R^n and strictly logarithmic concave in ri B.*

The first part of this statement is just Theorem 1. We shall begin with proving the strictly logarithmic concavity of the function I in $\operatorname{ri} B$ with subsidiary statements.

In this work the terminology has been taken from [2].

2. Auxiliary statements. We define the function $g: R^{n+m} \to R$ as follows

$$g(z) = -\ln f(z), \quad z = (x, y) \in R^{n+m}.$$

Received June 13, 1975.

Under the conditions imposed on f, g is a proper convex function with effective domain

$$\operatorname{dom} g = \{z \in R^{n+m} : g(z) < \infty\} = D.$$

We denote

$$f_*(z) = \limsup_{v \to z} f(v),\ v,\ z \in R^{n+m}.$$

Lemma 1. *For all* $z \in R^{n+m}$

$$(\operatorname{cl} g)(z) = -\ln f_*(z),$$

where cl g *is the closure of the convex function* g.

Proof. From the definition of cl g ([2] p. 67—68) and g we have

$$(\operatorname{cl} g)(z) = \liminf_{v \to z} g(v) = \liminf_{v \to z} [-\ln f(v)] = -\limsup_{v \to z} \ln f(v).$$

The continuity and strict monotonicity of the logarithm implies that

$$\limsup_{v \to z} \ln f(v) = \ln [\limsup_{v \to z} f(v)] = \ln f_*(z).$$

The lemma is proved.

Corollary 1. *The function* f_* *is logarithmic concave in* R^{n+m}.

Corollary 2. *The function* f *agrees with* f_* *in* R^{n+m} *except perhaps at relative boundary points of a convex set* D.

Corollaries 1 and 2 follow from Theorem 7.4 [2] and Lemma 1.

Lemma 2. *If* f *is upper semi-continuous on the closed bounded set* $D \subset R^k$, *then there exists* $z_0 \in D$ *such that*

$$\sup_{z \in D} f(z) = f(z_0).$$

Proof. Let $\sup_{z \in D} f(z) = C$ and $\varepsilon_n > 0$, $\varepsilon_n \to 0$ as $n \to \infty$. Then one can find a sequence $\{z_n\} \subset D$ such that for $n = 1, 2, \ldots$

$$f(z_n) > C - \varepsilon_n.$$

Since D is a bounded closed set without loss of generality we may assume that

$$z_n \to z_0 \quad \text{as} \quad n \to \infty, \quad z_0 \in D, \quad \text{and} \quad |z_n - z_0| \leqq \varepsilon_n \quad \text{for} \quad n = 1, 2, \ldots.$$

Hence the inequality

$$\text{(1)} \qquad \sup_{|z_0 - z| < \varepsilon_n} f(z) \geqq f(z_n) > C - \varepsilon_n, \quad n = 1, 2, \ldots$$

is valid. Taking into account the upper semi-continuity of the function f we get from (1) that

$$f(z_0) = \lim_{n \to \infty} \sup_{|z - z_0| < \varepsilon_n} f(z) \geqq C.$$

Thus $f(z_0) = C$. The lemma is proved.

Lemma 3. *Let* $z_1, z_2 \in R^{n+m}$ *and* $0<\lambda<1$. *If* f *is strictly logarithmic concave in* $\operatorname{ri} D \subset R^{n+m}$ *and* $\lambda z_1+(1-\lambda)z_2 \in \operatorname{ri} D$, *then the inequality*

$$f_*(\lambda z_1+(1-\lambda)z_2) > f_*^{\lambda}(z_1) f_*^{1-\lambda}(z_2) \tag{2}$$

is valid.

Proof. Two cases are possible.

(i) One of the points, either z_1 or z_2, does not belong to $\bar{D}$. In this case inequality (2) is obviously correct.

(ii) Let $z_1, z_2 \in \bar{D}$. Let us draw a straight line l across the points z_1 and z_2 and choose some point $z \in l \cap \operatorname{ri} D$. Let $\varphi(\mu)=g(\mu z_1+(1-\mu)z_2)$. Then $\operatorname{cl}\varphi$ is a proper strictly convex function on [0, 1]. From Theorems 7.4 and 7.5 of [2] it follows that $(\operatorname{cl}\varphi)(\mu)=\varphi(\mu)$ for $\mu \in (0, 1)$ and

$$(\operatorname{cl}\varphi)(1) = \lim_{\nu \uparrow 1} (\nu+(1-\nu)\mu_0) = \lim_{\nu \uparrow 1} g(\nu z_1+(1-\nu)z) = (\operatorname{cl} g)(z_1),$$

$$(\operatorname{cl}\varphi)(0) = \lim_{\nu \uparrow 1} (\mu_0-\nu\mu_0) = \lim_{\nu \uparrow 1} g(\nu z_2+(1-\nu)z) = (\operatorname{cl} g)(z_2),$$

where $z=\mu_0 z_1+(1-\mu_0)z_2$. This means that the function $\operatorname{cl} g$ is strictly convex on the set $1 \cap \bar{D}$, that is

$$(\operatorname{cl} g)(\lambda z_1+(1-\lambda)z_2) < \lambda(\operatorname{cl} g)(z_1)+(1-\lambda)(\operatorname{cl} g)(z_2), \quad 0<\lambda<1. \tag{3}$$

From (3) and Lemma 1 it can be seen that inequality (2) is true. The lemma is proved.

Corollary 3. *Let* $z_1, z_2 \in R^{n+m}$ *and* $0<\lambda<1$. *If* f *is strictly logarithmic concave in* $\operatorname{ri} D \subset R^{n+m}$ *and* $\lambda z_1+(1-\lambda)z_2 \in \operatorname{ri} D$, *then we have the inequality*

$$f(\lambda z_1+(1-\lambda)z_2) > f^{\lambda}(z_1) f^{1-\lambda}(z_2).$$

Lemma 4. *If* $x_0 \in \operatorname{ri} B$, $y_0 \in \operatorname{int} D(x_0)$, *then* $z_0=(x_0 y_0) \in \operatorname{ri} D$.

Proof. Let P be the projection $(x, y) \to x$ from R^{n+m} onto R^n. It can be shown that $B \subset PD$ and if B is not empty then the dimension of the set B agrees with that of PD. Hence $\operatorname{ri} B \subset \operatorname{ri}(PD)$ and the point $(x_0, y_0) \in \operatorname{ri} D$ by Theorem 6.8 of [2]. The lemma is proved.

3. Proof of Theorem 2. We denote

$$D_*(x) = \{y \in R^m : f_*(x, y) > 0\}.$$

For all $x \in \operatorname{ri} B$ the sets $D(x)$ and $D_*(x)$ have the same closure and the same interior (see Corollary 2).

Let $x_1, x_2 \in \text{ri } B$, $0<\lambda<1$ and $x_0=\lambda x_1+(1-\lambda)x_2$. We define the functions f_1 and f_2 as follows:

$$f_1(y)=f_*(x_1, y) \quad \text{if} \quad y\in\bar{A}, \quad \text{and} \quad f_1(y)=0 \quad \text{otherwise;}$$

$$f_2(y)=f_*(x_2, y) \quad \text{if} \quad y\in\bar{A}, \quad \text{and} \quad f_2(y)=0 \quad \text{otherwise.}$$

For given $y\in R^m$ and λ, $0<\lambda<1$, we shall denote by $S(y; \lambda)$ the set of points (u, v) such that $u, v\in R^m$, $\lambda u+(1-\lambda)v=y$.

It can be shown that for all $y\in R^m$

$$\sup_{S(y;\lambda)} f_*^{\lambda}(x_1, u)f_*^{1-\lambda}(x_2, v) \geqq \sup_{S(y;\lambda)} f_1^{\lambda}(u)f_2^{1-\lambda}(v)$$

and for $y\bar{\in}\bar{A}\cap\bar{D}(x_0)$

$$\sup_{S(y;\lambda)} f_1^{\lambda}(u)f_2^{1-\lambda}(v)=0.$$

Since f_* is logarithmic concave in R^{n+m} (Corollary 1), the following inequality will be valid for all $y\in R^m$:

$$f_*(x_0, y) \geqq \sup_{S(\lambda;y)} f_*^{\lambda}(x_1, u)f_*^{1-\lambda}(x_2, v).$$

We shall prove that for all $y\in\text{int } D(x_0)$ we have

$$f_*(x_0, y) > \sup_{S(y;\lambda)} f_*^{\lambda}(x_1, u)f_*^{1-\lambda}(x_2, v). \tag{4}$$

Suppose on the contrary that there could be found a $y_0\in\text{int } D(x_0)$ such that

$$f_*(x_0, y_0) = \sup_{S(y_0;\lambda)} f_*^{\lambda}(x_1, u)f_*^{1-\lambda}(x_2, v).$$

In this case $f_*(x_0, y_0)>0$ as $(x_0, y_0)\in\text{ri } D$ (Lemma 4). According to Lemma 2 there exists a point $(u_0, v_0)\in S(y_0; \lambda)$ such that

$$u_0\in\bar{D}(x_1), \quad v_0\in\bar{D}(x_2) \quad \text{and} \quad f_*(x_0, y_0)=f_*^{\lambda}(x_1, u_0)f_*^{1-\lambda}(x_2, v_0).$$

We have got a contradiction to Lemma 3. So, for all $y\in\text{int } D(x_0)$ inequality (4) is valid.

From the definition of the function I and from Corollary 2 we get

$$I(x_0)=\int_{\bar{A}} f(x_0, y)\,dy=\int_{\bar{A}\cap\bar{D}(x_0)} f_*(x_0, y)\,dy.$$

Taking into account (4) and Theorem 3 of [1] we obtain:

$$\int_{\bar{A}\cap\bar{D}(x_0)} f_*(x_0, y)\,dy > \int_{\bar{A}\cap\bar{D}(x_0)} \sup_{S(y;\lambda)} f_*^{\lambda}(x_1, u) f_*^{1-\lambda}(x_2, v)\,dy \geqq$$

$$\geqq \int_{D(x_0)\cap\bar{A}} \sup_{S(y;\lambda)} f_1^{\lambda}(u) f_2^{1-\lambda}(v)\,dy = \int_{R^m} \sup_{S(y;\lambda)} f_1^{\lambda}(u) f_2^{1-\lambda}(v)\,dy \geqq$$

$$\geqq \Big[\int_{R^m} f_1(y)\,dy\Big]^{\lambda}\Big[\int_{R^m} f_2(y)\,dy\Big]^{1-\lambda} = \Big[\int_{\bar{A}\cap\bar{D}(x_1)} f_*(x_1, y)\,dy\Big]^{\lambda}\Big[\int_{\bar{A}\cap\bar{D}(x_2)} f_*(x_2, y)\,dy\Big]^{1-\lambda} =$$

$$= [I(x_1)]^{\lambda}[I(x_2)]^{1-\lambda}.$$

The theorem is proved.

Corollary 4. *Let* $x_1, x_2 \in R^n$ *and* $0<\lambda<1$. *If* $\lambda x_1+(1-\lambda)x_2 \in \mathrm{ri}\, B$, *then the inequality*

$$I\big(\lambda x_1+(1-\lambda)x_2\big) > [I(x_1)]^{\lambda}[I(x_2)]^{1-\lambda} \tag{5}$$

is valid.

Proof. It follows from Theorem 2 and Corollary 3.

In conclusion the author expresses his gratitude to G. G. Pestov for his help in carrying out the present work.

References

[1] A. Prékopa, On logarithmic concave measures and functions, *Acta Sci. Math.*, **34** (1973), 335—343.

[2] Р. Рокафеллар, *Выпуклый анализ*, Изд-во Мир (Москва, 1973).

CHAIR OF MATH. ANALYSIS
UNIVERSITY TOMSK
TOMSK 634010, USSR

Acta Sci. Math., 38 (1976), 423—428

Similarity invariants for a class of nilpotent operators

L. R. WILLIAMS

In this note, all Hilbert spaces will be understood to be *complex*. If $\mathfrak{H}$ is a Hilbert space, we denote by $\mathfrak{L}(\mathfrak{H})$ the algebra of all bounded linear operators on $\mathfrak{H}$. If A belongs to $\mathfrak{L}(\mathfrak{H})$ and there is a positive integer n such that $A^n=0$ and $A^{n-1}\neq 0$, then we say A is a *nilpotent* operator of order n. If n is a positive integer, then the nilpotent operator acting on the direct sum of n copies of $\mathfrak{H}$ and defined by the $n\times n$ matrix $[A_{ij}]$ $(i,j=1, ..., n)$, where

$$A_{i,i+1} = 1_{\mathfrak{H}} \quad \text{for} \quad i = 1, ..., n-1 \quad \text{and} \quad A_{i,j} = 0_{\mathfrak{H}} \quad \text{for all other entries,}$$

is called a *Jordan block* operator of order n. (By definition, $0_{\mathfrak{H}}$, the zero operator on $\mathfrak{H}$, is a Jordan block operator or order one.) Let m be a positive integer. Suppose $\mathfrak{H}_1, ..., \mathfrak{H}_m$ are Hilbert spaces and $n_1, ..., n_m$ are positive integers. Let $\mathfrak{H}_k^{\sim}$ be the direct sum of n_k copies of $\mathfrak{H}_k$ and J_k be the Jordan block operator of order n_k acting on $\mathfrak{H}_k^{\sim}$, $k=1, 2, ..., m$. An operator of the form $J_1\oplus ...\oplus J_m$ acting on $\mathfrak{H}_1^{\sim}\oplus ...\oplus\mathfrak{H}_m^{\sim}$ is called a *Jordan operator*.

Recall that if $\mathfrak{K}_1$ and $\mathfrak{K}_2$ are Hilbert spaces and X: $\mathfrak{K}_1\to\mathfrak{K}_2$ is a bounded linear transformation such that kernel $X=$kernel $X^*=\{0\}$, then X is called a *quasiaffinity*. If $A_1\in\mathfrak{L}(\mathfrak{K}_1)$, $A_2\in\mathfrak{L}(\mathfrak{K}_2)$, and there exists a quasiaffinity X: $\mathfrak{K}_1\to\mathfrak{K}_2$ such that $XA_1=A_2X$, then we say A_1 is a *quasiaffine transform* of A_2. If A_1 and A_2 are quasiaffine transforms of each other, i.e., if there exist quasiaffinities X: $\mathfrak{K}_1\to\mathfrak{K}_2$ and Y: $\mathfrak{K}_2\to\mathfrak{K}_1$ such that $XA_1=A_2X$ and $YA_2=A_1Y$, then A_1 and A_2 are said to be *quasisimilar*. Recall also that if there exists an invertible bounded linear transformation Z: $\mathfrak{K}_1\to\mathfrak{K}_2$ such that $ZA_1=A_2Z$, then A_1 and A_2 are said to be *similar*.

It is a well-known theorem of linear algebra that every nilpotent operator on a finite dimensional Hilbert space is similar to a Jordan operator. Since every Jordan operator clearly has closed range, one cannot expect this theorem to be true on an

Received November 5, 1975.

This note is a part of the author's Ph. D. thesis written at the University of Michigan under the direction of Professor C. Pearcy.

infinite dimensional Hilbert space, but APOSTOL, DOUGLAS, and FOIAŞ [1] recently proved that the following weakened version of the theorem is valid on any Hilbert space.

Theorem 1. *Every nilpotent operator on a Hilbert space of arbitrary dimension is quasisimilar to a Jordan operator.*

The purpose of this note is two-fold. In the first place, we present below a proof of Theorem 1 that is somewhat simpler than the argument in [1]. Secondly, essentially the same proof establishes the following result.

Theorem 2. *A nilpotent operator T on a Hilbert space is similar to a Jordan operator if and only if the range of T^k is closed, $k=1, 2, \ldots$.*

It will be convenient to use the following notation. If $\mathfrak{K}_1$ and $\mathfrak{K}_2$ are Hilbert spaces, A belongs to $\mathfrak{L}(\mathfrak{K}_1)$, and B: $\mathfrak{K}_2 \to \mathfrak{K}_1$ is a bounded linear transformation, then we let $M(A, B)$ denote the operator

$$\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}$$

in $\mathfrak{L}(\mathfrak{K}_1 \oplus \mathfrak{K}_2)$. If A: $\mathfrak{K}_1 \to \mathfrak{K}_2$ is a bounded linear transformation, then we denote by $\mathfrak{K}(A)$ the kernel of A and by $\mathfrak{R}(A)$ the range of A.

We begin with the following lemma.

Lemma 1. *Suppose J is a Jordan operator acting on a Hilbert space $\mathfrak{H}$, and suppose there are a Hilbert space $\mathfrak{K}$ and an isometry V: $\mathfrak{K} \to \mathfrak{H}$ such that $\mathfrak{R}(V) = \mathfrak{H} \ominus \mathfrak{R}(J)$. Then the operator $M(J, V)$ in $\mathfrak{L}(\mathfrak{H} \oplus \mathfrak{K})$ is unitarily equivalent to a Jordan operator.*

Proof. To say that J is a Jordan operator on $\mathfrak{H}$ means that there exist Hilbert spaces $\mathfrak{H}_1, \ldots, \mathfrak{H}_m$ and positive integers $n_1, \ldots, n_m$ such that if we let $\mathfrak{H}_k^{\sim}$ be the direct sum of n_k copies of $\mathfrak{H}_k$ and J_k be the Jordan block operator of order n_k on $\mathfrak{H}_k^{\sim}$ $(k= =1, 2, \ldots, m)$, then $\mathfrak{H} = \mathfrak{H}_1^{\sim} \oplus \ldots \oplus \mathfrak{H}_m^{\sim}$ and $J = J_1 \oplus \ldots \oplus J_m$. Let $\mathfrak{H}_k^{\wedge} = \mathfrak{H}_k^{\sim} \ominus \mathfrak{R}(J_k)$, i.e. $\mathfrak{H}_k^{\wedge} = 0 \oplus \ldots \oplus 0 \oplus \mathfrak{H}_k$ $(k=1, 2, \ldots, m)$. It is easy to verify that $\mathfrak{R}(V) = \mathfrak{H} \ominus \mathfrak{R}(J) = =\mathfrak{H}_1^{\wedge} \oplus \ldots \oplus \mathfrak{H}_m^{\wedge}$. Let U_k be the natural Hilbert space isomorphism of $\mathfrak{H}_k$ onto $\mathfrak{H}_k^{\wedge}$. Define W_k: $\mathfrak{H}_k \to \mathfrak{H}_k^{\sim}$ by setting $W_k x = U_k x$ for each x in $\mathfrak{H}_k$. Let $U = U_1 \oplus \ldots \oplus U_m$ and $W = W_1 \oplus \ldots \oplus W_m$. Define V_0: $\mathfrak{K} \to \mathfrak{R}(V)$ by setting $V_0 x = Vx$ for each $x \in \mathfrak{K}$. The linear transformations U: $\mathfrak{H}_1 \oplus \ldots \oplus \mathfrak{H}_m \to \mathfrak{R}(V)$ and V_0: $\mathfrak{K} \to \mathfrak{R}(V)$ are unitary. Hence the linear transformation

$$1_{\mathfrak{H}} \oplus U^* V_0 : \mathfrak{H} \oplus \mathfrak{K} \to \mathfrak{H} \oplus (\mathfrak{H}_1 \oplus \ldots \oplus \mathfrak{H}_m)$$

is unitary and

$$(1_{\mathfrak{H}} \oplus U^* V_0) M(J, V) (1_{\mathfrak{H}} \oplus U^* V_0)^* = M(J, V V_0^* U) = M(J, W).$$

Furthermore, the operator

$$M(J,W) \quad \text{in} \quad \mathfrak{L}(\mathfrak{H}\oplus(\mathfrak{H}_1\oplus\ldots\oplus\mathfrak{H}_m))$$

is unitarily equivalent to the operator

$$M(J_1,W_1)\oplus\ldots\oplus M(J_m,W_m) \quad \text{in} \quad \mathfrak{L}((\tilde{\mathfrak{H}}_1\oplus\mathfrak{H}_1)\oplus\ldots\oplus(\tilde{\mathfrak{H}}_m\oplus\mathfrak{H}_m)).$$

Thus, in order to complete the proof, it suffices to show that the operators $M(J, W_k)$ $(k=1, 2, \ldots, m)$ are Jordan block operators. In order to do this, we observe that $W_k\colon \mathfrak{H}_k\to\tilde{\mathfrak{H}}_k=\mathfrak{H}_k\oplus\ldots\oplus\mathfrak{H}_k$ is defined by the $n_k\times 1$ matrix all of whose entries are $0_{\mathfrak{H}_k}$ except the last, which is $1_{\mathfrak{H}_k}$. Hence it is clear that the operator $M(J_k, W_k)$ is the Jordan block operator of order n_k+1 on the direct sum of n_k+1 copies of $\mathfrak{H}_k$. Thus the proof is complete.

Lemma 2. *Suppose T is a nilpotent operator of order $n>1$ on a Hilbert space $\mathfrak{H}$. Then there exist Hilbert spaces $\mathfrak{K}_1$ and $\mathfrak{K}_2$, a nilpotent operator A of order $n-1$ in $\mathfrak{L}(\mathfrak{K}_1)$, and a bounded linear transformation $B\colon \mathfrak{K}_2\to\mathfrak{K}_1$ such that T is unitarily equivalent to the operator $M(A, B)$ in $\mathfrak{L}(\mathfrak{K}_1\oplus\mathfrak{K}_2)$ and such that $(\mathfrak{R}(A)+\mathfrak{R}(B))^-=\mathfrak{K}_1$. Furthermore, if each $\mathfrak{R}(T^k)$ is closed $(k=1,2,\ldots)$, then $\mathfrak{R}(A^k)$ is closed $(k=1,2,\ldots)$, and in this case $\mathfrak{R}(A)+\mathfrak{R}(B)=\mathfrak{K}_1$.*

Proof. Let $\mathfrak{K}_1=\mathfrak{R}(T)^-$ and $\mathfrak{K}_2=\mathfrak{H}\ominus\mathfrak{R}(T)^-$. The operator T is clearly unitarily equivalent to some operator $M(A, B)$ in $\mathfrak{L}(\mathfrak{K}_1\oplus\mathfrak{K}_2)$ where $\mathfrak{R}(A)+\mathfrak{R}(B)=\mathfrak{R}(T)$. Hence we have $(\mathfrak{R}(A)+\mathfrak{R}(B))^-=\mathfrak{K}_1$. An elementary calculation shows that A is a nilpotent operator of order $n-1$. If $\mathfrak{R}(T^k)$ is closed, $k=1, 2, \ldots$, then it is clear that $\mathfrak{R}(A)+\mathfrak{R}(B)=\mathfrak{K}_1$ and it follows easily that $\mathfrak{R}([M(A, B)]^k)=\mathfrak{R}(A^{k-1})\oplus 0$ $(k=1, 2, \ldots)$. Hence $\mathfrak{R}(A^k)$ is closed, $k=1, 2, \ldots$, and the proof is complete.

Lemma 3. *Suppose T is a nilpotent operator on a Hilbert space $\mathfrak{H}$* [and $\mathfrak{R}(T^k)$ is closed $(k=1, 2, \ldots)$]. *Then T is a quasiaffine transform of* [similar to] *a Jordan operator.*

Proof. We prove the lemma by induction on the order n of T. If $n=1$, then T is the zero operator on $\mathfrak{H}$ and hence, by definition, T is a Jordan operator. So we assume $n>1$ and that the lemma is true for all nilpotent operators of order $n-1$. According to Lemma 2, T is unitarily equivalent to an operator $M(A, B)$ in $\mathfrak{L}(\mathfrak{K}_1\oplus\mathfrak{K}_2)$ for some Hilbert spaces $\mathfrak{K}_1$ and $\mathfrak{K}_2$, where A is a nilpotent operator of order $n-1$ and $(\mathfrak{R}(A)+\mathfrak{R}(B))^-=\mathfrak{K}_1$ [$\mathfrak{R}(A)+\mathfrak{R}(B)=\mathfrak{K}_1$ and each $\mathfrak{R}(A^k)$ is closed]. Thus, by the induction hypothesis, there exist a Jordan operator J on a Hilbert space $\mathfrak{H}_0$ and a quasiaffinity [an invertible bounded linear transformation] $X\colon \mathfrak{K}_1\to\mathfrak{K}_0$ such that $XA=JX$. The bounded linear transformation $X\oplus 1_{\mathscr{K}_2}\colon \mathfrak{K}_1\oplus\mathfrak{K}_2\to\mathfrak{H}_0\oplus\mathfrak{K}_2$ is a quasiaffinity [is invertible] and $(X\oplus 1_{\mathscr{K}_2})M(A, B)=M(J, C)(X\oplus 1_{\mathscr{K}_2})$ where $C=XB\colon \mathfrak{K}_2\to\mathfrak{H}_0$. It is easy to verify that $(\mathfrak{R}(J)+\mathfrak{R}(C))^-=\mathfrak{H}_0$ [$\mathfrak{R}(J)+\mathfrak{R}(C)=\mathfrak{H}_0$].

We observe that $\mathfrak{R}(J)$ is closed since J is a Jordan operator. Let E be the orthogonal projection onto $\mathfrak{R}(J)$. Then, of course, $\mathfrak{R}(EC)\subset\mathfrak{R}(J)$. It follows from a theorem of R. G. DOUGLAS ([2], Theorem 1) that there exists a bounded linear transformation $Y\colon \mathfrak{K}_2\to\mathfrak{H}_0$ such that $EC=JY$. The operator

$$\begin{pmatrix} 1_{\mathfrak{H}_0} & Y \\ 0 & 1_{\mathfrak{K}_2} \end{pmatrix}$$

in $\mathfrak{L}(\mathfrak{H}_0\oplus\mathfrak{K}_2)$ is invertible and

$$\begin{pmatrix} 1_{\mathfrak{H}_0} & Y \\ 0 & 1_{\mathfrak{K}_2} \end{pmatrix}\begin{pmatrix} J & C \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} J & D \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1_{\mathfrak{H}_0} & Y \\ 0 & 1_{\mathfrak{K}_2} \end{pmatrix}$$

where $D=-JY+C=-EC+C=(1_{\mathfrak{H}_0}-E)C$. A straight forward calculation shows that $\mathfrak{R}(D)^-=\mathfrak{H}_0\ominus\mathfrak{R}(J)$ $[R(D)=\mathfrak{H}_0\ominus\mathfrak{R}(J)]$.

Let $\mathfrak{K}_3=\mathfrak{K}_2\ominus\mathfrak{K}(D)$, $\mathfrak{K}_4=\mathfrak{K}(D)$, and let $D_0\colon\mathfrak{K}_3\to\mathfrak{H}_0$ be defined by $D_0x=Dx$ for each x in $\mathfrak{K}_3$. The operator $M(J, D)$ in $\mathfrak{L}(\mathfrak{H}_0\oplus\mathfrak{K}_2)$ is unitarily equivalent to the operator

$$\begin{pmatrix} J & D_0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

in $\mathfrak{L}(\mathfrak{H}_0\oplus\mathfrak{K}_3\oplus\mathfrak{K}_4)$. So in order to complete the proof of the lemma, it suffices to show that the operator $M(J, D_0)$ in $\mathfrak{L}(\mathfrak{H}_0\oplus\mathfrak{K}_3)$ is a quasiaffine transform of [similar to] a Jordan operator. We observe that $\mathfrak{K}(D_0)=(0)$ and $\mathfrak{R}(D_0)^-=\mathfrak{H}_0\ominus\mathfrak{R}(J)$ $[\mathfrak{R}(D_0)=\mathfrak{H}_0\ominus\mathfrak{R}(J)]$. Write $D_0=VP$, the polar decomposition of D_0. It follows that $V\colon \mathfrak{K}_3\to\mathfrak{H}_0$ is an isometry and $\mathfrak{R}(V)=\mathfrak{R}(D_0)^-=\mathfrak{H}_0\ominus\mathfrak{R}(J)$. The operator P in $\mathfrak{L}(\mathfrak{K}_3)$ is a quasiaffinity [an invertible operator] since P is positive and $\mathfrak{K}(P)=(0)$ [and $\mathfrak{R}(P)$ is closed]. Hence the operator $1_{\mathfrak{H}_0}\oplus P$ in $\mathfrak{L}(\mathfrak{H}_0\oplus\mathfrak{K}_3)$ is a quasiaffinity [an invertible operator] and $(1_{\mathfrak{H}_0}\oplus P)M(J, D_0)=M(J, V)(1_{\mathfrak{H}_0}\oplus P)$. The operator J and the linear transformation V satisfy the hypotheses of Lemma 1. Thus the operator $M(J, V)$ is unitarily equivalent to a Jordan operator, and hence the proof is complete.

Corollary 1. *Every nilpotent operator on a Hilbert space is quasisimilar to its adjoint.*

Proof. Suppose T is a nilpotent operator. By Lemma 3, there exist a quasiaffinity X and a Jordan operator J such that $XT=JX$. Then $T^*X^*=X^*J^*$. Since every Jordan operator is unitarily equivalent to its adjoint, we have $UJ=J^*U$ where U is a unitary operator. Combining these equations, we get $(X^*UX)T=T^*(X^*UX)$. Hence T is a quasiaffine transform of T^*. The same argument applied to T^* shows that T^* is a quasiaffine transform of T. Hence T and T^* are quasisimilar.

Corollary 2. *If T is a nilpotent operator on a Hilbert space and each $\mathfrak{R}(T^k)$ is closed* $(k=1, 2, \ldots)$, *then T is similar to its adjoint.*

Proof. By Lemma 3, there exist an invertible bounded linear transformation X and a Jordan operator J such that $XT=JX$. Now proceed as in the proof of Corollary 1 to obtain the equation $(X^*UX)T=T^*(X^*UX)$ where U is a unitary operator. Hence T and T^* are similar.

Proof of Theorem 1. Suppose T is a nilpotent operator on a Hilbert space. Then T^* is also a nilpotent operator. Thus, according to Lemma 3, there exist quasiaffinities X and Y and Jordan operators J_1 and J_2 such that $XT=J_1X$ and $YT^*=J_2Y$. Then $T^*X^*=X^*J_1^*$ and $TY^*=Y^*J_2^*$. Since J_1 and J_2 are Jordan operators, we have $UJ_1=J_1^*U$ and $VJ_2=J_2^*V$ where U and V are unitary operators. Combining these equations, we get $T(Y^*VYX^*U)=(Y^*VYX^*U)J_1$. Hence T and J_1 are quasisimilar.

Proof of Theorem 2. Let T be a nilpotent operator on a Hilbert space. If T is similar to a Jordan operator J, then T^k is similar to J^k for each positive integer k. It is clear that $\mathfrak{R}(J^k)$ is closed $(k=1, 2, \ldots)$. Hence $\mathfrak{R}(T^k)$ is closed, $k=1, 2, \ldots$. On the other hand if $\mathfrak{R}(T^k)$ is closed $(k=1, 2, \ldots)$, then we can conclude from Lemma 3 that T is similar to a Jordan operator.

Foiaş and Pearcy [3] proved that every nilpotent operator acting on a separable Hilbert space is quasisimilar to a compact operator. Below we give a different proof of this theorem based on the following lemma.

Lemma 4. *If T is a nilpotent operator on a separable Hilbert space $\mathfrak{H}$, then there exist a compact quasiaffinity Z and a compact operator K in $\mathfrak{L}(\mathfrak{H})$ such that $ZT=KZ$.*

Proof. We prove the lemma by induction on the order n of T. If $n=1$, then T is the zero operator on $\mathfrak{H}$ and the result is obvious. So we assume $n>1$ and that the lemma is true for all nilpotent operators of order $n-1$ acting on a separable Hilbert space. According to Lemma 2, the operator T is unitarily equivalent to an operator $M(A, B)$ in $\mathfrak{L}(\mathfrak{K}_1\oplus\mathfrak{K}_2)$ for some separable Hilbert spaces $\mathfrak{K}_1$ and $\mathfrak{K}_2$, where A is a nilpotent operator of order $n-1$ in $\mathfrak{L}(\mathfrak{K}_1)$. Thus by the induction hypothesis, there exist a compact quasiaffinity Z_0 and a compact operator K_0 in $\mathfrak{L}(\mathfrak{K}_1)$ such that $Z_0A=K_0Z_0$. Write $Z_0B=UP$, the polar decomposition of Z_0B. The operator P in $\mathfrak{L}(\mathfrak{K}_2)$ is positive and compact. Hence $P^{1/2}$ is compact. Let $\tilde{P}$ be any compact quasiaffinity in $\mathfrak{L}(\mathfrak{K}(P^{1/2}))$. We define a compact quasiaffinity P_0 on $\mathfrak{K}_2$ by setting $P_0x=\tilde{P}x$ for each x in $\mathfrak{K}(P^{1/2})$ and $P_0x=P^{1/2}x$ for each x in $\mathfrak{K}_2\ominus\mathfrak{K}(P^{1/2})$. Clearly $P=P^{1/2}P_0$. The operator $Z_0\oplus P_0$ is a compact quasiaffinity and the operator $M(K_0, UP^{1/2})$ is compact. An easy calculation shows that $(Z_0\oplus P_0)M(A, B)=M(K_0, UP^{1/2})(Z_0\oplus P_0)$, and hence the proof is complete.

Theorem 3. *Every nilpotent operator on a separable Hilbert space is quasisimilar to a compact operator.*

Proof. Suppose T is a nilpotent operator on a separable Hilbert space. According to Lemma 4, there exist a (compact) quasiaffinity Z and a compact operator K such that $ZT=KZ$. Then $T^*Z^*=Z^*K^*$. The operator K is necessarily nilpotent. Thus, by applying Corollary 1 to T and K, we can obtain quasiaffinities X and Y such that $TX=XT^*$ and $YK=K^*Y$. Combining these equations, we get $T(XZ^*Y)=$ $=(XZ^*Y)K$. Hence T and K are quasisimilar.

Bibliography

[1] C. Apostol, R. G. Douglas, and C. Foiaş, Quasisimilar models for nilpotent operators, *to appear*.

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

[3] C. Foiaş and C. Pearcy, A model for quasinilpotent operators, *Michigan Math. J.*, **21** (1974), 399—404.

DEPARTMENT OF MATH.
UNIVERSITY OF MICHIGAN
ANN ARBOR, MICH. 48104, USA

Bibliographie

M. Aigner, Kombinatorik I. Grundlagen und Zähltheorie (Hochschultext), XVII+409 pages. Springer-Verlag, Berlin—Heidelberg—New York, 1975.

The foundations of Combinatorics have developed very rapidly in the past years. A few decades ago combinatorics meant a collection of various enumeration problems, and there existed (as a separate discipline) several graph theoretical, statistical, geometrical results, problems and puzzles of combinatorial nature. We are witnessing the arousal of new notions, methods and theories of large unifying und theorem-proving power. Such are matroid theory (combinatorial geometries), the functional analysis treatment of generating functions, the theory of Moebius functions, categorial and lattice theoretical methods — just to mention those treated in the first volume of this nice book. In the light of these theories the enumerative and the "structural" parts of combinatorics turn out to be much closer related than thought before.

This book reflects these new changes. Although its subtitle is "Foundations and Enumeration", it treats parts of combinatorics which are of "structural" nature but play an important role in the enumerative theory as well (e.g. lattice theory or matroids). It is a first, and successful, attempt to present modern combinatorics and its relations to modern mathematics (algebra, functional analysis, category theory) in a textbook form. It goes into the material in a considerable depth (treating e.g. the Pólya Method), and remains easily readable and elegant. There are about 375 exercises, some of which contain further theoretical material.

It is the significance and novelty of this presentation that makes some criticism in order here. One, if not the most important, goal in deriving (sometimes rather complicated-looking) formulas and generating functions is to obtain asymptotic results. Pólya's famous paper, for example, carries through such a program: it derives generating functions and then, by the methods of function theory, obtains asymptotical formulas. The development in the methods for the first part of such an investigation has caused a tendency of forgetting the second, and I miss a mention of this in this book too.

L. Lovász (Szeged)

E. M. Alfsen, Compact Convex Sets and Boundary Integrals (Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 56), IX+210 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1971.

In the preface of his book the author says that "the integral representation theorems of Choquet and Bishop—de Leeuw together with the uniqueness theorem of Choquet inaugurated a new epoch in infinite-dimensional convexity". Although it has long been clear that convexity arguments are very fruitful in functional analysis, only with the advent of Choquet's theory a couple of decades ago did a comprehensive theory of infinite dimensional convex sets begin to exist. Now the original proofs of the basic results, initially considered technically difficult, are very much simplified. "Choquet

Theory provides a unified approach to integral representations in fields as diverse as potential theory, probability, function algebras, operator theory, group representations and ergodic theory." The book under review is an up to date introduction to Choquet Theory. It can be used as a text book for graduate students as well as a reference book for the working mathematician. It also tries to stimulate further study of the finer structure of infinite dimensional compact convex sets.

The book consists of two chapters. Chapter I: "Representations of Points by Boundary Measures". The paragraphs are: Distinguished Classes of Functions on a Compact Convex Set; Weak Integrals, Moments and Barycenters; Comparison of Measures on a Compact Convex Set; Choquet's Theorem; Abstract Boundaries Defined by Cones of Functions; Unilateral Representation Theorems with Application to Simplicial Boundary Measures. Chapter II: "Structure of Compact Convex Sets". The paragraphs in this chapter are: Order-unit and Base-norm Spaces; Elementary Embedding Theorems; Choquet Simplexes; Bauer Simplexes and the Dirichlet Problem of the Extreme Boundary; Order Ideals, Faces, and Parts; Split-faces and Facial Topology; The Concept of Center for $A(K)$; Existence and Uniqueness of Maximal Central Measures Representing Points of an Arbitrary Compact Convex Set.

As prerequisite, only some basic knowledge of functional analysis and integration theory is assumed on the part of the reader.

József Szűcs (Szeged)

R. Alletsee, G. Umhauer, Assembler I, II, III, Springer-Verlag, Berlin—Heidelberg—New York, 1974. 126, 150, 170 pages.

The books are useful for teaching or learning the IBM Assembly programming language. The student has to have only a limited preliminary knowledge about computer's hardware. Decimal, binary, floating point arithmetical, logical and branching machine instructions, furthermore the data and storage definition statements are treated. The Assembler instructions and the logical input/output macro instructions are not fully described. When finishing the course the student can write programs of one segment and one section with simple input/output activity. Numerous examples and excercises help to understand the notions and language elements. Test controls in the paragraphs qualify the books for using in assembler courses as a teacher's manual.

Árpád Makay (Szeged)

William Arveson, An Invitation to C^*-Algebras (Graduate Texts in Mathematics, Vol. 39), X+106 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1976.

This excellent book conveys to the reader the fundamentals of the representation theory of separable postliminal C^*-algebras, which are called by the author, after Kaplansky, GCR (generalized completely continuous representation) algebras. A GCR algebra is a C^*-algebra A having the following property: for every (two sided and closed) ideal J of A the quotient C^*-algebra A/J contains a non-zero C^*-algebra B such that the range of every irreducible $*$-representation of B on a Hilbert space consists of compact operators. It is known and proved in the book that the spectrum $\hat{A}$ of a separable GCR algebra A bears a standard Borel structure which makes it possible to uniquely decompose every separable, nondegenerate $*$-representation π of A as a direct integral of "orthogonal copies" of irreducible representations: $\pi \cong \int_{\hat{A}}^{\oplus} m(\xi)\xi \, d\mu(\xi)$, where μ is a finite positive Borel measure on $\hat{A}$ and m is an integral (possibly infinite) valued non-negative measurable function on $\hat{A}$ ($m(\xi)$ is the multiplicity of ξ in π and the decomposition is unique up to the equivalence class of μ). The complete proof of this last assertion is the main achievement of the book. It might seem so that the GCR property of A is a very strict stipulation. However, it is mentioned in the preface and text proper that "to this day no one has given a concrete parametric description of even the irreducible representations of any C^*-algebra which is not GCR" and "there is mathematical evidence which strongly suggests that no one ever will'. Thus, in spite of its specialization, the book is complete in this respect.

If the idea of a proof is clear in a special case, then the generalization is relegated to the exercises. There are four chapters. Chapter 1 contains the rudiments of the theory of C^*-algebras. The second chapter deals with multiplicity theory, typé I von Neumann algebras, and type I repr, sentations of C^*-algebras. It gives the multiplicity theory of normal operations of C^*-algebras. It gives that all representations of a GCR algebra are type I. Chapter 3 is a nice introduction to polish spaces, standard and analytic Borel structures and cross sections. Chapter 4 uses the results of the preceding chapter to prove the decomposition theorem for representations of (separable) GCR algebras. It also contains a section on elementary reduction theory, just enough to prove the decomposition theorem. There is a bibliography and index.

The text tries to serve a large variety of readers: different subject matters are treated as independently as possible. Only the knowledge of the basic results of functional analysis, measure theory, and Hilbert space are assumed.

József Szűcs (Szeged)

Alan Baker, Transcendental Number Theory, X+147 pages, Cambridge University Press, 1975.

The book under review provides "a comprehensive account of the recent major discoveries" in the theory of transcendental numbers. At the beginning the author discusses the historical aspects of the theory and gives a survey of the subject as it existed around the turn of the century. The text includes among others the latest theories relating to linear forms in the logarithms of algebraic numbers, Schmidt's generalization of the Thue-Siegel-Roth theorem, Shidlovsky's work on Siegel's E-functions and Sprındžuk's solution to the Mahler conjecture. As proofs in the subject are usually long and intricate, the author felt necessary to select for detailed treatment only those that led to fundamental results and wide application.

"The test has arisen from lectures delivered in Cambridge, America and elsewhere, and it has also formed the substance of an Adams Prize essay."

József Szűcs (Szeged)

Raymond Balbes—Philip Dwinger, Distributive Lattices, XIII+294 pages, Columbia, Missouri, University of Missouri Press, 1974.

The theory of distributive lattices is one of the oldest branches of lattice theory. The connections of distributive lattices and other fields of mathematics, especially topology, algebra and logic are the sources of a number of deep and important results. However, for a long time the theory consisted of separate topics; the general methods to handle distributive lattices originated from universal algebra and category theory, and have been developed only in the last two decades. The authors of this book are among the eminent specialists in those researches leading to this development. Their book under review presents the theory of distributive lattices in the framework of a homogeneous theory based on topology, univeral algebra and category theory. The book is excellent and up-to-date.

From the Preface: "In Chapter I all those elements of univeral algebra and category theory which the reader will need — and in addition, some notions of set theory — are presented... The fundamental theory of distributive lattices is developed in Chapters II—VII. Some highlights in these chapters are the prime ideal theory, the representation theory, free algebras, coproducts and extension theorems... The special classes of distributive lattices which are discussed in this book are pseudocomplemented distributive lattices (Chapter VIII), Heyting algebras (Chapter IX), Post algebras (Chapter X), de Morgan algebras and Lukasiewicz algebras (Chapter XI). Finally Chapter XII is entirely devoted to complete and α-complete distributive lattices, which may satisfy a higher degree of distributivity."

There are numerous exercises scattered throughout the book. The book is addressed to graduate students and to those mathematicians who work in the field or want to become acquainted with it.

We may add that this book is useful and enjoyable for anybody who studies lattice theory or is interested in the applications of universal algebra and category theory.

A. P. Huhn (Szeged)

Anatole Beck, Continuous Flows in the Plane (Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Band 201) X+462 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1974.

The study of continuous flows is an idealization of dynamical systems such as aerodynamics, hydrodynamics, electrodynamics etc. We imagine in the plane some sort of idealized particles which change position as time passes and after a time t, the particle which was at x will be transposed to the position $\varphi(t, x)$. After the definition of the moving points, fixed points, endpoints, stagnation points, regular and singular points, spirals, etc., the author gives a very geometrical description of the orbits. One of the basic results, the Gate Theorem, which simplifies the analysis of the orbits of any flow in the plane, is a generalization, in a sense, of the Jordan Curve Theorem.

If a flow φ is related to the flow ψ in such a way that every ψ-orbit is contained in a φ-orbit, we call ψ a reparametrization of φ. An important category of reparametrizations is the reparametrization by flow multiplers. In several chapters the author describes the important properties of these reparametrizations: canonical reparametrization, time measure of a quasi-reparametrization, algebraic combinations, etc. Every flow in the plane can be considered as a flow in the sphere which has ∞ as fixed point, every continuous flow in the sphere has at least one fixed point, thus the theories of flows in the plane and in the sphere are equivalent. In the chapters 6 and 7 the author concentrate on the problems: Given a flow φ on the boundary of a region, when does a continuous extension of this flow onto the given region exist? Let F be a compact subset of the sphere, and Y a subset of F. When does a continuous flow exist with fixpoints F and with stagnation points Y?

Let A and B be regions on the plane and φ a flow on A. Then for every homeomorphism f from A onto B this homeomorphism defines a flow $f\varphi$ on B. If $f\varphi$ is reparametrization of a flow ψ on B by a flow multiplier, then we say that φ and ψ are conjugate. It is examined in the last part of the book, when are the flows homeomorphic and when are they conjugate. The basic result of these analyses are the homeomorphism with an annular flow of standard type, the Theory of Kaplan and Markus, and the examination of the Kaplan diagramm.

The book only assumes a level of preparation equaivalent to first-year graduate courses, and it does not require any special knowledge of topology or differential equations. The work intended to serve as an introduction to the field of dynamics, particularly to readers with analytic training.

Z. I. Szabó (Szeged)

Norman Biggs, Finite Groups of Automorphisms (London Mathematical Society Lecture Notes Series 6), 117 pages, Cambridge University Press, 1971.

Since the beginnings of group theory, many important finite groups (especially, many simple ones) have been defined as automorphism groups of certain combinatorial structures. This book leads the reader through the main ideas of the development of this interrelation, starting with Galois and concluding with the quite recent discovery of new sporadic simple groups.

Chapter 1 is a brief introduction to permutation group theory.

Chapter 2 is devoted to the finite spaces and the finite linear groups. The simplicity of the projective linear groups and their relationship to projective geometries is shown. The symplectic, orthogonal and unitary groups are also introduced.

Chapter 3 introduces the t—(v, k, λ) designs. For symmetric designs (when the numbers of points and blocks are equal), the Bruck-Ryser-Chowla theorem is derived. Then, transitive extensions of permutation groups and extensions of designs are studied. The Mathieu groups and the corresponding designs are introduced this way (following Witt's treatment).

Chapter 4 is concerned with automorphism groups of distance transitive graphs. (A graph $G=(V, E)$ is *distance-transitive* if, given $x_1, \ldots, x_4 \in V$ such that the distances $d(x_1, x_2)$ and $d(x_3, x_4)$ are equal, there is an automorphism $\alpha \in$ Aut G such that $\alpha x_1 = x_3$. $\alpha x_2 = x_4$. The "intersection matrix" contains the information on the numerical regularity properties of such a graph. A beautiful theory, provoding very restrictive necessary conditions on the existence of distance-transitive graphs with given intersection matrix in terms of eigenvectors of this matrix is developed. In the case when these conditions are fulfilled, the matrix is said to be *feasible*. The feasibility in the case of diameter 2 and the absence of triangles is studied in detail. Then, the problem of realizability of feasible matrices with small parameters is investigated. Finally, as a coronation of the material presented, a distance-transitive, triangle free graph of degree 22 with any two non-adjacent vertices having 6 common neighbors is constructed, hence the celebrated rank 3 simple group of Higman and Sims.

As an Appendix, a list of parameters of new sporadic simple groups and another list of the feasibility and of the status of realizability of intersection matrices of distance transitive graphs of diameter 2 and degree $\leqq 16$ is added. The literature mentions 10 books and 13 papers.

The book requires introductory linear algebra and group theory courses only. The selection of material as well as its presentation are excellent. It should be a pleasure for mathematicians interested in *combinatorics, linear algebra* and *group theory* to read the book, and to base (advanced) courses on it (as did the reviewer).

L. Babai (Budapest),

Norman Biggs, Algebraic Graph Theory (Cambridge Tracts in Mathematics, 67), vii+170 pages, Cambridge University Press, 1974.

The term "algebraic" in the title refers to classical algebraic techniques (determinants, matrices, polynomials, groups). The book exhibits some important areas of graph theory where applications of such techniques have proved particularly fruitful. Classical results of Kirchhoff, Cayley, Whitney as well as the striking development of the last few decades are represented in a unified treatment.

In Part I ("Linear algebra and graph theory"), the basic concepts are introduced (incidence and adjacency matrices, characteristic polynomial, spectrum of a graph Γ). The circuit- and cutset-spaces (the homology of Γ) and the complexity (the number of spanning trees) are discussed. Various expansions of determinants, related to Γ, in terms of certain subgraphs, conclude Part I.

Part II ("Colouring problems") starts with inequalities, bounding the chromatic number in terms of the spectrum of Γ. Among others a highly non-trivial lower bound, due to A. J. Hoffman, is derived.

The rest of Part II is devoted to the study of the *chromatic polynomial* of Γ. For u a positive integer this is the number of colorings of the vertices of Γ by colors chosen from the set $\{1, \ldots, u\}$

such that adjacent vertices have different colors, which turns out to be a polynomial in u. Several expansions in terms of various families of subgraphs are derived. The useful "logarithmic transformation" is introduced and applied to obtain a multiplicative expansion, depending on a restricted family of subgraphs. The deepest result of Part II is Tutte's identity, relating the *Tutte-polynomial* of Γ (defined in terms of certain spanning trees) to the *rank polynomial* (defined in terms of ranks and co-ranks of subgraphs). This is then applied to obtain another expansion of the chromatic polynomial, in terms of these trees.

The central concept investigated in Part III ("Symmetry and regularity of graphs") is that of *automorphisms* of Γ. Γ is *t-transitive* ($t \geqq 1$) if for any two paths of length t, and any directions given on them, there is an automorphism α of Γ mapping one onto an other. An elegant proof of Tutte's deep theorem is given, stating that *if Γ is a trivalent t-transitive graph, then* $t \leqq 5$. A 5-transitive trivalent graph is also exhibited. By a *covering graph* construction, infinitely many such graphs are obtained from a single one.

Next, *distance-transitive* graphs are introduced (see the above review on Biggs' "Finite Groups of Automorphisms"). Γ is called *distance-regular* if for any two vertices u and v, the number s_{hij} of vertices w having distance h from u and distance i from v depends only on the distance j between u and v. A distance-transitive graph is clearly distance-regular. Powerful matrix techniques are developed to handle distance-regularity. Part III ends with the beautiful theory of (k, g)-graphs, also known as Moore-graphs or cages (these are graphs of degree k and girth g, whose cardinality attains a certain trivial lower bound on the number of vertices). The main result, obtained by investigation of the multiplicities of eigenvalues of the adjacency matrix, is the following: *for* $k, g \geqq 3$, *a (k, g)-graph exists only if either* $g \in \{3, 4, 6, 8, 12\}$, *or* $g=5$ *and* $k \in \{3, 7, 57\}$.

The bibliography contains 80 items.

A great deal of material is included in the form of well-chosen examples and results at the end of each of the 23 chapters.

The most valuable feature of the book is the *concise, clear, exceptionally aesthetic presentation* of a really *exciting material*, almost no part of which has yet appeared in book form. Most proofs represent essential simplifications of the original ones.

The reader is assumed to have a moderate knowledge of matrix theory and the basic concepts of graph and group theory only. It appeals to mathematicians in any field, and probably it will soon become one of the fundamental works. Everyone interested in graph theory, combinatorics and applications of matrix techniques should read the book.

L. Babai and *P. Komjáth* (Budapest)

Kai Lai Chung, Elementary Probability Theory with Stochastic Processes (Undergraduate Texts in Mathematics), X+325 pages, New York—Heidelberg—Berlin, Springer-Verlag, 1974.

This is the first volume of a new series and if the continuation will be so good as the beginning then this series will again be a new Springer-Verlag success. It is intended to be a very elementary introduction written by one of the outstanding experts of the field. A good deal of it does not even preassume calculus, but by brilliant organization, the author has succeeded in covering a wide range of topics, giving a real insight into the subject and preparing the reader for more advanced books. There are eight chapters: Set; Probability; Counting; Random variables; Conditioning and independence; Mean, variance and transformation; Poisson and normal distributions; From random walk to Markov chains; and three brief appendices: Borel fields and general random variables; Stirling's formula and DeMoivre-Laplace's theorem; Martingale. The body of each chapter also contains stimulating examples and at the end of each there are interesting classical and new problems for which solutions are also given at the end of the book. The emphasis is always on essential probabilistic reasoning, the style is inviting and at places humorous and all this is kept in good balance by the special intellectual power of the author. It can also stand up as a fine belletristic composition. Indeed, it is a book of great individuality.

S. Csörgő (Szeged)

N. S. M. Coxeter, Regular Complex Polytopes, X+185 pages, Cambridge University Press, Cambridge, 1974.

The very attentively constructed work gives a step by step introduction to the theme, beginning with plane and solid kinematics, through the geometrial description of the sixteen regular polytopes in four dimensional real Euclidean space and of finite multiplicative quaternion groups thereafter. (Chapters 1—7.) Meanwhile several devices and ideas which play central roles in the main Chapters are presented, such as free patterns, Cayley diagrams, the extended Schlőfli symbol, flags, Petrie polygons, Schwarz triangles, binary polyhedral groups, finite multiplicative quaternion groups etc.

In order to review the main sections of the book, let the corresponding part from the Preface be quoted: "The complete list of finite reflection groups in unitary n-space was complied in 1957 by Shephard and Todd, who found that there are many more of them in the plane than in any higher space. Chapter 10 checks their results (in the two dimensional case) by a new method: examining all the finite groups of unitary transformations and picking out those that are generated by reflections. In particular those that are generated by two reflections are the symmetry groups of the regular complex polygons. These are enumerated in Chapter 11. Somewhat surprisingly, it is possible to make real drawings of these imaginary figures, and in many cases such a drawing of one complex polygon serves as a *Cayley diagram* for the symmetry group of another. Chapters 12 and 13 deal with regular polytopes and honeycombs, using definitions suggested by Peter McMullen. There are interesting connections with certain projective configurations such as the 27 lines on the cubic surface. A remarkable presentation is found for the simple group of order 25920."

This book is an interesting and delectable reading both for research mathematicians and for students familiar with the material of the standard courses of elementary geometry and algebra. Most of the sections end with exercises; the solutions can be found at the end of the book. The beautiful presentation and the numerous figures also deserve special attention.

L. Stachó (Szeged)

Claude Dellacherie, Capacités et processus stochastiques (Ergebnisse der Mathematik und ihrer Grenzgebiete, 67) IX+155 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1972.

It is not an unfrequent opinion among mathematicians that the primary objects of probability theory are the distributions, and the sample space with its σ-fields constitutes only the necessary technical background. To avoid cumbrous measurability proofs some specialists prefer assuming sufficiently rich σ-fields to be given.

The author of the present book, a prominent member of the Strasbourg workshop of probability, does not share this opinion. On the contrary he shows that measurability properties of random processes with respect to some adequately defined σ-fields illuminate essential features of the processes and have deep connection with their sample path properties. The elegant general theory of stochastic processes elaborated in the book presents many classical questions (e.g. martingale decompositions) from a new unified view-pont. It can serve as a basis for a unified theory of stochastic integrals and can find important applications in statistics (filtration) of processes.

The book is divided into two parts. The first one contains the theory of Choquet capacity, the major tool of measurability proofs. This part bears interest not only for probabilists but for anyone working in measure theory. In the second part the main purpose of the book, the general theory of stochastic processes, is presented.

The whole exposition is brilliantly visual, its language is clear and easy-flowing.

D. Vermes (Szeged)

Joost Engelfriet, Simple Program Schemes and Formal Languages (Lecture Notes in Computer Science, Vol. 20), VI+254 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1974.

The aim of this book, as the author writes in the introduction, is "to fit a part of program scheme theory into a formal language theoretic framework in such a way that

(1) semantic properties of program schemes can be translated into syntactic properties of formal languages, and

(2) results from formal language theory need only to be 'rephrased' in order to be applicable to program schemes."

The book consists of three parts. In Part I formal languages are viewed as program schemes, called L-schemes. This is followed by the introduction of the following classes of program schemes: Ianov schemes, recursive systems of Ianov schemes, procedure parameter schemes, and μ-terms. The classes of L-schemes equivalent to these classes of program schemes are also given.

In Part II general properties of L-schemes, such as equivalence, semantic determinism and semantic regularity, are studied.

The general theory of L-schemes developed in Part II is used in Part III for investigating some specific problems concerning program schemes. Among the topics studied in Part III are the decidability of certain program scheme properties, translation of program schemes and program schemes with markers.

The book is self-contained with respect to the theory of program schemes. The reader is assumed to be familiar with the basic concepts of elementary set theory and elementary algebra as well as formal language theory.

The presentation of the material is very clear. The book is a valuable contribution to the literature of theoretical computer science.

Ferenc Gécseg (Szeged)

P. Erdős—J. Spencer, Probabilistic Methods in Combinatorics, 106 pages, Akadémiai Kiadó, Budapest, 1974.

This book describes a powerful method to prove theorems of combinatorial nature. The method, developed mainly by Erdős, is based on the following idea: often the existence of a certain structrure with some properties can be proved by selecting a structure at random and then showing that the probability that it has the desired property is positive. The method is, thus, non-constructive; somewhat surprisingly, it often gives much better results then any known constructive method.

The book illustrates the technique by solving a variety of combinatorial problems, some of very fundamental nature (e.g. Ramsey's Theorem, graph and hypergraph coloring etc.). In exercises several further results are listed, giving a good survey of the most recent status of these important researches. Several unsolved problems are stated as well. The treatment is elementary, it does not require any knowledge of probability theory, but it does require much computational skill in estimating binomial coefficients and in other techniques of "asymptotics".

It seems that the probabilistic method (with necessary modifications) may have a much wider range of application then found so far. Therefore, this nice book is most recommended to everyone learning, or working in, combinatorics or neighboring areas.

L. Lovász (Szeged)

Wendell H. Fleming—Raymond Richel, Deterministic and stochastic control theory (Applications of Mathematics, 1), 222 pages, New York—Heidelberg—Berlin, Springer-Verlag, 1975.

Control theory is generally referred to as a modern discipline of applied mathematics though its fundamental problem "How to reach a goal in the best possible way?" is older then mankind itself. To have a well-posed problem clearly one has to define the goal of the activity and to say what is ment by the word "best" (i.e. to specify an expense function). But the very essence of the problem is determined by the possible ways of reaching the aim. The processes by which we can achieve our purpose determine our restricted freedom in the choice and we have to make the best possible compromise, i.e. to use the optimal strategy. Also the underlying processes serve as a basis for the classification of control problems into classes like deterministic, stochastic continuous, discontinuous problems, etc.

The first half of the present book contains a well-written self-contained exposition of deterministic control problems governed by ordinary differential equations. (Calculus of variations, Pontrjagin's principle, dynamic programming, existence and continuity of optimal strategies.) The proofs are detailed, many examples help understanding the presented material and its applications.

In contrast with the deterministic problems, no closed, rounded up theory exists as yet for stochastic control, not even for the control of diffusion processes, the subject of the second half of the book. So this part aims rather to introduce the reader into this rapidly developing field (up to its stage at about 1970), and to enable him to solve concrete problems. The authors start with a list of definitions and (in part rather deep) theorems from the theory of Markov processes and partial differential equations, necessary for the further development. Proofs are omitted but several examples and precise references support the reader not to get bored. The last chapter contains one, (the authors' own) approach to optimal control of diffusion processes via partial differential equations. It culminates in a sufficient optimality condition and an existence theorem, which enable them to solve the linear regulator problem, the permanent example in stochastic control. The Kalman-Bucy filter and the separation principle for linear systems are presented as well.

An extensive bibliography helps the orientation in recent literature.

D. Vermes (Szeged)

Dale Husemoller, Fibre Bundles (Graduate Text in Mathematics, 20), Second edition, 327 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1974.

This book contains important chapters of the theory of fibre bundles. The author concentrates on the work of Milnor, Hirzebruch, Bott, Adams, Hopf, Chern, Stiefel, Whitney, Grothendieck Atiyah, Toda, etc. In this second edition the author has added a section on the Adams conjuncture and an appendix on the suspension theorems.

The book consists of three parts. Part I contains the general theory of fibre bundles; the Milnor construction of a universal fibre bundle for any topological group is also given. Part II gives the elements of K-theory, namely stability properties of vector bundles, relative K-theory, Bott periodicity in the complex case, Clifford algebras, the Adams oprations and representations, representation rings of classical groups, the Hopf invariant, vector fields on the sphere and stable homotopy. The proof of Atiyah on the nonexistence of elements with Hopf invariant 1 is also presented and the proof of the vector field problem is sketched. A systematic development of characteristic classes and their applications to manifolds is given in Part III and is based on the approach of Hirzbruch as modified by Grothendieck.

Reading the book claims a certain knowledge from topology and the theory of differentiable manifolds. It is a very instructive reading due in part to the large number of exercises and examples.

Z. I. Szabó (Szeged)

John G. Kemeny—J. Laurie Snell, Finite Markov Chains (Undergraduate Texts in Mathematics), IX+210 pages, New York—Heidelberg—Berlin, Springer-Verlag, 1976.

This book is a reprint of the 1960 edition published by D. Van Nostrand, Princeton, N. J., in the University Series in Higher Mathematics. No changes have been made of the first edition. It is a complete treatment of the theory of finite Markov chains and it has already proved its vitality in the last sixteen years. Suitable as an undergraduate introduction to probability theory or it can certainly replace a course in matrix calculus. Applications to learning theory and other socio-economic models (and to diffusion, genetics, sports, the Land of Oz and anything) are given. For a detailed review from such an authority as K. L. Chung see MR *22* (1961) #5998.

S. Csörgő (Szeged)

Rudolph Kurth, Elements of Analytical Dynamics (International Series in Pure and Applied Mathematics, Vol. 105), VIII+181 pages, Pergamon Press, Oxford—New York—Toronto—Sydney—Paris—Frankfurt, 1976.

This is a useful and easily readable textbook on analytical mechanics serving as a preparatory course to a profound study of topological dynamics for graduate students of mathematics. The reader is supposed to be familiar with some knowledge of calculus, general topology and differential geometry only. The mathematical structures occurring in the treatment of analytical dynamics are discussed in detail (e.g. the notion of differentiable manifold, elements of the theory of differential equations and of the calculus of variations). After the study of the Hamilton-Jacobi theory, Noether's theorem and the Liapunov stability theory the chapter "Jacobi's Geometric Interpretation of Dynamics" follows, which is a short introduction to Riemannian, Lagrangian and Finsler geometry.

P. T. Nagy (Szeged)

H. Elton Lacey, The isometric theory of classical Banach spaces (Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Band 208), X+272 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1974.

The main purpose of this book is to investigate structural questions for classical Banach spaces. A Banach space is called classical, if it is either linearly isometric to an $L^1(\mu)$ space (real or complex) for some measure μ and some $1 \leq p \leq \infty$ or its dual space is linearly isometric to an $L^1(\mu)$ space; in the last case we say that the space is an L^1-predual space. Various necessary and sufficient conditions are given for a Banach space to be a classical one. They are framed in terms of conditions on the norm, conditions on the dual spaces and on subspaces. In the investigation the vector lattice structure of classical spaces plays a basic role. The book is divided into 7 Chapters. Chapters 1 and 2 summarize the fundamental definitions and theorems concerning partially ordered Banach spaces, topology and regular Borel measures. Chapter 3 deals with the algebraic and Banach space characterization of the space of continuous functions. Chapter 4 contains embedding theorems for classical sequence spaces into continuous function spaces. Chapter 5 is devoted to representation theorems for spaces of type $L^p(\mu)$. Chapter 6 contains characterizations of abstract L^p spaces and measure algebras (abstract L^p spaces are Banach lattices with p-additive norm). Chapter 7 gives characterizations of L^1-predual spaces.

All the chapters end with exercises and some open problems. General topology, Banach spaces, and measure theory are assumed as prerequisites.

L. Gehér (Szeged)

Ernst G. Manes, Algebraic Theories (Graduate Text in Mathematics, No. 26), 356 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1976.

The following assignments are natural and often applied in mathematics: to each set S, assign its power set 2^S; to each element e, assign the one-element set $\{e\}$ (in this way one "inserts" S into 2^S); to each pair of relations, assign their relation-theoretical product (note that a relation between sets S and T may be considered as a mapping of S into 2^T). Formation of power set, insertion and product are connected by a few very simple laws; the same laws are observable, e.g., between formation of the free group $F(S)$ over S, insertion of free generators, and product of (homo)morphisms of free groups into one another (such a morphism of $F(S)$ into $F(T)$ may be considered as a mapping of S into $F(T)$). These assignments and laws lead to the notion of an *algebraic theory;* they furnish the "data" and "axioms" of this notion.

The above examples use the category of sets; however, algebraic theories can be defined over any category. The book we are concerned with develops a general theory of algebraic theories. This is the content of its main chapter, preceded by two big preparatory ones which are interesting also on their own right. The first of them presents a modern introduction to equational theory of algebras where infinitary operations are also allowed. The second chapter bears the attractive title "Trade Secrets of Category Theory", and, together with some paragraphs of the first chapter, it can serve as a mini-monograph on category theory for pure mathematicians. The last chapter deals with applications of algebraic theories to the following areas: topological dynamics, minimal realization of systems, theory of fuzzy automata. Since algebraic theories can be found in many further circumstances of algebra, topology and automata theory, the acquaintance with the third (main) chapter will be useful for everybody who is engaged in investigations in these fields.

The book is well-organized and well-readable; its style unites informality and exactness. The author helps the reader in several ways: every section is followed by historical notes and many exercises of various strength, while the entire book has useful indices and an abundant bibliography.

B. Csákány (Szeged)

P. McMullen—G. C. Shephard, Convex polytopes and the Upper Bound Conjecture (London Mathematical Society Lecture Notes Series 3), IV+184 pages, Cambridge University Press, 1971.

An outstanding problem in the theory of convex polytopes has been the Upper Bound Conjecture, describing which polytope (in d dimensions and with n vertices) has the largest possible number of faces. These notes were already in print when P. McMullen, one of the authors, succeeded to prove this famous conjecture. The solution was added to the book as a last chapter.

The book is devoted to the study of the combinatorial structure of convex polytopes. It describes the basic methods in this area: polarity, the Dehn-Sommerville questions, Gale diagrams, shelling. Many of these find their application in the solution of the upper bound conjecture. Although in great lines the presentation follows Grünbaum's well-known book "Convex Polytopes" (Wiley, 1967), there are several divergences, e.g. in the treatment of the support properties and in the proof of the Dehn-Sommerville equations. Also the authors manage to write up the material in a compact, and yet easily readable way. This book is well advised to all who want to learn, or do research in, the theory of convex polytopes.

L. Lovász (Szeged)

G. Pickert, Projektive Ebenen, Zweite Auflage (Die Grundlehren der mathematischen Wissenschaften, Band 80), IX+371 Seiten, Springer-Verlag, Berlin—Heidelberg—New York, 1975.

The first edition of this book in 1955 was the earliest in the mathematical literature giving a systematic treatment of the new domain of mathematics, called theory of projective planes, developed from the 1930's. The present book had a great effect encouraging the growth of the interest on this subject even beyond the area of foundation of geometries.

It is well known that the structure of projective planes has a greater variety then the structure of projective spaces, namely, Desargues's Theorem is not necessarily valid. Projective planes can be coordinatized by various not necessarily associative and distributive algebraic structures. Hence the projective planes provide models for algebraic structures, so they are useful in the study of questions of algebraic nature.

For the description of the structure of projective planes constructions and results from the geometry of webs (Geometrie der Gewebe) are used. This theory was introduced by Blaschke's school in the 1930's in connection with topological questions of differential geometry and developed later in algebraic and differential geometrical directions. A geometric web is three families of lines in the plane such that exactly one line of each family passes through each point. Very useful tools of the characterization of webs are the so-called "closure conditions", which are equivalent to identities for the coordinates of the plane.

The theory and classification of the finite and topological projective planes has made a very intensive progress in the last decades. The finite planes serve as standard models for combinatorial geometries, and the planes with topological and differentiable structures have a great interest in topological and differential geometry.

The book consists of 12 Chapters. The Chapters 1—2 serve as an introduction to the incidence structures and the theory of webs. Chapters 3—9 and 11 deal with planes satisfying various geometrical conditions and with algebraic investigations on the corresponding coordinate structures. In Chapters 10 and 11 a short introduction to the theory of topological and finite planes is given.

The book is recommended to mathematicians doing research in geometry, algebra or combinatorics and interested in problems connected with the theory of projective planes.

P. T. Nagy (Szeged)

G. Pólya—G. Szegő, Problems and Theorems in Analysis, Volumes I and II, (Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Band 193 and 216), XIX+389 and XI+391 pages respectively, Springer-Verlag, Berlin—Heidelberg—New York, 1972 and 1976.

A number of mathematicians has been brought up with the help of the famous and excellent problem-book *Aufgaben und Lehrsätze aus der Analysis*. The present book is not only an English translation of the German original. The original text has been enlarged by many new problems and there are some other changes. All the alterations amount to less than ten per cent of the text. The book also contains the solutions of the problems, which is of great help to the reader. These books are recommended to students and research workers who are interested in classical analysis problems.

L. Gehér (Szeged)

W. Rinow, Lehrbuch der Topologie, 724 pages, VEB Deutscher Verlag der Wissenschaften, Berlin, 1975.

The main text of this book is based on lectures in topology which have been held by the author since 1950 at Greifswald University. In accordance with this fact it is aimed to be a university textbook. The selection and style of the text show Professor Rinow's natural turn for the methods of instruction. In contrast with most modern topology books the text comprises general, combinatorial and algebraic topology. The book is divided into fifteen chapters. The first seven chapters lead the reader along the most significant parts of general topology, discussing all the usual concepts and problems like tracing and comparison of topologies, relativization, convergence, continuity of mappings, separation, compactness, metrization, uniform structures, etc. Chapter VIII gives a glance into combinatorial methods in topology and applies these to give a proof for the classical domain invariance theorem in Euclidean spaces. Chapter IX is devoted to a short survey of dimension theory. Chapter X introduces the concept of homotopy, studies mappings in spheres and proves the domain invariance theorem again. Jordan curve theorem and Schoenfliess theorem are also proved. The chapter ends with a short investigation into surface topology. The last five chapters deal with various homologies and cohomologies, with the connection between homologies and homotopy and with duality theorems.

The book is recommended to students and to anyone taking interest in topology.

L. Gehér (Szege)

C. P. Rourke and B. J. Sanderson, Iutroduction to Piecewise-Linear Topology (Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 69), VIII+123 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1972.

This book is an excellent introduction to modern geometric topology, treating the continuous and smooth topology as a unified subject. The generalization of many results of smooth topology is made possible by the application of the new technique of geometric topology, called the piecewise-linear (p.l.) topology.

Chapters 1—5 (Polyhedra and p.l. maps; Complexes; Regular neighbourhoods; Pairs of polyhedra and isotopies; General position and applications) serve as an undergraduate introductory course to p.l. topology. Here familiarity with the elementary notions of point-set topology is assumed only.

Chapters 6—7 (Handle theory; Applications) give an account of Smale's handle theory in a piecewise linear setting and of its applications to the Poincaré conjecture and the h-cobordism theo-theorem. Originally, this theory was developed using the technique of differentiable topology, in spite of the fact that these problems are of continuous topological nature.

The results of algebraic topology which are used are collected in Appendices. A bibliography of research papers is also included.

P. T. Nagy (Szeged)

S. Bouncrlstlano, C. P. Rourke and B. J. Sanderson, A Geometric Approach to Society (London Mathematical Society Lecture Note Series 18), VI+149 pages. Cambridge University Press, Cambridge—London—New York—Melbourne 1976.

From the introduction: "The purpose of these notes is to give a geometrical treatment of generalised homology and cohomology theories. The central idea is that of a 'mock bundle', which is the geometric cocycle of a general cobordism theory, and the main new result is that any homology theory is a generalised bordism theory. Thus every theory has both cycles and cocycles; the cycles are manifolds, with a pattern of singularities depending on the theory, and the cocycles are mock bundles with the same 'manifolds' as fibres."

In Chapter I the transition from functor on cell complexes to homotopy functor on polyhedra is axiomatised, the mock bundles of Chapter II being the principal example. In Chapter II, the simplest case of mock bundles, corresponding to p.l. (piecewise linear) cobordism, is treated, but the definitions and proofs all generalise to the more complicated setting of later chapters. Chapter III gives the geometric treatment of coefficients, where again only the simplest case, p.l. bordism, is treated. A geometric proof of functoriality for coefficients is given in this case. Chapter IV extends the previous work to a generalised bordism theory and includes the 'killing' process and a discussion of functoriality for coefficients in general (similar results to Hilton's treatment being obtained). Chapter V extends to the equivariant case and discusses the Z_2 operations on p.l. cobordism in detail. Chapter VI discusses sheaves, which work nicely in the cases when coefficients are functorial (for 'good' theories or for 2-torsion free abelian groups) and finally Chapter VII proves that a general theory is geometric.

P. T. Nagy (Szeged)

Joe Rosen, Symmetry discovered. Concepts and Applications in Nature and Science, 138 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1975.

This book, written with an excellent sense of didactics, introduces the reader to the examination of symmetry of geometrical objects, nature and science in a very light and witty style. Rosen starts his voyage of discovering the world of symmetry by explaining what symmetry is, and where and how to find it.

In the first part of the book the author describes the symmetry groups of forms in planar and 3-dimensional spaces with many examples and figures. But symmetry is not restricted to geometrical constructions alone. The author shows that physical operations are often symmetrical in nature, and he also gives an insight into symmetry provided by science and technology.

Reading the present work requires no special mathematical preparation. The reader is playfully introduced into the basic concepts and terminology of symmetry. For the readers who wish to pursue specific topics the author has supplied many references.

Z. I. Szabó (Szeged)

G. Segal, New Developments in Topology, (London Mathematical Society Lecture Note Series 11), 128 pages, Cambridge University Press, 1974.

In June 1972 a Symposium in Algebraic Topology was held in Oxford. The main theme of this Symposium was the K-theory: The present book contains eleven treatises on K-theory written by participants, based on their lectures. The familiarity of the reader with modern algebraic topology is required.

L. Gehér (Szeged)

D. J. Simms—N. M. J. Woodhouse, Lectures on Geometric Quantization (Lecture Notes in Physics, Vol. 53), II+166 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976.

These lectures are written in the spirit of the geometric quantization programme of B. Kostant and J.-M. Sourian. The aim of this programme is to formulate the procedure of quantization in differential geometric language. The systems of classical mechanics are modelled by symplectic geometries and Hamiltonian systems. The procedure of quantization is a construction of a Hilbert space H on which each classical observable (that is, each smooth function on the symplectic manifold M) is represented as an Hermitian operator in such a way that the Poisson bracket of classical observables is represented by the commutator of the corresponding operators. In the simplest case, the Hilbert space H consists of complex valued functions on the manifold M. In the case of more complicated systems (e.g. particles with internal degrees of freedom) H is constructed from the sections of a certain Hermitian line bundle over M. The described process of quantization is illustrated by very interesting examples.

The treatment assumes an experience in differential geometrical technique, especially in exterior calculus. In appendices a brief survey of the underlying mathematical theory is given: fibre bundles, Chern characteristic classes, and Lie algebra cohomology theory.

P. T. Nagy (Szeged)

Frank Spitzer, Principles of Random Walk (Graduate Texts in Mathematics 34), second edition, XIII+408 pages, New York—Heidelberg—Berlin, Springer-Verlag, 1976.

This is the second edition of a book (the first one was published by D. Van Nostrand, Princeton, N.J., in the University Series in Higher Mathematics, 1964) which can be safely called a classic. Classic, not in the sense that it would be old, but that it is fundamental and belongs to the group of best books ever published in probability theory. For an extensive and through-going review on the real mathematical content of the first edition we refer to MR 30(1965)#1521 by T. Watanabe. The book presents a complete and nearly self-contained treatment of random walk and certainly covers almost all major topics in the theory up to 1964. From the author's preface: "In this edition a large number of errors have been corrected, an occasional proof has been streamlined, and a number of references are made to recent progress". These new references (placed in brackets and footnotes) are to a supplementary bibliography, which contains 26 new items, and make the book again up-to-date. It is written mainly for probabilists and the prerequisite is, as described in the preface to the first edition, "some solid experience and interest in analysis, say, in two or three of the following areas: probability theory, real variables and measure, analytic functions, Fourier analysis, differential and integral operators". It has served as the main source for research in this area in the last twelve years, and it certainly will maintain this role for a long time to come.

S. Csörgő (Szeged)

Zhe-Xian Wan, Lie Algebras (International Series of Monographs in Pure and Applied Mathematics, Vol. 104), VIII+228 pages, Pergamon Press, Oxford—New York—Toronto—Sydney—Braunschweig, 1975.

This book is based on a series of lectures given in the seminar on Lie groups at the Institute of Mathematics of Academia Sinica (Peking) during the years 1961—1963. The purpose of the book "is to supply an elementary background to the theory of Lie algebras, together with sufficient material to provide a reasonable overview of the subject". In accord with its introductory character the book deals only with algebras over the complex field.

Chapters 1—4 present an introduction to the general theory of Lie algebras (nilpotency and solvability, Cartan subalgebras, Cartan's criterions). Chapters 5—8 deal with the structure and classification theory of semisimple Lie algebras and with their automorphisms. Chapters 9—11 serve as an introduction to the representation theory of semisimple Lie algebras. Chapters 12—15 contain selected topics on representation theory. Chapter 15 is devoted to the real forms of complex semisimple Lie algebras.

The book is well organized, the presentation is concise but always clear and well-readable, its format is nice.

P. T. Nagy (Szeged)

Bertram A. F. Wehrfritz, Infinite linear groups (Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 76), XIV+229 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1973.

A linear group is a group of invertible matrices with entries in a commutative field. Their study started in the early years of this century with the work of Burnside and Schur. In the last twently years infinite linear groups have been used increasingly in the theory of abstract groups. On the one hand, much of the work on linear groups is hard to read for group theorists, and on the other hand, many results on linear groups appeared under purely group-theoretic titles. The book under review is the first to gather all this material together.

Infinite linear groups are useful in group theory in several ways. First of all, they arise via the automorphism groups of certain types of abelian groups: free abelian groups of finite rank, torsion-free abelian groups of finite rank and divisible abelian p-groups of finite rank. Thanks to Mal'cev, infinite linear groups play, in these days, a central role in the theory of soluble groups satisfying various rank conditions and in the theory of the automorphism groups of these groups. It is a recent result, that "the automorphism groups of certain finitely generated soluble (in particular finitely generated metabelian) groups contain significant factors isomorphic to groups of automorphisms of finitely generated modules over certain commutative Noetherian rings". Linear groups also arise via the following theorem of Mal'cev: a group G is isomorphic to some linear group of degree n if and only if each of its finitely generated subgroups is isomorphic to a linear group of degree n. If one has some information about which linear groups are isomorphic to the finitely generated subgroups of G, then one can sometimes find a concrete linear group that is isomorphic to G. "This led to very important characterizations of certain groups such as PSL$(2, F)$ over locally finite fields F, which now play a crucial role in the theory of locally finite groups". In the author's opinion "to date we have only scratched the surface of the applications of infinite linear groups to locally finite groups."

Linear groups are also important in that they form a relatively accessible class of highly non-trivial, highly non-soluble groups, and, consequently, it is relatively easy to test conjectures on them. Moreover, it is quite common to solve a general problem for the linear case first. On the other hand, it sometimes happens that one ad-hoc knows that a group is isomorphic or related to a certain linear group.

The arrangement of the book is the following: the fundamentals are given in chapters 1, 5, 6, and, to some extent, 2. The basic material is split into two parts in order to present the theories of soluble linear groups and finitely generated linear groups in Chapters 3 and 4, before the reader gets bored. Roughly speaking, Chapter 1 is the ring theoretic and Chapters 5 and 6 are the geometric introduction. The rest of the 14 Chapters is devoted to the study of Jordan decomposition in linear groups, structure theorems for locally nilpotent linear groups, upper central series, locally super-soluble linear groups, periodic linear groups, groups of automorphisms of finitely generated modules over commutative rings, algebraic groups over algebraically closed fields. "Suggestions for Further Reading", a Bibliography, and Index close the book.

József Szűcs (Szeged)

LIVRES REÇUS PAR LA RÉDACTION

B. D. O. Anderson—M. A. Arbib—E. G. Manes, Foundations of system theory: Finitary and infinitary conditions (Lecture Notes in Economics and Mathematical Systems, Vol. 115), VII+93 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 18,—.

T. M. Apostol, Introduction to analytic number theory (Undergraduate Texts in Mathematics), XII+338 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 36,20.

A. V. Balakrishnan, Applied functional analysis (Applications of Mathematics, Vol. 3), X+309 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 48,40.

M. S. Bazaraa—C. M. Shetty, Foundations of optimization (Lecture Notes in Economics and Mathematical Systems. Mathematical Programming, Vol. 122), VI+193 pages, Berlin—Heidelberg—New York, 1976. — DM 20,—.

T. S. Blyth—M. F. Janowitz, Residuation theory (International Series of Monographs in Pure and Applied Mathematics, Vol. 102), IX+379 pages, Oxford—New York, Pergamon Press, 1972. — £ 7,50.

A. A. Borovkov, Stochastic processes in queueing theory (Applications of Mathematics, Vol. 4), Translated from the Russian by K. Wickwire, XI+280 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 32,—.

A. A. Borowkow, Wahrscheinlichkeitstheorie (Mathematische Lehrbücher und Monographien. I. Abteilung: Mathematische Lehrbücher, Bd. 27), XI+264 Seiten, Berlin, Akademia-Verlag, 1976, — 28,— M.

P. Branquart—J.-P. Cardinael—J.-P. Delescaille—J. Lewi—M. Vanbegin, An optimized translation process and its applications to ALGOL 68 (Lecture Notes in Computer Science, Vol. 38), IX+334 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM. 30,—.

S. Brehmer, Einführung in die Masstheorie (Wissenschaftliche Taschenbücher, 137), 186 Seiten, Berlin, Akademie-Verlag, 1975. — 8,— M.

P. J. Cameron, Parellelisms of complete designs (London Mathematical Society Lecture Note Series, 23), IV+144 pages, Cambridge University Press, 1976. — £ 3,95.

J. W. Cohen, On regenerative processes in queueing theory (Lecture Notes in Economics and Mathematical Systems, Vol. 121), IX+93 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 18,—.

Data Base Systems. Proceedings, 5th Informatik Symposium, Germany, Bad Homburg v. d. H., September 24—26, 1975. Edited by H. Hasselmeier and W. G. Spruth (Lecture Notes in Computer Science, Vol. 39), VI+386 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 32,—.

Diocles, In burning mirrors, The Arabic translation of the lost Greek original edition, with English translation and commentary by G. J. Toomer, IX+249 pages, IX+249 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 68,—.

J. E. Donahue, Complementary definitions of programming language semantics (Lecture Notes in Computer Science, Vol. 42), VII+172 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 20,—.

C. Faith, Algebra II. Ring theory (Die Grundlehren der mathematischen Wissenschaften — A Series of Comprehensive Studies in Mathematics, 191), XVIII+302 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 98,—.

L. Gillman—M. Jerison, Rings of continuous functions (Graduate Texts in Mathematics), Reprint of the 1960 ed. published by Van Nostrand, Princeton, XIII+300 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 36.20.

U. Grenander, Pattern synthesis. Lectures in pattern theory, Vol. 1 (Applied Mathematical Sciences, 18), VII+509 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 36,20.

R. C. Gunning, Riemann surfaces and generalized theta functions (Ergebnisse der Mathematik und ihrer Grenzgebiete — A Series of Modern Surveys in Mathematics, 91), XII+165 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 48,—.

Integer programming and related areas. A classified bibliography. Compiled at the Institut für Ökonometrie und Operations Research, University of Bonn. Edited by C. Kastning (Lecture Notes in Economics and Mathematical Systems. Operations Research, Vol. 128), XII+495 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 37,—.

N. Jacobson, Lectures in abstract algebra I. Basic concepts (Graduate Texts in Mathematics, Vol. 30), Reprint of the 1951—1964 ed. published by Van Nostrand, New York in The University Series in Higher Mathematics, XII+217 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 26,40.

N. Jacobson, Lectures in abstract algebra, Vol. 3: Theory of fields and Galois theory (Graduate Texts in Mathematics, Vol. 32), 2nd corrected printing. Originally published by Van Nostrand, New-York in The University Series in Higher Mathematics, XI+323 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 36,20.

P. Kall, Stochastic linear programming (Econometrics and Operations Research, Vol. 21), VI+95 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 38,—.

F. Kárteszi, Introduction to finite geometries (Disquisitiones Mathematicae Hungaricae, 7), XIII+266 pages, Budapest, Akadémiai Kiadó, 1976.

U. Kausmann—K. Lommatzsch—F. Nožička, Lineare parametrische Optimierung (Wissenschaftliche Taschenbücher, 166), 160 Seiten, Berlin, Akademie-Verlag, 1976. — 8,— M.

J. L. Kelley—I. Namioka, Linear topological spaces (Graduate Texts in Mathematics, Col. 36), 2nd corrected printing. Originally published by Van Nostrand Publishing Company, XV+256 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 36,20.

V. Kempe, Analyse stochastischer Systeme, Teil 1 (Wissenschaftliche Taschenbücher, 137), 200 Seiten, Berlin, Akademie-Verlag, 1976. — 8,— M.

A. A. Kirillov, Elements of the theory of representations (Grundlehren der mathematischen Wissenschaften — A Series of Comprehensive Studies in Mathematics, 220), Translated from the Russian by E. Hewitt, XI+315 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 85,—.

I. Lakatos, Proofs and refutations. The logic of mathematical discovery. Edited by J. Worrall and E. Zahar, XII+174 pages, Cambridge University Press, 1976. — £ 1,95.

P. Lax—S. Burstein—A. Lax, Calculus with applications and computing, Vol. 1 (Undergraduate Texts in Mathematics), XI+513 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 36,20.

H.-J. Lüthi, Komplementaritäts- und Fixpunktalgorithmen in der mathematischen Programmierung, Spieltheorie und Ökonomie (Lecture Notes in Economics and Mathematical Systems. Mathe-

matical Economics, Vol. 129), VII+145 Seiten, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 18,—.

J. E. Marsden—M. McCrasken, The Hopf bifurcation and its applications (Applied Mathematical Sciences, Vol. 19), XIII+408 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 36,20.

Mathematical Foundations of Computer Science 1976. Proceedings 5th Symposium, Gdansk, September 6—10, 1976. Edited by A. Mazurkiewicz (Lecture Notes in Computer Science, Vol. 45), XI+606 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 46,—.

Mathematical Systems Theory. Proceedings of the International Symposium, Udine, Italy, June 16—27, 1975. Edited by G. Marchesini and S. K. Mitter (Lecture Notes in Economics and Mathematical Systems. Systems Theory, Vol. 131), X+408 pages, Berlin—Heidelberg—New-York, Springer Verlag, 1976. — DM 35,—.

S. G. Michlin, Approximation auf dem kubischen Gitter (Mathematische Lehrbücher und Monographien. II. Abteilung: Mathematische Monographien, Bd. 40), IX+195 Seiten, Berlin, Akademie-Verlag, 1976. — 36,— M.

Multiple Criteria Decision Making, Kyoto, 1975. Edited by M. Zeleny (Lecture Notes in Economics and Mathematical Systems, Vol. 123), XXVII+345 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 32,—.

Multiple Criteria Decision Making. Proceedings of a Conference held at Jouy-en-Josas, France, May 21—23, 1975. Edited by H. Thiriez and S. Zionts (Lecture Notes in Economics and Mathematical Systems. Operations Research, Vol. 130), VI+409 pages, Berlin—Heidelberg—New-York, Springer-Verlag, 1976. — DM 35,—.

D. G. Northcott, Finite free resolutions (Cambridge Tracts in Mathematics, 71), XII+271 pages, Cambridge University Press, 1976. — £ 12,50.

J. T. Oden—J. N. Reddy, Variational methods in theoretical mechanics (Universitext), IX+302 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 29,80.

Optimization techniques modeling and optimization in the service of man, Part 1—2. Proceedings, 7th IFIP Conference, Nice, September 8—12, 1975, Edited by J. Cea (Lecture Notes in Computer Science, Vol. 40), XIV+854, XIII+852 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 62,— +62,—.

Proceedings of the lst Conference of the European Cooperation in Informatics, Amsterdam, August 9—12, 1976. Edited by K. Samelson (Lecture Notes in Computer Science, Vol. 44), VIII+322 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 30,—.

Revised Report on the Algorithmic Language ALGOL 68, Edited by A. van Wijngaarden, B. J. Mailloux, J. E. L. Peck, C. H. A. Koster, M. Sintzoff, C. H. Lindsey, L. G. L. T. Meertens, R. G. Fisker, II+236 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 24,—.

W. Schäfer, Theoretische Grundlagen der Stabilität technischer Systeme (Wissenschaftliche Taschenbücher, 169), 130 Seiten, Berlin, Akademie-Verlag, 1976. — 8,— M.

L. E. Sigler, Exercises in set theory, VII+134 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 9,70.

E. Specker—V. Strassen, Komplexität von Entscheidungsproblemen. Ein Seminar (Lecture Notes in Computer Science, Vol. 43), V+217 pages, Berlin—Heidelberg—New York, Springer-Verlag. — DM 23,—.

Traffic Equilibrium Methods. Proceedings of the Symposium held at the University of Montreal, November 21—23, 1974. Edited by M. A. Florian (Lecture Notes in Economics and Mathematical Systems. Operations Research, Vol. 118), XXIII+432 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 35,—.

W. Walter, Gewöhnliche Differentialgleichungen (Heidelberger Taschenbücher, Bd. 110), 2. korrigierte Auflage, XI+229 Seiten, Berlin—Heidelberg—New York, Springer-Verlag, — DM 18,80.

J. Wermer, Banach algebras and several complex variables (Graduate Texts in Mathematics, Vol. 35), 2nd edition, IX+161 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 36,20.

O. Zariski—P. Samuel, Commutative algebra ii (Graduate Texts in Mathematics, Vol. 29), Corrected printing. Originally published by Van Nostrand Publishing Company, X+414 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1976. — DM 36,20.

TOMUS XXXVIII — 1976 — 38. KÖTET

BIBLIOGRAPHIE

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)

INDEX: 26024

76-3085 — Szegedi Nyomda — F. v.: Dobó József igazgató

Felelős szerkesztő és kiadó: Szőkefalvi-Nagy Béla
A kézirat a nyomdába érkezet: 1976. január. 18
Megjelenés: 1976. december hó

Példányszám: 1200. Terjedelem: 19,25 (A/5) ív
Készült monószedéssel, íves magasnyomással,
az MSZ 5601-24 és az MSZ 5602-55 szabvány szerint